COMP-2702 – Data Management
Final Project Proposal

Data Analyst's Name: Mariah Anjannelle A. Quinquito

Project Name: eSports Earnings for Players & Teams by Game

Project Date: April 5, 2024

1. Project Description:

Using data from eSportsEarning.com, I will do a thorough analysis of earnings in the eSports sector for this project. Finding important information on how winnings are split between teams and players across a range of well-known games, such as League of Legends, Dota 2, Counterstrike: Global Offensive, Fortnite, and more, is my primary objective. I want to shed light on the patterns and trends that influence financial success in the realm of competitive gaming by looking at data on the highest-paid individuals and teams.

I'll be looking at things like the geographic distribution of earnings, trends in earnings over time, and any relationships between player/team performance, the popularity of the game, and regional effects on earning potential throughout this investigation. To convey the results in an understandable and perceptive way, I also intend to make use of statistical and visualization methods. In the end, I want to help advance knowledge of the economics of the eSports sector by offering insightful analysis to sponsors, teams, players, and other eSports ecosystem participants.

I chose this project because of my keen interest in the eSports industry. As a passionate follower of eSports, I'm intrigued by the rapid growth and evolution of this sector, particularly in terms of its economic dynamics. The opportunity to explore the earnings of players and teams across various games aligns perfectly with my curiosity about the financial aspects of competitive gaming. Additionally, I see this project as a chance to apply my data analysis skills to a domain I'm genuinely enthusiastic about, allowing me to contribute meaningfully to the understanding of the eSports ecosystem.

2. Business Case

This project addresses the growing need for comprehensive insights into the economic landscape of the eSports industry. As eSports continues to gain mainstream recognition and attract significant investments, there is a heightened interest in understanding the financial dynamics within this rapidly evolving sector. Analyzing earnings data of players and teams across various games provides valuable insights into the distribution of wealth, the factors influencing earning potential, and the overall economic viability of eSports as a profession. By delving into this data, we can uncover trends, identify opportunities, and inform strategic decision-making for players, teams, sponsors, investors, and other stakeholders within the eSports ecosystem.

To conduct this analysis, I've selected data from eSports Earnings in Kaggle as the primary source. This dataset is renowned for its comprehensive coverage of earnings in the eSports industry, providing detailed information on the highest-earning players and teams across a diverse array of games. The choice of this

source is grounded in its reliability, credibility, and depth of data it offers, enabling a thorough exploration of earnings trends spanning different games and regions. The target audience for this analysis encompasses eSports professionals, team managers, tournament organizers, sponsors, investors, and anyone keen on gaining insights into the financial intricacies of competitive gaming. By furnishing actionable insights derived from this dataset, the analysis endeavors to equip stakeholders with the knowledge to make informed decisions and navigate the eSports landscape more adeptly.
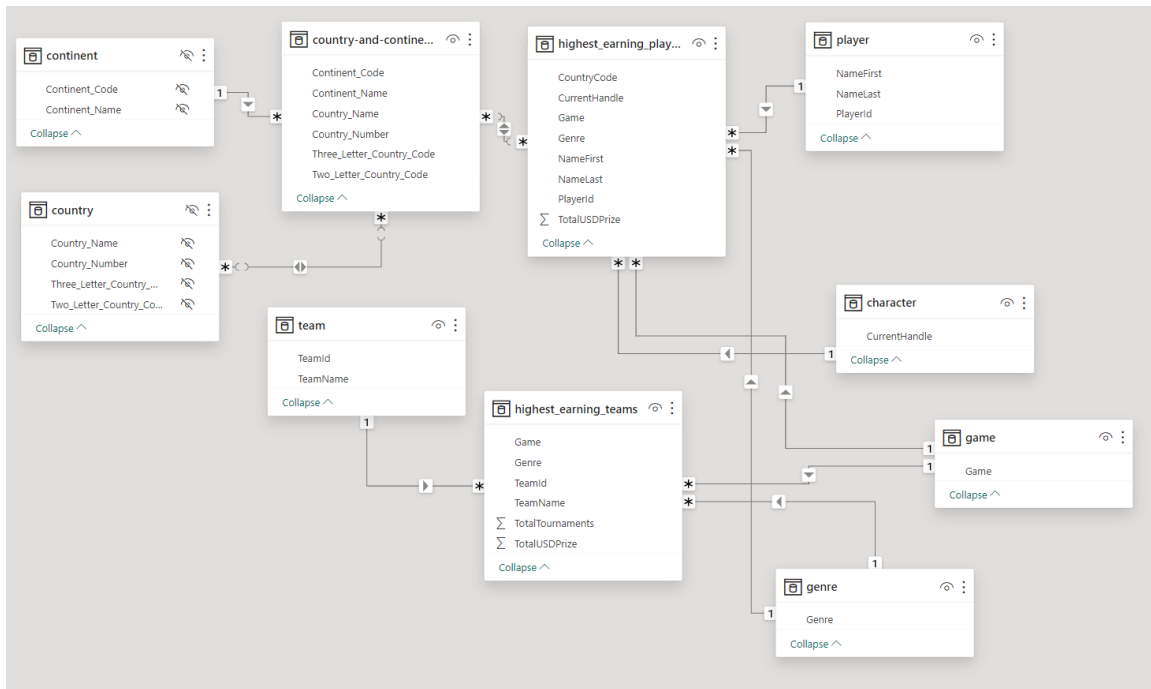
The primary data source for this analysis is sourced from Kaggle, specifically from the dataset titled "eSports Earnings for Players & Teams by Game," created by user jackdaoud. This dataset provides comprehensive information on earnings for both players and teams across various eSports games. It includes data on the highest-earning players and teams, covering a wide range of games such as Dota 2, Counterstrike: Global Offensive, Fortnite, League of Legends, and more.

