데이터 과학자 급여 분석: 신규 졸업생 및 구직자를 위한 인사이트

Analysis of Data Scientist Salaries: Insights for New Graduates and Job Seekers

지도교수 이주현

이 연구보고서를 공학석사학위
연구보곳서로 제출

2023년 05월

청주대학교 공과대학교

데이터사이언스학과

베네딕터스 에스라 헤르노오

목차

그림

1. 서론

1.1. 조사의 배경

Data science has become one of the most popular professions in the 21st century. The Harvard Business Review even dubbed it "The Sexiest Job of the 21st Century" (PatilThomas, 2012). This growth has been driven by the increasing importance of data in decision-making across various sectors, from finance and healthcare to technology and public policy.

Numerous companies are leveraging the power of data and technology to enhance their operation and services, showing the need of data processing skills which lead to high demand of data scientist job.

For Example, A Freeport industry case in 2018 located in Bagdad, Arizona, uses the technology of Data Science to analyze their mining process of copper and build their AI model to help find their equipment maximum capacity. The results show that the mine’s processing rate is 10 percent higher than before. These examples underscore the significant role and increasing demand for data scientists across different industries (mckinsey, 2020).

However, despite the high demand for data scientists, there's a lack of transparency and understanding about the salaries in this field. This is especially true for new graduates who are entering the job market for the first time. Given the wide range of industries that employ data scientists and the varying levels of expertise and specializations within the field, salaries can vary significantly.

This research aims to analyze data scientist salaries to provide valuable insights for new graduates about potential earnings, industry trends, and geographic influences on pay.

1.2. 조사의 목적

The objective of this research is to provide insights to new graduates or entry level employer about the variation in data scientist salaries based on experience level, job industry, and job location. By analyzing these factors, this study aims to assist students in receiving an informed job

seeking strategy, enabling them to target industries and locations that align with their career goals and expectations for compensation. This analysis will also help illuminate the career progression in the data science field, giving new graduates a clearer picture of potential earnings growth as they gain experience in their chosen industry and location.

Our research focuses on the following key aspects:

• Examining the impact of experience level on data scientist salaries, examining the salary gap between entry-level, mid-level, and senior professionals in the field.

• Analyze the impact of location on data scientist salaries and identify regional trends in salaries in different countries.

• Examines the impact of industry on data scientist salaries, highlighting pay patterns and differences across sectors.

• Correlation between a data scientist's experience level, country and industry, and salary to uncover potential synergies or interactions between these factors.

• Provides insight into strategies for advancing a data scientist career, identifying pathways to higher income potential and increased job satisfaction.

With these specific goals, this research project seeks to provide job seekers and graduate students with a comprehensive understanding and insight into data scientist salaries. The findings of this study will help inform career paths, negotiate effective salaries, and determine career strategies.

1.3. 조사의 필요성

The need for this research project can be found in the fast-growing field of data science, where the demand for skilled professionals is growing significantly. As the field evolves, graduate students in particular need reliable and up-to-date information to inform and make active career decisions. This project will provide useful insights into data scientist salaries by experience level, country, and industry, helping job seekers or graduate students understand the factors that affect salaries.

A summary statistical analysis can determine the average salary of data scientists at each level. This allows job seekers to negotiate salaries with confidence so that they are fairly compensated for their skills and expertise.

1.4. 조사의 방법론

We will follow a systematic implementation method that involves data collection, processing, and analysis using a variety of techniques, including regression analysis, feature selection and extraction, etc., to effectively undertake research projects on data scientist salaries. The implementation method consists of the following steps:

• Data Collection: We will import the dataset from Kaggle. The dataset will include salaries of data scientists at different experience levels, countries, and industries.

• Data preprocessing: Before performing the analysis, the dataset will be cleaned and preprocessed using R Studio. First, we will be categorizing the observation value of company location to continents since the variable was written by countries code, (for example “CA or Canada” = “America”). We also categorize job title into several groups labeled as their job title, (for example “Head of Data Science” = “Data Scientist”). The next step is to extract the data that we will only need to do the analysis. The last step is to handle missing values, and outliers to ensure that the dataset is accurate and suitable for analysis.

• Exploratory Data Analysis: We will perform exploratory data analysis (EDA) using R Studio to better understand the structure of a dataset and identify trends, patterns, or relationships among variables. We also going to be focusing on EDA for Entry Level experience.

• Regression analysis: We also will interpret regression analysis to make a prediction model for salary based on job industry, experience level and company location

• Interpretation and presentation of results: Finally, we will interpret the analysis results and present the insights from the project clearly and concisely using visualizations, tables and text.

2. 본론

2.1. 경험 수준별 데이터 사이언티스트 평균 급여 비교

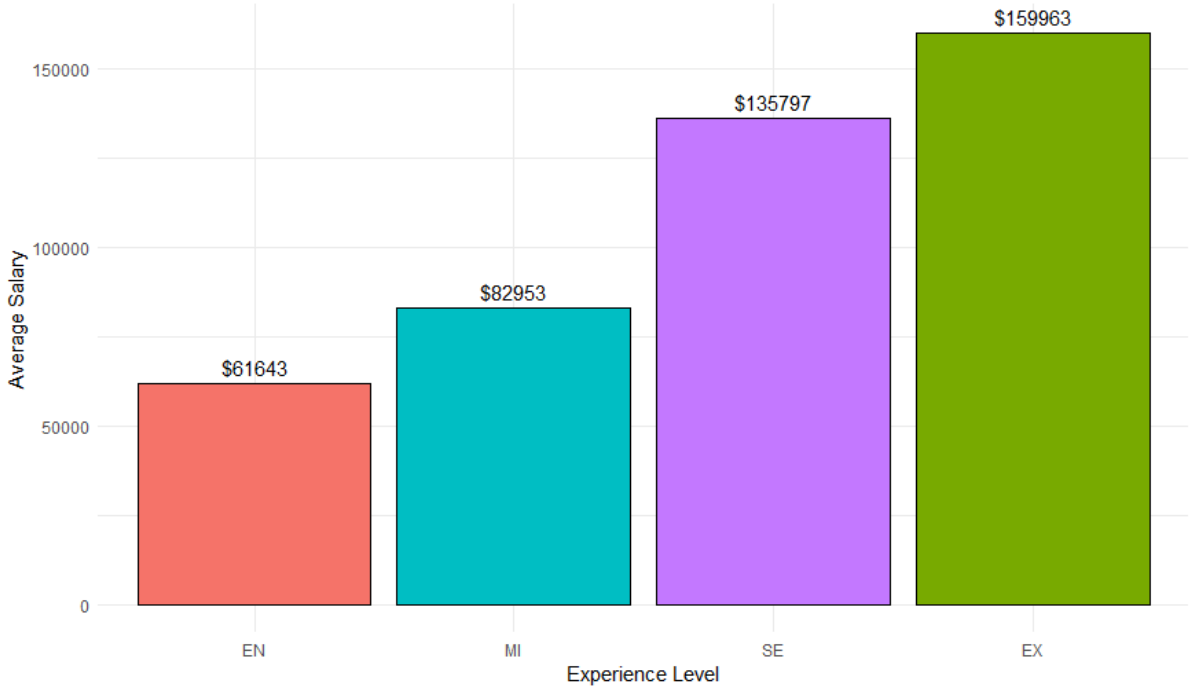

Figure 1: Mean by Experience Level

In our analysis of data scientist salaries, we observed significant differences based on experience level. Entry-Level (EN) position of data scientist earns an average salary of $61,643. As data scientists have advanced their career to Intermediate-Level (MI) position with more work year expertise and experience, their average salary is increase to $82,953. This shows that early-stage data scientist has a substantial potential for salary growth. Further progression to Senior or Expert (SE) position, it shows another significant salary increase, with an average salary of $135,797, reflecting the value of experience and expertise in this field. Finally, the Executive-level or Director (EX) position earns the highest average salary of $159,963, describing the leadership roles within data sciences. These findings indicate that having a career in data science entry level may start with lower salaries, but their earnings potential can increase as they ascend to higher position by gaining experience and advancing their skills.

2.2. 산업별 데이터 사이언티스트 평균 급여 비교

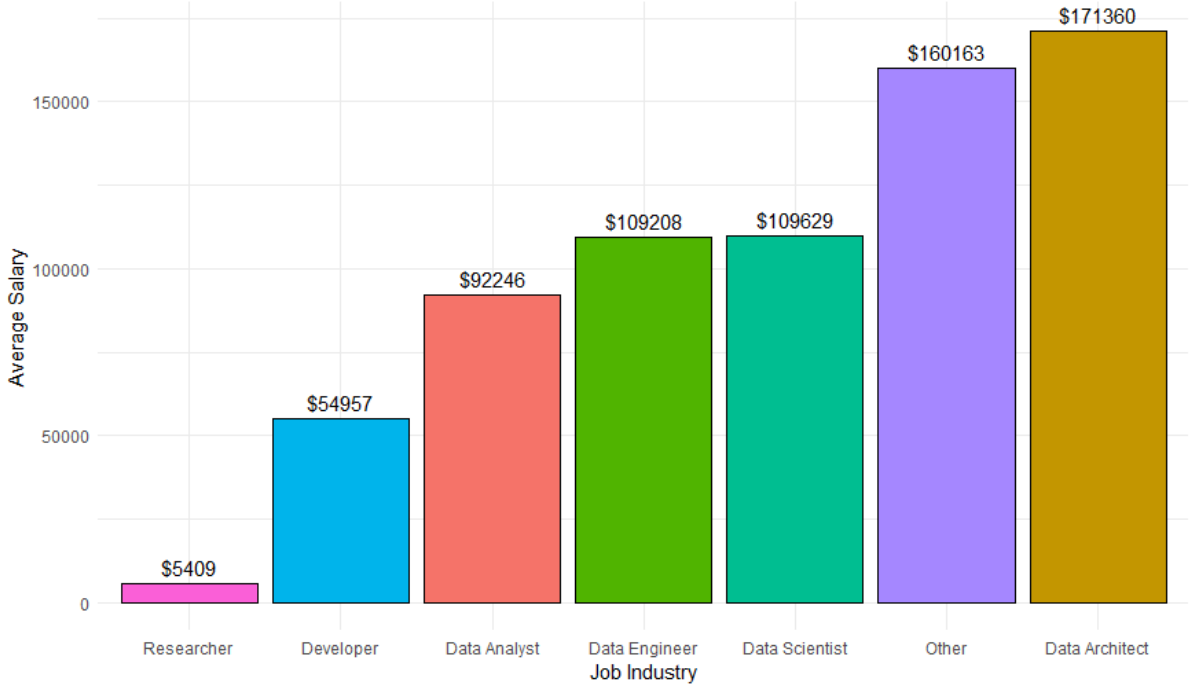

Figure 2: Mean by Industry

Our analysis of average salaries across different job industries in the field of data science shows significant variations. Our findings show that the highest job industry pays among all industries is Data Architect with an average salary of $171,360. They are responsible for designing, creating, and managing the company's data architecture. The second highest industry is labeled as "Other" industry with an average salary of $157,191. The term "Other" in the industry category refers to roles, such as "Head of Data" which are not categorized under the specified industries in the category. The third highest-paying industry is that of a Data Scientist. This job position are responsible for processing and interpreting complex data to aid in decision-making and earn an average salary of $109,629. The fourth highest-paying industry is held by Data Engineers, who earn a similar average salary to Data Scientists, coming in at $109,208. And followed by Data Analyst with $92,246, Developer with $54,957 and Researcher with an average of $5,409. These findings show a variation in salary range across different industries.

2.3. 지역별 데이터 사이언티스트 평균 급여 비교

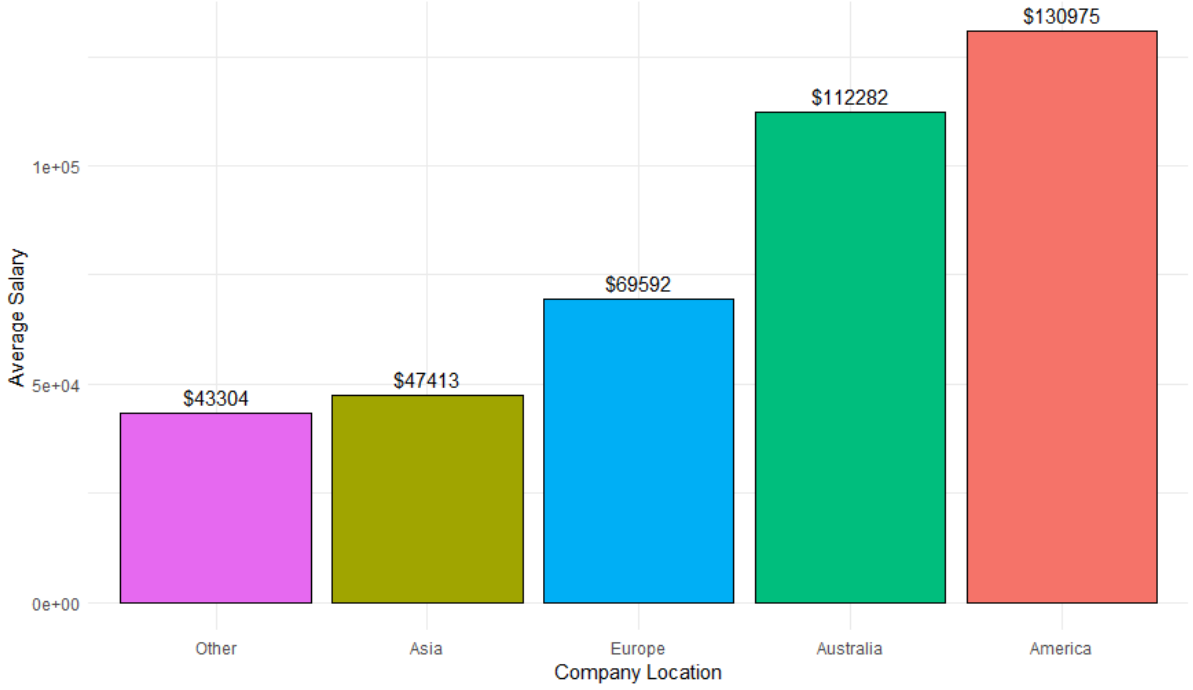

Figure 3: Mean by Job Location

Our findings about Data Scientist relating to Job Location sees that the region with highest paying is America with the average of $130,975 a year. This tells us that America has a strong demand for data science and data skills. Australia follows with the average salary of $112,282, showing a competitive pay for data professionals. In Europe, the average salary is considerably lower than America and Australia at $69,592, while in Asia it drops further to $47,413. The category labeled 'Other', which refers to region codes that could not be identified (example of region codes, 'US' stand for United States), has the lowest average salary of $43,304, slightly lower than Asia region. This finding tells the impact of geographical location on earning salary in the field of data science.

2.4. 전략적 분석**:** 초급 수준 경험을 위한 최적의 전략 탐색

Based on our strategic analysis for entry-level professionals in data science, we've identified several key factors that can influence salary outcomes. The key factors that we focused on Company's Location, Job Industry, and the Company's Size.

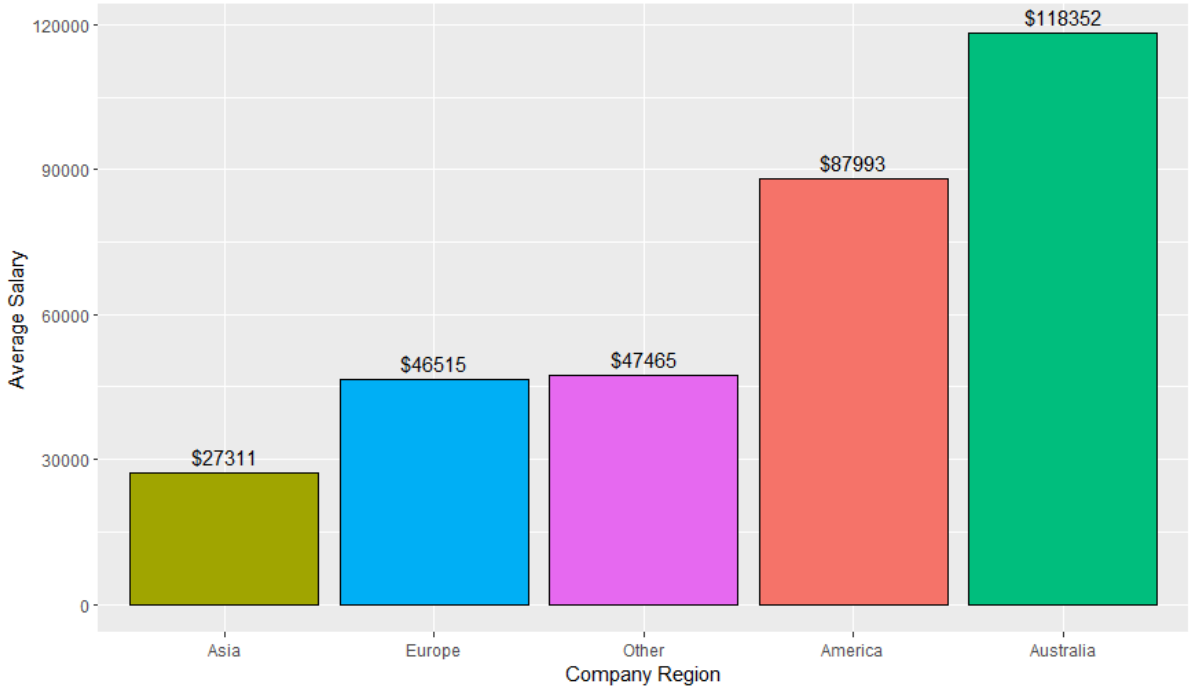

Figure 4: EN Level Average Salary based on region.

Geographically, entry-level salaries vary widely. The highest average salaries that entry level positions can possibly earn is in the company located in Australia region at an average of $118,352. There is a slight difference in average salary for the second region with an average of $87,993, which is the America Region. For the other three continents, there is a noticeable drop in average salaries. The 'Other' region has an average salary of $47,465, and Europe's average salary stands at $46,515, which is almost half the average salary found in the American region. The lowest average salary for entry level positions is Asian region at $27,311. The lower three regions may not be the most lucrative for individuals in entry-level positions seeking high salaries. Therefore, based on the data, a strategic approach for those aspiring to earn a higher salary would be to target opportunities in the Australian and American regions.

Entry Level Employees count on Company Size			
Company Size	S	M	L
Employees Count	29	29	30

Figure 5: EN Employees count on Company Size

Based on the analysis of the data, it can be observed that there is no significant concentration of entry-level positions within specific company sizes. The findings indicate a relatively even distribution of these positions across various company sizes. As a result, individuals seeking employment should not encounter significant difficulties when applying to companies based on their size.

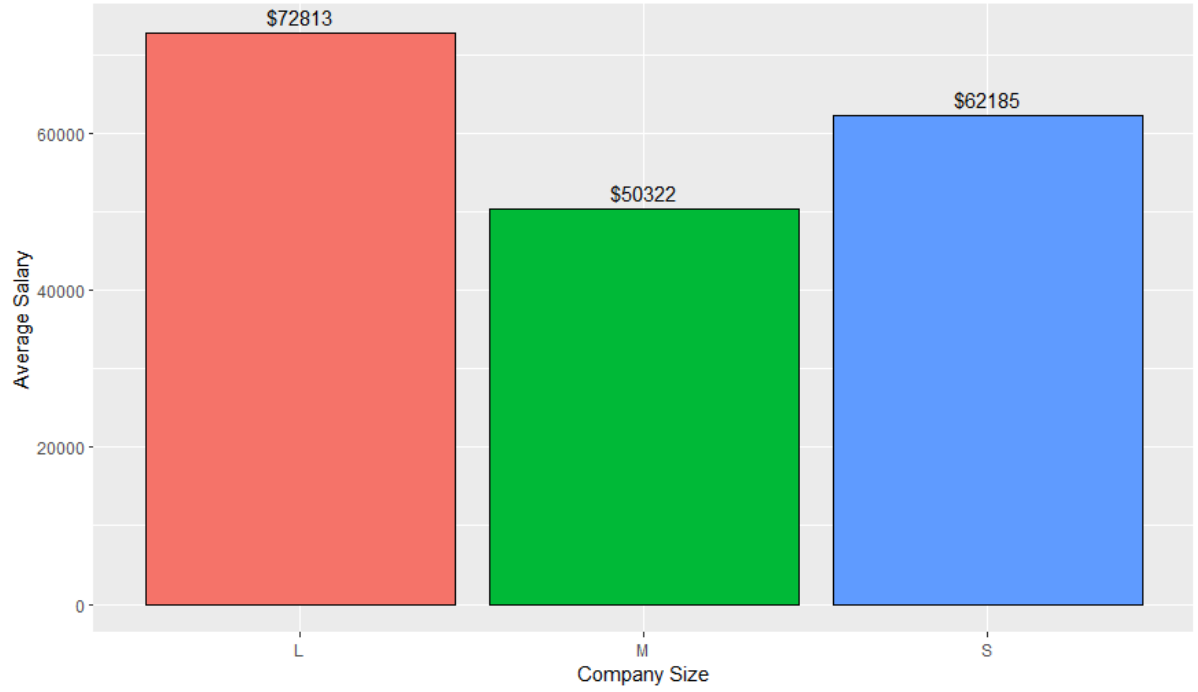

Figure 6: EN Average Salary based on Company Sizes

When talking about the salary based on company sizes, there is a significant difference in the average salary. Large companies show the highest average salary at $72,813, followed by small companies with an average salary of $62,185 and lastly the medium companies with an average salary of $50,322. In conclusion, with the data above about employment count, it will be better to apply for employment in whether the large company or small company if individuals are looking for higher salaries.

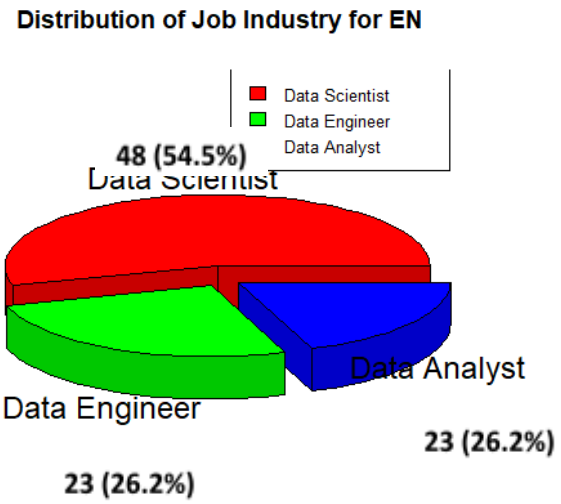

Figure 7: Distribution of EN's Job Industry

According to the job industry distribution analysis, the findings reveal that the Data Scientist category is the most popular job among professionals in the field of data science. This category comprises a total count of 48 individuals, which accounts for 54.5% of the sample population. Following closely, Data Engineers rank as the second most popular category with a total count of 23 individuals, representing 26.2% of the sample. Data Analyst-related industries come in third place, with a total count of 17 individuals, making up 19.3% of the sample.

Based on these findings, it is evident that the Data Scientist category exhibits the highest demand among entry-level positions in the field. Therefore, individuals aspiring to enter the data science field may find the Data Scientist category particularly suitable, considering its prominence and greater number of opportunities compared to other industries.

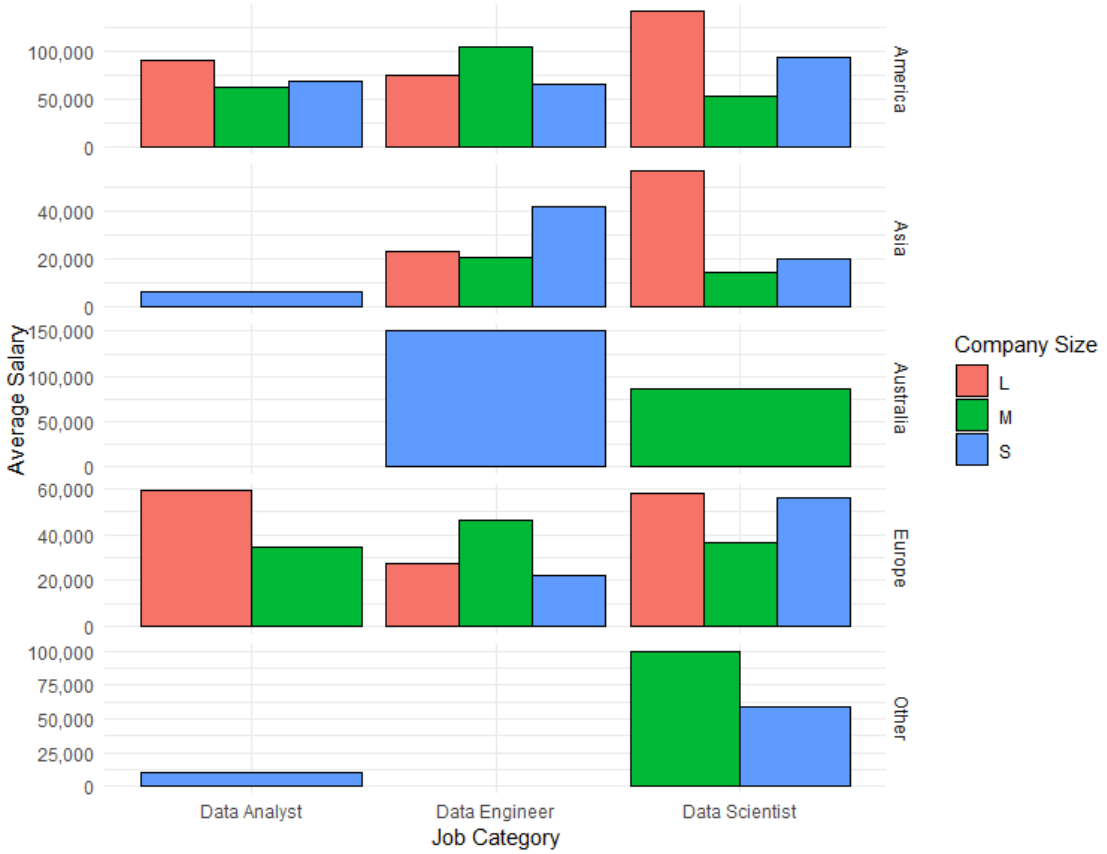

Figure 8: Entry Level Summarization

In the summary of the entry level salaries analysis across different job industries, company sizes, and continents, key patterns have emerged. Small companies in Australia offer the highest average for data engineer category industries at $150,000, followed by large companies in America with data scientist category industries with average of $105,000 and lastly the large companies in America also offering an average salary of $141,079 for Data scientist category industry.

On the other end of the spectrum, the lowest average salaries are found predominantly in Asia, regardless of the company size or job category. Particularly, Data Analysts in small companies earn an average salary of just $6,072. These findings show that the Job Location, job Industry, and company size all significantly influence the potential salary one can expect in the field of data science and related jobs.

Overall, the data indicates that those seeking to start their careers in the field of data science with entry-level positions but are also aiming for

higher salaries may consider targeting small companies in Australia or medium to large-sized companies in the American region.

3. 결론

This analysis has provided a comprehensive insight into data scientist's salaries based on several factors that contribute to salary variations. Factors mentioned such as Job Industry, Experience Level, Company Location, and Company Sizes were found to significantly have an impact on salary in the data science field.

For example, our findings show the positive impact of experience of salary, with a clear trend of increased earning as data scientist professional advancing their skills and experience in the field. Similarly, company location was also found to be significant determinant of salary, with data scientist in Australia and America earning is higher for either Entry Level position and Higher Position compared to their counterparts in Asia and Europe.

Additionally, our analysis of different job industries in data science field show a variation in salary, with data architects earning the highest salary in when talking about all data science experience level and data scientist is the highest earning for Entry level position. Lastly, the company size also influences salaries, with larger companies tending to offer higher average than the smaller companies.

4. 참고문헌

mckinsey. (2020, August 12). *How six companies are using technology and data to transform themselves*. Retrieved from McKinsey Digital : https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/how-six-companies-are-using-technology-and-data-to-transform-themselves

Patil, T. H. (2012, October). *Harvard Business Review*. Retrieved from Data Scientist: The Sexiest Job of the 21st Century. Harvard Business Review: https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century