**PROJECT 1: Ask A Manager Salary Survey Analysis**

**Introduction:** The Ask A Manager Salary Survey Analysis which includes age group, gender, industry, years of experience, annual salary, race type of the staff was used to provides insight into informed decision making for the organization.

**Objectives:**

The aim of this analysis is to gain insights into various salary-related factors, provide valuable information to professionals, job seekers, and employers about industry salary trends, experience-based salary growth, geographical variations, and demographic correlations.

**Dataset Overview:**

Source: [**https://docs.google.com/spreadsheets/d/1vB5eAhH\_93EKUyu-AQHQ68S3ULsEMO3zqkEf1LmIuoU/edit#gid=1625408792**](https://docs.google.com/spreadsheets/d/1vB5eAhH_93EKUyu-AQHQ68S3ULsEMO3zqkEf1LmIuoU/edit#gid=1625408792)

There were 27996 rows and 17 columns before cleaning

Initial data cleaning is required due to the presence of free-form text entries.

**Data cleaning:**

This dataset was extremely dirty and so the Data cleaning was done in excel and power query. Detailed steps are below:

1. Renamed all the columns to the appropriate names eg **What country do you work in? was renamed to “Country”**
2. Changed all the countries to a standard name eg **United States, America etc was renamed to US**
3. Values in Industry column was modified to a general name eg **Education (Higher Education) was modified to Education, Everything under academic and academia was modified to academia**
4. In the gender column, I modified all the **“Other or prefer not to answer” to “Prefer not to answer”**
5. I modified the race column to “**Race Type**” and named all those that belong to one race “**monoracial”**, those that belong to two races, “**biracial”** and those that belong to two or more “**multiracial”**
6. I also put N/A in all the blanks and removed duplicated
7. At the end of the data cleaning I had 27832 rows and 15 columns

**Analysis Focus Areas/Questions:**

Which Industry Pays the Most?

* Analyze and compare average salaries across different industries.
* Visualize industry-wise salary distributions.

Salary Growth with Years of Experience:

* Explore how salaries change with increasing years of experience.
* Identify trends in salary growth over the years.

Geographical Salary Variations:

* Compare salaries for the same role in different locations.
* Visualize regional salary differences.

Gender and Experience Impact on Salaries:

* Investigate how salaries differ based on gender and years of experience.
* Examine potential gender-based salary gaps.

Correlation of Race and Education with Salary:

* Assess the correlation between race, education level, and salary.
* Identify any patterns or disparities.

Total Work Experience vs. Field-Specific Experience:

* Determine if there's a "sweet spot" in the balance between total work experience and years in a specific field.
* Explore how this balance correlates with salary.

**Analysis and Findings/Insights:**

* Based on the analysis, IPR industry had the highest average salary.
* Employees between the ages of 21 – 30 had the highest average salary probably because they were more efficient and were promoted more often with pay rise.
* Indonesia had the highest average salary, specifically those in Regional Operations.
* Females had longer years of experience than males.
* Majority of the monoracials had college degree as their highest degree gotten.
* Those which PHD which is the highest level of education, have significantly higher salaries than others and they mostly belong to two races.

**Recommendations:**

Since it was observed that the females stayed longer in the organization, I’ll suggest that the males be given some incentives to encourage them to stay longer.

**Presentation:**

Visualization was done in PowerBi:

Link to the clean dataset: