Google Play Store Apps

-Abstract

The objective of this project is to make exploratory data analysis (EDA) models and prediction on the Google Play Store Apps dataset. These EDA and models will help the developers to understand the type of application people is preferred. I worked on a dataset founded through the Kaggle website. I used python libraries such as NumPy, pandas, and Matplotlib.

-Design

By Applying EDA the following questions will be answered:

- Will the price affect number of installations?
- What is the most downloaded app?
- What is the most famous category of the app?
- What are the top rated apps?
- TO WHICH INSTALLS LABEL THE MAXIMUM INSTALLS APPS BELONGS?
- WILL THE PRICE AFFECT NUMBER OF INSTALLATIONS?

-Dataset

The dataset contain over 1 million instances with 24 features include numerical and categorical types, such as rating ,developer. I have dropped unnecessary columns and that have NA values. Checked whether there is any duplicate data.

(https://www.kaggle.com/gauthamp10/google-playstore-apps)

-Algorithms

Feature Engineering:

- Drop some non-useful columns such as currency.
- Handle missing values in 'Minimum Installs' and ‘App Name' ,etc.
- Factorizing ContentRating ,Category, and Installs columns into integers values.

Model:

Regression model to predict the rate. By using Linear Regression, Decision Tree Regressor and Random Forest Regressor. The Decision Tree was the best.

-Tools

I will conduct the experiment by using:

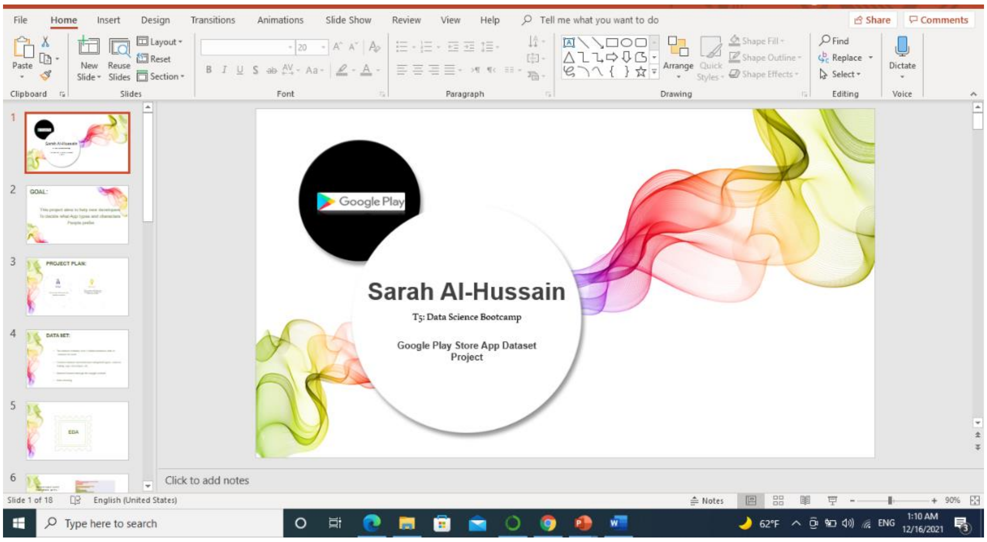
-Environment: Jupyter Notebook.

-Programming Language: Python.

-Libraries:

- NumPy
- Pandas
- Matplotlib and Seaborn to visualize the data.
- Sklearn to build the model.

