**Employee Data set Report**

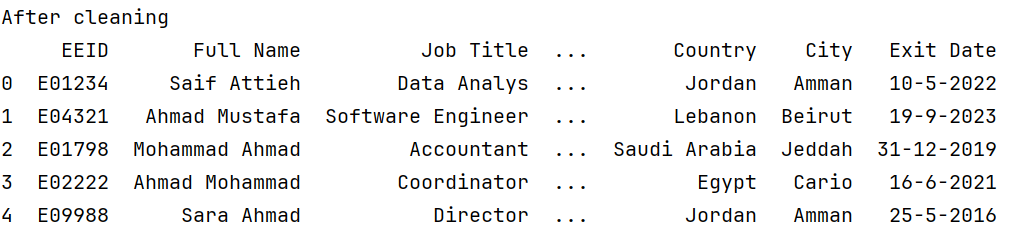
**Saif Attieh**

**Overview**

In this data set we have some information about the employees in one of the companies of the world and the data set contains the full name of each employee, job title and in which department he\she works, gender, age, hire date and exit date, Annual salary and bonus %, country and city.

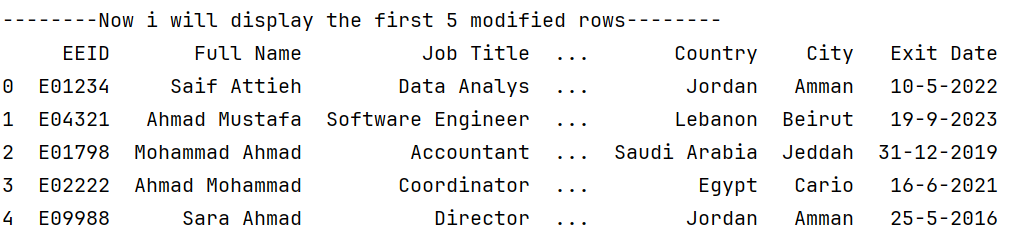
**Data Cleaning**

Data cleaning means cleaning wrong data format and cleaning empty cells and removing the duplicates and as you can see the result after cleaning in the image below I display the first 5 rows without any duplicates and wrong data format.



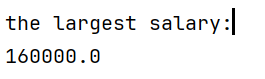
**Changing the five rows**

for changing the first five row we should create a dictionary and assign new values for each column and after finishing from dictionary we should put the the new inputs in data frame. Also, this is the result after I modified the first 5 row



