SWINBURNE UNIVERSITY OF TECHNOLOGY

MASTER OF DATA SCIENCE

COS80025 – Data Visualisation

Semester 2, 2024

Visualisation Critique

ANALYSING ICT JOB MARKET VISUALIZATIONS: INSIGHTS, DESIGN FLAWS, AND RECOMMENDATIONS

Due date: Sunday 8th September 2024

Submitted by: Thi Ngan Ha Do

Student ID: 103128918

Word count: 2496

Acknowledgement of Country

I respectfully acknowledge the Traditional Custodians of the land on which Swinburne University of Technology's Australian campuses are based and honour their Elders past and present.

Table of Contents

1.0 Introduction

The Information and Communication Technology (ICT) sector has long been central to global economic development but has recently faced challenges, particularly with the rise of "tech layoffs" exacerbated by the Covid-19 pandemic. To better understand the current ICT labour market and forecast future trends, this report examines three key areas: the demand for high-demand ICT jobs, the distribution of roles across technical and non-technical groups, and the essential skills employers seek in this evolving market.

The report evaluates data visualizations that depict ICT job trends, applying Edward Tufte's principles of effective visualization (Tufte 1997) and the design flaw categorization by (Lan & Liu 2024). In doing so, the report offers insights on the accessibility and clarity of these visualizations, with recommendations for improvement.

2.0 Visualisation Critique

2.1 Visualisation 1

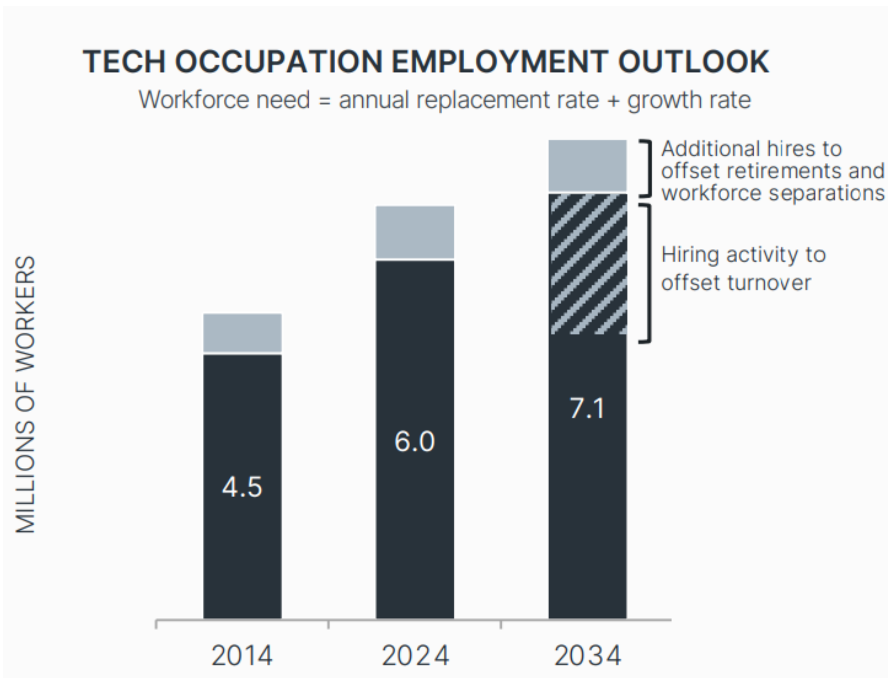

Figure 1 illustrates the tech occupation employment outlook (CompTIA 2024)

2.1.1. Context

This bar chart, titled "Tech Occupation Employment Outlook," depicts workforce demand in the technology sector at three key points: 2014, 2024, and 2034. Based on data from the U.S. Bureau of Labor Statistics and Lightcast, it shows the total number of tech workers needed and breaks this down into three components: additional hires to offset retirements, workforce separations, and hiring activities to address turnover.

The chart highlights both historical data and future projections, giving insights into areas requiring hires. The main question it addresses is: What are the tech workforce growth prospects over two decades?

Developed for the "State of the Tech Workforce 2024" report by CompTIA, the chart is intended for a broad audience, including students, career changers, and professionals, as a reference for tech hiring trends and workforce dynamics.

2.1.2 Visual Analysis

Data type

The x-axis represents three specific points in time: 2014, 2024, and 2034, while the y-axis measures the total number of tech workers in millions. For 2034, the workforce demand is calculated based on two factors: the annual replacement rate, which is projected to average around 6% annually (reflecting people leaving the workforce), and the growth rate, driven by the expansion of companies and the need to support new technologies. These two components combined offer a comprehensive projection of future workforce needs in the tech industry.

