Data Visualization Techniques for Brawl Stars Dataset Analysis: Data Exploration and Visualization

Lorena Danae Perez Lopez
Data Engineering
Universidad Politécnica de Yucatán
Km. 4.5. Carretera Mérida — Tetiz
Tablaje Catastral 4448. CP 97357
Ucú, Yucatán. México
Email: 2309179@upy.edu.mx

Julio Cesar De Aquino Castellanos
Data Engineering
Universidad Politécnica de Yucatán
Km. 4.5. Carretera Mérida — Tetiz
Tablaje Catastral 4448. CP 97357
Ucú, Yucatán. México
Email: 2309066@upy.edu.mx

Lester Stefhan Estrada Lopez
Universidad Politécnica de Yucatán
Km. 4.5. Carretera Mérida — Tetiz
Tablaje Catastral 4448. CP 97357
Ucú, Yucatán. México
Email: lester.estrada@upy.edu.mx

Abstract

This project focuses on analyzing and visualizing data from the mobile game Brawl Stars, specifically targeting the top 200 players and clubs worldwide and by country. Using datasets containing statistics like trophies, ranks, and character roles, the project aimed to explore the competitive dynamics of the game through data visualization and exploration techniques. The analysis employed Python, Jupyter notebooks, and libraries such as Pandas, Seaborn, and Matplotlib to preprocess and visualize the data. The preprocessing stage involved cleaning the datasets by addressing missing values, standardizing categorical data, and normalizing numerical variables to ensure consistency across the datasets. Exploratory Data Analysis (EDA) was then conducted to uncover trends and patterns in player performance, including trophy distributions and rank correlations. Various visualization techniques, such as histograms, scatter plots, and heatmaps, were applied to reveal insights about player behavior and club performance. The project also emphasized the importance of best data science practices by documenting the work in a GitHub repository, enabling version control using Git. Data visualizations provided clear insights into the relationships between variables, helping to understand factors influencing player and club rankings. This project demonstrated the power of data visualization in uncovering hidden trends and correlations, facilitating a deeper understanding of Brawl Stars' competitive landscape. The code, visualizations, and findings are accessible through the repository, ensuring reproducibility and transparency.

Index Terms

Data Preprocessing, Data Cleaning, Data Normalization, Data Visualization, Histogram, Box Plot, Correlation Matrix, Scatter, Count Plot, Line Chart, Bar Plot, Pair Plot, Pandas, Matplotlib, Seaborn, Sklearn.

Data Visualization Techniques for Brawl Stars Dataset Analysis: Data Exploration and Visualization

I. Introduction

In this project, the students explored some datasets that contains some statistics of the ranked enviroment of a mobile videogame called Brawl Stars. The main objective of this project was to analyze the competitive environment of Brawl Stars, from the representation of its statistics in graphs and data visualization tools, all this developed with Python, Jupyter and libraries such as Pandas, Science Kit Learn, Seaborn and Matplotlib. In addition, as a challenge of the members, they took the time to document the project with good MLOps practices and using a Github repository, so they also implemented the use of git from the terminal as a version control. With this in mind, it is important to mention that as the dataset is a different one from the previous project, the preprocessing and data cleaning stage could not be skipped.

II. Methodology

This project analyzed Brawl Stars gameplay data for the top 200 players and clubs globally and by country, focusing on player statistics like trophies, ranks, and character roles. The dataset, curated from public sources, was first cleaned by addressing missing values, standardizing categorical data, and normalizing numerical variables to ensure consistency. Exploratory Data Analysis (EDA) was then conducted to explore the distribution of trophies, rank patterns, and player preferences by character role, providing insights into competitive dynamics and player behavior.

Data visualization techniques, including distribution plots, heatmaps, and box plots, were employed to reveal trends in trophy accumulation and ranking. Visualizations were generated using Python libraries, such as Matplotlib and Seaborn, which allowed clear representation of these insights. The entire analysis was documented in Jupyter notebooks and made accessible through a GitHub repository to ensure transparency and reproducibility.

III. Data Preprocessing

The preprocessing phase was essential to ensure the consistency and quality of the data used for analysis in this Brawl Stars dataset project. For each dataset, specific cleaning and transformation steps were applied to standardize variables and handle missing values.