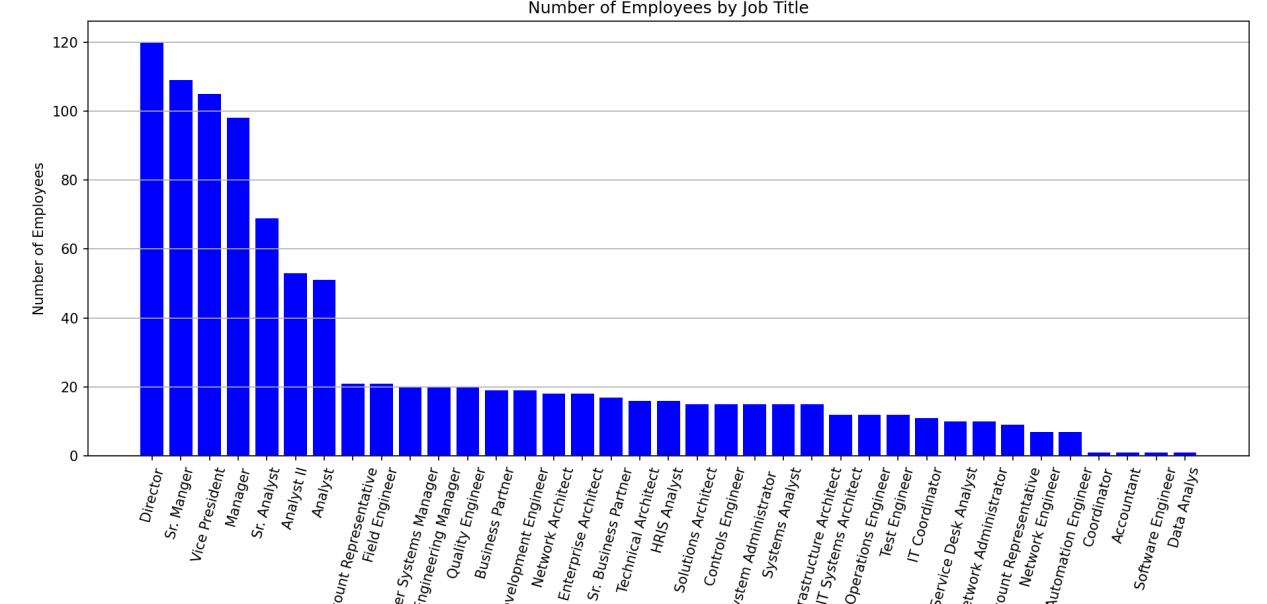
**Largest salary row**

To find the maximum annual salary in the organization we should use max() function and the maximum annual salary is 160000.0 and this indicates that the employee has many skills and qualifications.



**Number of employees by the job title**

To display the number of employees by job title we should use bar chart and this bar chart below. Also, I found that the highest number of employees are from directors (120 employee) and the lowest number of employees are from data analyst (~5 employee) and because the organization have the highest number of employees from directors this depends on several factors such as organization structure, strategy and industry norms.



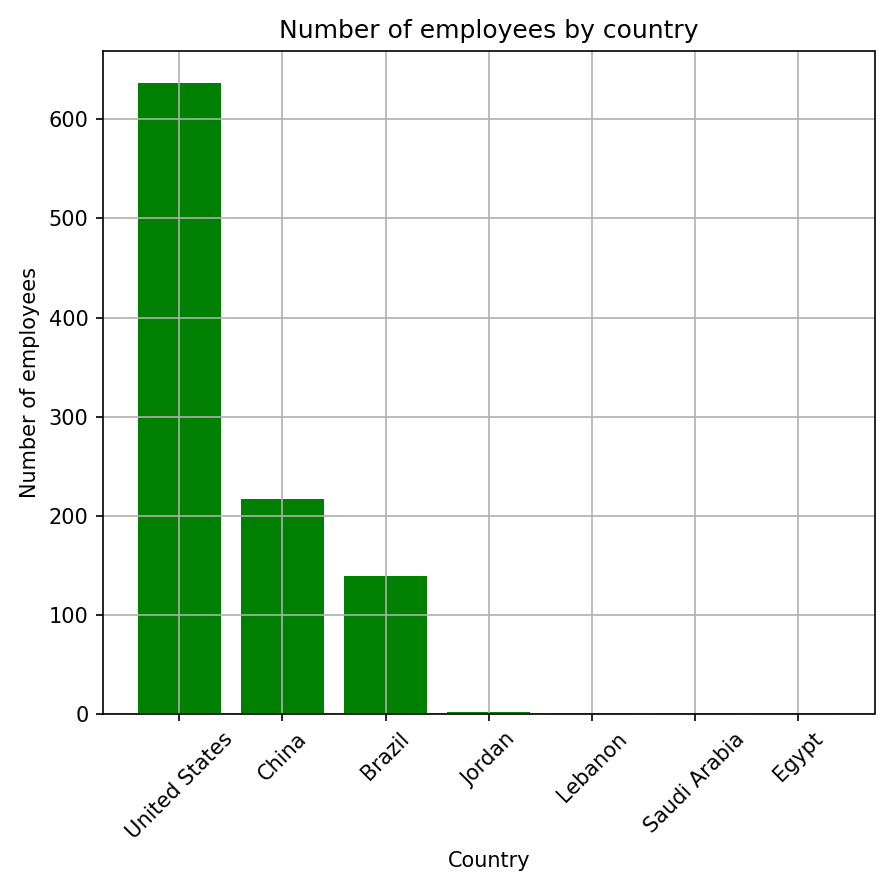
**Number of male and female employees**

Counting the number of males and females within an organization serves several important purposes like workforce diversity and inclusion , compliance and reporting. Moreover, as in this figure demonstrate that number of females are greater than males and number of females are approximately 510 employee and the number of males are approximately 490 employee.



**Number of employees in each country**

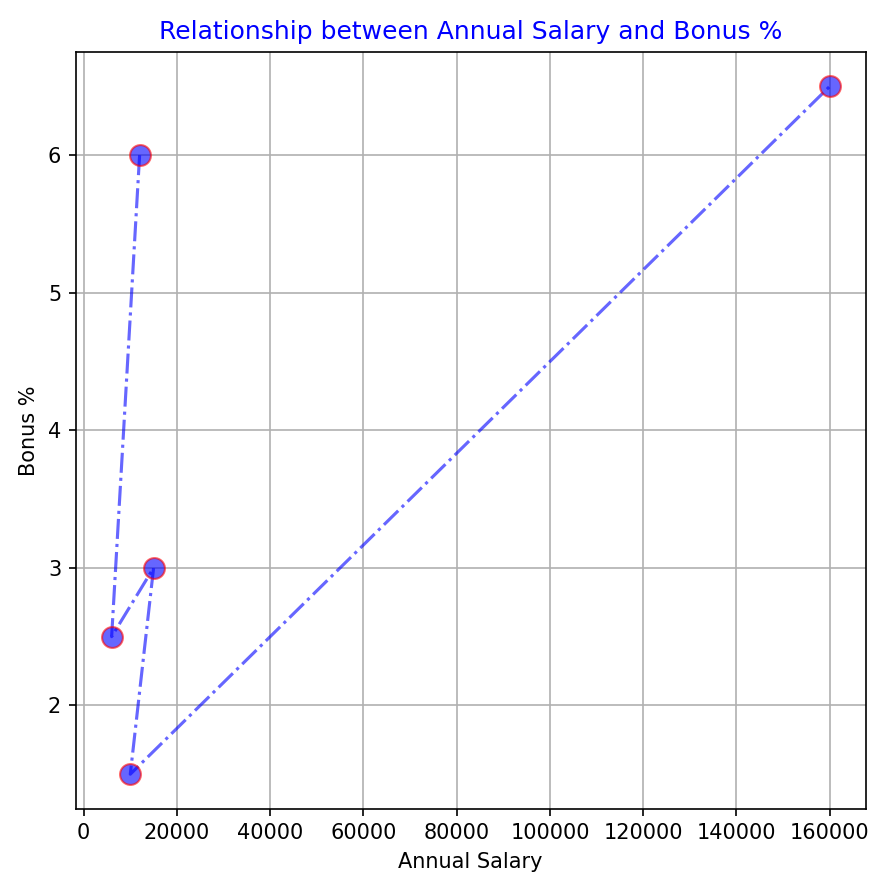
To determine the number of employees in each country we should count each employee in each country and after counting I present the result on bar chart.



After plotting the graph is saw that the highest number of employees are from United States and this indicates that the company’s head quarters in United States . Also, this depends on that United States large market size and operations, Industry dominance and cultural and language factors.

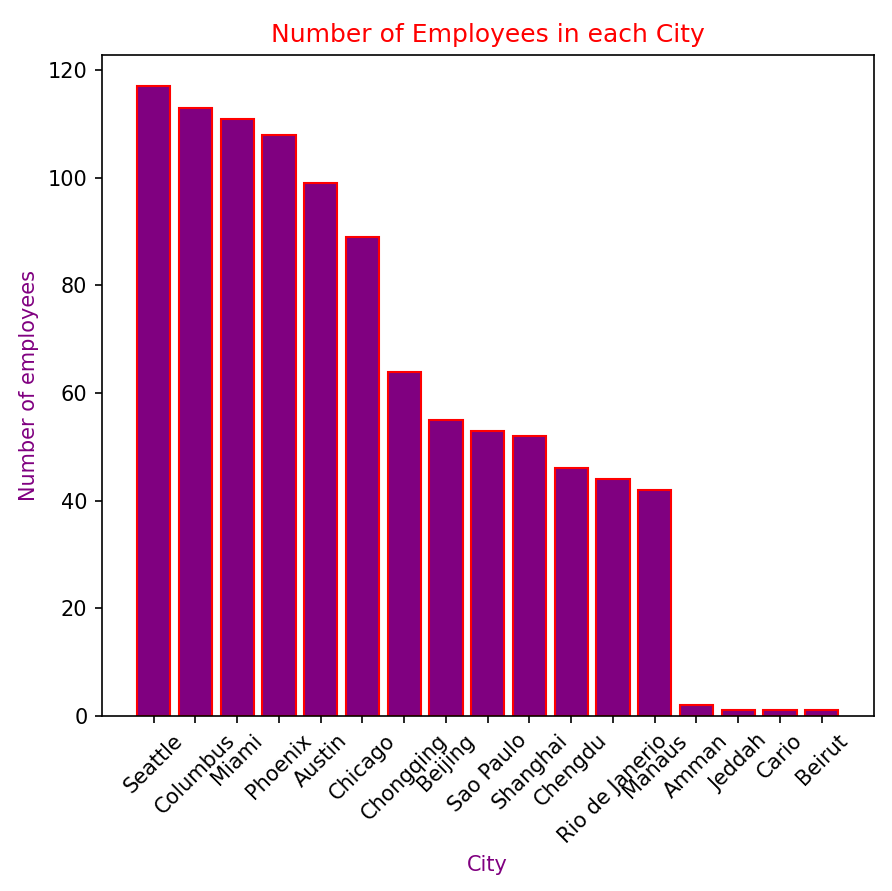
**Relationship between annual salary and bonus**

