An In-depth Analysis of Preferred Online Games Amongst Gamers in Year 2023

MEMBERS:

EMOCLING, ALLYSA M.
LUCIDO, JERMAINE B.
CANTAL, ZEDRICK S.J
MALIBIRAN, ROJEAN P.

Introduction

The Project aims to conduct an analysis about current gaming trends and preferences by analyzing the most popular online games among players. Data from surveys, online gaming platforms, and gaming forums are the sources of the data used in the research, which entails visualizing data linked to various aspects of online gaming.

Our Data visualization include:

Characteristics:
examining the age, gender, location, and other demographic data of players who have a preference for a particular game.

Platforms for gaming:
showing the player distribution for each preferred game on various gaming platforms (PC, console, and mobile).

Patterns of Play:
displaying the typical playtime patterns for each game, including with its peak times and days.

Participant Engagement:
analyzing data related to user engagement, including play frequency, in-game purchases, and community interactions.

Our analysis's main questions are the following:

- Which online games are the most played by players these days?
- For each game, who are the main target demographics?
- For what kinds of games, which gaming platforms are the favorites?
- What are the engagement levels and patterns of play for these games?

The project aims to provide a thorough overview of the gaming landscape by visualizing and analyzing this data. This will assist enthusiasts, marketers, and game creators in making decisions and staying up to date with the ever-changing gaming business.

Data

The Dataset consists of 32 rows each corresponding to a survey response. There a 9 columns, capturing various aspects of participant survey responses.

Dataset includes the following data types:
- **Integer:** Age, Average Hours Spent, Contact Number
- **String:** df, Gender, Location, Gaming Platform, Gaming Frequency, Preferred Gaming Time, Favorite Game(s) at the Moment
- **Datetime:** Time Stamp

There were missing values in the "contact number" column, representing participants whose chose not to disclose their contact information, we handled the missing values by putting "n/a" on the column.

Visualization Technique(s)

To make it easier for us to compare our data, we visualized our data using bar charts and pie charts. It is simpler for us to draw precise comparisons across categories when the length of the bars in our bar chart accurately reflects the quantitative value that they represent. A pie chart can be a good option if the main objective of the chart is to illustrate how each component relates to the overall.

Regarding the libraries and package, we employed Seaborn and Matplotlib. Because of Matplotlib's great flexibility, users can construct a vast array of plots and visualizations. It is appropriate for customized and specialized data visualizations since it offers fine-grained control over all plot elements. Statistical data visualization is the primary purpose of Seaborn. With little coding, it makes it easier to create intricate statistical displays like violin plots, pair plots, and heatmaps. Seaborn includes visually appealing story styles and basic color palettes that are pleasing. As a result, we can easily produce eye-catching visualizations without requiring a lot of modification.

Implementation in Google Colab

- Import Libraries:

Import in the necessary libraries, such as Seaborn for statistical visualizations, Matplotlib for basic charting, and Pandas for data manipulation.

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

- Load and Clean Data:

To read the dataset from a CSV file (pd.read_csv) and carry out any required data cleaning operations, including addressing missing values or choosing pertinent columns, use Pandas.

```
df = pd.read_csv('sample_data/survey.csv')

df.head()
```

- Create a visualization:

Matplotlib and Seaborn can be used to generate a basic visualization. We use;

```
sns.countplot(x='Age', data=df)
plt.title('Age Group')
plt.show()
```

- Customize and Style the Plot:

To make the plot more informative, add labels, titles, and other adjustments. Plot the size, label the axes, and add a title.

```
_df_15.groupby('Location').size().plot(kind='barh',
color=sns.palettes.mpl_palette('Dark2'))
plt.gca().spines[['top', 'right',]].set_visible(False)
```

- Show the Plot:

Finally, use plt.show() to display the constructed plot.

Result and Interpretation

- **Age group**

Approximately 30% of the individuals fall into the "18 - 24 years old" category.

About 3% are "Under 18 years old."

Around 2% belong to the "25 - 34 years old" group.

This information suggests that most individuals in our dataset are between 18 and 24 years old, followed by a smaller representation of those under 18 and individuals aged 25 to 34. The countplot provides a quick overview of the age distribution, helping you identify the most prevalent age groups in our data.

- **Gender**

Male: Represents the largest portion, constituting 48.4% of the dataset.

Female: Comprises 45.2% of the dataset.

Prefer not to say: Represents a smaller portion at 6.5%.

The pie chart visually conveys the proportional representation of each gender category. It is evident that males are the most prevalent gender, followed by females, with a smaller percentage preferring not to disclose their gender. This visualization provides a clear overview of the gender distribution within your data.

- **Location**

Gatid sta Cruz Laguna

Santa Cruz Laguna

Philippines

Nagcarlan Laguna

Laguna

Each horizontal bar represents the count of data points associated with a specific location. The longer the bar, the higher the count for that particular location. The color palette used is 'Dark2' from seaborn.

Additionally, the code attempts to remove the top and right spines of the plot, likely for aesthetic purposes, to create a cleaner appearance. By examining the chart, you can identify the distribution of data points across different locations, allowing for insights into the concentration or diversity of your dataset in these geographical areas.

- **Gaming platform**

PC Mobile phone: This category has the highest count, indicating that a significant portion of respondents use both PC and mobile phones for gaming.

Mobile phone: This is the platform with the second-highest count, suggesting a substantial number of respondents primarily use mobile phones for gaming.

PC, Mobile phone, playstation response: This combination of gaming platforms has the third-highest count.

PC: The response "PC" is the fourth most common gaming platform mentioned.

Mobile phone, playstation: This combination is the least common among the mentioned gaming platforms.

The plot allows us to quickly understand the distribution of preferences for gaming platforms in our dataset. It provides insights into which platforms are more popular or prevalent among the respondents, helping us understand the gaming habits of our surveyed audience.

- **Gaming frequency**

Daily: Represents 32.3% of the respondents, indicating a significant portion engages in gaming on a daily basis.

Monthly: Constitutes 9.7% of the respondents, suggesting a smaller percentage prefers a monthly gaming frequency.

Weekly: Accounts for 25.8% of the respondents, signifying a substantial portion engages in gaming on a weekly basis.

Occasionally: Represents 32.3% of the respondents, making it the same percentage as the "Daily" category. This suggests that a considerable number of respondents game either daily or occasionally.

This visualization provides a quick overview of the distribution of gaming frequencies among the surveyed individuals. It helps us understand the prevalence of different gaming habits, highlighting whether gaming is a daily, weekly, monthly, or occasional activity for the respondents in our dataset.

- **Average spent per week.**

Less than 5 hours: This category has the highest percentage, accounting for 61.3% of respondents. The majority of individuals in your dataset spend less than 5 hours per week on gaming.

10 - 20 hours: Represents 3.2% of respondents, indicating a smaller portion spends between 10 and 20 hours per week on gaming.

5 - 10 hours: Constitutes 35.5% of respondents, suggesting a significant proportion spends between 5 and 10 hours per week on gaming.

This visualization offers insights into the distribution of average weekly gaming hours among the surveyed individuals. It helps us understand the predominant patterns of gaming behavior, such as whether most respondents are casual gamers (spending less than 5 hours per week) or if there is a notable group dedicating more time to gaming.

- **Prepared gaming time**

Evening, midnight: 9 respondents prefer gaming during both the evening and midnight, with a specified duration of 9 hours.

Evening: 14 respondents have a preference for gaming in the evening, with a specified duration of 14 hours.

Midnight: 1 respondent prefers gaming during the midnight hours, with a specified duration of 1 hour.

Morning, afternoon, evening, midnight response: 2 respondents express a preference for gaming throughout the day and midnight, with a specified duration of 2 hours.

Morning, evening, midnight: 1 respondent prefers gaming during the morning, evening, and midnight, with a specified duration of 1 hour.

Morning, evening: 2 respondents prefer gaming during the morning and evening, with a specified duration of 2 hours.

After: 2 respondents, with the response cut off, suggest a preference for gaming in the afternoon with a specified duration of 2 hours.

This detailed information provides insights into both the preferred gaming times and the specific durations respondents allocate to gaming during those times. It gives a more nuanced understanding of the gaming habits and preferences of the surveyed individuals.

- **Favorite games**

ML (Mobile Legends), Valorant, Farlight: 1 respondent currently favors these three games.

ML (Mobile Legends): 14 respondents have Mobile Legends as their favorite game at the moment.

COD (Call of Duty), Valorant: 1 respondent prefers both Call of Duty and Valorant.

COD, ML response: 6 respondents have both Call of Duty and Mobile Legends as their favorite games at the moment.

ML, Valorant response: 2 respondents favor both Mobile Legends and Valorant.

ML, COD: 3 respondents have both Mobile Legends and Call of Duty as their favorite games.

LOL (League of Legends): 1 respondent currently favors League of Legends.

Valorant: 3 respondents have Valorant as their favorite game at the moment.

This visualization provides insights into the diversity of favorite games among the surveyed individuals. It helps us understand which games are more popular or preferred by respondents at the current time, highlighting the variety of gaming interests within our dataset.

Conclusion

Genre Popularity:

- Findings: Multiplayer online battle arena and mobile legends games continue to dominate the gaming landscape.

- Takeaway: Game developers and publishers should prioritize these genres in their future releases.

Cross-Platform Gaming:

- Findings: The trend of cross-platform gaming is on the rise, with gamers preferring titles that offer seamless experiences across multiple devices.

- Takeaway: Developers should invest in cross-platform compatibility to cater to the evolving preferences of the gaming community.

Esports Influence:

- Findings: Esports continues to grow in popularity, influencing gamers' choices towards titles with established competitive scenes.