The dataset can be accessed and downloaded from the following location on Kaggle: https://www.kaggle.com/datasets/jackdaoud/esports-earnings-for-players-teams-by-game
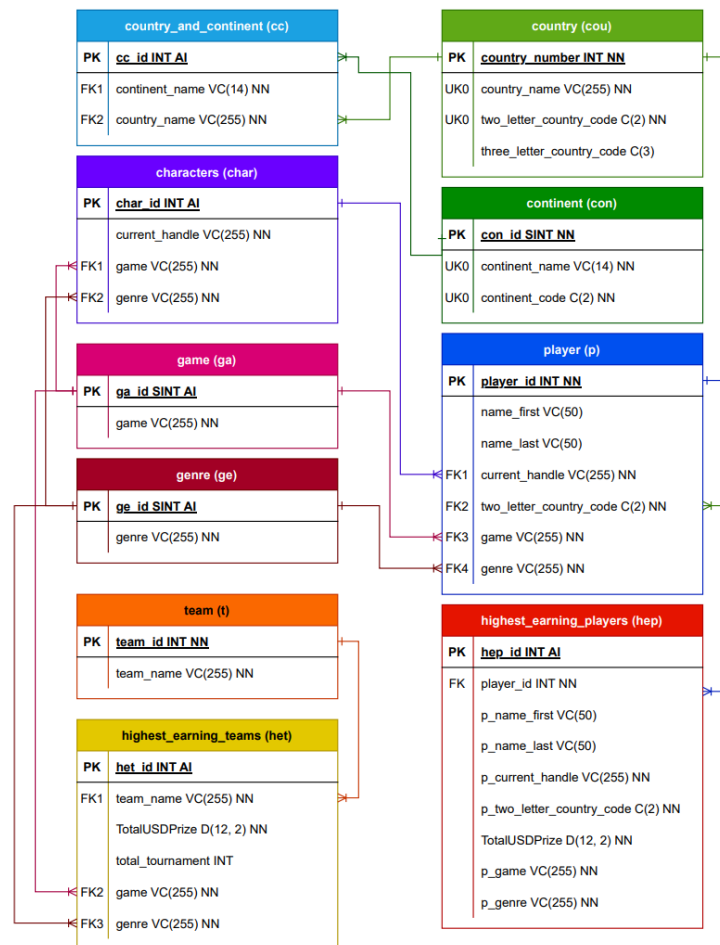
This dataset was chosen due to its credibility, reliability, and comprehensiveness in providing earnings data within the eSports industry. It offers valuable insights into the financial aspects of competitive gaming, making it a suitable data source for conducting in-depth analysis and exploration.

