**<Decoding the U.S. IT Job Market: Insights through Data Visualization>**

**Group 6: Jingheng Huan, Yantao Mei, Zhiyun Lu**

**Introduction**

As aspiring data scientists, many of us envision launching our careers in IT corporations. However, recent indicators point to a deceleration in growth within these firms and the broader tech sector, sparking widespread unease. Grasping the nuances of this shifting landscape is vital for anyone aiming to secure a role in IT. To demystify these trends, our project offers a comprehensive data visualization of the U.S. IT job market, providing actionable insights for informed career planning.

**Datasets**

To address the multifaceted research questions, a singular dataset is inadequate. We have employed a holistic approach, using LinkedIn job data as our cornerstone, which encompasses varied metrics such as location, employee count, and company specifics. This will be augmented by layoff figures sourced by web scraping Layoffs.fyi.

Supplemental datasets like stock market indicators further enrich our analysis. From pandas\_datareader library we can load the latest stock data from Yahoo which include all the necessary data we need.

In addition, we plan to integrate salary data from Comprehensive.io to fill potential voids in our primary LinkedIn dataset. It’s worth noting that the scope of our investigation may encounter data gaps, which we are fully equipped to address during our ongoing research.

**Methods and Sketches**

Our methodology employs Python for data wrangling and leverages the JavaScript library D3 for the visualization suite. We aim to offer five distinct interactive visualizations designed to probe the following research questions:

**Identifying Key Opinion Leaders (KOLs) on LinkedIn:**

Visualization: Nested Quotient Graph

Potential Interactions: Users can hover over nodes to reveal the KOL's name, industry, and influence score. Clicking on a node would expand it to display sub-nodes representing the KOL's professional network.

**Investigating Relationships between LinkedIn Activity and Corporate Affiliation:**

Visualization: Treemap

Potential Interactions: Hovering over a section will show tooltips detailing the department's average LinkedIn activity score. Clicking on a company sector expands it to reveal constituent departments.

**Exploring Common Job Titles in LinkedIn Profiles:**

Visualization: Word Cloud

Potential Interactions: Words will be clickable, leading to a pop-up that provides statistics like the average salary and estimated job openings for that title. Hovering can display a tooltip with the frequency of the job title.

**Comparing Tech Giants Across Multiple Metrics:**

Visualization: Radar Chart

Potential Interactions: Users can select different tech companies from a dropdown to dynamically update the radar chart. Hovering over a data point reveals detailed specs like market share or employee count.

**Mapping Layoffs in the U.S. IT Sector by State:**

Visualization: Geographical Map

Potential Interactions: Hovering over a state will display a tooltip with the number of layoffs. Clicking on a state zooms in to reveal city-level data.

**Timeline and milestones + Roles and responsibilities**

***Week 4: Final Project Proposal***

Jingheng: Finalize the roles and responsibilities of each group member. Compile the different sections into the final proposal document.

Yantao: Draft the Introduction section. Complete the Datasets section detailing sources, types of data, and their relevance.

Zhiyun: Outline the Analysis and Visualization Methods. Create sketches and references for the visualizations.

***Week 5: Data Collection, Acquisition, Data Cleaning and Preprocessing***

Jingheng: Collect the dataset from Kaggle or scrape the data. Begin preliminary data cleaning.

Yantao: Assist in data collection and take the lead in data preprocessing. Ensure data integrity and manage data storage.

Zhiyun: Start working on a codebase for analysis and visualizations. Validate the data's quality post-cleaning.

***Week 6: Data Visualizations***

Jingheng: Create two data visualizations focusing on initial data exploration and the representation of overall trends. Coordinate with team members to ensure that all visualizations are aligned in terms of style and data integrity.

Yantao: Work on creating two data visualizations, one focusing on user tasks and another based on sectoral analyses. Conduct usability tests on these visualizations to ensure they are intuitive and informative.

Zhiyun: Complete two additional data visualizations, concentrating on detailed data insights and potential predictive analytics. Fine-tune the code for these visualizations based on team feedback and usability tests.

***Week 7: Final Report and Presentation***

Jingheng: Compile the final report. Prepare for the final presentation including creating slides.

Yantao: Conduct a final review of all visualizations and data analysis. Assist in preparing the presentation.

Zhiyun: Contribute to the final report, particularly the Analysis and Visualization Methods section. Make sure the codebase is clean and well-documented.

**Appendix**

Figure 1: Lecture Slide 4A page 68

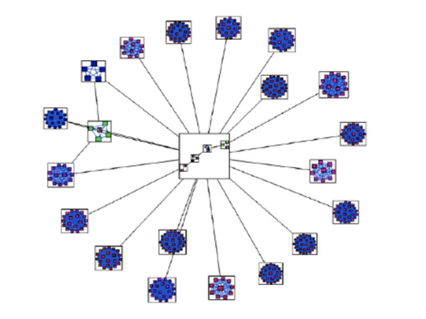
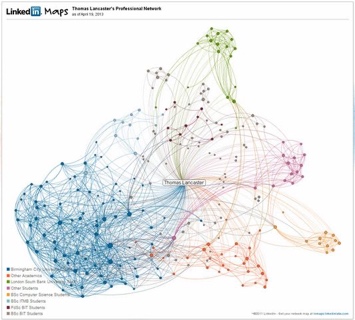
 

Figure 2: Lecture Slide 4A page 29

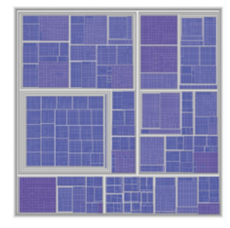
 

Figure 3: Lab 5

Figure 4: [Radar chart - Wikipedia](https://en.wikipedia.org/wiki/Radar_chart#/media/File:MER_Star_Plot.gif)

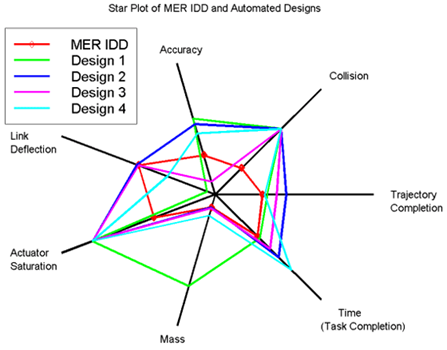
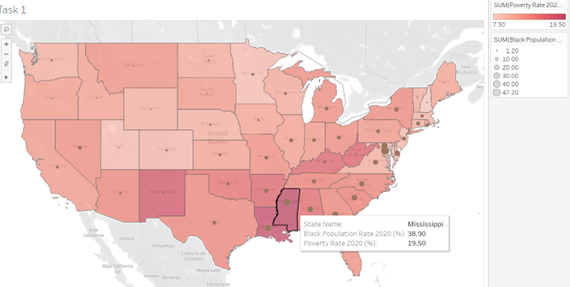


Figure 5: Lab 5

****