UNIVERSITI KUALA LUMPUR

Malaysian Institute of Industrial Technology

DATA

VISUALISATION

(IIB 40103)

Analysis Report on 2017-2021 Job Market Data in the US

Name	ID Number
WAN NADIRA BINTI TAJUDDIN	52224221156

Lecturer's Name

DIYANA BINTI AB

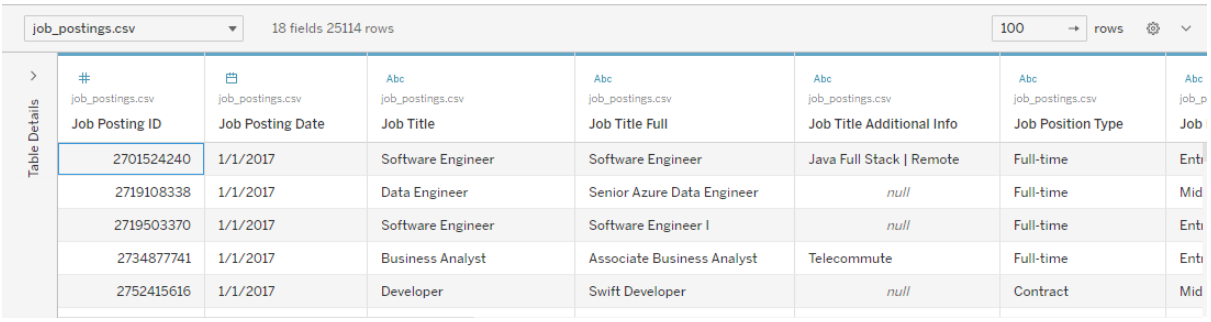
KADIR

Submission Date

18/01/2024

Part 1: Introduction

Every day, thousands of companies and individuals turn to LinkedIn in search of talent. This dataset contains a comprehensive record of 25,115 rows of job postings listed in the US between 2017 to 2021. Each individual posting contains 16 valuable attributes, including the job title (short name), full job title, job title additional information, job type either contract or full-time, job position level, years of experience, job skills, job location, minimum and maximum pay, pay rate, number of applicants in addition to company name, industry and size (number of employees) associated with each posting. The job in this dataset focuses solely on the IT industry, ranging from business analyst, data scientist, UI UX designer, researcher, software engineers all the way to a product manager as the job title imposes.



job_postings.csv | 18 fields 25114 rows | 100 rows

Table Details

job_postings.csv Job Posting ID	job_postings.csv Job Posting Date	job_postings.csv Job Title	job_postings.csv Job Title Full	job_postings.csv Job Title Additional Info	job_postings.csv Job Position Type	job_p Job
2701524240	1/1/2017	Software Engineer	Software Engineer	Java Full Stack | Remote	Full-time	Enti
2719108338	1/1/2017	Data Engineer	Senior Azure Data Engineer	*null*	Full-time	Mid
2719503370	1/1/2017	Software Engineer	Software Engineer I	*null*	Full-time	Enti
2734877741	1/1/2017	Business Analyst	Associate Business Analyst	Telecommute	Full-time	Enti
2752415616	1/1/2017	Developer	Swift Developer	*null*	Contract	Mid

Figure 1. There are 18 fields with 25,114 rows of data in the data source job_postings.csv file.

There are 119 distinct company industries in the .csv file as shown below:

