

Package ‘fasstr’

January 9, 2020

Title Analyze, Summarize, and Visualize Daily Streamflow Data

Version 0.3.1

Description The Flow Analysis Summary Statistics Tool for R, 'fasstr', provides various functions to clean and screen daily stream discharge data; calculate and visualize various summary statistics and metrics; and compute annual trending (using 'zyp' package methods <<https://CRAN.R-project.org/package=zyp>>) and volume frequency analyses (using methods similar to HEC-SSP (2019) <<https://www.hec.usace.army.mil/software/hec-ssp/>>). It features useful function arguments for filtering of and handling dates, customizing data and metrics, and the ability to pull daily data directly from the Water Survey of Canada hydrometric database (<<https://collaboration.cmc.ec.gc.ca/cmc/hydrometrics/www/>>).

Depends R (>= 3.2.0)

License Apache License 2.0

URL <https://github.com/bcgov/fasstr>,
<https://www2.gov.bc.ca/gov/content/environment/air-land-water/water>

BugReports <https://github.com/bcgov/fasstr/issues>

Encoding UTF-8

Imports dplyr (>= 0.8.1), e1071 (>= 1.7.0.1), fitdistrplus (>= 1.0.14), ggplot2 (>= 3.1.0), grDevices, lubridate (>= 1.7.4), PearsonDS (>= 1.1), plyr (>= 1.8.4), purrr (>= 0.3.2), RcppRoll (>= 0.3.0), scales (>= 1.0.0), tidyhydat (>= 0.4.0), tidyr (>= 0.8.3), openxlsx (>= 4.1.0), zyp (>= 0.10.1.1)

Suggests knitr, rmarkdown, testthat

LazyData true

RoxygenNote 7.0.2

VignetteBuilder knitr

NeedsCompilation no

Author Jon Goetz [aut, cre],
 Carl James Schwarz [aut],
 Robin Pike [ctb],
 Province of British Columbia [cph]

Maintainer Jon Goetz <jon.goetz@gov.bc.ca>

Repository CRAN

Date/Publication 2020-01-09 12:40:02 UTC

R topics documented:

add_basin_area	3
add_cumulative_volume	4
add_cumulative_yield	6
add_daily_volume	7
add_daily_yield	8
add_date_variables	10
add_rolling_means	11
add_seasons	13
calc_all_annual_stats	14
calc_annual_cumulative_stats	17
calc_annual_flow_timing	20
calc_annual_lowflows	22
calc_annual_outside_normal	25
calc_annual_stats	27
calc_daily_cumulative_stats	29
calc_daily_stats	32
calc_flow_percentile	35
calc_longterm_daily_stats	37
calc_longterm_mean	40
calc_longterm_monthly_stats	42
calc_longterm_percentile	45
calc_monthly_cumulative_stats	47
calc_monthly_stats	50
compute_annual_frequencies	53
compute_annual_trends	56
compute_frequency_analysis	59
compute_frequency_quantile	61
compute_full_analysis	64
compute_hydat_peak_frequencies	66
fill_missing_dates	69
plot_annual_cumulative_stats	70
plot_annual_flow_timing	73
plot_annual_lowflows	75
plot_annual_means	77
plot_annual_outside_normal	79
plot_annual_stats	81
plot_daily_cumulative_stats	84

plot_daily_stats	87
plot_data_screening	90
plot_flow_data	92
plot_flow_duration	94
plot_longterm_daily_stats	97
plot_longterm_monthly_stats	100
plot_missing_dates	103
plot_monthly_cumulative_stats	105
plot_monthly_stats	108
screen_flow_data	110
write_flow_data	113
write_full_analysis	115
write_objects_list	117
write_plots	118
write_results	120

Index	121
--------------	------------

add_basin_area	<i>Add a basin area column to daily flows</i>
----------------	---

Description

Add a column of basin areas to a streamflow dataset, in units of square kilometres.

Usage

```
add_basin_area(data, groups = STATION_NUMBER, station_number, basin_area)
```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using station_number argument.
groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using station_number argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires tidyhydat package and a HYDAT database. Leave blank if using data argument.
basin_area	Upstream drainage basin area, in square kilometres, to apply to observations. Three options: (1) Leave blank if groups is STATION_NUMBER with HYDAT station numbers to extract basin areas from HYDAT.

- (2) A single numeric value to apply to all observations.
- (3) List each basin area for each group/station in groups (can override HYDAT value if listed) as such `c("08NM116" = 795, "08NM242" = 10)`. If group is not listed the HYDAT area will be applied if it exists, otherwise it will be NA.

Value

A tibble data frame of the original source data with an additional column:

```
Basin_Area_sqkm
      area of upstream drainage basin area, in square kilometres
```

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Add the HYDAT basin area to a data frame with station numbers
flow_data <- tidyhydat::hy_daily_flows(station_number = "08NM116")
add_basin_area(data = flow_data)

# Add the HYDAT basin area to data from HYDAT
add_basin_area(station_number = "08NM116")

# Set a custom basin area
add_basin_area(station_number = "08NM116",
               basin_area = 800)

# Set multiple custom basin areas for multiple stations
add_basin_area(station_number = c("08NM116", "08NM242"),
               basin_area = c("08NM116" = 800, "08NM242" = 10))

}
```

`add_cumulative_volume` *Add daily cumulative volumetric flows on an annual basis*

Description

Add a column of rolling daily cumulative volumetric flows on an annual basis to a streamflow dataset. Adds the volumetric discharge from each day with the previous day(s) for each year, in units of cubic metres. The cumulative flows restart every year and are only calculated in years with complete data.

Usage

```
add_cumulative_volume(
  data,
  dates = Date,
```

```

  values = Value,
  groups = STATION_NUMBER,
  station_number,
  water_year_start = 1
)

```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.
groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using <code>station_number</code> argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires <code>tidyhydat</code> package and a HYDAT database. Leave blank if using data argument.
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.

Value

A tibble data frame of the source data with an additional column:

```

Cumul_Volume_m3
      cumulative volumetric flows for each day for each year, in units of cubic metres

```

Examples

```

# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

  # Add a column based on water years starting in August
  add_cumulative_volume(station_number = "08NM116",
                        water_year_start = 8)

}

```

add_cumulative_yield *Add daily cumulative runoff yield flows on an annual basis*

Description

Add a column of rolling daily cumulative runoff yield flows on an annual basis to a streamflow dataset. Adds the runoff yield discharge from each day with the previous day(s) for each year, in units of millimetres. Converts cumulative discharge to a depth of water based on the upstream drainage basin area. The cumulative flows restart every year and are only calculated in years with complete data.

Usage

```
add_cumulative_yield(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  basin_area,
  water_year_start = 1
)
```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.
groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using <code>station_number</code> argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires tidyhydat package and a HYDAT database. Leave blank if using data argument.
basin_area	Upstream drainage basin area, in square kilometres, to apply to observations. Three options:

- (1) Leave blank if groups is STATION_NUMBER with HYDAT station numbers to extract basin areas from HYDAT.
- (2) A single numeric value to apply to all observations.
- (3) List each basin area for each group/station in groups (can override HYDAT value if listed) as such `c("08NM116" = 795, "08NM242" = 10)`. If group is not listed the HYDAT area will be applied if it exists, otherwise it will be NA.

water_year_start

Numeric value indicating the month of the start of the water year for analysis.
Default 1.

Value

A tibble data frame of the source data with an additional column:

Cumul_Yield_mm cumulative yield flows for each day for each year, in units of millimetres

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Add a column based on water years starting in August
add_cumulative_yield(station_number = "08NM116",
                     water_year_start = 8)

# Add a column based on water years starting in August with a custom basin area to calculate yield
add_cumulative_yield(station_number = "08NM116",
                     water_year_start = 8,
                     basin_area = 800)

}
```

<code>add_daily_volume</code>	<i>Add daily volumetric flows</i>
-------------------------------	-----------------------------------

Description

Add a column of daily volumetric flows to a streamflow dataset, in units of cubic metres. Converts the discharge to a volume.

Usage

```
add_daily_volume(data, values = Value, station_number)
```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using station_number argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using station_number argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires tidyhydat package and a HYDAT database. Leave blank if using data argument.

Value

A tibble data frame of the source data with an additional column:

Volume_m3 daily total volumetric flow, in units of cubic metres

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Add a column of daily flow volumes
add_daily_volume(station_number = "08NM116")

}
```

add_daily_yield *Add daily volumetric runoff yields*

Description

Add a column of daily runoff yields to a streamflow dataset, in units of millimetres. Converts the discharge to a depth of water based on the upstream drainage basin area.

Usage

```
add_daily_yield(
  data,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  basin_area
)
```


Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.
groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using <code>station_number</code> argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires <code>tidyhydat</code> package and a HYDAT database. Leave blank if using <code>data</code> argument.
basin_area	Upstream drainage basin area, in square kilometres, to apply to observations. Three options: (1) Leave blank if groups is STATION_NUMBER with HYDAT station numbers to extract basin areas from HYDAT. (2) A single numeric value to apply to all observations. (3) List each basin area for each group/station in groups (can override HYDAT value if listed) as such <code>c("08NM116" = 795, "08NM242" = 10)</code> . If group is not listed the HYDAT area will be applied if it exists, otherwise it will be NA.

Value

A tibble data frame of the source data with an additional column:

Yield_mm daily runoff yield flow, in units of millimetres

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Add a column of yields based on HYDAT basin area
add_daily_yield(station_number = "08NM116")

# Add a column of yields based on a custom basin area
add_daily_yield(station_number = "08NM116",
                basin_area = 800)

}
```

add_date_variables *Add year, month, and day of year variables*

Description

Add columns of Year (YYYY), Month (MM), MonthName (e.g. 'Jan'), and DayofYear (1-365 or 366); and WaterYear (YYYY) and WaterDayofYear (1-365 or 366) if selected; to a data frame with a column of dates called 'Date'. Water years are designated by the year in which they end. For example, Water Year 1999 (starting Oct) is from 1 Oct 1998 (WaterDayofYear 1) to 30 Sep 1999 (WaterDayofYear 365).

Usage

```
add_date_variables(data, dates = Date, station_number, water_year_start = 1)
```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using station_number argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using station_number argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires tidyhydat package and a HYDAT database. Leave blank if using data argument.
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.

Value

A tibble data frame of the source data with additional columns:

CalendarYear	calendar year
Month	numeric month (1 to 12)
MonthName	month name (Jan-Dec)
WaterYear	year starting from the selected month start, water_year_start
DayofYear	day of the year from the selected month start (1-365 or 366)

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Add date variables using calendar years
add_date_variables(station_number = "08NM116")

# Add date variables using water years starting in August
add_date_variables(station_number = "08NM116",
                  water_year_start = 8)

}
```

add_rolling_means *Add rolling n-day averages*

Description

Adds selected n-day rolling means to a streamflow dataset. Based on selected n-days and alignment, the rolling mean for a given day is obtained by averaging the adjacent dates of daily mean values. For example, rolling days of '7' and 'right' alignment would obtain a mean of the given and previous 6 days of daily mean flow. Rolling mean values will not be calculated if there is less than the n-days provided.

Usage

```
add_rolling_means(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  roll_days = c(3, 7, 30),
  roll_align = "right"
)
```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using station_number argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using station_number argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using station_number argument.

groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using station_number argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires tidyhydat package and a HYDAT database. Leave blank if using data argument.
roll_days	Numeric values of the number of days to apply a rolling mean. Default c(3, 7, 30).
roll_align	Character string identifying the direction of the rolling mean from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'.

Value

A data frame of the source data with an additional column(s):

QnDay rolling means of the n-day flow values of the designated date and adjacent dates, direction of mean specified by roll_align

Default additional columns:

Q3Day rolling means of the 3-day flow values of the designated date and previous 2 days (roll_align = "right")

Q7Day rolling means of the 7-day flow values of the designated date and previous 6 days (roll_align = "right")

Q30Day rolling means of the 30-day flow values of the designated date and previous 29 days (roll_align = "right")

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Add default 3, 7, and 30-day rolling mean columns, with "right" alignment
add_rolling_means(station_number = "08NM116")

# Add custom 5 and 10-day rolling mean columns
add_rolling_means(station_number = "08NM116",
                  roll_days = c(5,10))

# Add default 3, 7, and 30-day rolling mean columns, with "left" alignment
add_rolling_means(station_number = "08NM116",
                  roll_align = "left")

}
```

add_seasons	<i>Add a column of seasons</i>
-------------	--------------------------------

Description

Adds a column of seasons identifiers to a data frame with a column of dates called 'Date'. The length of seasons, in months, is provided using the `seasons_length` argument. As seasons are grouped by months the length of the seasons must be divisible into 12 with one of the following season lengths: 1, 2, 3, 4, 6, or 12 months. The start of the first season coincides with the start month of each year; 'Jan-Jun' for 6-month seasons starting with calendar years or 'Dec-Feb' for 3-month seasons starting with water year starting in December.

Usage

```
add_seasons(
  data,
  dates = Date,
  station_number,
  water_year_start = 1,
  seasons_length
)
```

Arguments

<code>data</code>	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
<code>dates</code>	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
<code>station_number</code>	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires <code>tidyhydat</code> package and a HYDAT database. Leave blank if using <code>data</code> argument.
<code>water_year_start</code>	Numeric value indicating the month of the start of the water year for analysis. Default 1.
<code>seasons_length</code>	Numeric value indicating the desired length of seasons in months, divisible into 12. Required.

Value

A tibble data frame of the source data with additional column:

Season	season identifier labelled by the start and end month of the season
--------	---

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Add a column with four annual seasons starting in January
add_seasons(station_number = "08NM116",
             seasons_length = 4)

# Add a column with two annual seasons (of 6 months length) starting in October
add_seasons(station_number = "08NM116",
             water_year_start = 10,
             seasons_length = 6)

}
```

calc_all_annual_stats *Calculate all fasstr annual statistics*

Description

Calculates all annual statistics of daily flow values from a streamflow dataset from all annual fasstr functions. Calculates the statistics from all daily discharge values from all years, unless specified. Data is ideally long-term and continuous with minimal missing/seasonal data as annual statistics are calculated. Data calculated using the following functions:

- calc_annual_stats()
- calc_annual_lowflows()
- calc_annual_cumulative_stats()
- calc_annual_flow_timing()
- calc_monthly_stats()
- calc_annual_outside_normal()

Usage

```
calc_all_annual_stats(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  basin_area,
  water_year_start = 1,
  start_year,
  end_year,
  exclude_years,
  annual_percentiles = c(10, 90),
```

```

monthly_percentiles = c(10, 20),
stats_days = 1,
stats_align = "right",
lowflow_days = c(1, 3, 7, 30),
lowflow_align = "right",
timing_percent = c(25, 33, 50, 75),
normal_percentiles = c(25, 75),
transpose = FALSE,
ignore_missing = FALSE
)

```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.
groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using <code>station_number</code> argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires tidyhydat package and a HYDAT database. Leave blank if using <code>data</code> argument.
basin_area	Upstream drainage basin area, in square kilometres, to apply to observations. Three options: (1) Leave blank if groups is STATION_NUMBER with HYDAT station numbers to extract basin areas from HYDAT. (2) A single numeric value to apply to all observations. (3) List each basin area for each group/station in groups (can override HYDAT value if listed) as such <code>c("08NM116" = 795, "08NM242" = 10)</code> . If group is not listed the HYDAT area will be applied if it exists, otherwise it will be NA.
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.
start_year	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
end_year	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.

exclude_years	Numeric vector of years to exclude from analysis. Leave blank to include all years.
annual_percentiles	Numeric vector of percentiles to calculate annually. Set to NA if none required. Used for calc_annual_stats() function. Default c(10, 90).
monthly_percentiles	Numeric vector of percentiles to calculate monthly for each year. Set to NA if none required. Used for calc_monthly_stats() function. Default c(10, 20).
stats_days	Numeric vector of the number of days to apply a rolling mean on basic stats. Default c(1). Used for calc_annual_stats() and calc_monthly_stats() functions.
stats_align	Character string identifying the direction of the rolling mean on basic stats from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'. Used for calc_annual_stats(), calc_monthly_stats(), and calc_annual_outside_normal() functions.
lowflow_days	Numeric vector of the number of days to apply a rolling mean on lowflow stats. Default c(1, 3, 7, 30). Used for calc_lowflow_stats() function.
lowflow_align	Character string identifying the direction of the rolling mean on lowflow stats from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'. Used for calc_lowflow_stats() function.
timing_percent	Numeric vector of percents of annual total flows to determine dates. Used for calc_annual_flow_timing() function. Default c(25, 33.3, 50, 75).
normal_percentiles	Numeric vector of two values, lower and upper percentiles, respectively indicating the limits of the normal range. Default c(25, 75).
transpose	Logical value indicating if the results rows and columns are to be switched. Default FALSE.
ignore_missing	Logical value indicating whether dates with missing values should be included in the calculation. If TRUE then a statistic will be calculated regardless of missing dates. If FALSE then only statistics from time periods with no missing dates will be returned. Default FALSE.

Value

A tibble data frame with column "Year" and then 107 (default) variables from the fasstr annual functions. See listed functions above for default variables. Transposing data creates a column of "Statistics" and subsequent columns for each year selected.

See Also

[calc_annual_stats](#), [calc_annual_lowflows](#), [calc_annual_cumulative_stats](#), [calc_annual_flow_timing](#), [calc_monthly_stats](#), [calc_annual_outside_normal](#)

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Calculate all annual statistics from this package with default arguments
calc_all_annual_stats(station_number = "08NM116")

# Calculate all annual statistics from this package with default arguments
# with some default arguments shown to customize metrics
calc_all_annual_stats(station_number = "08NM116",
                      annual_percentiles = c(10,90),
                      monthly_percentiles = c(10,20),
                      stats_days = 1,
                      stats_align = "right",
                      lowflow_days = c(1,3,7,30),
                      lowflow_align = "right",
                      timing_percent = c(25,33,50,75),
                      normal_percentiles = c(25,75))

}
```

calc_annual_cumulative_stats

Calculate annual (and seasonal) cumulative flows

Description

Calculates annual and seasonal total flows, volumetric or runoff yield flows, from a streamflow dataset. Calculates the statistics from all daily discharge values from all years, unless specified. For water year and seasonal data, the designated year is the year in which the year or season ends. Two-seasons and four-seasons per year are calculated, with each 6 and 3-month seasons starting with the first month of the year (Jan for calendar year, specified for water year). Each season is designated by the calendar or water year in which it occurs.

