**Project Description**

1. **Aim of the Project:**

The aim of this project is to analyze and visualize the factors influencing career trajectories across global industries, with a particular focus on salary trends, compensation packages, and the role of experience, education, gender, and geography. By examining these factors in detail, the project seeks to:

* Identify disparities in salary and compensation across different job titles, industries, and locations.
* Understand how professional experience and educational attainment impact career progression and salary levels.
* Explore potential gender-based pay gaps and the distribution of additional monetary compensation.
* Provide a comprehensive, data-driven view of how various factors intersect to shape career growth and earning potential.

Ultimately, the goal is to empower organizations, employees, and policymakers with insights that can inform decisions on compensation strategies, career development, and workforce planning.

1. **Business Problem or Problem Statement:**

Organizations across industries face challenges in understanding and optimizing compensation strategies to ensure equitable pay, attract top talent, and retain experienced professionals. Despite having access to extensive data on employee salaries, many businesses struggle to identify key factors that contribute to variations in earnings, such as gender disparities, education level, geographic location, and years of professional experience.

This project addresses the need for a data-driven approach to analyze career trajectories and compensation patterns across global industries. By examining salary and compensation trends, the business can better understand the factors influencing pay gaps and career growth. The goal is to uncover insights that can help organizations develop more competitive and equitable compensation frameworks, align salaries with industry standards, and make informed decisions about talent management and workforce planning.

Key business questions include:

* How do salary and compensation vary by industry, gender, education level, and location?
* What role does professional experience play in salary growth and job title advancement?
* Are there any significant gender-based pay gaps that need to be addressed?
* How can organizations improve their compensation strategies to attract and retain talent effectively?

By addressing these challenges, the project aims to provide actionable insights that contribute to building a fairer, more competitive, and data-informed approach to employee compensation and career development.

1. **Project Description:**

This project aims to analyze career trajectories and salary trends across global industries using a salary survey dataset. The primary goal is to understand how various factors, such as job title, gender, education, professional experience, and geographic location, influence salary and compensation packages.

The project begins with data cleaning and preprocessing, ensuring the dataset is free from missing values, inconsistent entries, and outliers. After preparing the data, it is imported into MySQL for storage and query execution. Key queries are then run to explore salary patterns across different industries, comparing variables such as gender, education level, experience, and location. These queries focus on metrics such as average salaries, compensation breakdowns, and salary distributions, providing a clear view of how each factor affects earnings.

The analysis culminates in the creation of a dynamic dashboard in Tableau. This dashboard visualizes the key insights derived from the queries, making it easier to identify trends and draw conclusions. The visualizations allow stakeholders to explore salary distributions, gender disparities, and the impact of education and experience on compensation in an interactive manner.

Ultimately, the project seeks to provide businesses with actionable insights into salary structures and career development patterns, helping them make informed decisions about compensation strategies, talent management, and workforce planning. The analysis will also highlight any potential salary disparities, including gender-based pay gaps, and offer a deeper understanding of how industry, location, and professional experience contribute to career growth and earnings.

**Objectives**

The objective of this project is to analyze and understand career trajectories across global industries by exploring key factors that influence salary and compensation. By leveraging a salary survey dataset, the project aims to identify patterns and trends related to:

1. **Salary Disparities**: Investigating how salary varies across different industries, job titles, experience levels, education, gender, and geographic locations.
2. **Gender and Compensation**: Analyzing potential gender-based discrepancies in salaries and additional compensation.
3. **Impact of Education**: Examining how the highest level of education completed correlates with salary and overall career growth.
4. **Experience and Career Growth**: Analyzing how professional experience (both overall and in the specific field) influences salary progression and job titles.
5. **Geographic Influence**: Understanding the variations in salaries based on location (country, state, city) within different industries and job roles.
6. **Comprehensive Overview**: Gaining insights into how multiple factors (such as gender, education, and experience) intersect and influence overall compensation packages.

By creating a data-driven dashboard in Tableau, the project aims to provide actionable insights that can guide professionals, employers, and policymakers in understanding career trajectories, compensation trends, and potential areas for improvement in workforce planning and compensation strategies.

