**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 play an integral part of our culture today. During this time, the supposed demographic / target audience for video games has drastically changed due to evolving societal norms - one being the decrease of gaming being a male dominated space[1], where nearly half of gamers in the US identify as female. Video games are consistently stereotyped to a degree, whether it be the increase of violence in youth or that of the gaming space being inherently misogynistic. Although these opinions are rooted in an element of truth, they are not conclusive.

In this proposal we investigate what factors play into the success of chosen genres of video games, how they adapt over time and how they fluctuate with current trends. We explore the rise of certain genres in recent years[2], relating to an increase of female identifying players in the last decade. We want to examine which video game genres are experiencing an increase in popularity / sales based on changes in customer demographics due to varying social norms.

Additionally, we discuss the connection between the increase of violent video games in recent years and if they mirror spikes in minor / young adult violence in the West. Many have challenged whether the rise of violence in minors ties to easily obtainable video games portraying varying levels of aggression, and if it factorises the spike of egregious violence carried out by minors, with emphasis on the US. By referencing data science related articles that showcase challenges and observations within the gaming industry, backed up data sets cited below, we will accurately portray our findings.

**Objectives:**

We wish to achieve this by:

1. Collect datasets relevant to video game sales, genre popularity & research social norms around video game consumption.

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

1.2) 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:
   1. Review findings to determine what genres are experiencing increases in "popularity velocity" and could development studios target them for future games.
   2. Review genre popularity velocity over time with respect to youth violence rates.

**Problem Definition:**

Video game popularity and consumer demographics are changing over time as social norms change around their consumption. What has historically been a male consumer/niche hobby is no longer true and our Capstone project aims to analyse the potential influence genres have on consumers behaviour as well as potential influence of consumers on genres popularity.

In essence we aim to analyse the relationship between consumers and genres.

**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.

* The changes in genre popularity over time in the context of increase inclusivity of consumers.
* 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 to inform our analysis.

3) Perform a preliminary review of project to exclude potential personal biases

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

5) Determine what methods might be best suited for our variables/metrics (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.
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.
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).

References:

[1] Leonhardt, M. and Overå, S. (2021). Are There Differences in Video Gaming and Use of Social Media among Boys and Girls?—A Mixed Methods Approach. *International Journal of Environmental Research and Public Health*, 18(11), p.6085. doi:https://doi.org/10.3390/ijerph18116085.

‌[2] Veltri, N., Krasnova, H., Baumann, A. and Kalayamthanam, N., 2014. Gender differences in online gaming: A literature review.