International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 2, Mar-Apr-2026, ISSN (Online): 2395-566X

Exploring Trends in Job Postings and Salaries Across Different Industries

Ms.C. Harivarshini[1], Ms.M. Shubhashree[2], Dr.R. Karthik[3]
[1]Student of III BSC Data Science, Sri Krishna Adithya College of Arts and Science, Kovaipudur, Coimbatore-641042
[2]Assistant Professor of Data Science Department, Sri Krishna Adithya College of Arts and Science, Kovaipudur, Coimbatore-641042,

Abstract: The global workforce is undergoing rapid evolution. Current data-driven research into worldwide employment trends thus has become a pressing need. The objective of the present study was to conduct a comprehensive analysis of job advertisements and salary trends by reviewing 999 job records collected from 213 countries; these included a total of 13 data points. As part of its analytic process, the present study utilized a data preprocessing pipeline that involved the passing of data through multiple stages – data cleansing, data type conversion, aggregation, data partitioning, normalization, etc., prior to submission to various data visualisation techniques; these included bar charts, histograms, box plots, scatter plots, correlation heat maps, skills frequency heat maps, pie charts, violin plots, and choropleth maps. Among the most significant findings of the present study were the following: job advertisements show evidence of consistent salary levels based on both level of education and type of job; however, geographic region and industry sector appear to play an important role in determining salary levels. Additionally, the study concluded that the most common skills necessary for attaining such salaries are as follows: management skills; analytical skills; design skills; communication skills; and technical/data oriented skills. Consequently, the authors propose a framework that can be used to better understand the trends of employment and provide actionable insights into the employment market.

Keywords — Job Market Analysis, Data Preprocessing, Salary Trends, Data Visualization, Employment Analytics, Python, Machine Learning Preprocessing

I. INTRODUCTION

Over the past 30 years, we've seen several changes that have occurred in our economy. One of the biggest things we've seen happen is the World Wide Web's transformation of how people communicate with one another. We also see that technology continues to evolve and impact how we work. Through globalisation, many regions have been able to participate in the global economy than we've ever seen before and at a rate we've never seen. The types of changes we're seeing in how companies are structured and how they can hire workers from all over the world are now presenting a whole slew of new challenges to those individuals looking for work via job postings available on the internet. Both of these things combined offer new challenges for people to use structured data to analyse labour markets within the framework of data-oriented analysis; however, they also provide new challenges to derive useful analyses from structured data analysis using organised scientific methods.

Understanding the various patterns related to job postings is very important to all stakeholders. For the job seeker, learning about wage patterns across different qualifications, jobs, and regions is critical in making constructive career choices. Employers gain valuable knowledge about how to compete for talent by also knowing different skill-set wage patterns. For governments, geographical wage and opportunity disparities allow them to create appropriate educational investments.

This research is an in-depth exploratory assessment of a curated job posting dataset consisting of 999 job postings from 213 countries, including 13 primary characteristics of transactional data, users, and organizations. This study has multiple objectives: to appropriately process the dataset for the purpose of analysis, to develop transformations that facilitate comparative assessments of the datasets, to create a comprehensive set of visual representations of the datasets in terms of salaries, geographical areas, types of

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 2, Mar-Apr-2026, ISSN (Online): 2395-566X

jobs performed, and required skill sets, and to identify the primary drivers of salaries and job opportunities.

The remaining parts of this article contain the following information: Section II describes details of our data set including its distinctive characteristics; Section III describes the methodology used to preprocess the data; Section IV describes a software suite used to visualize results produced from analyzing our data set; Section V contains a discussion on the results produced from analyzing our data set including their potential impact upon future research; and Section VI summarizes the above content as well as provides an outline of future work related to this topic.

II. DATASET DESCRIPTION

Data were collected from publicly available job posting databases, yielding 999 records of unique job postings. The 999 records contain 13 key attributes representing a variety of characteristics related to transactional, user, and organizational aspects of job postings. A summary of the dataset's attributes and corresponding descriptions is presented in Table I.

1. **Job Id –** Represents the unique identifier assigned to each job posting.
2. **Experience –** Indicates the required experience range for the job position (e.g., 2–12 years).
3. **Qualifications –** Refers to the minimum educational qualification required for the job such as MBA, PhD, or BCA.
4. **Salary Range –** Describes the salary bracket offered for the position (e.g., $59K–$99K).
5. **Country –** Specifies the country where the job is located.
6. **Work Type** – Refers to the type of employment such as Intern, Full-Time, Part-Time, Contract, or Temporary.
7. **Company Size –** Represents the total number of employees working in the company.
8. **Preference –** Indicates the gender preference for the job role, if specified.
9. **Job Title –** Refers to the name of the job position.
10. **Role –** Describes the specific role or specialization within the job title.
11. **Skills –** Represents the skills required to perform the job effectively.
12. **Responsibilities –** Describes the main duties and tasks associated with the role.
13. **Company** – Indicates the name of the organization offering the job.

All of the attributes have 999 entries that are non-null (no missing values). The original data appears to be missing no data at this time. The dataset represents a global sample of employment distribution data with at least 213 countries included geographically. The x values that make up the attribute of "salary range" appear as the string "$min−$max", so those values require more suitable parsing for further analysis using a more standard format. The values for "company size" in the data set range from 12,677 to 134,519.

III. DATA PREPROCESSING

Data Cleaning

The foundation of the data preprocessing pipeline resulted from the data cleaning phase. During this stage, we attempted to accomplish the goal of eliminating quality problem areas in order to ensure that the analysis results can be valid without any issues associated with the dataset of 999 non-null records, zero missing records across all 13 attributes and by eliminating any possible duplicate job postings from the dataset.

In addition, Company Size records were reviewed for outlier conditions. Company Size records greater than 500,000 employees were identified as statistical outliers, and those records were deleted from the dataset. Candidate Qualification records, Country records, and Job Title records were also standardized due to inconsistencies in capitalization and spaces within these datasets (resulting in the presence of counterfeit categories after aggregating and grouping records).

Conversion of Data Types

The Salary Range attribute, which was originally stored as a string in the format of "$min-$max," was converted to a numeric data type using regular expression extraction to create two new independent columns, the Min Salary and Max Salary, as numeric data types that can be used for quantitative statistical analysis of salaries. The Min Salary's range is $55,000 to $65,000; the average is $59,931; as a result, the Min Salary has a relatively thin and highly concentrated salary range. The Company Size attribute was also converted from a string to a numeric data type.

Data Aggregation

The analysis included numerous aggregation calculations made to provide high level comparison analysis across categorical variables. Aggregation was performed on a groupby operation by country for the collection and summarization of job postings and mean salary for each of the 213 countries contained in the data file. Additional aggregation aggregations were performed for the Work Type variables. Aggregation of Postings and nominal salary range were carried out using the Country and Qualification variables while creating cross dimensional analysis of salary range as a function of employment type. The aggregation of total job postings by employment type contained as part of the Work Type aggregate found that the largest percentage of employment types consisted of Part-Time at 22.0%, Temporary at 19.9%, Contract at 19.8%, Full-Time at 19.4% and Intern at 18.8% This finding indicates that there is a relatively uniform distribution over the various employment types.

Data Division

To build prediction models with analytics, the dataset was randomly divided into 80/20 training and testing datasets, with the training dataset containing 799 records/ observations and the testing dataset containing 200 records/observations. Additional categorical splits were made by 'Country' and 'Qualifications' to aid in the subgroup analysis. The use of stratified splits by 'Types of Work' produced equal proportions of the five types of work. A random 30% subsample of the entire dataset was used for both rapid exploratory analysis and visualization prototype development.

Feature Normalization and Scaling

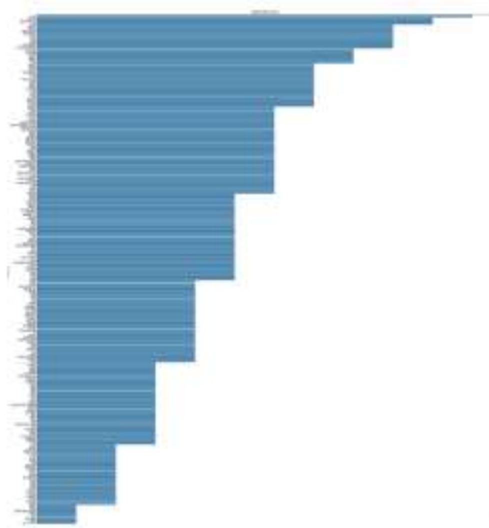
Feature normalization/scaling transformations for three continuous numerical features were performed on the values of each of the 3 continuous numerical features in the dataset. The' Min Salary' & 'Max Salary' features were transformed through the application of Min-Max normalization to the normalized range of [0,1]. The data were Standardized by removing the mean from all observations and dividing observations by the standard deviation (i.e., Normalized to Zero Mean, Unit Variance). This standardization of the dataset results in a mean of 0 and a standard deviation of 1 for the continuous numerical features. The use of Logarithmic Transformation via the numpy.log1p function on the dataset reduced the influence of extreme value distribution on overall distribution.

