**Analyze Mobile App Data on various platforms and generate recommendations based on the analysis and various metrics available in the datasets using Python.**

**Task - 1**

* **Downloading Data, Reading and Saving CSV Data into Dataframes, Summarizing Data using Python**

Steps performed in downloading and storing data into DataFrame

**Downloading the dataset**:

First we’ll have to figure out the type of data needed and format.

Here I downloaded the AppStore and Play Store data (CSV format) and saved inside my directory and for further analysis we would have to download certain packages such as NumPy, Pandas and import them inside the code directory.

#pip install pandas

#pip install NumPy

**To read the dataset in python:**

To read the dataset in python first have to load the pandas packages and by using read\_csv method we can load our dataset in python working interface.

* Importing the necessary packages

Import pandas as pd

* Storing the dataset inside app\_data variable

app\_store = pd.read\_csv('C:\\Users\\Mithunsomu\\Downloads\\AppleStore.csv',sep=',')

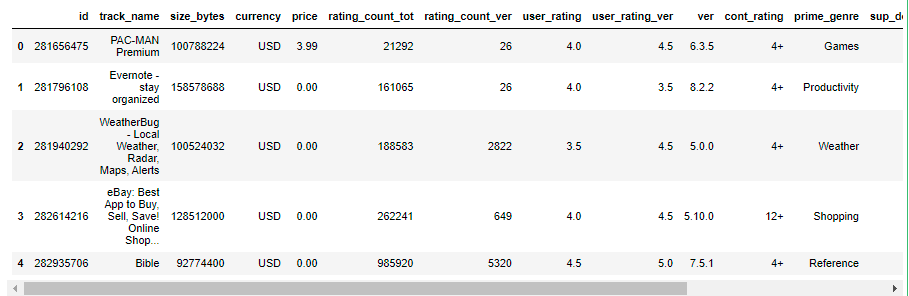
play\_store = pd.read\_csv('C:\\Users\\Mithunsomu\\Downloads\\googleplaytore.csv',sep=',')

**Summarizing the datasets and print out some rows from each to get an idea of the data:**

To read and understand data let’s print the first 5 rows to get a clear picture about the data.

* Printing the starting 5 header rows of the dataset

Print(app\_store.head())



**Task - 2**

* **Null Value Treatment, Data Validation and Processing**

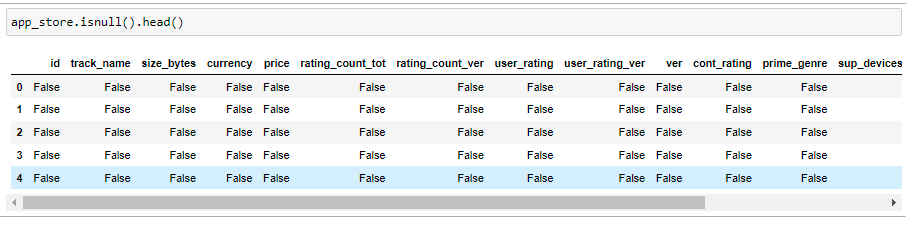
**Examine if the Datasets have NULL values and proceed with NULL value treatment**

To know if there are any NULL values in the dataset we use isnull() method

Let’s try it with the app\_store data:-

* checking if the dataset is having any NULL values

App\_store.isnull()

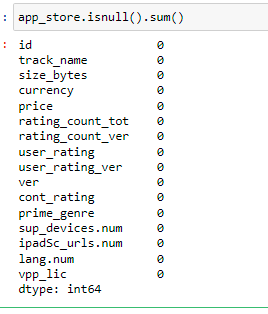


Here we’ll get a Boolean values (True or False), where True represents the NULL value and False represents NotNULL values.

Here we cannot find True values so to be clear I’ll use .sum() method to summarize every NULL vales if there are any in the dataset.

* Summarizing the dataset to find out NULL Values

App\_store.isnull ().sum()



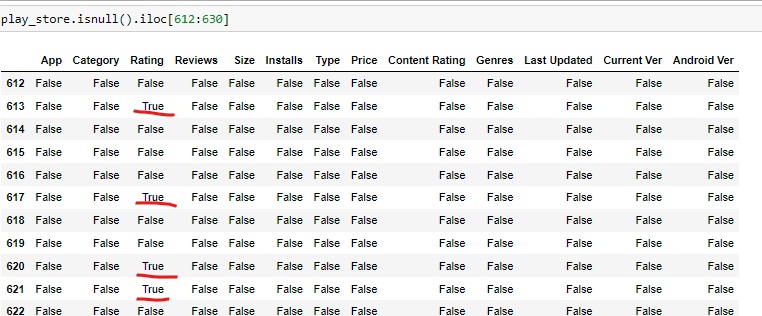
Here I have summarized all the NULL vales and it is showing 0 for every column from that we can be sure that there are no NULL values in this dataset.