For the Jessie Top 200 Players dataset, we addressed missing values by replacing undefined club information with the label ”No-club-info.” Additionally, the extraneous ”Unnamed: 0” column was removed to streamline the data structure, and

Unnamed: 0	tag	name	nameColor	icon	trophies	rank	club
0	#VL0GPPJV	Physic Gear 5	0xfff9c908	28000014	1401	1	Eclipse
1	#2Y92C2Q8V	NOMAD|かーず。❄❄	0xfff9c908	28000119	1399	2	NaN
2	#VLQPVPY	Naz❤️Hyra	0xffff8afb	28000303	1399	3	Fenernazçe
3	#JQVQYVY	Code: VTzim	0xffff8afb	28000339	1309	4	Eclipse
4	#RRVVPG2Y	Brosta卡拉	0xff1ba5f5	28000206	1307	5	vMoove
5	#89UPQJJVV	+BraboX	0xffa8e132	28000236	1302	6	Ghost Team
6	#VYGVV08Q	domates	0xffffffff	28000261	1302	7	NaN
7	#2Y2Y99GQR	VizeZ	0xffffffff	28000230	1302	8	Ex0tic-E-Sport
8	#8GV2CP2J8	ZORLU	0xffcb5aff	28000003	1301	9	<c0> L O N E L Y★</c>
9	#2G90YQ9CQ	ʸᴹᴵᴿᴬᴺᴰᴵᴺᴴᴬᵛᵛ多	0xffa8e132	28000174	1301	10	Wolf gang

Fig. 1. Top 200 Jessie Players Raw Table (Anomalies highlighted in yellow)

	tag	name	nameColor	icon	trophies	rank	club	normalized_trophies	normalized_rank
0	#VL0GPPJV	Physic Gear 5	0xfff9c908	28000014	1401	1	Eclipse	1.000000	0.000000
1	#2Y92C2Q8V	NOMAD|かーず。❄❄	0xfff9c908	28000119	1399	2	no-team-info	0.993289	0.005025
2	#VLQPVPY	Naz❤️Hyra	0xffff8afb	28000303	1399	3	Fenernazçe	0.993289	0.010050
3	#JQVQYVY	Code: VTzim	0xffff8afb	28000339	1309	4	Eclipse	0.691275	0.015075
4	#RRVVPG2Y	Brosta卡拉	0xff1ba5f5	28000206	1307	5	vMoove	0.684564	0.020101
5	#89UPQJJVV	+BraboX	0xffa8e132	28000236	1302	6	Ghost Team	0.667785	0.025126
6	#VYGVV08Q	domates	0xffffffff	28000261	1302	7	no-team-info	0.667785	0.030151
7	#2Y2Y99GQR	VizeZ	0xffffffff	28000230	1302	8	Ex0tic-E-Sport	0.667785	0.035176
8	#8GV2CP2J8	ZORLU	0xffcb5aff	28000003	1301	9	<c0> L O N E L Y★</c>	0.664430	0.040201
9	#2G90YQ9CQ	ʸᴹᴵᴿᴬᴺᴰᴵᴺᴴᴬᵛᵛ多	0xffa8e132	28000174	1301	10	Wolf gang	0.664430	0.045226

Fig. 2. Top 200 Jessie Players Preprocessed

the ”trophies” and ”rank” columns were scaled using min-max normalization to keep values within a consistent range.

In the Global Clubs Ranking dataset, missing descriptions were filled with the placeholder ”no-description,” and entries without team members were marked as ”no-team-member” to prevent gaps in analyses involving club composition. The ”Unnamed: 0” column was renamed to ”rank” for clarity, while the ”trophies” and ”trophiesrequired” columns underwent min-max scaling to normalize the dataset.

```
Missing values per column:
 Unnamed: 0              0
tag                      0
name                     0
description              1
type                     0
                        ..
member_29_trophies       9
member_30_tag           39
member_30_name          39
member_30_role          39
member_30_trophies      39
Length: 127, dtype: int64
Total missing values in the dataset: 201
```

Fig. 3. Missing Values in Top 200 Global Clubs

Finally, the Top 200 Mexican Clubs dataset required minimal adjustments due to its completeness, with no missing or

duplicate values detected. We replaced the ”Unnamed: 0” column with ”rank” to match the other datasets and applied min-max normalization to the ”trophies” and ”trophiesrequired” columns to ensure uniformity across datasets.

