REPORT SERIES WITH DLOOKR

Exploratory Data Analysis Report

Author:
dlookr package

Version:
0.4.0

June 5, 2023

Contents

Chapter 1

Introduction

The EDA Report provides exploratory data analysis information on objects that inherit data.frame and data.frame.

1.1 Information of Dataset

The dataset that generated the EDA Report is an 'data.frame' object. It consists of 456 observations and 9 variables.

1.2 Information of Variables

Table 1.1: Information of Variables

variables	types	missing_count	missing_percent	unique_count	unique_rate
V1	character	0	0.00	456	1.000
V2	character	0	0.00	3	0.007
V3	character	0	0.00	4	0.009
V4	character	0	0.00	3	0.007
V5	character	1	0.22	40	0.088
V6	character	287	62.94	170	0.373
V7	character	0	0.00	149	0.327
V8	character	0	0.00	123	0.270
V9	character	56	12.28	103	0.226

The target variable of the data is 'NULL', and the data type of the variable is NULL(You did not specify a target variable).

1.3 About EDA Report

EDA reports provide information and visualization results that support the EDA process. In particular, it provides a variety of information to understand the relationship between the target variable and the rest of the variables of interest.

Chapter 2

Univariate Analysis

2.1 Descriptive Statistics

edaData
9 Variables 456 Observations

V1

n	missing	distinct
456	0	456

```
lowest : 1   10  100 101 102, highest: 96  97  98  99  id
```

V2

n	missing	distinct
456	0	3

Value	Ibuprofen	Placebo	treat
Frequency	224	231	1
Proportion	0.491	0.507	0.002

V3

n	missing	distinct
456	0	4

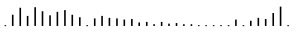
Value	African American	Other	race	White
Frequency	130	32	1	293
Proportion	0.285	0.070	0.002	0.643

V4

n	missing	distinct
456	0	3

Value	Alive	Dead	fate
Frequency	279	176	1
Proportion	0.612	0.386	0.002

V5



n	missing	distinct
455	1	39

```
lowest : 0      10     11     12     13    , highest: 6      7      8      9      apache
```

V6

n	missing	distinct
169	287	169

```
lowest : 1002.119995 1010.809998 1014.299988 1018.030029 1019.669983
highest: 994.4199829 995.2299805 997.2600098 997.7700195 o2del
```

V7

n	missing	distinct
456	0	149

```
lowest : 1        10       103      104      11       , highest: 90       91       94       96       followup
```

V8

n	missing	distinct
456	0	123

```
lowest : 33.10000102 33.9999983  34.29999881 34.49999915 34.61111281
highest: 41.0000017  41.11111111 41.38888889 41.66666667 temp0
```

V9

n	missing	distinct
400	56	102

```
lowest : 33.5000017  33.89999814 33.9999983  34.19999864 34.88889058
highest: 39.49999915 39.59999932 39.77777693 40           temp10
```

2.2 Normality Test of Numerical Variables

2.2.1 Statistics and Visualization of (Sample) Data

There is no numeric variable.

Chapter 3

Relationship Between Variables

3.1 Correlation Coefficient

3.1.1 Correlation Coefficient by Variable Combination

Number of numerical variables is less than 2.

3.1.2 Correlation Plot of Numerical Variables

Number of numerical variables is less than 2.

Chapter 4

Target based Analysis

4.1 Grouped Descriptive Statistics

4.1.1 Grouped Numerical Variables

There is no target variable.

4.1.2 Grouped Categorical Variables

There is no target variable.

4.2 Grouped Relationship Between Variables

4.2.1 Grouped Correlation Coefficient

There is no target variable.

4.2.2 Grouped Correlation Plot of Numerical Variables

There is no target variable.