Assignment: 8.3 Final Project Step 1

Raghuwanshi, Prashant

2021-08-01

Introduction

As Credit card transaction increases in india cities, here i am going to ### perform

some EDA on the top of credit card transaction data. and have a plan to

split out the data and try to found the correlation , dependant variales, ### independent variables

, outliers and biased for the given dataset. this will help the researches

to predit the which card type is best fit for which expense type and more

Research questions

Do we have any coorelation between card type and exp_type

Are there any relationships between Cities and exp_type behavious

Do the cities Tiers types affetcs exp_type behavious

summary for each variables

find out dependant and multiple independent varibales

find out the best fit

multiple regression between dependent and multiple independent varibales

prediting the expense amount by using city, card type and exp_type predecting ### variables

Approach

How your approach addresses (fully or partially) the problem.

my approach will focus on below lsited three points

1 Getting a better understanding of data

2 Identifying various data patterns

3 Getting a better understanding of the problem statement

below steps which requried to address above points

Checking Introductory Details About Data

Statistical Insight

Data cleaning

Checking Duplicates

Data Visualization

Multi-Variate analysisVarious Plots

Here EDA approach will give the right inofrmation for dependent and multiple ### independent varibales

which is going to help me in identifying the closes fit independent variable

which help researcher to recommend the card types to the cities and gender ### types

Data

https://neostencil.com/list-of-tier-1-and-tier-2-cities-in-india

https://data.world/ash018/credit-card-transaction/workspace

file name : Credit card transactions - India - Simple.csv

second dataset: india city tier list

```
library(ggplot2)
theme_set(theme_minimal())
library(readxl)
library(dplyr)
library(QuantPsyc)
library(gridExtra)
library(car)
```

Plots and Table Needs

Correlation Matrix

Scatterplot

pairplot

Histogram

Boxplot

Questions for future steps

findout more datasets to check spending patters for credit cards

add more question related to predections

need to perfrom EDA by using R programming