Usage

```
calc_annual_cumulative_stats(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  use_yield = FALSE,
  basin_area,
  water_year_start = 1,
  start_year,
  end_year,
  exclude_years,
```

```

months = 1:12,
include_seasons = FALSE,
transpose = FALSE
)

```

Arguments

<code>data</code>	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
<code>dates</code>	Column in <code>data</code> that contains dates formatted YYYY-MM-DD. Only required if <code>dates</code> column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
<code>values</code>	Column in <code>data</code> that contains numeric flow values, in units of cubic metres per second. Only required if <code>values</code> column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.
<code>groups</code>	Column in <code>data</code> that contains unique identifiers for different data sets, if applicable. Only required if <code>groups</code> column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using <code>station_number</code> argument.
<code>station_number</code>	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires <code>tidyhydat</code> package and a HYDAT database. Leave blank if using <code>data</code> argument.
<code>use_yield</code>	Logical value indicating whether to use yield runoff, in mm, instead of volumetric. Default FALSE.
<code>basin_area</code>	Upstream drainage basin area, in square kilometres, to apply to observations. Three options: (1) Leave blank if <code>groups</code> is STATION_NUMBER with HYDAT station numbers to extract basin areas from HYDAT. (2) A single numeric value to apply to all observations. (3) List each basin area for each group/station in <code>groups</code> (can override HYDAT value if listed) as such <code>c("08NM116" = 795, "08NM242" = 10)</code> . If group is not listed the HYDAT area will be applied if it exists, otherwise it will be NA.
<code>water_year_start</code>	Numeric value indicating the month of the start of the water year for analysis. Default 1.
<code>start_year</code>	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
<code>end_year</code>	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
<code>exclude_years</code>	Numeric vector of years to exclude from analysis. Leave blank to include all years.
<code>months</code>	Numeric vector of months to include in analysis (e.g.. 6:8 for Jun-Aug). Leave blank to summarize all months (default 1:12).

include_seasons	Logical value indication whether to include seasonal yields and total discharges. Default TRUE.
transpose	Logical value indicating if the results rows and columns are to be switched. Default FALSE.

Value

A tibble data frame with the following columns, ending with `'_Volume_m3'` or `'_Yield_mm'` based on selection:

Year	calendar or water year selected
Total_*	annual (or selected months) total flow, in m3 or mm

Default seasonal columns:

MMM-MMM_*	first of two season total flows, in m3 or mm
MMM-MMM_*	second of two season total flows, in m3 or mm
MMM-MMM_*	first of four season total flows, in m3 or mm
MMM-MMM_*	second of four season total flows, in m3 or mm
MMM-MMM_*	third of four season total flows, in m3 or mm
MMM-MMM_*	fourth of four season total flows, in m3 or mm

Transposing data creates a column of `'Statistics'` and subsequent columns for each year selected.

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Calculate annual total volumetric flow statistics
calc_annual_cumulative_stats(station_number = "08NM116")

# Calculate annual total yield statistics with default HYDAT basin area
calc_annual_cumulative_stats(station_number = "08NM116",
                             use_yield = TRUE)

# Calculate annual total yield statistics with a custom basin area
calc_annual_cumulative_stats(station_number = "08NM116",
                             use_yield = TRUE,
                             basin_area = 800)

}
```

 calc_annual_flow_timing

Calculate annual flow timing

Description

Calculates annual the timing (day of year) and date of occurrence of portions of total annual flow of daily flow values from a streamflow dataset. Calculates the statistics from all daily discharge values from all years with complete annual data, unless specified.

Usage

```
calc_annual_flow_timing(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  percent_total = c(25, 33.3, 50, 75),
  water_year_start = 1,
  start_year,
  end_year,
  exclude_years,
  transpose = FALSE
)
```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using station_number argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using station_number argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using station_number argument.
groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using station_number argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires tidyhydat package and a HYDAT database. Leave blank if using data argument.

percent_total	Numeric vector of percents of total annual flows to determine dates. Default c(25, 33.3, 50, 75).
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.
start_year	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
end_year	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
exclude_years	Numeric vector of years to exclude from analysis. Leave blank to include all years.
transpose	Logical value indicating if the results rows and columns are to be switched. Default FALSE.

Value

A tibble data frame with the following columns:

Year	calendar or water year selected
DoY_'n'pct_TotalQ	day of year for each n-percent of total volumetric discharge
Date_'n'pct_TotalQ	date (YYYY-MM-DD) for each n-percent of total volumetric discharge

Default columns:

DoY_25pct_TotalQ	day of year of 25-percent of total volumetric discharge
Date_25pct_TotalQ	date (YYYY-MM-DD) of 25-percent of total volumetric discharge
DoY_33.3pct_TotalQ	day of year of 33.3-percent of total volumetric discharge
Date_33.3pct_TotalQ	date (YYYY-MM-DD) of 33.3-percent of total volumetric discharge
DoY_50pct_TotalQ	day of year of 50-percent of total volumetric discharge
Date_50pct_TotalQ	date (YYYY-MM-DD) of 50-percent of total volumetric discharge
DoY_75pct_TotalQ	day of year of 75-percent of total volumetric discharge
Date_75pct_TotalQ	date (YYYY-MM-DD) of 75-percent of total volumetric discharge

Transposing data creates a column of 'Statistics' (just DoY, not Date values) and subsequent columns for each year selected.

References

- Barnett, T.P., Pierce, D.W., Hidalgo, H.G., Bonfils, C., Santer, B.D., Das, T., Bala, G., Wood, A.W., Nozawa, T., Mirin, A.A., Cayan, D.R., Dettinger, M.D., 2008. Human-Induced Changes in the Hydrology of the Western United States. *Science* 319, 1080-1083.

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Calculate annual flow timings with default percent of annual totals
calc_annual_flow_timing(station_number = "08NM116")

# Calculate annual flow timings with custom percent of annual totals
calc_annual_flow_timing(station_number = "08NM116",
                        percent_total = 50)

}
```

calc_annual_lowflows *Calculate annual lowflows*

Description

Calculates annual n-day minimum values, and the day of year and date of occurrence of daily flow values from a streamflow dataset. Calculates the statistics from all daily discharge values from all years, unless specified.

Usage

```
calc_annual_lowflows(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  roll_days = c(1, 3, 7, 30),
  roll_align = "right",
  water_year_start = 1,
  start_year,
  end_year,
  exclude_years,
  months = 1:12,
  transpose = FALSE,
  ignore_missing = FALSE
)
```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.
groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using <code>station_number</code> argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires <code>tidyhydat</code> package and a HYDAT database. Leave blank if using <code>data</code> argument.
roll_days	Numeric value of the number of days to apply a rolling mean. Default 1.
roll_align	Character string identifying the direction of the rolling mean from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'.
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.
start_year	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
end_year	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
exclude_years	Numeric vector of years to exclude from analysis. Leave blank to include all years.
months	Numeric vector of months to include in analysis (e.g.. 6:8 for Jun-Aug). Leave blank to summarize all months (default 1:12).
transpose	Logical value indicating if the results rows and columns are to be switched. Default FALSE.
ignore_missing	Logical value indicating whether dates with missing values should be included in the calculation. If TRUE then a statistic will be calculated regardless of missing dates. If FALSE then only statistics from time periods with no missing dates will be returned. Default FALSE.

Value

A tibble data frame with the following columns:

Year calendar or water year selected

Min_'n'_Day annual minimum for each n-day rolling mean, direction of mean specified by roll_align

Min_'n'_Day_DoY day of year for each annual minimum of n-day rolling mean

Min_'n'_Day_Date date (YYYY-MM-DD) for each annual minimum of n-day rolling mean

Default columns:

Min_1_Day annual 1-day mean minimum (roll_align = right)

Min_1_Day_DoY day of year of annual 1-day mean minimum

Min_1_Day_Date date (YYYY-MM-DD) of annual 1-day mean minimum

Min_3_Day annual 3-day mean minimum (roll_align = right)

Min_3_Day_DoY day of year of annual 3-day mean minimum

Min_3_Day_Date date (YYYY-MM-DD) of annual 3-day mean minimum

Min_7_Day annual 7-day mean minimum (roll_align = right)

Min_7_Day_DoY day of year of annual 7-day mean minimum

Min_7_Day_Date date (YYYY-MM-DD) of annual 7-day mean minimum

Min_30_Day annual 30-day mean minimum (roll_align = right)

Min_30_Day_DoY day of year of annual 30-day mean minimum

Min_30_Day_Date date (YYYY-MM-DD) of annual 30-day mean minimum

Transposing data creates a column of 'Statistics' and subsequent columns for each year selected. 'Date' statistics not transposed.

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Calculate annual 1, 3, 7, and 30-day (default) low flows with
# default alignment ('right')
calc_annual_lowflows(station_number = "08NM116")

# Calculate custom 3 and 7-day annual low flows with 'center' alignment
calc_annual_lowflows(station_number = "08NM116",
                      roll_days = c(3,7),
                      roll_align = "center")

}
```

 calc_annual_outside_normal

Calculate annual days above and below normal

Description

Calculates the number of days per year outside of the 'normal' range (typically between 25 and 75th percentiles) for each day of the year. Upper and lower-range percentiles are calculated for each day of the year of from all years, and then each daily flow value for each year is compared. All days above or below the normal range are included. Calculates the statistics from all daily discharge values from all years, unless specified.

Usage

```
calc_annual_outside_normal(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  normal_percentiles = c(25, 75),
  roll_days = 1,
  roll_align = "right",
  water_year_start = 1,
  start_year,
  end_year,
  exclude_years,
  months = 1:12,
  transpose = FALSE
)
```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if <code>dates</code> column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if <code>values</code> column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.
groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if <code>groups</code> column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using <code>station_number</code> argument.

<code>station_number</code>	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires tidyhydat package and a HYDAT database. Leave blank if using data argument.
<code>normal_percentiles</code>	Numeric vector of two values, lower and upper percentiles, respectively indicating the limits of the normal range. Default <code>c(25, 75)</code> .
<code>roll_days</code>	Numeric value of the number of days to apply a rolling mean. Default 1.
<code>roll_align</code>	Character string identifying the direction of the rolling mean from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'.
<code>water_year_start</code>	Numeric value indicating the month of the start of the water year for analysis. Default 1.
<code>start_year</code>	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
<code>end_year</code>	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
<code>exclude_years</code>	Numeric vector of years to exclude from analysis. Leave blank to include all years.
<code>months</code>	Numeric vector of months to include in analysis (e.g.. 6:8 for Jun-Aug). Leave blank to summarize all months (default 1:12).
<code>transpose</code>	Logical value indicating if the results rows and columns are to be switched. Default FALSE.

Value

A tibble data frame with the following columns:

<code>Year</code>	calendar or water year selected
<code>Days_Below_Normal</code>	number of days per year below the daily normal (default 25th percentile)
<code>Days_Above_Normal</code>	number of days per year above the daily normal (default 75th percentile)
<code>Days_Outside_Normal</code>	number of days per year below and above the daily normal (default 25/75th percentile)

Transposing data creates a column of "Statistics" and subsequent columns for each year selected.

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Calculate statistics with default limits of normal (25 and 75th percentiles)
calc_annual_outside_normal(station_number = "08NM116")
```

```
# Calculate statistics with custom limits of normal
calc_annual_outside_normal(station_number = "08NM116",
                           normal_percentiles = c(10,90))

}
```

calc_annual_stats *Calculate annual summary statistics*

Description

Calculates annual mean, median, maximum, minimum, and percentiles of daily flow values from a streamflow dataset. Calculates the statistics from all daily discharge values from all years, unless specified.

Usage

```
calc_annual_stats(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  roll_days = 1,
  roll_align = "right",
  percentiles = c(10, 90),
  water_year_start = 1,
  start_year,
  end_year,
  exclude_years,
  months = 1:12,
  transpose = FALSE,
  ignore_missing = FALSE
)
```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using station_number argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using station_number argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using station_number argument.

groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using station_number argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires tidyhydat package and a HYDAT database. Leave blank if using data argument.
roll_days	Numeric value of the number of days to apply a rolling mean. Default 1.
roll_align	Character string identifying the direction of the rolling mean from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'.
percentiles	Numeric vector of percentiles to calculate. Set to NA if none required. Default c(10,90).
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.
start_year	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
end_year	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
exclude_years	Numeric vector of years to exclude from analysis. Leave blank to include all years.
months	Numeric vector of months to include in analysis (e.g.. 6:8 for Jun-Aug). Leave blank to summarize all months (default 1:12).
transpose	Logical value indicating if the results rows and columns are to be switched. Default FALSE.
ignore_missing	Logical value indicating whether dates with missing values should be included in the calculation. If TRUE then a statistic will be calculated regardless of missing dates. If FALSE then only statistics from time periods with no missing dates will be returned. Default FALSE.

Value

A tibble data frame with the following columns:

Year	calendar or water year selected
Mean	annual mean of all daily flows for a given year
Median	annual median of all daily flows for a given year
Maximum	annual maximum of all daily flows for a given year
Minimum	annual minimum of all daily flows for a given year
P'n'	each annual n-th percentile selected of all daily flows

Default percentile columns:

P10 annual 10th percentile of all daily flows for a given year
 P90 annual 90th percentile of all daily flows for a given year

Transposing data creates a column of "Statistics" and subsequent columns for each year selected.

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Calculate annual statistics from a data frame using the data argument
flow_data <- tidyhydat::hy_daily_flows(station_number = "08NM116")
calc_annual_stats(data = flow_data)

# Calculate annual statistics using station_number argument
calc_annual_stats(station_number = "08NM116")

# Calculate annual statistics regardless if there
# is missing data for a given year
calc_annual_stats(station_number = "08NM116",
                  ignore_missing = TRUE)

# Calculate annual statistics for water years starting in October
calc_annual_stats(station_number = "08NM116",
                  water_year_start = 10)

# Calculate annual statistics filtered for custom years
calc_annual_stats(station_number = "08NM116",
                  start_year = 1981,
                  end_year = 2010,
                  exclude_years = c(1991,1993:1995))

# Calculate annual statistics for 7-day flows for July-September
# months only, with 25 and 75th percentiles
calc_annual_stats(station_number = "08NM116",
                  roll_days = 7,
                  months = 7:9,
                  percentiles = c(25,75))

}
```

calc_daily_cumulative_stats

Calculate cumulative daily flow statistics

Description

Calculate cumulative daily flow statistics for each day of the year of daily flow values from a stream-flow dataset. Calculates the statistics from all daily discharge values from all years, unless specified. Defaults to volumetric cumulative flows, can use `use_yield` and `basin_area` to convert to runoff yield.

Usage

```
calc_daily_cumulative_stats(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  percentiles = c(5, 25, 75, 95),
  use_yield = FALSE,
  basin_area,
  water_year_start = 1,
  start_year,
  end_year,
  exclude_years,
  transpose = FALSE
)
```

Arguments

<code>data</code>	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
<code>dates</code>	Column in <code>data</code> that contains dates formatted YYYY-MM-DD. Only required if <code>dates</code> column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
<code>values</code>	Column in <code>data</code> that contains numeric flow values, in units of cubic metres per second. Only required if <code>values</code> column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.
<code>groups</code>	Column in <code>data</code> that contains unique identifiers for different data sets, if applicable. Only required if <code>groups</code> column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using <code>station_number</code> argument.
<code>station_number</code>	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires <code>tidyhydat</code> package and a HYDAT database. Leave blank if using <code>data</code> argument.
<code>percentiles</code>	Numeric vector of percentiles to calculate. Set to NA if none required. Default <code>c(5, 25, 75, 95)</code> .
<code>use_yield</code>	Logical value indicating whether to use yield runoff, in mm, instead of volumetric. Default FALSE.
<code>basin_area</code>	Upstream drainage basin area, in square kilometres, to apply to observations. Three options: (1) Leave blank if <code>groups</code> is <code>STATION_NUMBER</code> with HYDAT station numbers to extract basin areas from HYDAT. (2) A single numeric value to apply to all observations.

(3) List each basin area for each group/station in groups (can override HYDAT value if listed) as such `c("08NM116" = 795, "08NM242" = 10)`. If group is not listed the HYDAT area will be applied if it exists, otherwise it will be NA.

<code>water_year_start</code>	Numeric value indicating the month of the start of the water year for analysis. Default 1.
<code>start_year</code>	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
<code>end_year</code>	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
<code>exclude_years</code>	Numeric vector of years to exclude from analysis. Leave blank to include all years.
<code>transpose</code>	Logical value indicating if the results rows and columns are to be switched. Default FALSE.

Value

A data frame with the following columns, default units in cubic metres, millimetres if `use_yield` and `basin_area` provided:

Date	date (MMM-DD) of daily cumulative statistics
DayofYear	day of year of daily cumulative statistics
Mean	daily mean of all cumulative flows for a given day of the year
Median	daily mean of all cumulative flows for a given day of the year
Maximum	daily mean of all cumulative flows for a given day of the year
Minimum	daily mean of all cumulative flows for a given day of the year
P'n'	each daily n-th percentile selected of all cumulative flows for a given day of the year

Default percentile columns:

P5	daily 5th percentile of all cumulative flows for a given day of the year
P25	daily 25th percentile of all cumulative flows for a given day of the year
P75	daily 75th percentile of all cumulative flows for a given day of the year
P95	daily 95th percentile of all cumulative flows for a given day of the year

Transposing data creates a column of "Statistics" and subsequent columns for each year selected.

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Calculate annual daily cumulative volume statistics
calc_daily_cumulative_stats(station_number = "08NM116")

# Calculate annual daily cumulative yield statistics
```

```

# with default HYDAT basin area
calc_daily_cumulative_stats(station_number = "08NM116",
                             use_yield = TRUE)

# Calculate annual daily cumulative yield statistics
# with custom basin area
calc_daily_cumulative_stats(station_number = "08NM116",
                             use_yield = TRUE,
                             basin_area = 800)

}

```

calc_daily_stats	<i>Calculate daily summary statistics</i>
------------------	---

Description

Calculates the daily mean, median, maximum, minimum, and percentiles for each day of the year of daily flow values from a streamflow dataset. Calculate the statistics from all daily discharge values from all years, unless specified. Can determine statistics of rolling mean days (e.g. 7-day flows) using the `roll_days` argument. Note that statistics are based on the numeric days of year (1-365) and not the date (Jan 1 - Dec 31) of year so day of year values for days after Feb 29 in leap years will be one value higher than non-leap years.

