

# Package ‘puzzle’

November 28, 2019

**Type** Package

**Title** Assembling Data Sets for Non-Linear Mixed Effects Modeling

**Version** 0.0.1

**Maintainer** Mario Gonzalez Sales <mario@modelinggreatsolutions.com>

**Description**

To Simplify the time consuming and error prone task of assembling complex data sets for non-linear mixed effects modeling. Users are able to select from different absorption processes such as zero and first order, or a combination of both. Furthermore, data sets containing data from several entities, responses, and covariates can be simultaneously assembled.

**License** GPL-3

**Encoding** UTF-8

**LazyData** true

**Imports** utils, lubridate, stats, readxl, reshape, reshape2, sqldf,  
kableExtra, plyr, dplyr, tidyverse, readr

**Suggests** rmarkdown, knitr, devtools, testthat

**RoxygenNote** 6.1.1

**URL** <https://github.com/syneoshealth/puzzle>

**BugReports** <https://github.com/syneoshealth/puzzle/issues>

**NeedsCompilation** no

**Author** Olivier Barriere [aut],  
Mario Gonzalez Sales [aut, cre]

**Repository** CRAN

**Date/Publication** 2019-11-28 16:10:02 UTC

## R topics documented:

|                                       |   |
|---------------------------------------|---|
| df_cov . . . . .                      | 2 |
| df_cov_start . . . . .                | 3 |
| df_cov_time_dependent_start . . . . . | 3 |
| df_dose . . . . .                     | 4 |

|   |    |
|---|----|
| df_dose_datetime . . . . .                | 4  |
| df_dose_evid4 . . . . .                   | 5  |
| df_dose_optional_columns . . . . .        | 5  |
| df_dose_start . . . . .                   | 6  |
| df_extra_times . . . . .                  | 6  |
| df_extra_times_datetime . . . . .         | 7  |
| df_extra_times_metabolite_evid4 . . . . . | 7  |
| df_extra_times_parent_evid4 . . . . .     | 8  |
| df_extra_times_time . . . . .             | 8  |
| df_metabolite_evid4 . . . . .             | 9  |
| df_parent_evid4 . . . . .                 | 9  |
| df_pd_start . . . . .                     | 10 |
| df_pk . . . . .                           | 10 |
| df_pk_datetime . . . . .                  | 11 |
| df_pk_metabolite . . . . .                | 11 |
| df_pk_optional_columns . . . . .          | 12 |
| df_pk_parent . . . . .                    | 12 |
| df_pk_start . . . . .                     | 13 |
| df_response1 . . . . .                    | 13 |
| df_response2 . . . . .                    | 14 |
| df_response3 . . . . .                    | 14 |
| puzzle . . . . .                          | 15 |

|              |           |
|--------------|-----------|
| <b>Index</b> | <b>17</b> |
|--------------|-----------|

---

|        |                              |
|--------|------------------------------|
| df_cov | <i>A covariate data set.</i> |
|--------|------------------------------|

---

## Description

A dataset containing covariate information.

## Usage

df\_cov

## Format

A tibble with 12 rows and 4 variables:

**ID** Individual

**TIME** Time, in hours

**VARIABLE** Variable

**VALUE** Value of the variable

---

|              |                                     |
|--------------|-------------------------------------|
| df_cov_start | <i>Starting covariate data set.</i> |
|--------------|-------------------------------------|

---

**Description**

A dataset containing covariate information.

**Usage**

df\_cov\_start

**Format**

A data frame with 4 rows and 3 variables:

**ID** Individual

**VARIABLE** Variable

**VALUE** Value of the variable

---

|                             |  |
|-----------------------------|--|
| df_cov_time_dependent_start | <i>A covariate data set to be used with time dependent covariates.</i> |
|-----------------------------|--|

---

**Description**

A dataset containing time dependent covariates.

**Usage**

df\_cov\_time\_dependent\_start

**Format**

A data frame with 6 rows and 4 variables:

**ID** Individual

**VARIABLE** Variable

**VALUE** Value of the variable

**TIME** Time, in hours

---

|         |                         |
|---------|-------------------------|
| df_dose | <i>A dose data set.</i> |
|---------|-------------------------|

---

**Description**

A dataset containing dose information.

**Usage**

df\_dose

**Format**

A data frame with 12 rows and 3 variables:

**ID** Individual

**TIME** Time, in weeks

**AMT** Dose, in mg

---

|                  |   |
|------------------|---|
| df_dose_datetime | <i>A dose data set including datetimes.</i> |
|------------------|---|

---

**Description**

A dataset containing dose information in datetime format.

