Macy Chen

April 12th, 2024

Analyzing Tech Salaries: the Influence of Experience on Compensation

This paper explores variations and factors affecting tech salaries, leveraging a 2016 Hacker News survey dataset about salaries and bonuses. The primary focus of the research is on analyzing the effect experience has on tech salary. The research question guiding this investigation is: "To what extent does experience influence salaries in the tech industry?" Through data analysis and visualization, this research aims to uncover patterns and correlations between salary and experience that can contribute to a more nuanced understanding of compensation trends within the tech sector.

Previous research shows that for the average individual, most past experiences in the field, such as time spent at a different job, generally offer similar returns in wages. However, skills obtained in a different industry or occupation may be less valuable than other types of experience. For example, a study by Goldsmith and Veum (2002) shows that while diverse experiences within the same industry or occupation are valuable, skills acquired from experiences in entirely different contexts might not provide as much benefit. Additionally, research conducted by Williams (1991) indicates that as employees accumulate more time (tenure) at a specific job, their wages tend to increase, especially during the initial years of employment. However, this effect diminishes over time, implying that while job tenure initially contributes significantly to wage growth, its impact becomes less pronounced as employees remain in the same job for more extended periods. On the other hand, general labour market experience has a smaller immediate impact on wages, yet its cumulative effect over a career is more significant than job-specific tenure.

Furthermore, according to Lazear (1976), Young employees typically receive around one-third of their overall employment compensation through human capital. As a result, the influence of present total work experience on determining salary is substantial. Bagger and et al. (2014) provide further research on this, stating that human capital accumulation (improvement of skills and knowledge) and job search activities influence the shape of salary-experience models. Specifically, within the first ten years of a career, the effect of job search on salary growth, whether within or between jobs, decreases. These studies provide valuable insights into the nuanced dynamics of the labour market, laying the groundwork for this paper's research into the effects of experience in the tech industry specifically.

This paper contributes to the existing literature by concentrating specifically on the tech industry, offering a specialized examination of the relationship between experience and salaries within this sector. While prior research has provided insights into the general dynamics of experience and wages across various industries, this study seeks to provide a more tailored understanding within the context of technology. By narrowing the focus to the tech sector, this research can uncover insights that may not be directly transferable to other industries due to the unique characteristics of technology-driven roles, including rapid technological advancements and specialized skill requirements. Through data visualization and statistical analysis, this research aims to enhance the understanding of compensation trends and inform strategies for career advancement within the tech industry.

DATA

The data is obtained from Kaggle Notebook, with the original source by Brandon Telle [source]. This dataset includes details about salaries in the tech industry in 2016, including information such as employer, location, job title, experience level, and compensation details. After cleaning, the dataset contains 1,512 observations and seven variables.

Total salary is the dependent variable. Total salary is the annual base salary plus annual bonuses, excluding any other benefits such as stocks. It serves as the dependent variable in the analysis because it represents the respondent's core earnings and is a key metric in understanding the global tech salary landscape. This allows for analysis of how experience influences salary trends.

The following variables are the independent variables: country, total experience years, experience years with current employer, and job category. The country variable represents the country in which the respondent is employed. This variable is essential for understanding salary disparities across different regions and countries. Economic conditions, cost of living, and market demand for tech professionals vary greatly between countries, directly influencing salary levels. The nature of the job, whether it's software, data, engineering, or management, can affect the salary level. Categorizing job roles helps examine how distinct fields in the tech industry contribute to salary variations at different career stages. The total number of years of professional experience the respondent has accumulated is an expected influencing factor on salary. Individuals with more experience tend to command higher pay due to their acquired skills, expertise, and experience in the industry. Examining the experience level can reflect the market value of an individual's expertise in different locations, providing a more comprehensive understanding of the dynamics influencing compensation in the tech industry. Employer experience years represents the number of years the individual has worked for current employer. With more experience working for an employer, employees may become eligible for annual pay raises or advancement opportunities. Long-term employees may be valued for their corporate knowledge and loyalty, resulting in higher compensation to retain their expertise and commitment. Analyzing how years with the employer correlate with annual total salary can provide insights into career progressions and the value of seniority within the industry.

SUMMARY STATISTICS

Table 1.0: Summary statistics table of qualitative variables.

	location_name	location_state	location_country	job_title_category
count	1512	1101	1512	1512
unique	442	50	61	8
top	san francisco	CA	USA	Software
freq	165	408	1067	823

This table presents the qualitative variables in the dataset related to tech jobs across different employers and locations. With a total of 1,512 entries after cleaning, the dataset encompasses a diverse range of 61 unique countries. The United States (US) dominates with 1,067 occurrences. This represents that the majority of respondents work in the US, with San Francisco, California being the common. This finding aligns with the United States being known as a global leader in the technology industry, with the well-known concentration of tech companies and innovation hubs in the broader California region. In terms of job titles, "Software Engineer" stands out as the most frequently occurring title at 287 times. In line with this result, out of the eight job categories available (engineering, software, web, data, management, operations, applied science, other), "Software" is most commonly worked in, appearing 823 times. This high popularity in software-related jobs likely results in higher competition in this category, and its effect on the experience-salary relationship is an area of exploration for my research.

Table 2.0: Summary statistics table of quantitative variables.

	total_experience_years	employer_experience_years	annual_base_pay	annual_bonus	total_salary
count	1512.000000	1512.000000	1.512000e+03	1512.000000	1.512000e+03
mean	6.568671	2.521733	1.094018e+05	7109.686242	1.165115e+05
std	5.294033	2.695830	3.600806e+05	27601.150643	3.630274e+05
min	0.000000	0.000000	3.015246e-02	0.000000	3.359846e-02
25%	3.000000	1.000000	5.800000e+04	0.000000	5.937124e+04
50%	5.000000	2.000000	9.418736e+04	0.000000	9.828384e+04
75%	10.000000	3.000000	1.260000e+05	7000.000000	1.332500e+05
max	40.000000	30.000000	1.028000e+07	750000.000000	1.028000e+07

This table includes quantitative variables in the dataset, including compensation and experience. The mean total experience is approximately 6.6 years, with the standard deviation at 5.3 years. The minimum of 0 years suggests the presence of respondents with minimal experience in the industry, while the maximum is 40 years. The average years with the current employer is approximately 2.5 years, with the standard deviation at 2.7 years. Similar to total experience years, the minimum of 0 years and a maximum of 30 years reflect a large difference in years with the current employer. This notable range between maximum and minimum experience years indicates diverse experience levels within the survey respondents.

The average annual base pay is $109401.80 USD, but the high standard deviation of $360080.60 USD indicates substantial salary variability. Notably, some entries report no base pay, and there is a wide range of salaries, including potential outliers with extremely high values. Whether this is related to the wide range of years of experience can be explored later in my research. The annual bonus, with a mean of $7109.69 USD and a significant standard deviation of $27601.15 USD, reveals wide-ranging bonus amounts, and a substantial proportion of individuals do not receive a bonus, as indicated by the 25th and 50th percentiles that are $0 USD.

Similar to annual base pay, the difference between the maximum and minimum yearly bonus recorded is also very large, which could also indicate potential outliers. The total salary is annual base pay plus annual bonus. The average total salary is $116511.50 USD, with a standard deviation of $363027.40 USD, which is slightly higher than the annual base pay alone. This is expected due to a significant standard deviation in annual bonuses.

While investigating the relationship between salary and experience, further exploration is warranted to understand the outliers in annual base pay and bonus. I plan to explore factors contributing to this variation, such as experience level, location and job category. While additional data exploration and statistical analysis are needed to draw more robust conclusions about the factors influencing salary variation, the summary statistics tables provide a foundation for more in-depth analysis and indications of areas of interest.

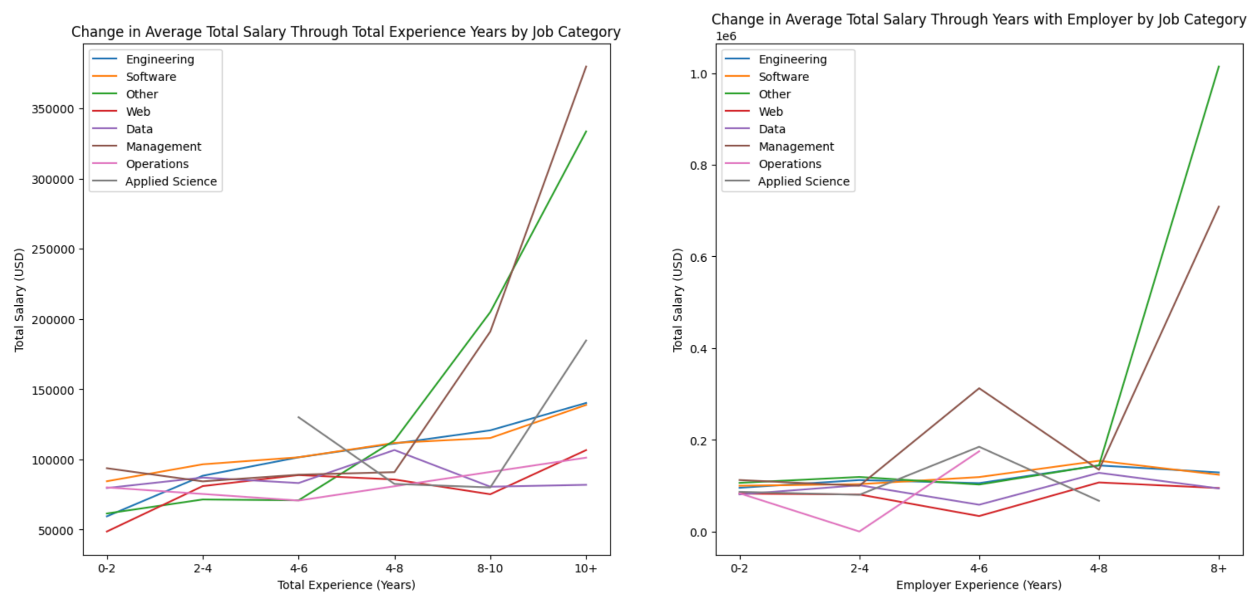

Figure 1.0: Change in average total salary in each job category by total experience.

Looking at the salaries trend by total and employer experience, it would be expected for every category to trend upward as time goes on for both graphs. However, the graphs show that some fields have significantly larger increases than others.

For salaries trend by total experience, web and engineering jobs experience overall growth. In contrast, operations and data jobs appear to stagnate in salary, remaining in the high five figures throughout the years. Management has the highest entry-level salary (0 total experience years) and experiences a significant increase after eight years of total experience, ending up as the highest paying category when the total experience years is high. While the "other" category starts with lower salaries early in the career, it experiences a similar boost to the management category after eight years. This supports the results from the previous highest-paying job graph. Freelancers or self-employed individuals are often unable to earn high salaries at the beginning due to immense investment, but once successful, they see great returns.

Interestingly, despite software being the most popular and highest-paying category in many countries, while salary does increase steadily throughout the years, it is not significant compared to other fields. Despite it having the second highest entry-level salary, just after management, its salary doesn't depend heavily on total experience. A reason behind this could be that the software field evolves rapidly, and in some cases, employees may reach a point where their skill set and experience have plateaued, and they are not acquiring new, in-demand skills. Due to this, this category values new talents, which are attracted by its high starting salary, resulting in the high supply and demand in this category.

For salaries trend by years with current employer, while the starting salary appears to be quite similar for all categories, only the management and "other" category sees significant growth as years with current employer increases. All other categories are relatively stagnant, experiencing minimal growth. A clear divide emerges between roles dependent on company experience versus those reliant on technical proficiency. Management and the "other" category exhibit substantial salary growth as tenure increases, reflecting the significance of trust, loyalty, and accumulated experience within an organization. In contrast, technical roles experience minimal growth, as pay is primarily determined by industry-specific skills rather than employer-specific tenure. This suggests that while technical expertise is an important determinant for salary in certain categories in the tech industry, pay for less technical roles often correlates with organizational longevity and leadership capabilities.

Analyzing salary trends by total and employer experience reveals interesting insights into the dynamics of various job categories. While one might expect a uniform upward salary trajectory over time, the data showcases distinct patterns across different fields. The graphs offer insights into how experience influences salary and which job categories exhibit more prominent pay growth with increased experience, both in total and with current employer. These trends underscore the multifaceted nature of salary dynamics within different job categories, influenced by industry demand, skill evolution, and corporate culture. Understanding these nuances is crucial for employers and employees to navigate career trajectories and compensation strategies effectively.

