# Exploring Demographics and Relationships in Engagement, Ohio (VAST 2022)

Ankit Gurudutt Basrur
*Arizona State University*
Tempe, USA
abasrur@asu.edu

## I. INTRODUCTION

The town of Ohio is preparing for substantial growth, prompting an urban planning initiative to understand its current state and potential areas for expansion. About 1000 representative residents voluntarily contributed data through the city's urban planning app, detailing their visited places, social interactions, spending habits, and more. Our main project goal is to gain crucial insights into the town's demographics, social connections, and business landscape.

To achieve this, visual analytics tools were developed which aimed at comprehending the city's neighborhoods and business scenario. This involves uncovering correlations among various factors like age, interests, and social interactions. Additionally, to identify popular recreational spots and assess the prevalence of businesses in different areas of the city this project could be used.

## II. SOLUTION

This project, uses six visualizations techniques to explore the demographics and relationships of the town of Ohio.

*a. Line Chart:*

This chart aims to display the average joviality and age across ten interest groups. Users can interact by clicking on specific interest group lines to explore them further through the Parallel Coordinates Plot. Using this visualization, the end-user can understand the happiness index (joviality) in the population and its relationship with age for an interest group.

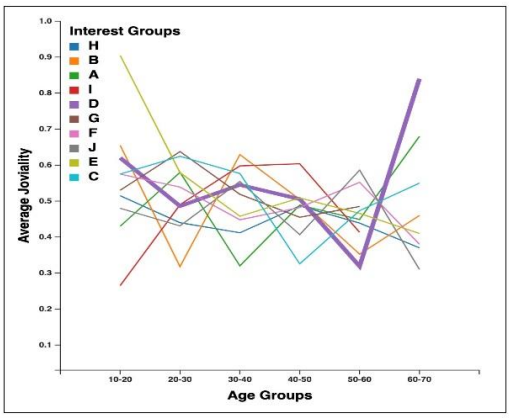


*Figure 1: Line Chart*

*b. Parallel Coordinates Plot:*

This graph showcases distinct trends within a chosen interest group, presenting useful information about attributes like Joviality, Age, Number of kids, Education Level, and Household size across various axes. The interaction scope was further enhanced by implementing hover effects that highlight a specific line while blurring the others. Furthermore, this graph includes axis brushing to select multiple lines throughout the chart. Overall, this chart aids in comprehending the town's demographics more effectively.

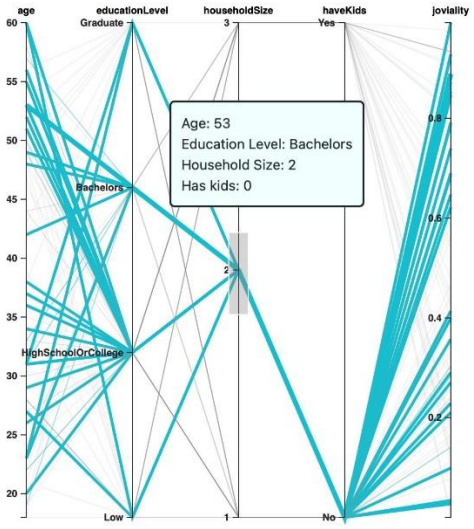


*Figure 2: Parallel Coordinates Chart*

*c. Network Graph:*

This visualization is being used to find social interactions between participants from various interest groups through this graph [1] [2]. It specifically displays interactions between two individuals from different interest groups. Our interface includes a control panel enabling the selection of these two groups, along with a calendar to choose specific date ranges. This setup allows us to track trends, observing how connections between these groups evolve over time and pinpoint periods of heightened interaction, such as during festivals. This chart serves as a tool to analyze the shifting trends in social interactions within the town.

