Task Report: Data Visualization

# Introduction

The objective of the first task is to create a **chart**, either a **histogram** or a bar **chart**, to visualize the distribution of a variable in a dataset. This variable can be either categorical (such as gender: male/female) or continuous (such as the age of a population).

# Tools Used

I chose the Kaggle platform to generate the dataset and selected the 'Salary Prediction dataset.' I used the following tools: Jupyter and the Python programming language to complete the first task.

# Steps to Accomplish the Task

* Import the necessary libraries for data analysis and visualization.
* Read the CSV file of the 'SalaryData' dataset.
* Generate the header of the DataFrame named Salary\_data.
* Generate descriptive statistics for the Salary\_data DataFrame.
* Display summary information about the Salary\_data DataFrame, including data types and non-null values.
* Display the data types of each column in the Salary\_data DataFrame.
* Display the 'Age' column of the Salary\_data DataFrame.
* Count the number of missing (null) values in the 'Age' column of the Salary\_data DataFrame.
* Display all rows in the Salary\_data DataFrame where the 'Age' column has missing (null) values.
* Create a copy of the 'Age' column from the Salary\_data DataFrame and store it in the variable Copy\_Data.
* Remove any missing (null) values from the Copy\_Data series.
* Plot a histogram of the Copy\_Data series with 30 bins.