

# Project 4

## Play Video Games But In A Different Format

This time we have to work on a different dataset called as Video-Game-Sales. As usual this dataset has some requirements too, but in two phases.

### 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 that what you did in which command and describe charts as much as you can.

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

=====

by Cognixia®