MATIKS - DATA ANALYST TASK

Project Objective – To design an actionable insights from raw data, visualize it clearly, and communicate your findings effectively using Visualization tool.

Tools used:- Jupyter Notebook, Python(Pandas, Matplotlib, Seaborn), Microsoft Power BI.

Explore Data:-

Read the data in Pandas Dataframe to have an overview of entire database

```
In [1]: import pandas as pd
        import matplotlib.pyplot as plt
        import seaborn as sns
        %matplotlib inline
```

```
In [2]: df = pd.read_csv('Matiks - Data Analyst Data - Sheet1.csv')
        df.head()
```

From **df.shape** and **df.info()** we get that dataset have 10000 rows and 20 columns and data types of column entries and by **df.duplicated().sum()** and **df.isnull().sum(),** we get that dataset have no duplicate and null values.

By which we can say that data is clean ready to use for further operation.

Some insights being taken out at Jupyter Notebook itself using Matlplotlib and Seaborn.

Future scope:-

We can check datatypes of column entries and can change in python for further use.

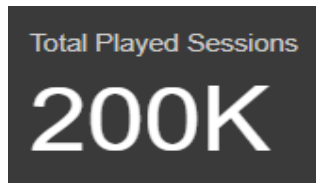From this we can save clean file in .csv, .xlxs, .sql format for any further operation.

We can integrate MYSQL Workbench using connectors like pymysql and sqlalchemy for tabular view.
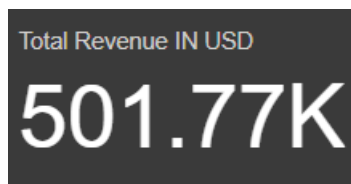
Visualization of Data:-

Microsoft Power BI is being used for interactive Visualization by importing clean file in it.

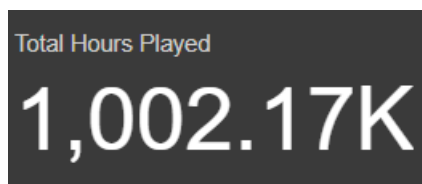Insights taken out from visualizations are where all parameters varies as other parameters being selected:
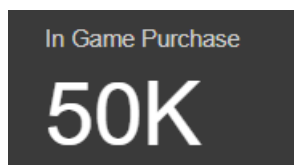
1. Total Played Sessions

Total Played Sessions
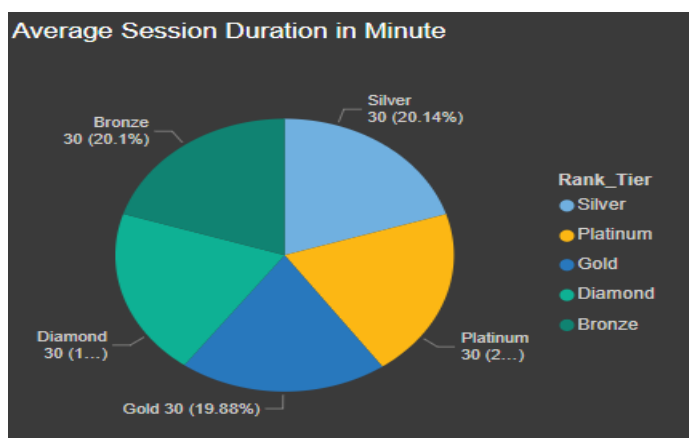**200K**

2. Total Revenue is USD

Total Revenue IN USD
501.77K

3. Total hours played

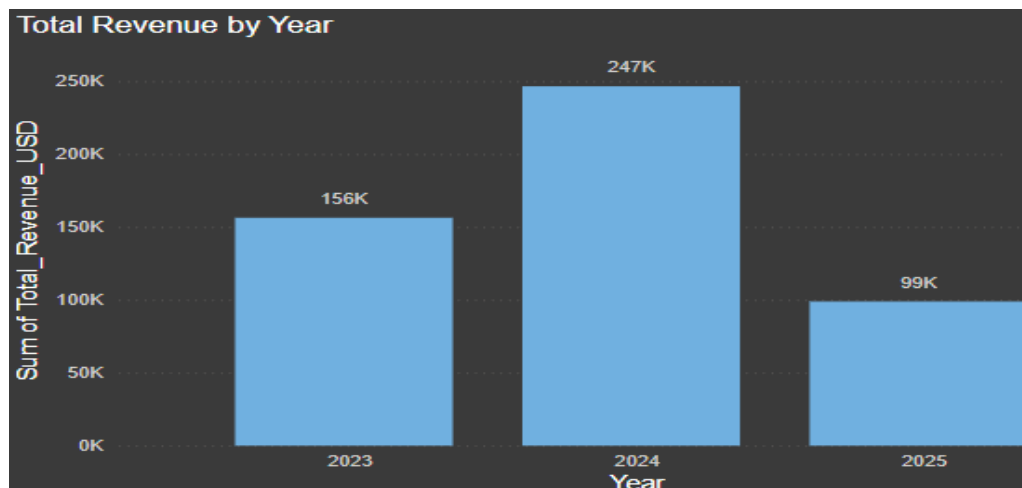**Total Hours Played**
1,002.17K

4. In game Purchase

In Game Purchase
50K

5. Average Session Duration in Minute

6. Total Revenue by Year

Parameters made are that can varies above parameters :

Link for the Dashboard:- https://drive.google.com/file/d/1z6Ig9GYt3vn4bb7AUac8fh-iLWemicLq/view?usp=sharing

Due to short duration of time, I made the visualization on basic level, but I am going to make it more interactive and also going to visualize it in MYSQL Workbench and will add it to my GitHub Repository.