**4 . Functionalities:**

1. **Data Cleaning and Preprocessing**:
   * **Handling Missing Values**: The project includes functions to detect and manage missing data by either filling in gaps with mean/median values or removing rows with missing values.
   * **Standardizing Data Types**: It ensures that each dataset attribute has the correct data type (e.g., numerical values for salaries, categorical values for job titles and industries).
   * **Normalization of Categorical Data**: This function corrects inconsistencies in categorical data (e.g., ensuring uniform capitalization of job titles and industries).
   * **Outlier Detection and Handling**: Identifies extreme values in numerical columns (such as salary) and decides whether to remove or adjust these values based on domain knowledge.
2. **Database Integration (MySQL)**:
   * **Database Creation and Management**: The project includes functionality to create and manage a MySQL database specifically for storing the salary survey data.
   * **Table Design and Data Upload**: A table is created in MySQL, and the cleaned dataset is uploaded using tools like LOAD DATA to populate the database efficiently.
   * **Efficient Query Execution**: The system supports running specific SQL queries to retrieve data based on different criteria (e.g., average salary by industry, salary distribution by education level).
3. **Data Querying and Analysis**:
   * **Salary by Industry and Gender**: A query calculates the average salary within each industry, broken down by gender, to reveal potential gender-based salary gaps.
   * **Total Compensation by Job Title**: A query to sum base salary and additional monetary compensation for each job title, identifying which roles have the highest total compensation.
   * **Salary Distribution by Education Level**: A query shows how salary varies with the highest level of education completed, providing insights into the impact of education on salary.
   * **Employees by Industry and Experience**: This query helps analyze how many employees are in each industry and how their experience levels vary.
   * **Median Salary by Age Range and Gender**: A query that calculates median salary within different age ranges and gender groups, helping to explore salary trends across demographic segments.
   * **Top-Paying Jobs by Country**: A query identifies the highest-paying job titles in different countries, offering insights into geographic salary trends.
   * **Salary by City and Industry**: A query calculates the average salary for each combination of city and industry, helping to identify higher-paying cities for specific industries.
   * **Compensation by Job Title and Experience**: This query calculates the total compensation (salary + additional compensation) based on years of professional experience in specific job titles.
   * **Salary Analysis by Industry, Gender, and Education**: A complex query that compares salary variations across multiple factors (industry, gender, education level).
4. **Data Export and Visualization Preparation**:
   * **Export Query Results to CSV**: After running the queries, the results are saved as CSV files to be used in the Tableau dashboard for further visualization.
   * **Tableau Data Import**: The exported CSV files are then imported into Tableau for the creation of dynamic visualizations and dashboards.
5. **Dashboard Creation in Tableau**:
   * **Salary Distribution Visualizations**: Interactive charts that showcase how salaries are distributed across various industries, job titles, and education levels.
   * **Gender Pay Gap Insights**: Visualizations comparing salaries for men and women within industries, highlighting any gender-based pay discrepancies.
   * **Experience and Compensation**: A series of visualizations showing how years of professional experience influence total compensation and career advancement.
   * **Geographic Salary Insights**: Maps or bar charts that display salary variations across different cities, states, and countries, showing location-based pay trends.
   * **Education’s Role in Salary Progression**: Visualizations that help illustrate how educational attainment correlates with salary and compensation.
   * **Interactive Filtering and Drill-Downs**: The dashboard allows users to filter data based on various dimensions (e.g., gender, age range, industry) to gain more specific insights.
6. **Insights and Reporting**:
   * **Automated Reporting**: Once the analysis is complete, the project generates a detailed report summarizing the key insights derived from the data, such as the impact of gender, education, and experience on salary.
   * **Actionable Insights**: The system provides recommendations to businesses on how to improve compensation strategies, address gender pay gaps, and attract and retain talent based on salary trends.

