FOG

Christophe Mpaga

Contents

Executive summary

In this project, we will be using data from kaggle.We aim to analysis time to freezing of gait (FOG). FOG is a pattern occurring in patient with Parkinson diseases. It indicates kinetic inability and impairment during gait for instance. Some indicative events like walking hesitation, turning body could be observed and help to detect FOG occurence.

We have 3 tables to analyse :
- events
- subjects
- tasks

Data analysis : 1D EDA

events

data structure

```
## # A tibble: 3 x 5
##   Id          Init Completion Type  Kinetic
##   <chr>      <dbl>      <dbl> <chr>   <dbl>
## 1 003f117e14  8.61       14.8 Turn        1
## 2 009ee11563 11.4        41.2 Turn        1
## 3 009ee11563 54.7        58.8 Turn        1
```

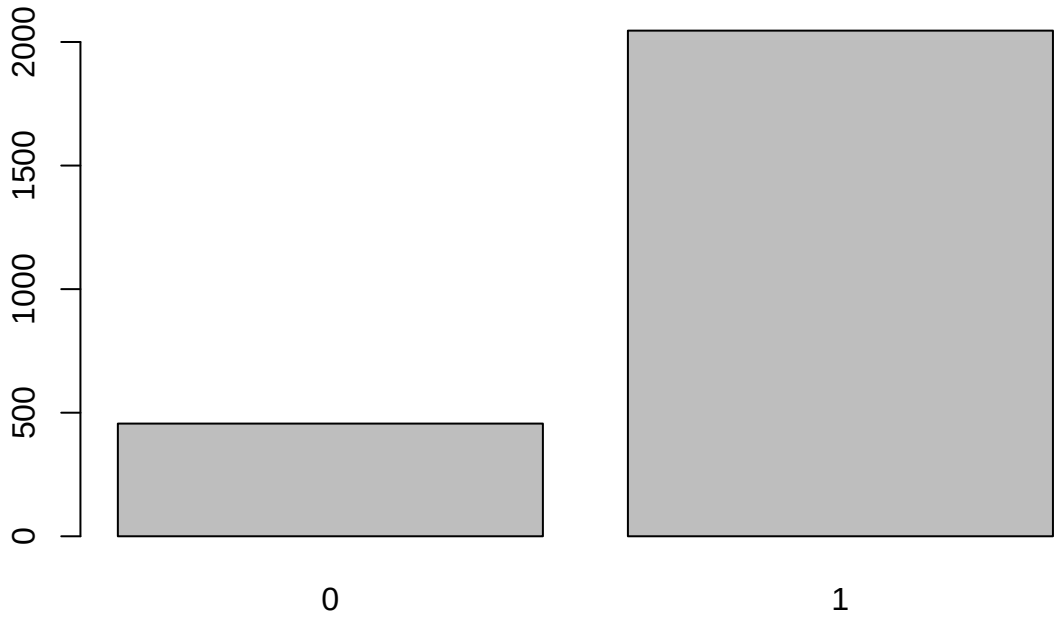
```
## # A tibble: 3 x 5
##   Id          Init Completion Type  Kinetic
##   <chr>      <dbl>      <dbl> <chr>   <dbl>
## 1 f9fc61ce85  924.       926. Turn        1
## 2 f9fc61ce85  983.       984. Turn        0
## 3 f9fc61ce85 1173.      1173. Turn        1
```

summary stats

```
##       Id                  Init           Completion          Type
##  Length:3544        Min.   : -30.67   Min.   : -29.72   Length:3544
##  Class :character   1st Qu.:  39.52   1st Qu.:  48.61   Class :character
##  Mode  :character   Median : 768.66   Median : 774.26   Mode  :character
##                     Mean   : 956.30   Mean   : 964.49
##                     3rd Qu.:1570.30   3rd Qu.:1576.70
##                     Max.   :4381.22   Max.   :4392.74
##
##     Kinetic
##  Min.   :0.0000
##  1st Qu.:1.0000
##  Median :1.0000
##  Mean   :0.8177
##  3rd Qu.:1.0000
##  Max.   :1.0000
##  NA's   :1042
```

Kinetic has 1042 NA.

Kinetic graphical summary



Most kinetic event status is "1", i.e. many events occurs instaed of beeing censored.