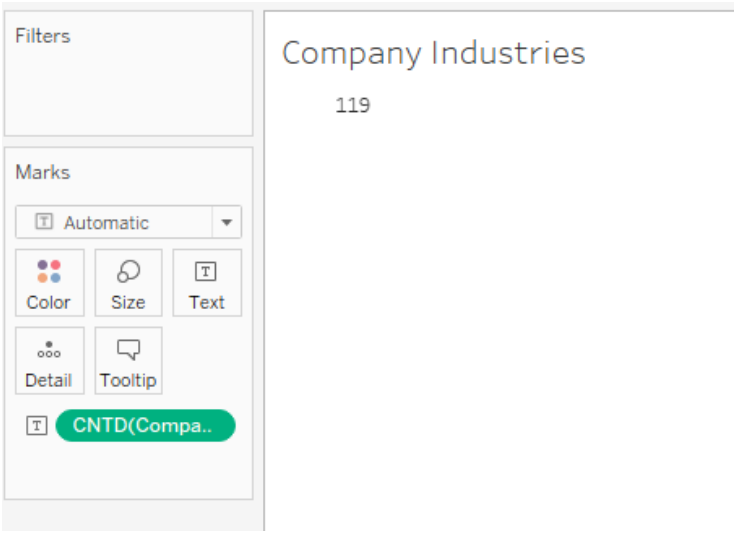


Figure 2. Count Distinct of Company Industries on LinkedIn.

Those 119 company industries were the ones posting one or more job advertisements on LinkedIn and this report will present the market insight analysis from 2017 to 2021 in regard to the job market in the USA.

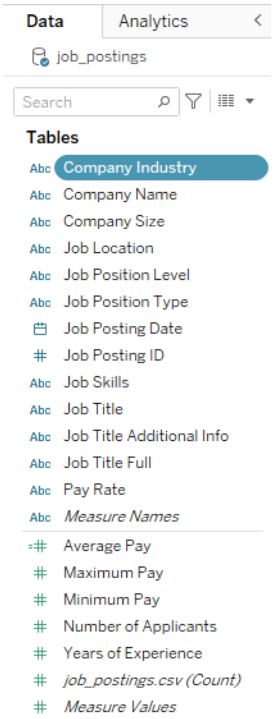


Figure 3. Dimensions and Measures in job_postings.csv in Tableau.

This dataset was downloaded from Kaggle, the world's largest data science community with powerful tools and resources to help people achieve their data science goals. As a subsidiary of Google, it is an exclusive online community of data scientists and machine learning engineers and various other career fields to work together to learn Big Data. Moreover, Kaggle also allows its users to find datasets they want to use in building AI models, publish datasets, work with other data scientists and machine learning engineers, and enter competitions to solve data science challenges. Kaggle got its start in 2010 by offering machine learning and data science competitions as well as offering a public data and cloud-based business platform for data science and AI education.

Part 2: Objectives and Goals

The main objectives of this report are:

1) To investigate the trend and demand of the job market in the IT industry that specializes in Big Data. In the filter section, only 11 job titles are ticked: Data Analyst, Data Architect, Data Engineer, Data Infrastructure, Data Quality Analyst, Data Science Engineer, Data Science Manager, Data Scientist, Data Warehouse Engineer, Database Engineer and Machine Learning Engineer.
2) To examine the relationship between the job position level, salary and the years of experience of an employee.

Additionally, we are also going to find out the top industries and companies that are requesting these roles and the job skills relevant to that position. Since we are using Tableau to visualize our findings, we can also find job titles that heavily consider Tableau as an important skill so that fresh graduates can gain insights from this analysis report. Salary analysis for different job position level is also an important aspect for job seekers who are curious about the job opportunities in the US.