3. **Business Rules/Assumptions**
   a) A player can belong to only one team at a given time.
   b) A team can consist of multiple players.
   c) A player may have earnings from participating in multiple games.
   d) A team can earn income from participating in multiple games.
   e) A player's earnings within a specific game should be recorded as individual entries.
   f) A team's earnings within a specific game should be recorded as individual entries.
   g) A player's earnings can vary depending on their performance in different tournaments or events within a game.
   h) A team's earnings can vary depending on their performance in different tournaments or events within a game.
   i) Each entry in the dataset should include the player's or team's name, the game they participated in, the earnings amount, and other relevant details.
   j) Players or teams may have earnings in different currencies, which should be standardized for analysis.
   k) The dataset should include information on the date or period during which the earnings were accrued.
   l) Each player or team entry should have a unique identifier to avoid duplication.
   m) The dataset should accurately represent earnings data without any missing or incomplete entries.
   n) The dataset should be regularly updated to reflect the latest earnings information within the eSports industry.
   o) The analysis should consider regional differences in earnings due to varying tournament prize pools, sponsorships, and other factors.

4. Entity Relationship Diagram (ERD)

country_and_continent (cc)

PK	cc_id INT AI
FK1	continent_name VC(14) NN
FK2	country_name VC(255) NN

country (cou)

PK	country_number INT NN
UK0	country_name VC(255) NN
UK0	two_letter_country_code C(2) NN
	three_letter_country_code C(3)

characters (char)

PK	char_id INT AI
	current_handle VC(255) NN
FK1	game VC(255) NN
FK2	genre VC(255) NN

continent (con)

PK	con_id SINT NN
UK0	continent_name VC(14) NN
UK0	continent_code C(2) NN

game (ga)

PK	ga_id SINT AI
	game VC(255) NN

player (p)

PK	player_id INT NN
	name_first VC(50)
	name_last VC(50)
FK1	current_handle VC(255) NN
FK2	two_letter_country_code C(2) NN
FK3	game VC(255) NN
FK4	genre VC(255) NN

genre (ge)

PK	ge_id SINT AI
	genre VC(255) NN

team (t)

PK	team_id INT NN
	team_name VC(255) NN

highest_earning_players (hep)

PK	hep_id INT AI
FK	player_id INT NN
	p_name_first VC(50)
	p_name_last VC(50)
	p_current_handle VC(255) NN
	p_two_letter_country_code C(2) NN
	TotalUSDPrize D(12, 2) NN
	p_game VC(255) NN
	p_genre VC(255) NN

highest_earning_teams (het)

PK	het_id INT AI
FK1	team_name VC(255) NN
	TotalUSDPrize D(12, 2) NN
	total_tournament INT
FK2	game VC(255) NN
FK3	genre VC(255) NN

a) Entity Description and Relationships:
   o Country and Continent (cc_table): Represents the relationship between countries and continents. One country can belong to only one continent, but a continent can have multiple countries. Symbols: One-to-many relationship.
   o Player (player_table): Contains information about individual players. Players can belong to only one country but can participate in multiple games and genres. Symbols: One-to-many relationship.
   o Highest Earning Players (hep_table): Stores data on players with the highest earnings. It is related to the player_table. Symbols: One-to-many relationship.
   o Team (team_table): Contains information about gaming teams. A team can have multiple players. Symbols: One-to-many relationship.
   o Highest Earning Teams (het_table): Stores data on teams with the highest earnings. It is related to the team_table. Symbols: One-to-many relationship.
   o Character (char_table): Contains information about characters used by players. A character can be associated with multiple games and genres. Symbols: One-to-many relationship.
   o Game (game_table): Represents different games that players can participate in. A game can have multiple characters associated with it. Symbols: One-to-many relationship
   o Genre (genre_table): Represents different genres of games. A genre can have multiple characters associated with it. Symbols: One-to-many relationship.
b) One-to-Many Relationship:
   o This symbol indicates that there is a one-to-many relationship between two entities.
   o It signifies that one record in the "one" entity can be related to multiple records in the "many" entities.
   o For example, in the relationship between the player_table and the country_table, one player can belong to only one country (one side), but one country can have multiple players (many sides).
c) Bridge or Joiner Tables:
   o No explicit bridge or joiner tables are created in this dataset.
d) Table Creation Explanation:
   o Each table was created to organize and manage specific types of data separately, facilitating efficient querying and analysis.
   o Normalization principles were followed to avoid data redundancy and ensure data integrity.
e) Primary Key (PK) in Joiner Table:
   o Since there are no explicit joiner tables, there's no need for a primary key specific to joiner tables. Primary keys were assigned to each table based on unique identifiers for that entity.
   o Primary keys were chosen based on attributes that uniquely identify each record within the table.
f) Foreign Keys (FKs) in Joiner Table:
   o In the absence of joiner tables, foreign keys were utilized in other tables to establish relationships between entities.

