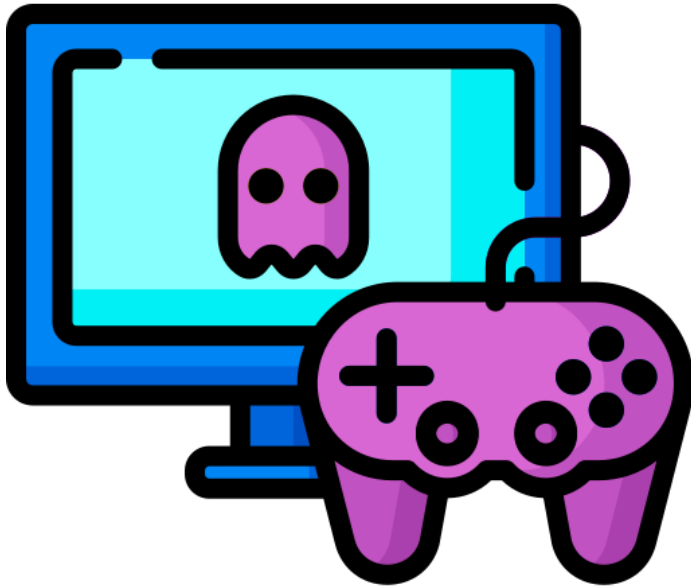


Video Game Sales Data Analysis



Video games are a hugely popular pastime and hobby for many. Especially over the past 20 years, game sales have increased as more and more diverse games have come on the market.

For this project, you will be given a dataset on video game sales called *Video_Games_Sales_as_at_22_Dec_2016.csv*.

Phase 1 (Data Cleaning Phase)

1. Import required libraries
2. Read the csv file
3. Show first five results
4. Show if columns have any NaN,N/A or null values in boolean form
5. Check how many null,empty,N/A values are there in each column
6. Show null values of each column with the help of Heatmap
7. Drop outliers (columns which are not useful according to you)
8. Drop the null values containing rows from the object which you created in point 7
9. Check the dataframe again if there is anything which you want to remove more.
10. Export dataframe to your desired location with no index column

Phase 2 (EDL - Data Analysis Phase)

1. Open new Jupyter notebook
2. Import required libraries
3. Read your cleaned csv which you did in phase 1
4. Shape the data
5. Show first five records
6. Show all columns of the dataframe
7. Show last 5 five rows
8. Describe the data.
9. Group the sales data of each region on Year_of_Release and plot the same (It should show which region sales was at peak on which year of release)
10. Group North American, Japan, European & Other region sales data as per Genre and plot it (It should show highest of region according to game category , ex- NA_Sales is highest in Adventure category , plot kind should be bar)
11. Group North American, Japan, European & Other region sales data as per Platform and plot it (Should show sales as per "Platform genre" , plot kind should be bar)
12. Plot the Sales in all regions as per Platform , you need to create a boxplot for that which will show the distribution of sales as per platform
13. Find out which Genre games are sold in highest numbers (create a countplot of seaborn which could do that)
14. Try to find out global sales as per Genre ,(create a barplot of seaborn for that)

*** Comments are super important, don't forget to explain what you did in your commands and describe the charts as much as you can

Bonus: Try to clean this data set using other ways too. Extract some more information about the video game sales. Feel free to go beyond the scope of this project. Show some more in-depth data using some good plots.

Due Date

You have **1 week** to complete this project, after which you will present your project.