**Stamic Groups 2019 Dashboard**

**1. Introduction**

**Purpose**

The primary purpose of the Stamic Groups 2019 Dashboard is to evaluate job trends in data-related professions by analyzing critical factors such as salaries, experience levels, employment types, and company sizes. This comprehensive analysis aims to provide a clear understanding of the job market dynamics within the data industry.

**Objective of the Project**

The objective of this project is to identify which job roles, experience levels, and company sizes yield the highest salaries and most opportunities. By doing so, the dashboard serves as a valuable tool for job seekers and HR professionals, aiding in strategic decision-making regarding career paths and recruitment strategies.

**Problem Being Addressed**

Job seekers and recruiters often lack access to data-driven insights into the job market landscape, particularly in data-driven fields. This analysis addresses the need for clarity on compensation, job type demand, and employment structures, enabling both job seekers and employers to make informed decisions.

**Key Datasets and Methodologies**

* Datasets Used: The analysis utilizes a variety of datasets, including job roles, salary data, experience levels, employment types, company sizes, and geographic locations.
* Excel Methods: Various Excel functionalities are employed, including Pivot Tables, Bar and Donut Charts, and Filters/Slicers for interactive visualization, to present the data in an accessible manner.

**Story of Data**

**Purpose**

The purpose of this section is to explore and visualize job market dynamics for technical roles in 2019, providing insights into the landscape of data-related professions.

Data Source

Data was aggregated from multiple sources, including online job boards, HR databases, and company salary disclosures, ensuring a comprehensive representation of the job market.

Data Structure

* Rows: Each row represents a unique job listing.
* Columns: Key variables include job title, salary, company size, employment type, location, and experience level.

**Important Features and Their Significance**

* Job Title: Specifies the technical role, allowing for targeted analysis.
* Salary: A key metric for understanding compensation trends across roles.
* Company Size: Reflects the structure of employers (Small, Medium, Large), which can influence salary offerings.
* Experience Level: A critical factor affecting expected salary and job opportunities.
* Employment Type: Includes Full-Time (FT), Part-Time (PT), Contract (CT), and Freelance (FL), providing insights into job stability and preferences.
* Location: Essential for analyzing geographic salary variations and job availability.

**Data Limitations or Biases**

* The dataset is limited to 2019, which restricts year-over-year trend analysis.
* There may be underrepresentation of certain employment types, particularly Part-Time and Freelance roles.
* The geographical focus is primarily on the United States, with minimal representation from global markets.

**3. Data Splitting and Preprocessing**

Data Cleaning

* Duplicate Entries: Removed to maintain data integrity.
* Standardization: Job titles and experience codes were standardized (e.g., SE for Senior, MI for Mid-level, EN for Entry-level, EX for Executive).

**Handling Missing Values**

* Missing values were omitted from aggregation to preserve the integrity of the analysis.
* Filters were applied to exclude undefined data points, ensuring clarity in the visualizations.

**Data Transformations**

* Aggregation: Salaries were aggregated by job title, company size, and location to facilitate comparative analysis.
* New Fields: Created fields such as “Experience Level and Salary” to enhance analytical depth.

**Data Splitting**

* Dependent Variables: Salary and Job Count.
* Independent Variables: Job Title, Company Size, Location, and Experience Level.

**Industry Context**

The focus of this analysis is within the Tech/Data Industry, highlighting roles such as Data Engineer, Data Scientist, Machine Learning Engineer, and others.

**Stakeholders**

**Key stakeholders include:**

* Job Seekers: Individuals seeking opportunities in data-related fields.
* HR Departments: Professionals responsible for hiring and talent acquisition.
* Career Counselors: Advisors guiding job seekers in their career paths.
* Workforce Analysts: Individuals analyzing labor market trends and dynamics.

**Value to the Industry**

This analysis supports talent acquisition strategies and career path optimization for data professionals, helping organizations attract and retain top talent.

**4. Pre-Analysis**

**Purpose**

To provide initial insights into job role prevalence and salary ranges within the data industry.

Identify Key Trends

* Most In-Demand Role: Data Engineer is identified as the highest-paid and most sought-after position in the market.
* Employment Type Dominance: Full-Time (FT) positions dominate the job market with a total of 16,454 job listings.
* Experience Level Insights: Senior (SE) level positions yield the highest salaries and opportunities.

**Potential Correlations**

* A higher experience level (SE, MI) correlates positively with higher salaries.
* Larger company sizes (Medium and Large) are associated with better compensation packages.
* The United States location correlates with the majority of high-salary job opportunities.