Data encoding

Attributes	Data type	Mark	Channel	Encoding
Year	Categorical	Bars	Position along x-axis	Encoded using position on the x-axis, showing time progression.
Number of workers	Quantitative		Height (y-axis), text labels	Encoded using the height of the bars, where higher bars represent a greater number of workers
Additional hires	Quantitative		Colour (light gray)	Encoded using light gray color
Turnover hires	Quantitative		Pattern (stripes)	Encoded using a striped pattern in the 2034 bar
Workforce growth	Quantitative		Colour (dark gray)	Encoded using dark gray color

Table 1. Data encoding for visualization 1

The visual progression of increasing bar heights suggests steady growth in the tech workforce over the decades. The chart utilizes a dark base color for the core workforce, a lighter gray for additional hires to offset retirements and separations, and a striped pattern to represent turnover. This color-coding is logical, as it visually distinguishes between the natural workforce growth and the workers required to replace those leaving the industry. However, instead of a traditional legend, the chart provides a written explanation on the right-hand side, adhering to a minimalist design.

2.1.3 Critique

The visualization succeeds in answering key questions about workforce growth over two decades. It provides a clear comparison of tech workforce numbers across three time points, making it useful for understanding long-term trends. However, the design could be improved to enhance the precision and clarity of the information presented.

Strengths

The bar chart is an effective choice for visualizing changes in workforce demand over time. Bar charts are suitable for comparing discrete values across time periods (Guerra-Gomez et al. 2013), with equal spacing between intervals that make it easy to see trends. The increasing bar heights clearly highlight the tech workforce growth. Additionally, the colour variations provide more context, showing the breakdown of workforce components beyond just the total numbers. The inclusion of numeric labels on each bar further enhances clarity, eliminating any ambiguity and ensuring viewers can quickly grasp the exact figures.

Limitations and recommendations

The chart's primary issue is the absence of a legend to explain the color shading and patterns used in the 2034 bar. While a written explanation is provided, a visual legend would be more intuitive and make it easier for users to quickly understand the data. Without this, the numeric label on the 2034 bar could be misinterpreted as only referring to the dark segment, causing confusion about the total workforce estimate. Additionally, the y-axis truncation, which is hidden, may exaggerate differences in workforce numbers over time.

Design flaws and recommendations

- Low readability: The small font size and weak color contrast make the workforce numbers difficult to read, especially on smaller screens or for visually impaired viewers. Increasing text size and contrast would greatly improve readability and accessibility.
- Low coherence: The 2034 bar's segmentation is inconsistent with the 2014 and 2024 bars, complicating comparisons. A uniform structure or clear explanation for the segmentation would enhance coherence.
- Oversimplification: The truncated y-axis and lack of a legend reduce data clarity. Adding a full y-axis and a legend to explain the color-coded categories would improve context and prevent misinterpretation.