**Overall Functionalities:**

* **Data Collection**: Import salary survey data into a structured database.
* **Data Cleaning**: Prepare data for analysis by handling missing values, inconsistencies, and outliers.
* **Data Querying**: Generate queries to explore salary patterns, job titles, compensation, and other career factors.
* **Data Exporting**: Export query results into CSV files for use in Tableau.
* **Dashboard Creation**: Visualize the insights using Tableau, creating an interactive dashboard with multiple filters and drill-down options.
* **Insights Reporting**: Generate a comprehensive report that summarizes key findings, including salary disparities, compensation trends, and the effects of experience and education on salaries.

These functionalities combined will help in deriving actionable insights that can assist businesses, organizations, and policymakers in making informed decisions regarding compensation, workforce planning, and career development.

1. **Queries Implementation:**

**1. Average Salary by Industry and Gender**

**Query Purpose**: This query calculates the average salary within each industry, split by gender. It helps identify if there are any significant gender-based salary discrepancies within each industry.

**Explanation**:

* We group data by **Industry** and **Gender**, and then calculate the **average salary** for each combination.
* This query highlights any gender-based pay gaps, such as whether women or men tend to earn more in specific industries.

SELECT Industry, Gender, AVG(Annual\_Salary) AS Avg\_Salary

FROM salary\_data

GROUP BY Industry, Gender;

**2. Total Salary Compensation by Job Title**

**Query Purpose**: This query computes the total monetary compensation (base salary + additional monetary compensation) for each job title. It identifies which roles have the highest overall compensation.

**Explanation**:

* We combine **Annual Salary** and **Additional Monetary Compensation** for each **Job Title** and calculate the sum.
* This helps to assess the total earnings of different roles, including base salary and bonuses/commissions.

SELECT Job\_Title, SUM(Annual\_Salary + Additional\_Monetary\_Compensation) AS Total\_Compensation

FROM salary\_data

GROUP BY Job\_Title;

**3. Salary Distribution by Education Level**

**Query Purpose**: This query examines how salary varies with the highest level of education attained, providing insights into how education impacts salary.

**Explanation**:

* We group data by **Highest Level of Education Completed** and calculate the **average**, **minimum**, and **maximum** salary for each education level.
* This provides a clear understanding of how education influences earning potential.

SELECT Highest\_Level\_of\_Education\_Completed,

AVG(Annual\_Salary) AS Avg\_Salary,

MIN(Annual\_Salary) AS Min\_Salary,

MAX(Annual\_Salary) AS Max\_Salary

FROM salary\_data

GROUP BY Highest\_Level\_of\_Education\_Completed;

**4. Number of Employees by Industry and Years of Experience**

**Query Purpose**: This query determines how many employees are working in each industry, broken down by their years of professional experience. This can help assess the experience distribution across industries.

**Explanation**:

* We group the data by **Industry** and **Years of Professional Experience** and then count the number of employees in each group.
* This query helps identify which industries employ more experienced professionals.

SELECT Industry, Years\_of\_Professional\_Experience\_Overall, COUNT(\*) AS Num\_Employees

FROM salary\_data

GROUP BY Industry, Years\_of\_Professional\_Experience\_Overall;

**5. Median Salary by Age Range and Gender**

**Query Purpose**: This query calculates the **median salary** for different age ranges and genders. It helps explore salary trends across different age groups and gender.

**Explanation**:

* For each **Age Range** and **Gender**, the query calculates the **median salary**.
* It helps understand how salary progresses with age and whether there are any gender disparities within each age group.

SELECT Age\_Range, Gender, MEDIAN(Annual\_Salary) AS Median\_Salary

FROM salary\_data

GROUP BY Age\_Range, Gender;

**6. Job Titles with the Highest Salary in Each Country**

**Query Purpose**: This query identifies the highest-paying job titles in each country, providing insight into global salary trends.

**Explanation**:

* We group the data by **Country** and find the **Job Title** with the highest salary within each country.
* This query helps identify which positions are the most lucrative in different countries.

SELECT Country, Job\_Title, MAX(Annual\_Salary) AS Max\_Salary

FROM salary\_data

GROUP BY Country;

**7. Average Salary by City and Industry**

**Query Purpose**: This query calculates the average salary for each combination of **City** and **Industry**. It helps identify which cities offer the highest salaries within specific industries.