	Unnamed: 0	tag	name	badgeId	trophies	rank	memberCount
0	0	#2G0UCPY80	Бурый Медведь	8000021	4026880	1	30
1	1	#282RJGGGY	WOLFMX	8000016	1984538	2	30
2	2	#2CLCYL202	Moonlight	8000052	1833116	3	30
3	3	#2GU9902G0	家DarkMoon|	8000055	1775840	4	30
4	4	#29JRP22G9	HeadShoterZ	8000023	1773993	5	30
5	5	#C09LG8YL	minuta:3	8000035	1752235	6	30
6	6	#2RCQ9CGQQ	STMN Esports	8000015	1697266	7	29
7	7	#2GURLVQ29	AT|GALAXY XI	8000010	1696536	8	30
8	8	#QRVJP92R	Brawl Con Limón	8000039	1670500	9	30
9	9	#Q8J09VPQ	AxE Gaming	8000027	1670097	10	30

Fig. 4. Top 200 Mexican Clubs Raw Table (Anomalies highlighted in yellow)

	Rank	tag	name	badgeId	trophies	memberCount	trophies_normalized	member_count_normalized
1	1	#282RJGGGY	WOLFMX	8000016	1984538	30	1.000000	1.0
2	2	#2CLCYL202	Moonlight	8000052	1833116	30	0.811088	1.0
3	3	#2GU9902G0	家DarkMoon|	8000055	1775840	30	0.739631	1.0
4	4	#29JRP22G9	HeadShoterZ	8000023	1773993	30	0.737327	1.0
5	5	#C09LG8YL	minuta:3	8000035	1752235	30	0.710182	1.0
6	6	#2RCQ9CGQQ	STMN Esports	8000015	1697266	29	0.641603	0.8
7	7	#2GURLVQ29	AT|GALAXY XI	8000010	1696536	30	0.640692	1.0
8	8	#QRVJP92R	Brawl Con Limón	8000039	1670500	30	0.608210	1.0
9	9	#Q8J09VPQ	AxE Gaming	8000027	1670097	30	0.607707	1.0
10	10	#2RGV8QCJR	LEGENDARYS	8000029	1668577	29	0.605811	0.8

Fig. 5. Top 200 Mexican Clubs Preprocessed Table

This preprocessing stage facilitated subsequent analysis by creating a clean, consistent dataset that allowed for accurate visualizations and data exploration. All the codes implemented could be seen on Appendix A.

IV. Justification of Visualization

In this section, we´re gonna justify the uses of the graphs for the visualization of our datasets. In this case, since we have three datasets we´re gonna be using at least three graphs for each one.

A. Jessie Players Ranking Dataset

1) Histogram of Trophies: A histogram is used to show the distribution of trophies among Jessie players. It helps to identify how evenly or unevenly the trophies are distributed.

2) Top Clubs by Player Count: This bar chart visualizes the clubs with the most Jessie players. It provides insights into the popularity and reach of different clubs within the Jessie player base.

3) Normalized Trophies vs. Normalized Rank: A scatter plot helps visualize the relationship between normalized trophies and rank, showing if higher trophies correlate with higher ranks.

4) Trophies by Player's: A box plot is used to display the distribution of trophies across different categories (here, the nameColor). It's useful for identifying the central tendency, spread, and any potential outliers.

B. Clubs Ranking Dataset

1) Club Trophy Distribution: A histogram or box plot is used to display how club trophies are distributed. It helps to understand how successful clubs are and the variation in their trophy counts.

2) Required Trophies vs. Total Trophies: This scatter plot compares the required trophies to join a club with the total trophies of that club. It helps to understand if clubs with higher required trophies also perform better.

3) Roles Distribution: This bar chart displays the distribution of roles among the top members of the clubs, providing insights into the typical composition of successful clubs.

4) Normalized Trophies vs. Rank: This line plot shows how normalized trophies correlate with rank across clubs. It helps visualize the relationship between performance (normalized trophies) and ranking.

5) Required Trophies vs. Total Trophies for Clubs: This heatmap helps visualize the correlation between two continuous variables, in this case, required trophies and total trophies. It can show patterns or correlations in a color-coded format.

C. Mexico Clubs Ranking Dataset

1) Trophy and Member Count Distributions: Two histograms display the distribution of trophies and member counts among Mexican clubs. This helps understand the size and performance variability of clubs in the region.

2) Normalized Trophies vs. Member Count: A scatter plot shows whether clubs with more members tend to have higher normalized trophies, indicating the relationship between size and success.

