7COM1079-2024 Student Group No: A21

Names of Student Attendees:

1) Dasari Purna Satesh
2) Bhanu PrakashReddy
3) Rahul Dulla
4) B. Upendra Yadav
5) Chandra Sekhar Chintapalli

Group Name:

Name of Student Presenting: Dasari Purna Satesh

Research Question –

Tutorial Presentation for Feedback

Date: 17th November 2024

University of Hertfordshire UH

Team Research Project - Data Analyst Jobs

University of Hertfordshire UH

Dataset Overview

Dataset Source: Provided from Kaggle.com.
Dataset Name:Data Analyst Jobs
Dataset Code: DS178

Introduction of dataset:

This dataset contains job postings for data analyst positions, including information such as **job titles, salary estimates, company ratings, location, industry, and size**. It helps explore trends in job market demand, salary variations, and the skills required for data analyst roles. By analyzing company details like revenue and competitors, the dataset offers a comprehensive view of the data analyst job landscape across various sectors and regions.

Dataset Information

Variables:

Independent Variables: Job Title, Location, Size, Industry, Type of ownership.
Dependent Variable: Salary Estimate.
Dataset Size: Total Rows: 2253

Data Snapshot:

	Job Title	Location	Size	Industry	Type of ownership	Salary Estimate
0	Immigration and Justice (CIJ)	New York, NY	201 to 500 employees	Social Assistance	Nonprofit Organization	37K – 66K (Glassdoor est.)
1	Quality Data Analyst	New York, NY	10000+ employees	Health Care Services & Hospitals	Nonprofit Organization	37K – 66K (Glassdoor est.)
2	Team [Customer Operations]	New York, NY	1001 to 5000 employees	Internet	Company - Private	37K – 66K (Glassdoor est.)
3	Data Analyst	New York, NY	201 to 500 employees	IT Services	Subsidiary or Business Segment	37K – 66K (Glassdoor est.)
4	Reporting Data Analyst	New York, NY	501 to 1000 employees	Sports & Recreation	Company - Private	37K – 66K (Glassdoor est.)

Research Objective & Question

- **Analyze the salary variations** based on different **job titles**.
- **Compare salary estimates** across different **locations** (cities or countries).
- Investigate whether **size** affects salary estimates for data analyst positions.
- Assess how **industry** and **type of ownership** influence salary differences.
- Provide insights for job seekers to understand salary trends based on job title and company factors.

Research Question:

Is there a difference in the **Salary Estimate** (dependent variable) based on **Job Title**, **Location**, **Size**, **Industry**, and **Type of Ownership** (independent variables)?

Hypothesis

Null Hypothesis (H_0):
There is no significant difference in **Salary Estimate** based on **Job Title, Location, Company Size, Industry, or Type of Ownership**.

Alternative Hypothesis (H_1):
There is a significant difference in **Salary Estimate** based on **Job Title, Location, Company Size, Industry, or Type of Ownership**.

Note: After performing statistical analysis, we will either accept or reject the null hypothesis based on p-values.

University of Hertfordshire UH

Analysis Methods

Comparison of Means: Analyze the average salary differences across Job Titles, Locations, and other factors.

Correlation: Study relationships between Salary Estimate and variables like Company Size or Experience.

Comparison of Proportions: Evaluate how salary proportions differ across categories..

Statistical Tools: Python (Pandas, NumPy, SciPy) will be used for data manipulation, statistical analysis, and hypothesis testing.

Expected Output

- **Salary Differences**: Identifying significant variations in **Salary Estimates** across different **Job Titles**, **Locations**, **Company Size**, **Industry**, and **Type of Ownership**.
- **Factors Impacting Salary**: Determining which variables (such as **Job Title**, **Industry**, **Location**) significantly affect salary levels.
- **Insight for Job Seekers**: Providing actionable insights to help job seekers make informed decisions about **job titles**, **company types**, and **locations** based on salary trends.
- **Statistical Findings**: Statistical tests (e.g., ANOVA) will highlight which factors are most influential.