Exploration of the Common Factors Among the Highest Paying Information Technology and STEM Jobs

Crispin Corpuz, Ken Drucker, Connor Moorhous, Josh Reynoso, Alexis Walther

February 27, 2023

Introduction

As technology continues to spread into countless parts of everyday life, it is easy to marvel at the fact that there are numerous people, teams, and organizations behind its development. While it is true that the jobs and organizations in the tech industry have become lucrative, it is also true that not all jobs are created equal in terms of salary. This raises the question about what factors contribute to the higher salaries in tech and STEM, in addition to what factors contribute to wage discrimination and quality of life for professionals. For context, some jobs in these fields could possibly pay higher salaries than others according to a multitude of factors that could include education level, years of experience, supply and demand for individuals qualified for these jobs, and many more.

The research in this project seeks to answer three questions. First, are there common factors between tech jobs that significantly raise their salary, and what are they? Finding categorical trends among the salaries of IT professionals that can provide deeper insight than the surface level stereotypes of moving upward through the corporate chain or small, startup company successes would be particularly valuable. Second—which is almost a subcategory of first question because of its likely proximity to the same data—is wage discrimination present in STEM and the tech industry? While there are social aspects that confound the information needed to draw a conclusion to answer this question and that are difficult to quantify in data sets, there is value for emerging professionals in being able to provide insight into this topic from an objective data analysis perspective. Third, what are the best cities for STEM jobs in terms of salary, cost of living, and other quality of life factors? This question is in the same spirit as the first question in that building trends and profiles is valuable for quick information digestion. However, it differs in the sense that answering this question identifies specific tech industry hotspots that can help young professionals discover feasible paths on which to start and build their career.

The answers and conclusions generated from this research and data analysis, as mentioned, will be a useful tool for a wide range of young individuals from high school students to college students and graduates to individuals looking to change careers into the tech industry. Interaction with this information can range from a detailed report to a simple web application in order to educate this demographic on the nuances behind jobs in the tech industry and STEM.

About the Data

The first data source for this report, titled "Data Science and STEM Salaries,"[1] is a compilation of data from levels.fyi, a website used for reporting and researching salaries, that is located on kaggle.com. The data set, with 29 columns and over 62,000 entries, contains salary data – including experience, company, location, and demographic information including gender, race, and education – for 15 STEM roles between

[1] https://www.kaggle.com/datasets/jackogozaly/data-science-and-stem-salaries

2017 and 2021. This large and diverse data set aids the discovery of any common factors of the highest paid positions in STEM careers. Additionally, the demographic information in this set helps find the presence of any wage discrimination based on gender or race, and the location information helps determine the best cities for STEM jobs.

The second source is the U.S. Bureau of Labor Statistics' "Occupational Employment and Wage Statistics."[2] This employment data from May 2021 contains occupational and wage information for over 1,000 various jobs. These data sets are broken down by geographic region into distinct categories such as: employment per thousand jobs, annual mean wage, and location quotient. This group of datasets will serve as a foundation for determining the relationship between geographic location, occupation, and salaries.

The final source is teleport.org[3], a site that calculates and reports quality of life metrics for cities across the globe. The site ranks cities in 17 categories including cost of living, commute, safety, and more. The data found on this site, which can be accessed using a public API, provides another perspective that benefits the evaluation of the best cities for STEM careers.

Biweekly Timeline

Date	Person	Task
2/27	Entire Group	Complete Project Proposal
2/27	Crispin Corpuz	Submit proposal
3/13	Crispin Corpuz, Joshua Reynoso	Clean and summarize "Data Science and STEM Salaries" dataset
3/13	Ken Drucker, Alexis Walther	Clean and summarize "Occupational Employment and Wage Statistics" dataset
3/13	Connor Moorhous, Teddy Novotny	Scrape, clean and summarize "Teleport.org" data
3/27	Entire Group	- Finalize midterm report - Explain how data was cleaned and reshaped - Explain basic data analysis and visualizations done - Describe obstacles met and how they were handled
4/10	Entire Group	Determine the project's data product scope
4/10	Joshua Reynoso, Ken Drucker, Alexis Walther	Create data visualizations for first research question
4/10	Crispin Corpuz, Connor Moorhous, Teddy Novotny	Create data visualizations for second research question
4/24	Entire Group	- Finish final revisions on project based on feedback from midterm report - Combine final conclusions and findings together into a rough draft for the final report - Identify topics for each group member to organize a portion to present about for the final presentation
5/8	Entire Group	Complete final draft of report; Record and combine individual presentations into one 10-minute long presentation

[2]https://www.bls.gov/oes/current/oes_nat.htm
[3]https://teleport.org/cities/