3) Ranking Analysis: A scatter plot shows the relationship between club rank and normalized trophies in Mexican clubs, helping analyze how performance affects ranking.

4) Trophies, Member Count, and Normalized Trophies: A pair plot shows the relationships between multiple variables and is useful for detecting correlations, trends, or clusters in a dataset with more than one numeric feature.

V. Data Visualization

In this section, we are gonna be able to see the results of the codes for the graphics.

A. Jessie Players Ranking Dataset

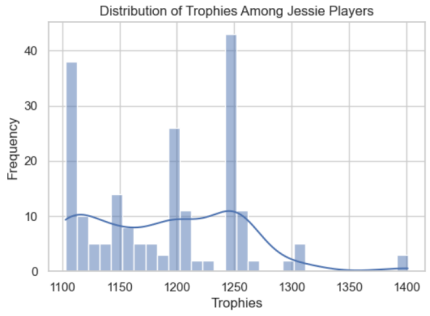

Fig. 6. Histogram of Trophies.

1) Histogram of Trophies:

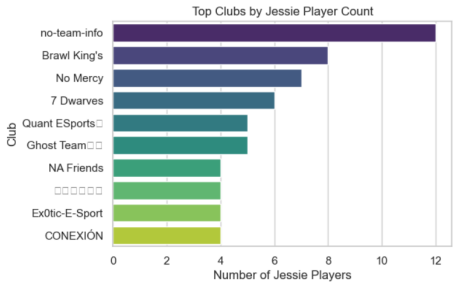

Fig. 7. Top Clubs by Player Count.

2) Top Clubs by Player Count:

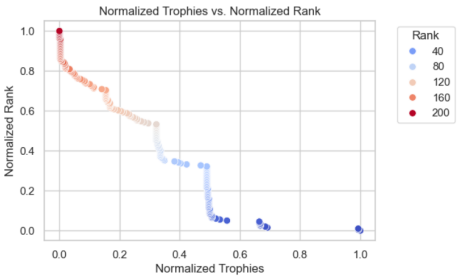

Fig. 8. Normalized Trophies vs. Normalized Rank.

3) Normalized Trophies vs. Normalized Rank:

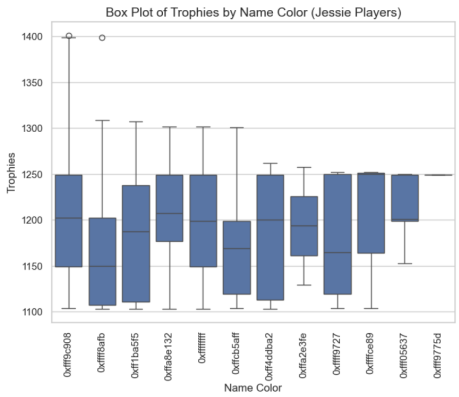

Fig. 9. Trophies by Player's.

4) Trophies by Player's:

B. Club Ranking Dataset

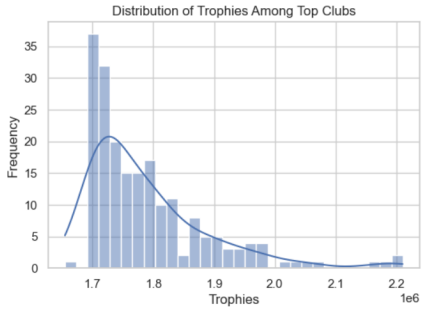

Fig. 10. Club Trophy Distribution.

1) Club Trophy Distribution:

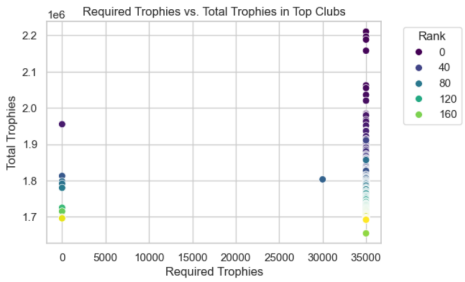

Fig. 11. Required Trophies vs. Total Trophies.

2) Required Trophies vs. Total Trophies:

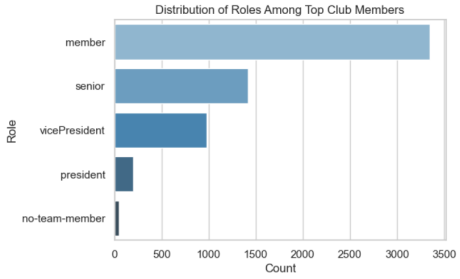

Fig. 12. Roles Distribution.

