# Team Project Phase II

M124020029 黃政曦
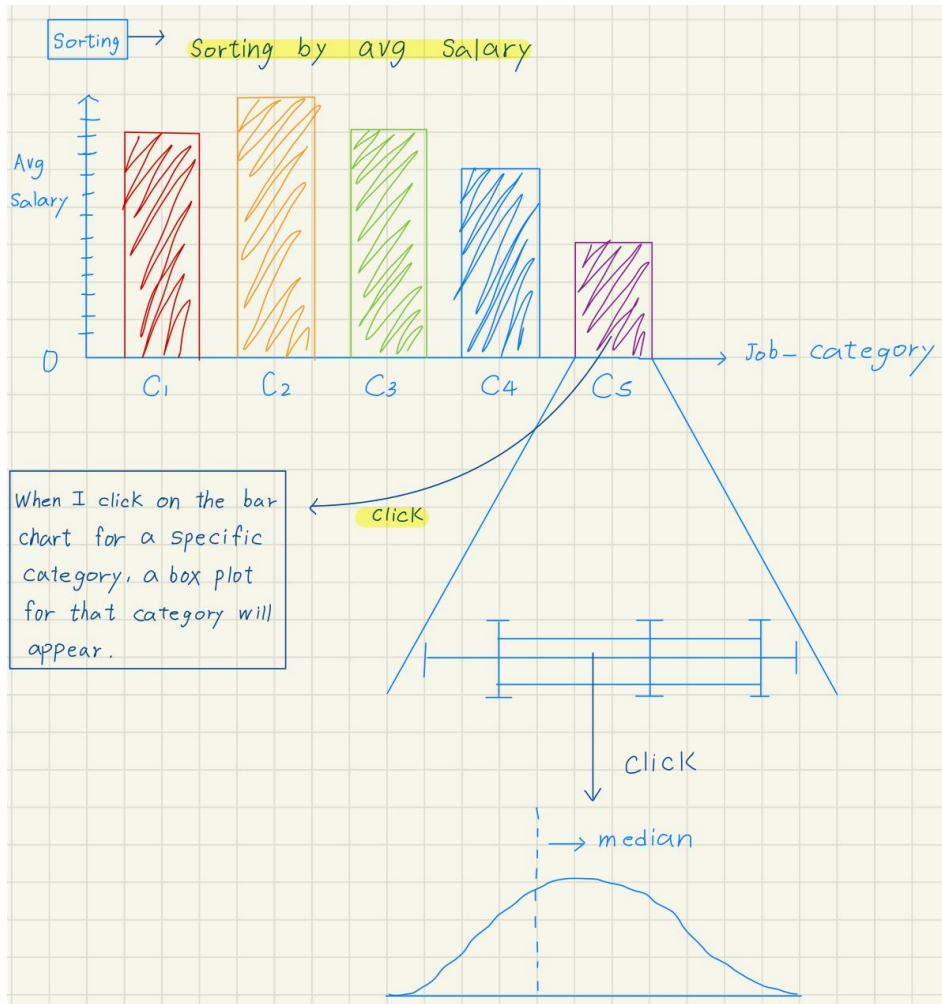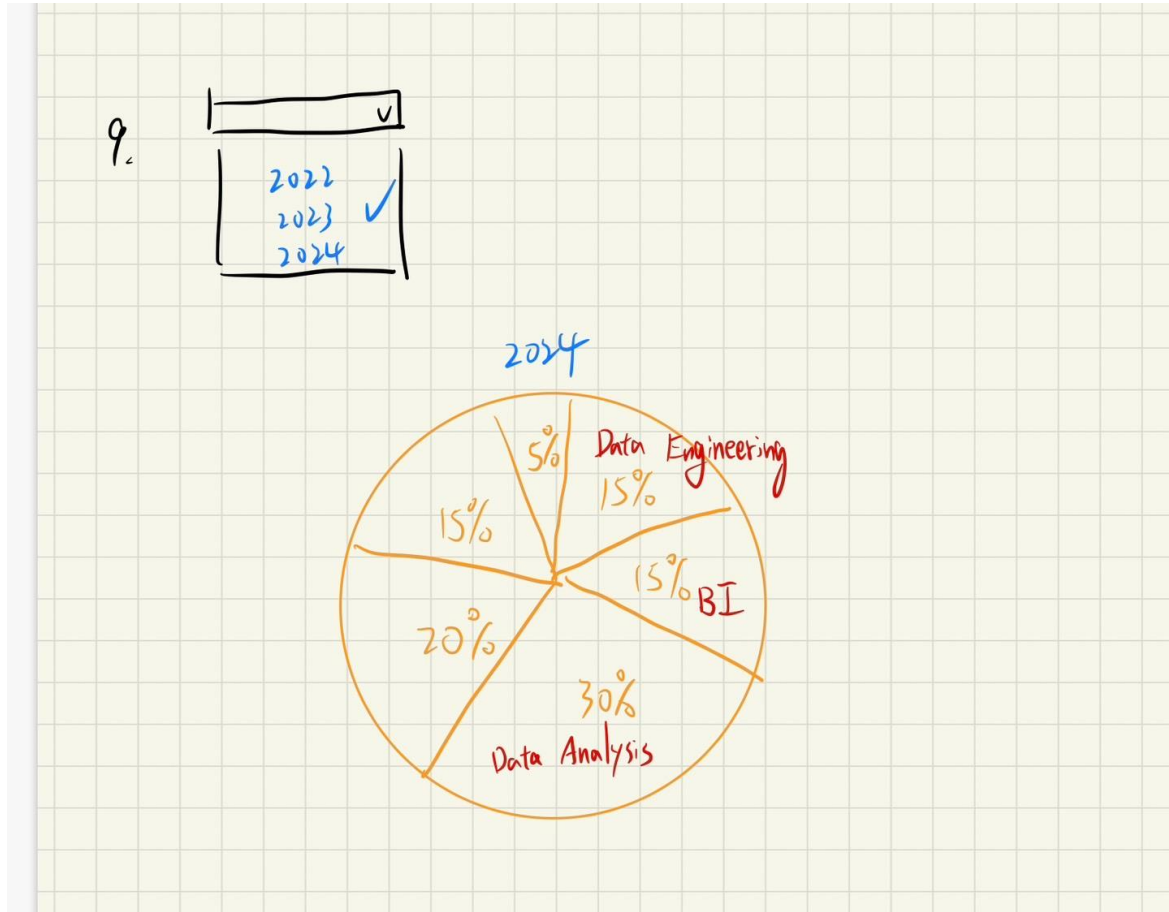
M124020033 毛天楷

