Mandatory assignment 1

a) Use the Grocery_dataset.csv file provided to answer questions in this assignment

b) Use Code Cell in Jupyter Notebook to write your codes.

c) Work in groups 2-4 and submit the Jupyter Notebook in .ipynb format on Canvas.

1. **Read the CSV file in Pandas and create a DataFrame named Grc_df. What is the number of rows and columns in Grc_df? Print the first 10 and last 10 rows of Grc_df.**

2. **Are there any null values in the Grc_df? If yes, then in which columns and how many? Finally, handle these null values using any strategy shown during the labs.**

3. **How many unique Outlet Sizes are there in the Grc_df? Which outlet size is maximum, and which is minimum?**

4. **How many unique Item Fat Content types are in the Grc_df? List them. Do you see any issues with the Item Fat Content types? If yes, then handle this issue.**

5. **Drop the columns having index values of 0, 6 and create a new DataFrame Grc_new_df.**

6. **Using different Supermarket type listed in the column Outlet_Type create two different DataFrames from Grc_new_df. Name these DataFrames as SupType_1 and SupType_2.**

7. **Using Seaborn ("ggplot style") create a (2,1) subplot of a box plot showing 5-point summary of the column Item_MRP for SupType_1 and SupType_2. Which Outlet Type has a higher median MRP? Are there any outliers?**

8. **Concatenate the DataFrames SupType_1 and SupType_2 to create a new DataFrame Grc_Concat_df and sort it in ascending order based on the column Item_Outlet_Sales. What is the Outlet_Location_type of the store with the lowest sales?**

9. **Using columns Outlet_Size and Outlet_Location_Type to create a multiple index for Grc_Concat_df.**

10. **Cut the column Item_Weight of Grc_Concat_df into 10 buckets and compute each bucket's mean, minimum, maximum, and count.**