Part 3: Findings and Elaboration

1. The number of applicants for data jobs across various industries:

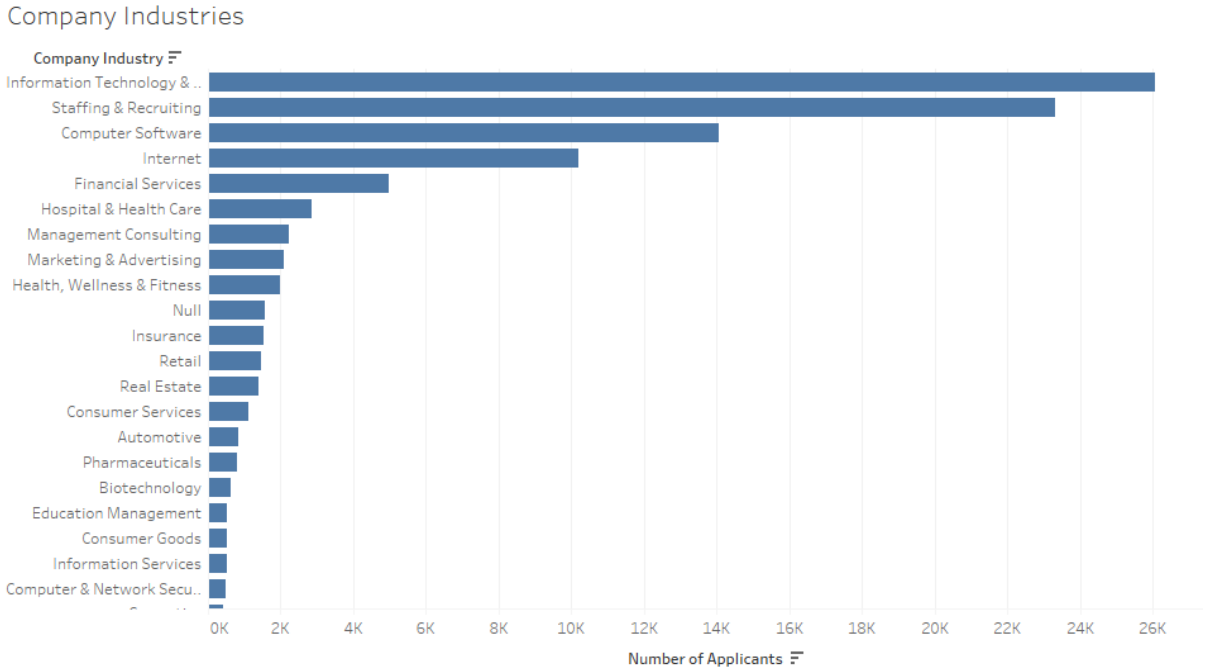


Figure 1- Horizontal bar chart for Company Industries VS Number of job applicants in LinkedIn (sorted in descending order):

The figure above illustrates that the top three company industries applied by job applicants are Information Technology & Services, followed by Staffing and Recruiting industry and Computer Software.

2. Next, to get an initial idea of the relationship between years of experience and the position level for 11 titles of job related to data, the appropriate bar chart is created:

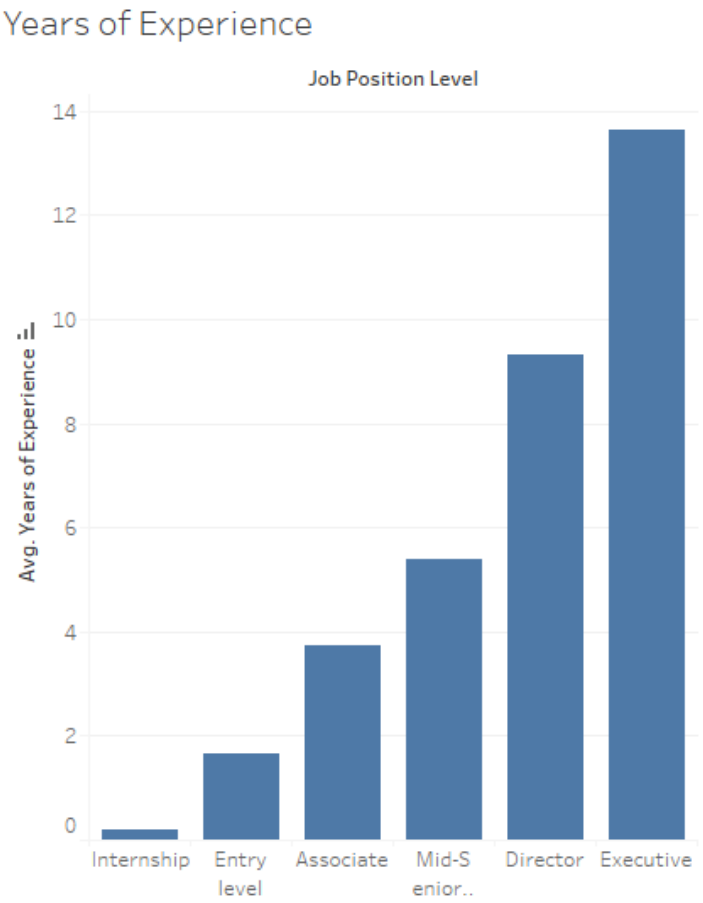


Figure 2- Average years of experience VS job position level for data jobs.

