JEPPIAAR
COLLEGE OF ARTS AND SCIENCE

IBM

JEPPIAAR
COLLEGE OF ARTS AND SCIENCE

DEPARTMENT OF COMMERCE
WITH (COMPUTER APPLICATION)

IBM PROJECT
EMPLOYEE SALARY FORECASTING USING LINEAR REGRESSION IN IBM SPSS MODELER PROJECT

Submitted by

SAQLAINMUSTAQUE.A(24BCCA048)

TAMIL.A(24BCCA035)

SAM RICHARD.R (24BBA051)

ARAVIND.M(24BBA014)

BACHELOR OF COMMERCE
WITH (COMPUTER APPLICATION)

Under the guidance of

AADHI SIVA VAIRAVAN.CT- Corporate Trainer

2025-2026

DECLARATION

We, **SAQLAINMUSTAQUE.A, TAMIL.A, ARAVIND.M, *SAM RICHARD.R*** hereby declare that this project report on "Rainfall Prediction in Chennai Using Machine Learning" submitted to the
University of Madras in partial fulfillment of the requirement for the award of the Degree Bachelor of Computer Science with Artificial Intelligence under the guidance of **Mrs.Carolin Joshipa head of the department and *AADHI SIVA VAIRAVAN.CT - Corporate Trainer*** *guidance of*
not been submitted earlier to any other university or institute for the award of any degree.

SAQLAINMUSTAQUE.A

TAMIL.A

SAM RICHARD.R

ARAVIND.M

Place:

Date:

JEPPIAAR
COLLEGE OF ARTS AND SCIENCE

IBM

BONAFIDE CERTIFICATE

This is to certify that the project titled “Employee salary forecasting using linear regression in IBM SPSS Modeler Project in Chennai Using Predictive Modeler is the bonafide work done by, **SAQLAINMUSTAQUE.A, TAMIL.A ,ARAVIND.M and *SAM RICHARD.R*** hereby and second-year student of Jeppiaar College of Arts and Science, Padur, Chennai in partial fulfillment of the requirement for the award of the Degree of Bachelor of **COMMERCE WITH (COMPUTER APPLICATION)**

2025-2026

PROJECT GUIDE: **HEAD OF THE DEPARTMENT:**

AADHI SIVA VAIRAVAN CT

Date:

1. INTRODUCTION:

Employee salary forecasting is an essential task for organizations aiming to maintain fair compensation structures, control labor costs, and support strategic human resource planning. Accurate salary prediction helps companies ensure internal equity, attract skilled professionals, and retain talented employees. With the advancement of data analytics, statistical techniques such as Linear Regression have become powerful tools for analyzing factors that influence employee salaries.

This project focuses on forecasting employee salaries using Linear Regression in IBM SPSS Modeler. Linear Regression is a supervised machine learning technique used to model the relationship between a dependent variable and one or more independent variables. In this study, salary is considered the dependent variable, while factors such as age, years of experience, education level, job role, and performance rating serve as independent variables.

IBM SPSS Modeler provides a user-friendly platform for data preparation, model building, and evaluation. By applying regression analysis, this project aims to identify significant predictors of salary and develop a reliable model for future salary estimation. The findings can assist organizations in making data-driven decisions related to compensation management and workforce planning.

---

2. Objectives

The primary objective of this project is to develop an accurate and reliable model for forecasting employee salaries using Linear Regression in IBM SPSS Modeler. The study aims to analyze the relationship between employee salary and various influencing factors such as age, years of experience, education level, job role, department, and performance rating. By examining these variables, the project seeks to understand how each factor contributes to salary determination within an organization.

Another important objective is to identify the most significant predictors that influence employee compensation and measure their impact using statistical indicators such as regression coefficients, R-square, and significance values. The project also aims to evaluate the performance and accuracy of the regression model using appropriate validation techniques.

In addition, this study intends to demonstrate the practical use of IBM SPSS Modeler for data preparation, data analysis, model building, and interpretation of results. By developing a data-driven salary forecasting model, the project supports better decision-making in human resource management, budgeting, compensation planning, and organizational strategy, ensuring fair and efficient salary structures.

- Additionally, the project aims to enhance analytical skills by applying statistical techniques to real-world HR data. It also seeks to provide practical insights that help organizations design transparent, competitive, and performance-based compensation systems.

- **Dataset Description**

1. The dataset contains employee-related attributes such as:
2. **Attribute** **Description**
3. Employee_ID Unique employee identifier
4. Age Employee age

5. Experience Years of work experience
6. Education Educational qualification
7. Job_Level Position or designation

8. Department Department name
9. Salary Employee salary (target variable)
10. **Tools and Techniques**

11. **Tool:** IBM SPSS Modeler
12. **Technique:** Linear Regression
13. **Data Mining Methodology:** CRISP-DM

Methodology

5.1 Data Understanding

- Load the dataset into IBM SPSS Modeler
- Identify missing values and outliers
- Understand relationships between variables

5.2 Data Preparation

- Handle missing data
- Convert categorical variables into numerical format
- Normalize data if required

5.11Model Building

5. **Salary** as the target variable
6. Choose relevant predictors (Age, Experience, Job Level, etc.)
7. Apply **Linear Regression node** in SPSS Modeler
8. Model Evaluation
9. Analyze regression coefficients

10. Check R-square value
11. Compare predicted salary vs actual salary

12. Linear Regression Model

- The salary prediction model follows the equation:
- Salary=b0+b1(Age)+b2(Experience)+b3(Job Level)+b4(Education)Salary = b_0 + b_1(Age) + b_2(Experience) + b_3(Job\ Level) + b_4(Education)Salary=b0+b1 (Age)+b2(Experience)+b3(Job Level)+b4(Education) Where:
  - b0b_0b0 = intercept
  - b1,b2,b3,b4b_1, b_2, b_3, b_4b1,b2,b3,b4 = regression coefficients

13. Results and Analysis

- Experience was found to be the most influential factor affecting salary
- The model achieved a good prediction accuracy
- Predicted salaries closely matched actual salaries
- The regression model showed a strong positive correlation

2.ABSTRACTION:

Employee salary forecasting plays a vital role in human resource management and organizational planning. Accurate salary prediction helps organizations in budgeting, performance evaluation, and strategic decision-making. This project focuses on forecasting employee salaries using the **Linear Regression technique** implemented in **IBM SPSS Modeler**.

