Project Title: Web Scraping

Deadline: Lab exam Week (in lab class)

Project Overview:

~~You have to collect data from a website (you decide which website) and then store the date into a csv file. Make sure you are collecting data which have multiple variables (at least four) and a good number of records (rows) to build a decent dataset.~~

Now, you have to apply the pre-processing techniques to prepare the dataset for data analysis.

To prepare a cleaned dataset, perform the following tasks of data pre-processing using R language:

1. Data cleaning:
    a. Smooth Noisy Data
    b. Handling Missing Data
    c. Data Wrangling or Munging
2. Data Integration
3. Data Transformation
4. Data Reduction
5. Data Discretization

Next, you have to describe the data by applying descriptive statistics.

Finally, you have to draw graphs/plot to visualize the data. Make sure you explain each of the graphs you mentioned in the report.

Note that you must mention the corresponding R code where necessary in your project report.

Project Report:

You must submit a complete report mentioning every step of the web scraping project in detail.

The project report should contain the following information:

- A standard cover-page
- Project overview: Write details about the project
- Project solution design: Write how do you design the solution to complete this project
- Data Collection via Web Scraping: Write how do you collect the data via web scraping along with your necessary R code.
- Data pre-processing: Detailed description of each tasks of data pre-processing to prepare a cleaned dataset along with R code for the project with output (attach screenshot of the output where necessary with description).
- Descripting statistics: Describe your dataset using descriptive statistics (with R code).
- Data visualization: Draw the required graphs/plot based on your dataset to visualize the data with explanation (also mention the R code used for data visualization).
- Discussion and Conclusion: Discuss about the project outcomes and draw your concluding remarks about the project.

Remember that you must submit printed version of your project report and face the viva to complete the project.