



Internship Project Report

Title: Data Analyst Job Market Analysis

Internship Organization: Unified Mentor Pvt. Ltd.

Intern Name: Himanshu Kaushik

Domain: Data analyst Intern

Duration: 3 months

Tools & Technologies Used: Python, Pandas, NumPy, Matplotlib, Seaborn, Plotly, Scikit-learn, Excel, SQL

Dataset Source: unified mentor



1. Objective

This project analyzes job listings for data analyst roles to:

- Explore salary distribution, trends, and job title demands.
 - Predict salary estimates based on company size, rating, and tech skills.
 - Identify patterns in job requirements across sectors and locations.
-



2. Dataset Overview

- **Source:** Unified Mentor Pvt. Ltd.
 - **Size:** 2253 rows × 13 columns.
 - **Features Include:**
 - Job Title
 - Salary Estimate
 - Job Description
 - Company Rating
 - Company Name
 - Location & Headquarters
 - Size, Type of Ownership
 - Industry, Sector, Revenue
 - Easy Apply Option
-



3. Data Exploration & Cleaning

- Removed unnecessary columns (Unnamed, Founded, Competitors).
- Handled missing values using:
 - Median for numerical (Rating)
 - Forward fill for categorical columns

- Cleaned and extracted:
 - Min Salary, Max Salary → computed Avg Salary
 - Parsed Job Title for standardization
 - Extracted City and State from Location
 - Renamed columns for consistency
-

4. Exploratory Data Analysis (EDA)

Top 5 Most Frequent Job Titles:

Job title	counts
Data analyst	405
Senior Data Analyst	131
Junior Data Analyst	58
Business Data Analyst	28
Data Quality Analyst	17

Salary Distribution:

- **Avg Salary Range:** \$60K – \$80K
- **Max Observed Salary:** Above \$140K
- **Top Paying Role:** Lead Data Analyst

Location Insights:

- **Most Jobs Posted In:** New York, NY
- **Top 5 Highest Salary Cities:** Newark, CA; Daly City, CA; Marin City, CA; Los Gatos, CA; Pico Rivera, CA

Company Size & Salary:

- Most listings: Companies with 51–200 employees
- No strong correlation between company size and average salary

Company Ratings:

- Ratings range between **3.0 – 4.0**, indicating moderate satisfaction

Type of Ownership:

- Most companies are **Private** followed by **Public**

Sector Insights:

- **Top Hiring Sectors:**
 1. Information Technology
 2. Business Services
 3. Finance
 4. Healthcare

5. Education
- **Top Paying Sectors:**
1. Biotech & Pharmaceuticals
2. Real Estate
3. Entertainment & Recreation
4. Legal
5. Information Technology

5. Feature Engineering

- Binary encoding for tech skills (Python, Excel) from job descriptions
- Created a Tech_Skills score combining skill flags
- Encoded company size, location for modeling

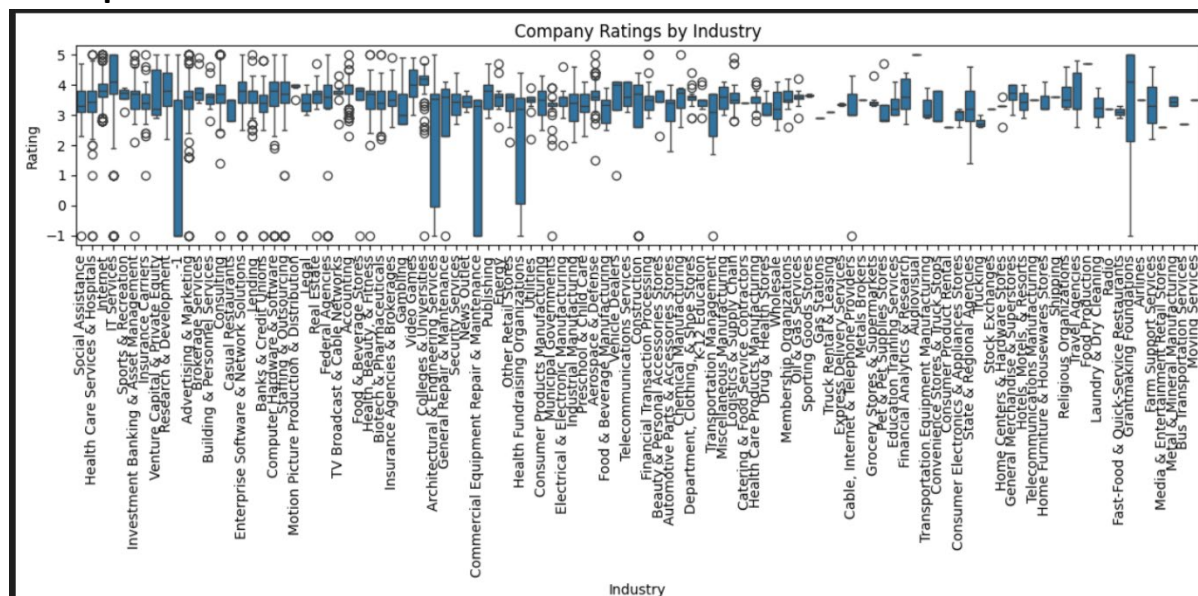
6. Salary Prediction Model

- **Model Used:** Random Forest Regressor
- **Features:** Rating, Tech_Skills, Size, Year Founded

7. Conclusion

- There's high demand for Data Analysts with proficiency in Python and Excel.
- The average pay is promising, especially in California and sectors like Biotech.
- Salary depends more on skills and industry than company size.
- The model built helps estimate expected salary, aiding job seekers and HR planning.

8. Snapshots



Salary Estimate Distribution

