Data Analysis Report

IT Data and Salary Survey (EU Region) 2019- 2021

An anonymous unguided data and salary survey has been conducted annually since 2015 among European IT specialists with a stronger focus on Germany. The dataset contains rich information about the salary and data patterns among IT professionals in the EU region and offers some great insights.

Special thanks to Viktor Shcerban and Sergey Vasilyev for collaborating on the survey for 2019 and 2020. Furthermore, the tale of the unguided data and its analysis will be uncovered in the analysis.

- **Data:** Each year of this datasets has a comma-seperated value(csv) file which contains information about the survey in a spreadsheet format, where each column of the spreadsheet has a different story on its own, for example, the city column tells the particular location in European Union where the IT professional plies their skills.

- **Method/Analysis:** First, extracted the dataset from Kaggle,skimmed through it on Excel and found loads of missing values and filled them using the fill method on Excel since one can not retort to any one for the missing values.

    Next, uploaded it to my analysis notebook, imported all necessary libraries for analysis that would help in the analysis of these datasets, cleaned the data,also cut the columns down to tailor it to my analysis before going full on into the analysis. Since the data is unguided,went on to ask and answer the following questions to appease my curiosity.

- **Questions:**

    1. The count of both gender and what gender has the higher count.
    2. The minimum and maximum salary earnings per annum.
    3. What job position has the highest and lowest frequency in the survey.
    4. The top three highest paid salary earners in the whole survey.
    5. The city with the highest and lowest frequency in the datasets.
    6. The youngest and oldest IT professionals.
    7. The language popularly used in the IT industry of the EU Region.
    8. The count of people in the range of 20-30,31-40 and 41 and above years in IT.

9. .The count of professionals that earn 50 thousand euros or below and those that make 50 thousand and above.
10. The noticeable difference in the IT industry in the three years of survey.

● **Results:**

1. The survey shows that there are 2702 males and 492 females that took part in this exercise,hence there is a higher ratio of male to female.
2. The minimum annual salary is 6000.00€ and the maximum annual salary earned by an IT professional is 850000.00€.
3. Software Engineering role has the highest frequency and AI management has the least frequency in regards to Job Positions in the survey.
4. The top three salaries; A software engineer earns 850000€, next is a DevOps engineer who earns 300000€ and finally another software engineer that racks up 250000€ per annum.
5. Berlin and Aachen are the cities with the highest and least frequency respectively.
6. The oldest IT professional according to the survey is a 69 year old male product analyst while the youngest is a 20 year old male machine learning engineer.
7. English stands as the most popularly known language among the IT professionals in the EU region followed by German.
8. The survey proffers a count of 1263 persons in the range of 20-30,a whooping 1693 for 31-40 years and 242 for persons in the range of 41 years and above.
9. The analysis shows that there are 406 professionals who make less than 50000 euros per annum and 2790 that make more than 50000 euros.
10. In the three years in which survey, there was a rise in IT professionals in 2020 which was obviously due to the Covid-19 pandemic, analysis shows that in 2019,that the survey could count 970 professionals but by 2020 it skyrocketed to 1251 professional evidently due to lockdown which made more people to transition into IT and it dropped down to 974 in 2021 when the lockdown was lifted and most people returned back to their old jobs.

    This also affected the total income per annum in the yearly count,where 70.3 million euros was the total salary earned per annum in 2019 and then there was a surge to 86.6 million in 2020 before a drop to 70.7 million at the end of 2021.

- **Conclusion:**

 The results of these analysis and their visualisation can all be found in the analysis notebook and visualisation worksheet which are all available in the project folder and looking forward to the end of the year 2022 to analyze its IT survey.

- **Appendices**

 – Microsoft Excel for data exploration
 – Python/ SQL for my analysis.
 – Tableau for data visualisation.
 – Google Docs for Report Writing.

 .