3) Roles Distribution:

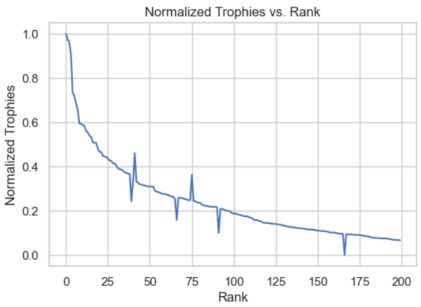

Fig. 13. Normalized Trophies vs. Rank.

4) Normalized Trophies vs. Rank:

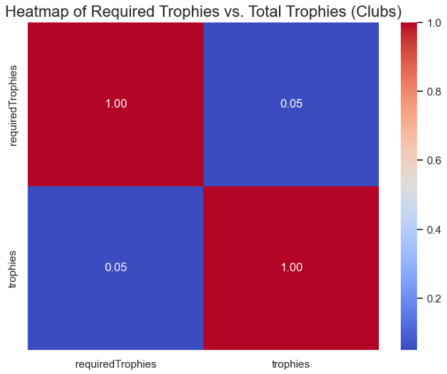

Fig. 14. Required Trophies vs. Total Trophies for Clubs.

5) Required Trophies vs. Total Trophies for Clubs:

C. Mexico Clubs Ranking Dataset

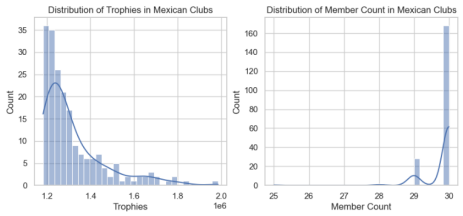

Fig. 15. Trophy and Member Count Distributions.

1) Trophy and Member Count Distributions:

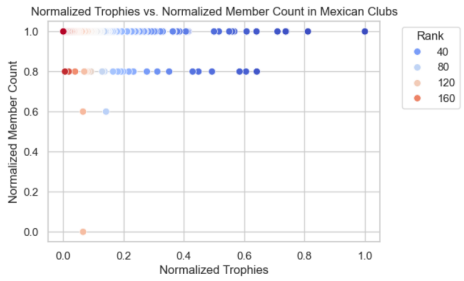

Fig. 16. Normalized Trophies vs. Member Count.

2) Normalized Trophies vs. Member Count:

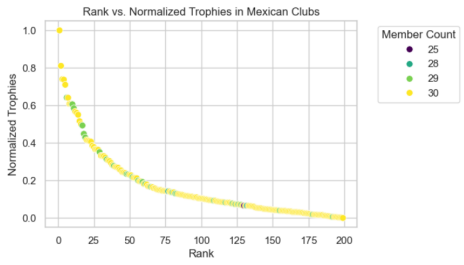

Fig. 17. Ranking Analysis.

3) Ranking Analysis:

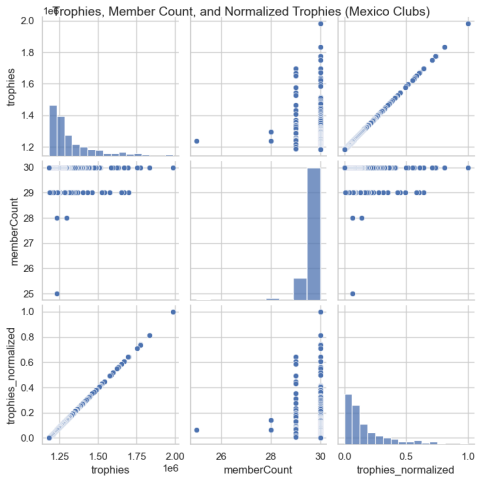

Fig. 18. Trophies, Member Count, and Normalized Trophies.

4) Trophies, Member Count, and Normalized Trophies:

VI. Interpretation of Results

In this section we´re gonna give a short interpretation of the results we can see on the graphs.

A. Jessie Players Ranking Dataset

1) Histogram of Trophies: If the histogram shows a concentration around certain values, it could indicate that most players are clustered in specific trophy ranges. A skewed distribution could suggest a few top players with very high trophies, and many players with lower trophies.

