“Mobile Games Success And Failure: Mining The Hidden Factors”

NAME: Aditya Mani Tripathi

Registration No.: 12017318

Roll No.: RK20KT10

"On what Project I'm working on and why I choose it"

I'm working on **"Mobile games success and failure: Mining the hidden factors"** dataset, And the reason behind choosing this project is, it's almost 2022, and our lives are greatly influenced by the pandemic. In most of the cases the effect has been negative, but for some industries specially the ones who are easily and virtually accessible by people it was boon, One of these industry is **Mobile Gaming Industry** in recent years we saw a lot of growth and revenue in this industry during this pandemic the role of the mobile games was to help and relief people, Witnessing the growth of this industry during these years also got me interested in it .

"Brief Description of Domain"

My dataset belongs to Video Game Industry, the **video game industry** is the industry involved in the development, marketing, and monetization of games. A mobile game is a video game that is typically played on a mobile phone. (The video game industry has grown from focused markets to mainstream in recent years. As of 2018, July, video games generated sales of US$134.9 billion annually worldwide. In the US, it earned about $9.5 billion in 2007, $11.7 billion in 2008, and US$25.1 billion in 2010, according to the ESA annual report. Research from Ampere Analysis indicated three points: the sector has consistently grown since at least 2015 and expanded 26% from 2019 to 2021, to a record $191 billion; the global games and services market is forecast to shrink 1.2% annually to $188 billion in 2022; the industry is not recession-proof. (**Source: Wikipedia**))

"Dataset Column description"

This Dataset contains 18 col and 17008 rows.

Column Name	Data Type
URL	String
ID	Numerical
Name	String
Subtitle	String
Icon URL	String
Average User Rating	Numerical
User Rating Count	Numerical
Price	Numerical
In-app Purchases	Numerical
Description	String
Developer	String
Age Rating	Categorical
Language	Categorical
Size	Numerical

Primary Genre	Categorical
Genres	Categorical
Original Release Date	Date
Current Version Release Date	Date

"What will be my approach, and what are the insights I want to retrieve from this data"

As we can see the dataset contains more the 17 thousand rows and each row contain 18 columns, in which some columns can be identify as unnecessary columns like URL, Icon URL, ID and subtitle because these columns has no relation in success or failure of game so first I will drop unnecessary columns second in Primary genre column more then 95% rows contain same value as Game and it also has no relation with failure and success of game, so in short first I will clean my dataset to make it more clear and precise for analysis, after cleaning the dataset I will target the columns which can give me the most appropriate data for my analysis in this step I will analyze the data for example relation rating of the game and the prize of the game, also I can analyze with size and rating, what is the relation between rating, prize and size of the game, from this analysis I can get the success and failure relation with size, price, genre and average rating, I will visualize all the relation so we can get some interesting graphs also.