Updated: -

Created a clean dashboard after working on it after submitting work on MYSQL

Link to updated dashboard: -
https://drive.google.com/file/d/1OUB03Zgf2QSP_zzgdKMe13zmvhtnuJv9/view?usp=sharing

MATIKS - DATA ANALYST TASK

(Updated)

Updates done:-

1. Connected MYSQL Server with Jupyter notebook and get varies information through SQL Queries.

```python
In [1]: import pandas as pd
        import matplotlib.pyplot as plt
        import seaborn as sns
        %matplotlib inline

        import pymysql
        from sqlalchemy import create_engine
```

Pymysql works as a SQL Driver that makes connections between Python and Workbench.

Sqlalchemy is a SQL toolkit that can use pymysql in backend.

2. Connection Query of Python to MYSQL Workbench

```python
In [23]: engine_mysql = create_engine("mysql+pymysql://root:akashwadibhasme@localhost:3306/matiks_db")

         try:
             engine_mysql
             print('Success')
         except:
             print('Fail')

         Success
```

```python
In [24]: df.to_sql(name='matiks', con=engine_mysql, if_exists='append', index=False)
Out[24]: 10000
```

Matiks_db is the database name for MYSQL with table name as **matik**.

3. Varies insights being taken out from SQL.

Insights: -

- Total Count of **Mobile users** with **Preferred Game Mode** as **Co-op**, **Solo** and **Multiplayer**
  1. Total **Mobile users** are – 6602 with username miguel84 from Finland has highest Achievement score of 9995.
  2. Total **Mobile user**s with **Co-op** as Preferred Game Mode – 2170 with username miguel84 from Finland has highest Achievement score of 9995 in Co-op mode.
  3. Total **Mobile users** with **Solo** as Preferred Game Mode – 2184 with username ndaniel from Belize has highest Achievement score of 9977.
  4. Total **Mobile users** with **Multiplayer** as Preferred Game Mode – 2248 with username ywerner from Rwanda has highest Achievement score of 9990.

- Total Count of **PC users** with **Preferred Game Mode** as **Co-op**, **Solo** and **Multiplayer**
  1. Total **PC users** are – 6608 with username cindy64 from Vietnam has highest Achievement score of 9997.
  2. Total **PC users** with **Co-op** as Preferred Game Mode – 2324 with username cindy64 from Vietnam has highest Achievement score of 9997 in Co-op mode.
  3. Total **PC users** with **Solo** as Preferred Game Mode – 2114 with username emma83 from Cameroon has highest Achievement score of 9988.
  4. Total **PC users** with **Multiplayer** as Preferred Game Mode – 2170 with username amysmith from Australia has highest Achievement score of 9965.

- Total Count of **Console users** with **Preferred Game Mode** as **Co-op, Solo** and **Multiplayer**
    1. Total **Console users** are – 6790 with username audrey66 from Maldives has highest Achievement score of 9997.
    2. Total **Console users** with **Co-op** as Preferred Game Mode – 2208 with username hstokes from Salvador has highest Achievement score of 9985 in Co-op mode.
    3. Total **Console users** with **Solo** as Preferred Game Mode – 2238 with username douglasstephanie from Niue has highest Achievement score of 9984.
    4. Total **Console users** with **Multiplayer** as Preferred Game Mode – 2344 with username audrey66 from Maldives has highest Achievement score of 9997.

- Count of different Referral Sources
  Total 4 Referral Sources as: -
    1. **Social media** with **Male** as 1616, **Female** as 1728 and **others** 1566.
    2. **Ads** with **Male** as 1666, **Female** as 1610 and **others** 1704.
    3. **Organic** with **Male** as 1744, **Female** as 1640 and **others** 1758.
    4. **Friend** with **Male** as 1706, **Female** as 1678 and **others** 1584.
- Count of different Rank Tiers
  Total 5 Tier Ranks as: -
    1. **Bronze Tier** with **Male** as 1318, **Female** as 1332 and **others** 1386.
    2. **Platinum Tier** with **Male** as 1302, **Female** as 1306 and **others** 1260.
    3. **Diamond Tier** with **Male** as 1416, **Female** as 1382 and **others** 1356.
    4. **Gold Tier** with **Male** as 1376, **Female** as 1344 and **others** 1294.
    5. **Silver Tier** with **Male** as 1320, **Female** as 1292 and **others** 1316.

Conclusion taken out as: -

1. Many Users prefer Console with Multiplayer Mode.
2. Most user are here with Organic Referral Source with Male being most in number.
3. Most player are in Diamon Tier with Male being most in number.

This task is uploaded to my GitHub for further reference

I know as per instructions, I have to create a dashboard, but for in-depth analysis, MYSQL give freehand and most accurate outcome. This is the extra part I have done along with Dashboard Visualization.

GitHub: - https://github.com/akash95189/MATIKS---DATA-ANALYST-TASK.git

Made By: -

Akash Wadibhasme

📞 :- 9518925655

📧 : - akashwadibhasme1005@gmail.com

🔗 Portfolio: - https://akash-wadibhasme-portfolio.my.canva.site/

LinkedIn: - http://www.linkedin.com/in/akash-wadibhasme

Resume: - https://drive.google.com/file/d/1p_PGzkMuVs3e7S9tpYbjkdGqF8SMYJ9I/view?usp=sharing