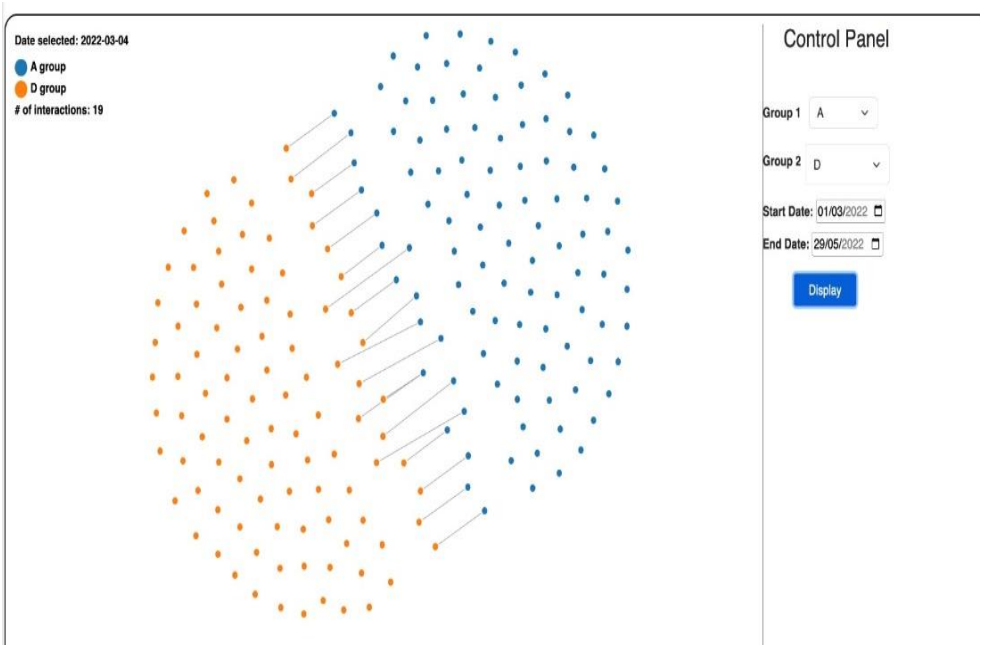


*Figure 3: Node-link Graph*

d. *Scatter Map:*

This visualization maps popular recreational spots like Restaurants and Pubs across Ohio, utilizing dataset information. Varied circle sizes represent the popularity of each place, while color luminance indicates their average cost. Clicking on any bubble allows viewing a separate heat map or flowmap, illustrating the visitor's starting locations for that particular Restaurant or Pub. This map serves as a valuable tool to identify the predominant business base within the town.

