

INFSCI 2415 Information Visualization

Final Report

GitHub: https://github.com/Chandler-He/Infor_Visual.git

Name – **Peilin He**

ID – **4447252**

Date – 11/28/2022

Email – peh53@pitt.edu

Cyber Security Experience Levels Salary Analyzation

Abstract

Cybersecurity is the practice of deploying people, policies, processes, and technologies to protect organizations, their critical systems and sensitive information from digital attacks. It is the protection of computer systems and networks from information leakage, theft, damage or spread out of the whole network. The field has become of significance due to the expanded reliance on computer systems, the Internet, and wireless network standards such as Bluetooth and Wi-Fi, and due to the growth of smart devices, including smartphones, televisions, and the various devices that constitute the Internet of things (IoT). Cybersecurity is also one of the significant challenges in the contemporary world, due to the complexity of information systems, both in terms of political usage and technology. Its primary goal is to ensure the system's dependability, integrity, and data privacy. In this project, I will demonstrate the salary for the level of experience in Cyber Security based on location, job title, residency, company size, etc.

Keyword: cybersecurity, salary, job title, position, level of experience.

Data Visualization

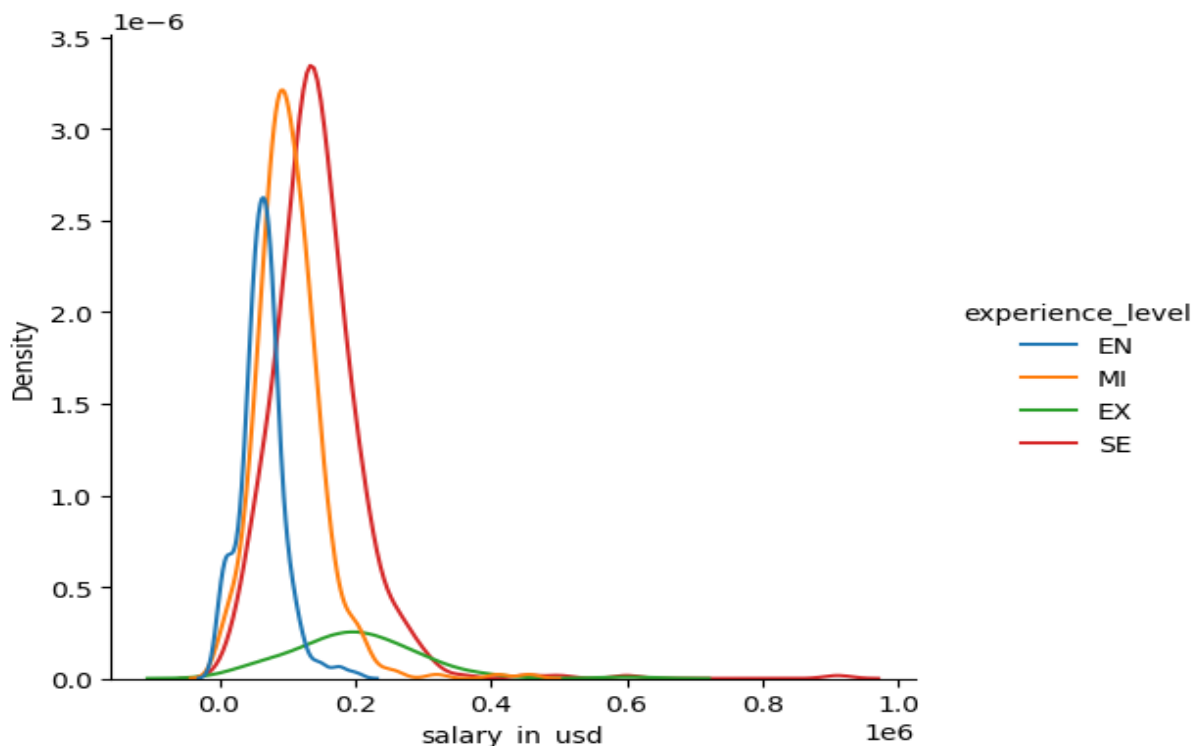


Fig. 1

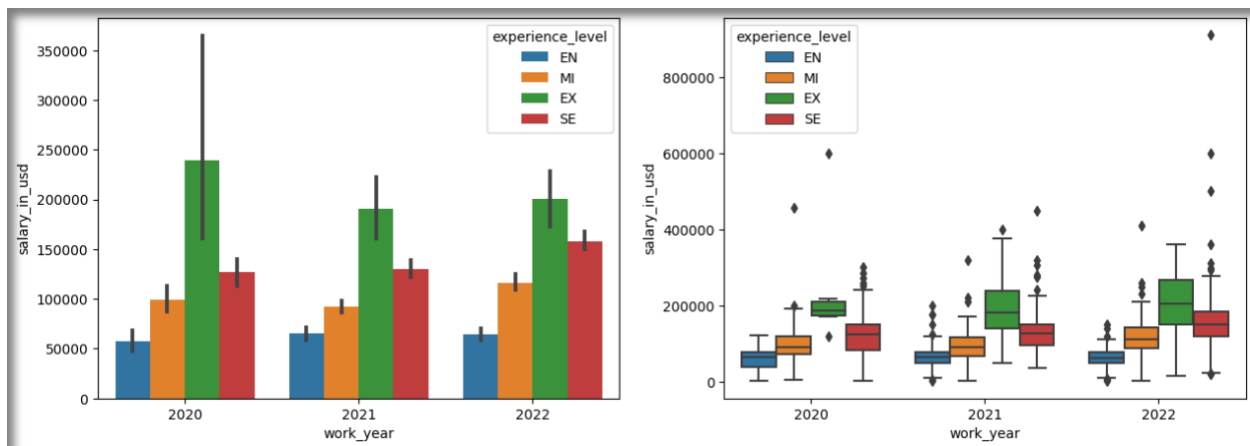


Fig. 2

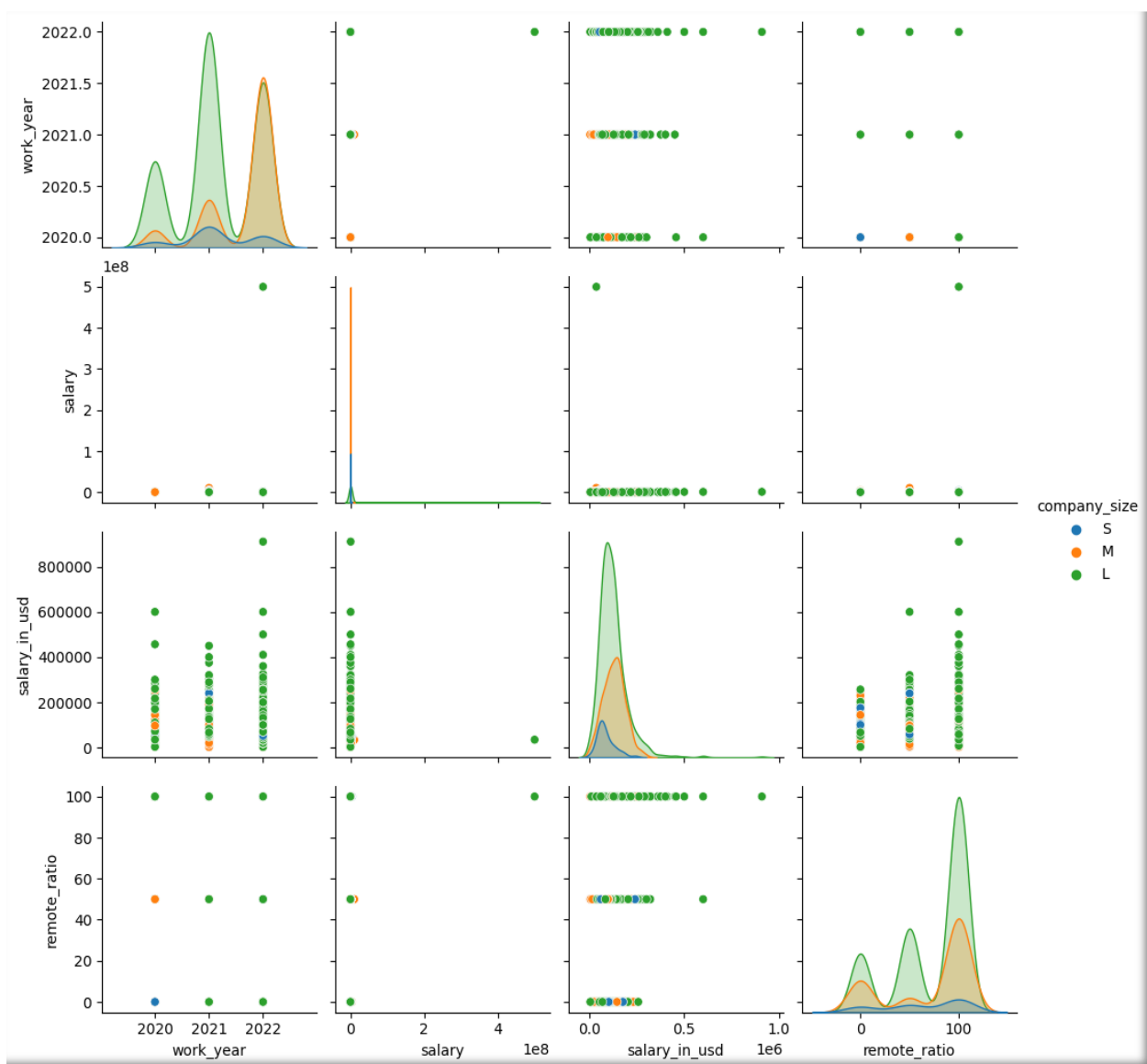


Fig. 3

Data Analyzation

