

# SALARY PREDICTION USING DUMMY AND CONTINUOUS VARIABLES IN EXCEL

## 1. Background and Problem Statement:

An HR consulting agency managing compensation benchmarking for multiple corporate clients wanted a flexible, Excel-based salary estimation tool that considers both **quantitative factors** (like experience) and **categorical variables** (like education and department). The goal was to create a transparent and data-driven approach to salary prediction without using complex statistical software.

## 2. Objectives:

- Build a multiple linear regression model in Excel using both **continuous and dummy variables**
- Estimate salary based on experience, education level, and department
- Quantify the impact of each variable on salary
- Develop an interactive Excel calculator for HR teams to simulate salary ranges

## 3. Methodology:

### 3.1 Dataset Overview

- **Sample Size:** 250 employees from 5 departments
- **Variables Collected:**
  - Salary (in INR lakhs per annum)
  - Experience (in years)
  - Education Level: Postgraduate or Not
  - Department: IT, HR, Sales, Finance, Operations

### 3.2 Data Preparation

- Created **dummy variables**:
  - Education\_PostGrad (1 = Postgraduate, 0 = Others)
  - Department dummies: IT, HR, Sales, Finance (Operations used as reference)

- Removed outliers using interquartile range filters
- Structured all input variables in Excel tables with named ranges

### 3.3 Regression Execution in Excel

- Used **Excel Data Analysis Toolpak → Regression**
- Dependent Variable: Salary
- Independent Variables: Experience, Education\_PostGrad, Dept\_IT, Dept\_HR, Dept\_Sales, Dept\_Finance
- Verified regression assumptions using residual scatter plots and p-values

## 4. Results and Interpretation:

### 4.1 Regression Coefficients

Variable	Coefficient	p-Value	Interpretation
Intercept	4.15	0.001	Base salary for Operations dept with no PG, no experience
Experience	0.43	<0.001	Each additional year of experience adds ₹43,000
Education_PostGrad	1.95	0.012	PG employees earn ₹1.95L more on average
Dept_IT	3.20	0.002	IT earns ₹3.2L more than Ops, all else constant
Dept_HR	0.90	0.042	HR earns ₹0.9L more than Ops
Dept_Sales	1.15	0.030	Sales earns ₹1.15L more than Ops
Dept_Finance	2.25	0.005	Finance earns ₹2.25L more than Ops

- **$R^2 = 0.81$**
- All variables significant at 5% level
- No multicollinearity observed among dummies (checked by manual  $R^2$  for each predictor)

### 4.2 Key Insight

- **Education and department type significantly influence salary**, even after accounting for experience

- IT and Finance are top-paying departments
- PG qualification adds notable premium

## 5. Deliverables in Excel:

- **Salary Estimator Tool**
  - User selects department and education level from dropdown
  - Inputs years of experience
  - Predicted salary auto-generated using regression formula
- Summary tables:
  - Regression coefficients
  - Interpretation sheet
- Visualizations:
  - Salary distribution by department
  - Bar charts of predicted salary by education and department

## 6. Recommendations:

- Use this model to **audit pay equity** across departments
- Reassess compensation strategy for Operations department
- Encourage upskilling/PG education among HR and Sales staff to bridge gaps
- Refresh model quarterly to adjust for market shifts

## 7. Stakeholder Relevance:

### Academic:

- Demonstrates multiple linear regression using both continuous and dummy variables
- Introduces Excel modeling for HR analytics with real-world structure

### Corporate:

- Offers a scalable Excel solution for salary benchmarking and budgeting
- Builds trust through transparent, formula-driven predictions