**Usage**

df\_dose\_datetime

**Format**

A data frame with 5 rows and 12 variables:

**ID** Individual

**TRT** Treatment label

**DOSE** Dose, in mg

**PERIOD** Period

**DAY** Day of administration

**AMT** Dose, in mg

**DATETIME** Date in datetime format

**TIMEPOINT** Timepoint

**COHORT** Cohort

**FORM** Drug form

**TREATMENT** Treatment

**FOOD** Food status

---

|               |  |
|---------------|--|
| df_dose_evid4 | <i>A dose data set to be used with EVID=4.</i> |
|---------------|--|

---

**Description**

A dataset containing dosing information.

**Usage**

df\_dose\_evid4

**Format**

A data frame with 418 rows and 10 variables:

**ID** Individual

**PERIOD** Period

**TIMEPOINT** Timepoint

**TIME** Time, in hours

**AMT** Dose, in mg

**TRT** Treatment label

**DAY** Day of administration

**SEQUENCE** Sequence

**TRT2** Treatment

**EVID** Evid value

---

|                          |  |
|--------------------------|--|
| df_dose_optional_columns | <i>A dose data set to be used with optional columns.</i> |
|--------------------------|--|

---

**Description**

A dataset containing dosing information.

**Usage**

df\_dose\_optional\_columns

**Format**

A data frame with 4 rows and 6 variables:

**ID** Individual

**TIME** Time, in hours

**AMT** Dose, in mg

**OCC** Occasion

**TIMEPOINT** Timepoint

**TRT** Treatment

---

df\_dose\_start            *A dose data set example.*

---

**Description**

A dataset containing dosing information.

**Usage**

df\_dose\_start

**Format**

A data frame with 4 rows and 3 variables:

**ID** Individual

**TIME** Time, in hours

**AMT** Dose, in mg

---

df\_extra\_times            *An extra times data set example.*

---

**Description**

A dataset containing extra times.

**Usage**

df\_extra\_times

**Format**

A data frame with 251 rows and 1 variable:

**TIME** Time, in hours

---

df\_extra\_times\_datetime

*An extra times data set example with datetime format.*

---

**Description**

A dataset containing extra times in datetime format.

**Usage**

df\_extra\_times\_datetime

**Format**

A data frame with 20 rows and 1 variable:

**ID** Individual

**DATETIME** Datetime

**TIMEPOINT** Timepoint

---

df\_extra\_times\_metabolite\_evid4

*An extra times metabolite data set to be used with EVID=4.*

---

**Description**

A dataset containing extra times for an hypothetical metabolite.

**Usage**

df\_extra\_times\_metabolite\_evid4

**Format**

A data frame with 770 rows and 3 variable:

**PERIOD** Period

**TIMEPOINT** Timepoint

**TIME** Time, in hours

---

df\_extra\_times\_parent\_evid4

*An extra times parent data set to be used with EVID=4.*

---

### Description

A dataset containing extra times for an hypothetical parent drug.

### Usage

df\_extra\_times\_parent\_evid4

### Format

A data frame with 770 rows and 3 variable:

**PERIOD** Period

**TIMEPOINT** Timepoint

**TIME** Time, in hours

---

df\_extra\_times\_time *An extra times data set example.*

---

### Description

A dataset containing extra times.

### Usage

df\_extra\_times\_time

### Format

A data frame with 1040 rows and 3 variable:

**ID** Individual

**TIME** Time, in hours

**TIMEPOINT** Timepoint



---

df\_metabolite\_evid4     *A pharmacokinetic metabolite data set to be used with EVID=4.*

---

**Description**

A dataset containing pharmacokinetic information for an hypothetical metabolite.

**Usage**

df\_metabolite\_evid4

**Format**

A data frame with 1359 rows and 7 variables:

**ID** Individual

**PERIOD** Period

**TIMEPOINT** Timepoint

**TIME** Time, in hours

**DV** Drug concentration, in mg/L

**TIMEDAY** Timeday

**DAY** Day of administration

---

df\_parent\_evid4     *A pharmacokinetic parent data set to be used with EVID=4.*

---

**Description**

A dataset containing pharmacokinetic information for an hypothetical parent drug.

**Usage**

df\_parent\_evid4

**Format**

A data frame with 1359 rows and 7 variables:

**ID** Individual

**PERIOD** Period

**TIMEPOINT** Timepoint

**TIME** Time, in hours

**DV** Drug concentration, in mg/L

**TIMEDAY** Timeday

**DAY** Day of administration

---

|             |  |
|-------------|--|
| df_pd_start | <i>An starting pharmacodynamic data set example.</i> |
|-------------|--|

---

**Description**

A dataset containing pharmacodynamic observations.