2.2 Visualisation 2

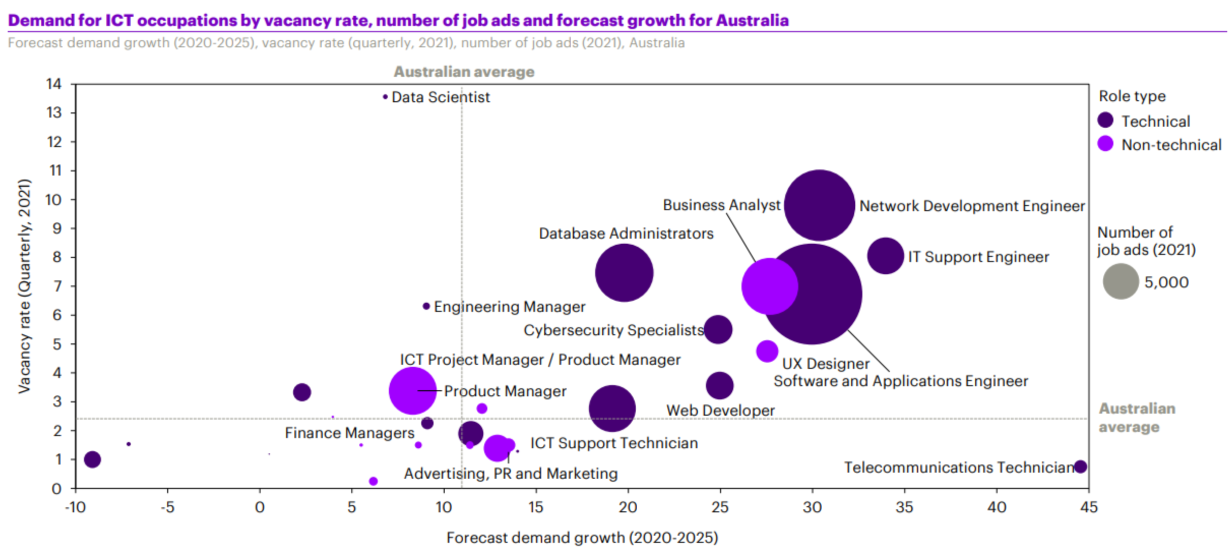

Figure 2 illustrates the demand for ICT occupations by vacancy rate, number of job ads and forecast growth for Australia (NSW Skills Board 2023)

2.2.1 Context

The bubble chart, titled "Demand for ICT occupations by vacancy rate, number of job ads, and forecast growth for Australia," compares demand, hiring difficulty, and growth prospects for technical and non-technical ICT roles. It contrasts these metrics with the national average, highlighting roles with high vacancy rates and strong forecast growth.

The chart provides an overview of current ICT job market conditions and future trends, answering key questions such as which positions are in high demand, where vacancies exist, and how technical roles compare to non-technical ones in terms of demand and hiring challenges. As part of the "ICT Industry Landscape Report 2023," the chart is designed for a broad audience with varying technical expertise.

2.2.2 Visual Analysis

This visualization incorporates multiple data dimensions requiring deeper calculation. The y-axis represents the vacancy rate, the proportion of unfilled ICT positions within a quarter (Baydur 2017), while the x-axis displays forecast demand growth, a projected percentage change in ICT job demand from 2020 to 2025. Bubble size reflects the number of job ads across various ICT roles in Australia, with roles categorized as technical or non-technical. Despite the data's complexity, the chart uses effective visual encodings. Bubble size represents job ad counts, and the legend clearly distinguishes technical from non-technical roles. The inclusion of a scale linking bubble size to 5,000 job ads further simplifies interpretation for a broad audience.

Attributes	Data type	Mark	Channel	Encoding
Vacancy rate	Quantitative	Bubbles	Position along x-axis	Encoded using position on the y-axis, where higher positions indicate higher rates
Forecast demand growth	Quantitative		Position along y-axis	Encoded using position on the x-axis, where positions further to the right indicate higher growth
Role type	Categorical (nominal)		Colour	Encoded using color, with technical roles (darker purple) and non-technical roles (lighter purple)
Number of job ads	Quantitative		Size of bubble	Encoded using the size of the bubbles, where larger bubbles represent occupations with more job ads in 2021.
Occupation title	Categorical		Text label	Represented as a text label on each bubble

Table 2. Data encoding for visualization 2

2.2.3 Critique

The visualization effectively addresses key questions about the ICT job market, particularly in identifying positions that will remain relevant over the next five years. It successfully highlights the rising presence of non-technical roles, though they remain a minority. With over 20 occupations displayed, readers can quickly assess the position of each job title based on vacancy rates and forecast demand growth. The use of lines for three specific positions emphasizes their high growth rates, making them notable for further attention.

Strengths

The chart efficiently handles a complex dataset, conveying a large amount of information through a clear and technically proficient design. The inclusion of national average lines on both axes divides the chart into four quadrants, allowing users to categorize ICT job roles into four groups: high vacancy and high growth, low vacancy and high growth, high vacancy and declining demand, and low vacancy with declining demand. This quadrant system makes the chart particularly informative and useful for comparisons. Additionally, while the x-axis starting from negative values might typically be considered a visual distortion, it is well-justified in this case as it reflects roles expected to experience negative growth.

Limitations and recommendations

However, the chart sacrifices some clarity due to the density of information it presents. Overlapping bubbles, particularly in the center, make it difficult to distinguish between roles and job counts, complicating comparisons. The use of bubble size to encode job ads is effective but lacks precision when comparing smaller bubbles, as differences may be exaggerated, leading to potential misinterpretations. A more granular size scale or clear labels for each bubble could improve clarity. Furthermore, the chart's complexity demands significant effort from viewers to interpret, which may not be ideal for a general audience, particularly those unfamiliar with data visualization techniques.

Design flaws and recommendations

- Low Readability: Overlapping data points create clutter and reduce readability. The similar shades of purple for technical and non-technical roles lack contrast, especially for visually impaired users. Higher contrast and colorblind-friendly colors would enhance clarity.
- Overcomplexity: The combination of color, bubble size, and position results in a dense visual. The wide x-axis range complicates interpretation, particularly for negative growth. Clearer labels and explanatory notes would improve simplicity.
- Ambiguity: The bubble size scale is unclear, making it difficult to interpret values, especially for smaller bubbles. Negative growth values are also not intuitive. Adding a clear legend for bubble sizes and explanations would enhance clarity.

2.3 Visualisation 3

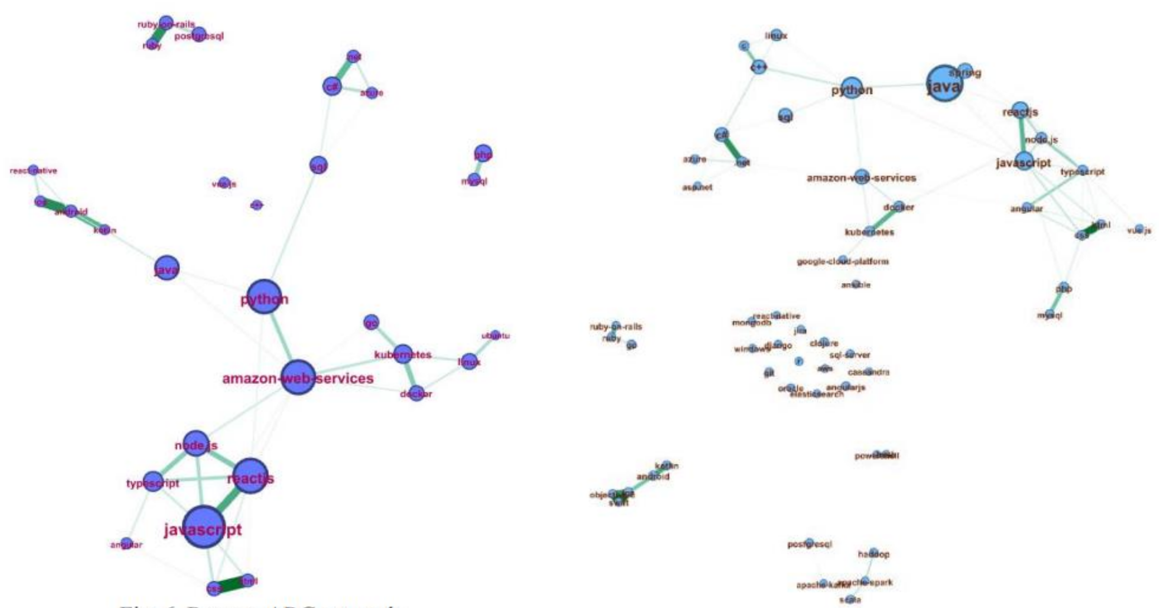

Fig. 6. Remote ARG network

Fig. 7 On-Site ARG network

Figure 3 illustrates the ARG network of in-demand technological skills by availability (Apatsidis et al. 2021)

2.3.1 Context

This Association Rule Graph (ARG) network visualizes relationships between technological skills based on job availability (remote or on-site) and demand. Originally proposed by (Yao et al. 2012), it uses a tagging system evaluated by tag, frequency, and confidence metrics to illustrate interconnected job posting requirements.

The visualization highlights the demand for technologies in remote and on-site jobs, addressing key questions from (Apatsidis et al.) about the most requested technologies and their interconnections. Presented at the 2021 47th Euromicro Conference on Software Engineering and Advanced Applications (SEAA), it is primarily intended for researchers, academics, and professionals in advanced technical fields.

2.3.2 Visual Analysis

This visualization incorporates a combination of categorical and numerical data. The categorical data consists of technologies derived from a thorough analysis of on-site and remote job postings, highlighting the top ten most in-demand technologies in the ICT job market. Numerically, the frequency of each technology is calculated as the ratio of the number of times it appears in job postings to the total number of postings. Additionally, the confidence metric, which is essential for ARG network graphs, is calculated by dividing the support metric between two tags by the frequency of the first tag.

The choice to represent the frequency of technologies with varying node sizes is intuitive, namely bubble charts or tree maps, where larger nodes represent more frequent occurrences. The ARG network format effectively visualizes the interconnections between technologies and their significance in the job market. A key takeaway from this network is the centrality of specific technologies, indicating which skills are essential for job seekers to stay competitive.

Attributes	Data type	Mark	Channel	Encoding
Technology names	Categorical	Nodes and lines	Text label	Represented as a text label on each node
Connection between technologies	Quantitative		Connecting lines	Encoded using lines between the nodes
Frequency of technology	Quantitative		Size of nodes	Encoded using nodes, where larger circles indicate a higher frequency of occurrence in job postings
Strength of connection	Quantitative		Thickness of lines	Encoded by the thickness of lines, thicker lines indicate stronger co-occurrences between the technologies
Availability (on-site or remote)	Categorical (nominal)		Colour	Encoded using color, with on-site roles (teal blue) and remote roles (light purple)

Table 3. Data encoding for visualization 3

The side-by-side comparison of on-site and remote job postings highlights demand differences. However, the similar cold color palettes for remote and on-site postings are hard to distinguish. Additionally, some connecting edges between nodes are faint, making relationships between technologies unclear.

2.3.3 Critique

While the visualisation provides useful insights into high-demand technologies and their interconnections, it falls short in a few areas. The visualization emphasizes individual technologies rather than broader groups of skills such as cloud services or data management.

In an ever-evolving tech industry, focusing only on "trendy" tools risks overlooking more foundational skills that remain valuable over time. Additionally, non-technical skills like leadership, teamwork, and communication, which are crucial for career growth even in tech-centric fields, are not represented.

Strengths

The chart is rooted in sound academic principles and mathematical foundations, making it a reliable reflection of the job market. It accurately highlights the importance of central skills like Python and AWS, offering a strategic advantage to job seekers who possess these skills due to their wide applicability across different roles. The side-by-side comparison of on-site versus remote jobs is also insightful, clearly showing the concentration of front-end development in remote roles and a more diverse skill set required for on-site roles.

Limitations and recommendations

Despite its strong foundation, the chart's complexity requires a certain level of familiarity with ARG networks, which limits accessibility for the general audience. Using more common chart types like a tree map or Gantt view could communicate similar insights in a more accessible manner. Additionally, some nodes are disconnected from the main network, suggesting weaker relationships between certain technologies. Grouping related technologies and labeling these clusters could enhance understanding and provide a clearer sense of the skills landscape.

Design flaws and recommendations

- Low Readability: Overlapping data marks in denser areas of the chart make it difficult to distinguish between technologies. Increasing spacing between nodes, applying size constraints, and enlarging text would improve readability.
- Low Contrast: The color scheme used for on-site and remote jobs is too similar, which creates confusion, particularly for colorblind users. Adopting a color-blind-friendly palette and using more distinct colors or patterns for different categories could improve differentiation.
- Oversimplification: The chart lacks a clear explanation of how node size, edge thickness, and color interact with each other. Including a detailed legend to clarify these relationships would make the visualization easier to interpret.

3.0 Conclusion

In conclusion, this report analysed three visualizations related to ICT job opportunities, focusing on their effectiveness in conveying data, strengths, and limitations. From these analyses, it is evident that while visualizations can convey complex data, there must be a balance clarity and complexity to be truly effective. In future visualizations, incorporating clear legends, improving contrast, and ensuring better accessibility for diverse audiences will be key takeaways for more effective design.

4.0 Reference List

Apatsidis, I, Georgiou, K, Mittas, N & Angelis, L 'A study of remote and on-site ICT labor market demand using job offers from stack overflow,' IEEE, 252-259.

Baydur, I 2017, 'Worker selection, hiring, and vacancies', *American Economic Journal: Macroeconomics*, vol. 9, no. 1, pp. 88-127.

Guerra-Gomez, J, Pack, ML, Plaisant, C & Shneiderman, B 2013, 'Visualizing change over time using dynamic hierarchies: TreeVersity2 and the StemView', *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 12, pp. 2566-2575.

Lan, X & Liu, Y 2024, '" I Came Across a Junk": Understanding Design Flaws of Data Visualization from the Public's Perspective', *arXiv preprint arXiv:2407.11497*,

Tufte, ER 1997, Graphic Press.

Yao, J, Cui, B, Cong, G & Huang, Y 2012, 'Evolutionary taxonomy construction from dynamic tag space', *World Wide Web*, vol. 15, pp. 581-602.