Kinetic numerical summary

```
##
```

```
##    0    1
## 0.18 0.82
```

82 % of events has been observed during this experiment.

events type graphical summary

events type numerical summary

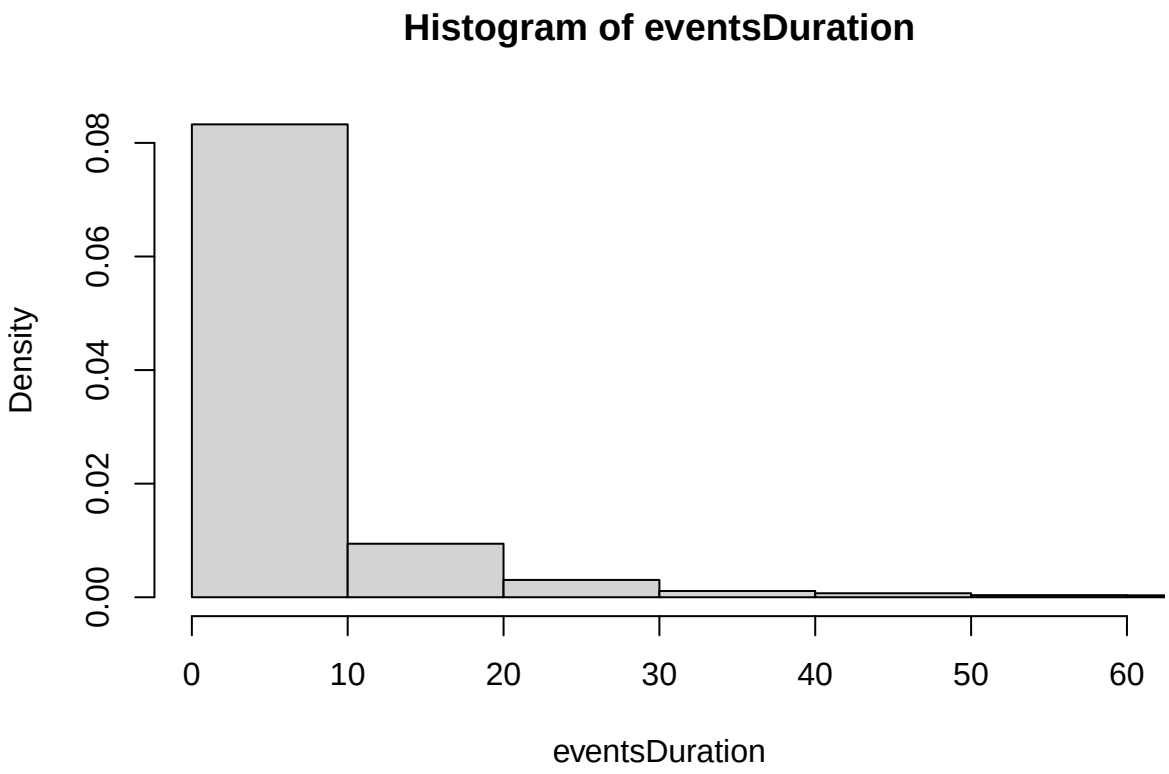
```
##
## StartHesitation            Turn         Walking
##            0.04            0.80            0.16
```

The most frequent events is Turn. (80% of the observed events).

Adding duration column : Completion - Init

```
## # A tibble: 3 x 6
##   Id          Init Completion Type  Kinetic eventsDuration
##   <chr>      <dbl>      <dbl> <chr>   <dbl>          <dbl>
## 1 003f117e14  8.61       14.8 Turn        1           6.16
## 2 009ee11563 11.4        41.2 Turn        1          29.8
## 3 009ee11563 54.7        58.8 Turn        1           4.12
```

events Duration graphical summary.



Events duration is assymetric.

EventsDuration Numerical summary

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.119   1.210   2.680   8.191   6.920 581.980
```

`EventsDuration` has median 2.7 seconds , and has range 0 .11 to 581 seconds.

subjects

data structure

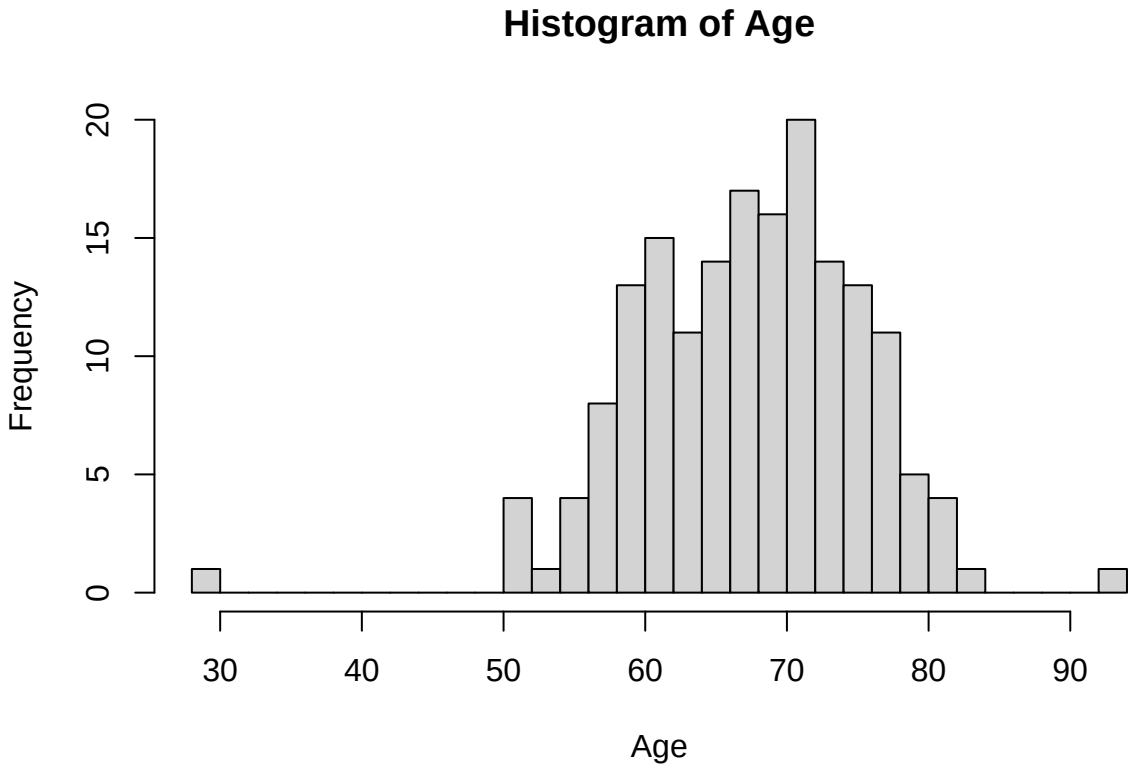
```
## # A tibble: 3 x 8
##   Subject Visit   Age Sex   YearsSinceDx UPDRSIII_On UPDRSIII_Off NFOGQ
##   <chr>   <dbl> <dbl> <chr>        <dbl>       <dbl>        <dbl> <dbl>
## 1 00f674      2    63 M               27          43           49    24
## 2 00f674      1    63 M               27          31           30    26
## 3 02bc69     NA    69 M                4          21           NA    22
```