The bar chart above shows that the higher the job position level is for an employee, the higher the average years of experience. Generally, after gaining the first few years of experience in a specific field, employees become qualified to start applying to associate-level or mid-senior jobs because with more level of competence and knowledge obtained, the employer will be less likely to supervise them closely, and they also may find more opportunities to work independently.

3. The number of job postings that are related to Big Data on LinkedIn are trending upward over time from 2017 to 2021. With the exception of a couple of brief periods of decline in job postings in both early and end of 2020; overall, job postings for data roles are rising in demand, which is a very good thing for job seekers who are looking for a job in that domain. The job postings in the area chart in Figure 4 included all the job position levels (e.g. entry level, mid senior level, etc.)

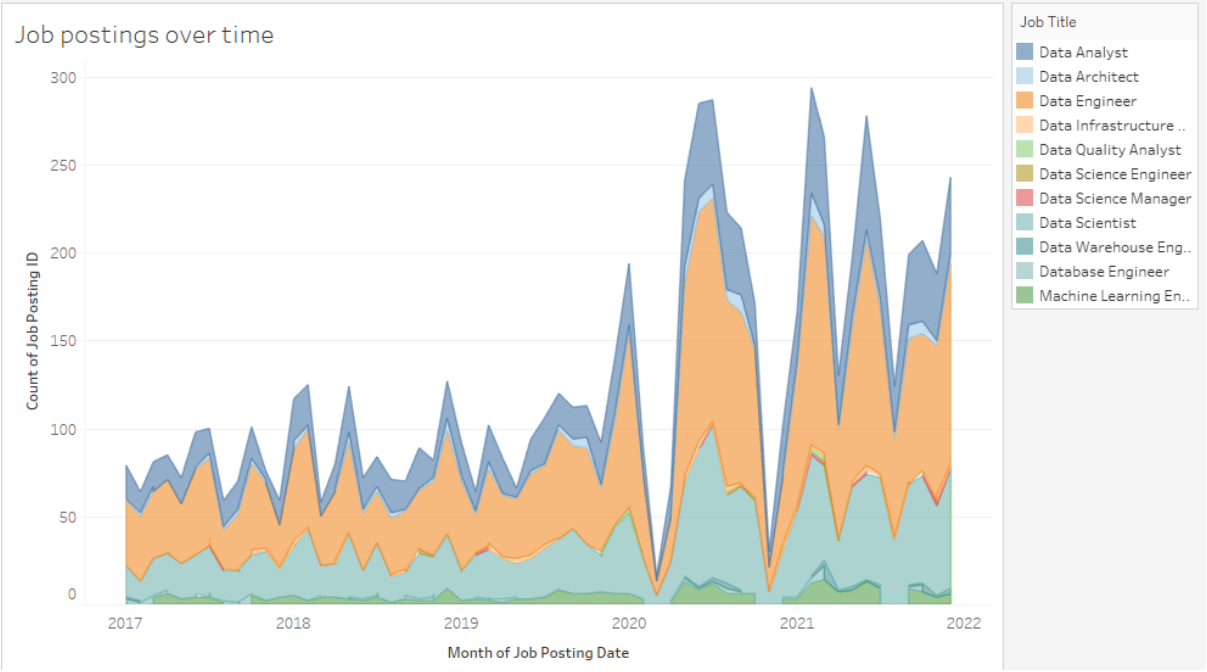


Figure 3- The trend for job postings throughout the year using area chart (continuous) for 11 job titles.

