Web Scraping and EDA Project

This project helps you to apply python, numpy, pandas, matplotlib, regular expressions, web scraping and EDA skill sets. This project will help you in business understanding, data cleaning and data visualization in a life cycle of machine learning projects.

After the completion of the project add it to your profiles such as LinkedIn, GitHub and CV/Resume which gives more skill volume.

Steps Involved in this project:

1. Problem Statement
- *Define the Domain problem statement*

2. Search for Websites
- *Find the websites which are allow to given access*

3. Extract the data
- *Confirm the variables (columns) which are useful to your problem statement. To do the better analysis and draw the more conclusions about your business problem it would be suggestable to scrap the minimum number of variables (columns) are 8, minimum number of rows are 400*

4. Create a Data Frame
- *Convert the scraping variable observations into Data Frame*

5. Export into .csv format
- *Export the data frame into .csv format*

6. Read csv file
- *After exporting data frame you need to import the data frame to perform the various steps which are given below*

7. Clean the data

In this section you have to clean the data like:
- *removing the special characters,*
- *imputing the null values,*

- *converting the data types*

8. Data Analysis

- *Uni-variate Analysis*
- *Bi-variate Analysis*
- *Multi-variate Analysis using Group by concept and describing the statistical analysis*

9. Data Visualization (Uni, Bi and Multi-variate)

Plot the every variable

- *Plots for Numerical and Categorical variables*
- *Plots for Numerical to Categorical variable*
- *Plots for Categorical to Numerical variables*

[Box-plot, Bar-plot, count plot, Pie chart, Scatter-plot, violin-plot, Distribution-plot, Heat map, Histogram and kde-plot etc., Use all plots for individual variables]

10. Interpretations

- *Write the interpretations for Data Analysis and Data Visualizations*

11. Conclusion

- *At last give the interpretations for what you infer about your problem statement*

"Based on the project completion you have to give the presentations about what you infer from your business problem. We will Schedule a presentation once the Project is submitted. Please Contact Our Mentors for the Presentation Schedule.

Submission:

After completion of the project Zip the web scraping code file, csv file and data analysis and data visualization file upload the zip file with your name and batch number.

NOTE:

For the doubts clarifications you are welcome to join our One to One mentorship session