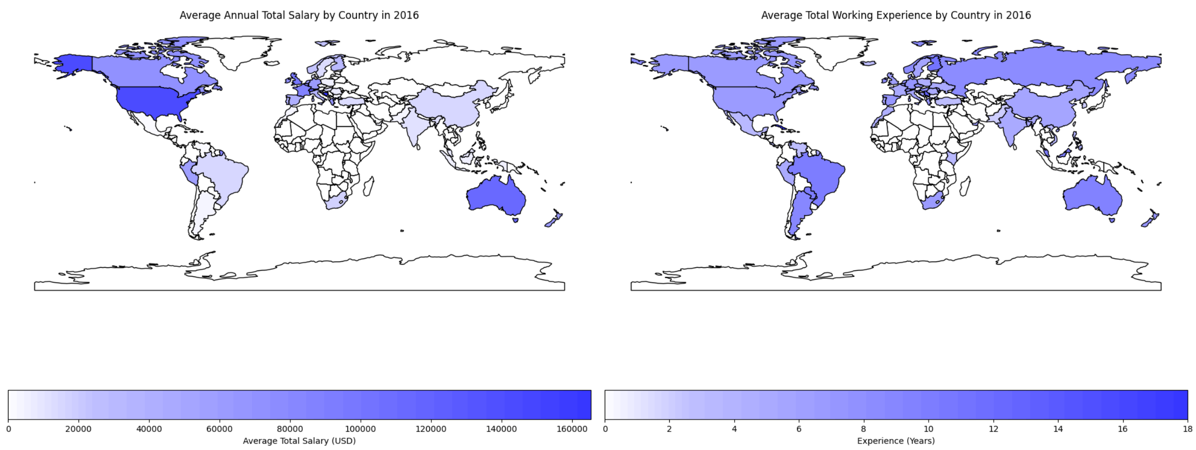

Figure 2.0: Average total salary and working experience by country.

The maps compare the average annual total salary with the average total working experience in countries worldwide. Higher average annual salaries can be mainly seen in countries in North America and Europe as well as Australia, as indicated by the darker purple shades on the left map. The map on the right, which depicts mean working experience in years, shows a similar concentration of higher values in such regions. Through the two maps, there appears to be a correlation between annual salary and working experience, as areas with higher average annual salaries also tend to have higher average total working experience. This implies that there may be a positive relationship between salary and experience in the tech industry, suggesting that individuals with more experience tend to command higher salaries.

This finding is supported by Figure 1.0, which shows an overall upward trend in salary in the majority of job categories as the experience years increase. The correlation could be attributed to several factors. Firstly, as individuals gain more experience in the tech industry, they often acquire specialized skills and knowledge for their job category that make them more valuable to employers, thereby justifying higher compensation. Moreover, companies may prioritize retaining experienced employees through competitive compensation, reinforcing the correlation between salary and experience.

However, Russia stands out as a notable outlier, with a low average salary but high average working experience. One possible explanation is Russia's economy's challenges and fluctuations, including currency devaluation, sanctions, and geopolitical tensions, which can impact salary levels and overall job market conditions. Additionally, the cost of living in Russia may be lower compared to other regions with higher average tech salaries, which could influence salary levels in the industry. This will be explored in the following map on the cost of living index by country.

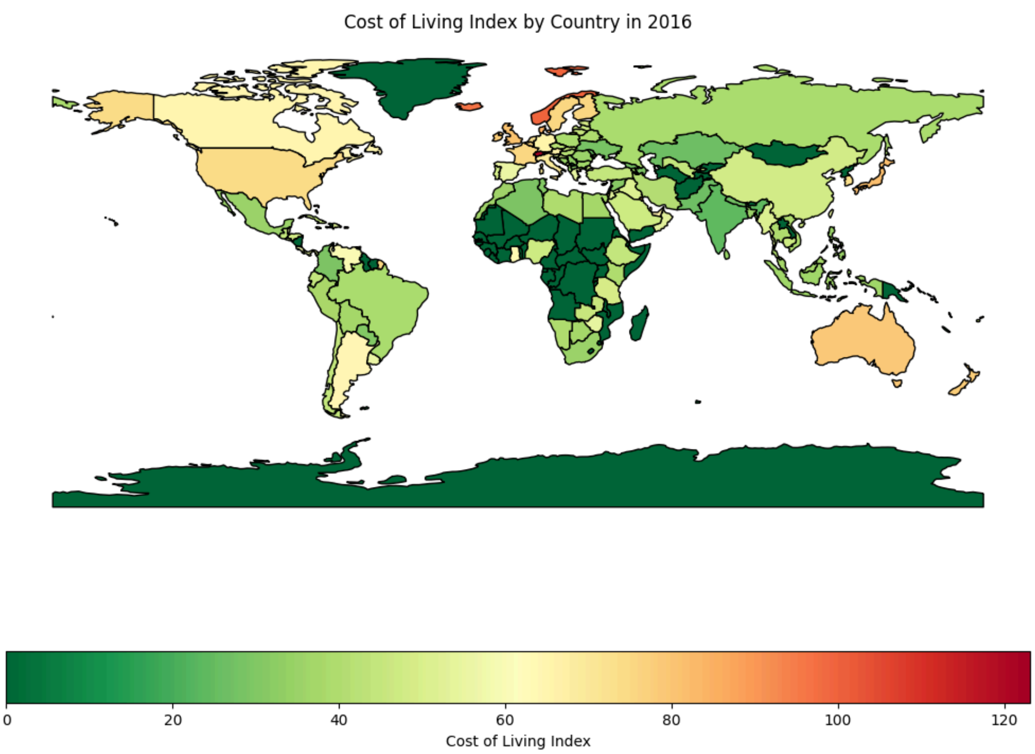

Figure 3.0: Cost of living index by country in 2016.

The maps illustrate the cost of living index by country worldwide in 2016. Areas in dark green, where the cost of living index is 0, represent countries for which the dataset does not contain data.

Upon examination, the highest cost of living is seen in countries/territories in Europe, notably Bermuda, Switzerland and Norway. Outside of Europe, countries such as the US, Australia, Argentina, Ghana, and Japan are among the countries with the highest cost of living index in their respective continents. The relationship between the cost of living and salary impacts the standard of living of individuals in that country. It is expected that tech salaries should be higher in countries with a higher cost of living index to compensate for the increased expenses associated with living in those regions and to ensure that employees can maintain a comparable standard of living relative to their counterparts in areas with lower costs of living.

Looking at Russia, the cost of living is relatively low. This provides evidence for the previously mentioned idea that the cost of living may be a reason why the country is an outlier in terms of its tech salary and experience levels. With a low cost of living index, necessities such as groceries and utilities are generally more affordable relative to other global tech hubs. As a result, salary offerings are lower due to the lower cost of living, even for highly skilled and experienced tech professionals. This results in low average base salaries compared to countries with higher living expenses.

Employers often adjust salaries based on local economic conditions and the availability of skilled workers, which determines the level of influence experience years can have on annual base salary. In regions with a high cost of living and intense competition for talent, experience may carry more weight in determining salary levels. For example, in countries where the cost of living is high, tech professionals with more years of experience or specialized skills may command significantly higher salaries compared to those in less competitive markets. Additionally, the availability of annual bonuses can vary depending on the region and the competitiveness of the job market. In high-cost areas, employers may offer more generous bonuses as part of their compensation strategy to attract and retain top talent. Overall, the cost of living index provides valuable insight into regional economic conditions, which shapes how much weight years of experience carry in determining the annual salary.