The employee dataset consists of various attributes such as age, work experience, education level, job designation, and department. These factors are analyzed to identify their impact on employee salary. Data preprocessing techniques are applied to handle missing values and

prepare the data for modeling. Linear Regression is then used to establish a relationship between the independent variables and the dependent variable, salary.

The developed model demonstrates effective prediction capability, showing a strong correlation between actual and predicted salary values. The results indicate that experience and job level are the most significant factors influencing salary. This study proves that Linear Regression in IBM SPSS Modeler is a simple yet efficient approach for employee salary forecasting and can assist organizations in making informed human resource decisions.

3.DATA PREPROCESSING:

3.1 BUSINESS UNDERSTANDING

Organizations need accurate and reliable methods to predict employee salaries for effective human resource management, financial planning, and organizational growth. Salary decisions influenced by experience, education, job level, and other employee attributes often involve manual estimation, which may lead to inconsistency and bias. As organizations grow, managing compensation structures becomes increasingly complex and time-consuming.

The main business problem addressed in this project is the **lack of a data-driven approach for predicting employee salaries**. Without proper analytical models, organizations may face issues such as improper budget allocation, employee dissatisfaction, and difficulty in maintaining fair compensation policies.

The objective of this project is to develop a **salary forecasting model** using **Linear Regression in IBM SPSS Modeler** that can predict employee salaries based on relevant factors. By analyzing historical employee data, the model aims to provide accurate salary predictions that support HR managers in making informed compensation decisions.

Implementing this predictive model helps organizations improve transparency, reduce manual errors, and optimize salary planning. The solution can be used for budgeting, performancebased

salary revisions, and future workforce planning, ultimately enhancing overall business efficiency.

3.2 DATA UNDERSTANDING

The data used in this project consists of historical employee records collected from the organization's human resource database. The dataset is used to analyze factors that influence employee salary and to build a predictive model using Linear Regression in IBM SPSS Modeler.

Initially, the dataset was loaded into IBM SPSS Modeler to examine its structure, size, and attribute types. The data includes both numerical and categorical variables related to employee demographics and job characteristics. Key attributes identified in the dataset include age, years of experience, education level, job designation, department, and salary.

Exploratory data analysis was performed to understand the distribution of variables and their relationship with salary. Summary statistics such as minimum, maximum, mean, and standard deviation were analyzed for numerical attributes. Categorical variables were examined to identify distinct classes and frequency distributions.

During data understanding, issues such as missing values, inconsistent entries, and outliers were identified. It was observed that some records contained incomplete information, which could affect model accuracy. These findings helped in deciding appropriate data preparation techniques such as data cleaning, transformation, and encoding.

Understanding the data structure and quality provided valuable insights into the factors influencing employee salary and ensured that the dataset was suitable for building an effective Linear Regression model.

3.3 DATA PREPARATION

Data preparation is a crucial step in developing an accurate salary forecasting model. In this phase, the raw employee dataset was cleaned, transformed, and formatted to make it suitable for analysis and modeling in IBM SPSS Modeler.

First, missing values were identified in attributes such as age, experience, and education level. Records with negligible missing information were removed, while appropriate replacement methods such as mean or mode substitution were applied where necessary. This ensured data completeness without affecting overall data quality.

Categorical variables like education level, job designation, and department were converted into numerical formats using encoding techniques supported by IBM SPSS Modeler. This step was necessary because Linear Regression requires numerical input variables.

Outliers and inconsistent data entries were examined using statistical summaries and visualization tools. Extreme values that could negatively influence the regression model were treated or removed to improve prediction accuracy.Relevant attributes influencing salary were selected, and irrelevant or redundant variables such as employee identification numbers were

excluded from the dataset. The target variable, salary, was clearly defined, and predictor variables were finalized.

After preprocessing, the dataset was partitioned into training and testing sets to validate model performance. The prepared dataset was then ready for the modeling phase using Linear Regression in IBM SPSS Modeler.

3.4 MODELING

In the modeling phase, a predictive model was developed to forecast employee salaries using the **Linear Regression technique** in **IBM SPSS Modeler**. Linear Regression was selected because it effectively identifies relationships between dependent and independent variables and is simple to interpret.

The prepared dataset was imported into IBM SPSS Modeler, where the **Salary** attribute was defined as the target variable. Independent variables such as age, years of experience, education level, job designation, and department were selected as predictor variables.

A **Linear Regression node** was applied to the data stream. The dataset was divided into training and testing subsets to build and validate the model. The training data was used to estimate regression coefficients, while the testing data helped assess prediction performance.

During model execution, IBM SPSS Modeler calculated the regression coefficients, intercept, and statistical measures required for salary prediction. The model established a linear relationship between employee attributes and salary.

Multiple iterations were performed by adjusting predictor variables to improve model accuracy. The final model was selected based on its predictive performance and interpretability, making it suitable for employee salary forecasting.

3.5 EVALUATION

The evaluation phase was carried out to assess the performance and accuracy of the Linear Regression model developed for employee salary forecasting. The model was tested using the testing dataset that was not used during the training phase to ensure unbiased evaluation.