M124020044 林聖皓

# Dataset

- We choose：Jobs and Salaries in Data field

- Reason：
  1. Compared with the spotify data set, this data set has more years of data.
  2. Compared with the spotify data set, this data set is better visualized
  3. The multi-dimensional information provided by this dataset.
  4. We are more interested in this data set than the spotify data set

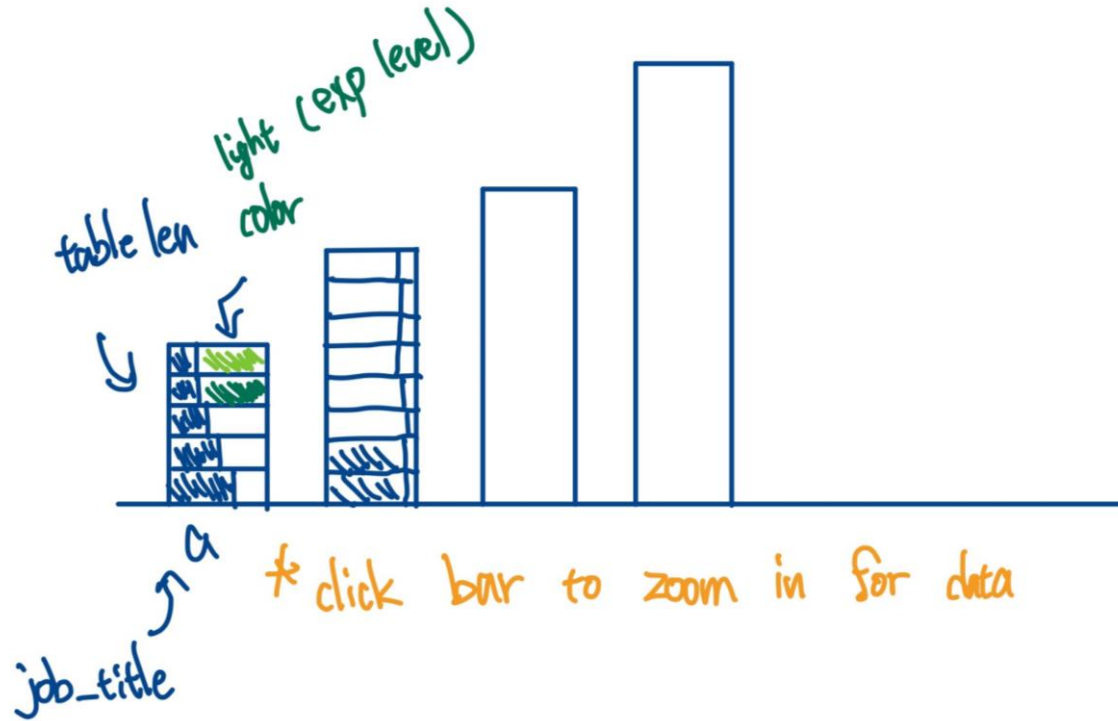# Q：Display the average salary and median salary for each Job_category.



- Critique：

  - pro：Using box plots can help ensure that outliers do not affect the results.

  - Con：Additional space is required to display box plots.

# Q. To show how job_category changes with different work_year



- Critique：
  - pro：Using pie chart is easy for comprehending the propotions of job categories based on different years

  - con：If there isn't enough data, then there may be some bias through this proportion

# Q. What the job_title distribution in a single job_category



- Visualization：
  x axis - job_title
  (table lens with a categorical variable)
- y axis – number of people

# Members' Contribution

- Each member has to figure 3 visualization designs in detail and indicate which questions they can answer.

- Choose three appropriate designs from these 9 version of figures.

| jobs | loading |
|---|---|
| Explain this topic we chose | 1 member |
| 3 version of designs | 3 members |
| Finish the ppt | 1 member |

# References

- Data:

https://www.kaggle.com/datasets/murilozangari/jobs-and-salaries-in-data-field-2024