IV. DATA VISUALIZATION

Fundamental Visuals

Job Postings Globally: Bar Graph

The job postings by country were illustrated with a bar graph that indicated how job opportunities are distributed geographically throughout the 213 countries in the dataset. The layout appears to show that there exists a geographic concentration of job postings because only a few of the countries create a disproportionately high number of job postings compared to their total employment. This geographic concentration may result from the distribution and acceptance of online platforms for job posting and the variances in economic activity in each individual country.

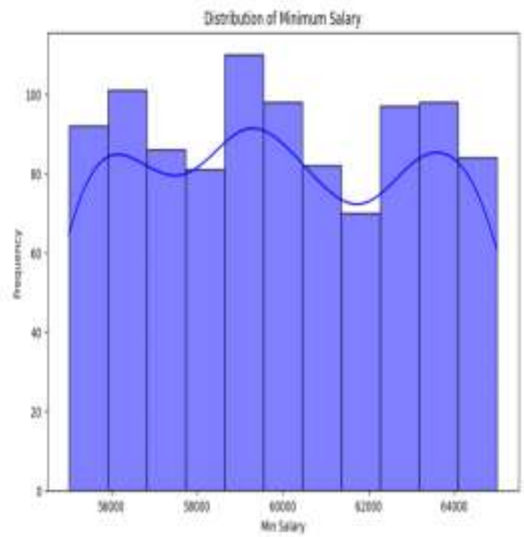


Salary Distributions through Histograms

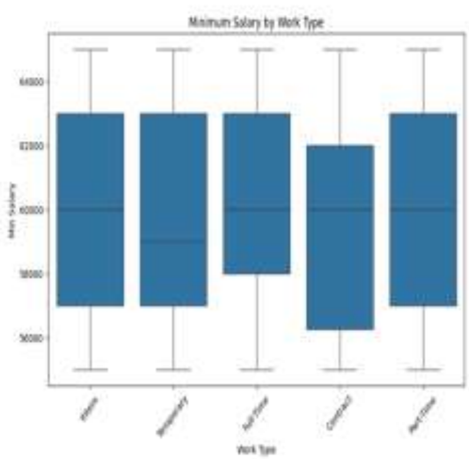
To understand the distribution of both Min and Max salary variables, histograms were created. The histograms showed that there were numerous peaks in the hall of the Min and Max salary distributions and that the Min salary histogram peak occurred between $55,000-$65,000 while the Max salary peaked out much more broadly. The fact that both salary distributions had multi-modal shapes suggest that

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 2, Mar-Apr-2026, ISSN (Online): 2395-566X

clusters of salaries exist at different levels of employment, industry, and geography.