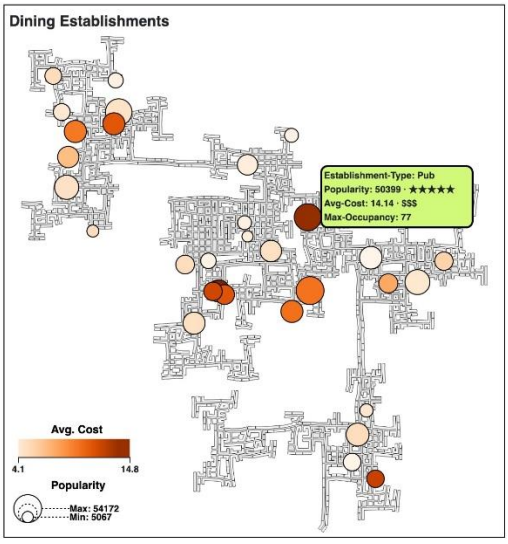


*Figure 4: Scatter Map*

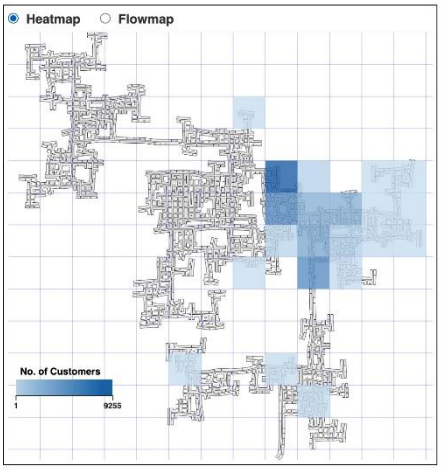


*Figure 5: Heat map*

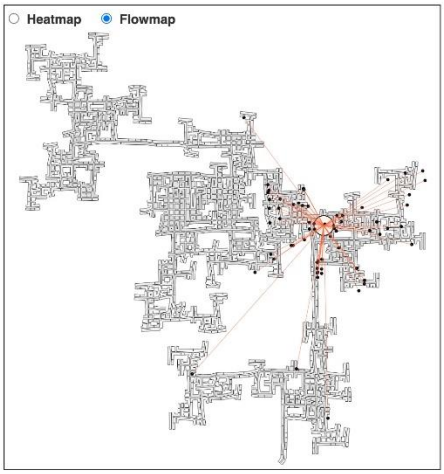


*Figure 5: Flow map*

*e. Radial Bubble Chart:*

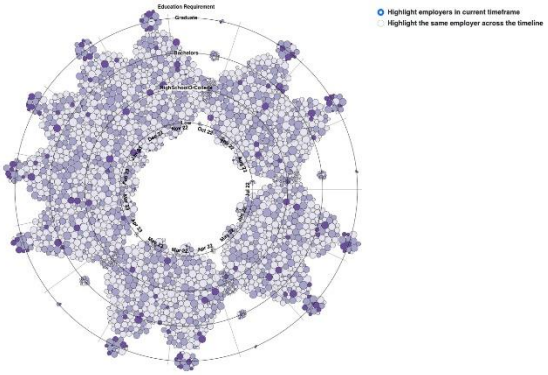


*Figure 6: Innovative radial bubble chart*

This graph showcases our innovative approach to visualizing employees and their educational qualifications across a range of months. The bubble size corresponds to the number of employees at each employer. Users have the option to view either the current employer for a specific timeframe or visualize all employers within the current timeframe.

## III. RESULTS

The final product is a tool that allows the end-user to analyse the demographics and relationships of the city of Ohio. It also enables them to use our project to explore and study interactive visualizations [3].

## IV. CONTRIBUTIONS

I have been involved in the project right from the brainstorming step to the final day poster presentation activity.

My specific contribution in this project involves the following:

- Researched different project topics and comprehend the dataset provided and conduct feasibility study on them
- Explore different visualization techniques and implementation methodology to solve the chosen problem statement
- Collaborated in preparing the project proposal, presentation, bi-weekly reports, and poster write-up
- Pre-processed the Participants.csv and SocialNetworks.csv files to remove redundant or missing values and store the processed records in MySQL database
- Implemented Node-link diagram to reveal inter-interest group trends. Designed and developed a control panel to switch between different interest groups and change the calendar
- Helped teammates in resolving their bugs

This project's success owes much to our collaborative approach in distributing the workload effectively. I gained significant insights, particularly in handling extensive CSV files with over 7 million rows. I implemented a strategy to display social interactions day by day instead of all at once, allowing users to observe interaction patterns between interest groups gradually over time. This approach required learning data pre-processing in Python and employing MySQL to push processed data into the database.

Moreover, I ventured into implementing the Network Graph using D3.js force simulation. I explored a new methodology of separating two groups of points based on centroids, which added a fresh dimension to my learning experience.

## REFERENCES

[1] U. Brandis, P. Kenis and J. Raab, "Through Network Visualization," vol. 2, no. 1, 2006.

[2] P. Chakraborty and A. Nath, "Application of Parallel Co-ordinate System to visualize Multivariate dataset," *International Journal of Innovative Research in Advanced Engineering,* vol. 1, no. 11, November 2014.

[3] J. Hullman and N. Diakopoulos, "Visualization Rhetoric: Framing Effects in Narrative Visualization," *IEEE,* vol. 17, no. 12, pp. 2231-2240, 2011.

[4] F. Levesque and T. Hurtut, "MuzLink: Connected beeswarm timelines for visual analysis of musical adaptations and artist relationships," *Information Visualization,* vol. 20, no. 2-3, pp. 170-191, 2021.

[5] E. Hehman and S. Y. Xie, "Doing Better Data Visualization. Advances in Methods and Practices in Psychological Science," *Sage Journals,* 2021.

[6] M. Waldner, A. Diehl, D. Gračanin, R. Splechtna, C. Delrieux and K. Matković, "A comparison of radial and linear charts for visualizing daily patterns," *IEEE transactions on visualization and computer,* vol. 26, no. 1, pp. 1033-1042, 2019.

[7] H. Y. Wu, K. Klein and D. Yan, "Network Visualization and Graph Data Management," *IEEE Computer Graphics and Applications,* vol. 43, no. 3, pp. 10-11, 2023.