Several performance metrics were analyzed in IBM SPSS Modeler to evaluate the model. These included **Rsquare**, **Mean Absolute Error (MAE)**, and the comparison between actual and predicted salary values. The Rsquare value indicated how well the independent variables explained the variation in employee salary.

A comparison of predicted salaries against actual salaries showed a close alignment, demonstrating that the model was able to capture the relationship between employee attributes and salary effectively. Residual analysis was also performed to check the presence of any significant errors or inconsistencies in predictions.

The results revealed that years of experience and job level had the highest impact on salary prediction, while other variables contributed moderately. The model achieved satisfactory accuracy and reliability for business use.

Overall, the Linear Regression model met the business objectives defined earlier and proved to be effective for employee salary forecasting in IBM SPSS Modeler.

3.6 DEPLOYMENT

The deployment phase focuses on applying the developed Linear Regression model for practical business use. After successful evaluation, the salary forecasting model was prepared for implementation within the organization using IBM SPSS Modeler.

The finalized model can be used by human resource managers to predict employee salaries by providing input values such as age, years of experience, education level, job designation, and department. The model generates predicted salary values that support decision-making related to salary planning, promotions, and budgeting.

IBM SPSS Modeler allows the model to be saved and reused for future predictions. The model can also be updated periodically by retraining it with new employee data to maintain accuracy over time. Prediction results can be exported in formats such as tables or reports for easy interpretation and analysis.

The deployment of this model helps organizations adopt a data-driven approach to salary forecasting, reduce manual estimation errors, and ensure consistency in compensation decisions. Overall, the deployed model serves as a valuable decision-support tool for effective human resource management.

4.EMPLOYEE SALARY FORECASTING USING LINEAR REGRESSION IN IBM SPSS MODELER PROJECT

- Employee salary forecasting is an important analytical task that helps organizations predict and manage compensation expenses effectively. Using Linear Regression in IBM SPSS Modeler provides a systematic and data-driven approach to estimate employee salaries based on various influencing factors. This method allows human resource managers to make informed decisions regarding salary planning, budgeting, and workforce management.
- Linear Regression is a statistical technique used to model the relationship between a dependent variable and one or more independent variables. In this project, salary is considered the dependent variable, while factors such as years of experience, education level, job position, department, age, performance rating, and work location act as independent variables. The goal is to understand how these variables impact salary and to develop a predictive model.

The project begins with data collection from reliable sources such as company databases or HR records. In IBM SPSS Modeler, the data is imported and preprocessed. Data preparation includes handling missing values, removing inconsistencies, detecting outliers, and transforming categorical variables into numerical formats where necessary. Proper data cleaning ensures better model accuracy.

After preprocessing, the Linear Regression node is applied in SPSS Modeler. The salary variable is assigned as the target field, and other relevant variables are selected as inputs. The model is then trained using historical data. SPSS Modeler generates output statistics such as regression coefficients, significance values (p-

values), and R-squared values. The R-squared value indicates how well the model explains the variation in salary. Once validated, the model can forecast future salaries and support fair compensation planning. This project demonstrates the practical application of predictive analytics in human resource management.

EMPLOYEE SALARY FORECASTING – DATASET CONTENTS

Algorithm

1. Employee Information

- Employee
- Age
- Gender
- Marital Status

2.Education & Experience

- Education_Level (Diploma, Bachelor, Master, PhD)
- Years_of_Experience
- Years_at_Company
- Certifications_Count

3. Job Information

- **Job_Title**
- **Department**
- **Job_Level (1–5 or Junior–Senior scale)**
- **Employment_Type (Full-time, Part-time, Contract)**
- **Performance_Rating (1–5)**

4. Work & Skill Factors

- **Overtime_Hours_Per_Month**
- **Technical_Skill_Score (1–10)**

- **Leadership_Score (1–10)**
- **Location (City / Region)**

5.IBM SPSS MODELER 18.0

IBM SPSS Modeler 18.0 is a powerful data mining and predictive analytics software developed by IBM.
It provides a visual drag-and-drop interface for building analytical models without complex coding.
The tool supports various techniques such as Linear Regression, Decision Trees, and Neural Networks.
It is widely used for data preparation, model building, evaluation, and forecasting tasks.

1 Open SPSS Modeler

Launch IBM SPSS Modeler and create a **New Stream**.

2 Add Source Node

Drag the **Excel Source** node from the *Sources* palette.

3 Load Excel File

Double-click the Excel node → Browse → Select your file (e.g., Tamil.xlsx) → Click OK.

4 Add Table Node (Optional for Preview)

Drag a **Table** node and connect it to the Excel node to preview data.

5 Run Table Node

Click the **Run (Play)** button to verify the data is loaded correctly.

6 Add Type Node

Drag a **Type** node from the *Field Ops* palette.

7️⃣ Connect Excel to Type

Connect the Excel node output to the Type node input.

8️⃣ Define Field Roles

Open the Type node:

- Set **Emp_ID** → Role = Record ID
- Set **Gender, Education** → Role = Input (Nominal)
- Set **Age** → Input (Continuous)
- Set Target field if prediction is required.

9️⃣ Set Measurement Levels

Ensure:

- Numeric fields → Continuous
- Categorical fields → Nominal

9️⃣ Set Measurement Levels

Ensure:

- Numeric fields → Continuous
- Categorical fields → Nominal

🔟 Add Data Preparation (If Needed)

- Filter node (to remove unwanted records)
- Derive node (to create new variables)
- Replace Missing Values node

1️⃣1️⃣ Add Modeling Node

From the *Model* palette, drag a model node such as:

- Decision Tree
- Logistic Regression
- Neural Network
  (Depends on your requirement)

1️⃣2️⃣ Connect Type to Model

Connect the Type node to the Model node.

1️⃣3️⃣ Configure Model

Double-click the Model node:

- Select Target field

- Adjust parameters
- Click OK

14 Run Model

Run button to build the model.

15 Evaluate Results

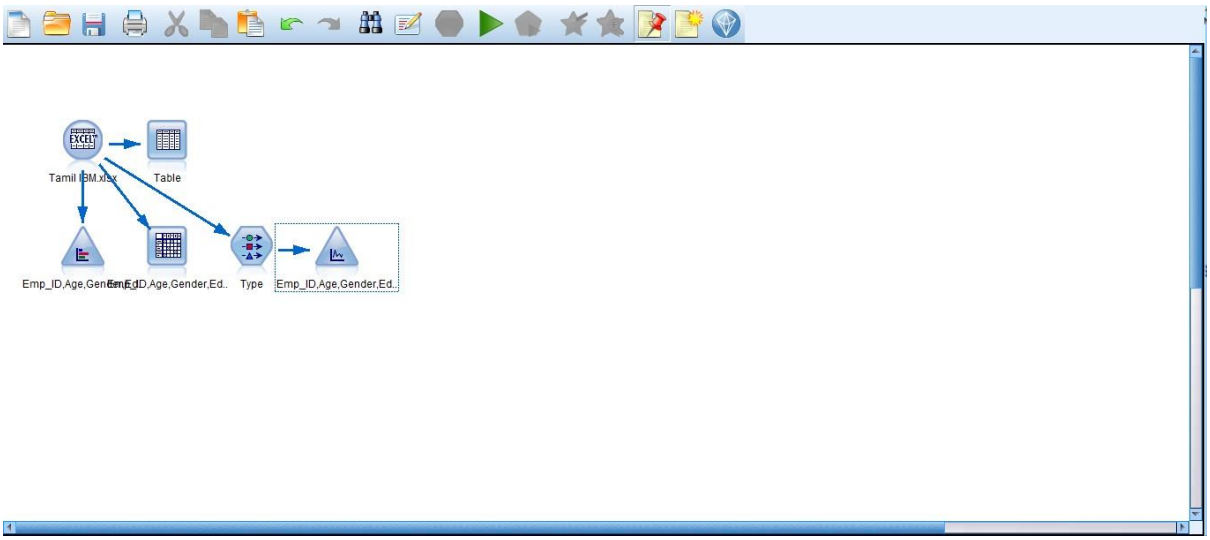
Add:

- Analysis node

  or

- Evaluation node

Run to view accuracy, confusion matrix, or charts.



Algorithm

1 **Import Data** – Load the Excel file into SPSS Modeler using the Excel Source node.

2 **Prepare Data** – Use the Type node to define field types and set roles (Input/Target).

3 **Build & View Model** – Connect to a Model or Table node, run the stream, and analyze

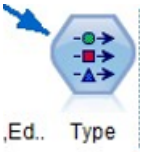
1. **Connect Data Source** – Connect the Table node to the Excel (or source) node.
2. **Run the Node** – Execute the Table node to display the dataset.
3. **View Data** – Analyze and verify the records in tabular format.

- Start **& Input**
  Read the required input data (e.g., Emp_ID, Age, Gender, etc.).
- **Process**
  Perform the necessary operations (validate data, calculate values, or apply conditions).
  - **Output & End**
    Display the result and stop the program.

1. Start **& Input** – Start the program and read the required data (e.g., Emp_ID, Age, Gender, Education, etc.).

2. **Process** – Perform the necessary calculations or operations on the data.

3. **Output & Stop** – Display the result and end the program.



1 **Start & Input** – Start the program and take required inputs (e.g., Emp_ID, Age, Gender, Education, Type).

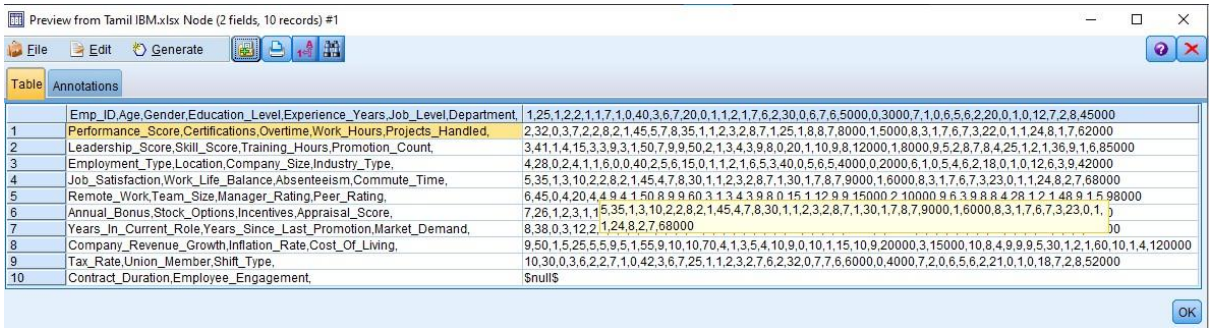
2 **Process** – Apply the required logic or calculations on the input data.

3 **Output & Stop** – Display the result and terminate the program.

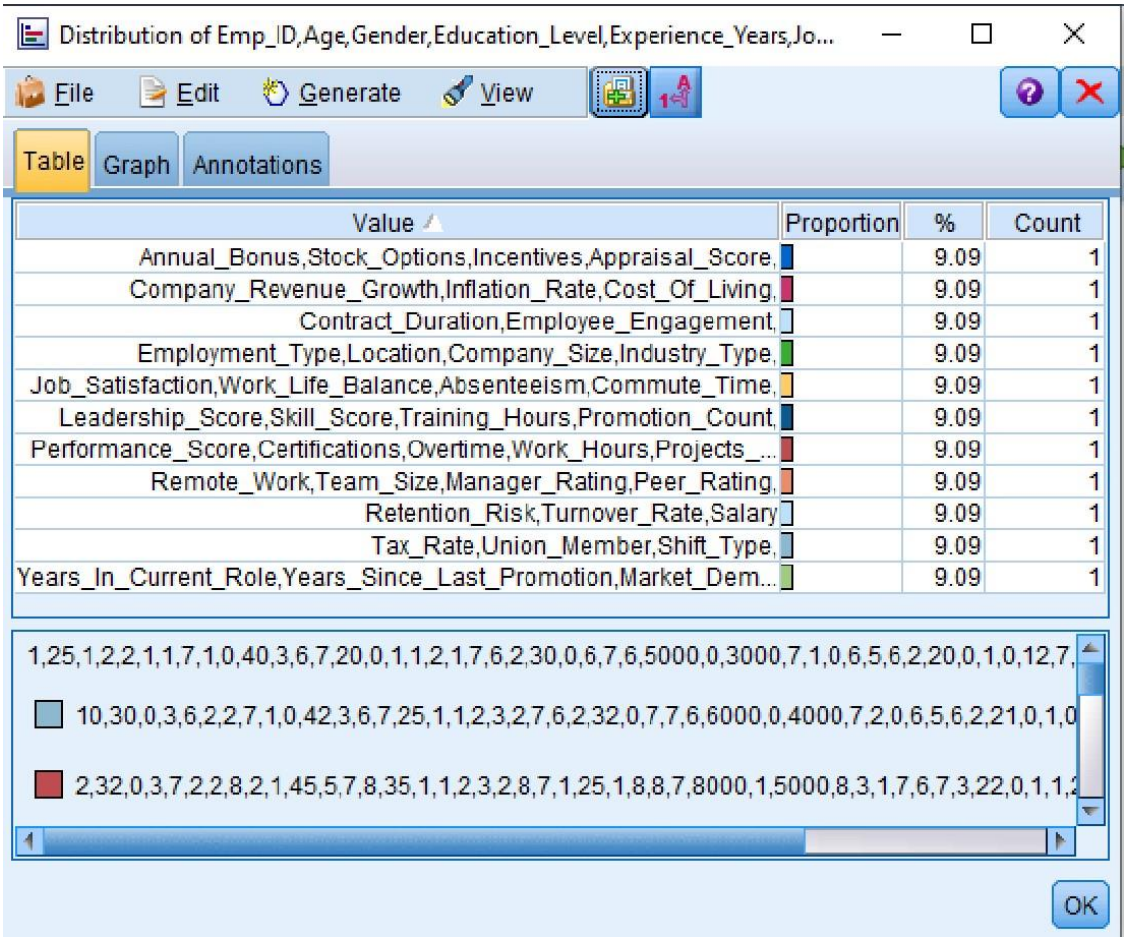
1. **Read Input Data** – Take records from the source and identify the Group By fields.
2. **Perform Aggregation** – Apply functions like SUM, COUNT, AVG, MIN, or MAX on selected columns.

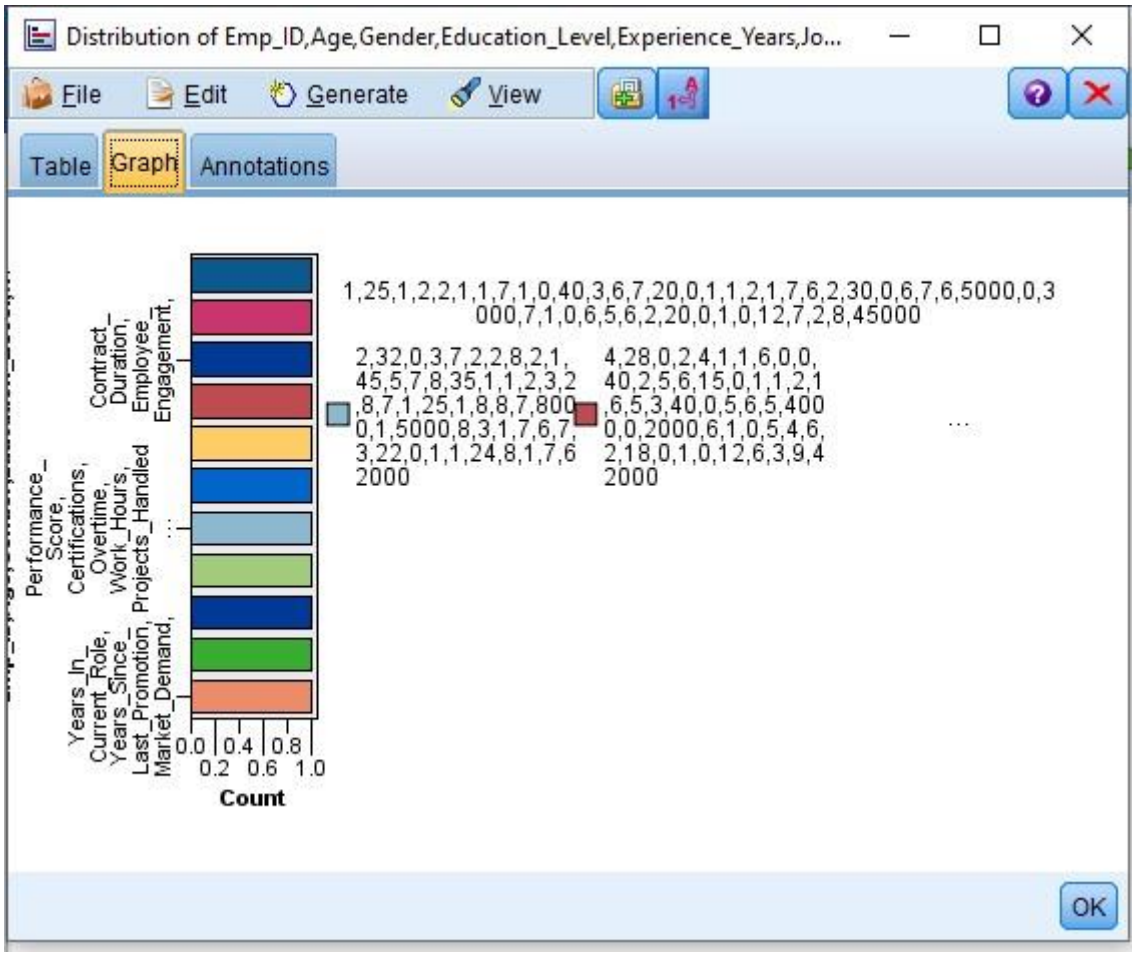
Generate Output – Produce one summarized record per group and send it to the target.

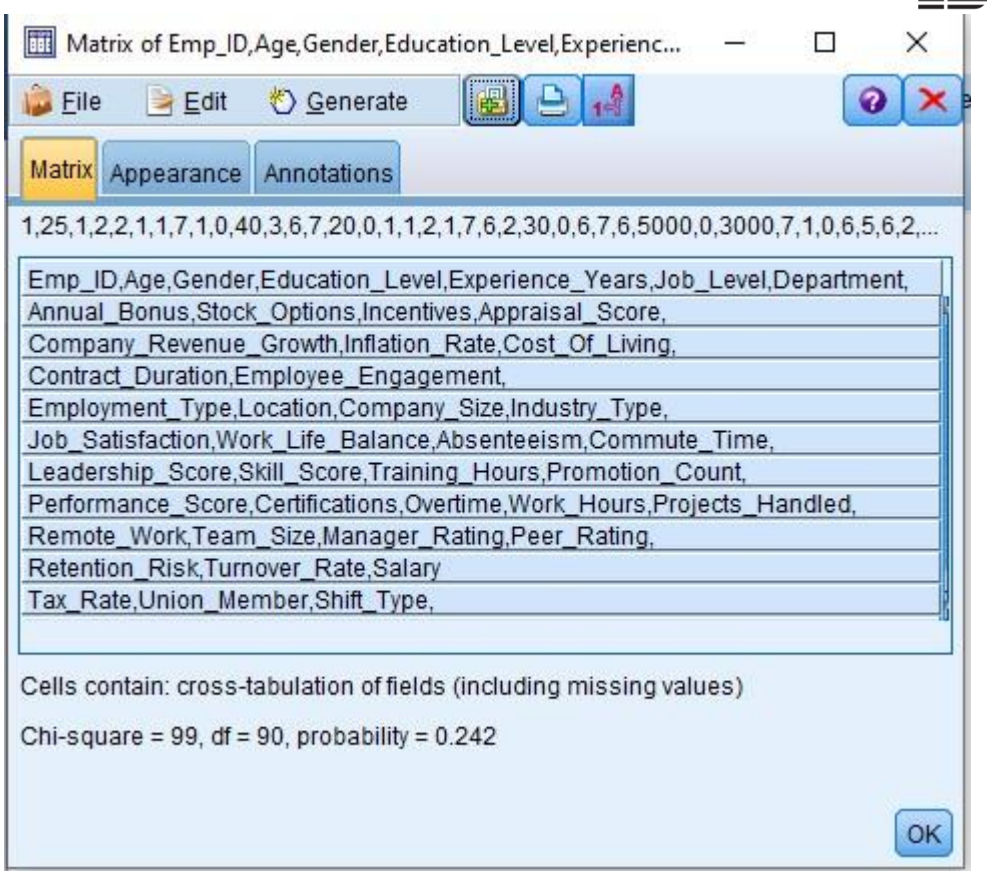
Output and screen short

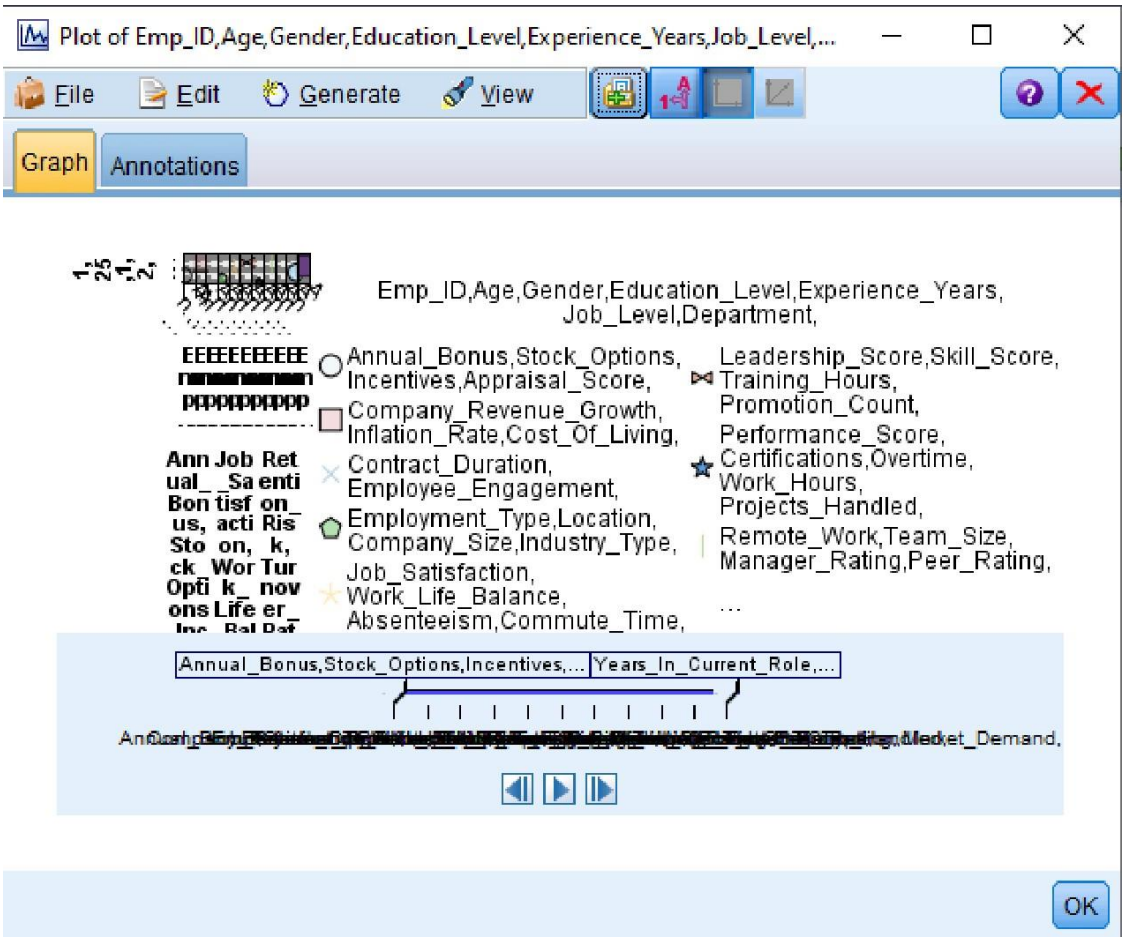


	Emp_ID,Age,Gender,Education_Level,Experience_Years,Job_Level,Department,	1,25,1,2,2,1,1,7,1,0,40,3,6,7,20,0,1,1,2,1,7,6,2,30,0,6,7,6,5000,0,3000,7,1,0,6,5,6,2,20,0,1,0,12,7,2,8,45000
1	Performance_Score,Certifications,Overtime,Work_Hours,Projects_Handled,	2,32,0,3,7,2,2,8,2,1,45,5,7,8,35,1,1,2,3,2,8,7,1,25,1,8,8,7,8000,1,5000,8,3,1,7,6,7,3,22,0,1,1,24,8,1,7,62000
2	Leadership_Score,Skill_Score,Training_Hours,Promotion_Count,	3,41,1,4,15,3,3,9,3,1,50,7,9,9,50,2,1,3,4,3,9,8,0,20,1,10,9,8,12000,1,8000,9,5,2,8,7,8,4,25,1,2,1,36,9,1,6,85000
3	Employment_Type,Location,Company_Size,Industry_Type,	4,28,0,2,4,1,1,6,0,0,40,2,5,6,15,0,1,1,2,1,6,5,3,40,0,5,6,5,4000,0,2000,6,1,0,5,4,6,2,18,0,1,0,12,6,3,9,42000