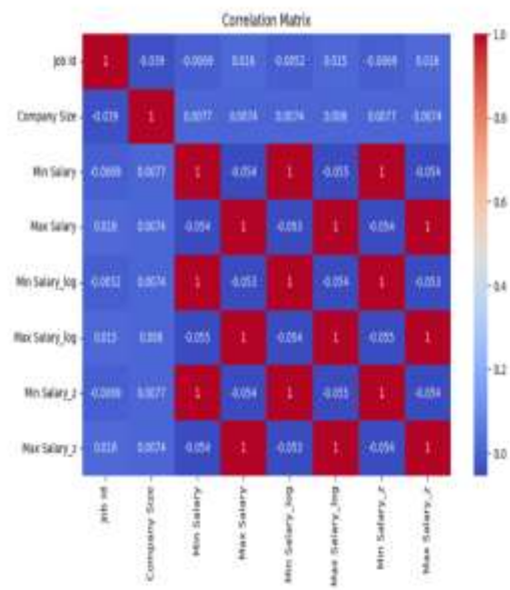
Box Plots - Salary by Work Type



Box plots of the Min Salary and Max Salary distributions were made for the five different work arrangement types to help compare their compensation distributions. Upon viewing the box plots it is clear that the median salaries were fairly consistent between all five work types. For example, the median Min Salary for each of the five work types ranges from $56,000 to $63,000 and the median Max Salary ranges from $90,000 to $121,000. Thus the lack of variance in both the Min and Max Salary for each of the five work types indicates that the type of formal employment arrangement - whether it be Full-time, Part-time, Contract, Temporary or Intern - does not appear to be a key factor in determining the level of salary for any of the five work types. There must be other factors that affect the salary than those associated with the different types of work arrangements.

Experience Level - Salary Scatter Plot

As illustrated in the graph above, experience level as determined by job title shows little or no correlation with the minimum pay an employee would receive for their position. As a result, it is possible to have a high level of experience in one job and receive a lower salary than someone with less experience in another job. Although having experience is a factor when determining how much income to assign to an employer; experience is not the only factor but one of many factors affecting salaries.

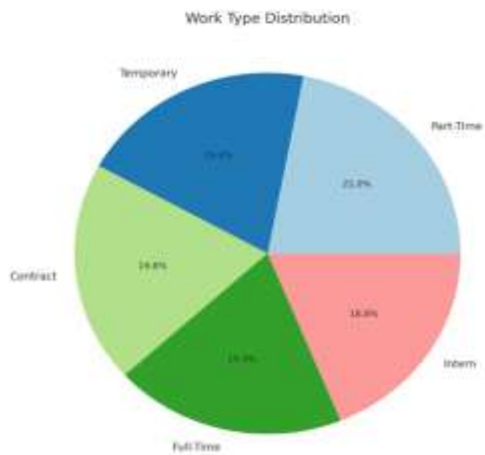


Pie Chart — Work Type Distribution

In the aggregation phase, the pie chart was also used to visually represent the distribution of the identified work types. Part Time (22.0 percent), Temporary(19.9 percent), Contract(19.8 percent), Full Time(19.4 percent), and Interns(18.8 percent) are all very similar and almost equally distributed in the data set making them one-fifth of the overall data set. The near equal distribution was identified as a significant trend within the data set and supported the

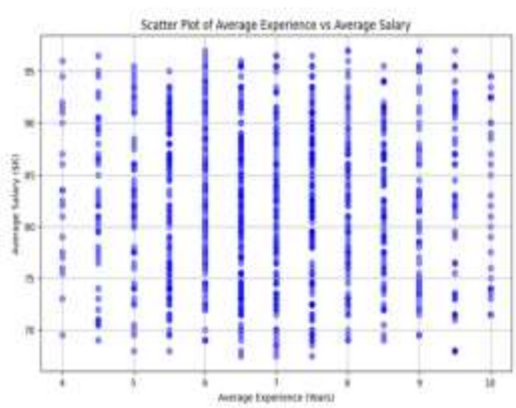
International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 2, Mar-Apr-2026, ISSN (Online): 2395-566X

fact that there are many different types of employment types represented in the data.



Choropleth Map — Global Job Distribution

A choropleth map was created with Plotly Express to demonstrate the total number of jobs per country on an interactive map, using a Viridis color scheme, for all 213 countries in the dataset. In reviewing the choropleth map, it is clear that each country had a wide range of job posts from 1-10 postings; this showed that the dataset was generally uniformly sampled around the world. Several countries in Sub-Saharan Africa and Asia appeared to have low job postings compared to other areas. Simultaneously, several smaller countries (Albania and American Samoa) had higher average maximum salaries (e.g. > $100,000).



Advanced Visualizations

Correlation Heatmap

A correlation heatmap was produced for all continuous numerical variables in the dataset to ascertain linear relationships between those variables. The correlation between Min Salary and Max Salary displayed a weak, negative linear relationship (r=-0.054). Contrary to how we would expect them both to be positively correlated, the data suggests there is no correlation between Minimum Salary and Maximum Salary: Higher Minimum Salary jobs offer no guarantee of higher Maximum Salaries. Similar to the previous example, Company Size had a very low linear relationship with both Minimum and Maximum Salaries. This further supports that Company Size is not an influential determinant of Salary. The weak linear relationships also suggest that even if there were multiple variables included in this analysis, the determination of Salaries for the job postings was based on nonlinear variables.

