

explore_data

Dave Jongenelen

17-10-2021

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(df)
```

```
##           id           listing_url           scrape_id           last_scraped
##  Min.      : 73282   Length:1835      Min.      :2.021e+13   Length:1835
##  1st Qu.:16197502   Class :character   1st Qu.:2.021e+13   Class :character
##  Median :32759483   Mode  :character   Median :2.021e+13   Mode  :character
##  Mean   :30421509           Mean   :2.021e+13
##  3rd Qu.:45138187           3rd Qu.:2.021e+13
##  Max.   :52512196           Max.   :2.021e+13
##
##           name           description           neighborhood_overview picture_url
##  Length:1835   Length:1835   Length:1835   Length:1835
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##           host_id           host_url           host_name           host_since
##  Min.      : 14288   Length:1835   Length:1835   Length:1835
##  1st Qu.: 12886487   Class :character   Class :character   Class :character
##  Median : 28318759   Mode  :character   Mode  :character   Mode  :character
##  Mean   : 89363889
##  3rd Qu.:131225786
##  Max.   :420370006
##
##  host_location   host_about   host_response_time host_response_rate
##  Length:1835     Length:1835   Length:1835     Length:1835
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##
```

```

## host_acceptance_rate host_is_superhost host_thumbnail_url host_picture_url
## Length:1835          Length:1835          Length:1835          Length:1835
## Class :character      Class :character      Class :character      Class :character
## Mode :character        Mode :character        Mode :character        Mode :character
##
##
##
## host_neighbourhood host_listings_count host_total_listings_count
## Length:1835          Min. : 0.00          Min. : 0.00
## Class :character      1st Qu.: 1.00          1st Qu.: 1.00
## Mode :character        Median : 1.00          Median : 1.00
##                        Mean : 21.35          Mean : 21.35
##                        3rd Qu.: 7.00          3rd Qu.: 7.00
##                        Max. :252.00          Max. :252.00
##                        NA's :3              NA's :3
## host_verifications host_has_profile_pic host_identity_verified
## Length:1835          Length:1835          Length:1835
## Class :character      Class :character      Class :character
## Mode :character        Mode :character        Mode :character
##
##
##
## neighbourhood neighbourhood_cleansed neighbourhood_group_cleansed
## Length:1835          Length:1835          Length:1835
## Class :character      Class :character      Class :character
## Mode :character        Mode :character        Mode :character
##
##
##
## latitude longitude property_type room_type
## Min. :47.32 Min. :8.461 Length:1835 Length:1835
## 1st Qu.:47.37 1st Qu.:8.519 Class :character Class :character
## Median :47.38 Median :8.532 Mode :character Mode :character
## Mean :47.38 Mean :8.533
## 3rd Qu.:47.39 3rd Qu.:8.548
## Max. :47.43 Max. :8.605
##
## accommodates bathrooms bathrooms_text bedrooms
## Min. : 0.000 Mode:logical Length:1835 Min. :1.000
## 1st Qu.: 2.000 NA's:1835 Class :character 1st Qu.:1.000
## Median : 2.000 Mode :character Median :1.000
## Mean : 2.605 Mean :1.274
## 3rd Qu.: 3.000 3rd Qu.:1.000
## Max. :16.000 Max. :7.000
## NA's :283
## beds amenities price minimum_nights
## Min. : 0.000 Length:1835 Length:1835 Min. : 1.000
## 1st Qu.: 1.000 Class :character Class :character 1st Qu.: 1.000
## Median : 1.000 Mode :character Mode :character Median : 3.000
## Mean : 1.473 Mean : 7.833
## 3rd Qu.: 2.000 3rd Qu.: 5.000

```