- Takeaway: Game developers and sponsors should focus on fostering esports communities around their titles for increased engagement.

Free-to-Play Model:

- Findings: Free-to-play (F2P) games with microtransactions maintain their appeal, attracting a larger player base.

- Takeaway: Publishers should consider F2P models with strategic in-game purchases to maximize revenue while keeping games accessible.

Emergence of Virtual Reality (VR):

- Findings: VR gaming experiences are gaining traction, offering a more immersive and engaging gameplay environment.

- Takeaway: Developers should explore VR integration to tap into the growing interest in immersive gaming experiences.

Community and Social Features:

- Findings: Games with robust community and social features are more likely to retain players.

- Takeaway: Emphasize the development of in-game social systems, guilds, and communication tools to enhance player retention.

Content Updates and Events:

- Findings: Games that regularly introduce new content, events, and updates maintain player interest over time.

- Takeaway: Continuous support through regular content drops and events is crucial for long-term player engagement.

Inclusivity and Diversity:

- Findings: Games that prioritize inclusivity and diversity in character representation and themes are well-received.

- Takeaway: Developers should focus on creating inclusive and diverse gaming experiences to resonate with a broader audience

Potential future improvements or extensions to the project.

Real-Time Data Integration:

- Improve the project's capability to integrate real-time data. This involves setting up systems to constantly collect and analyze the most recent gaming trends and preferences.

Machine Learning for Predictive Analysis:

- Implement machine learning algorithms to predict future gaming trends based on historical data. This could help in anticipating shifts in preferences and staying ahead of the curve.

Global and Regional Breakdowns:

- Expand the analysis to provide a more granular view of preferences, taking into account regional variations. Different gaming cultures and preferences may emerge in various parts of the world.

Social Media Sentiment Analysis:

- Integrate social media sentiment analysis to gauge how gamers discuss and feel about specific games. This can provide a qualitative layer to the quantitative data.

User Surveys and Feedback Integration:

- Incorporate user surveys and feedback directly into the analysis. Understanding player sentiments and preferences through direct feedback can offer valuable insights.

Gaming Industry Partnerships:

- Establish partnerships with gaming industry stakeholders, such as game developers, publishers, and esports organizations, to gain access to exclusive data and insights.

In-Depth Player Profiles:

- Develop more detailed player profiles, considering factors such as age, gender, gaming platform preferences, and gaming history. This can help create a more nuanced understanding of different player segments.

Emerging Technologies Analysis:

- Include an analysis of emerging technologies that might impact the gaming landscape, such as augmented reality (AR), blockchain in gaming, or other innovative trends.

Security and Privacy Measures:

- Strengthen the project's security measures to ensure the protection of sensitive gaming data and user privacy, especially if the analysis involves personally identifiable information.

Visual Representation Enhancement:

- Improve data visualization techniques to make the findings more accessible and engaging. This could involve interactive dashboards, dynamic charts, and multimedia elements.

Appendix

Variable Name	Description	Data Type	Allowed Values/Range	Units	Missing Values	Example	Source
Age	Age of the respondent	Categorical	Under 18, 18-24, 25-34, 35-44, 55+		No	25-34	Surveys
Contact Number	Contact number of the respondent	String			Yes	+1234567890	Surveys
Location	Location of the respondent	String			Yes	New York	Surveys
Gender	Gender of the respondent	Categorical	Male, Female, Prefer not to say		Yes	Female	Surveys
Gaming Platform	Gaming platform used by respondent	Categorical	PC, Mobile Phone, Playstation, Xbox, Nintendo		Yes	Playstation	Surveys
Gaming Frequency	Frequency of gaming	Categorical	Daily, Weekly, Monthly, Occasionally		Yes	Weekly	Surveys
Average Hours Spent per Week	Average hours spent on gaming per week	Categorical	Less Than 5 Hours, 5-10 Hours, 10-20 Hours, 20+ Hours		Yes	10-20 Hours	Surveys
Preferred Gaming Time	Preferred time for gaming	Categorical	Morning, Afternoon, Evening, Midnight		Yes	Evening	Surveys
Favorite Game(s) at the Moment	Current favorite game(s)	Categorical	Mobile Legends (ML), Call of Duty Mobile (COD), League of Legends (LOL), Valorant, Farlight		Yes	League of Legends (LOL)	Surve

```
# Check for missing values
print(data.isnull().sum())  # This will show the count of missing values in each column

# Handle missing values
data = data.dropna()  # This will drop rows with any missing values
# Or you can fill missing values with a specific value
# data['column_name'].fillna(value, inplace=True)

# Convert data types if needed
```

```python
# data['column_name'] = pd.to_numeric(data['column_name'])
# data['date_column'] = pd.to_datetime(data['date_column'])

# Remove duplicates if necessary
data = data.drop_duplicates()

# Perform additional data cleaning and processing steps based on your requirements

# Save the cleaned data to a new CSV file
```