Usage

```

calc_daily_stats(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  percentiles = c(5, 25, 75, 95),
  roll_days = 1,
  roll_align = "right",
  water_year_start = 1,
  start_year,
  end_year,
  exclude_years,
  complete_years = FALSE,
  months = 1:12,
  transpose = FALSE,
  ignore_missing = FALSE
)

```


Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using station_number argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using station_number argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using station_number argument.
groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using station_number argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires tidyhydat package and a HYDAT database. Leave blank if using data argument.
percentiles	Numeric vector of percentiles to calculate. Set to NA if none required. Default c(5, 25, 75, 95).
roll_days	Numeric value of the number of days to apply a rolling mean. Default 1.
roll_align	Character string identifying the direction of the rolling mean from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'.
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.
start_year	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
end_year	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
exclude_years	Numeric vector of years to exclude from analysis. Leave blank to include all years.
complete_years	Logical values indicating whether to include only years with complete data in analysis. Default FALSE.
months	Numeric vector of months to include in analysis (e.g.. 6:8 for Jun-Aug). Leave blank to summarize all months (default 1:12).
transpose	Logical value indicating if the results rows and columns are to be switched. Default FALSE.
ignore_missing	Logical value indicating whether dates with missing values should be included in the calculation. If TRUE then a statistic will be calculated regardless of missing dates. If FALSE then only statistics from time periods with no missing dates will be returned. Default FALSE.

Value

A tibble data frame with the following columns:

Date	date (MMM-DD) of daily statistics
DayofYear	day of year of daily statistics
Mean	daily mean of all flows for a given day of the year
Median	daily mean of all flows for a given day of the year
Maximum	daily mean of all flows for a given day of the year
Minimum	daily mean of all flows for a given day of the year
P'n'	each daily n-th percentile selected of all flows for a given day of the year

Default percentile columns:

P5	daily 5th percentile of all flows for a given day of the year
P25	daily 25th percentile of all flows for a given day of the year
P75	daily 75th percentile of all flows for a given day of the year
P95	daily 95th percentile of all flows for a given day of the year

Transposing data creates a column of "Statistics" and subsequent columns for each year selected.

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Calculate daily statistics using data argument with defaults
flow_data <- tidyhydat::hy_daily_flows(station_number = "08NM116")
calc_daily_stats(data = flow_data,
                 start_year = 1980)

# Calculate daily statistics using station_number argument with defaults
calc_daily_stats(station_number = "08NM116",
                 start_year = 1980)

# Calculate daily statistics regardless if there is missing data for a given day of year
calc_daily_stats(station_number = "08NM116",
                 ignore_missing = TRUE)

# Calculate daily statistics using only years with no missing data
calc_daily_stats(station_number = "08NM116",
                 complete_years = TRUE)

# Calculate daily statistics for water years starting in October between 1980 and 2010
calc_daily_stats(station_number = "08NM116",
                 start_year = 1980,
                 end_year = 2010,
                 water_year_start = 10)

# Calculate daily statistics with custom years and removing certain years
```

```

calc_daily_stats(station_number = "08NM116",
                 start_year = 1981,
                 end_year = 2010,
                 exclude_years = c(1991,1993:1995))

# Calculate daily statistics for 7-day flows for July-September months only,
# with 25 and 75th percentiles starting in 1980
calc_daily_stats(station_number = "08NM116",
                 start_year = 1980,
                 roll_days = 7,
                 months = 7:9,
                 percentiles = c(25,75))

}

```

calc_flow_percentile *Calculate the percentile rank of a flow value*

Description

Calculates the percentile, or percentile rank, of a discharge value compared to all flow values of a streamflow dataset. Looks up the value in the distribution (stats::ecdf() function) of all daily discharge values from all years, unless specified.

Usage

```

calc_flow_percentile(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  roll_days = 1,
  roll_align = "right",
  flow_value,
  water_year_start = 1,
  start_year,
  end_year,
  exclude_years,
  complete_years = FALSE,
  months = 1:12
)

```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using station_number argument.
------	---

dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using station_number argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using station_number argument.
groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using station_number argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires tidyhydat package and a HYDAT database. Leave blank if using data argument.
roll_days	Numeric value of the number of days to apply a rolling mean. Default 1.
roll_align	Character string identifying the direction of the rolling mean from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'.
flow_value	A numeric flow value of which to determine the percentile rank. Required.
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.
start_year	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
end_year	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
exclude_years	Numeric vector of years to exclude from analysis. Leave blank to include all years.
complete_years	Logical values indicating whether to include only years with complete data in analysis. Default FALSE.
months	Numeric vector of months to include in analysis (e.g.. 6:8 for Jun-Aug). Leave blank to summarize all months (default 1:12).

Value

A tibble data frame, or a single numeric value if no station number provided, of the percentile rank of a given flow value.

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Calculate the percentile rank of a 10-cms flow value from a full record
calc_flow_percentile(station_number = "08NM116",
```

```

        flow_value = 10)

# Calculate the percentile rank of a 10-cms flow value from years with no missing data
calc_flow_percentile(station_number = "08NM116",
                    complete_years = TRUE,
                    flow_value = 10)

# Calculate the percentile rank of a 10-cms flow value for June from years with no missing data
calc_flow_percentile(station_number = "08NM116",
                    complete_years = TRUE,
                    months = 6,
                    flow_value = 10)

}

```

calc_longterm_daily_stats

Calculate long-term summary statistics from daily mean flows

Description

Calculates the long-term and long-term monthly mean, median, maximum, minimum, and percentiles of daily flow values from a streamflow dataset. Calculates the statistics from all daily values from all years, unless specified.

Usage

```

calc_longterm_daily_stats(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  percentiles = c(10, 90),
  roll_days = 1,
  roll_align = "right",
  water_year_start = 1,
  start_year,
  end_year,
  exclude_years,
  months = 1:12,
  complete_years = FALSE,
  include_longterm = TRUE,
  custom_months,
  custom_months_label,
  transpose = FALSE,
  ignore_missing = FALSE
)

```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using station_number argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using station_number argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using station_number argument.
groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using station_number argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires tidyhydat package and a HYDAT database. Leave blank if using data argument.
percentiles	Numeric vector of percentiles to calculate. Set to NA if none required. Default c(10,90).
roll_days	Numeric value of the number of days to apply a rolling mean. Default 1.
roll_align	Character string identifying the direction of the rolling mean from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'.
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.
start_year	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
end_year	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
exclude_years	Numeric vector of years to exclude from analysis. Leave blank to include all years.
months	Numeric vector of months to include in analysis (e.g.. 6:8 for Jun-Aug). Leave blank to summarize all months (default 1:12).
complete_years	Logical values indicating whether to include only years with complete data in analysis. Default FALSE.
include_longterm	Logical value indicating whether to include longterm calculation of all data. Default TRUE.
custom_months	Numeric vector of months to combine to summarize (ex. 6:8 for Jun-Aug). Adds results to the end of table. If wanting months that overlap calendar years (ex. Oct-Mar), choose water_year_start that begins before the first month listed. Leave blank for no custom month summary.

custom_months_label	Character string to label custom months. For example, if choosing months 7:9 you may choose "Summer" or "Jul-Sep". Default "Custom-Months".
transpose	Logical value indicating if the results rows and columns are to be switched. Default FALSE.
ignore_missing	Logical value indicating whether dates with missing values should be included in the calculation. If TRUE then a statistic will be calculated regardless of missing dates. If FALSE then only statistics from time periods with no missing dates will be returned. Default FALSE.

Value

A tibble data frame with the following columns:

Month	month of the year, included 'Long-term' for all months, and 'Custom-Months' if selected
Mean	mean of all daily data for a given month and long-term over all years
Median	median of all daily data for a given month and long-term over all years
Maximum	maximum of all daily data for a given month and long-term over all years
Minimum	minimum of all daily data for a given month and long-term over all years
P'n'	each n-th percentile selected for a given month and long-term over all years

Default percentile columns:

P10	annual 10th percentile selected for a given month and long-term over all years
P90	annual 90th percentile selected for a given month and long-term over all years

Transposing data creates a column of "Statistics" and subsequent columns for each year selected.

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Calculate long-term statistics using data argument with defaults
flow_data <- tidyhydat::hy_daily_flows(station_number = "08NM116")
calc_longterm_daily_stats(data = flow_data,
                           start_year = 1980)

# Calculate long-term statistics using station_number argument with defaults
calc_longterm_daily_stats(station_number = "08NM116",
                           start_year = 1980)

# Calculate long-term statistics regardless if there is missing data for a given year
calc_longterm_daily_stats(station_number = "08NM116",
                           ignore_missing = TRUE)

# Calculate long-term statistics for water years starting in October
calc_longterm_daily_stats(station_number = "08NM116",
                           start_year = 1980,
```

```

        water_year_start = 10)

# Calculate long-term statistics with custom years
calc_longterm_daily_stats(station_number = "08NM116",
                          start_year = 1981,
                          end_year = 2010,
                          exclude_years = c(1991,1993:1995))

# Calculate long-term statistics for 7-day flows for July-September months only,
# with 25 and 75th percentiles
calc_longterm_daily_stats(station_number = "08NM116",
                          roll_days = 7,
                          months = 7:9,
                          percentiles = c(25,75),
                          ignore_missing = TRUE,
                          include_longterm = FALSE) # removes the Long-term numbers

# Calculate long-term statistics and add custom stats for July-September
calc_longterm_daily_stats(station_number = "08NM116",
                          start_year = 1980,
                          custom_months = 7:9,
                          custom_months_label = "Summer")

}

```

calc_longterm_mean *Calculate the long-term mean annual discharge*

Description

Calculates the long-term mean annual discharge of a streamflow dataset. Averages all daily discharge values from all years, unless specified.

Usage

```

calc_longterm_mean(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  roll_days = 1,
  roll_align = "right",
  water_year_start = 1,
  start_year,
  end_year,
  exclude_years,
  complete_years = FALSE,
  months = 1:12,

```



```

    percent_MAD,
    transpose = FALSE
  )

```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using station_number argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using station_number argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using station_number argument.
groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using station_number argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires tidyhydat package and a HYDAT database. Leave blank if using data argument.
roll_days	Numeric value of the number of days to apply a rolling mean. Default 1.
roll_align	Character string identifying the direction of the rolling mean from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'.
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.
start_year	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
end_year	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
exclude_years	Numeric vector of years to exclude from analysis. Leave blank to include all years.
complete_years	Logical values indicating whether to include only years with complete data in analysis. Default FALSE.
months	Numeric vector of months to include in analysis (e.g.. 6:8 for Jun-Aug). Leave blank to summarize all months (default 1:12).
percent_MAD	Numeric vector of percents of long-term mean annual discharge to add to the table (ex. 20 for 20 percent MAD). Leave blank for no values to be calculated.
transpose	Logical value indicating if the results rows and columns are to be switched. Default FALSE.

Value

A tibble data frame of numeric values of a long-term mean (and percent of long-term mean if selected) of selected years and months.

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Calculate the long-term mean annual discharge (MAD) using only years with no missing data
calc_longterm_mean(station_number = "08NM116",
                   complete_years = TRUE)

# Calculate the long-term MAD and 5, 10 and 20-percent MADs using only years with no missing data
calc_longterm_mean(station_number = "08NM116",
                   complete_years = TRUE,
                   percent_MAD = c(5,10,20))

}
```

calc_longterm_monthly_stats

Calculate long-term summary statistics from annual monthly mean flows

Description

Calculates the long-term and long-term monthly mean, median, maximum, minimum, and percentiles of monthly mean flow values from a streamflow dataset. Calculates the statistics from all annual monthly mean values from all years, unless specified.

Usage

```
calc_longterm_monthly_stats(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  percentiles = c(10, 90),
  roll_days = 1,
  roll_align = "right",
  water_year_start = 1,
  start_year,
  end_year,
  exclude_years,
  months = 1:12,
  complete_years = FALSE,
```

```

include_annual = TRUE,
custom_months,
custom_months_label,
transpose = FALSE,
ignore_missing = FALSE
)

```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using station_number argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using station_number argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using station_number argument.
groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using station_number argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires tidyhydat package and a HYDAT database. Leave blank if using data argument.
percentiles	Numeric vector of percentiles to calculate. Set to NA if none required. Default c(10,90).
roll_days	Numeric value of the number of days to apply a rolling mean. Default 1.
roll_align	Character string identifying the direction of the rolling mean from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'.
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.
start_year	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
end_year	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
exclude_years	Numeric vector of years to exclude from analysis. Leave blank to include all years.
months	Numeric vector of months to include in analysis (e.g.. 6:8 for Jun-Aug). Leave blank to summarize all months (default 1:12).

complete_years	Logical values indicating whether to include only years with complete data in analysis. Default FALSE.
include_annual	Logical value indicating whether to include annual calculation of all months. Default TRUE.
custom_months	Numeric vector of months to combine to summarize (ex. 6:8 for Jun-Aug). Adds results to the end of table. If wanting months that overlap calendar years (ex. Oct-Mar), choose water_year_start that begins before the first month listed. Leave blank for no custom month summary.
custom_months_label	Character string to label custom months. For example, if choosing months 7:9 you may choose "Summer" or "Jul-Sep". Default "Custom-Months".
transpose	Logical value indicating if the results rows and columns are to be switched. Default FALSE.
ignore_missing	Logical value indicating whether dates with missing values should be included in the calculation. If TRUE then a statistic will be calculated regardless of missing dates. If FALSE then only statistics from time periods with no missing dates will be returned. Default FALSE.

Value

A tibble data frame with the following columns:

Month	month of the year, included 'Annual' for all months, and 'Custom-Months' if selected
Mean	mean of all annual monthly means for a given month over all years
Median	median of all annual monthly means for a given month over all years
Maximum	maximum of all annual monthly means for a given month over all years
Minimum	minimum of all annual monthly means for a given month over all years
P'n'	each n-th percentile selected for annual monthly means for a given month over all years

Default percentile columns:

P10	annual 10th percentile selected for annual monthly means for a given month over all years
P90	annual 90th percentile selected for annual monthly means for a given month over all years

Transposing data creates a column of "Statistics" and subsequent columns for each year selected.

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Calculate long-term monthly statistics using data argument with defaults
flow_data <- tidyhydat::hy_daily_flows(station_number = "08NM116")
calc_longterm_monthly_stats(data = flow_data,
```

```

        start_year = 1980)

# Calculate long-term monthly statistics using station_number argument with defaults
calc_longterm_monthly_stats(station_number = "08NM116",
                           start_year = 1980)

# Calculate long-term monthly statistics regardless if there is missing data for a given year
calc_longterm_monthly_stats(station_number = "08NM116",
                           ignore_missing = TRUE)

# Calculate long-term monthly statistics for water years starting in October
calc_longterm_monthly_stats(station_number = "08NM116",
                           start_year = 1980,
                           water_year_start = 10)

# Calculate long-term monthly statistics with custom years
calc_longterm_monthly_stats(station_number = "08NM116",
                           start_year = 1981,
                           end_year = 2010,
                           exclude_years = c(1991,1993:1995))

# Calculate long-term monthly statistics for 7-day flows for July-September months only,
# with 25 and 75th percentiles
calc_longterm_monthly_stats(station_number = "08NM116",
                           roll_days = 7,
                           months = 7:9,
                           percentiles = c(25,75),
                           ignore_missing = TRUE,
                           include_annual = FALSE) # removes the Long-term numbers

# Calculate long-term monthly statistics and add custom stats for July-September
calc_longterm_monthly_stats(station_number = "08NM116",
                           start_year = 1980,
                           custom_months = 7:9,
                           custom_months_label = "Summer")

}

```

calc_longterm_percentile

Calculate the long-term percentiles

Description

Calculates the long-term percentiles discharge of a streamflow dataset. Averages all daily discharge values from all years, unless specified.

Usage

```
calc_longterm_percentile(
```

```

data,
dates = Date,
values = Value,
groups = STATION_NUMBER,
station_number,
percentiles,
roll_days = 1,
roll_align = "right",
water_year_start = 1,
start_year,
end_year,
exclude_years,
complete_years = FALSE,
months = 1:12,
transpose = FALSE
)

```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.
groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using <code>station_number</code> argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires <code>tidyhydat</code> package and a HYDAT database. Leave blank if using <code>data</code> argument.
percentiles	Numeric vector of percentiles to calculate. Required.
roll_days	Numeric value of the number of days to apply a rolling mean. Default 1.
roll_align	Character string identifying the direction of the rolling mean from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'.
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.
start_year	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.

end_year	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
exclude_years	Numeric vector of years to exclude from analysis. Leave blank to include all years.
complete_years	Logical values indicating whether to include only years with complete data in analysis. Default FALSE.
months	Numeric vector of months to include in analysis (e.g., 6:8 for Jun-Aug). Leave blank to summarize all months (default 1:12).
transpose	Logical value indicating if the results rows and columns are to be switched. Default FALSE.

Value

A tibble data frame of a long-term percentile of selected years and months.

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Calculate the 20th percentile flow value from a flow record
calc_longterm_percentile(station_number = "08NM116",
                        percentile = 20)

# Calculate the 90th percentile flow value with custom years
calc_longterm_percentile(station_number = "08NM116",
                        start_year = 1980,
                        end_year = 2010,
                        percentile = 90)

}
```

calc_monthly_cumulative_stats

Calculate cumulative monthly flow statistics

Description

Calculate cumulative monthly flow statistics for each month of the year of daily flow values from a streamflow dataset. Calculates the statistics from all daily discharge values for each month from all years, unless specified. Defaults to volumetric cumulative flows, can use use_yield and basin_area to convert to runoff yield.

Usage

```
calc_monthly_cumulative_stats(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  percentiles = c(5, 25, 75, 95),
  use_yield = FALSE,
  basin_area,
  water_year_start = 1,
  start_year,
  end_year,
  exclude_years,
  transpose = FALSE
)
```

Arguments

<code>data</code>	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
<code>dates</code>	Column in <code>data</code> that contains dates formatted YYYY-MM-DD. Only required if <code>dates</code> column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
<code>values</code>	Column in <code>data</code> that contains numeric flow values, in units of cubic metres per second. Only required if <code>values</code> column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.
<code>groups</code>	Column in <code>data</code> that contains unique identifiers for different data sets, if applicable. Only required if <code>groups</code> column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using <code>station_number</code> argument.
<code>station_number</code>	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires <code>tidyhydat</code> package and a HYDAT database. Leave blank if using <code>data</code> argument.
<code>percentiles</code>	Numeric vector of percentiles to calculate. Set to NA if none required. Default <code>c(5, 25, 75, 95)</code> .
<code>use_yield</code>	Logical value indicating whether to use yield runoff, in mm, instead of volumetric. Default FALSE.
<code>basin_area</code>	Upstream drainage basin area, in square kilometres, to apply to observations. Three options: (1) Leave blank if <code>groups</code> is <code>STATION_NUMBER</code> with HYDAT station numbers to extract basin areas from HYDAT. (2) A single numeric value to apply to all observations.

