Individual Assignment 1

# **Descriptive Statistics**

1. **Dataset**

The dataset contains the information of employees including Employee ID (EEID), Full Name, Job Title, Department, Business Unit, Gender, Ethnicity, Age, Hire Date, Annual Salary, Bonus %, Country, City, Exit Date.

1. **Required Analysis**

Using the descriptive analytics tools you learned, answer the following questions.

* 1. Describe the data types for the following columns: **Gender**, **Age**, **Hire Date**, **City**.
  2. What are the 3 most repeated **Job Title** and **Business Unit** and the number of repetitions?
  3. Provide **pie chart**(s) for **Gender** and **Ethnicity** distributions. (Hint: you can use the pivot table)
  4. Which **Ethnicity** has the highest ratio of Female-to-Male employees? (Hint: you can use the pivot table)
  5. What is the average **Annual** **Salary** for each **Country**? (Hint: you can use the pivot table)
  6. What is the average Age for each Ethnicity? (Hint: you can use the pivot table)
  7. Calculate [**mean**, **mean**, **median**, **standard deviation**, **range]** for **Annual Salary** column.
  8. Provide a **histogram** plot of **Age** column.
  9. Provide a **scatter plot** the **Annual** **Salary** versus **Bonus %.** Calculate the **correlation** between these two columns and provide a discussion based on the calculated correlation and the scatter plot.
  10. Categorize the **Annual** **Salary** column into the bins of $40,000 width and plot the result.

1. **Methodology**

* Mention the assumptions/techniques/calculations used to analyze the data (when necessary).
* Tabulate the calculations for each question to simplify the understanding of the results.
* Provide necessary explanations/discussions (when necessary).

1. **Grading Rubric**

|  |  |
| --- | --- |
| **Key Points** | **Grade Allocation (%)** |
| Format (font type, size, table, formulas) | 20 |
| Results, analysis and assumptions | 60 |
| Overall content, including references if required (APA Style) | 20 |

N.B. Failure to comply with the above would result in low grades.

1. **Word Limit**

**600 words** (excluding tables, and appendix)

Submission: a word document

Deadline: 23:59, Jan 27, 2023