**Usage**

df\_pd\_start

**Format**

A tibble with 6 rows and 3 variable:

**ID** Individual

**TIME** Time, in hours

**DV** Response, ng/mL

---

|       |                                    |
|-------|------------------------------------|
| df_pk | <i>A pharmacokinetic data set.</i> |
|-------|------------------------------------|

---

**Description**

A dataset containing pharmacokinetic information.

**Usage**

df\_pk

**Format**

A tibble with 132 rows and 4 variable:

**ID** Individual

**TIMEPOINT** Timepoint

**TIME** Time, in hours

**DV** Drug concentration, ng/mL

---

df\_pk\_datetime      *A pharmacokinetic data set example in datetime format.*

---

**Description**

A dataset containing pharmacokinetic information.

**Usage**

df\_pk\_datetime

**Format**

A data frame with 65 rows and 7 variable:

**ID** Individual

**DV** Response, ng/mL

**DATETIME** Datetime

**TIMEPOINT** Timepoint

**DAY** Day

**PERIOD** Period

**BLQ** I a BLQ?

**LLOQ** Lower limit of quantification, ng/mL

---

df\_pk\_metabolite      *A pharmacokinetic data set of metabolite data.*

---

**Description**

A dataset containing pharmacokinetic information for an hypothetical metabolite.

**Usage**

df\_pk\_metabolite

**Format**

A data frame with 10 rows and 3 variable:

**ID** Individual

**TIME** Time, in hours

**DV** Drug concentration, ng/mL

---

df\_pk\_optional\_columns

*A pharmacokinetic data set to be used with optional columns.*

---

**Description**

A dataset containing pharmacokinetic information.

**Usage**

df\_pk\_optional\_columns

**Format**

A data frame with 12 rows and 5 variable:

**ID** Individual

**TIME** Time, in hours

**DV** Drug concentration, ng/mL

**OCC** Occasion

**TIMEPOINT** Timepoint

---

df\_pk\_parent

*A pharmacokinetic data set for an hypothetical parent drug.*

---

**Description**

A dataset containing pharmacokinetic information.

**Usage**

df\_pk\_parent

**Format**

A data frame with 12 rows and 3 variable:

**ID** Individual

**TIME** Time, in hours

**DV** Drug concentration, ng/mL

---

|             |  |
|-------------|--|
| df_pk_start | <i>A pharmacokinetic data set example.</i> |
|-------------|--|

---

**Description**

A dataset containing pharmacokinetic information.

A dataset containing pharmacokinetic information.

**Usage**

```
df_pk_start
```

```
df_pk_start
```

**Format**

A tibble with 12 rows and 3 variable:

**ID** Individual

**TIME** Time, in hours

**DV** Response, ng/mL

---

|              |                                    |
|--------------|------------------------------------|
| df_response1 | <i>A pharmacodynamic data set.</i> |
|--------------|------------------------------------|

---

**Description**

A dataset containing pharmacodynamic information for response 1.

**Usage**

```
df_response1
```

**Format**

A data frame with 6 rows and 3 variable:

**ID** Individual

**TIME** Time, in hours

**DV** Response, ng/mL

---

|              |                                    |
|--------------|------------------------------------|
| df_response2 | <i>A pharmacodynamic data set.</i> |
|--------------|------------------------------------|

---

**Description**

A dataset containing pharmacodynamic information for response 2.

**Usage**

df\_response2

**Format**

A data frame with 6 rows and 3 variable:

**ID** Individual

**TIME** Time, in hours

**DV** Response, seconds

---

|              |                                    |
|--------------|------------------------------------|
| df_response3 | <i>A pharmacodynamic data set.</i> |
|--------------|------------------------------------|

---

**Description**

A dataset containing pharmacodynamic information for response 3.

**Usage**

df\_response3

**Format**

A data frame with 6 rows and 3 variable:

**ID** Individual

**TIME** Time, in hours

**DV** Response, in hours

---

puzzle                      *puzzle*

---

## Description

Build pharmacometric data sets from basic tabulated files

## Usage

```
puzzle(directory = NULL, order, coercion = list(name = NULL, sep =
  ","), optionalcolumns = NULL, pk = list(name = NULL, data = NULL),
  dose = list(name = NULL, data = NULL), cov = list(name = NULL, data =
  NULL), pd = list(name = NULL, data = NULL), extratimes = list(name =
  NULL, data = NULL), nm = list(name = NULL), fillcolumns = NULL,
  nocoercioncolumns = NULL, norepeatcolumns = NULL, initialindex = 0,
  na.strings = "N/A", arrange = "ID,TIME,CMT,desc(EVID)",
  datetimeformat = "%Y-%m-%d %H:%M:%S", timeunits = "hours",
  timezone = Sys.timezone(), ignore = "C", missingvalues = ".",
  parallel = TRUE, verbose = FALSE, username = NULL)
```