(3) List each basin area for each group/station in groups (can override HYDAT value if listed) as such `c("08NM116" = 795, "08NM242" = 10)`. If group is not listed the HYDAT area will be applied if it exists, otherwise it will be NA.

<code>water_year_start</code>	Numeric value indicating the month of the start of the water year for analysis. Default 1.
<code>start_year</code>	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
<code>end_year</code>	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
<code>exclude_years</code>	Numeric vector of years to exclude from analysis. Leave blank to include all years.
<code>transpose</code>	Logical value indicating if the results rows and columns are to be switched. Default FALSE.

Value

A tibble data frame with the following columns, default units in cubic metres, or millimetres if `use_yield` and `basin_area` provided:

Month	month (MMM-DD) of cumulative statistics
Mean	monthly mean of all cumulative flows for a given month of the year
Median	monthly mean of all cumulative flows for a given month of the year
Maximum	monthly mean of all cumulative flows for a given month of the year
Minimum	monthly mean of all cumulative flows for a given month of the year
P'n'	each monthly n-th percentile selected of all cumulative flows for a given month of the year

Default percentile columns:

P5	monthly 5th percentile of all cumulative flows for a given month of the year
P25	monthly 25th percentile of all cumulative flows for a given month of the year
P75	monthly 75th percentile of all cumulative flows for a given month of the year
P95	monthly 95th percentile of all cumulative flows for a given month of the year

Transposing data creates a column of "Statistics" and subsequent columns for each year selected.

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Calculate annual monthly cumulative volume statistics
calc_monthly_cumulative_stats(station_number = "08NM116")

# Calculate annual monthly cumulative volume statistics with default HYDAT basin area
calc_monthly_cumulative_stats(station_number = "08NM116",
```

```

        use_yield = TRUE)

# Calculate annual monthly cumulative volume statistics with custom basin area
calc_monthly_cumulative_stats(station_number = "08NM116",
                             use_yield = TRUE,
                             basin_area = 800)

}

```

calc_monthly_stats *Calculate monthly summary statistics*

Description

Calculates monthly mean, median, maximum, minimum, and percentiles for each month of all years of daily flow values from a streamflow dataset. Calculates the statistics from all daily discharge values from all years, unless specified.

Usage

```

calc_monthly_stats(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  percentiles = c(10, 90),
  roll_days = 1,
  roll_align = "right",
  water_year_start = 1,
  start_year,
  end_year,
  exclude_years,
  months = 1:12,
  transpose = FALSE,
  spread = FALSE,
  ignore_missing = FALSE
)

```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using station_number argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using station_number argument.

values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using station_number argument.
groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using station_number argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires tidyhydat package and a HYDAT database. Leave blank if using data argument.
percentiles	Numeric vector of percentiles to calculate. Set to NA if none required. Default c(10,90).
roll_days	Numeric value of the number of days to apply a rolling mean. Default 1.
roll_align	Character string identifying the direction of the rolling mean from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'.
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.
start_year	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
end_year	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
exclude_years	Numeric vector of years to exclude from analysis. Leave blank to include all years.
months	Numeric vector of months to include in analysis (e.g.. 6:8 for Jun-Aug). Leave blank to summarize all months (default 1:12).
transpose	Logical value indicating if each month statistic should be individual rows. Default FALSE.
spread	Logical value indicating if each month statistic should be the column name. Default FALSE.
ignore_missing	Logical value indicating whether dates with missing values should be included in the calculation. If TRUE then a statistic will be calculated regardless of missing dates. If FALSE then only statistics from time periods with no missing dates will be returned. Default FALSE.

Value

A tibble data frame with the following columns:

Year	calendar or water year selected
Month	month of the year
Mean	mean of all daily flows for a given month and year

Median	median of all daily flows for a given month and year
Maximum	maximum of all daily flows for a given month and year
Minimum	minimum of all daily flows for a given month and year
P'n'	each n-th percentile selected for a given month and year

Default percentile columns:

P10	10th percentile of all daily flows for a given month and year
P90	90th percentile of all daily flows for a given month and year

Transposing data creates a column of 'Statistics' for each month, labeled as 'Month-Statistic' (ex "Jan-Mean"), and subsequent columns for each year selected. Spreading data creates columns of Year and subsequent columns of Month-Statistics (ex 'Jan-Mean').

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Calculate statistics using a data frame and data argument with defaults
flow_data <- tidyhydat::hy_daily_flows(station_number = "08NM116")
calc_monthly_stats(data = flow_data,
                   start_year = 1980)

# Calculate statistics using station_number argument with defaults
calc_monthly_stats(station_number = "08NM116",
                   start_year = 1980)

# Calculate statistics regardless if there is missing data for a given year
calc_monthly_stats(station_number = "08NM116",
                   ignore_missing = TRUE)

# Calculate statistics for water years starting in October
calc_monthly_stats(station_number = "08NM116",
                   start_year = 1980,
                   water_year_start = 10)

# Calculate statistics with custom years
calc_monthly_stats(station_number = "08NM116",
                   start_year = 1981,
                   end_year = 2010,
                   exclude_years = c(1991,1993:1995))

# Calculate statistics for 7-day flows, with 25 and 75th percentiles
calc_monthly_stats(station_number = "08NM116",
                   roll_days = 7,
                   percentiles = c(25,75),
                   ignore_missing = TRUE)

}
```

 compute_annual_frequencies

Perform an annual low or high-flow frequency analysis

Description

Performs a flow volume frequency analysis on annual statistics from a streamflow dataset. Defaults to low-flow frequency analysis using annual minimums. Use `use_max` for annual high flow frequency analyses. Calculates the statistics from all daily discharge values from all years, unless specified. Function will calculate using all values in the provided data (no grouped analysis). Analysis methodology replicates that from [HEC-SSP](#).

Usage

```
compute_annual_frequencies(
  data,
  dates = Date,
  values = Value,
  station_number,
  roll_days = c(1, 3, 7, 30),
  roll_align = "right",
  use_max = FALSE,
  use_log = FALSE,
  prob_plot_position = c("weibull", "median", "hazen"),
  prob_scale_points = c(0.9999, 0.999, 0.99, 0.9, 0.5, 0.2, 0.1, 0.02, 0.01, 0.001,
    1e-04),
  fit_distr = c("PIII", "weibull"),
  fit_distr_method = ifelse(fit_distr == "PIII", "MOM", "MLE"),
  fit_quantiles = c(0.975, 0.99, 0.98, 0.95, 0.9, 0.8, 0.5, 0.2, 0.1, 0.05, 0.01),
  plot_curve = TRUE,
  water_year_start = 1,
  start_year,
  end_year,
  exclude_years,
  months = 1:12,
  ignore_missing = FALSE
)
```

Arguments

<code>data</code>	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
<code>dates</code>	Column in <code>data</code> that contains dates formatted YYYY-MM-DD. Only required if <code>dates</code> column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.

values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using station_number argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires tidyhydat package and a HYDAT database. Leave blank if using data argument.
roll_days	Numeric value of the number of days to apply a rolling mean. Default 1.
roll_align	Character string identifying the direction of the rolling mean from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'.
use_max	Logical value to indicate using annual maximums rather than the minimums for analysis. Default FALSE.
use_log	Logical value to indicate log-scale transforming of flow data before analysis. Default FALSE.
prob_plot_position	Character string indicating the plotting positions used in the frequency plots, one of "weibull", "median", or "hazen". Points are plotted against $(i-a)/(n+1-a-b)$ where i is the rank of the value; n is the sample size and a and b are defined as: ($a=0$, $b=0$) for Weibull plotting positions; ($a=.2$; $b=.3$) for Median plotting positions; and ($a=.5$; $b=.5$) for Hazen plotting positions. Default "weibull".
prob_scale_points	Numeric vector of probabilities to be plotted along the X axis in the frequency plot. Inverse of return period. Default $c(.9999, .999, .99, .9, .5, .2, .1, .02, .01, .001, .0001)$.
fit_distr	Character string identifying the distribution to fit annual data, one of "PIII" (Pearson Log III distribution) or "weibull" (Weibull distribution). Default "PIII".
fit_distr_method	Character string identifying the method used to fit the distribution, one of "MOM" (method of moments) or "MLE" (maximum likelihood estimation). Selected as "MOM" if <code>fit_distr=="PIII"</code> (default) or "MLE" if <code>fit_distr=="weibull"</code> .
fit_quantiles	Numeric vector of quantiles to be estimated from the fitted distribution. Default $c(.975, .99, .98, .95, .90, .80, .50, .20, .10, .05, .01)$.
plot_curve	Logical value to indicate plotting the computed curve on the probability plot. Default TRUE.
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.
start_year	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
end_year	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
exclude_years	Numeric vector of years to exclude from analysis. Leave blank to include all years.
months	Numeric vector of months to include in analysis (e.g.. 6:8 for Jun-Aug). Leave blank to summarize all months (default 1:12).


```
## End(Not run)
```

```
compute_annual_trends Calculate prewhitened nonlinear annual trends on streamflow data
```

Description

Calculates prewhitened nonlinear trends on annual streamflow data. Uses the 'zyp' package to trend. Review 'zyp' to understand its methodology. All annual statistics calculated using the calc_all_annual_stats() function which uses the following fasstr functions:

- calc_annual_stats()
- calc_annual_lowflows()
- calc_annual_cumulative_stats()
- calc_annual_flow_timing()
- calc_monthly_stats()
- calc_annual_outside_normal()

Usage

```
compute_annual_trends(  
  data,  
  dates = Date,  
  values = Value,  
  groups = STATION_NUMBER,  
  station_number,  
  zyp_method,  
  basin_area,  
  water_year_start = 1,  
  start_year,  
  end_year,  
  exclude_years,  
  annual_percentiles = c(10, 90),  
  monthly_percentiles = c(10, 20),  
  stats_days = 1,  
  stats_align = "right",  
  lowflow_days = c(1, 3, 7, 30),  
  lowflow_align = "right",  
  timing_percent = c(25, 33, 50, 75),  
  normal_percentiles = c(25, 75),  
  ignore_missing = FALSE,  
  include_plots = TRUE,  
  zyp_alpha  
)
```


Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using station_number argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using station_number argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using station_number argument.
groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using station_number argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires tidyhydat package and a HYDAT database. Leave blank if using data argument.
zyp_method	Character string identifying the prewhitened trend method to use from 'zyp', either "zhang" or "yuepilon". Required.
basin_area	Upstream drainage basin area, in square kilometres, to apply to observations. Three options: (1) Leave blank if groups is STATION_NUMBER with HYDAT station numbers to extract basin areas from HYDAT. (2) A single numeric value to apply to all observations. (3) List each basin area for each group/station in groups (can override HYDAT value if listed) as such c("08NM116" = 795, "08NM242" = 10). If group is not listed the HYDAT area will be applied if it exists, otherwise it will be NA.
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.
start_year	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
end_year	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
exclude_years	Numeric vector of years to exclude from analysis. Leave blank to include all years.
annual_percentiles	Numeric vector of percentiles to calculate annually. Set to NA if none required. Used for calc_annual_stats() function. Default c(10, 90).
monthly_percentiles	Numeric vector of percentiles to calculate monthly for each year. Set to NA if none required. Used for calc_monthly_stats() function. Default c(10, 20).

stats_days	Numeric vector of the number of days to apply a rolling mean on basic stats. Default <code>c(1)</code> . Used for <code>calc_annual_stats()</code> and <code>calc_monthly_stats()</code> functions.
stats_align	Character string identifying the direction of the rolling mean on basic stats from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'. Used for <code>calc_annual_stats()</code> , <code>calc_monthly_stats()</code> , and <code>calc_annual_outside_normal()</code> functions.
lowflow_days	Numeric vector of the number of days to apply a rolling mean on lowflow stats. Default <code>c(1, 3, 7, 30)</code> . Used for <code>calc_lowflow_stats()</code> function.
lowflow_align	Character string identifying the direction of the rolling mean on lowflow stats from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'. Used for <code>calc_lowflow_stats()</code> function.
timing_percent	Numeric vector of percents of annual total flows to determine dates. Used for <code>calc_annual_flow_timing()</code> function. Default <code>c(25, 33.3, 50, 75)</code> .
normal_percentiles	Numeric vector of two values, lower and upper percentiles, respectively indicating the limits of the normal range. Default <code>c(25, 75)</code> .
ignore_missing	Logical value indicating whether dates with missing values should be included in the calculation. If TRUE then a statistic will be calculated regardless of missing dates. If FALSE then only statistics from time periods with no missing dates will be returned. Default FALSE.
include_plots	Logical value indicating if annual trending plots should be included. Default TRUE.
zyp_alpha	Numeric value of the significance level (ex. 0.05) of when to plot a trend line. Leave blank for no line.

Value

A list of tibbles and optional plots from the trending analysis including:

Annual_Trends_Data

a tibble of the annual statistics used for trending

Annual_Trends_Results

a tibble of the results of the zyp trending analysis

Annual_*

each ggplot2 object for each annual trended statistic

References

References from zyp package:

- Wang, X.L. and Swail, V.R., 2001. Changes in extreme wave heights in northern hemisphere oceans and related atmospheric circulation regimes. *Journal of Climate*, 14: 2204-2221.
- Yue, S., P. Pilon, B. Phinney and G. Cavadias, 2002. The influence of autocorrelation on the ability to detect trend in hydrological series. *Hydrological Processes*, 16: 1807-1829.
- Zhang, X., Vincent, L.A., Hogg, W.D. and Niitsoo, A., 2000. Temperature and Precipitation Trends in Canada during the 20th Century. *Atmosphere-Ocean* 38(3): 395-429.

- Sen, P.K., 1968. Estimates of the Regression Coefficient Based on Kendall's Tau. Journal of the American Statistical Association Vol. 63, No. 324: 1379-1389.

See Also

[zyp-package](#), [zyp.trend.dataframe](#), [calc_all_annual_stats](#)

Examples

```
## Not run:

# Working examples:

# Compute trends statistics using a data frame and data argument with defaults
flow_data <- tidyhydat::hy_daily_flows(station_number = "08NM116")
trends <- compute_annual_trends(data = flow_data,
                               zyp_method = "yuepilon")

# Compute trends statistics using station_number with defaults
trends <- compute_annual_trends(station_number = "08NM116",
                               zyp_method = "yuepilon")

# Compute trends statistics and plot a trend line if the significance is less than 0.05
trends <- compute_annual_trends(station_number = "08NM116",
                               zyp_method = "yuepilon",
                               zyp_alpha = 0.05)

# Compute trends statistics and do not plot the results
trends <- compute_annual_trends(station_number = "08NM116",
                               zyp_method = "yuepilon",
                               include_plots = FALSE)

## End(Not run)
```

compute_frequency_analysis

Perform a custom annual frequency analysis

Description

Performs a flow volume frequency analysis from a streamflow dataset. Defaults to ranking by minimums; use `use_max` for to rank by maximum flows. Calculates the statistics from events and flow values provided. Columns of events (years), their values (mins or maxs), and identifiers (lowflows, highflows, etc), Function will calculate using all values in the provided data (no grouped analysis). Analysis methodology replicates that from [HEC-SSP](#).

Usage

```
compute_frequency_analysis(
  data,
  events = Year,
  values = Value,
  measures = Measure,
  use_max = FALSE,
  use_log = FALSE,
  prob_plot_position = c("weibull", "median", "hazen"),
  prob_scale_points = c(0.9999, 0.999, 0.99, 0.9, 0.5, 0.2, 0.1, 0.02, 0.01, 0.001,
    1e-04),
  fit_distr = c("PIII", "weibull"),
  fit_distr_method = ifelse(fit_distr == "PIII", "MOM", "MLE"),
  fit_quantiles = c(0.975, 0.99, 0.98, 0.95, 0.9, 0.8, 0.5, 0.2, 0.1, 0.05, 0.01),
  plot_curve = TRUE
)
```

Arguments

<code>data</code>	A data frame of flow data that contains columns of events, flow values, and measures (data type).
<code>events</code>	Column in data that contains event identifiers, typically year values. Default "Year".
<code>values</code>	Column in data that contains numeric flow values, in units of cubic metres per second. Default "Value".
<code>measures</code>	Column in data that contains measure identifiers (example data: '7-day low' or 'Annual Max'). Can have multiple measures (ex. '7-day low' and '30-day low') in column if multiple statistics are desired. Default "Measure".
<code>use_max</code>	Logical value to indicate using annual maximums rather than the minimums for analysis. Default FALSE.
<code>use_log</code>	Logical value to indicate log-scale transforming of flow data before analysis. Default FALSE.
<code>prob_plot_position</code>	Character string indicating the plotting positions used in the frequency plots, one of "weibull", "median", or "hazen". Points are plotted against $(i-a)/(n+1-a-b)$ where i is the rank of the value; n is the sample size and a and b are defined as: ($a=0$, $b=0$) for Weibull plotting positions; ($a=.2$; $b=.3$) for Median plotting positions; and ($a=.5$; $b=.5$) for Hazen plotting positions. Default "weibull".
<code>prob_scale_points</code>	Numeric vector of probabilities to be plotted along the X axis in the frequency plot. Inverse of return period. Default <code>c(.9999, .999, .99, .9, .5, .2, .1, .02, .01, .001, .0001)</code> .
<code>fit_distr</code>	Character string identifying the distribution to fit annual data, one of "PIII" (Pearson Log III distribution) or "weibull" (Weibull distribution). Default "PIII".
<code>fit_distr_method</code>	Character string identifying the method used to fit the distribution, one of "MOM" (method of moments) or "MLE" (maximum likelihood estimation). Selected as "MOM" if <code>fit_distr=="PIII"</code> (default) or "MLE" if <code>fit_distr=="weibull"</code> .

`fit_quantiles` Numeric vector of quantiles to be estimated from the fitted distribution. Default `c(.975, .99, .98, .95, .90, .80, .50, .20, .10, .05, .01)`.

`plot_curve` Logical value to indicate plotting the computed curve on the probability plot. Default `TRUE`.

Value

A list with the following elements:

`Freq_Analysis_Data` Data frame with computed annual summary statistics used in analysis.

`Freq_Plot_Data` Data frame with co-ordinates used in frequency plot.

`Freq_Plot` `ggplot2` object with frequency plot

`Freq_Fitting` List of fitted objects from `fitdistrplus`.

`Freq_Fitted_Quantiles` Data frame with fitted quantiles.