4. From all the 11 data-related jobs to investigate, Data Engineers, Scientists, and Analysts are some of the most in demand jobs based on this treemap:

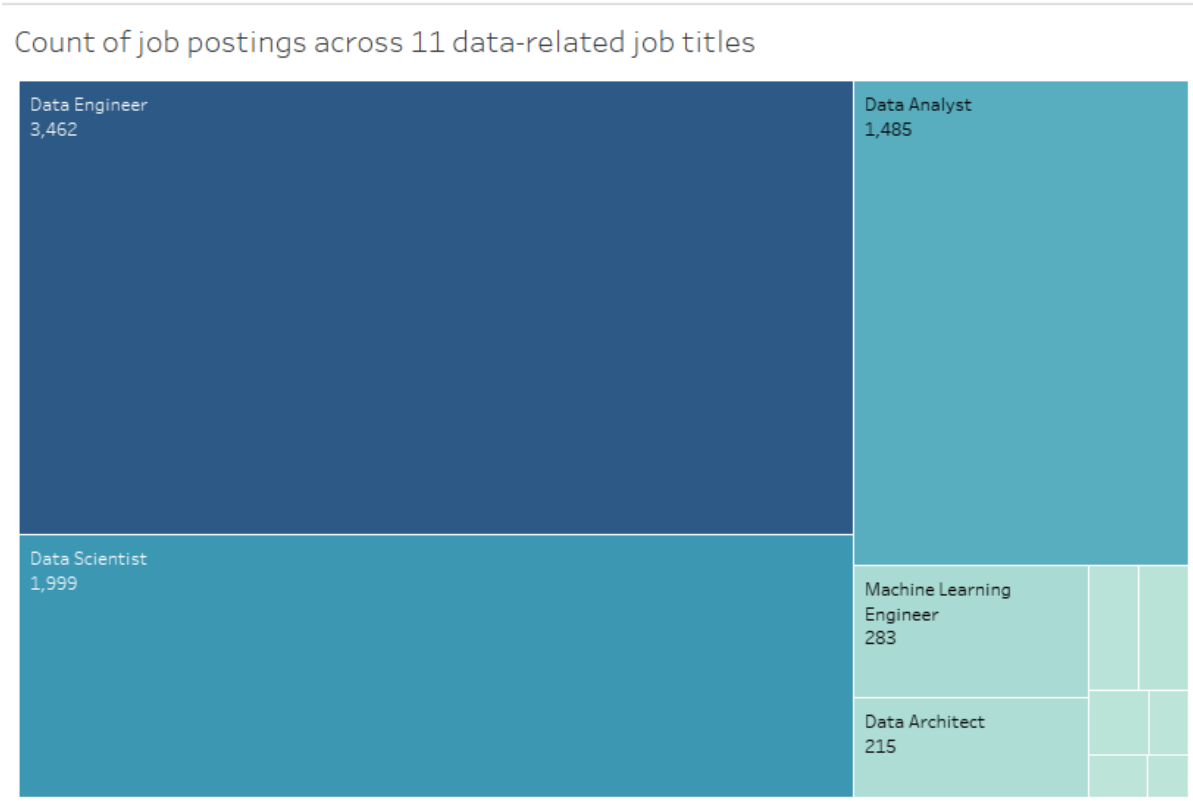
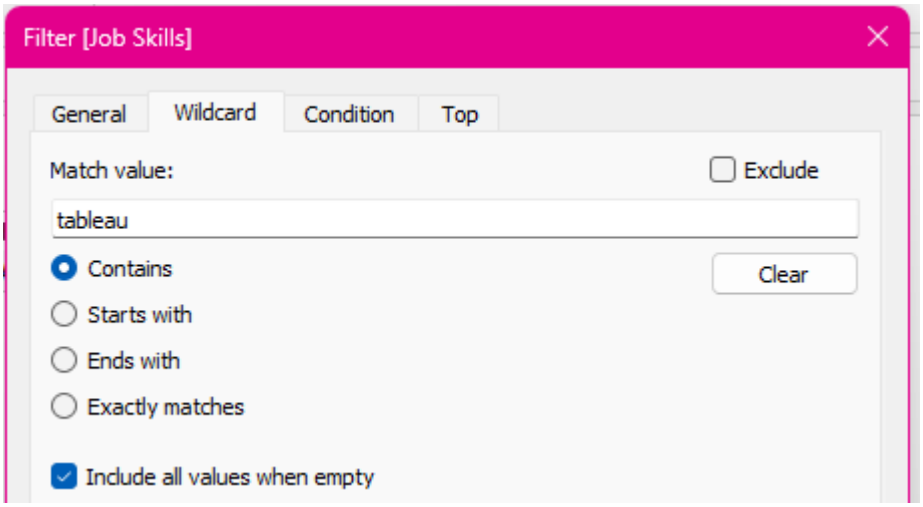


Figure 4- The treemap showing the number of job postings advertised for the data job titles.

