**Final Project Proposal**

The purpose of the final project is to:

* give you a chance to research a topic that is of personal interest to you
* allow you to process the data for this project using the Python skills you have developed during the course

This proposal will allow you to begin making plans for your final project. You will probably make some adjustments to this plan as you become more familiar with your project. That’s fine. However, you should still adhere to the overall requirements presented below..

**Part 1: Exploration**

1. What is the specific question or questions you are trying to answer through this data wrangling project?

**The specific question(s) I am trying to answer through data wrangling is: What are the top 10 games for each system, Do the top 10 games by sales match the top 10 games by critic review, what are the top 10 games by country, top 5 games per year, Do the video games with the top user rating match the top games based on ESRB rating, top 3 overall video game genres, and is there a correlation between the top video game genre and overall sales?**

1. What data sources will you use to answer these questions (please identify at least two data sources).

**The data sources I will be using will both come from Kaggle.com. The links to the data sources are provided down below.**

[**https://www.kaggle.com/gregorut/videogamesales**](https://www.kaggle.com/gregorut/videogamesales)

[**https://www.kaggle.com/rush4ratio/video-game-sales-with-ratings**](https://www.kaggle.com/rush4ratio/video-game-sales-with-ratings)

1. Please explain why you chose those data sources and how they will help you answer your research questions

**The reason I choose these data sources is because both of these data sources contain 5 or more years of data. In addition, to carrying the list of video games from 5 years or more, these data sets contain: ratings from overall sales, sales by country, score by critics, user score, the number of users who scored, and a rating from ESRB. These data sets allow me to compare how video games performed not just by sales, but also by critic, company, and consumer review.**

**Part 2: Analysis**

1. Please identify statistical or informational techniques that you will use to process the data. (You must use at least three of the following).
   1. correlation statistics
   2. grouping into categories
   3. simple counts
   4. trends over time
   5. boxplots
   6. normal curve comparison
   7. simple averages comparison

**The statistical and informational techniques I will use to process my data are: correlation statistics, trends over time, and grouping into categories**

1. Please identify why you are choosing each of those approaches and how they will help you to answer your project questions.

**I am choosing these approaches to use during my project because when analyzing this data I am trying to approach it like a company executive or marketing analyst would. When developing and creating a video game, it is important to decide which projects need to get more of the budget. This can be determined by looking at trends over various years. In addition to looking at the various trends, data becomes easier to understand if it is broken down into categories. The main categories I will be using to separate my data is either by: year, system, or genre. Finally, trying to find some form of a correlation between these variables can also have a great impact as to what video game projects should be pushed forward, and what projects should be discontinued.**

**Part 3: Presentation**

1. Please identify the presentation approach you will use (select one):
   1. Creating a Django website
   2. Creating a Powerpoint deck
   3. Creating an Adobe Spark Page
   4. Creating a Jupyter notebook

**The presentation approach I will be using is a Jupyter notebook.**

1. Why did you choose this presentation approach over the others?

**I chose to use a Jupyter notebook instead of the other methods, because of my experience with Jupyter notebooks. While I have only been using Jupyter notebooks for a short while, I feel like I am more confident, comfortable, and familiar with using this program instead of the other presentation approaches. Jupyter notebook allows me to not only process my data, but it also allows me to quickly make notes and connections I have found in my data quickly, and within the same window. I do not want to try and experiment with these other presentation approaches just because it will add more stress to me while I work on my project.**