- ⑩ Fig. 1 uses Normal Distribution to demonstrate the density of four experience levels which write as: Entry (EN), Mid-L (MI), Senior-L (SE) and Executive (EX) corresponding to the color in the graph. The x-axis is representing the salary in USD (per 100K). We can infer that as the level elevating, the co-efficient becomes bigger but the density drops. It adopts **distplot** imported from **seaborn** package and **dense** from **tensorflow** to emulate the trends in density as a function of salary. From this normal distribution there are also significant differences in salaries for different experience levels of jobs.
- ⑩ Fig. 2 is the derivative from Fig. 1, it specifies the year ranging from 2020 – 2022, demonstrates the salary distribution for these four levels of experience. Salary in every year by experience level use **bar plot** and **box plot** from **matplotlib.pyplot**. In box plot can easily detect the outlier. From the derivative we can incur that the tendency is the same with Fig. 1, and the salary for all the categories are all increasing, besides, the percentage of executive is rising.
- ⑩ Fig. 3 which is the main outline graph from the visualization part, it analyzes the entire dataset by using **pairplot** and **tensorflow** as well. It compares from work_year, salary, remote_ratio to company_size, etc. There are three lines represent with blue, orange, and green with respect to the big, mid, and small company size. It is clear that the correlation amongst each other.

References:

1. https://www.sas.com/en_us/insights/big-data/data-visualization.html
2. <https://socviz.co/>
3. https://en.wikipedia.org/wiki/Data_and_information_visualization
4. <https://www.sisense.com/blog/10-useful-ways-visualize-data-examples/>