2) Top Clubs by Player Count: A few clubs with large numbers of players might suggest these clubs are highly popular or have more resources, while others with fewer members could indicate newer or less active clubs.

3) Normalized Trophies vs. Normalized Rank: If the plot shows an upward trend, it suggests that players with higher trophies tend to have better ranks. A random scatter could imply no strong correlation between the two variables.

4) Trophies by Player's: The box plot allows you to see if certain name colors correlate with higher or lower trophies. If some colors consistently have higher median trophies, it could imply a correlation between player aesthetics and their success.

B. Clubs Ranking Dataset

1) Club Trophy Distribution: A wide spread of trophy counts suggests diversity in club performance, with some clubs significantly outperforming others. A concentration around a certain range suggests that most clubs are performing at similar levels.

2) Required Trophies vs. Total Trophies: If the plot shows a positive correlation, clubs with higher entry requirements tend to have more total trophies, suggesting that tougher clubs attract stronger players.

3) Roles Distribution: A dominant role (e.g., leaders or strategists) could suggest that most clubs are structured with specific positions to optimize performance. A more balanced distribution indicates a versatile or less role-dependent structure.

4) Normalized Trophies vs. Rank: A line moving upward indicates that clubs with higher normalized trophies tend to achieve better ranks. Flat or erratic lines suggest no clear trend between trophies and rank.

5) Required Trophies vs. Total Trophies for Clubs: A strong color correlation (either positive or negative) between the required trophies and total trophies indicates a meaningful relationship. A lack of correlation (neutral color) suggests that these variables might not be closely related.

C. Mexico Clubs Ranking Dataset

1) Trophy and Member Count Distributions: Clubs with higher trophies might have fewer members, or large clubs may be less successful. The shapes of the distributions give insight into how clubs are structured in terms of achievements and size.

2) Normalized Trophies vs. Member Count: If a positive correlation is found, larger clubs might have more success, potentially due to having more resources or players to contribute to the overall performance. A lack of correlation suggests that club size doesn't necessarily impact performance.

3) Ranking Analysis: A positive correlation indicates that clubs with higher trophies tend to rank better. A scattered distribution might suggest that rank is not solely dependent on trophy count but may involve other factors like club activity or player contributions.

4) Trophies, Member Count, and Normalized Trophies: The pair plot will help identify which variables are most correlated with each other (e.g., member count and trophies) and provide insights into how these attributes relate to each other across clubs in Mexico.

VII. Conclusion

Through the data visualization techniques explored in this project, we learned how to effectively represent and analyze information to gain key insights. Visualizations such as histograms, scatter plots, bar charts, and line graphs provide a clear understanding of variable distributions, relationships between them, and overall trends within the data. It was also observed how graphical tools help identify patterns, correlations, and potential anomalies that wouldn't be easily detected through numerical statistics alone. In particular, the ability to combine multiple visualizations, such as pair plots and heatmaps, proved valuable for exploring complex relationships between variables and gaining a better understanding of data behavior. In summary, visualization techniques are essential for communicating analysis results and making informed decisions based on data observations.