```

## Max. :10.000 Max. :365.000
## NA's :28
## maximum_nights minimum_minimum_nights maximum_minimum_nights
## Min. : 1.0 Min. : 1.000 Min. : 1.000
## 1st Qu.: 75.0 1st Qu.: 1.000 1st Qu.: 2.000
## Median :1125.0 Median : 2.000 Median : 3.000
## Mean : 704.1 Mean : 7.042 Mean : 8.912
## 3rd Qu.:1125.0 3rd Qu.: 5.000 3rd Qu.: 7.000
## Max. :9999.0 Max. :365.000 Max. :365.000
## NA's :1 NA's :1
## minimum_maximum_nights maximum_maximum_nights minimum_nights_avg_ntm
## Min. : 1.0 Min. :1.000e+00 Min. : 1.000
## 1st Qu.: 90.0 1st Qu.:3.650e+02 1st Qu.: 1.625
## Median :1125.0 Median :1.125e+03 Median : 3.000
## Mean : 764.7 Mean :4.685e+06 Mean : 8.350
## 3rd Qu.:1125.0 3rd Qu.:1.125e+03 3rd Qu.: 6.000
## Max. :9999.0 Max. :2.147e+09 Max. :365.000
## NA's :1 NA's :1 NA's :1
## maximum_nights_avg_ntm calendar_updated has_availability availability_30
## Min. : 1 Mode:logical Length:1835 Min. : 0.000
## 1st Qu.: 365 NA's:1835 Class :character 1st Qu.: 0.000
## Median : 1125 Mode :character Median : 3.000
## Mean : 16832 Mean : 9.119
## 3rd Qu.: 1125 3rd Qu.:19.000
## Max. :14709784 Max. :30.000
## NA's :1
## availability_60 availability_90 availability_365 calendar_last_scraped
## Min. : 0.00 Min. : 0.00 Min. : 0.0 Length:1835
## 1st Qu.: 0.00 1st Qu.: 0.00 1st Qu.: 24.0 Class :character
## Median :21.00 Median :39.00 Median : 91.0 Mode :character
## Mean :23.42 Mean :39.17 Mean :148.6
## 3rd Qu.:46.00 3rd Qu.:73.00 3rd Qu.:283.0
## Max. :60.00 Max. :90.00 Max. :365.0
##
## number_of_reviews number_of_reviews_ltm number_of_reviews_l30d
## Min. : 0.0 Min. : 0.000 Min. : 0.0000
## 1st Qu.: 1.0 1st Qu.: 0.000 1st Qu.: 0.0000
## Median : 5.0 Median : 1.000 Median : 0.0000
## Mean : 24.9 Mean : 4.406 Mean : 0.6343
## 3rd Qu.: 21.0 3rd Qu.: 5.000 3rd Qu.: 1.0000
## Max. :790.0 Max. :143.000 Max. :16.0000
##
## first_review last_review review_scores_rating
## Length:1835 Length:1835 Min. :0.000
## Class :character Class :character 1st Qu.:4.600
## Mode :character Mode :character Median :4.830
## Mean :4.671
## 3rd Qu.:5.000
## Max. :5.000
## NA's :377
## review_scores_accuracy review_scores_cleanliness review_scores_checkin
## Min. :1.000 Min. :1.000 Min. :1.000
## 1st Qu.:4.720 1st Qu.:4.630 1st Qu.:4.820
## Median :4.900 Median :4.855 Median :4.960

```

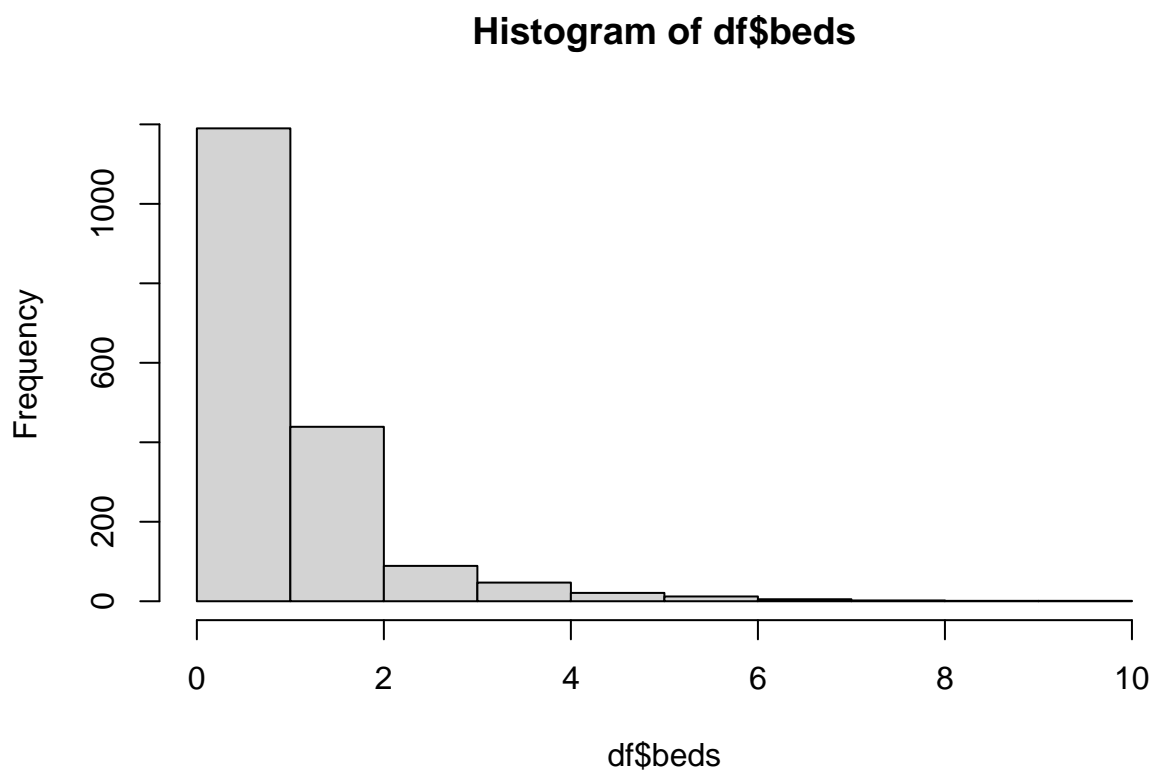
```