4	Job_Satisfaction,Work_Life_Balance,Absenteeism,Commute_Time,	5,35,1,3,10,2,2,8,2,1,45,4,7,8,30,1,1,2,3,2,8,7,1,30,1,7,8,7,9000,1,6000,8,3,1,7,6,7,3,23,0,1,1,24,8,2,7,68000
5	Remote_Work,Team_Size,Manager_Rating,Peer_Rating,	6,45,0,4,20,4,4,9,4,1,50,8,9,9,60,3,1,3,4,3,9,8,0,15,1,12,9,9,15000,2,10000,9,6,3,9,8,8,4,28,1,2,1,48,9,1,5,98000
6	Annual_Bonus,Stock_Options,Incentives,Appraisal_Score,	7,26,1,2,3,1,1,5,35,1,3,10,2,2,8,2,1,45,4,7,8,30,1,1,2,3,2,8,7,1,30,1,7,8,7,9000,1,6000,8,3,1,7,6,7,3,23,0,1,
7	Years_In_Current_Role,Years_Since_Last_Promotion,Market_Demand,	8,38,0,3,12,2,1,24,8,2,7,68000
8	Company_Revenue_Growth,Inflation_Rate,Cost_Of_Living,	9,50,1,5,25,5,5,9,5,1,55,9,10,10,70,4,1,3,5,4,10,9,0,10,1,15,10,9,20000,3,15000,10,8,4,9,9,9,5,30,1,2,1,60,10,1,4,120000
9	Tax_Rate,Union_Member,Shift_Type,	10,30,0,3,6,2,2,7,1,0,42,3,6,7,25,1,1,2,3,2,7,6,2,32,0,7,7,6,6000,0,4000,7,2,0,6,5,6,2,21,0,1,0,18,7,2,8,52000
10	Contract_Duration,Employee_Engagement,	$null$

