Hannah Tieman

Intro to Statistics - CIST 2500

**Gender Representation in Video Games (2012-2022)**

**Introduction**

Throughout the past twenty years, we’ve seen a massive change in the way that different genders are represented in media, including video games. The 80s and 90s flourished with the idea that women were prizes to be won, or a quest to be sought after. With the introduction of more inclusive ideas being shared among different facets of the media in our society, there are more people willing to tell the stories of people who are not men, and do not follow the normalized gender binary.

With that being said, I chose to do my semester project on the representation (or lack thereof) of women and non-gender-conforming folk in video games from 2012 to 2022, including the roles they play and if they are sexualized or not. This is an interesting topic to look at because not only has the societal standards for men and women changed, we are also recognizing that some people do not identify as either a man or a woman, usually referred to as non-binary or genderqueer. By showing these people in media, it not only educates the public on social issues occurring, but also normalizes the idea that some people are different or identify outside the “normal” gender binary – and that’s okay.

In addition, it’s important to note that the games that these characters are featured in may or may not sexualize said characters. The sexualization of these groups of people not only fetishizes their very identity, but also creates the assumption that they solely exist to live in “normal”, gender-conforming people’s fantasies, and do not have a story of their own to tell, which is simply irresponsible.

**Dataset Description**

The data set I’m choosing to look at was compiled by Brisa Palomar earlier this month, so the data is fresh and relevant. This information was gathered and organized into three different files (games, characters, and sexualization), but I will shorten the data to what is relevant to this term paper.

The ‘games’ category covers the type of game (genre, topic, theme), which is the fundamental consensus of what the game is. The ‘characters’ category looks specifically at the characters involved in the game and tackles the gender identity and role they play in the storyline of the game as a whole. And lastly, the ‘sexualization’ category covers the sexualization or inappropriate recognition of certain non-male characters in the games mentioned.

The data set follows a set of predetermined guidelines provided by the author, which I’ll include below:

**GAME SELECTION CRITERIA**

1. The games must have a storyline. To analyze the paper of each character is essential that games had a plot where characters had a role assigned (even if this role might change based on the player's choices). This excludes games like:

- Puzzle games: Tetris, Candy Crush, Minesweepers, etc.

- Racing games: Gran Turismo, Formula 1, Mario Kart, etc.

- Social Simulators: Animal Crossing, The Sims, etc.

- MMORPGs: World of Warcraft, Skyrim, etc.

- Shooters with no story mode: Fortnite, Valorant

- Other popular games with no storyline: Minecraft, Roblox, League of Legends

2. For games that offer a story and multiplayer modes (like Call of Duty, GTA V…), just the story mode is taken into consideration for this analysis.

3. The games were selected for being top-selling or best-rated games of the year.

4. At least 5 games were selected for each year

In addition to these preliminary rules, there are multiple characterizations of the data, which will be summarized at the end of this report. These terms allow the viewer to organize the information into different sects, including how they might relate to others in the set. Given that this dataset is a collection of three spreadsheets, I’ll be combining all of the terms into one section for brevity.

**Descriptive Statistics (Chapters 2-3)**

**Inferential Statistics (Chapters 6-13)**

**Regression (Chapters 14-15) – Extra Credit**

**Discussion**

**Conclusion**

**Definitions**

**Game ID** - A unique set of letters and numbers that identifies a game.

**Title** - the title of the game release

**Date** - The date when the game was first released.

**Series** - Series where the game belongs, if any.

**Country** - Country of the game developer

**Platform** - Platforms where the game is available. In the case of the game being available on other platforms years later, just the original is noted.

**PEGI** - the Pan-European Game Information rating for that game. It indicated the minimum age recommendation for a video game.

**Customization** - If the game offers the option of customizing or not the character. It contains three values: ‘Yes’, ‘No’, and ‘Non-Binary’.

**Protagonist** - number of protagonists for that game.

**Protagonist\_non\_**male - number of non-male protagonists. This is, female, non-binary, and customizable characters.

**Relevant\_males** - male characters in the game.

**Relevant\_no\_males** - non-male characters in the game

**Percentage\_non\_male** - the percentage of non-male characters

**Criteria** - criteria for selecting the game. Contains three values:

- ‘TR’ - top rated

- ‘MS’ - most sold

- ‘SR’ - sales and rating.

**Director** - game director gender. Can contain 4 values:

- ‘M’ - male

- ‘F’ - female

- ‘NB’ - non-binary

- ‘B’ - both in the case that there is more than one director of different genders.

**Total\_team** - number of main people involved in the game creation. Includes main programmers, developers, directors, producers, artists, and designers.

**Female\_team** - number of team integrants that are female.

**Team\_percentage** - the percentage of women in the team.

**Metacritic** - punctuation out of ten given to the game by Metacritic.

**Destructoid** - punctuation out of ten given to the game by Destructoid.

**IGN** - punctuation out of ten given to the game by IGN.

**GameSpot** - punctuation out of ten given to the game by GameSpot.

**Avg\_reviews** - the average of the four previous columns.

**Character’s Name** - the name of the character

**Gender** - the gender of the character. It contains 4 different values:

- ‘Female’ (she/her)

- ‘Male’ (he/him)

– ‘Non-binary’ (they/them) - characters whose gender is purposely left ambiguous, those that due to their nature don’t have a gender and no gender has been assigned to them (animated object, plants, animals…), and those who self-identify as non-binary.