Examples

```
## Not run:

# Working example:

# Calculate some values to use for a frequency analysis
# (requires years, values for those years, and the name of the measure/metric)
low_flows <- calc_annual_lowflows(station_number = "08NM116",
                                start_year = 1980,
                                end_year = 2000,
                                roll_days = 7)

low_flows <- dplyr::select(low_flows, Year, Value = Min_7_Day)
low_flows <- dplyr::mutate(low_flows, Measure = "7-Day")

# Compute the frequency analysis using the default parameters
results <- compute_frequency_analysis(data = low_flows,
                                     events = Year,
                                     values = Value,
                                     measure = Measure)

## End(Not run)
```

compute_frequency_quantile

Calculate an annual frequency analysis quantile

Description

Performs a volume frequency analysis on annual statistics from a streamflow dataset and calculates a statistic based on the provided mean n-days and return period of the statistic, defaults to minimum flows. For example, to determine the 7Q10 of a dataset, set the `roll_days` to 7 and the `return_period` to 10. Calculates the statistic from all daily discharge values from all years and months, unless specified. Function will calculate using all values in the provided data (no grouped analysis). Analysis methodology replicates that from [HEC-SSP](#).

Usage

```
compute_frequency_quantile(
  data,
  dates = Date,
  values = Value,
  station_number,
  roll_days = NA,
  roll_align = "right",
  return_period = NA,
  use_max = FALSE,
  use_log = FALSE,
  fit_distr = c("PIII", "weibull"),
  fit_distr_method = ifelse(fit_distr == "PIII", "MOM", "MLE"),
  water_year_start = 1,
  start_year,
  end_year,
  exclude_years,
  months = 1:12,
  ignore_missing = FALSE
)
```

Arguments

<code>data</code>	A data frame of flow data that contains columns of events, flow values, and measures (data type).
<code>dates</code>	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
<code>values</code>	Column in data that contains numeric flow values, in units of cubic metres per second. Default "Value".
<code>station_number</code>	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires tidyhydat package and a HYDAT database. Leave blank if using <code>data</code> argument.
<code>roll_days</code>	Numeric value of the number of days to apply a rolling mean. Required.
<code>roll_align</code>	Character string identifying the direction of the rolling mean from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'.

return_period	Numeric vector of the estimated time interval, in years, between flow events of a similar size, inverse of probability, used to estimate the frequency statistic. Required.
use_max	Logical value to indicate using annual maximums rather than the minimums for analysis. Default FALSE.
use_log	Logical value to indicate log-scale transforming of flow data before analysis. Default FALSE.
fit_distr	Character string identifying the distribution to fit annual data, one of "PIII" (Pearson Log III distribution) or "weibull" (Weibull distribution). Default "PIII".
fit_distr_method	Character string identifying the method used to fit the distribution, one of "MOM" (method of moments) or "MLE" (maximum likelihood estimation). Selected as "MOM" if fit_distr=="PIII" (default) or "MLE" if fit_distr=="weibull".
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.
start_year	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
end_year	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
exclude_years	Numeric vector of years to exclude from analysis. Leave blank to include all years.
months	Numeric vector of months to include in analysis (e.g., 6:8 for Jun-Aug). Leave blank to summarize all months (default 1:12).
ignore_missing	Logical value indicating whether dates with missing values should be included in the calculation. If TRUE then a statistic will be calculated regardless of missing dates. If FALSE then only statistics from time periods with no missing dates will be returned. Default FALSE.

Value

A numeric value of the frequency analysis quantile, given the roll_days and return_period

See Also

[compute_frequency_analysis](#)

Examples

```
## Not run:

# Working example:

# Compute the annual 7-day flow value with a 1 in 10 year return interval
compute_frequency_quantile(station_number = "08NM116",
                           roll_days = 7,
                           return_period = 10)
```

```
## End(Not run)
```

```
compute_full_analysis Compute a suite of tables and plots from various fasstr functions
```

Description

Calculates tables and plots from a suite of statistics from fasstr functions. The statistics are grouped into 7 analysis groups (see 'analyses' argument) which are stored in lists in the object. Due to the number of tables and plots to be made, this function may take several minutes to complete. If using `ignore_missing = FALSE` (default) and there is missing data, some tables and plots may be empty and produce warnings. Use `ignore_missing = TRUE` to ignore the missing values or filter your data to complete years.

Usage

```
compute_full_analysis(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  analyses = 1:7,
  basin_area,
  water_year_start = 1,
  start_year,
  end_year,
  exclude_years,
  ignore_missing = FALSE,
  zyp_method = "yuepilon",
  zyp_alpha
)
```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.

groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using station_number argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires tidyhydat package and a HYDAT database. Leave blank if using data argument.
analyses	Numeric vector of analyses to run (default is all (1:7)): <ul style="list-style-type: none"> • 1: Screening • 2: Long-term • 3: Annual • 4: Monthly • 5: Daily • 6: Annual Trends • 7: Low-flow Frequencies
basin_area	Upstream drainage basin area, in square kilometres, to apply to observations. Three options: <ol style="list-style-type: none"> (1) Leave blank if groups is STATION_NUMBER with HYDAT station numbers to extract basin areas from HYDAT. (2) A single numeric value to apply to all observations. (3) List each basin area for each group/station in groups (can override HYDAT value if listed) as such c("08NM116" = 795, "08NM242" = 10). If group is not listed the HYDAT area will be applied if it exists, otherwise it will be NA.
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.
start_year	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
end_year	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
exclude_years	Numeric vector of years to exclude from analysis. Leave blank to include all years.
ignore_missing	Logical value indicating whether dates with missing values should be included in the calculation. If TRUE then a statistic will be calculated regardless of missing dates. If FALSE then only statistics from time periods with no missing dates will be returned. Default FALSE.
zyp_method	Character string identifying the prewhitened trend method to use from 'zyp', either "zhang" or "yuepilon". Only required if analysis group 6 is included. Default "yuepilon".
zyp_alpha	Numeric value of the significance level (ex. 0.05) of when to plot a trend line. Leave blank for no line.

Value

A list of lists of tibble data frames and ggplot2 objects from various fasstr functions organized by the analysis groups as listed above.

See Also

plot_flow_data, screen_flow_data, plot_data_screening, plot_missing_dates, calc_longterm_monthly_stats, plot_longterm_monthly_stats, calc_longterm_daily_stats, plot_longterm_daily_stats, plot_flow_duration, calc_annual_stats, plot_annual_stats, calc_annual_cumulative_stats, plot_annual_cumulative_stats, calc_annual_flow_timing, plot_annual_flow_timing, calc_annual_outside_normal, plot_annual_outside_normal, calc_annual_lowflows, plot_annual_lowflows, plot_annual_means, calc_monthly_stats, plot_monthly_stats, calc_monthly_cumulative_stats, plot_monthly_cumulative_stats, calc_daily_stats, plot_daily_stats, calc_daily_cumulative_stats, plot_daily_cumulative_stats, compute_annual_trends, compute_annual_frequencies, write_flow_data, write_plots

Examples

```
## Not run:
```

```
# Working examples:
```

```
# Compute a full analysis will all the analyses
results <- compute_full_analysis(station_number = "08NM116",
                                start_year = 1980,
                                end_year = 2010)
```

```
# Compute a full analysis with only Annual (3) and Daily (5) analyses
results <- compute_full_analysis(station_number = "08NM116",
                                start_year = 1980,
                                end_year = 2010,
                                analyses = c(3,5))
```

```
## End(Not run)
```

```
compute_hydat_peak_frequencies
```

Perform a flow frequency analysis on annual statistics

Description

Performs a volume frequency analysis on annual statistics from a streamflow dataset. Calculates the statistics from all daily discharge values from all years, unless specified. Function will calculate using all values in the provided data (no grouped analysis). Analysis methodology replicates that from **HEC-SSP**.

Usage

```

compute_hydat_peak_frequencies(
  station_number,
  use_max = FALSE,
  use_log = FALSE,
  prob_plot_position = c("weibull", "median", "hazen"),
  prob_scale_points = c(0.9999, 0.999, 0.99, 0.9, 0.5, 0.2, 0.1, 0.02, 0.01, 0.001,
    1e-04),
  fit_distr = c("PIII", "weibull"),
  fit_distr_method = ifelse(fit_distr == "PIII", "MOM", "MLE"),
  fit_quantiles = c(0.975, 0.99, 0.98, 0.95, 0.9, 0.8, 0.5, 0.2, 0.1, 0.05, 0.01),
  start_year,
  end_year,
  exclude_years,
  plot_curve = TRUE
)

```

Arguments

station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract annual peak minimum or maximum instantaneous streamflow data from a HYDAT database. Requires tidyhydat package and a HYDAT database.
use_max	Logical value to indicate using annual maximums rather than the minimums for analysis. Default FALSE.
use_log	Logical value to indicate log-scale transforming of flow data before analysis. Default FALSE.
prob_plot_position	Character string indicating the plotting positions used in the frequency plots, one of "weibull", "median", or "hazen". Points are plotted against $(i-a)/(n+1-a-b)$ where i is the rank of the value; n is the sample size and a and b are defined as: ($a=0$, $b=0$) for Weibull plotting positions; ($a=.2$; $b=.3$) for Median plotting positions; and ($a=.5$; $b=.5$) for Hazen plotting positions. Default "weibull".
prob_scale_points	Numeric vector of probabilities to be plotted along the X axis in the frequency plot. Inverse of return period. Default $c(.9999, .999, .99, .9, .5, .2, .1, .02, .01, .001, .0001)$.
fit_distr	Character string identifying the distribution to fit annual data, one of "PIII" (Pearson Log III distribution) or "weibull" (Weibull distribution). Default "PIII".
fit_distr_method	Character string identifying the method used to fit the distribution, one of "MOM" (method of moments) or "MLE" (maximum likelihood estimation). Selected as "MOM" if $fit_distr=="PIII"$ (default) or "MLE" if $fit_distr=="weibull"$.
fit_quantiles	Numeric vector of quantiles to be estimated from the fitted distribution. Default $c(.975, .99, .98, .95, .90, .80, .50, .20, .10, .05, .01)$.
start_year	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.

fill_missing_dates	<i>Fill dates of missing flow values with NA</i>
--------------------	--

Description

Adds rows of dates with missing flow values to a streamflow dataset with daily flow values of NA. Missing dates will be filled in gaps between data and completely fill the first and last years (calendar or water year if selected).

Usage

```
fill_missing_dates(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  water_year_start = 1
)
```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.
groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using <code>station_number</code> argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires <code>tidyhydat</code> package and a HYDAT database. Leave blank if using <code>data</code> argument.
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.

Value

A tibble data frame of the source data with additional rows of filled values of missing dates.

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Fill missing dates with NA using calendar years
fill_missing_dates(station_number = "08NM116")

# Fill missing dates with NA using water years starting in August
fill_missing_dates(station_number = "08NM116",
                   water_year_start = 8)

}
```

plot_annual_cumulative_stats

Plot annual and seasonal total flows

Description

Plots annual and seasonal total flows, volumetric or runoff yield flows, from a streamflow dataset. Calculates the statistics from all daily discharge values from all years, unless specified. Data calculated from `plot_annual_cumulative_stats()` function. For water year and seasonal data, the designated year is the year in which the year or season ends. For example, if using water years with a start month of 11, the OND season is designated by the water year which starts in November (designated by the calendar year in which it ends).

Usage

```
plot_annual_cumulative_stats(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  use_yield = FALSE,
  basin_area,
  water_year_start = 1,
  start_year,
  end_year,
  exclude_years,
  include_seasons = FALSE,
  log_discharge = FALSE,
  include_title = FALSE
)
```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.
groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using <code>station_number</code> argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires tidyhydat package and a HYDAT database. Leave blank if using data argument.
use_yield	Logical value indicating whether to use yield runoff, in mm, instead of volumetric. Default FALSE.
basin_area	Upstream drainage basin area, in square kilometres, to apply to observations. Three options: (1) Leave blank if groups is STATION_NUMBER with HYDAT station numbers to extract basin areas from HYDAT. (2) A single numeric value to apply to all observations. (3) List each basin area for each group/station in groups (can override HYDAT value if listed) as such <code>c("08NM116" = 795, "08NM242" = 10)</code> . If group is not listed the HYDAT area will be applied if it exists, otherwise it will be NA.
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.
start_year	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
end_year	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
exclude_years	Numeric vector of years to exclude from analysis. Leave blank to include all years.
include_seasons	Logical value indication whether to include seasonal yields and total discharges. Default TRUE.
log_discharge	Logical value to indicate plotting the discharge axis (Y-axis) on a logarithmic scale. Default FALSE.
include_title	Logical value to indicate adding the group/station number to the plot, if provided. Default FALSE.

```
plot_annual_flow_timing
      Plot annual flow timing
```

Description

Plots annual the timing (day of year) of occurrence of portions of total annual flow of daily flow values from a streamflow dataset. Calculates the statistics from all daily discharge values from all years, unless specified. Data calculated using `calc_annual_flow_timing()` function.

Usage

```
plot_annual_flow_timing(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  percent_total = c(25, 33.3, 50, 75),
  water_year_start = 1,
  start_year,
  end_year,
  exclude_years,
  include_title = FALSE
)
```

Arguments

<code>data</code>	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
<code>dates</code>	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
<code>values</code>	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.
<code>groups</code>	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using <code>station_number</code> argument.
<code>station_number</code>	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires <code>tidyhydat</code> package and a HYDAT database. Leave blank if using <code>data</code> argument.

<code>percent_total</code>	Numeric vector of percents of total annual flows to determine dates. Default <code>c(25, 33.3, 50, 75)</code> .
<code>water_year_start</code>	Numeric value indicating the month of the start of the water year for analysis. Default 1.
<code>start_year</code>	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
<code>end_year</code>	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
<code>exclude_years</code>	Numeric vector of years to exclude from analysis. Leave blank to include all years.
<code>include_title</code>	Logical value to indicate adding the group/station number to the plot, if provided. Default FALSE.

Value

A list of `ggplot2` objects with the following for each station provided:

`Annual_Flow_Timing`

a plot that contains each n-percent of total volumetric discharge

Default plots on each object:

`DoY_25pct_TotalQ`

day of year of 25-percent of total volumetric discharge

`DoY_33.3pct_TotalQ`

day of year of 33.3-percent of total volumetric discharge

`DoY_50pct_TotalQ`

day of year of 50-percent of total volumetric discharge

`DoY_75pct_TotalQ`

day of year of 75-percent of total volumetric discharge

References

- Barnett, T.P., Pierce, D.W., Hidalgo, H.G., Bonfils, C., Santer, B.D., Das, T., Bala, G., Wood, A.W., Nozawa, T., Mirin, A.A., Cayan, D.R., Dettinger, M.D., 2008. Human-Induced Changes in the Hydrology of the Western United States. *Science* 319, 1080-1083.

See Also

[calc_annual_flow_timing](#)

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Plot annual flow timing statistics with default percent totals
plot_annual_flow_timing(station_number = "08NM116")
```

```
# Plot annual flow timing with custom percent totals
plot_annual_flow_timing(station_number = "08NM116",
                        percent_total = 50)

}
```

plot_annual_lowflows *Plot annual lowflows*

Description

Plot annual n-day minimum values, and the day of year of occurrence of daily flow values from a streamflow dataset. Plots statistics from all daily discharge values from all years, unless specified. Data calculated from `calc_annual_lowflows()` function.

Usage

```
plot_annual_lowflows(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  roll_days = c(1, 3, 7, 30),
  roll_align = "right",
  water_year_start = 1,
  start_year,
  end_year,
  exclude_years,
  months = 1:12,
  ignore_missing = FALSE,
  log_discharge = FALSE,
  include_title = FALSE
)
```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.

groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using station_number argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires tidyhydat package and a HYDAT database. Leave blank if using data argument.
roll_days	Numeric value of the number of days to apply a rolling mean. Default 1.
roll_align	Character string identifying the direction of the rolling mean from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'.
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.
start_year	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
end_year	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
exclude_years	Numeric vector of years to exclude from analysis. Leave blank to include all years.
months	Numeric vector of months to include in analysis (e.g.. 6:8 for Jun-Aug). Leave blank to summarize all months (default 1:12).
ignore_missing	Logical value indicating whether dates with missing values should be included in the calculation. If TRUE then a statistic will be calculated regardless of missing dates. If FALSE then only statistics from time periods with no missing dates will be returned. Default FALSE.
log_discharge	Logical value to indicate plotting the discharge axis (Y-axis) on a logarithmic scale. Default FALSE.
include_title	Logical value to indicate adding the group/station number to the plot, if provided. Default FALSE.

Value

A list of ggplot2 objects with the following for each station provided:

Annual_Minimums

ggplot2 object of annual minimums of selected n-day rolling means

Annual_Minimums_Days

ggplot2 object of the day of years of annual minimums of selected n-day rolling means

See Also

[calc_annual_lowflows](#)

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Plot annual 1, 3, 7, and 30-day (default) low flow statistics with default alignment
plot_annual_lowflows(station_number = "08NM116")

# Plot annual custom 3 and 7-day low flow statistics with "center" alignment
plot_annual_lowflows(station_number = "08NM116",
                      roll_days = c(3,7),
                      roll_align = "center")

}
```

plot_annual_means	<i>Plot annual means compared to the long-term mean</i>
-------------------	---

Description

Plot annual means using the long-term annual mean as the point of reference to compare wet vs. dry years. Plots the statistics from all daily discharge values from all years, unless specified. Data calculated using `calc_annual_stats()` function.

Usage

```
plot_annual_means(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  roll_days = 1,
  roll_align = "right",
  water_year_start = 1,
  start_year,
  end_year,
  exclude_years,
  months = 1:12,
  ignore_missing = FALSE,
  include_title = FALSE
)
```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
------	--

<code>dates</code>	Column in data that contains dates formatted YYYY-MM-DD. Only required if <code>dates</code> column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
<code>values</code>	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if <code>values</code> column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.
<code>groups</code>	Column in data that contains unique identifiers for different data sets, if applicable. Only required if <code>groups</code> column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using <code>station_number</code> argument.
<code>station_number</code>	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires <code>tidyhydat</code> package and a HYDAT database. Leave blank if using <code>data</code> argument.
<code>roll_days</code>	Numeric value of the number of days to apply a rolling mean. Default 1.
<code>roll_align</code>	Character string identifying the direction of the rolling mean from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'.
<code>water_year_start</code>	Numeric value indicating the month of the start of the water year for analysis. Default 1.
<code>start_year</code>	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
<code>end_year</code>	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
<code>exclude_years</code>	Numeric vector of years to exclude from analysis. Leave blank to include all years.
<code>months</code>	Numeric vector of months to include in analysis (e.g.. 6:8 for Jun-Aug). Leave blank to summarize all months (default 1:12).
<code>ignore_missing</code>	Logical value indicating whether dates with missing values should be included in the calculation. If TRUE then a statistic will be calculated regardless of missing dates. If FALSE then only statistics from time periods with no missing dates will be returned. Default FALSE.
<code>include_title</code>	Logical value to indicate adding the group/station number to the plot, if provided. Default FALSE.

Value

A list of `ggplot2` objects for with the following plots for each station provided:

`Annual_Means` a plot that contains annual means with the long-term mean as the x-axis intercept

See Also

[calc_annual_stats](#)

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Plot annual means
plot_annual_means(station_number = "08NM116")

# Plot mean flows from July-September
plot_annual_means(station_number = "08NM116",
                  months = 7:9)

}
```

```
plot_annual_outside_normal
```

Plot annual days above and below normal

Description

Plots the number of days per year outside of the 'normal' range (typically between 25 and 75th percentiles) for each day of the year. Upper and lower-range percentiles are calculated for each day of the year of from all years, and then each daily flow value for each year is compared. All days above or below the normal range are included. Calculates the statistics from all daily discharge values from all years, unless specified. Data calculated using `calc_annual_outside_normal()` function.

Usage

```
plot_annual_outside_normal(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  normal_percentiles = c(25, 75),
  roll_days = 1,
  roll_align = "right",
  water_year_start = 1,
  start_year,
  end_year,
  exclude_years,
  months = 1:12,
  include_title = FALSE
)
```

Arguments

<code>data</code>	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
<code>dates</code>	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
<code>values</code>	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.
<code>groups</code>	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using <code>station_number</code> argument.
<code>station_number</code>	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires <code>tidyhydat</code> package and a HYDAT database. Leave blank if using <code>data</code> argument.
<code>normal_percentiles</code>	Numeric vector of two values, lower and upper percentiles, respectively indicating the limits of the normal range. Default <code>c(25, 75)</code> .
<code>roll_days</code>	Numeric value of the number of days to apply a rolling mean. Default 1.
<code>roll_align</code>	Character string identifying the direction of the rolling mean from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'.
<code>water_year_start</code>	Numeric value indicating the month of the start of the water year for analysis. Default 1.
<code>start_year</code>	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
<code>end_year</code>	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
<code>exclude_years</code>	Numeric vector of years to exclude from analysis. Leave blank to include all years.
<code>months</code>	Numeric vector of months to include in analysis (e.g.. 6:8 for Jun-Aug). Leave blank to summarize all months (default 1:12).
<code>include_title</code>	Logical value to indicate adding the group/station number to the plot, if provided. Default FALSE.

Value

A list of `ggplot2` objects with the following for each station provided:

`Annual_Days_Outside_Normal`

a plot that contains the number of days outside normal

Default plots on each object:

Days_Below_Normal
 number of days per year below the daily normal (default 25th percentile)

Days_Above_Normal
 number of days per year above the daily normal (default 75th percentile)

Days_Outside_Normal
 number of days per year below and above the daily normal (default 25/75th percentile)

See Also

[calc_annual_outside_normal](#)

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Plot annual statistics with default limits of normal (25 and 75th percentiles)
plot_annual_outside_normal(station_number = "08NM116")

# Plot annual statistics with custom limits of normal
plot_annual_outside_normal(station_number = "08NM116",
                           normal_percentiles = c(10,90))

}
```

plot_annual_stats *Plot annual summary statistics*

Description

Plot annual mean, median, maximum, minimum, and percentiles of daily flow values from a stream-flow dataset. Plots the statistics from all daily discharge values from all years, unless specified. Data calculated using `calc_annual_stats()` function.

Usage

```
plot_annual_stats(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  percentiles,
  roll_days = 1,
  roll_align = "right",
```

```

water_year_start = 1,
start_year,
end_year,
exclude_years,
months = 1:12,
ignore_missing = FALSE,
log_discharge = FALSE,
include_title = FALSE
)

```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.
groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using <code>station_number</code> argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires <code>tidyhydat</code> package and a HYDAT database. Leave blank if using <code>data</code> argument.
percentiles	Numeric vector of percentiles to calculate. Set to NA if none required. Default NA.
roll_days	Numeric value of the number of days to apply a rolling mean. Default 1.
roll_align	Character string identifying the direction of the rolling mean from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'.
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.
start_year	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
end_year	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
exclude_years	Numeric vector of years to exclude from analysis. Leave blank to include all years.

months	Numeric vector of months to include in analysis (e.g. 6:8 for Jun-Aug). Leave blank to summarize all months (default 1:12).
ignore_missing	Logical value indicating whether dates with missing values should be included in the calculation. If TRUE then a statistic will be calculated regardless of missing dates. If FALSE then only statistics from time periods with no missing dates will be returned. Default FALSE.
log_discharge	Logical value to indicate plotting the discharge axis (Y-axis) on a logarithmic scale. Default FALSE.
include_title	Logical value to indicate adding the group/station number to the plot, if provided. Default FALSE.

Value

A list of ggplot2 objects for with the following plots (percentile plots optional) for each station provided:

Annual_Stats a plot that contains annual statistics

Default plots on each object:

Mean	annual mean of all daily flows
Median	annual median of all daily flows
Maximum	annual maximum of all daily flows
Minimum	annual minimum of all daily flows

See Also

[calc_annual_stats](#)

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Plot annual statistics using a data frame and data argument with defaults
flow_data <- tidyhydat::hy_daily_flows(station_number = "08NM116")
plot_annual_stats(data = flow_data)

# Plot annual statistics using station_number argument with defaults
plot_annual_stats(station_number = "08NM116")

# Plot annual statistics regardless if there is missing data for a given year
plot_annual_stats(station_number = "08NM116",
                  ignore_missing = TRUE)

# Plot annual statistics for water years starting in October
plot_annual_stats(station_number = "08NM116",
                  water_year_start = 10)

# Plot annual statistics with custom years
```

```

plot_annual_stats(station_number = "08NM116",
                  start_year = 1981,
                  end_year = 2010,
                  exclude_years = c(1991,1993:1995))

# Plot annual statistics for 7-day flows for July-September months only,
# with 25 and 75th percentiles
plot_annual_stats(station_number = "08NM116",
                  roll_days = 7,
                  months = 7:9,
                  percentiles = c(25,75))

# Plot annual statistics with the a log-scale Discharge axis
plot_annual_stats(station_number = "08NM116",
                  start_year = 1981,
                  end_year = 2010,
                  log_discharge = TRUE)

# Plot annual statistics and include a title with the grouping (default by STATION_NUMBER)
plot_annual_stats(station_number = "08NM116",
                  start_year = 1981,
                  end_year = 2010,
                  include_title = TRUE)

}

```

```
plot_daily_cumulative_stats
```

Plot cumulative daily flow statistics

Description

Plot the daily cumulative mean, median, maximum, minimum, and 5, 25, 75, 95th percentiles for each day of the year from a streamflow dataset. Plots the statistics from all daily cumulative values from all years, unless specified. Data calculated using `calc_daily_cumulative_stats()` function. Can plot individual years for comparison using the `add_year` argument. Defaults to volumetric cumulative flows, can use `use_yield` and `basin_area` to convert to runoff yield.

Usage

```

plot_daily_cumulative_stats(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  use_yield = FALSE,
  basin_area,
  water_year_start = 1,

```

```

    start_year,
    end_year,
    exclude_years,
    log_discharge = FALSE,
    include_title = FALSE,
    add_year
  )

```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.
groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using <code>station_number</code> argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires <code>tidyhydat</code> package and a HYDAT database. Leave blank if using data argument.
use_yield	Logical value indicating whether to use yield runoff, in mm, instead of volumetric. Default FALSE.
basin_area	Upstream drainage basin area, in square kilometres, to apply to observations. Three options: (1) Leave blank if groups is STATION_NUMBER with HYDAT station numbers to extract basin areas from HYDAT. (2) A single numeric value to apply to all observations. (3) List each basin area for each group/station in groups (can override HYDAT value if listed) as such <code>c("08NM116" = 795, "08NM242" = 10)</code> . If group is not listed the HYDAT area will be applied if it exists, otherwise it will be NA.
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.
start_year	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
end_year	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.

<code>exclude_years</code>	Numeric vector of years to exclude from analysis. Leave blank to include all years.
<code>log_discharge</code>	Logical value to indicate plotting the discharge axis (Y-axis) on a logarithmic scale. Default FALSE.
<code>include_title</code>	Logical value to indicate adding the group/station number to the plot, if provided. Default FALSE.
<code>add_year</code>	Numeric value indicating a year of daily flows to add to the daily statistics plot. Leave blank for no years.

Value

A list of ggplot2 objects with the following for each station provided:

`Daily_Cumulative_Stats`

a plot that contains daily cumulative flow statistics

Default plots on each object:

Mean daily cumulative mean

Median daily cumulative median

Min-5 Percentile Range

a ribbon showing the range of data between the daily cumulative minimum and 5th percentile

5-25 Percentiles Range

a ribbon showing the range of data between the daily cumulative 5th and 25th percentiles

25-75 Percentiles Range

a ribbon showing the range of data between the daily cumulative 25th and 75th percentiles

75-95 Percentiles Range

a ribbon showing the range of data between the daily cumulative 75th and 95th percentiles

95 Percentile-Max Range

a ribbon showing the range of data between the daily cumulative 95th percentile and the maximum

'Year' Flows (optional) the daily cumulative flows for the designated year

See Also

[calc_daily_cumulative_stats](#)

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Plot annual daily volume statistics
plot_daily_cumulative_stats(station_number = "08NM116")
```

```

# Plot annual daily yield statistics with default HYDAT basin area
plot_daily_cumulative_stats(station_number = "08NM116",
                             use_yield = TRUE)

# Plot annual daily yield statistics with custom basin area
plot_daily_cumulative_stats(station_number = "08NM116",
                             use_yield = TRUE,
                             basin_area = 800)

}

```

plot_daily_stats *Plot daily summary statistics*

Description

Plot the daily mean, median, maximum, minimum, and percentiles for each day of the year of daily flow values from a streamflow dataset. Plots the statistics from all daily discharge values from all years, unless specified. Can determine statistics of rolling mean days (e.g. 7-day flows) using the `roll_days` argument. The Maximum-Minimum band can be removed using the `'include_extremes'` argument and the percentile bands can be customized using the `'inner_percentiles'` and `'outer_percentiles'` arguments. Data calculated using `calc_daily_stats()` function.

Usage

```

plot_daily_stats(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  roll_days = 1,
  roll_align = "right",
  water_year_start = 1,
  start_year,
  end_year,
  exclude_years,
  complete_years = FALSE,
  months = 1:12,
  ignore_missing = FALSE,
  include_extremes = TRUE,
  inner_percentiles = c(25, 75),
  outer_percentiles = c(5, 95),
  add_year,
  log_discharge = TRUE,
  include_title = FALSE
)

```

Arguments

<code>data</code>	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
<code>dates</code>	Column in <code>data</code> that contains dates formatted YYYY-MM-DD. Only required if <code>dates</code> column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
<code>values</code>	Column in <code>data</code> that contains numeric flow values, in units of cubic metres per second. Only required if <code>values</code> column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.
<code>groups</code>	Column in <code>data</code> that contains unique identifiers for different data sets, if applicable. Only required if <code>groups</code> column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using <code>station_number</code> argument.
<code>station_number</code>	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires <code>tidyhydat</code> package and a HYDAT database. Leave blank if using <code>data</code> argument.
<code>roll_days</code>	Numeric value of the number of days to apply a rolling mean. Default 1.
<code>roll_align</code>	Character string identifying the direction of the rolling mean from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'.
<code>water_year_start</code>	Numeric value indicating the month of the start of the water year for analysis. Default 1.
<code>start_year</code>	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
<code>end_year</code>	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
<code>exclude_years</code>	Numeric vector of years to exclude from analysis. Leave blank to include all years.
<code>complete_years</code>	Logical values indicating whether to include only years with complete data in analysis. Default FALSE.
<code>months</code>	Numeric vector of months to include in analysis (e.g.. 6:8 for Jun-Aug). Leave blank to summarize all months (default 1:12).
<code>ignore_missing</code>	Logical value indicating whether dates with missing values should be included in the calculation. If TRUE then a statistic will be calculated regardless of missing dates. If FALSE then only statistics from time periods with no missing dates will be returned. Default FALSE.
<code>include_extremes</code>	Logical value to indicate plotting a ribbon with the range of daily minimum and maximum flows. Default TRUE.

inner_percentiles	Numeric vector of two values of two percentiles indicating the lower and upper limits of the inner percentiles ribbon for plotting. Default c(25, 75), set to NULL for no inner ribbon.
outer_percentiles	Numeric vector of two values of two percentiles indicating the lower and upper limits of the outer percentiles ribbon for plotting. Default c(5, 95), set to NULL for no outer ribbon.
add_year	Numeric value indicating a year of daily flows to add to the daily statistics plot. Leave blank for no years.
log_discharge	Logical value to indicate plotting the discharge axis (Y-axis) on a logarithmic scale. Default FALSE.
include_title	Logical value to indicate adding the group/station number to the plot, if provided. Default FALSE.

Value

A list of ggplot2 objects with the following for each station provided:

Daily_Stats a plot that contains daily flow statistics

Default plots on each object:

Mean daily mean

Median daily median

25-75 Percentiles

a ribbon showing the range of data between the daily 25th and 75th percentiles

5-95 Percentiles

a ribbon showing the range of data between the daily 5th and 95th percentiles

Minimum-Maximum

a ribbon showing the range of data between the daily minimum and maximums

'Year'

(on annual plots) the daily flows for the designated year

See Also

[calc_daily_stats](#)

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Plot daily statistics using a data frame and data argument with defaults
flow_data <- tidyhydat::hy_daily_flows(station_number = "08NM116")
plot_daily_stats(data = flow_data,
                 start_year = 1980)

# Plot daily statistics using only years with no missing data
plot_daily_stats(station_number = "08NM116",
```

```

        complete_years = TRUE)

# Plot daily statistics and add a specific year's daily flows
plot_daily_stats(station_number = "08NM116",
                 start_year = 1980,
                 add_year = 1985)

# Plot daily statistics for 7-day flows for July-September months only
plot_daily_stats(station_number = "08NM116",
                 start_year = 1980,
                 roll_days = 7,
                 months = 7:9)

# Plot daily statistics without a log-scale Discharge axis
plot_daily_stats(station_number = "08NM116",
                 start_year = 1981,
                 end_year = 2010,
                 log_discharge = FALSE)

}

```

plot_data_screening *Plot annual summary statistics for data screening*

Description

Plots the mean, median, maximum, minimum, standard deviation of annual flows. Plots the statistics from all daily discharge values from all years, unless specified. Data calculated using `screen_flow_data()` function.

Usage

```

plot_data_screening(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  roll_days = 1,
  roll_align = "right",
  water_year_start = 1,
  months = 1:12,
  start_year,
  end_year,
  include_title = FALSE
)

```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.
groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using <code>station_number</code> argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires tidyhydat package and a HYDAT database. Leave blank if using <code>data</code> argument.
roll_days	Numeric value of the number of days to apply a rolling mean. Default 1.
roll_align	Character string identifying the direction of the rolling mean from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'.
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.
months	Numeric vector of months to include in analysis (e.g.. 6:8 for Jun-Aug). Leave blank to summarize all months (default 1:12).
start_year	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
end_year	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
include_title	Logical value to indicate adding the group/station number to the plot, if provided. Default FALSE.

Value

A list of ggplot2 objects with the following for each station provided:

`Data_Screening` a plot that contains annual summary statistics for screening

Default plots on each object:

Minimum	annual minimum of all daily flows for a given year
Maximum	annual maximum of all daily flows for a given year
Mean	annual mean of all daily flows for a given year
StandardDeviation	annual 1 standard deviation of all daily flows for a given year

See Also[screen_flow_data](#)**Examples**

```

# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Plot screening statistics using a data frame and data argument with defaults
flow_data <- tidyhydat::hy_daily_flows(station_number = "08NM116")
plot_data_screening(data = flow_data)

# Plot screening statistics using station_number argument with defaults
plot_data_screening(station_number = "08NM116")

# Plot screening statistics for water years starting in October
plot_data_screening(station_number = "08NM116",
                    water_year_start = 10)

# Plot screening statistics for 7-day flows for July-September months only
plot_data_screening(station_number = "08NM116",
                    roll_days = 7,
                    months = 7:9)

}

```

`plot_flow_data`*Plot daily mean streamflow*

Description

Plot the daily mean flow values from a streamflow dataset. Plots the statistics from all daily discharge values from all years, unless specified. Can choose specific dates to start and end plotting. Can choose to plot out each year separately. Data calculated using `calc_daily_stats()` function. Multiple groups/stations can be plotted if provided with the `groups` argument.

Usage

```

plot_flow_data(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  roll_days = 1,
  roll_align = "right",
  water_year_start = 1,
  start_year,

```

```

    end_year,
    exclude_years,
    start_date,
    end_date,
    log_discharge = FALSE,
    plot_by_year = FALSE,
    one_plot = FALSE,
    include_title = FALSE
  )

```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.
groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using <code>station_number</code> argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires <code>tidyhydat</code> package and a HYDAT database. Leave blank if using <code>data</code> argument.
roll_days	Numeric value of the number of days to apply a rolling mean. Default 1.
roll_align	Character string identifying the direction of the rolling mean from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'.
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.
start_year	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
end_year	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
exclude_years	Numeric vector of years to exclude from analysis. Leave blank to include all years.
start_date	Date (YYYY-MM-DD) of first date to consider for plotting. Leave blank if all years are required.

<code>end_date</code>	Date (YYYY-MM-DD) of last date to consider for plotting. Leave blank if all years are required.
<code>log_discharge</code>	Logical value to indicate plotting the discharge axis (Y-axis) on a logarithmic scale. Default TRUE.
<code>plot_by_year</code>	Logical value to indicate whether to plot each year of data individually. Default FALSE.
<code>one_plot</code>	Logical value to indicate whether to plot all groups/stations on one plot. Default FALSE.
<code>include_title</code>	Logical value to indicate adding the group/station number to the plot, if provided. Default FALSE.