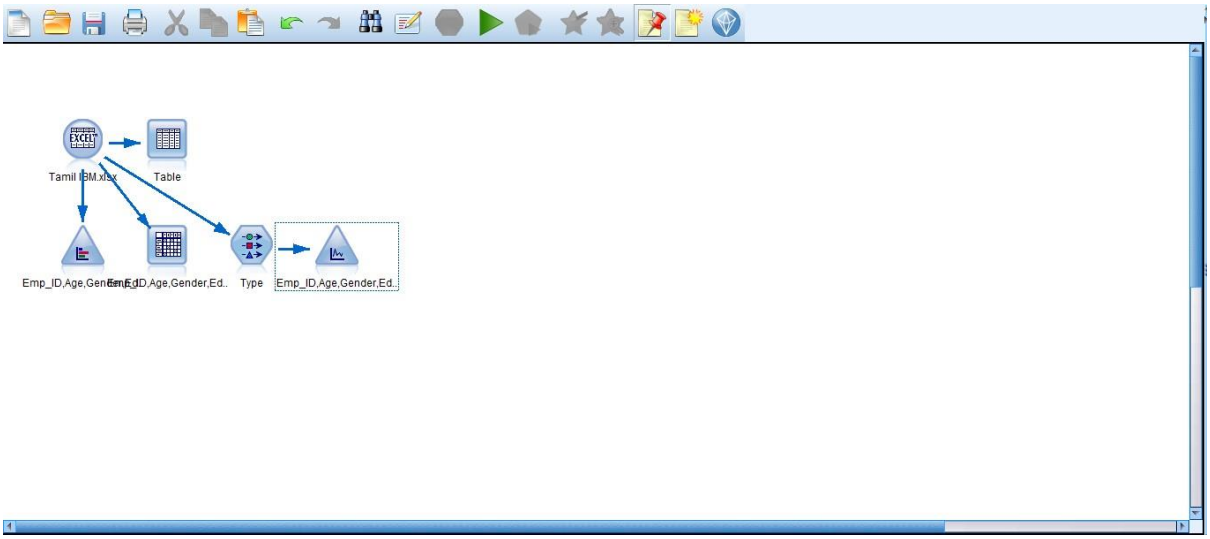

Distribution of Emp_ID,Age,Gender,Education_Level,Experience_Years,Jo...
File
Edit
Generate
View
Table
Graph
Annotations
Value
Proportion
%
Count
Annual_Bonus,Stock_Options,Incentives,Appraisal_Score, 9.09 1
Company_Revenue_Growth,Inflation_Rate,Cost_Of_Living, 9.09 1
Contract_Duration,Employee_Engagement, 9.09 1
Employment_Type,Location,Company_Size,Industry_Type, 9.09 1
Job_Satisfaction,Work_Life_Balance,Absenteeism,Commute_Time, 9.09 1
Leadership_Score,Skill_Score,Training_Hours,Promotion_Count, 9.09 1
Performance_Score,Certifications,Overtime,Work_Hours,Projects_... 9.09 1
Remote_Work,Team_Size,Manager_Rating,Peer_Rating, 9.09 1
Retention_Risk,Turnover_Rate,Salary 9.09 1
Tax_Rate,Union_Member,Shift_Type, 9.09 1
Years_In_Current_Role,Years_Since_Last_Promotion,Market_Dem... 9.09 1
1,25,1,2,2,1,1,7,1,0,40,3,6,7,20,0,1,1,2,1,7,6,2,30,0,6,7,6,5000,0,3000,7,1,0,6,5,6,2,20,0,1,0,12,7,
10,30,0,3,6,2,2,7,1,0,42,3,6,7,25,1,1,2,3,2,7,6,2,32,0,7,7,6,6000,0,4000,7,2,0,6,5,6,2,21,0,1,0
2,32,0,3,7,2,2,8,2,1,45,5,7,8,35,1,1,2,3,2,8,7,1,25,1,8,8,7,8000,1,5000,8,3,1,7,6,7,3,22,0,1,1,
OK


Distribution of Emp_ID,Age,Gender,Education_Level,Experience_Years,Jo...
File
Edit
Generate
View
Table
Graph
Annotations
Performance_Score,
Certifications,
Overtime,
Work_Hours,
Projects_Handled
Years_In_Current_Role,
Years_Since_Last_Promotion,
Market_Demand,
Contract_Duration,
Employee_Engagement,
0.0
0.2
0.4
0.6
0.8
1.0
Count
1,25,1,2,2,1,1,7,1,0,40,3,6,7,20,0,1,1,2,1,7,6,2,30,0,6,7,6,5000,0,3000,7,1,0,6,5,6,2,20,0,1,0,12,7,2,8,45000
2,32,0,3,7,2,2,8,2,1,45,5,7,8,35,1,1,2,3,2,8,7,1,25,1,8,8,7,8000,1,5000,8,3,1,7,6,7,3,22,0,1,1,24,8,1,7,62000
4,28,0,2,4,1,1,6,0,0,40,2,5,6,15,0,1,1,2,1,6,5,3,40,0,5,6,5,4000,0,2000,6,1,0,5,4,6,2,18,0,1,0,12,6,3,9,42000
...
OK


Matrix of Emp_ID,Age,Gender,Education_Level,Experienc...
File
Edit
Generate
Matrix
Appearance
Annotations
1,25,1,2,2,1,1,7,1,0,40,3,6,7,20,0,1,1,2,1,7,6,2,30,0,6,7,6,5000,0,3000,7,1,0,6,5,6,2,...
Emp_ID,Age,Gender,Education_Level,Experience_Years,Job_Level,Department,
Annual_Bonus,Stock_Options,Incentives,Appraisal_Score,
Company_Revenue_Growth,Inflation_Rate,Cost_Of_Living,
Contract_Duration,Employee_Engagement,
Employment_Type,Location,Company_Size,Industry_Type,
Job_Satisfaction,Work_Life_Balance,Absenteeism,Commute_Time,
Leadership_Score,Skill_Score,Training_Hours,Promotion_Count,
Performance_Score,Certifications,Overtime,Work_Hours,Projects_Handled,
Remote_Work,Team_Size,Manager_Rating,Peer_Rating,
Retention_Risk,Turnover_Rate,Salary
Tax_Rate,Union_Member,Shift_Type,
Cells contain: cross-tabulation of fields (including missing values)
Chi-square = 99, df = 90, probability = 0.242
OK



Final output



Conclusion

The Employee Salary Forecasting project using Linear Regression in IBM SPSS Modeler successfully demonstrated how statistical modeling can be used to predict employee salaries based on key factors such as age, experience, education, and other relevant variables. The linear regression model identified significant predictors influencing salary and established a measurable relationship between independent variables and salary outcomes.

The model showed reliable performance with acceptable accuracy levels, indicating that linear regression is an effective method for salary prediction when the relationship between variables is linear. The results can help organizations make data-driven decisions in workforce planning, budgeting, and compensation management.

Overall, this project highlights the importance of predictive analytics in human resource management and demonstrates how IBM SPSS Modeler can be effectively used to build, evaluate, and interpret forecasting models for real-world business applications.

REFERENCE: