Salary About the Data Related Job Investigation

Project Background

This study analyzes data-related jobs by extracting 969 job records from a dataset containing over 6,500 records. The goal is to provide insights into the number of positions, educational requirements, years of prior work experience, and salaries for various data-related roles. This analysis aims to help individuals interested in the data industry gain a better understanding of the job market and receive tailored recommendations for their career planning and job selection based on their backgrounds.

The dataset is sourced from Kaggle and is compiled from multiple sources, including surveys, job posting sites, and other publicly available data.

Executive Summary

This study analyzes **969 data-related job records**, focusing on **Leadership, Data Analyst, and Data Science roles** due to data limitations in other categories.

Data Science has the highest representation (**515 individuals**) with an **average salary of $164,099**, primarily held by **PhD graduates (71.26%)**. **Data Analysts (391 individuals)** earn an **average of $120,606**, with most holding **a Bachelor's degree (79.8%)**. **Leadership roles (58 individuals)** have the **highest salary at $204,827**, with **100% holding a PhD**.

Salary growth is tied to experience, with **Data Scientists progressing from $108,636 (5 years) to $240,000 (25 years)**. Promotions in **Data Analysis** show a **200% salary increase** from **Junior to Senior roles**. **Leadership salaries rise steadily with experience, reaching $220,000 at 20 years**.

The **most common job title is Data Scientist (453 records), followed by Data Analyst (363 records)**. Senior and leadership roles appear less frequently but offer **higher earning potential**.

Data Preparation

Null Data

Since the data was transferred from Excel to the SQL dataset, all empty rows were removed beforehand. Therefore, this analysis is based on a clean dataset containing **6,699 entries**.

Correcting Import Name Errors

During the data import process, an issue occurred with the `age` column due to encoding errors, resulting in garbled text. This error was corrected, and the feature name was restored to its original form, age.

Inconsistent Values

While inspecting the `Education Level` column, inconsistent values were identified where different terms represented the same meaning, such as "Bachelor's Degree" and "Bachelor's". To ensure consistency, these values were standardized across the dataset.

Duplicate Entries

Upon checking the dataset, **4,913 duplicate entries** were detected, accounting for approximately **70%** of the data. Due to the high proportion, these entries could not be removed without compromising the integrity of the analysis. Although duplicates should not exist in a clean dataset, they have been retained for this analysis and treated as reference data.

Insights Deep-Dive or Summary of Insights

1. **Educational Distribution Across Job Functions**
    - PhD degrees are dominant in high-paying roles, particularly in Leadership (100%) and Data Science (71.26%).
    - Data Analysts typically hold a bachelor's degree (79.8%), making it the most accessible entry-level role.
2. **Salary Growth and Career Progression**
    - Experience significantly impacts salary growth:
        - Data Scientists' salaries grow from $108,636 (5 years) to $240,000 (25 years).
        - Senior Data Analysts earn 3 times more than Junior Analysts, with an increase from $51,360 to $155,000.
        - Chief Data Officers earn $220,000 with an average of 17.67 years of experience.
    - Leadership salaries increase steadily, reaching $220,000 in 20 years.
3. **Most Common Job Titles and Their Trends**
    - Data Scientist (453 records) and Data Analyst (363 records) are the most frequent roles, indicating high market demand.
    - Senior and leadership positions appear less frequently but command higher salaries and require extensive experience.

Recommendations

Based on the analysis, the following recommendations are provided to help individuals navigate career paths in data-related fields. These insights can assist in making informed decisions regarding education, skill development, and long-term career planning.

- **Establish a Strong Foundation in Data Analysis:**
  Pursuing a bachelor's degree is the most common entry point into Data Analyst roles, offering a solid foundation for career growth. This path provides an accessible entry into the industry with opportunities for progression.
- **Leverage Advanced Degrees for Data Science and Leadership:**
  Data Science roles often require a PhD, but they offer higher earning potential from the start. Mid-career professionals should consider upskilling through a Master's or PhD to transition into Data Science or Leadership positions, which command higher salaries and greater career stability.
- **Maximize Career Growth Through Experience and Promotions:**
  Gaining 5+ years of experience can significantly increase earning potential in Data Science ($108,636 → $164,099). For long-term career advancement, Leadership roles offer the highest salaries ($204,827+), but typically require 15+ years of experience and a PhD.
- **Capitalize on Market Demand and Strategic Job Planning:**
  Data Science and Data Analyst roles are in high demand, making them strategic career choices for professionals entering the field. Senior and leadership positions are less frequent but offer lucrative salaries, emphasizing the importance of long-term career planning and strategic role transitions.

Assumptions and Caveats

- The analysis assumes that **salary trends are influenced primarily by education level and years of experience**, though other factors such as company size, location, and industry specialization may also contribute.

- The dataset includes **969 job records** extracted from a larger dataset of **6,500+ records**, which may not fully represent all data-related job markets.

- **Leadership and Data Engineering roles are underrepresented**, with only **58 and 4 records**, respectively, which may limit the accuracy of insights in these categories.

- Salary progression is based on **averages across job titles and experience levels**, and **individual career trajectories may vary** depending on specific job responsibilities and company policies.

Technical Project Information

The dataset used for this analysis was sourced from Kaggle, containing data from multiple sources, including surveys, job postings, and publicly available records. The dataset includes details on job titles, salaries, education levels, years of experience, and job functions.

Data cleaning and transformation were performed in SQL, including handling missing values and standardizing job categories. The final dataset was then analyzed and visualized using Tableau, providing key insights into salary distribution, career progression, and job market trends.