```
## [1] 173   8
```

subjects has 173 rows and 8 columns.

```
## [1] 173   8
```

There are 173 unique subjects. Age graphical summary.

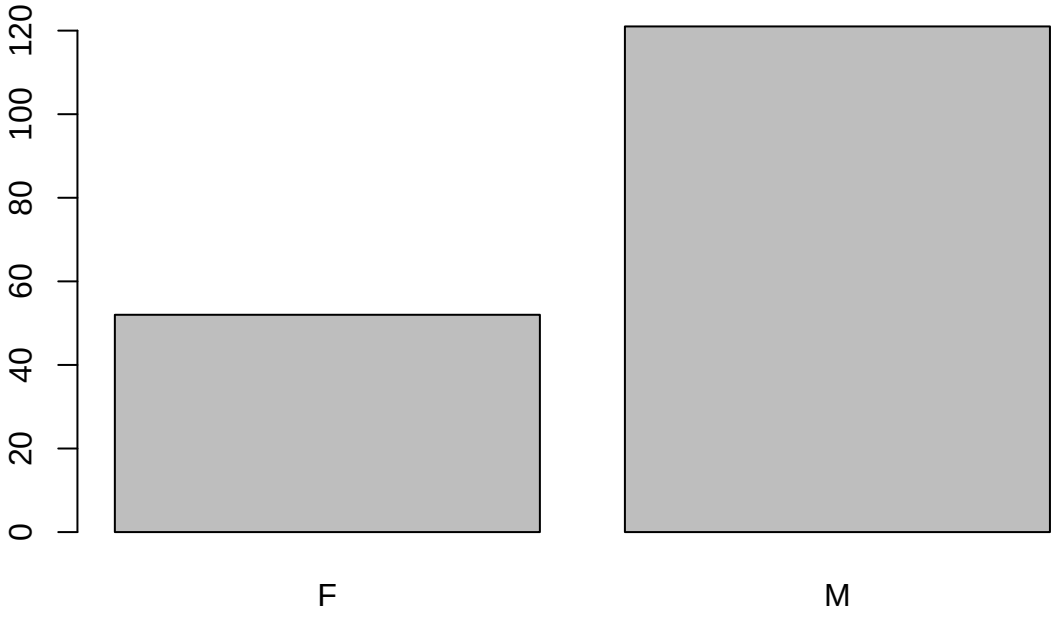


Age numerical summary.

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   28.00   62.00   68.00   67.76   73.00   94.00
```

There are 94 NA in Age. Age has median of 68 years old and presents some outliers.

sex graphical summary.



sex numerical summary

```
## Sex
##   F   M
## 0.3 0.7
```

There is almost 70% of men in this cohort.

tasks

data structure

```
## spc_tbl_ [2,817 x 4] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
##  $ Id   : chr [1:2817] "02ab235146" "02ab235146" "02ab235146" "02ab235146" ...
##  $ Begin: num [1:2817] 10 211 506 578 701 ...
##  $ End  : num [1:2817] 190 272 522 595 715 ...
##  $ Task : chr [1:2817] "Rest1" "Rest2" "4MW" "4MW-C" ...
##  - attr(*, "spec")=
##   .. cols(
##   ..   Id = col_character(),
##   ..   Begin = col_double(),
##   ..   End = col_double(),
##   ..   Task = col_character()
##   .. )
##  - attr(*, "problems")=<externalptr>
```

REPRISE