## Mean      :4.763          Mean      :4.727          Mean      :4.845
## 3rd Qu.:5.000          3rd Qu.:5.000          3rd Qu.:5.000
## Max.     :5.000          Max.      :5.000          Max.      :5.000
## NA's     :391           NA's      :391           NA's      :391
## review_scores_communication review_scores_location review_scores_value
## Min.      :1.00          Min.      :1.000          Min.      :1.000
## 1st Qu.:4.80          1st Qu.:4.710          1st Qu.:4.500
## Median :4.96          Median :4.890          Median :4.710
## Mean      :4.82          Mean      :4.796          Mean      :4.622
## 3rd Qu.:5.00          3rd Qu.:5.000          3rd Qu.:4.910
## Max.      :5.00          Max.      :5.000          Max.      :5.000
## NA's     :391          NA's      :391          NA's      :391
## license      instant_bookable  calculated_host_listings_count
## Mode:logical Length:1835        Min.      : 1.00
## NA's:1835     Class :character  1st Qu.: 1.00
##              Mode  :character  Median : 2.00
##              Mean      :12.21
##              3rd Qu.:10.00
##              Max.      :91.00
##
## calculated_host_listings_count_entire_homes
## Min.      : 0.00
## 1st Qu.: 0.00
## Median : 1.00
## Mean      :11.17
## 3rd Qu.: 6.00
## Max.      :91.00
##
## calculated_host_listings_count_private_rooms
## Min.      : 0.00
## 1st Qu.: 0.00
## Median : 0.00
## Mean      : 1.02
## 3rd Qu.: 1.00
## Max.      :17.00
##
## calculated_host_listings_count_shared_rooms reviews_per_month
## Min.      :0.00000          Min.      : 0.010
## 1st Qu.:0.00000          1st Qu.: 0.200
## Median :0.00000          Median : 0.560
## Mean      :0.02125          Mean      : 1.187
## 3rd Qu.:0.00000          3rd Qu.: 1.420
## Max.      :3.00000          Max.      :18.260
##              NA's      :377

```

Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.