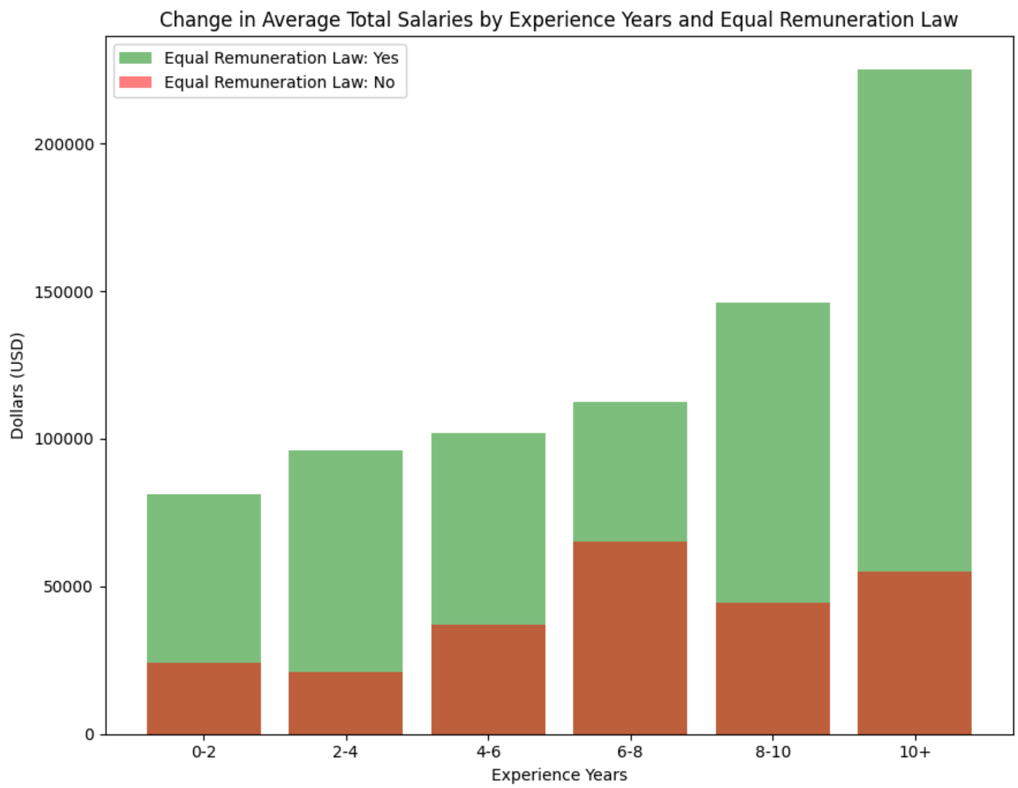

Figure 4.0: Change in Average total salaries by experience years, grouped by remuneration law.

The graph indicates the influence of equal remuneration laws on salaries within the tech industry, shedding light on the relationship between experience and compensation. Specifically, salaries for individuals working in countries with legislation ensuring equal pay for males and females exhibit a distinct pattern of growth. The growth follows an exponential trajectory as individuals accumulate more experience in the tech sector. This suggests that regulatory measures promoting pay equity positively impact salary advancement as the experience years increase.

Equal remuneration laws help to address gender pay gaps by ensuring that individuals with similar levels of experience and qualifications receive comparable pay, regardless of their gender. In the tech industry, where gender disparities in pay and representation have been well-documented, these laws can help to level the playing field and ensure that women are compensated fairly for their experience and expertise, which raises the average total salary. By ensuring that individuals are rewarded fairly for their experience and qualifications, equal remuneration laws can incentivize tech professionals to invest in skill development and career advancement, driving innovation and productivity growth, which tends to increase salary.

Conversely, in countries lacking such legislation, the correlation between experience and salary growth is less apparent, indicating that without regulatory frameworks, salary progression may not align with increasing experience in the field. Moreover, the data reveals that salaries for individuals in countries with equal remuneration laws consistently surpass those in nations without such regulations. Furthermore, the salary disparity between these two groups of countries widens as individuals accumulate more years of experience in the tech industry. This underscores the significant role of equal remuneration laws in not only shaping salary growth trends but also in maintaining higher overall compensation levels.

While experience undoubtedly influences salaries in the tech sector across all regions, the presence of equal remuneration laws amplifies this effect, ensuring fair and equitable compensation practices and resulting in more robust and sustained salary increases over time.

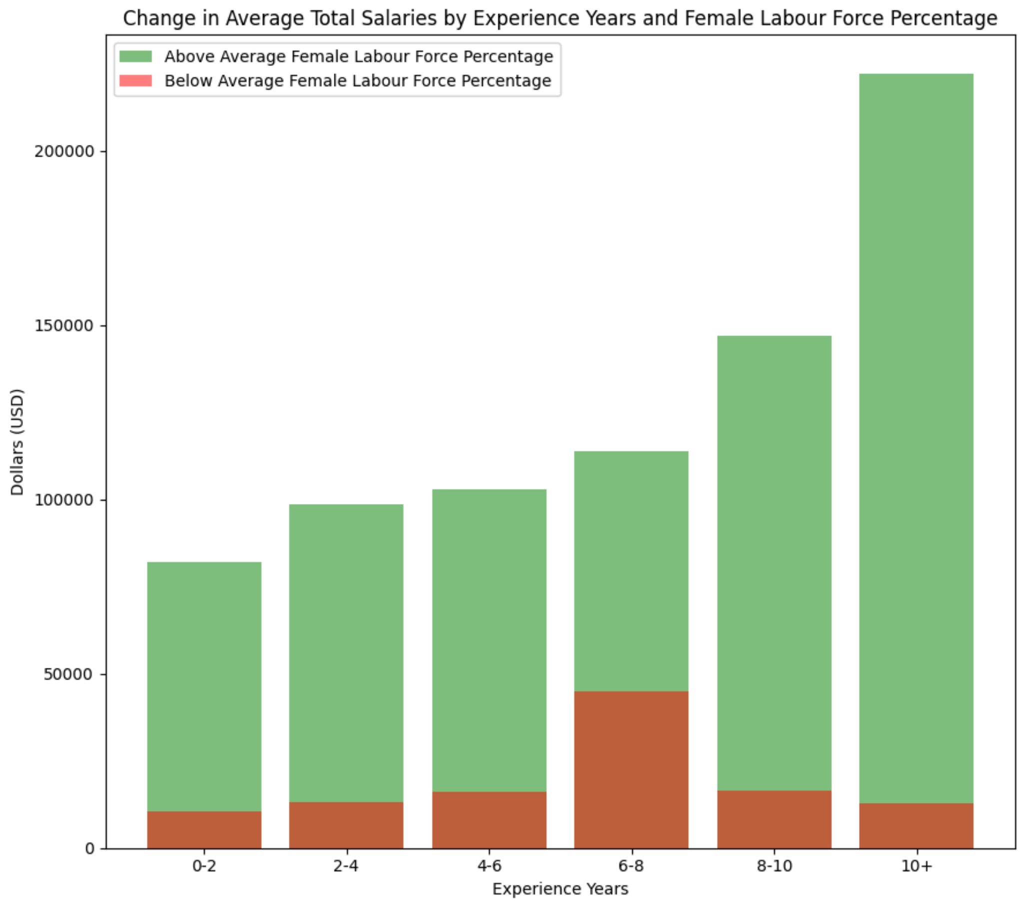

Figure 5.0: Change in Average total salaries by experience years, grouped by female labour force percentage.