Value

A ggplot2 object of daily flows from `flow_data` or HYDAT flow data provided

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Plot data from a data frame and data argument
flow_data <- tidyhydat::hy_daily_flows(station_number = "08NM116")
plot_flow_data(data = flow_data)

# Plot data directly from HYDAT
plot_flow_data(station_number = "08NM116")

# Plot statistics with custom years
plot_flow_data(station_number = "08NM116",
               start_year = 1981,
               end_year = 2010,
               exclude_years = c(1991,1993:1995))

# Plot data between specific dates
plot_flow_data(station_number = "08NM116",
               start_date = "1990-01-01",
               end_date = "1990-06-01")

# Plot data multiple groups on one plot
plot_flow_data(station_number = c("08NM241", "08NM242"),
               one_plot = TRUE)

}
```

Description

Plots flow duration curves, percent time a flow value is equalled or exceeded, for a streamflow dataset. Plots statistics from all daily discharge values from all years, unless specified. Data calculated using `calc_longterm_stats()` function then converted for plotting.

Usage

```
plot_flow_duration(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  roll_days = 1,
  roll_align = "right",
  water_year_start = 1,
  start_year,
  end_year,
  exclude_years,
  complete_years = FALSE,
  custom_months,
  custom_months_label,
  ignore_missing = FALSE,
  months = 1:12,
  include_longterm = TRUE,
  log_discharge = TRUE,
  include_title = FALSE
)
```

Arguments

<code>data</code>	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
<code>dates</code>	Column in <code>data</code> that contains dates formatted YYYY-MM-DD. Only required if <code>dates</code> column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
<code>values</code>	Column in <code>data</code> that contains numeric flow values, in units of cubic metres per second. Only required if <code>values</code> column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.
<code>groups</code>	Column in <code>data</code> that contains unique identifiers for different data sets, if applicable. Only required if <code>groups</code> column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using <code>station_number</code> argument.
<code>station_number</code>	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT

	database. Requires tidyhydat package and a HYDAT database. Leave blank if using data argument.
roll_days	Numeric value of the number of days to apply a rolling mean. Default 1.
roll_align	Character string identifying the direction of the rolling mean from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'.
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.
start_year	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
end_year	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
exclude_years	Numeric vector of years to exclude from analysis. Leave blank to include all years.
complete_years	Logical values indicating whether to include only years with complete data in analysis. Default FALSE.
custom_months	Numeric vector of months to combine to summarize (ex. 6:8 for Jun-Aug). Adds results to the end of table. If wanting months that overlap calendar years (ex. Oct-Mar), choose water_year_start that begins before the first month listed. Leave blank for no custom month summary.
custom_months_label	Character string to label custom months. For example, if choosing months 7:9 you may choose "Summer" or "Jul-Sep". Default "Custom-Months".
ignore_missing	Logical value indicating whether dates with missing values should be included in the calculation. If TRUE then a statistic will be calculated regardless of missing dates. If FALSE then only statistics from time periods with no missing dates will be returned. Default FALSE.
months	Numeric vector of month curves to plot. NA if no months required. Default 1:12.
include_longterm	Logical value indicating whether to include longterm curve of all data. Default TRUE.
log_discharge	Logical value to indicate plotting the discharge axis (Y-axis) on a logarithmic scale. Default FALSE.
include_title	Logical value to indicate adding the group/station number to the plot, if provided. Default FALSE.

Value

A list of ggplot2 objects with the following for each station provided:

Flow_Duration a plot that contains flow duration curves for each month, long-term, and (option) customized months

See Also[calc_longterm_daily_stats](#)**Examples**

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Plot flow durations using a data frame and data argument with defaults
flow_data <- tidyhydat::hy_daily_flows(station_number = "08NM116")
plot_flow_duration(data = flow_data,
                  start_year = 1980)

# Plot flow durations using station_number argument with defaults
plot_flow_duration(station_number = "08NM116",
                  start_year = 1980)

# Plot flow durations and add custom stats for July-September
plot_flow_duration(station_number = "08NM116",
                  start_year = 1980,
                  custom_months = 7:9,
                  custom_months_label = "Summer")

}
```

plot_longterm_daily_stats

Plot long-term summary statistics from daily mean flows

Description

Plots the long-term monthly mean, median, maximum, minimum, and percentiles of daily flow values for a given month from a streamflow dataset. Plots statistics from discharge values from all years, unless specified. The Maximum-Minimum band can be removed using the 'include_extremes' argument and the percentile bands can be customized using the 'inner_percentiles' and 'outer_percentiles' arguments. Data calculated using the calc_longterm_daily_stats() function.

Usage

```
plot_longterm_daily_stats(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  roll_days = 1,
  roll_align = "right",
  water_year_start = 1,
```

```

start_year,
end_year,
exclude_years,
complete_years = FALSE,
ignore_missing = FALSE,
include_extremes = TRUE,
inner_percentiles = c(25, 75),
outer_percentiles = c(5, 95),
add_year,
log_discharge = TRUE,
include_title = FALSE
)

```

Arguments

<code>data</code>	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
<code>dates</code>	Column in <code>data</code> that contains dates formatted YYYY-MM-DD. Only required if <code>dates</code> column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
<code>values</code>	Column in <code>data</code> that contains numeric flow values, in units of cubic metres per second. Only required if <code>values</code> column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.
<code>groups</code>	Column in <code>data</code> that contains unique identifiers for different data sets, if applicable. Only required if <code>groups</code> column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using <code>station_number</code> argument.
<code>station_number</code>	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires <code>tidyhydat</code> package and a HYDAT database. Leave blank if using <code>data</code> argument.
<code>roll_days</code>	Numeric value of the number of days to apply a rolling mean. Default 1.
<code>roll_align</code>	Character string identifying the direction of the rolling mean from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling <code>n</code> -day group of observations. Default 'right'.
<code>water_year_start</code>	Numeric value indicating the month of the start of the water year for analysis. Default 1.
<code>start_year</code>	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
<code>end_year</code>	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
<code>exclude_years</code>	Numeric vector of years to exclude from analysis. Leave blank to include all years.

complete_years	Logical values indicating whether to include only years with complete data in analysis. Default FALSE.
ignore_missing	Logical value indicating whether dates with missing values should be included in the calculation. If TRUE then a statistic will be calculated regardless of missing dates. If FALSE then only statistics from time periods with no missing dates will be returned. Default FALSE.
include_extremes	Logical value to indicate plotting a ribbon with the range of daily minimum and maximum flows. Default TRUE.
inner_percentiles	Numeric vector of two values of two percentiles indicating the lower and upper limits of the inner percentiles ribbon for plotting. Default c(25, 75), set to NULL for no inner ribbon.
outer_percentiles	Numeric vector of two values of two percentiles indicating the lower and upper limits of the outer percentiles ribbon for plotting. Default c(5, 95), set to NULL for no outer ribbon.
add_year	Numeric value indicating a year of daily flows to add to the daily statistics plot. Leave blank for no years.
log_discharge	Logical value to indicate plotting the discharge axis (Y-axis) on a logarithmic scale. Default FALSE.
include_title	Logical value to indicate adding the group/station number to the plot, if provided. Default FALSE.

Value

A list of ggplot2 objects with the following for each station provided:

Long-term_Monthly_Statistics
a plot that contains long-term flow statistics

Default plots on each object:

Monthly Mean	mean of all annual monthly means for a given month over all years
Monthly Median	median of all annual monthly means for a given month over all years
25-75 Percentiles Range	a ribbon showing the range of data between the monthly 25th and 75th percentiles
5-95 Percentiles Range	a ribbon showing the range of data between the monthly 5th and 95th percentiles
Max-Min Range	a ribbon showing the range of data between the monthly minimum and maximums

See Also

[calc_longterm_daily_stats](#)

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Plot longterm daily statistics using data argument with defaults
flow_data <- tidyhydat::hy_daily_flows(station_number = "08NM116")
plot_longterm_daily_stats(data = flow_data,
                          start_year = 1980)

# Plot longterm daily statistics for water years starting in October
plot_longterm_daily_stats(station_number = "08NM116",
                          start_year = 1980,
                          end_year = 2010,
                          water_year_start = 10)

}
```

```
plot_longterm_monthly_stats
```

Plot long-term summary statistics from annual monthly mean flows

Description

Plots the long-term monthly mean, median, maximum, minimum, and 5, 25, 75, and 95 percentiles of annual monthly mean flow values from a single streamflow dataset. Plots statistics from discharge values from all years, unless specified. The Maximum-Minimum band can be removed using the 'include_extremes' argument and the percentile bands can be customized using the 'inner_percentiles' and 'outer_percentiles' arguments. Data calculated using the `calc_longterm_monthly_stats()` function.

Usage

```
plot_longterm_monthly_stats(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  roll_days = 1,
  roll_align = "right",
  water_year_start = 1,
  start_year,
  end_year,
  exclude_years,
  complete_years = FALSE,
  ignore_missing = FALSE,
  include_extremes = TRUE,
  inner_percentiles = c(25, 75),
```

```

    outer_percentiles = c(5, 95),
    add_year,
    log_discharge = TRUE,
    include_title = FALSE
  )

```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.
groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using <code>station_number</code> argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires <code>tidyhydat</code> package and a HYDAT database. Leave blank if using <code>data</code> argument.
roll_days	Numeric value of the number of days to apply a rolling mean. Default 1.
roll_align	Character string identifying the direction of the rolling mean from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'.
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.
start_year	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
end_year	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
exclude_years	Numeric vector of years to exclude from analysis. Leave blank to include all years.
complete_years	Logical values indicating whether to include only years with complete data in analysis. Default FALSE.
ignore_missing	Logical value indicating whether dates with missing values should be included in the calculation. If TRUE then a statistic will be calculated regardless of missing dates. If FALSE then only statistics from time periods with no missing dates will be returned. Default FALSE.

<code>include_extremes</code>	Logical value to indicate plotting a ribbon with the range of daily minimum and maximum flows. Default TRUE.
<code>inner_percentiles</code>	Numeric vector of two values of two percentiles indicating the lower and upper limits of the inner percentiles ribbon for plotting. Default <code>c(25, 75)</code> , set to NULL for no inner ribbon.
<code>outer_percentiles</code>	Numeric vector of two values of two percentiles indicating the lower and upper limits of the outer percentiles ribbon for plotting. Default <code>c(5, 95)</code> , set to NULL for no outer ribbon.
<code>add_year</code>	Numeric value indicating a year of daily flows to add to the daily statistics plot. Leave blank for no years.
<code>log_discharge</code>	Logical value to indicate plotting the discharge axis (Y-axis) on a logarithmic scale. Default FALSE.
<code>include_title</code>	Logical value to indicate adding the group/station number to the plot, if provided. Default FALSE.

Value

A list of `ggplot2` objects with the following for each station provided:

`Long-term_Monthly_Statistics`
a plot that contains long-term flow statistics

Default plots on each object:

`Monthly Mean` mean of all annual monthly means for a given month over all years
`Monthly Median` median of all annual monthly means for a given month over all years
`25-75 Percentiles Range`
a ribbon showing the range of data between the monthly 25th and 75th percentiles
`5-95 Percentiles Range`
a ribbon showing the range of data between the monthly 5th and 95th percentiles
`Max-Min Range` a ribbon showing the range of data between the monthly minimum and maximums

See Also

[calc_longterm_monthly_stats](#)

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Plot longterm monthly statistics using data argument with defaults
flow_data <- tidyhydat::hy_daily_flows(station_number = "08NM116")
plot_longterm_monthly_stats(data = flow_data,
```

```

        start_year = 1980)

# Plot longterm monthly statistics using station_number argument with defaults
plot_longterm_monthly_stats(station_number = "08NM116",
                           start_year = 1980)

# Plot longterm monthly statistics and add a specific year's daily flows
plot_longterm_monthly_stats(station_number = "08NM116",
                           start_year = 1980,
                           add_year = 1985)

}

```

plot_missing_dates *Plot annual and monthly missing dates*

Description

Plots the number of missing data for each month of each year. Calculates the statistics from all daily discharge values from all years, unless specified. Data calculated using `screen_flow_data()` function.

Usage

```

plot_missing_dates(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  roll_days = 1,
  roll_align = "right",
  water_year_start = 1,
  start_year,
  end_year,
  months = 1:12,
  include_title = FALSE
)

```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.

values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using station_number argument.
groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using station_number argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires tidyhydat package and a HYDAT database. Leave blank if using data argument.
roll_days	Numeric value of the number of days to apply a rolling mean. Default 1.
roll_align	Character string identifying the direction of the rolling mean from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'.
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.
start_year	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
end_year	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
months	Numeric vector of months to include in analysis (e.g.. 6:8 for Jun-Aug). Leave blank to summarize all months (default 1:12).
include_title	Logical value to indicate adding the group/station number to the plot, if provided. Default FALSE.

Value

A list of ggplot2 objects with the following for each station provided:

Missing_Dates a plot that contains the number of missing dates for each year and month

See Also

[screen_flow_data](#)

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Plot missing dates using a data frame and data argument with defaults
flow_data <- tidyhydat::hy_daily_flows(station_number = "08NM116")
plot_missing_dates(data = flow_data)
```



```

# Plot missing dates using station_number argument with defaults
plot_missing_dates(station_number = "08NM116")

# Plot missing dates for water years starting in October
plot_missing_dates(station_number = "08NM116",
                   water_year_start = 9)

# Plot missing dates for 7-day flows for July-September months only
plot_missing_dates(station_number = "08NM116",
                   roll_days = 7,
                   months = 7:9)

}

```

```
plot_monthly_cumulative_stats
```

Plot cumulative monthly flow statistics

Description

Plot the monthly cumulative mean, median, maximum, minimum, and 5, 25, 75, 95th percentiles for each month of the year from a streamflow dataset. Plots the statistics from all monthly cumulative values from all years, unless specified. Data calculated using `calc_monthly_cumulative_stats()` function. Can plot individual years for comparison using the `add_year` argument. Defaults to volumetric cumulative flows, can use `use_yield` and `basin_area` to convert to runoff yield.

Usage

```

plot_monthly_cumulative_stats(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  use_yield = FALSE,
  basin_area,
  water_year_start = 1,
  start_year,
  end_year,
  exclude_years,
  log_discharge = FALSE,
  include_title = FALSE,
  add_year
)

```

Arguments

<code>data</code>	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
<code>dates</code>	Column in <code>data</code> that contains dates formatted YYYY-MM-DD. Only required if <code>dates</code> column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
<code>values</code>	Column in <code>data</code> that contains numeric flow values, in units of cubic metres per second. Only required if <code>values</code> column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.
<code>groups</code>	Column in <code>data</code> that contains unique identifiers for different data sets, if applicable. Only required if <code>groups</code> column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using <code>station_number</code> argument.
<code>station_number</code>	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires <code>tidyhydat</code> package and a HYDAT database. Leave blank if using <code>data</code> argument.
<code>use_yield</code>	Logical value indicating whether to use yield runoff, in mm, instead of volumetric. Default FALSE.
<code>basin_area</code>	Upstream drainage basin area, in square kilometres, to apply to observations. Three options: (1) Leave blank if <code>groups</code> is <code>STATION_NUMBER</code> with HYDAT station numbers to extract basin areas from HYDAT. (2) A single numeric value to apply to all observations. (3) List each basin area for each group/station in <code>groups</code> (can override HYDAT value if listed) as such <code>c("08NM116" = 795, "08NM242" = 10)</code> . If group is not listed the HYDAT area will be applied if it exists, otherwise it will be NA.
<code>water_year_start</code>	Numeric value indicating the month of the start of the water year for analysis. Default 1.
<code>start_year</code>	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
<code>end_year</code>	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
<code>exclude_years</code>	Numeric vector of years to exclude from analysis. Leave blank to include all years.
<code>log_discharge</code>	Logical value to indicate plotting the discharge axis (Y-axis) on a logarithmic scale. Default FALSE.
<code>include_title</code>	Logical value to indicate adding the group/station number to the plot, if provided. Default FALSE.
<code>add_year</code>	Numeric value indicating a year of daily flows to add to the daily statistics plot. Leave blank for no years.

Value

A list of ggplot2 objects with the following for each station provided:

Monthly_Cumulative_Stats

a plot that contains monthly cumulative flow statistics

Default plots on each object:

Mean monthly cumulative mean

Median monthly cumulative median

Min-5 Percentile Range

a ribbon showing the range of data between the monthly cumulative minimum and 5th percentile

5-25 Percentiles Range

a ribbon showing the range of data between the monthly cumulative 5th and 25th percentiles

25-75 Percentiles Range

a ribbon showing the range of data between the monthly cumulative 25th and 75th percentiles

75-95 Percentiles Range

a ribbon showing the range of data between the monthly cumulative 75th and 95th percentiles

95 Percentile-Max Range

a ribbon showing the range of data between the monthly cumulative 95th percentile and the maximum

'Year' Flows (optional) the monthly cumulative flows for the designated year

See Also

[calc_monthly_cumulative_stats](#)

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Plot annual cumulative volume statistics
plot_monthly_cumulative_stats(station_number = "08NM116")

# Plot annual cumulative yield statistics with default HYDAT basin area
plot_monthly_cumulative_stats(station_number = "08NM116",
                              use_yield = TRUE)

# Plot annual cumulative yield statistics with custom basin area
plot_monthly_cumulative_stats(station_number = "08NM116",
                              use_yield = TRUE,
                              basin_area = 800)

}
```

plot_monthly_stats *Plot monthly summary statistics*

Description

Plots monthly mean, median, maximum, minimum, and percentiles for each month of all years of daily flow values from a streamflow dataset. Plots the statistics from all daily discharge values from all years, unless specified. Data calculated using the `calc_monthly_stats()` function.

Usage

```
plot_monthly_stats(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  percentiles,
  roll_days = 1,
  roll_align = "right",
  water_year_start = 1,
  start_year,
  end_year,
  exclude_years,
  months = 1:12,
  ignore_missing = FALSE,
  log_discharge = FALSE,
  include_title = FALSE
)
```

Arguments

<code>data</code>	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
<code>dates</code>	Column in <code>data</code> that contains dates formatted YYYY-MM-DD. Only required if <code>dates</code> column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
<code>values</code>	Column in <code>data</code> that contains numeric flow values, in units of cubic metres per second. Only required if <code>values</code> column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.
<code>groups</code>	Column in <code>data</code> that contains unique identifiers for different data sets, if applicable. Only required if <code>groups</code> column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using <code>station_number</code> argument.

station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires tidyhydat package and a HYDAT database. Leave blank if using data argument.
percentiles	Numeric vector of percentiles to calculate. Set to NA if none required. Default NA.
roll_days	Numeric value of the number of days to apply a rolling mean. Default 1.
roll_align	Character string identifying the direction of the rolling mean from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'.
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.
start_year	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
end_year	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
exclude_years	Numeric vector of years to exclude from analysis. Leave blank to include all years.
months	Numeric vector of months to include in analysis (e.g.. 6:8 for Jun-Aug). Leave blank to summarize all months (default 1:12).
ignore_missing	Logical value indicating whether dates with missing values should be included in the calculation. If TRUE then a statistic will be calculated regardless of missing dates. If FALSE then only statistics from time periods with no missing dates will be returned. Default FALSE.
log_discharge	Logical value to indicate plotting the discharge axis (Y-axis) on a logarithmic scale. Default FALSE.
include_title	Logical value to indicate adding the group/station number to the plot, if provided. Default FALSE.

