# Lab Exercise Questions - Pandas

1. Create a Pandas DataFrame named `df1` of dimension $10 \times 5$ whose entries are normally distributed random numbers. Named the columns as *'col1, col2, col3, col4, col5'* and rows as *row1* to *row10*.
2. Find the `shape, size` and `description` of the DataFrame `df`1.
3. Also find the `mean, sum, cumsum` along columns and rows of `df1`.
4. Create another Pandas DataFrame named `df2` of dimension $5 \times 5$ whose entries are random integers between $10$ and $50$. Named the columns as *col1, col2, col3, col4, col5* and rows as *row0, row2, row4, row6, row8*.
5. Create a DataFrame `df3` by appending `df2` to `df1`.
6. Describe the DataFrame `df3` and then create (i) a view by dropping the rows *row0, row8* (ii) a view by dropping the rows *col3*
7. Using the appropriate function techniques, add a column named *col6* to the DataFrame `df3` which contains the values $3 \times \sqrt{|col3|} + (col5)^2 + 10$.
8. Store the DataFrame `df3` into a file named `data.csv` by using the command `df3.to_csv('data.csv')`.
9. Apply the function $|min(x) - max(x)|$ to DataFrame `df3` where $x$ is the rows of `df3` and add these values as *col7* and store the resulting DataFrame as `df4`. Further add a column named *col8* which holds the double the values of *col7*.
10. Add a new row to this `df4` whose entries are nothing but the $mean(x)$ where $x$ is the columns of the DataFrame `df4`.
11. Store the DataFrame `df4` into a file named `myData.xlsx` by using the command `df4.to_excel('myData.xlsx')`.
12. Read the data in the *myData.xlsx* into a DataFrame named `Data` and perform the following plotting operations.
    - Plot the columns as a line plot.
    - Plot the columns as area plot with each column in a subplot and set the figure size to (6, 18)
    - Plot the column *col7* as a bar plot.
    - Plot the scatter plot of *col3* vs *col6*
    - Plot the box plot of *col2*
    - Plot the second row of the `Data` as a box plot.