Appendix A
Data Preprocessing Python Scripts

```python
import pandas as pd
import numpy as np
from sklearn.preprocessing import StandardScaler, MinMaxScaler, OneHotEncoder
from sklearn.impute import SimpleImputer

df_jessie = pd.read_csv(
"../data/raw/dataset/brawler_ranking/JESSIE_ranking.csv"
)
df_jessie.head(10)
df_jessie.info()
df_jessie.describe
print(df_jessie.columns)

null_values = df_jessie.isnull().sum()
print("Missing values per column:\n", null_values)
total_nulls = df_jessie.isnull().sum().sum()
print(f'Total missing values in the dataset: {total_nulls}')
null_percentage = (df_jessie.isnull().sum() / len(df_jessie)) * 100
print("Percentage of missing values per column:\n", null_percentage)

df_jessie['club'] = df_jessie['club'].fillna("no-team-info")
df_jessie.head(10)
df_jessie.drop(columns=['Unnamed: 0'], inplace=True)

scaler = MinMaxScaler()
df_jessie['normalized_trophies'] = scaler.fit_transform(df_jessie[['trophies']])
df_jessie['normalized_rank'] = scaler.fit_transform(df_jessie[['rank']])
df_jessie.head(10)
df_jessie.to_csv(
'../data/preprocessed/cleaned_jessie_ranking.csv',
index=False
)

df_club = pd.read_csv(
"../data/raw/dataset/global_club_info.csv"
)
df_club.head(10)
df_club.info()
df_club.columns
df_club.describe

null_values = df_club.isnull().sum()
print("Missing values per column:\n", null_values)
total_nulls = df_club.isnull().sum().sum()
print(f'Total missing values in the dataset: {total_nulls}')
df_club['description'] = df_club['description'].fillna("no-description")
df_club = df_club.fillna("no-team-member")
df_club[df_club['description'] == 'no-description']
df_club = df_club.rename(columns={'Unnamed: 0': 'Rank'})
df_club.head(10)

df_club['trophies_normalized'] = scaler.fit_transform(df_club[['trophies']])
df_club['requiredTrophies_normalized'] = scaler.fit_transform(
   df_club[['requiredTrophies']]
   )
df_club.head()
df_club.to_csv(
'../data/preprocessed/cleaned_club_ranking.csv',
index=False
)

df_mexico_clubs = pd.read_csv(
"../data/raw/dataset/country_club_rankings/Mexico_club_rankings.csv"
)
df_mexico_clubs.head(10)
df_mexico_clubs = df_mexico_clubs[df_mexico_clubs['Unnamed: 0'] != 0]
df_mexico_clubs.drop(columns=['rank'], inplace=True)
df_mexico_clubs = df_mexico_clubs.rename(columns={'Unnamed: 0': 'Rank'})
df_mexico_clubs

null_values = df_mexico_clubs.isnull().sum()
print("Missing values per column:\n", null_values)
total_nulls = df_mexico_clubs.isnull().sum().sum()
print(f'Total missing values in the dataset: {total_nulls}')

df_mexico_clubs['trophies_normalized'] = scaler.fit_transform(
df_mexico_clubs[['trophies']]
)
df_mexico_clubs['member_count_normalized'] = scaler.fit_transform(
df_mexico_clubs[['memberCount']]
)
df_mexico_clubs.head(10)
df_mexico_clubs.to_csv(
"../data/preprocessed/cleaned_mexico_clubs_ranking.csv"
)
```

