**Title:** "Video game genre popularity: An analysis of genre popularity over time with respect to changing cultural norms and societal impact"

**Introduction:**

Video games have played a very important role in our lives for the better half of the last 50 years. From the humble beginnings of ‘Pong’ and arcade machines to now, being a multi-billion dollar industry, video games have played an integral part of our culture today. During this time, the supposed demographic or target audience for video games have drastically changed due to evolving societal norms. One being the decrease of it being a male dominated space, where nearly half of games in the US identify as female. Video games have always been stigmatised to a certain degree, whether it be the increase of violence in youth or the supposed stereotype of the gaming space being inherently misogynistic. Although these opinions are rooted in an element of truth, they are not fully conclusive.

In this proposal we will be looking into what factors play into the success of a certain genre of video games, how these adapt as time goes on and how they change with current trends. We want to explore the rise of certain genres in recent years in relation to the influx of female players in the last decade. We want to examine what video game genres are experiencing an increase in popularity/sales based on a change in customer demographics due to changing social norms.

Additionally, we want to explore the connection between the increase of violent video games in recent years and whether or not they have any influence over the spike in minor/young adult violence in the West. Many people have challenged whether or not the rise of violent video games being made easily attainable by the youth is a huge factor playing into the spike of egregious violence carried out by minors, specifically in the US. With reference to data science related articles showcasing challenges and observations within the gaming industry and the data sets cited below, we wish to accurately portray our findings.

**Objectives:**

We wish to achieve this by;

1. Finding a topic of interest with which to base the capstone project on, that also has potential application in a business setting.
2. Collect datasets relevant to video game sales, genre popularity & research social norms around video game consumption.

2.1) Analyse these datasets using methods demonstrated in classes to date.

2.2) Review sociological studies around video game consumption by gender to determine a basis for the problem statement.

1. Perform EDA on datasets to determine what variables will be analysed to differentiate games by genre and what possible analytical methods can be used to predict future popularity/sales potential.
2. Clean, merge if necessary, compare and analyse the data to determine trends in genre popularity over time.
3. Collect findings from analysis and review as follows:

5.1) Review findings to determine what genres are experiencing increases in "popularity velocity" and could development studios target them for future games.

5.2) Review genre popularity velocity over time with respect to youth violence rates.

**Problem Definition:**

Video game consumer demographics are evolving as social norms change around their consumption. What historically has been a male dominated consumerbase is becoming more inclusive. However the genres enjoyed by newer consumers differ to that of the historical consumer. This project aims to analyse the changes in genre trends over time with regards to sales figures recorded for games of each genre type.

From this analysis of games sales over time we will determine what genres are experiencing greater rates of increased popularity (popularity velocity). This greater "velocity" of popularity could be used as a metric for video game producers to cater to these growing markets.

**Scope:**

The scope of work to be done for this project is as follows: After determining a topic of mutual interest, in this case Video Games, we highlighted what topics within that area of interest we would like to analyse. Steven O'Sullivan - The changes in genre popularity over time with respect to increased inclusivity of the consumerbase. Courtney Tracey - Investigate genre popularity over time in the context of youth-violence and misogyny.

Shared metric with which to pursue areas of interest: Genre popularity over time.

In semester one we will:

1) Research datasets relating to the video game industry for suitability.

2) Research sociological studies relating to video game consumption with which to inform our analysis.

3) Perform a preliminary review of findings to determine possible influences of bias within our selection of datasets/research papers/articles.

4) Perform EDA on selected datasets to determine key variables/metrics for analysis.

5) Determine what methods might be best suited for our respective analytical focus(Classification vs Regression).

In semester two we will:

1) Apply the decided upon method from Semester One, Step 5 to datasets.

2) Review findings obtained from initial analysis to determine if methods require adjustment.

2.1) Review precision of analysis with respect to method used (e.g. if using semi-supervised regression model).

3) Collect findings and analyse them within the context of area of interest to determine if problem definition is applicable (e.g. consumerbase influences genre popularity or genre popularity influences youth violence rates).

**Data Sources:**

Below are 3 major datasets we will be analysing throughout this project, taking comparable columns paired with societal and circumstantial issues and seeing if there are correlations between the three.

1. [Video Games Data](https://www.kaggle.com/datasets/ghassenkhaled/video-games-data?select=Video_Games.csv): Our base dataset, containing a list of each factor of the most popular video games, such as the name, the year it was published, the publishing company, the genre of said game, what platform this game was released on and the number of sales for North America, Europe, Japan and Other (not previously mentioned locations). This Dataset is free to use and provided by CC0: Public Domain. All data presented is ethically sourced, copyright free and gathered consensually.
2. [Video Game Data with EDA ratings](https://www.kaggle.com/datasets/imohtn/video-games-rating-by-esrb): A dataset containing a list of the most popular games in recent years, similarly to the previous mentioned dataset, with additional information. It includes titles, the ERSB rating of each game and the nature of graphic content within that game. This Dataset is free to use and provided by CC0: Public Domain. All data presented is ethically sourced, copyright free and gathered consensually.
3. [US school shootings/minor violence 1990 - 2023](https://www.kaggle.com/datasets/joebeachcapital/school-shootings?select=school-shootings-data.csv): This dataset provides a wide range of information in the realm of gun violence and mass shootings that have taken place in the US. Necessary data taken from this dataset in relation to our research would be the age of the perpetrators and the increase of frequency in which these attacks take place. This data set is licensed and provided by CC BY-NC-SA 4.0. All data presented is ethically sourced, copyright free and gathered consensually.

**Ethical Considerations:**

All datasets used throughout our capstone project are public knowledge and produced either by the companies mentioned themselves through end of the year reports and published statistics. All participants have been granted anonymity, with all defining characteristics and identifying information being redacted. The datasets we have chosen to work with have all been ethically sourced and only provide the specific details needed to conduct our research (gender, broadened location, age).