## Arguments

|                   |  |
|-------------------|--|
| directory         | path to your directory   |
| order             | define the absorption order, can be 0, 1, c(0,1), or c(1,1)  |
| coercion          | define name for coercion file  |
| optionalcolumns   | define optional columns  |
| pk                | define the required file containing the pk information. It can be a .csv or an .xlsx file                      |
| dose              | define the required file containing the dose information. It can be a .csv, an .xlsx file or an R object.      |
| cov               | define the optional file containing the covariate information. It can be a .csv, an .xlsx file or an R object. |
| pd                | define the optional file containing the pd information. It can be a .csv, or a .xlsx file.                     |
| extratimes        | define the optional file containing the additional times. It can be a .csv, or a .xlsx file.                   |
| nm                | name of output file generated by puzzle  |
| fillcolumns       | define columns to be filled  |
| nocoercioncolumns | define columns to be dropped from the coercion file  |
| norepeatcolumns   | define columns not to be repeated  |
| initialindex      | define the lower category of categorical covariates  |

|                |   |
|----------------|---|
| na.strings     | define value for na                           |
| arrange        | define how the columns should be arranged     |
| datetimeformat | define format for date times                  |
| timeunits      | define time units if needed                   |
| timezone       | define timezone                               |
| ignore         | define ignore value                           |
| missingvalues  | define missing value                          |
| parallel       | define parallel zero + first order absorption |
| verbose        | define verbose                                |
| username       | define person performing the assembling       |

**Value**

a pharmacometrics ready data set

**Examples**

```
## Not run:
nm = list(pk = list(parent=as.data.frame(puzzle::df_pk_start)),
          dose=as.data.frame(puzzle::df_dose_start),
          cov=as.data.frame(puzzle::df_cov_start))
puzzle(directory=file.path(tempdir()),
        order=c(0),
        pk=list(data=nm$pk),
        dose=list(data=nm$dose),
        cov=list(data=nm$cov))

## End(Not run)
```



# Index

## \*Topic **datasets**

df\_cov, [2](#)  
df\_cov\_start, [3](#)  
df\_cov\_time\_dependent\_start, [3](#)  
df\_dose, [4](#)  
df\_dose\_datetime, [4](#)  
df\_dose\_evid4, [5](#)  
df\_dose\_optional\_columns, [5](#)  
df\_dose\_start, [6](#)  
df\_extra\_times, [6](#)  
df\_extra\_times\_datetime, [7](#)  
df\_extra\_times\_metabolite\_evid4, [7](#)  
df\_extra\_times\_parent\_evid4, [8](#)  
df\_extra\_times\_time, [8](#)  
df\_metabolite\_evid4, [9](#)  
df\_parent\_evid4, [9](#)  
df\_pd\_start, [10](#)  
df\_pk, [10](#)  
df\_pk\_datetime, [11](#)  
df\_pk\_metabolite, [11](#)  
df\_pk\_optional\_columns, [12](#)  
df\_pk\_parent, [12](#)  
df\_pk\_start, [13](#)  
df\_response1, [13](#)  
df\_response2, [14](#)  
df\_response3, [14](#)

df\_metabolite\_evid4, [9](#)  
df\_parent\_evid4, [9](#)  
df\_pd\_start, [10](#)  
df\_pk, [10](#)  
df\_pk\_datetime, [11](#)  
df\_pk\_metabolite, [11](#)  
df\_pk\_optional\_columns, [12](#)  
df\_pk\_parent, [12](#)  
df\_pk\_start, [13](#)  
df\_response1, [13](#)  
df\_response2, [14](#)  
df\_response3, [14](#)

puzzle, [15](#)

df\_cov, [2](#)  
df\_cov\_start, [3](#)  
df\_cov\_time\_dependent\_start, [3](#)  
df\_dose, [4](#)  
df\_dose\_datetime, [4](#)  
df\_dose\_evid4, [5](#)  
df\_dose\_optional\_columns, [5](#)  
df\_dose\_start, [6](#)  
df\_extra\_times, [6](#)  
df\_extra\_times\_datetime, [7](#)  
df\_extra\_times\_metabolite\_evid4, [7](#)  
df\_extra\_times\_parent\_evid4, [8](#)  
df\_extra\_times\_time, [8](#)