Video Game Sales Exploratory Data Analysis

Abstract: Who doesn't like video games? Video games have become a favorite pastime for most of us. Video games are a billion-dollar business and have been for many years. Now imagine being able to analyze the sales trends for the popular Nintendo games and many more such platforms. How cool would it be!

Design: This project is one of the T5 Data Science Boot Camp requirements. Data provided by Kaggle & data. world have been used in this project. In this module, we will be laying the foundation for our analysis by processing and exploring a large amount of data on video game sales. The dataset contains information regarding the sales of video games across various regions like North America, Europe, Japan and globally, while also giving information regarding the Names, Publishers and Platforms. This dataset has been made available thanks to Kaggle which is the home for many such datasets and competitions. Get the data here.

Understanding the dataset:

This dataset contains a list of video games with sales greater than 100,000 copies. In this data set there are 14 columns. Their names and data types as follows: Fields include

Column	Description
Rank	The overall rank of the game
Name	The name of the game
Year	The year in which the game was released.
User Review	The review of game by user
Platform	The platform on which the game was released
Genre	The genre of the game
Publisher	The publisher of the game
Meta Score	The score from 1 to 100 by the review aggregator Metacritic.
Release Date	The date of game released
NA Sales, JP Sales, EU Sales , Other Sales, Global Sales.	Sales made by a particular game in North America, Europe, Japan, Other regions and Globally (in millions).

Algorithms

1- Use panda's library to analyze different features of the dataset, which includes,
2- Read the dataset.
3- Cleaning the data and remove null values
4- Remove the outlier from dataset
5- Analyze the sales for different regions with respect to various features.
6- Use seaborn and matplotlib library to visualize the given results.
7- Plot graphs like bar graphs and pie charts

Tools

•Pandas for data manipulation

• IQR for discover outliers

• Remove Duplicate or unnecessary data

• Matplotlib for plotting

• Seaborn for visualizations

Communication

• The slides will be provided here, Feel free to any pull requests besides details are provided at the readme of the project