**Project Information**

|  |
| --- |
| Title : Exploratory Data Analysis on Laptop Sales Dataset |
| Name : Meganathan A |
| DA/DS : May-2025 |
| Batch Number : RP-36 |
| Online/Offline : Offline |
| Roll Number : B2025057352 |

**Github link :** [Exploratory datat analysis on laptop saless datatset](https://github.com/Meganathan2405/pythonproject)

**Table of Contents**

|  |  |
| --- | --- |
| **S.No** | **INDEX** |
| **1** | Project Overview |
| **2** | Objective |
| **3** | Dataset Description |
| **4** | Data Preparation |
| **5** | SQL Queries Executed |
| **6** | Excel Dashboard Development |
| **7** | Key Visual Insights |
| **8** | Screenshots |
| **9** | Insights Summary |
| **10** | Challenges Faced |
| **11** | Tools Used |
| **12** | Final Outcome |
| **13** | Conclusion |

Salary Data Analysis Project Report

1. Project Overview

This project focuses on analyzing a salary dataset to understand trends and patterns across various demographic and professional attributes, including industry, gender, job title, education, location, and experience. The data was processed, analyzed using MySQL, and visualized through Excel dashboards to extract actionable insights.

2. Objective

The objective of this project is to:

* Analyze salary distributions across industries and job roles.
* Understand compensation differences by gender and location.
* Evaluate the influence of education and experience on salaries.
* Identify high-paying sectors and roles.
* Create a dashboard for visual insights.

3. Dataset Description

* File Name: miles.xlsx
* Total Records: ~500+
* Key Columns:
  + Industry
  + Gender
  + Job\_Title
  + Annual\_Salary(INR)
  + Additional\_Monetary\_Compensation(INR)
  + Highest\_Level\_of\_Education\_Completed
  + Years\_of\_Professional\_Experience\_Overall
  + Country, City, Age\_Range

The dataset was cleaned to ensure consistency, handle missing values, and correct data types.

4. Data Preparation

1. Cleaning:
   * Renamed inconsistent column headers.
   * Removed null or incorrect entries.
   * Ensured proper data types (numeric for salary columns).
2. Conversion:
   * Excel file converted to CSV for MySQL import.
3. Database Setup:
   * Database Name: sal
   * Table Name: mmm
   * Imported using:
   * LOAD DATA INFILE 'path\_to\_file.csv'
   * INTO TABLE mmm
   * FIELDS TERMINATED BY ','
   * ENCLOSED BY '"'
   * LINES TERMINATED BY '\n'
   * IGNORE 1 ROWS;

5. SQL Queries Executed

1. Average Salary by Industry and Gender
2. Total Compensation by Job Title
3. Salary Statistics by Education Level
4. Employee Count by Industry and Experience
5. Median Salary by Age Range and Gender
6. Top Earners in Each Country
7. Average Salary by City and Industry
8. Percentage of Employees with Additional Compensation by Gender
9. Total Compensation by Job and Experience
10. Average Salary by Industry, Gender, and Education

Each query result was exported as a .csv file.

6. Excel Dashboard Development

Workbook Name: miles\_analysis.xlsx

* 10 sheets for each query result.
* Dashboard sheet summarizes visual insights:
  + Bar charts for salary comparisons.
  + Pie chart for compensation distribution by gender.
  + Line graph for salary vs. experience.
  + Heatmap by city and industry.

7. Key Visual Insights

* Tech and Finance dominate in average salary.
* Males receive more additional compensation than females.
* Data Scientist and Product Manager are top-paying jobs.
* Bangalore leads in average city salary.
* Master’s and Doctorate holders earn higher than Bachelor's and High School graduates.

8. Screenshots

* Screenshot 1: Cleaned Excel Dataset
* Screenshot 2: MySQL table view of mmm
* Screenshot 3: Query results in Excel
* Screenshot 4: Bar chart of salary by industry
* Screenshot 5: Pie chart of gender-based compensation
* Screenshot 6: Final dashboard view

9. Insights Summary

* Industries: Tech, Finance, Healthcare pay better.
* Job Titles: Managerial and technical roles have higher compensation.
* Gender Gap: Exists, especially in bonuses.
* Education Impact: Advanced degrees correlate with higher salaries.
* Experience Matters: Steady rise in compensation with years of experience.
* City Influence: Metro cities offer significantly higher packages.

10. Challenges Faced

* Inconsistent column names in raw data.
* Null values in critical fields like salary and compensation.
* Need to normalize educational and job titles.
* Time spent on formatting and cleaning data.

11. Tools Used

* MySQL: For structured querying and aggregation.
* Excel: For cleaning, analysis, and visualization.
* PowerPoint: To prepare presentation slides.

12. Final Outcome

* 10 well-documented SQL queries with insights.
* Cleaned and formatted Excel workbook.
* Rich visual dashboard for summary view.
* Insights ready for HR and strategic use.

13. Presentation Overview

The final PowerPoint includes:

1. Title Slide
2. Objective
3. Dataset Description
4. Methodology
5. SQL Highlights
6. Key Results
7. Charts and Visuals
8. Dashboard Snapshots
9. Insights Summary
10. Thank You/Q&A

14. Conclusion

This project provided a complete end-to-end data analysis pipeline—from raw data cleaning to advanced SQL querying and dashboard creation. It highlighted the real-world application of database systems and business intelligence in workforce analytics. The learnings and insights gained here can be applied to HR analytics, industry benchmarking, a