5. To find the number of people who have sent their job application to data-related jobs on LinkedIn that listed Tableau as their job skill, filtering out the relevant keyword is necessary in our study. In this case, the first step is to create a text table of job posting ID (count) against the number of applicants (sum). Then, apply the filter 'tableau' to the Job Skills.



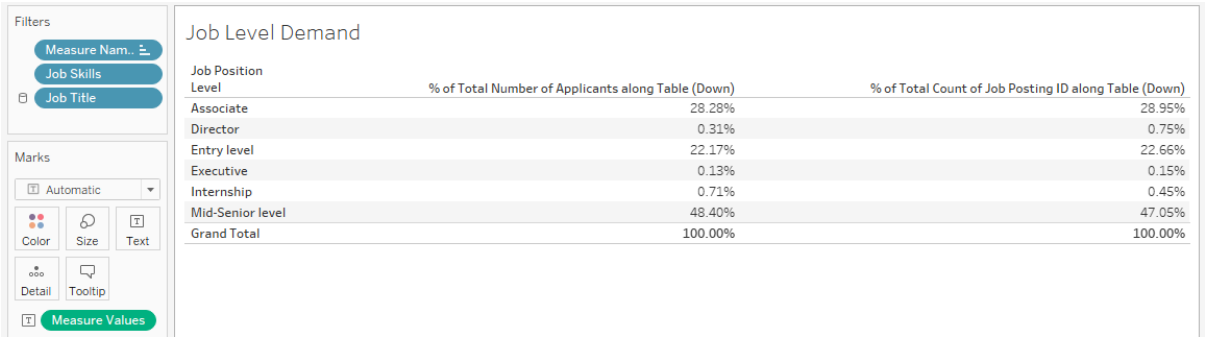
As a result, the text table has shown that there are a total of 27,579 applicants that were applying for 1,337 job postings with Tableau listed as one of the job skills must-haves throughout the years. This also meant that the competition for data-related jobs in the US was already going strong since few years back because there are so many applicants eyeing for very few data-related positions available on that platform.

Tableau Skills Metrics

Number of Applicants	27,597
Number of Job Postings	1,337

Figure 6- Table showing Tableau Skills Metrics; for our next analysis.

6. Continuing from the above finding, it is also possible to study the requirement for having skills in Tableau by experience level. This helps to understand the experience level of applicants that are looking for particular roles in the Big Data industry. However, the table below shows that there is no correlation between the seniority level and the demand for Tableau job skill that has been calculated using total percentage of the column – table down.



Job Position Level	% of Total Number of Applicants along Table (Down)	% of Total Count of Job Posting ID along Table (Down)
Associate	28.28%	28.95%
Director	0.31%	0.75%
Entry level	22.17%	22.66%
Executive	0.13%	0.15%
Internship	0.71%	0.45%
Mid-Senior level	48.40%	47.05%
Grand Total	100.00%	100.00%

Figure 7- Text table showing the total percentage of number of applicants and total percentage of count of job posting IDs along the table (down).

7. Next, we can also do investigation on job postings that have Tableau listed as one of the top three required skills. If they rank Tableau either first or second or third one in the order of the job skills for that particular job posting, then we can assume that those kinds of jobs may require more Tableau expertise in job applicants rather than other jobs who list Tableau only as their 4th, 10th or even 12th skill in the job poster.

Figure 8- Calculated field formula for a new column called "Tableau Top 3".

Hence, from these circle sizes, it can be seen that among the 11 data job titles, the one that listed Tableau the most as the top 3 required skills is Data Analyst with 182 count of job posting IDs on LinkedIn.