- o Foreign keys were used to reference primary keys in related tables, maintaining data integrity and enforcing referential integrity constraints.
- o For example, in the player_table, the two_letter_country_code serves as a foreign key referencing the country_table, establishing a relationship between players and their respective   countries.

5. Power BI Dashboard

Overview:

a) Funnel Chart:
   - o Category: Characters (CurrentHandle)
   - o Measure: Players Details
b) Donut Chart:
   - o Legend: Genres
   - o Values: Team Name and Player Details
c) Pie Chart:
   - o Legend: Game
   - o Value: Team Name and Player Details
d) Map:
   - o Location: Country Code (Two-Letter-Country)
   - o Legend: Players Details

Players Dashboard:

e) Slicers:
   - o Field: Players Details, Genre, Game, Current Handle
f) Aster Plot:
   - o Visualization: Pie chart illustrating distribution of earnings by genre or game.
   - o Value: Proportion of earnings contributed by each genre or game.
   - o Slicer: Continent, country, or any other relevant category for filtering.
g) Scatter Plot:
   - o Category: Genre
   - o Y-Axis: Total USD Prize
h) Guage:
   - o Value: Total USD Prize
i) Pie Chart:
   - o Legend: Game
   - o Value: Total USD Prize
j) Tree Map:
   - o Category: Current Handle
   - o Details: Players Details
   - o Values: Total USD Prize
k) Clustered Column Chart:
   - o X-Axis: Players Details
   - o Y-Axis Total USD Prize

l) Table:
  o Columns: Player Id, Name First, Name Last, Genre, Game, Current Handle, Country Code, Total USD Prize
m) Filled Map:
  o Location: Country Code (Two-Letter-Country)
  o Legend: Player Id

Teams Dashboard:

n) Card:
  o Fields: Total Tournaments
o) Guage:
  o Value: Total USD Prize
p) Pie Chart:
  o Legend: Game
  o Value: Total USD Prize
q) Slicer:
  o Field: Team Name, Game, Genre
r) Stacked Column Chart:
  o X-Axis: Team Name
  o Y-Axis: Total USD Prize
s) Table:
  o Column: Team Id, Team Name, Total USD Prize, Total Tournaments, Game, Genre
t) Treemap:
  o Category: Team Name
  o Details: Team Id
  o Values: Count of Genre
u) Enlighten Aquarium:
  o Fish: Team Name
  o Fish Size: Total USD Prize

6. **Annexure**
  - Conducted preliminary analysis of the eSports Earnings dataset from Kaggle to understand the structure and content of the data.
  - Conducted initial exploratory data analysis to understand the structure and content of the dataset thoroughly.
  - Adjusted business rules where necessary to align them with the characteristics of the dataset and project objectives.
  - Identified key variables such as Player Name, Team Name, Game, Genre, Earnings, Country, and Continent
  - Established a clear data schema based on the provided ERD, including entities such as Player, Team, Game, Genre, Country, and Continent.
  - Cleaned and preprocessed the dataset to ensure uniformity and consistency in data representation.
  - Updated the ERD as necessary to accommodate any changes or additions to the data model.

- Planned out Power BI Dashboard Visualizations.
- Edited the Power BI Dashboard according to Final Deliverable output.