

Employee Data Analysis using Excel

STUDENT NAME: VALARMATHI.B

NM ID: asunm110312201416

DEPARTMENT: COMMERCE

COLLEGE : DRBCCC HINDU COLLEGE

AGENDA

- Problem Statements
- Overview
- End Users
- Our Solution and Proposition
- Dataset Descriptions
- Modelling Approach
- Results and Discussion
- Conclusion

Problem Statements

1. Salary Comparison Based on Education Level:

- Compare the average salary across different education levels (Bachelors, Masters, and PhD).
- Determine whether higher education levels correlate with higher salaries across the locations.

2. Geographical Salary Distribution:

- Investigate the salary distribution across the three cities: Bangalore, Pune, and New Delhi.
- Identify which city offers the highest average salary and whether there is a significant difference between cities.

3. Impact of Age on Salary:

- Examine the relationship between age and salary.
- Determine if there is a particular age group that earns significantly more or less than others.

4. Education and Salary Trends Over Time:

- Analyze how the salary trends vary with age for different educational qualifications.
- Identify any patterns or anomalies in salary growth with respect to education level and age.

The goal is to understand the key factors that influence salary variations and to provide insights for stakeholders on how education, age, and location impact earnings.

Project overview

Education Level

- The dataset includes individuals with Bachelor's degrees, Master's degrees, and PhDs.
- Majority of the entries are for individuals with Bachelor's degrees

Age Distribution

- Ages range from 22 to 40 years.

Salary Distribution

- Salaries range from as low as ₹12,400 to as high as ₹69,000.
- Higher salaries are generally associated with individuals in Bangalore and -New Delhi.

Location Insights

- Bangalore: Appears to offer a wider salary range, with some of the highest salaries in the dataset (up to ₹69,000).
- Pune: Generally shows lower salaries, with most entries between ₹12,500 and ₹54,000.

End Users

1. Human Resources Professionals
2. Job Seekers
3. Recruitment Agencies
4. Market Analysts and Consultants
5. Educational Institutions
6. Government Agencies
7. Employers and Companies

These end users would benefit from this data by making more informed decisions related to employment, education, and economic planning.



Solution

ANALYSIS OF SALARY BASED ON EDUCATION LEVEL

We could calculate the average salary for each education level (Bachelors, Masters, PhD) to validate this hypothesis.

2. ANALYSIS OF SALARY BASED ON LOCATION

We can calculate and compare the average salaries across different cities to see if this proposition holds true.

3. AGE AND SALARY CORRELATION

We could examine the correlation between age and salary to see if older individuals earn more, irrespective of their education level or location.

Proposed timeline

IMPACT OF REPEATED ENTRIES ON DATA INTEGRITY

Duplicate entries could distort the average salary calculations and lead to incorrect assumptions.

SALARY RANGE ANALYSIS

There could be a significant variance in salaries within the same educational level or location.



ANALYSIS OF SALARY BASED ON LOCATION

Salaries might vary significantly depending on the location. Bangalore, as a tech hub, might offer higher salaries compared to Pune or New Delhi.

■ EDUCATION LEVEL

Educational qualification of the individual.

- Bachelors - Masters - PhD

■ AGE

The age of the individual in years, ranging from 22 to 40.

■ SALARY

The salary of the individual in local currency (assuming INR based on the Indian city locations). The salary values range from 12,400 to 69,000.

■ LOCATION

The city where the individual resides. The dataset includes the following cities:

- Bangalore - Pune - New Delhi

Dataset Description





Modelling Approach

The modeling approach depends on your goal, which could be:

A. Predicting Salary (Regression Task)

For predicting the salary based on education, age, and location:

- Regression Algorithms:

- Linear Regression
- Decision Trees/Random Forest
- Gradient Boosting (XGBoost, LightGBM)
- Support Vector Regression (SVR)

- Evaluation Metrics: Use metrics like Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared for evaluating model performance.

B. Salary Group Classification (Classification Task)

If you want to classify individuals into salary groups (e.g., low, medium, high):

- Classification Algorithms:

- Logistic Regression
- Decision Trees/Random Forests
- Support Vector Machines (SVM)
- K-Nearest Neighbors (KNN)

- Evaluation Metrics: Use accuracy, precision, recall, F1-score, and confusion matrix for classification performance.



Result



Discussion



1. Education and Salary:

- Bachelor's Degree: ₹12,500 to ₹69,000. The highest-paying in Bangalore and New Delhi.
- Master's Degree: ₹12,000 to ₹45,000. Bangalore earn more compared to those in New Delhi.
- PhD: Both earn ₹25,350.

2. Age and Salary:

- Younger Workers: Some younger workers (e.g., 22-23 years old) with a bachelor's degree earn relatively high salaries, especially in Bangalore and New Delhi.
- Older Workers: Those aged 30 and above tend to have more varied salaries. Some earn significantly more, while others earn less depending on the city and industry.

3. City Comparison:

- Bangalore: Offers higher salaries overall, especially for bachelor's and master's degree holders.
- Pune: Generally offers lower salaries compared to Bangalore.
- New Delhi: Shows a wide range of salaries, from ₹12,000 to ₹66,000, depending on the role and education level.

Conclusion

Salaries depend on a combination of factors such as education, experience, and city. Bangalore generally offers higher-paying jobs, especially for individuals with a bachelor's degree, while New Delhi and Pune show more variation.



Thank you