- ‘Custom’ - those characters that the game offers the option to customize their gender. They are addressed according to the gender chosen by the player.

**Age** - the age of the character during the game events.

**Age\_range** - a ranking categorization of the ages. The values are as follows:

- Infant - 0 to 5 years old

– Child - 6 to 14 years old

- Teenager - 15 to 17 years old

- Young-Adult - 18 to 24 years old

- Adult - 25 to 39 years old

- Middle-Aged - 40 to 64 years old

- Elderly - older than 65 years old

- Unknown - characters whose age is unknown

**Playable** - if the character is playable. It can be either: - 0 - No - 1 - Yes

**Sexualization** - a punctuation out of 4 determined by the data frame ‘Sexualization’.

**Species** - the species of the character.

- Human - all those characters with human-like appearances. This includes humans but also other fantasy races whose appearance is basically that of a human with very little physical divergences like gods, elves, dwarfs…

- Humanoid - characters whose appearance is anthropomorphic but their physical appearance makes it obvious that they are not human.

- Animal-Humanoids do not belong here as they have their own category.

- Humanoid-Animal - characters that are a combination between a human and a real-life animal. The animal features must be recognizable enough to know the type of animal, otherwise, they are considered Humanoids.

- Android-Robot - androids, robots that have a human appearance.

- Robots - robots that do not have a human appearance. This includes artificial intelligences without a physical body.

- Animated object - characters whose physical appearance is that of an object, but that has been given consciousness and some human features such as a face, arms or legs.

- Animated plant - plants that keep their plant appearance but that have been given some human features such as a face, arms, or legs.

- Animal - characters that are physically recognizable real-life animals. They might have some special features such as speech ability or magic powers, but it does not affect their looks.

- Creatures - all those creatures that do not fit in any of the above categories. It includes monsters whose physical appearance is too distant from that of a human to be considered humanoids, fantastic animals that do not look like any real-life animal, mythological creatures…

- Unknown: those characters' whole physical appearance is never revealed in the game.

**Side** - the side in which a character is.

- P - on the protagonist's side. This includes characters that are supportive or neutral towards the protagonist.

- A – antagonists

- B - characters that are or can be on both of the above sides.

**Relevance** - the relevance of a character in a game and in relation to the protagonist

- PA - protagonist - the most important person(s) in the game. In case of is more than one both of them must have the same importance in the plot.

- DA - deuteragonist. The second most important character(s) in the plot. In case of being more than one, both must have the same relevance.

- SK - sidekick. Those characters accompany the protagonist during all, or most of, the story. They offer constant support to the protagonist by giving advice, battle support, exploration aid, etc. They differ from the deuteragonist as they usually have little to no relevance in the storyline.

- MC - the main character. A character that is relevant throughout all, or most part of, the story. This category can include antagonists.

- SC - secondary character. A character that is important in the storyline but whose relevance is occasional, be it because their plot just lasts a short amount of time or because they are mentioned throughout the game but barely appear in-game. This category can include antagonists.

- MA - main antagonist. The main antagonist(s) of the game. It is relevant throughout the game.

**Romantic interest** - If that character is, or can be, the romantic interest of the protagonist. This includes one-sided romances as long as the one who holds the romantic feelings is the protagonist and also love-less relationships as long as the characters are married. A protagonist only can be a romantic interest if there is more than one protagonist and they are each other’s romantic interest.

- No - this character is no romantic interest of the protagonist

- Yes - this character is unavoidably the romantic interest or partner of the main character.

- Opt - this character can be dated or can have romantic or sexual interactions with the protagonist based on the player’s choice.

**Sexualized clothing** - if the clothes of a character are sexualized (1) or not (0). They just need to meet one of these criteria to be TRUE:

- The character is wearing clothes that are not age-appropriate: too sexualized adult clothes on children or children-like sexualized clothes on adults.

- The character is wearing clothes that make no sense: high-heels for running and fighting, too revealing clothes in extreme weather conditions such as heavy rain or cold, and armor that does not cover vital organs.

- Just the female character(s) wears clothes that reveal sexualized areas of the human body such as cleavage, upper tights, lower belly, or buttstock, while her male counterparts don’t.

**Trophy** - the character is a trophy for a main character (1) or not (0). The only purpose of the character in the game is to be ‘won’ as a prize for a male character. The character is objectified.

**Damsel in distress** - the character has the role of a damsel in distress (1) or not (0). The character in-game exists to be rescued or saved by the protagonist or the main character. If the character has this role but their age is that of a child or an infant this category is considered as negative, as a child needing the protection or help of an adult is not considered a sexist role.

**Sexualized cut scenes** - some or all the cut scenes where this character appears are sexualized by portraying them in a suggestive way or unnecessarily focusing on areas of the body that are usually considered sexually appealing such as cleavage, breasts, pubic area, and buttstock.

**Research question(s) – the intention behind the analysis.**

* **Example: Are there gender differences in income? Does number of bedrooms impact home prices? Are there differences between states when looking at soda preferences?**

**Extra Credit Opportunities**

* **Using software other than Excel for analysis (R, Tableau, SAS, Minitab, etc.). For extra credit, you must submit code, files, etc.**
* **Regression**