This graph shows that individuals working in countries with above-average female labour force participation exhibit a similar pattern of salary growth to those with equal pay legislation, with salaries following an exponential trajectory as the experience years increase. This suggests that social dynamics supporting gender equality in the workforce contribute to significant salary progression. Conversely, in countries with below-average female labour force participation, salaries remain consistently lower across all experience years in countries with below-average female labour force participation, echoing the disparities observed in nations without equal pay legislation.

In correlation with the graphs regarding equal remuneration laws, these findings collectively emphasize the critical role of regulatory frameworks and societal factors, such as female labour force participation rates, in shaping salary growth trends within the tech industry. Countries with progressive legislation and higher female labour force participation tend to exhibit more favourable salary trajectories as experience years increase, whereas those without such measures experience slower or less consistent salary progression.

RESULTS

The economic relationship between annual total salary and experience in the tech field is likely to be non-linear rather than linear. Economic theories such as human capital theory and diminishing marginal returns provide insight into this phenomenon.

Human capital theory suggests that individuals accumulate skills and knowledge as experience increases, leading to an increase in productivity and, consequently, higher pay. However, this accumulation of skills is not necessarily linear; rather, it tends to exhibit diminishing returns as individuals reach a certain level of expertise in their field. The theory can be applied to certain categories within the tech field that rely on technical expertise, such as software, engineering and data. Due to rapid technological advancements and evolving skill requirements in these categories, employees who have plateaued in their skills may experience difficulty commanding higher pay, even with years of experience in the industry, as seen in the change in the average total salary graph by category. The relationship between experience and pay also varies across these different paths, with individuals in management roles typically commanding higher salaries as experience increases. These leadership roles depend heavily on trust and loyalty within an organization, which is accumulated through experience.

Local economic conditions within a country also play a significant role in shaping the relationship between annual pay and experience in the tech industry by influencing the cost of living, industry composition and government policies. In regions with a high cost of living, employers may offer higher salaries to attract and retain tech talent, regardless of experience level. This can lead to disparities in pay between regions, even for individuals with similar levels of experience, which can be seen in the map comparisons of the average annual total salary with the average total working experience and cost of living index across the world.

Government policies and regulations, such as equal remuneration laws, significantly impact the relationship between annual pay and experience in the tech industry. These laws mandate that employers provide equal pay for equal work, regardless of gender, race, or other protected characteristics. By addressing gender pay gaps and promoting transparency in salary practices, equal remuneration laws empower tech professionals to negotiate fair compensation based on their experience and qualifications. This fosters a more skilled and diverse workforce, driving competition and higher average salaries in the tech sector.

Human capital theory, diminishing marginal returns, local economic conditions, and government policies all play a role in shaping the tech labour market. As a result, the economic relationship between annual total salary and experience in the tech field is complex and likely to be non-linear rather than linear.

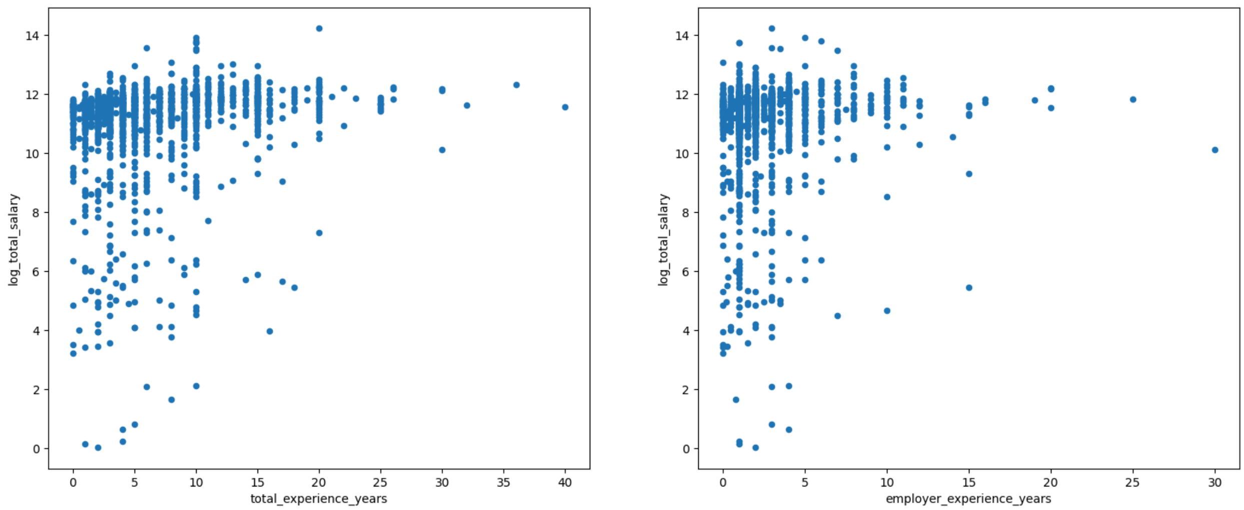

Figure 6.0: Relationship between total/employer experience years and log of total salary without outliers.

Using a scatterplot to see whether there exists an obvious relationship between total and employer experience years and total salary, two notable outliers with significantly higher total salaries were identified. While these outliers may represent valid data points, they were removed from the analysis to enhance the performance of the following regressions. This process ensures that the data used in the regressions is more representative of the overall trend. Due to the relationship between annual total salary and experience in the tech field being likely to be non-linear, the log of total salary is taken and used in regressions.

Linear Regressions

Table 3.0: Linear regression results, part 1.

					Dependent variable: log_total_salary
	(1)	(2)	(3)	(4)	(5)
const	10.685***	10.693***	10.518***	10.298***	10.319***
	(0.068)	(0.070)	(0.112)	(0.146)	(0.146)
emp_exp*job_type					0.057
					(0.038)
employer_experience_years		-0.009	-0.007	-0.011	-0.044
		(0.019)	(0.019)	(0.019)	(0.029)
job_type			0.203**	0.509***	0.466***
			(0.102)	(0.165)	(0.167)
tot_exp*job_type				-0.041**	-0.057***
				(0.018)	(0.020)
total_experience_years	0.051***	0.053***	0.055***	0.084***	0.094***
	(0.008)	(0.009)	(0.009)	(0.015)	(0.017)
Observations	1510	1510	1510	1510	1510
R^2	0.025	0.026	0.028	0.032	0.033
Adjusted R^2	0.025	0.024	0.026	0.029	0.030
Residual Std. Error	1.655 (df=1508)	1.655 (df=1507)	1.653 (df=1506)	1.651 (df=1505)	1.650 (df=1504)
F Statistic	39.258*** (df=1; 1508)	19.741*** (df=2; 1507)	14.506*** (df=3; 1506)	12.307*** (df=4; 1505)	10.302*** (df=5; 1504)
Note:					*p<0.1; **p<0.05; ***p<0.01

The regressions conducted aim to understand the relationship between experience (measured in total_experience_years and employer_experience_years), job type and total salary in the tech industry. Job type categorizes the 8 job categories into two types, technical and non-technical roles. Data, engineering, operations, software, web, and applied science are technical roles, while management and other category are considered non-technical. The regressions interact with experience and job type to compare the difference in salary between job types as total/employer experience increases. The regressions seek to quantify the value and impact of job category and experience on salaries for different roles within the tech industry.

Looking at the performance assessment measures, R^2 indicates the proportion of the variance in the dependent variable (log_total_salary) that is explained by the independent variable. Looking at regression (1), R^2 is quite low at 0.025, suggesting that only 2.5% of the variability in log_total_salary is explained by total_experience_years and the R^2 of 0.026 for regression (2) suggests that only 2.6% of the variability in log_total_salary is explained by total_experience_years and employer_experience_years. The low R^2 value indicates that experience alone explains only a small portion of the variability in salaries within the tech industry, and other factors not included in the models also significantly influence salaries. The R^2 is slightly higher when job_type is incorporated into the model, but not by a significant amount. However, the p-value associated with the F-statistic is highly significant for all 5 regressions, indicating that the regression models as a whole are statistically significant. Therefore regression results suggest that while experience and job type do have a statistically significant impact on salaries in the tech industry, it explains only a small portion of the overall variation. Therefore, to fully understand the determinants of salaries in this industry, it's essential to consider other factors alongside experience and job type.

The coefficient for tot_exp * job_type and emp_exp * job_type indicates the difference in salary between job types as years of total/employer experience increase. The negative coefficient value for tot_exp * job_type indicates that the wage gap between technical and non-technical jobs is greater for workers with lower total experience. The positive coefficient value for emp_exp * job_type means that the salary difference increases with years of experience with the employer, meaning that the wage gap (between job types) is greater for workers with a higher number of years at the company.

Table 4.0: Linear regression results, part 2.

	Dependent variable: log_total_salary		
	(1)	(2)	(3)
High_COL	2.125***	2.451***	2.432***
	(0.102)	(0.167)	(0.173)
const	8.938***	8.656***	8.672***
	(0.104)	(0.155)	(0.160)
emp_exp*High_COL			0.020
			(0.046)
employer_experience_years	-0.004	-0.004	-0.020
	(0.016)	(0.016)	(0.042)
tot_exp*High_COL		-0.052**	-0.057**
		(0.021)	(0.024)
total_experience_years	0.047***	0.093***	0.097***
	(0.008)	(0.020)	(0.022)
Observations	1510	1510	1510
R^2	0.244	0.247	0.247
Adjusted R^2	0.243	0.245	0.245
Residual Std. Error	1.458 (df=1506)	1.456 (df=1505)	1.456 (df=1504)
F Statistic	162.137*** (df=3; 1506)	123.508*** (df=4; 1505)	98.789*** (df=5; 1504)
Note:			*p<0.1; **p<0.05; ***p<0.01

The 3 regressions above aim to understand the relationship between experience, cost of living, and total salary in the tech industry. The cost of living is an important factor affecting salaries, especially in the tech industry, where valued professionals are likely to be concentrated in high-cost areas. The 'high_COL' variable represents countries where the cost of living index is greater than the average cost of living index across all countries in the dataset. The regressions interact the cost of living index with experience to compare the difference in salary in countries with higher and lower costs of living as total/employer experience increases. The regressions seek to quantify the value and impact of the cost of living and experience on salaries within the tech industry.

For regression (1), the R^2 value of 0.244 indicates that this model explains 24.4% of the variability in log_total_salary, which is substantially higher than the regression models in the previous table. This suggests that cost of living index may have a stronger explanatory power than job type. The p-value associated with the F-statistic for all 3 models is highly significant, meaning that the regression models as a whole are statistically significant. The coefficient for High_COL is 2.125 (***), indicating that being in a high-cost-of-living area is associated with a substantial increase in log_total_salary compared to areas with lower costs of living. The regression results suggest that location cost of living is a significant factor influencing salaries in the tech industry. This aligns with previous findings as professionals in high-cost-of-living areas tend to earn substantially more to compensate for the higher living expenses, and there is greater competition between companies to attract/retain talent.

Looking at regressions (2) and (3), the coefficient for tot_exp * High_COL and emp_exp * High_COL indicates the difference in salary between locations with high and low cost of living as years of total/employer experience increase. The negative coefficient value for tot_exp * High_COL indicates that if years of total experience increase, the salary gap between workers in high and low-cost-of-living locations decreases. The negative coefficient value for tot_exp * High_COL indicates that if years of total experience increase, it decreases the salary gap between workers in high and low cost of living locations. In other words, the effect of total experience on salary is such that it tends to equalize the salary discrepancy between high and low cost of living areas. This could suggest that experienced workers are better able to negotiate salaries that reflect the cost of living, regardless of location. The positive coefficient value for emp_exp * High_COL means that the salary difference increases with years of experience with the employer, meaning that the wage gap (between high and low cost of living locations) is greater for workers with a higher number of years at the company. The longer an employee stays with a company, the more pronounced the salary gap becomes between high and low cost of living areas. This could indicate that companies reward tenure with higher salaries, but these increases are proportionally larger in high cost of living areas, thereby exacerbating the wage gap between different locations.

Table 5.0: Linear regression results, part 3.

	Dependent variable: log_total_salary			
	(1)	(2)	(3)	(4)
const	9.046^{***}	4.403^{***}	8.170^{***}	7.832^{***}
	(0.150)	(0.536)	(0.590)	(0.618)
employer_experience_years	-0.008	-0.008	-0.008	-0.007
	(0.018)	(0.017)	(0.016)	(0.016)
female_labour_force		0.120^{***}	-0.008	0.007
		(0.013)	(0.016)	(0.018)
high_female_labour_force			2.741^{***}	2.196^{***}
			(0.217)	(0.368)
remuneration_law	1.817^{***}	0.975^{***}	0.455^{***}	-0.051
	(0.149)	(0.172)	(0.169)	(0.323)
remuneration_law*high_female_labour_force				0.707^{*}
				(0.385)
total_experience_years	0.049^{***}	0.046^{***}	0.043^{***}	0.043^{***}
	(0.009)	(0.009)	(0.008)	(0.008)
Observations	1510	1510	1510	1510
R^2	0.114	0.159	0.240	0.241
Adjusted R^2	0.112	0.157	0.237	0.238
Residual Std. Error	1.579 (df=1506)	1.539 (df=1505)	1.463 (df=1504)	1.462 (df=1503)
F Statistic	64.281^{***} (df=3; 1506)	71.079^{***} (df=4; 1505)	94.771^{***} (df=5; 1504)	79.661^{***} (df=6; 1503)
Note:				$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 5.0 explores the influence of several factors, including equal remuneration law, female labour force percentage, experience and the interaction between remuneration laws and a high female labour force on total salary within the tech industry. The 'high_female_labour_force' variable represents countries where the female labour force percentage is greater than the average percentage across all countries in the dataset. By including gender-related variables, the

regressions aim to understand how gender-related dynamics and regulatory frameworks affect salary outcomes alongside factors like experience.

The R^2 value of 0.114 for regression (1) indicates that this model explains 11.4% of the variability in log_total_salary, which is higher than regression (3) in Table 3.0 with experience and job type but lower than regression (1) in Table 4.0 with experience and cost of living index. This suggests that equal_remuneration_law has some explanatory power, however not as notably as cost of living. The p-value associated with the F-statistic is highly significant for all 4 models, indicating that the regression models as a whole are statistically significant. Looking at the R^2 value for regression (2), it is higher than in regression (1), suggesting that female labour force percentage is also important. Interestingly, in regressions (3) and (4), with the added dummy variable and the interaction term, female_labour_force and remuneration_law become no longer statistically significant. This may indicate that these two variables hold the most explanatory power when looked at together, and whether the female labour force percentage is relatively high or low is more important than the exact value. The p-value associated with the F-statistic is highly significant for all 4 models, indicating that the regression models as a whole are statistically significant.

In regression (4), the coefficients for female_labour_force and high_female_labour_force are both positive, indicating that higher female labour force participation, particularly in industries with a high female presence, is associated with higher log_total_salary. The interaction term remuneration_law * high_female_labour_force is also positive and statistically significant, indicating that the effect of remuneration laws on log_total_salary is more pronounced in industries with a high female labour force. This aligns with what is expected, as in countries with equal remuneration laws, higher female labour force participation should lead to higher salaries. The regression results suggest that gender-related dynamics and regulatory frameworks together, along with factors like experience, play a role in shaping salaries within the tech industry.

Regression Tree

The objective function of the regression tree is the following:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(\log(\text{total\_salary}_i) - (\beta_0 + \beta_1 x_{1,i} + \beta_2 x_{2,i} + \beta_3 x_{3,i} + \beta_4 x_{4,i} + \beta_5 x_{5,i} + \beta_6 x_{6,i} + \beta_7 x_{7,i} + \beta_8 x_{8,i}))^2$$

Where N is the total number of data points, log(total_salary$_i$) represents the actual total salary for the i-th data point, $x_{1,\, i}$, $x_{2,\, i}$, …, $x_{8,\, i}$ represent the values of the independent variables (location latitude, location longitude, total experience years, employer experience years, cost of living index, remuneration law, female labour force, job type, respectively) for the i-th data point. The beta coefficients are the regression coefficients representing the effect or weight of the corresponding independent variable on the dependent variable.

The MSE objective function calculates the average squared difference between the actual total salaries of the data points and their predicted total salaries by the regression tree model. This function is used to measure how well the model is performing in terms of minimizing the errors in its predictions. The goal of the regression tree algorithm is to find the tree structure that minimizes this MSE, effectively creating splits in the data based on the given predictor variables to predict total salary as accurately as possible.

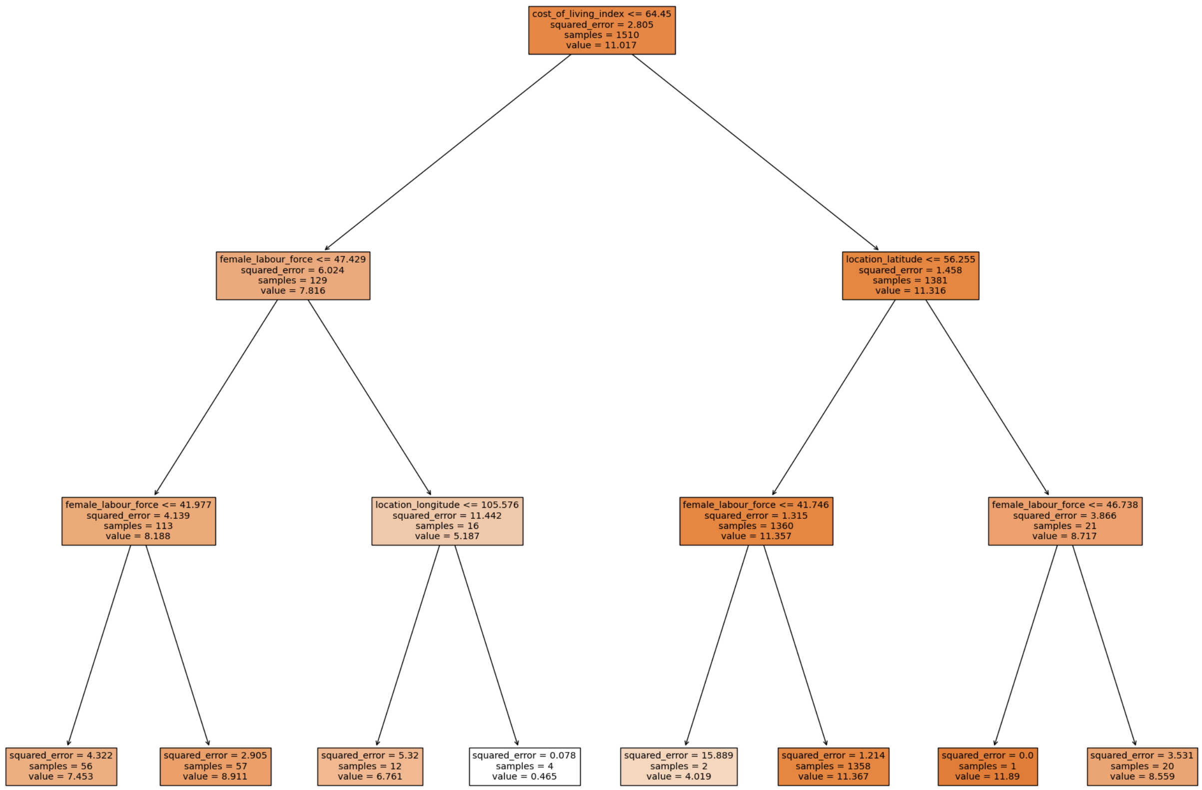

Figure 7.0: Regression tree.

Looking at the regression tree, it can be seen that for workers in countries with the cost of the living index being less than or equal to 64.45, the total salary is then determined by the female labour force percentage in the country, with the longitude of the location also being a contributing factor when female labour force percentage is greater than 47.429. For workers in countries with a cost of living index greater than 64.45, the salary is then further split by latitude of the location and female labour force percentage. Overall, the regression tree shows that the cost of living index, female labour force, longitude, and latitude of location are the main variables that determine salary.

Random Forest

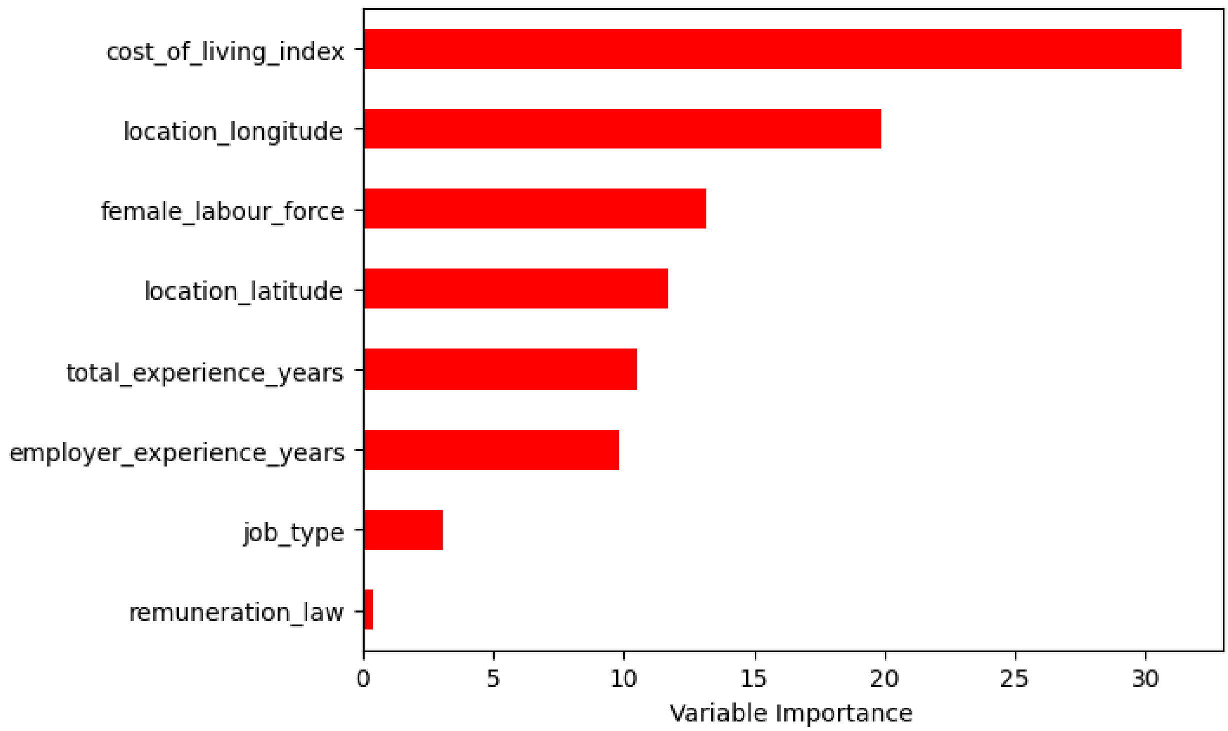

Figure 8.0: Variable importance matrix.

The importance matrix using a random forest model shows the cost of living index as the most important variable in determining total salary, followed by location longitude and latitude, female labour force percentage, and experience years. Equal remuneration law and job type are the least important compared to other variables, with variable importances below 5 for both. It can be seen that while experience does impact total salary to some extent, it is not the most statistically significant factor in determining total salary in the tech industry.

CONCLUSION

This paper addresses the economic question: "To what extent does experience influence salaries in the tech industry?" By focusing specifically on the tech sector, analyzing diverse types of experience, and considering temporal dynamics such as job tenure and the influence of human capital accumulation and job search activities, this research has provided valuable insights into the nuanced relationship between experience and salaries within technology-driven roles.

Through the application of a random forest model, this paper identified several key determinants of salary levels, shedding light on the complex interplay of economic, geographical, and demographic factors in shaping compensation within the sector. Notably, the analysis highlights the importance of the cost of living index, underscoring the significant influence of regional economic conditions on salary levels. Moreover, the inclusion of variables such as location longitude and latitude emphasizes the role of geographic considerations in salary determination, reflecting the diverse landscape of tech hubs and regional disparities in compensation. Surprisingly, the female labour force percentage emerges as a noteworthy factor, suggesting potential gender-related dynamics in salary differentials within the industry. Importantly, while experience years exhibit some impact on total salary, the findings indicate that they are not the most statistically significant factor, challenging conventional assumptions about the centrality of experience in salary determination. Additionally, variables such as the equal remuneration law and job type demonstrate relatively lower importance, suggesting a limited influence on salary levels within this analysis framework.

This study contributes to a nuanced understanding of salary dynamics in the tech industry, highlighting the need for a comprehensive approach that considers a broad array of economic, geographic, and demographic factors. Future research endeavours should consider how education factors such as years of schooling influence salary, especially at the early stages of one's career. Additionally, comparative analyses should be conducted with similar studies on salaries in the tech industry to verify the integrity and accuracy of the results in this study. Looking at larger datasets will offer a broader perspective, contributing to a more robust and reliable understanding of the relationship between experience and tech salaries.

Works Cited

Bagger, J., Fontaine, F., Postel-Vinay, F., & Robin, J.-M. (2014). Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics. The American Economic Review, 104(6), 1551–1596. http://www.jstor.org/stable/42920859

Goldsmith, A. H., & Veum, J. R. (2002). Wages and the Composition of Experience. Southern Economic Journal, 69(2), 429–443. https://doi.org/10.2307/1061681

Lazear, E. (1976). Age, Experience, and Wage Growth. The American Economic Review, 66(4), 548–558. http://www.jstor.org/stable/1806695

Williams, N. (1991). Reexamining the Wage, Tenure and Experience Relationship. The Review of Economics and Statistics, 73(3), 512–517. https://doi.org/10.2307/2109577