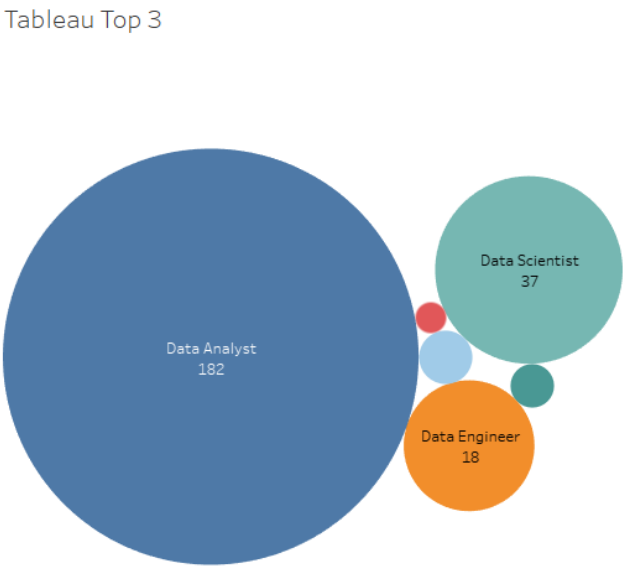


Figure 9- Different circle size illustrating the top 3 data jobs that emphasized Tableau job skill on their job postings.

8. By going through the key job descriptions, sometimes candidates in the Big Data field might have longer years of experience despite not being represented across job position levels. We can study this theory by investigating the relationship between job roles and years of experience.

