Project Proposal for STAT 4185

Title: Comparative Analysis of Gaming Console Prices and Global Sales

Objective Statement:

This project aims to conduct a comprehensive analysis of various gaming consoles, focusing on their prices and global sales figures. The goal is to understand the relationship between the pricing strategies of different consoles and their success in the global market. Insights from this analysis could provide valuable information for future pricing and marketing strategies in the gaming industry.

Data Collection:

Data will be sourced from reliable gaming and sales websites, such as VGChartz and official console websites. The dataset will include console names, their prices, and global sales figures. Efforts will be made to ensure that the dataset is comprehensive, covering a wide range of consoles from various manufacturers over a significant timespan.

Data Cleaning:

Data cleaning and preprocessing will be performed using Python's pandas library. This process will include handling missing values, normalizing data formats, and verifying the accuracy and reliability of the data.

Data Visualization:

Visualizations will be created using Python's matplotlib and seaborn libraries. These visualizations will include scatter plots, bar charts, and line graphs to depict the relationship between console prices and global sales, trends over time, and comparisons between different console brands. All visualizations will be clearly labeled and annotated for ease of understanding.

Data Modeling and Analysis:

The project will employ statistical models and machine learning algorithms, such as linear regression, decision trees, and hyperparameter tuning techniques like GridSearchCV and RandomizedSearchCV. The aim is to model and predict sales figures based on console prices and to uncover underlying patterns and trends in the data.

Submission and Documentation:

The project will be thoroughly documented in a series of Jupyter notebooks, which will include detailed code, explanations, and interpretations of the analyses and visualizations. A final report summarizing the findings, insights, and conclusions will be provided. All project materials will be hosted on a GitHub repository.

GitHub Repository Link: https://github.com/DarkMunchkin/Stat4185.git

This proposal aligns with the course guidelines and aims to leverage data science methodologies to yield meaningful insights into the gaming console market, emphasizing the analysis of pricing strategies and global sales. It demonstrates a practical application of data collection, cleaning, visualization, modeling, and analysis techniques.