Appendix B
Data Visualization Plots in Python

```python
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

jessie_data = pd.read_csv(
'../data/preprocessed/cleaned_jessie_ranking.csv'
```

```
)
club_data = pd.read_csv(
'../data/preprocessed/cleaned_club_ranking.csv'
)
mexico_data = pd.read_csv(
'../data/preprocessed/cleaned_mexico_clubs_ranking.csv'
)

sns.set(style="whitegrid")

plt.figure(figsize=(6, 4))
sns.histplot(jessie_data['trophies'], bins=30,
    kde=True)
plt.title("Distribution of Trophies Among Jessie
    Players")
plt.xlabel("Trophies")
plt.ylabel("Frequency")
plt.show()

top_clubs =
    jessie_data['club'].value_counts().nlargest(10)
plt.figure(figsize=(6, 4))
sns.barplot(x=top_clubs.values, y=top_clubs.index,
    palette="viridis")
plt.title("Top Clubs by Jessie Player Count")
plt.xlabel("Number of Jessie Players")
plt.ylabel("Club")
plt.show()

plt.figure(figsize= (6, 4))
sns.scatterplot(data=jessie_data,
    x='normalized_trophies', y='normalized_rank',
    hue='rank', palette="coolwarm", s=50)
plt.title("Normalized Trophies vs. Normalized Rank")
plt.xlabel("Normalized Trophies")
plt.ylabel("Normalized Rank")
plt.legend(title="Rank", bbox_to_anchor=(1.05, 1),
    loc='upper left')
plt.show()

plt.figure(figsize=(8,6))
sns.boxplot(x='nameColor', y='trophies',
    data=jessie_data)
plt.title('Box Plot of Trophies by Name Color (Jessie
    Players)', fontsize=14)
plt.xlabel('Name Color', fontsize=12)
plt.ylabel('Trophies', fontsize=12)
plt.xticks(rotation=90)
plt.show()

plt.figure(figsize=(6, 4))
sns.histplot(club_data['trophies'], bins=30, kde=True)
plt.title("Distribution of Trophies Among Top Clubs")
plt.xlabel("Trophies")
plt.ylabel("Frequency")
plt.show()

plt.figure(figsize=(6, 4))
sns.scatterplot(data=club_data, x='requiredTrophies',
    y='trophies', hue='Rank', palette="viridis", s=50)
plt.title("Required Trophies vs. Total Trophies in Top
    Clubs")
plt.xlabel("Required Trophies")
plt.ylabel("Total Trophies")
plt.legend(title="Rank", bbox_to_anchor=(1.05, 1),
    loc='upper left')
plt.show()

role_columns = [col for col in club_data.columns if
    'member_' in col and '_role' in col]
roles = pd.concat([club_data[col] for col in
    role_columns])
role_counts = roles.value_counts()
plt.figure(figsize=(6, 4))
sns.barplot(x=role_counts.values, y=role_counts.index,
    palette="Blues_d")
plt.title("Distribution of Roles Among Top Club
    Members")
plt.xlabel("Count")
plt.ylabel("Role")
plt.show()

plt.figure(figsize=(6, 4))
sns.lineplot(data=club_data, x='Rank',
    y='trophies_normalized')
plt.title("Normalized Trophies vs. Rank")
plt.xlabel("Rank")
plt.ylabel("Normalized Trophies")
plt.show()

corr_matrix = club_data[['requiredTrophies',
    'trophies']].corr()

plt.figure(figsize=(8,6))
sns.heatmap(corr_matrix, annot=True, cmap='coolwarm',
    fmt='.2f', cbar=True)
plt.title('Heatmap of Required Trophies vs. Total
    Trophies (Clubs)', fontsize=16)
plt.show()

fig, ax = plt.subplots(1, 2, figsize=(10, 4))
sns.histplot(mexico_data['trophies'], bins=30,
    kde=True, ax=ax[0])
ax[0].set_title("Distribution of Trophies in Mexican
    Clubs")
ax[0].set_xlabel("Trophies")

sns.histplot(mexico_data['memberCount'], bins=30,
    kde=True, ax=ax[1])
ax[1].set_title("Distribution of Member Count in
    Mexican Clubs")
ax[1].set_xlabel("Member Count")
plt.show()

plt.figure(figsize=(6, 4))
sns.scatterplot(data=mexico_data,
    x='trophies_normalized',
    y='member_count_normalized', hue='Rank',
    palette="coolwarm", s=50)
plt.title("Normalized Trophies vs. Normalized Member
    Count in Mexican Clubs")
plt.xlabel("Normalized Trophies")
plt.ylabel("Normalized Member Count")
plt.legend(title="Rank", bbox_to_anchor=(1.05, 1),
    loc='upper left')
plt.show()

plt.figure(figsize=(6, 4))
sns.scatterplot(data=mexico_data, x='Rank',
    y='trophies_normalized', hue='memberCount',
    palette="viridis", s=50)
plt.title("Rank vs. Normalized Trophies in Mexican
    Clubs")
plt.xlabel("Rank")
plt.ylabel("Normalized Trophies")
plt.legend(title="Member Count", bbox_to_anchor=(1.05,
    1), loc='upper left')
plt.show()

sns.pairplot(mexico_data[['trophies', 'memberCount',
    'trophies_normalized']])
plt.suptitle('Trophies, Member Count, and Normalized
    Trophies (Mexico Clubs)', fontsize=14)
plt.show()
```

REFERENCES

[1] A. Vidal, "Brawl stars - Rankings, players and clubs," *Kaggle*, Dec. 2023. [Online]. Available: https://www.kaggle.com/datasets/albertovidalrod/brawl-stars-rankings-players-and-clubs

[2] J. De Aquino and L. Perez, "Data Visualization Techniques for Brawl Stars Dataset Analysis: Data Exploration and Visualization," *GitHub*, Nov. 2024. [Online]. Available: https://github.com/pyrawn/Data-Visualization-Techniques-for-Brawl-Stars-Dataset-Analysis-Data-Exploration-And-Visualization

[3] L. Estrada, "SD24-ID-3B, Class Lecture: Data Preprocessing, Unit II: Basic Data Visualization Techniques, Topic: 'Introduction to Data Visualization'," Data Engineering Program, Universidad Politécnica de Yucatán, Oct. 28, 2024.

[4] L. Estrada, "SD24-ID-3B, Class Lecture: Data Preprocessing, Unit II: Basic Data Visualization Techniques, Topic: 'Applications of Data Visualization in Data Preprocessing'," Data Engineering Program, Universidad Politécnica de Yucatán, Nov. 3, 2024.