Final Project Plan of Action Form

# What are you trying to achieve & learn with this project?

I’m trying to make a system that will be helpful on the business side of my capstone project that can find stats for video game sales. I’m trying to learn how to do some data science as I believe it’ll be helpful to me in the future

# Where is the official documentation for the libraries you will be learning?

<https://pandas.pydata.org/docs/>

# What tutorials, websites, books and resources will you utilize?

1. [Pandas Main website](https://pandas.pydata.org/)
2. <https://medium.com/analytics-vidhya/statistical-analysis-in-python-using-pandas-27c6a4209de2>
3. <https://realpython.com/pandas-python-explore-dataset/>

# What starter project or tutorial will you complete for Week 5?

<https://www.kaggle.com/learn/pandas>

I’ve gone through 4 out of the 6 exercises so far

<https://www.kaggle.com/code/samdansereau/notebook4516885103/edit>

<https://www.kaggle.com/code/samdansereau/exercise-indexing-selecting-assigning/edit>

<https://www.kaggle.com/code/samdansereau/exercise-summary-functions-and-maps/edit>

<https://www.kaggle.com/code/samdansereau/exercise-grouping-and-sorting/edit>

# List the major classes that exist in your library

.DataFrame(), .Series(), The summary functions and the indexing Functions also are quite important

# Identify one of the most important classes, for this class identify:

1. DataFrame is the most important and making a dataframe is easy: pd.DataFrame({'Apples': [30], 'Bananas': [21]}), you could also do things like change the index or where data is placed in the dataframe
2. So this class stores a data table of whatever data you put in it whether that’s manually or with a table
3. To\_csv and read\_csv are the two most important at least for what I’m gonna want to do

# What am I going to build for my final project?

An application that can be used to help determine a market for your game based on data from the current/past games market with a possibility of trying to determine the future of the market in certain years as well

# How can I build this in phases? Try to break the project into some smaller parts aiming ideally for 3 phases (you should have at least 3 phases and not more than 6). What will you accomplish in each of these phases?

Phase 1: Start building out a base and find an initial database for sales data

Phase 2: Implement a lot more databases so that I just have more data to run through, and try to help make decisions on your game based on these databases using ui for stuff like inputting genres, multiplayer, features so on and so forth.

Phase 3: Finalize the project with comparison methods of multiple data frames for example one dataframe is a list of games that released a remaster and the another that has all time sales of all games

Phase 4: If I have time I want to try and implement something that determines the future of the market and if certain games would be profitable based on predictions or steam wishlist data. I do think this is ultimately where I want the project to be when I’m finished but will only do it if I have time for it