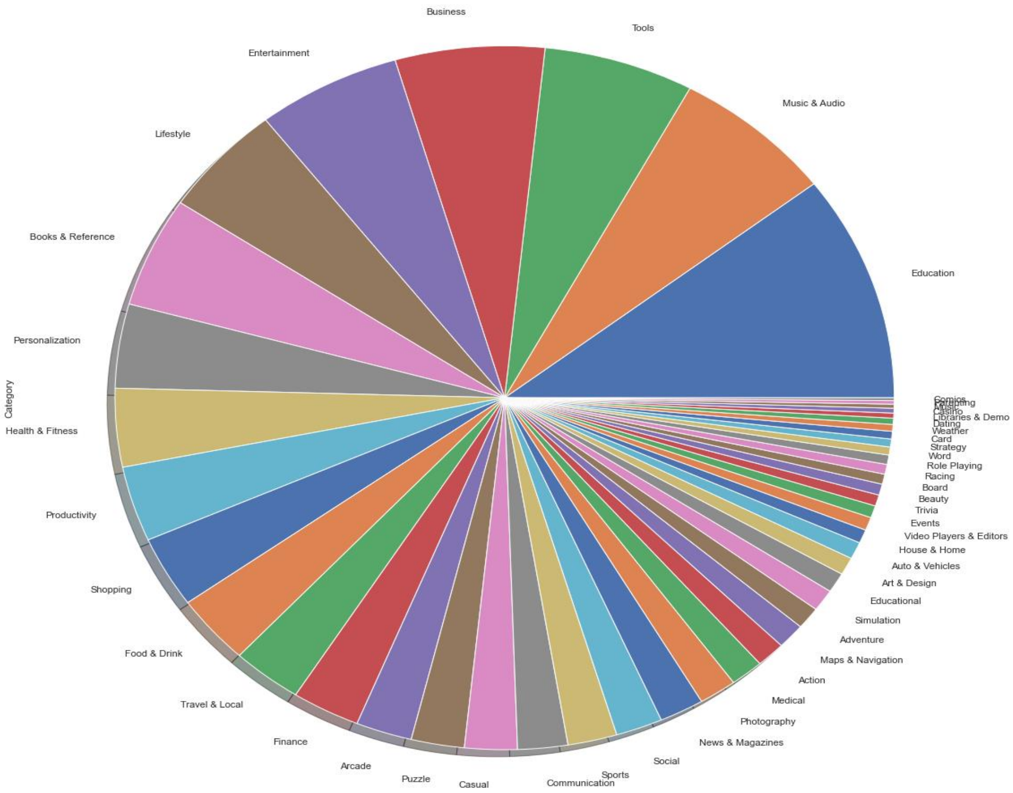
-Communication

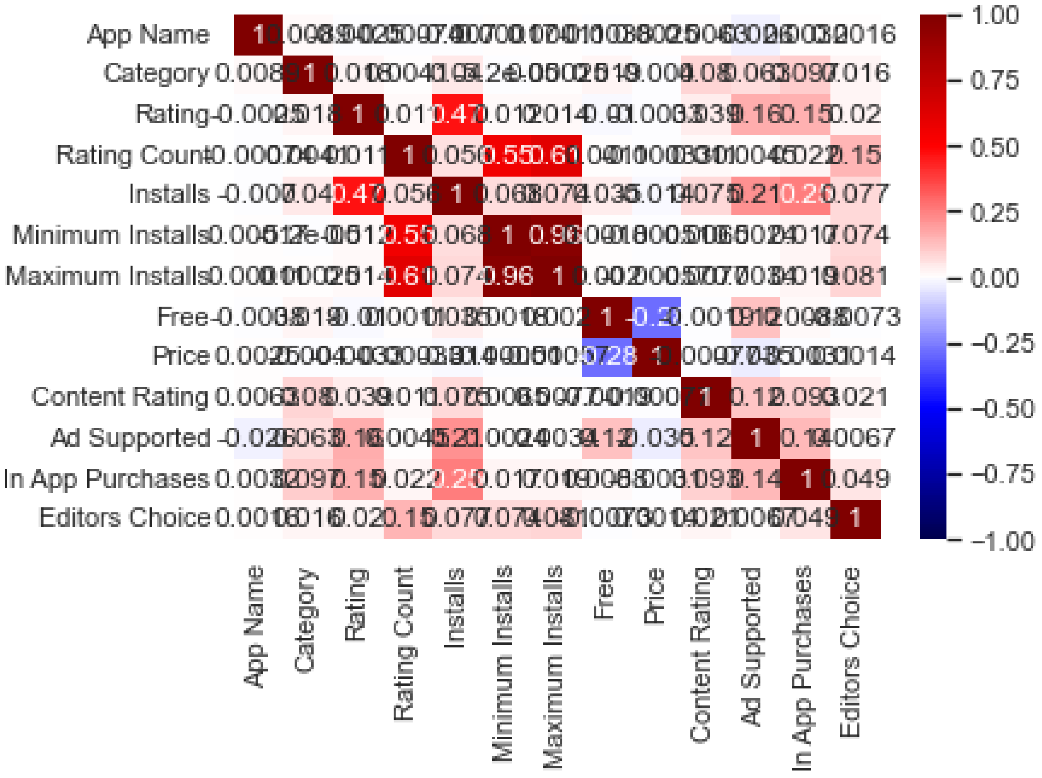


Sarah Al-Hussain

T5: Data Science Bootcamp

Google Play Store App Dataset Project


Business
Tools
Entertainment
Music & Audio
Lifestyle
Books & Reference
Education
Personalization
Category
Health & Fitness
Comics
Parenting
Music
Casino
Libraries & Demo
Dating
Weather
Card
Strategy
Word
Role Playing
Racing
Board
Beauty
Trivia
Events
Video Players & Editors
House & Home
Auto & Vehicles
Art & Design
Educational
Simulation
Adventure
Maps & Navigation
Action
Medical
Photography
News & Magazines
Social
Sports
Communication
Casual
Puzzle
Arcade
Finance
Travel & Local
Food & Drink
Shopping
Productivity


App Name
Category
Rating
Rating Count
Installs
Minimum Installs
Maximum Installs
Free
Price
Content Rating
Ad Supported
In App Purchases
Editors Choice
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00