V. RESULTS AND DISCUSSION

Qualification Level and Salary

The types of education/diplomas required by a job do not create a linear correlation with salary compensation. Violin plots showed that salary compensation is the same across all levels of education/diplomas so therefore the level of education/diploma required for a job posting does not appear to be a significant factor in determining salary compensation. Overall, the general research on the labour market suggests that the independent variables which directly impact salary compensation are: sector; geographic location of the job; and type / function of job (i.e. accounting). Therefore, it can be concluded that employers usually pay market rate salaries for a job position as opposed to paying a salary based upon the level of education / diploma that is required for that job.

A. Geographic Disparities

A choroplethic map of various geographical locations reveals differences in compensation across the globe. An examination of several countries shows that Albania and American Samoa have an average maximum salary over \$100,000 – substantially above the country means. Comparatively, the maximum salaries in Sub-Saharan Africa and other areas of Asia are less than average maximum compensation levels for similar jobs. It should be noted that all findings are correlated to macroeconomic conditions in each country and therefore demonstrate the importance of geography when determining compensation levels.

Work Type and Company Size

Each category of work type in the dataset contains about one-fifth of the total postings, so this data set reflects a balanced cross-section of employment flexibility available today across various forms of employment. The employee number range for all companies represented within this data set is from 12,677 (small) to 134,519 (large), indicating that flexible work arrangements are available at all sizes of companies. Additionally, flexible work arrangements such as Part Time and Contract are found in both small and large companies.

Demand Patterns

An analysis of the frequency of requested skills across industries and job types indicated that the top five requested skills are management, analysis, design, communication, and data. Therefore, this skill demand profile reflects the combination of technical and interpersonal skills needed for today's workforce. The addition of data literacy to the traditional management/analysis/communication trio underscores the increasing importance of data literacy as a cross-disciplinary professional skill; i.e., data literacy is not merely a technical skill, but also a non-technical skill.

VI. CONCLUSION

This research paper provides a meticulous and thorough examination of a global dataset of job postings and demonstrates that thorough data preprocessing and visualization techniques provide invaluable data for the labor market. The preprocessing techniques (e.g., data cleaning, conversion, aggregation, splitting, scaling) were necessary to establish the integrity of the analysis. The visualization techniques helped uncover complex multi-dimensional relationships between salary, education/qualification, type of work and geography. The key findings of this research paper are the indication that nearly identical distributions exist in relation to salary paid for types of work and level of qualification; the presence of wage differentials across between geographic areas, which reflect the macro-economic conditions of those areas; and the determination of the required skills for the labor market. Overall, this research paper is of great value to companies interested in optimizing their pay structures, educators interested in to assist develop the necessary skills, and individuals interested in aligning their skills with available positions. Future work could include temporal analysis by using time-series job posting datasets, salary prediction using supervised learning methodologies, and additional macro-economic variables to provide greater insight into geographic wage disparities.

REFREENCES

1. R. Singhrana, 'Job Description Dataset,' Kaggle,2024. https://www.kaggle.com/datasets/ravindrasinghrana/job-description-dataset
2. StatSo, 'Jobs Dataset,' StatSo Analytics. [Online]. Available: https://statso.io/jobs-dataset/
3. M. Lawrence, 'LinkedIn Tech Job Data,' GitHub Repository, 2021. [Online]. Available:https://github.com/Mlawrence95/LinkedIn-Tech-Job-Data
4. Madhab, 'Job Posts Dataset, Kaggle, 2017.[Online].Available: https://www.kaggle.com/datasets/madhab/jobposts
5. W. McKinney, 'Data Structures for Statistical Computing in Python,' Proceedings of the 9th Python in Science Conference, 2010, pp. 56– 61.
6. F. Pedregosa et al., 'Scikit-learn: Machine Learning in Python,' Journal of Machine Learning Research, vol. 12, pp. 2825– 2830, 2011
7. M. Waskom, 'seaborn: Statistical Data Visualization,' Journal of Open Source Software, vol. 6, no. 60, p. 3021, 2021.