The graph below depicts the relationship between annual salary and bonus and it shows a generally positive trend indicating that as the annual salary increases, the bonus percentage also to increase. Also, there are some fluctuations at the lower salary levels and this suggests that higher salaries are often associated with higher bonus percentages.



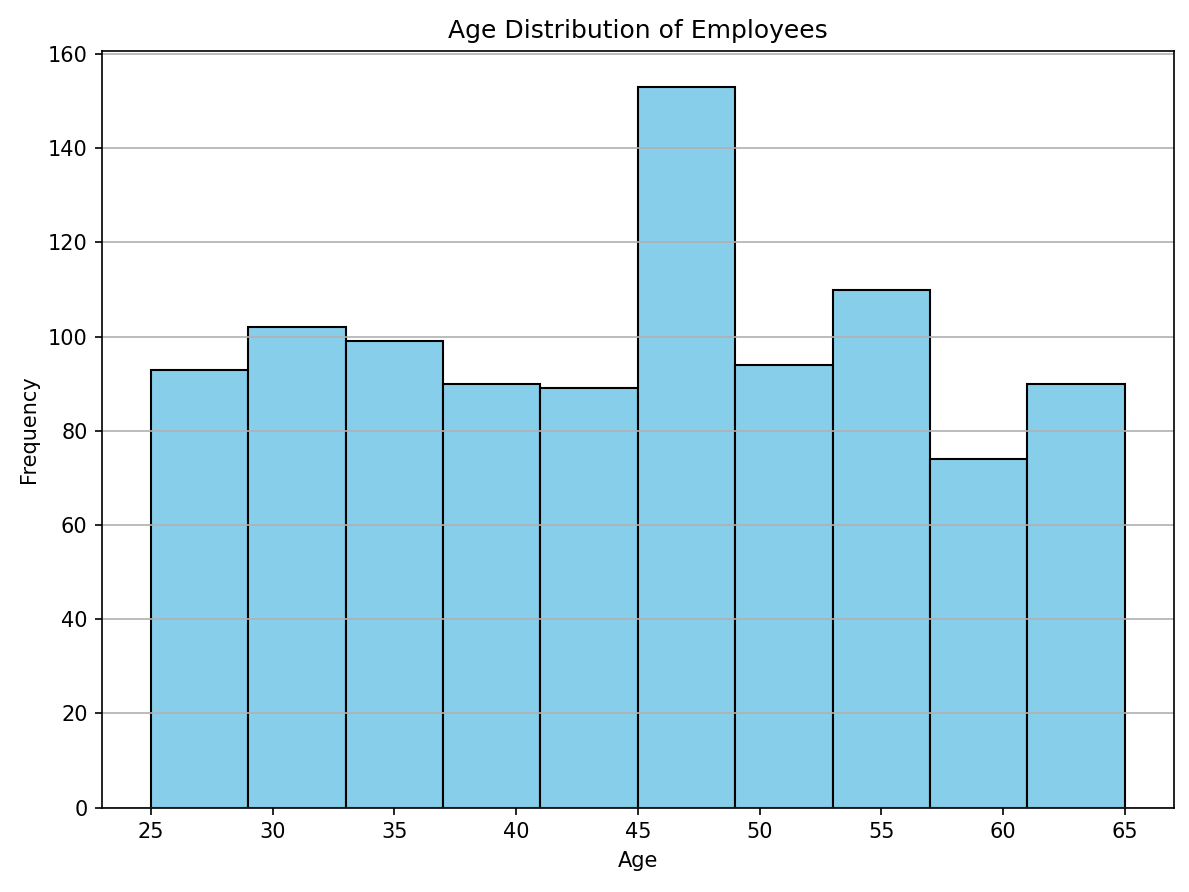
**Number of employees in each city**

As in the figure below we have 17 city and the highest number of employees is from Seattle which locates in United States and the number approximately is 120 employee. Also, as the figure demonstrate that most of cities are from United States and I explained why United States include many employees in number of employees of each country.



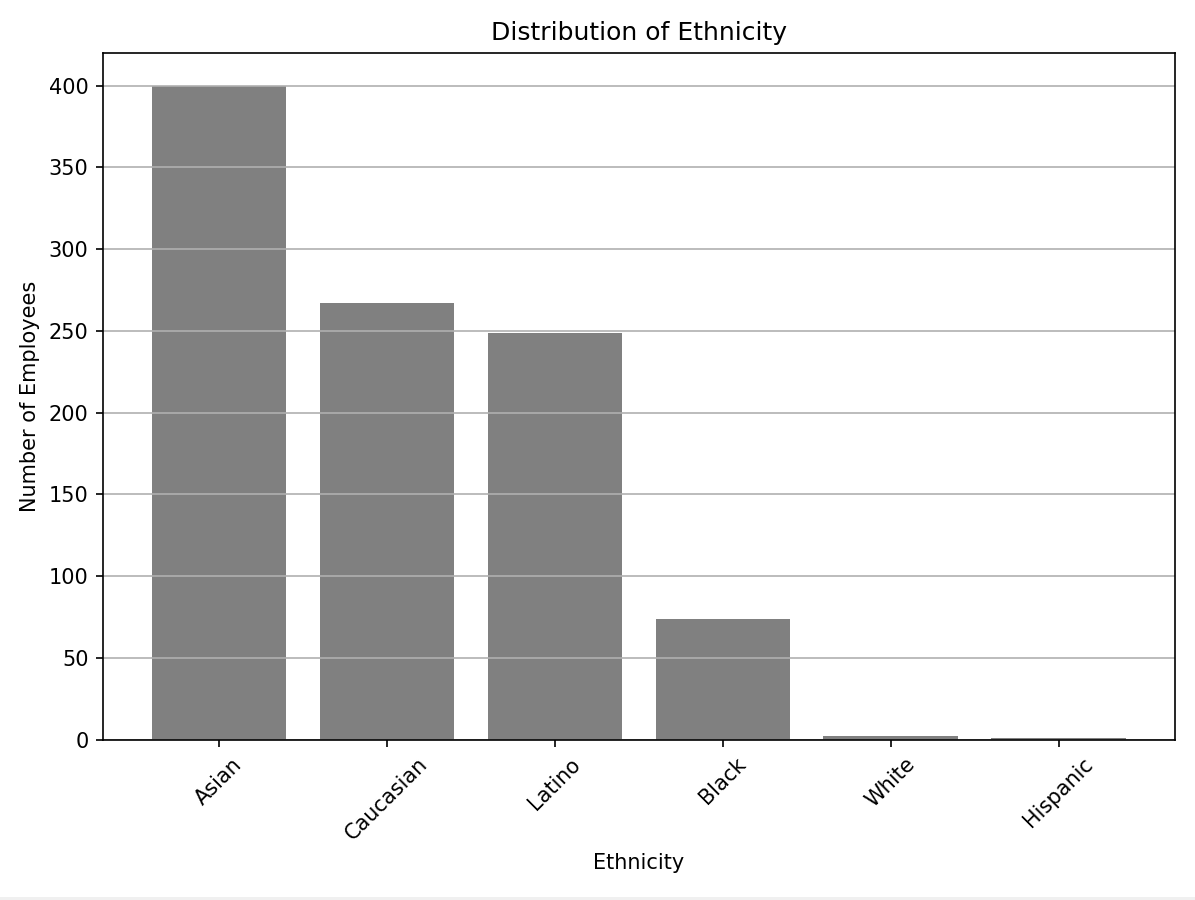
**Age Distribution**

To know the distribution of ages, a histogram must be used. If we look at the figure below, the majority of employees are between 45 and 50 years old, and this indicates that their experience and knowledge are great because they have broad skills in their field of work.



**Distribution of ethnicity**

In this figure each bar indicates the number of employees belonging to each ethnic group. The highest of the bars suggest the relative representation of each group among the employees: The “Asian” category has the highest number of employees, approaching 400 and “Caucasian” and “Latino” categories have a similar number of employees, both around 250. Also, The “Black” category has fewer employees, around 150. Moreover, “White” and “Hispanic” categories seem to have significantly fewer employees compared to the others.

****