**Initial Insights**

* There is a clear emphasis on the value of experience, with senior roles commanding higher salaries.
* Full-Time jobs are overwhelmingly preferred and employed, indicating job stability.
* Emerging roles, such as Machine Learning Engineer and Analytics Engineer, still have lower representation in the market.

**5. In-Analysis**

**Purpose**

This section focuses on deeper pattern discovery and validation of early insights derived from the data.

Unconfirmed Insights

* Roles such as Data Analyst and Machine Learning Engineer show moderate earnings but may see increases in compensation as demand grows.
* Small companies (S) display lower salary distributions compared to their medium and large counterparts.

**Recommendations**

* For Job Seekers: Target Data Engineering roles and aim for Full-Time and Senior positions to maximize salary potential and job opportunities.
* For HR Professionals: Focus on retaining senior professionals and actively recruiting for high-demand roles such as Data Engineer and Data Scientist.

**Analysis Techniques Used in Excel**

* Pivot Tables: Utilized for summarizing salary data by experience level and job title.
* Bar Charts: Employed to visualize job counts and experience level distributions.
* Donut Charts: Used to illustrate company size distribution in the job market.
* Slicers: Implemented to enable interactive filtering by location, salary, and job level for more granular analysis.

**6. Post-Analysis and Insights**

**Purpose**

To finalize and summarize the data-backed conclusions drawn from the analysis.

**Key Findings**

* Top Job: Data Engineer, with an average salary of 3464 units, stands out as the most lucrative position.
* Best Job Type: Full-Time roles dominate the market with 16,454 listings.
* Most Experience Demand: Senior Level positions account for 10,670 jobs, indicating a strong demand for experienced professionals.
* Top Company Size: Medium-sized firms dominate salary distributions, providing competitive compensation.
* Top Location: The United States leads with 14,518 job listings, reflecting a concentrated job market for data roles.

**Comparison with Initial Findings**

Initial assumptions regarding the prominence of Data Engineers and Full-Time roles were strongly validated. Additionally, the importance of experience level was found to be more substantial than initially expected, emphasizing the need for skilled professionals in the industry.

**7. Data Visualizations & Charts**

Purpose

To communicate insights clearly and intuitively through visual representations of the data.

Charts and Graphs

* Line Graph: Displays the top 10 jobs by salary, illustrating the salary hierarchy within the data industry.
* Bar Charts: Show employment type distribution, experience levels, and job distribution by company size, providing a comprehensive overview of market dynamics.
* Donut Chart: Breaks down salary distributions by company size, highlighting the contribution of different employer types to overall compensation.

**8. Recommendations and Observations**

Actionable Insights

* For Candidates: Focus on pursuing Data Engineer roles, seek Full-Time opportunities at medium to large firms, and aim to build toward a Senior experience level for optimal career growth.
* For Employers: Offer competitive salaries for Senior roles and recognize that Full-Time contracts are favored by the job market.

**Optimizations or Business Decisions**

* Resource allocation should prioritize hiring for Senior positions to retain top talent.
* Invest in career growth programs to support the development of high-potential mid-level employees, ensuring a robust talent pipeline.

**Unexpected Outcomes**

* Despite the hype surrounding Machine Learning and Research Engineer roles, these positions lag in compensation compared to more established roles.
* Contract and freelance roles (CT, FL) show negligible volume in the job market, suggesting a preference for stable employment types.

**9. Conclusion**

**Key Learnings**

* Experience and job type are the strongest influencers on salary within the data industry.
* The Data Engineer role is identified as the most valued position in 2019, reflecting industry demand.
* Company size and geographic location significantly impact pay structures, with larger firms generally offering better compensation.

**Limitations**

* The analysis lacks time-series or growth trend analysis, limiting insights into market changes over time.
* Geographic diversity is limited, with a primary focus on the US job market.
* There is a lack of role-specific competency data or performance KPIs, which could provide additional context for salary variations.

**Future Research**

* Future analyses should include year-over-year growth for better trend tracking and market forecasting.
* Expanding the dataset to incorporate benefits and perks would provide a more holistic view of job compensation.

**References**

* Data sources include hypothetical job platforms such as LinkedIn, Glassdoor, and Indeed, which provide valuable insights into job market trends.
* Excel dashboarding and analysis techniques employed include Pivot Tables, Filters, and Chart Tools for effective data visualization.

**Appendices**

* Pivot Table Summary Sheets: Detailed summaries of the data analysis.