Value

A list of ggplot2 objects for each monthly statistic for each station provided that contain:

Monthly Mean Flows

mean of all daily flows for a given month and year

Monthly Median Flows

median of all daily flows for a given month and year

Monthly Maximum Flows

maximum of all daily flows for a given month and year

Monthly Minimum Flows

minimum of all daily flows for a given month and year

Monthly P'n' Flows

(optional) each n-th percentile selected for a given month and year

See Also[calc_monthly_stats](#)**Examples**

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Plot monthly statistics using a data frame and data argument with defaults
flow_data <- tidyhydat::hy_daily_flows(station_number = "08NM116")
results <- plot_monthly_stats(data = flow_data,
                             start_year = 1980,
                             percentiles = 10)

# Plot monthly statistics for water years starting in October
results <- plot_monthly_stats(station_number = "08NM116",
                             start_year = 1980,
                             end_year = 2010,
                             water_year_start = 10,
                             percentiles = 10)

}
```

screen_flow_data

Calculate annual summary and missing data statistics

Description

Calculates mean, median, maximum, minimum, standard deviation of annual flows and data availability and missing data statistics for each year and month of each year. Calculates the statistics from all daily discharge values from all years, unless specified.

Usage

```
screen_flow_data(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  roll_days = 1,
  roll_align = "right",
  water_year_start = 1,
  start_year,
  end_year,
  months = 1:12,
  transpose = FALSE
)
```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.
groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using <code>station_number</code> argument.
station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires <code>tidyhydat</code> package and a HYDAT database. Leave blank if using <code>data</code> argument.
roll_days	Numeric value of the number of days to apply a rolling mean. Default 1.
roll_align	Character string identifying the direction of the rolling mean from the specified date, either by the first ('left'), last ('right'), or middle ('center') day of the rolling n-day group of observations. Default 'right'.
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.
start_year	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
end_year	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
months	Numeric vector of months to include in analysis (e.g.. 6:8 for Jun-Aug). Leave blank to summarize all months (default 1:12).
transpose	Logical value indicating if the results rows and columns are to be switched. Default FALSE.

Value

A tibble data frame with the following columns:

Year	calendar or water year selected
n_days	number of days per year
n_Q	number of days per year with flow data
n_missing_Q	number of days per year with no flow data
Minimum	annual minimum of all daily flows for a given year

Maximum annual maximum of all daily flows for a given year
 Mean annual mean of all daily flows for a given year
 Median annual median of all daily flows for a given year
 StandardDeviation annual 1 standard deviation of all daily flows for a given year

and the following monthly missing columns (order will depend on water_year_month):

Jan_missing_Q number of Jan days per year with no flow data
 Feb_missing_Q number of Feb days per year with no flow data
 Mar_missing_Q number of Mar days per year with no flow data
 Apr_missing_Q number of Apr days per year with no flow data
 May_missing_Q number of May days per year with no flow data
 Jun_missing_Q number of Jun days per year with no flow data
 Jul_missing_Q number of Jul days per year with no flow data
 Aug_missing_Q number of Aug days per year with no flow data
 Sep_missing_Q number of Sep days per year with no flow data
 Oct_missing_Q number of Oct days per year with no flow data
 Nov_missing_Q number of Nov days per year with no flow data
 Dec_missing_Q number of Dec days per year with no flow data

Transposing data creates a column of "Statistics" and subsequent columns for each year selected.

Examples

```
# Run if HYDAT database has been downloaded (using tidyhydat::download_hydat())
if (file.exists(tidyhydat::hy_downloaded_db())) {

# Calculate screening statistics using a data frame and data argument with defaults
flow_data <- tidyhydat::hy_daily_flows(station_number = "08NM116")
screen_flow_data(data = flow_data)

# Calculate screening statistics using station_number argument with defaults
screen_flow_data(station_number = "08NM116")

# Calculate screening statistics for water years starting in October
screen_flow_data(station_number = "08NM116",
                 water_year_start = 9)

# Calculate screening statistics for 7-day flows for July-September months only
screen_flow_data(station_number = "08NM116",
                 roll_days = 7,
                 months = 7:9)

}
```

write_flow_data	<i>Write a streamflow dataset as a .xlsx, .xls, or .csv file</i>
-----------------	--

Description

Write a streamflow dataset to a directory. Can fill missing dates or filter data by years or dates before writing using given arguments. Just list data frame or HYDAT station number to write its entirety. Can write as .xls, .xlsx, or .csv file types. Writing as Excel file type uses the 'writexl' package.

Usage

```
write_flow_data(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  water_year_start = 1,
  start_year,
  end_year,
  start_date,
  end_date,
  file_name,
  fill_missing = FALSE,
  digits
)
```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.
groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'. Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using <code>station_number</code> argument.

station_number	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires tidyhydat package and a HYDAT database. Leave blank if using data argument.
water_year_start	Numeric value indicating the month of the start of the water year for analysis. Default 1.
start_year	Numeric value of the first year of data to write. Leave blank to use the first year of the source data.
end_year	Numeric value of the last year of data to write. Leave blank to use the last year of the source data.
start_date	Date (YYYY-MM-DD) of first date of data to write. Leave blank if all dates required.
end_date	Date (YYYY-MM-DD) of last date of data to write. Leave blank if all dates required.
file_name	Character string naming the output file. If none provided, a default file name (with .xlsx) is provided (see "Successfully created" message when using function for file name).
fill_missing	Logical value indicating whether to fill dates with missing flow data with NA. Default FALSE.
digits	Integer indicating the number of decimal places or significant digits used to round flow values. Use follows that of base::round() digits argument.

Examples

```
## Not run:

# Working examples:

# Write data from a data frame
flow_data <- tidyhydat::hy_daily_flows(station_number = "08NM116")
write_flow_data(data = flow_data,
                file_name = "Mission_Creek_daily_flows.xlsx")

# Write data directly from HYDAT
write_flow_data(station_number = "08NM116",
                file_name = "Mission_Creek_daily_flows.xlsx")

# Write data directly from HYDAT and fill missing dates with NA
write_flow_data(station_number = "08NM116",
                file_name = "Mission_Creek_daily_flows.xlsx",
                fill_missing = TRUE)

## End(Not run)
```

write_full_analysis *Write a suite of tables and plots from various fasstr functions into a directory*

Description

Calculates and writes tables and plots from a suite of statistics from fasstr functions into an Excel workbook, and accompanying plot files for certain analyses. Due to the number of tables and plots to be made, this function may take several minutes to complete. If using `ignore_missing = FALSE` (default) and there is missing data, some tables and plots may be empty and produce warnings. Use `ignore_missing = TRUE` to ignore the missing values or filter your data to complete years.

Usage

```
write_full_analysis(
  data,
  dates = Date,
  values = Value,
  groups = STATION_NUMBER,
  station_number,
  analyses = 1:7,
  basin_area,
  water_year_start = 1,
  start_year,
  end_year,
  exclude_years,
  ignore_missing = FALSE,
  zyp_method = "yuepilon",
  zyp_alpha,
  file_name,
  plot_filetype = "pdf"
)
```

Arguments

data	A data frame of daily data that contains columns of dates, flow values, and (optional) groups (e.g. station numbers). Leave blank if using <code>station_number</code> argument.
dates	Column in data that contains dates formatted YYYY-MM-DD. Only required if dates column name is not 'Date' (default). Leave blank if using <code>station_number</code> argument.
values	Column in data that contains numeric flow values, in units of cubic metres per second. Only required if values column name is not 'Value' (default). Leave blank if using <code>station_number</code> argument.
groups	Column in data that contains unique identifiers for different data sets, if applicable. Only required if groups column name is not 'STATION_NUMBER'.

	Function will automatically group by a column named 'STATION_NUMBER' if present. Remove the 'STATION_NUMBER' column beforehand to remove this grouping. Leave blank if using <code>station_number</code> argument.
<code>station_number</code>	A character string vector of seven digit Water Survey of Canada station numbers (e.g. "08NM116") of which to extract daily streamflow data from a HYDAT database. Requires <code>tidyhydat</code> package and a HYDAT database. Leave blank if using <code>data</code> argument.
<code>analyses</code>	Numeric vector of analyses to run (default is all (1:7)): <ul style="list-style-type: none"> • 1: Screening • 2: Long-term • 3: Annual • 4: Monthly • 5: Daily • 6: Annual Trends • 7: Low-flow Frequencies
<code>basin_area</code>	Upstream drainage basin area, in square kilometres, to apply to observations. Three options: <ol style="list-style-type: none"> (1) Leave blank if <code>groups</code> is <code>STATION_NUMBER</code> with HYDAT station numbers to extract basin areas from HYDAT. (2) A single numeric value to apply to all observations. (3) List each basin area for each group/station in <code>groups</code> (can override HYDAT value if listed) as such <code>c("08NM116" = 795, "08NM242" = 10)</code>. If <code>group</code> is not listed the HYDAT area will be applied if it exists, otherwise it will be NA.
<code>water_year_start</code>	Numeric value indicating the month of the start of the water year for analysis. Default 1.
<code>start_year</code>	Numeric value of the first year to consider for analysis. Leave blank to use the first year of the source data.
<code>end_year</code>	Numeric value of the last year to consider for analysis. Leave blank to use the last year of the source data.
<code>exclude_years</code>	Numeric vector of years to exclude from analysis. Leave blank to include all years.
<code>ignore_missing</code>	Logical value indicating whether dates with missing values should be included in the calculation. If TRUE then a statistic will be calculated regardless of missing dates. If FALSE then only statistics from time periods with no missing dates will be returned. Default FALSE.
<code>zyp_method</code>	Character string identifying the prewhitened trend method to use from 'zyp', either "zhang" or "yuepilon". Only required if analysis group 6 is included. Default "yuepilon".
<code>zyp_alpha</code>	Numeric value of the significance level (ex. 0.05) of when to plot a trend line. Leave blank for no line.
<code>file_name</code>	Character string of the name of the Excel Workbook (and folder for plots if necessary) to create on drive to write all results.
<code>plot_filetype</code>	Image type to write. One of "png", "eps", "ps", "tex", "pdf", "jpeg", "tiff", "bmp", or "svg". If not "pdf" then individual plots will be created instead of a combined PDF. Default "pdf".

See Also

compute_full_analysis, screen_flow_data, plot_data_screening, plot_missing_dates, calc_longterm_monthly_stats, plot_longterm_monthly_stats, calc_longterm_daily_stats, plot_longterm_daily_stats, plot_flow_duration, calc_annual_stats, plot_annual_stats, calc_annual_cumulative_stats, plot_annual_cumulative_stats, calc_annual_flow_timing, plot_annual_flow_timing, calc_annual_outside_normal, plot_annual_outside_normal, calc_annual_lowflows, plot_annual_lowflows, plot_annual_means, calc_monthly_stats, plot_monthly_stats, calc_monthly_cumulative_stats, plot_monthly_cumulative_stats, calc_daily_stats, plot_daily_stats, calc_daily_cumulative_stats, plot_daily_cumulative_stats, compute_annual_trends, compute_annual_frequencies, write_flow_data, write_plots

Examples

```
## Not run:

# Working examples:

# Save a full analysis with all the analyses
write_full_analysis(station_number = "08NM116",
                    file_name = "Mission Creek",
                    start_year = 1980,
                    end_year = 2010)

# Save a full analysis with only Annual and Daily analyses
write_full_analysis(station_number = "08NM116",
                    file_name = "Mission Creek",
                    start_year = 1980,
                    end_year = 2010,
                    analyses = c(3,5))

## End(Not run)
```

write_objects_list *Write all data frames and plots from a list of objects into a directory*

Description

Write a list of tables (data.frames) and plots (ggplots; as used by fasstr) into a directory. Objects that are not class "data.frame" or "gg" will not be saved. Each table and plot will be named by the object name in the list.

Usage

```
write_objects_list(
  list,
  folder_name,
  table_filetype,
  plot_filetype,
```

```

    width,
    height,
    units = "in",
    dpi = 300
  )

```

Arguments

<code>list</code>	List of data frames and plots to write to disk.
<code>folder_name</code>	Name of folder to create on disk (if it does not exist) to write each plot from list. If using <code>combined_pdf</code> argument, then it will be the name of the PDF document.
<code>table_filetype</code>	Table file type to write. One of "csv", "xls", or "xlsx".
<code>plot_filetype</code>	Image type to write. One of "png", "eps", "ps", "tex", "pdf", "jpeg", "tiff", "bmp", or "svg". Image type will be overwritten if using <code>combined_pdf</code> is used.
<code>width</code>	Numeric plot width in units. If not supplied, uses the size of current graphics device.
<code>height</code>	Numeric plot height in units. If not supplied, uses the size of current graphics device.
<code>units</code>	Character string plot height and width units, one of "in", "cm", or "mm". Default "in".
<code>dpi</code>	Numeric resolution of plots. Default 300.

Examples

```

## Not run:

# Working examples:

# Example list of tables and plots to save
frequency <- compute_annual_frequencies(station_number = "08NM116")

# Write objects in a folder
write_objects_list(list = frequency,
                  folder_name = "Frequency Analysis",
                  table_filetype = "xlsx",
                  plot_filetype = "png")

## End(Not run)

```

Description

Write a list of plots (ggplots; as used by `fasstr`) into a directory or PDF document. When writing into a named directory each plot will be named by the plot name listed in the list; uses `ggplot2::ggsave` function. When writing into a PDF document (`combined_pdf == TRUE`) the plot names will not appear; uses `grDevices:pdf` function.

Usage

```
write_plots(  
  plots,  
  folder_name,  
  plot_filetype,  
  width,  
  height,  
  units = "in",  
  dpi = 300,  
  combined_pdf = FALSE  
)
```

Arguments

<code>plots</code>	List of plots to write to disk.
<code>folder_name</code>	Name of folder to create on disk (if it does not exist) to write each plot from list. If using <code>combined_pdf</code> argument, then it will be the name of the PDF document.
<code>plot_filetype</code>	Image type to write. One of "png", "eps", "ps", "tex", "pdf", "jpeg", "tiff", "bmp", or "svg". Image type will be overwritten if using <code>combined_pdf</code> is used.
<code>width</code>	Numeric plot width in units. If not supplied, uses the size of current graphics device.
<code>height</code>	Numeric plot height in units. If not supplied, uses the size of current graphics device.
<code>units</code>	Character string plot height and width units, one of "in", "cm", or "mm". Default "in".
<code>dpi</code>	Numeric resolution of plots. Default 300.
<code>combined_pdf</code>	Logical value indicating whether to combine list of plots into one pdf document. Default FALSE.

Examples

```
## Not run:  
  
# Working examples:  
  
# Example plots to save  
plots <- plot_annual_lowflows(station_number = "08NM116")  
  
# Write the plots as "png" files  
write_plots(plots = plots,
```

```

        folder_name = "Low Flow Plots",
        plot_filetype = "png")

# Write the plots as a combined "pdf" document
write_plots(plots = plots,
            folder_name = "Low Flow Plots",
            combined_pdf = TRUE)

## End(Not run)

```

write_results *Write a data frame as a .xlsx, .xls, or .csv file*

Description

Write a data frame to a directory with all numbers rounded to specified digits. Can write as .xls, .xlsx, or .csv file types. Writing as .xlsx or .xls uses the 'writexl' package.

Usage

```
write_results(data, file_name, digits)
```

Arguments

data	Data frame to be written to a directory.
file_name	Character string naming the output file. Required.
digits	Integer indicating the number of decimal places or significant digits used to round flow values. Use follows that of base::round() digits argument.

Examples

```

## Not run:

# Working examples:

# Example data to write
data_results <- calc_longterm_daily_stats(station_number = c("08HA002", "08HA011"),
                                         start_year = 1971, end_year = 2000)

# Write the data and round numbers to 1 decimal place
write_results(data = data_results,
             file_name = "Cowichan River Long-term Flows (1971-2000).xlsx",
             digits = 1)

## End(Not run)

```


Index

add_basin_area, 3
add_cumulative_volume, 4
add_cumulative_yield, 6
add_daily_volume, 7
add_daily_yield, 8
add_date_variables, 10
add_rolling_means, 11
add_seasons, 13

calc_all_annual_stats, 14, 59
calc_annual_cumulative_stats, 16, 17, 66, 72, 117
calc_annual_flow_timing, 16, 20, 66, 74, 117
calc_annual_lowflows, 16, 22, 66, 76, 117
calc_annual_outside_normal, 16, 25, 66, 81, 117
calc_annual_stats, 16, 27, 66, 78, 83, 117
calc_daily_cumulative_stats, 29, 66, 86, 117
calc_daily_stats, 32, 66, 89, 117
calc_flow_percentile, 35
calc_longterm_daily_stats, 37, 66, 97, 99, 117
calc_longterm_mean, 40
calc_longterm_monthly_stats, 42, 66, 102, 117
calc_longterm_percentile, 45
calc_monthly_cumulative_stats, 47, 66, 107, 117
calc_monthly_stats, 16, 50, 66, 110, 117
compute_annual_frequencies, 53, 66, 117
compute_annual_trends, 56, 66, 117
compute_frequency_analysis, 55, 59, 63, 68
compute_frequency_quantile, 61
compute_full_analysis, 64, 117
compute_hydat_peak_frequencies, 66

fill_missing_dates, 69

plot_annual_cumulative_stats, 66, 70, 117
plot_annual_flow_timing, 66, 73, 117
plot_annual_lowflows, 66, 75, 117
plot_annual_means, 66, 77, 117
plot_annual_outside_normal, 66, 79, 117
plot_annual_stats, 66, 81, 117
plot_daily_cumulative_stats, 66, 84, 117
plot_daily_stats, 66, 87, 117
plot_data_screening, 66, 90, 117
plot_flow_data, 66, 92
plot_flow_duration, 66, 94, 117
plot_longterm_daily_stats, 66, 97, 117
plot_longterm_monthly_stats, 66, 100, 117
plot_missing_dates, 66, 103, 117
plot_monthly_cumulative_stats, 66, 105, 117
plot_monthly_stats, 66, 108, 117

screen_flow_data, 66, 92, 104, 110, 117

write_flow_data, 66, 113, 117
write_full_analysis, 115
write_objects_list, 117
write_plots, 66, 117, 118
write_results, 120

zyp.trend.dataframe, 59