**Explanation**:

* We group the data by **City** and **Industry**, and calculate the **average salary** for each combination.
* This helps highlight cities with higher-paying industries.

SELECT City, Industry, AVG(Annual\_Salary) AS Avg\_Salary

FROM salary\_data

GROUP BY City, Industry;

**8. Percentage of Employees with Additional Monetary Compensation by Gender**

**Query Purpose**: This query calculates the percentage of employees within each gender who receive additional monetary compensation (e.g., bonuses, commissions).

**Explanation**:

* We calculate the percentage of employees who receive **Additional Monetary Compensation** within each **Gender**.
* This helps identify whether there are gender-based differences in receiving bonuses or other forms of additional compensation.

SELECT Gender,

(COUNT(CASE WHEN Additional\_Monetary\_Compensation > 0 THEN 1 END) / COUNT(\*)) \* 100 AS Percentage\_With\_Compensation

FROM salary\_data

GROUP BY Gender;

**9. Total Compensation by Job Title and Years of Experience**

**Query Purpose**: This query calculates the total compensation (salary + additional compensation) for each **Job Title**, broken down by **Years of Professional Experience**.

**Explanation**:

* For each **Job Title** and each **Experience Level**, we sum **Annual Salary** and **Additional Monetary Compensation**.
* This helps to identify how compensation changes with experience levels within different job titles.

SELECT Job\_Title, Years\_of\_Professional\_Experience\_Overall,SUM(Annual\_Salary + Additional\_Monetary\_Compensation) AS Total\_Compensation FROM salary\_data

GROUP BY Job\_Title, Years\_of\_Professional\_Experience\_Overall;

**10. Average Salary by Industry, Gender, and Education Level**

**Query Purpose**: This query analyzes how salary varies by **Industry**, **Gender**, and **Education Level**. It helps understand how multiple factors collectively influence salary.

**Explanation**:

* We group the data by **Industry**, **Gender**, and **Highest Level of Education Completed**, and calculate the **average salary**.
* This comprehensive analysis helps to see the combined impact of industry, gender, and education on salary.

SELECT Industry, Gender, Highest\_Level\_of\_Education\_Completed, AVG(Annual\_Salary) AS Avg\_Salary

FROM salary\_data

GROUP BY Industry, Gender, Highest\_Level\_of\_Education\_Completed;

**7 . Results and Outcomes:**

The analysis of career trajectories and salary patterns across global industries revealed significant insights. Gender-based salary discrepancies persist, especially in high-paying industries like **Technology** and **Finance**, where men tend to earn more than women. Higher education levels, particularly **Collage Degree**, correlate with higher salaries, especially in sectors like **Technology**. **Experience** plays a crucial role in salary growth, with senior roles offering significantly higher compensation. Geographic differences show that cities like **Albuqerque** and **Ann Abror**  provide higher salaries, particularly in **Tech** and **Finance**. Gender disparities were also noted in **non-salary compensations** such as bonuses and stock options, with men receiving more. Overall, the findings highlight the need for businesses to focus on **gender pay equity**, competitive compensation strategies, and offering equal opportunities for **higher education** and **experience growth**.

Top of Form

Bottom of Form

**8 . Conclusion:**

This project provides a comprehensive analysis of career trajectories and salary patterns across global industries, uncovering important trends and disparities. The findings highlight the ongoing issue of **gender-based salary discrepancies**, particularly in high-paying sectors like **Technology** and **Finance**, suggesting a need for targeted efforts to achieve **pay equity**. The impact of **education** and **experience** on salary is clear, with higher education and professional experience leading to higher compensation, especially in specialized fields. Geographic differences further emphasize the premium on salaries in cities like **Albuqerque** and **Ann Abror**.

The analysis also reveals that men are more likely to receive **additional monetary compensation** such as bonuses, raising concerns about equity in non-salary benefits. Overall, this project underscores the importance of addressing **gender disparities**, offering fair **compensation** for all levels of experience, and ensuring that **education** and **career development opportunities** are accessible to all employees. These insights provide actionable guidance for companies seeking to improve their compensation strategies and foster more equitable and competitive work environments.