Now let’s check the playstore data:-

* checking if the dataset is having any NULL values

play\_store.isnull()

**Outcome:-**



Here we can find few True values that represents that the areas showing True are the areas having NULL values.

To get a clear idea of NULL values let’s summarize the data

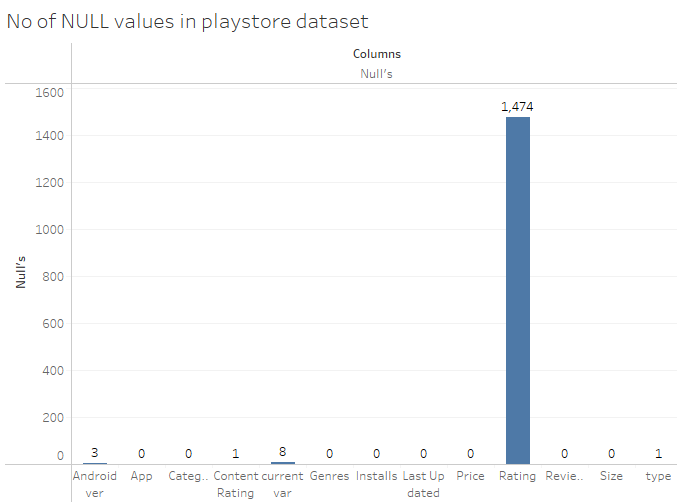
* Summarizing the dataset to find out NULL Values

Play\_store.isnull().sum()



**Visual representation of no of null values**

**Outcome:-**



Here we got a clear picture of how many NULL values are there in the respective columns

**Types of handling Missing Values (NULL):-**

* Removing the data (rows)
* Replacing the missing values with mean, median or mode
* Assigning an unique category

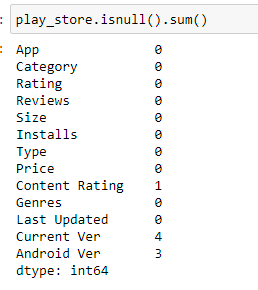
We can apply above types according to the type of analysis we need.

Here let’s remove the null data

* Removing the rows that has NULL value inside

play\_store=play\_store[~play\_store.Rating.isnull()]

Outcome:-

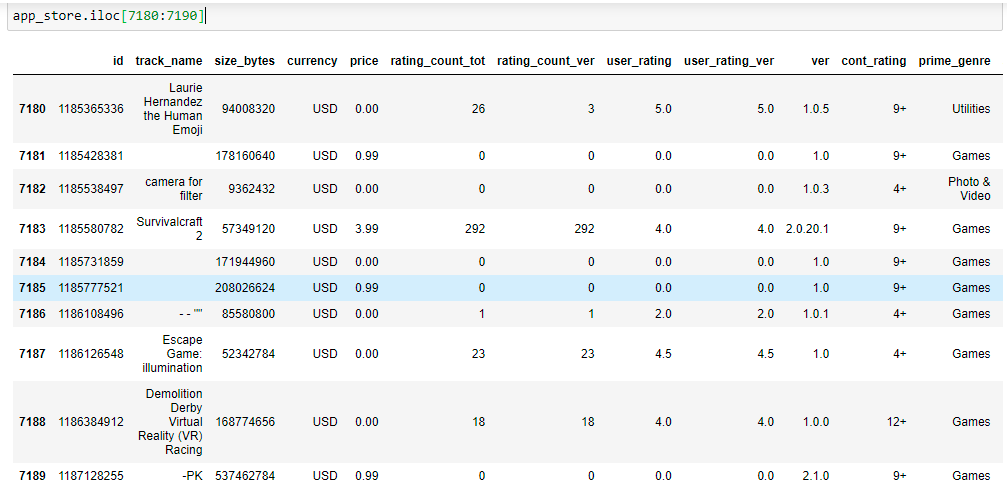
 

**Removing Non English Apps:**

For removing Non English data we use encoding and decoding of ASCII using lambda function.

app\_store['track\_name'] = app\_store['track\_name'].apply(lambda row: row.encode('ascii',errors='ignore').decode())





----------------------------------------------------------Thank You-------------------------------------------------------------------