Exploratory Data Analysis

Submitted by: Reet

Index:

About exploratory data analysis

Steps for Exploratory data analysis

1. Variable Identification: Numerical/Categorical
2. Summary of Numerical Variables
3. Statistical Summary of dataset
4. Graphical analysis
5. Summary or conclusion from the data

Analysing the csv files

A. a.csv

- Variable Identification: log_time, phone, status, type, product, pay_mode, marker
  Categorical: status, product, pay_mode, log_time
  Numerical: phone, type, marker

- Summary of Numerical Variables: to create an efficient algorithm, we should try to convert all categorical data ie object data type. I converted log_time into date_time.

```
log_time     object
phone       float64
status       object
type          int64
product      object
pay_mode     object
marker        int64
dtype: object
```

- Statistical summary of dataset:

```
(998822, 7)
                phone           type         marker
count   998814.000000  998822.000000  998822.000000
mean    260397.183060    1244.267067       1.623012
std     187624.002369     457.458217       2.865491
min          0.000000    1001.000000     -99.000000
25%      89754.250000    1001.000000       1.000000
50%     244574.500000    1002.000000       1.000000
75%     423747.000000    1002.000000       1.000000
max     607732.000000    2209.000000      10.000000
```

- Graphical analysis of data: