Link for the medium blogpost

https://medium.com/@chonelpatrice/unlocking-the-secrets-behind-developer-salaries-e5d9c6e9c587

link for the GitHub repo

https://github.com/PatriceChonel/NanodegreeProject1

**Unlocking the Secrets Behind Developer Salaries**

Table of contents

Hi everyone, my name is Patrice Chonel. I immigrated to Canada a few years ago, transitioning from a background in finance and accounting. After moving, I decided to make a career change and earned a degree in IT tech support. Although I’m relatively new to the field of IT—especially coding—I’ve always been captivated by the creative and logical aspects of programming. That fascination inspired me to take yet another leap and pursue a career in Data.

One day, my wife asked me a simple yet thought-provoking question: "What salary should you expect after graduating?" To be honest, I hadn’t really considered it before. Her question sparked an idea for my very first project as part of my Udacity Nanodegree. If you’re curious about this topic too, stick around—I hope to uncover some valuable insights about what influences the salary of a developer.

For my project, I chose the **Tech Careers: StackOverflow Developer Survey** as the dataset to analyze. Specifically, I used the 2024 survey conducted by StackOverflow, which can be accessed [here](https://survey.stackoverflow.co/). My goal was to answer a fundamental question: *What factors impact the salary of a developer?*

* Is it the programming language you work with?
* The number of years of professional experience?
* The location where you live?

Spoiler alert: the answer is far from straightforward!

After diving into the dataset, I discovered that over **65,437 developers** from diverse countries participated in the survey. With such a robust dataset, we have a great opportunity to derive meaningful and representative results. Before jumping into the analysis, I had to clean the data first. There were many columns unrelated to my study, as well as rows with missing salary values. After carefully filtering the dataset, I narrowed it down to the rows and columns that are most relevant to answering my questions.

Now, let’s start by examining the country with the most respondents to this survey.

A graph of the country's population

AI-generated content may be incorrect.

Survey data reveals which geographic areas contained participants who entered salary information.  
I had to remove the respondents from the data who failed to submit any salary information.  
The United States dominates the survey results by contributing 4,663 respondents which is a significantly higher number than any other nation. StackOverflow surveys show that USA developers demonstrate higher engagement which could be due to increased platform usage or active participation in international tech discussions.  
The European representation is robust with Germany attracting 2,043 respondents and Ukraine with 1,464 participants while the United Kingdom contributed 1,383 respondents.

Then out of the top 20 countries I took the top 10 to analyze their median salary

A graph of blue rectangular columns with white text

AI-generated content may be incorrect.

And again, we see that the United States has the highest median income on the list, reaching the high end of the graph's range (140,000 USD). This reinforces the picture of the U.S. as a high-paying market for tech professionals,  
Other nations with high-paying wages like Canada, the United Kingdom, and Germany follow, offering strong median compensation. They are world technology hubs with advanced industries and high-paying wages.  
Is already answering that the sole determinant for a high pay is the live in the USA sufficient? No it is not though It's clear from this graph that place is a vital determinant of salary. This conforms to wider patterns where salaries have a strong connection to economic situation, local demand for technical skill, and maturity of industry

I then analyze the influence of the number of years of experience on the salary.

A graph showing the difference between a number of numbers

AI-generated content may be incorrect.

Here I filtered to show only salaries between 5000 and 300000 USD since some of the salaries were outliers. Below are the key takeaway that we obtain from this plot.  
there is a positive correlation because the salary increases with the years of experience. The greatest salary jumps occurst in the initial 5-10 years (or say from junior to mid senior levels).

Then I wanted to analyze if the type of degree influence the salary

A graph of a salary

AI-generated content may be incorrect.

I then try to analyze the median salary by language

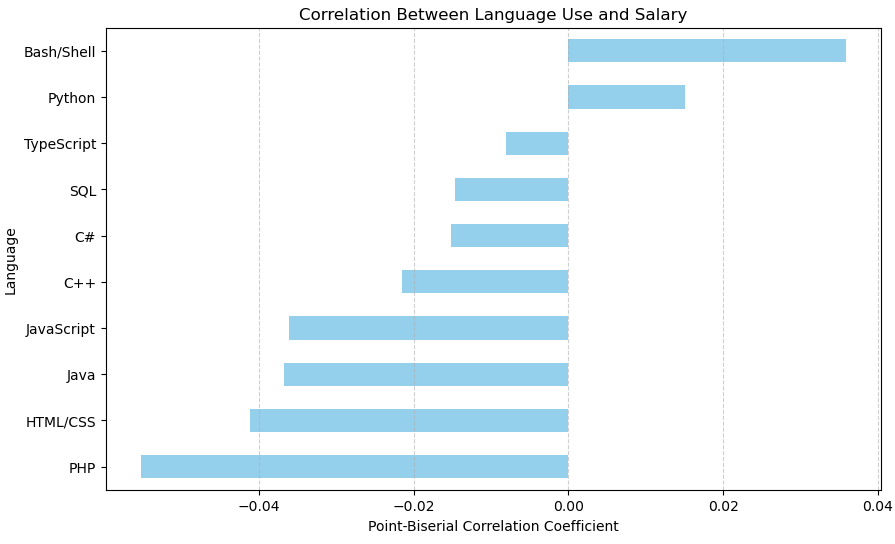
A graph of a number of people

AI-generated content may be incorrect.

We can see tha Bash/Shell has the highest median salary and we understand why because they are often used by senors engineers with devops and sys administrators who are specialized roles.

Then in second place we have python because It is used a lot in different development such as data science , AI , data analytic.

Now I could try to see if there is a correlation between the type of language and the salary.



The horizontal bar plot displays the strength or direction of the correlation between knowing a language and the salary. The stronger the positive value is the higher the correlation is with salary.

The weaker the negative value is the lower the correlation with salary. Those languages near 0 have no clear relationship.

A chart of a salary distribution

AI-generated content may be incorrect.

Here this boxplot for salary distribution shows the salary distribution for developers who use or do not use each language.

The key takeaway for this two plots is that there are not strong correlation with the language alone that could influence the salary

As we have not seen that one single factor was not the sole influence on the salary we will now try to see if all together combined (location, years of experience, language and degree).

A close-up of a graph

AI-generated content may be incorrect.



Here I ran a multi linear model because there were several independent variables that could have an effect on the target that is the salary. We performed the analysis by aggregating the number of years, education level, location, and language.  
We observe that with relative significance of the four determinants in determining the salary however we have a Model R² measure: 0.31. This indicates that our model explains approximately 31% of the variation in salaries and implies that there are other important factors in order to be able to determine the salary.  
Example of other potential factors to influence the compensation, negotiation skills, soft skills (management), bonuses or equity participation, quality of experience (no years), specific job titles.

Conclusion

This study exposed us to the complex terrain of programmer compensation. Salary is not solely influence by years of experience, educational attainment, location, and level of programming expertise. Although we observed each variable's relative significance, Model R² rating of 0.31 reminds us of a harsh reality: salaries are influenced by numerous other factors outside of the current study.  
Factors such as negotiation skills, leadership skills, bonuses or stock, titles, and the type of work experience significantly influence salary expectations. These qualitative factors emphasize the manner in which real-world outcomes are typically decided on a mix of technical competence and human elements.  
So is there an "only" correct answer as to what controls developer salaries? No. What this project reveals is that it's a combination of skills, location, experience, and opportunity. For potential future developers such as myself, seeing this revelation is motivation not to just focus on technical expertise, but also on networking, interpersonal skills, and familiarity with evolving technology environments.  
Thanks for coming along with me today. I hope this analysis provokes questions and gives you insight into the forces at work in developers' wages. If you're on a similar path or would like to chat, I look forward to your feedback. We'll continue this discussion as we navigate the interesting terrain of tech careers!