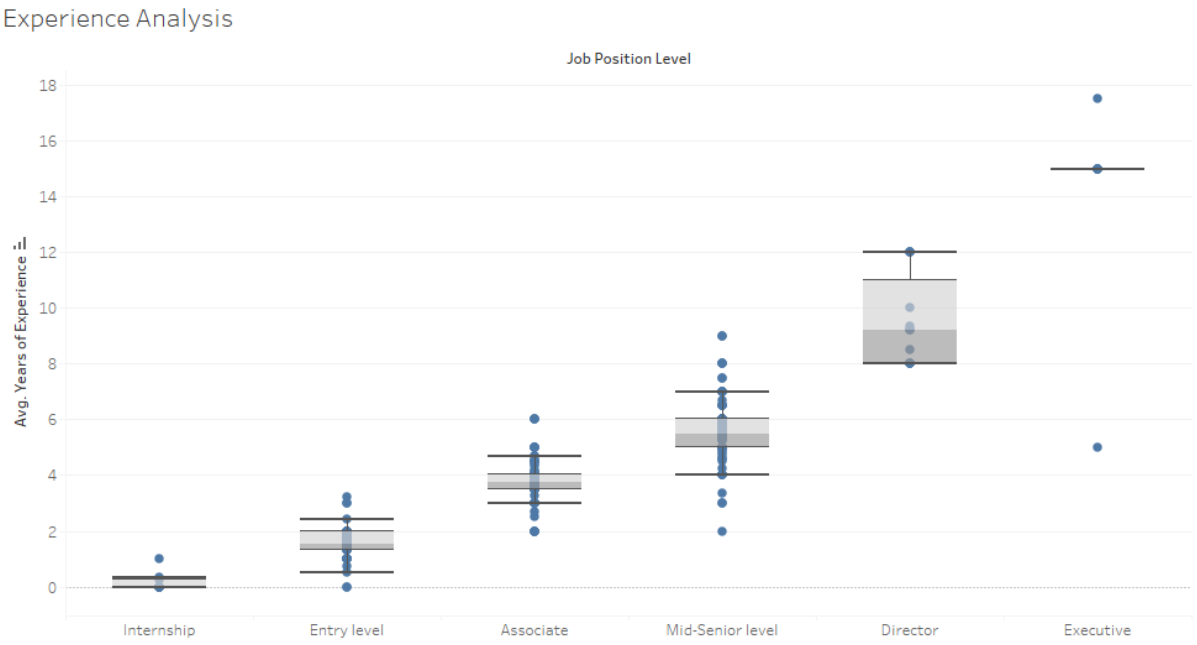
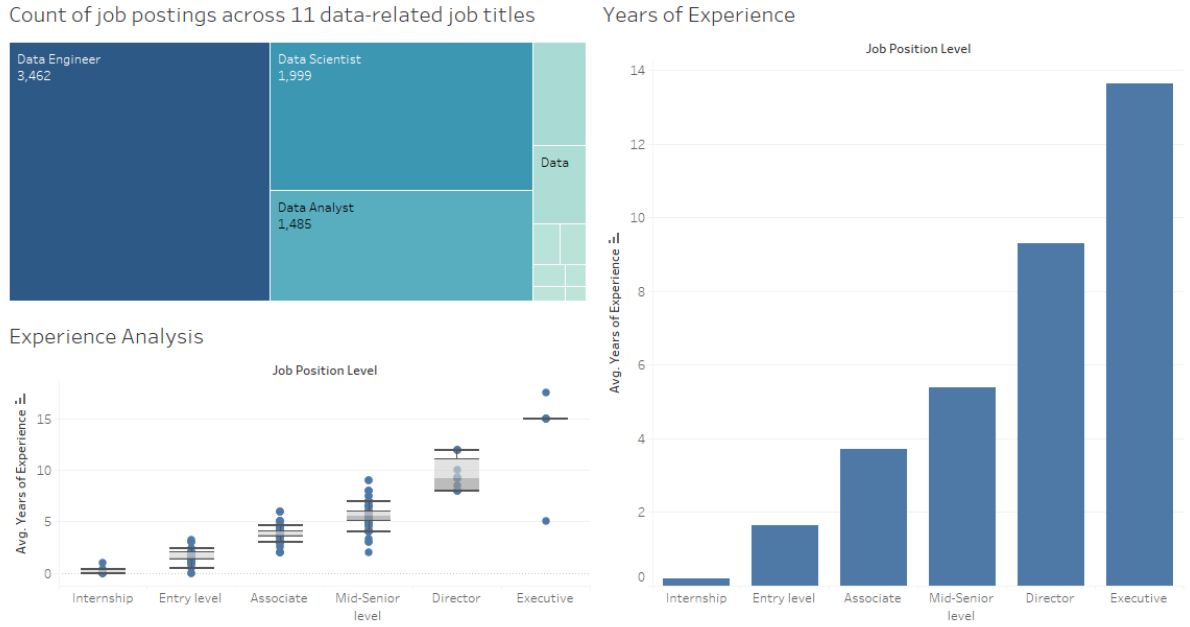


Figure 10- Box plots showing the average years of experience by job position level in different company industries.

The above box plot has been sorted in ascending order of average years of experience and according to it, the higher the position level of a job, the more average years of experience it takes for someone to make a career progression. There is also a huge gap difference of average years of experience for executive level across different company industries. To get to the executive level for example, it takes someone an average of 17.5 years of experience to reach that level in the IT Services industry whereas the Staffing and Recruiting industry takes only 5 years for someone to be promoted to the executive level.

Tableau Dashboard for the Job Market Data in the US 2017-2021



The above dashboard was made by dragging 3 sheets which has helped us to analyze the data job market. The first one is the treemap count of job postings, second is the years of experience against job position level in Big Data industry and lastly, the experience analysis. This interactive dashboard can be used as tool to discover the average years of experience for a specific data job title by applying the filter on the treemap and then selecting any job title that we are interested in knowing. The variable that is kept the same across the sheets is the 11 job titles which are Data Analyst, Data Architect, Data Engineer, Data Infrastructure, Data Quality Analyst, Data Science Engineer, Data Science Manager, Data Scientist, Data Warehouse Engineer, Database Engineer and Machine Learning Engineer.

By navigating through the job title, we can also observe the job market by looking at the number of job postings and predict how many years it takes to specialize in that niche and the appropriate job position level for that seniority. Besides, we can also relate that to the average pay in the other sheet that is not in the dashboard – Salary Analysis.