Overview and Motivation

The University of Utah's Marković Research Group is trying to convince transportation organizations to purchase services for managing trajectory data. Part of the service they're trying to propose involves creating a tool that people in operations research and transportation planning can use to better understand traffic patterns. They would like to have a prototype of that tool and we would like to make it for them. One of the most important factors in transportation planning and maintenance is having a good understanding what traffic occurs between traffic analysis zones (TAZ.) We're creating a prototype of the tool that would be used to visualize TAZ flow by using counties in the place of TAZ.

Related Work

The idea to visualize trips using a Chord diagram came from Mike Bostock who used one for visualizing Uber Rides by Neighborhood (https://bost.ocks.org/mike/uberdata/.) Seth was shown this image as an example visualization in his quantitative methods of transportation operations class taught by Nikola Marković. The idea for the sortable bar charts came from Homework 4 of our data visualization course, and the idea for the interactive map came from Homework 3.

Questions

We're not exactly trying to answer questions. We're more accurately trying to create a prototype for a tool that different departments of transportation can use to visualize transit between analysis zones. Because we began with such a clear objective, our questions haven't changed over the course of the project. We still have some questions for what an extension would be of our design. We've made a clear visualization for trips between regions, but we still must answer the question of how to visualize link demand contribution for each origin destination pair. To do this would we select a link and generate a new chord diagram and bar charts for that link specifically? Would we select two regions and have links thickness shown to indicate travel demand? These questions of course extend beyond the scope of the project for this course.

Data

Trajectory Data

INRIX is a company that provides trajectory data. Their data contain approximate longitude-latitude coordinates with timestamps for different devices as well as classifications for the device. They use multiple classifications, including the category of device ownership, which is split into corporate, personal, and governmental. The University of Utah's Marković Research Group, which includes one of our team members, Seth, purchased data from this company for the month of September 2018, and we'll be using that data for our project. Our data came separated into two sets, trip overviews (origin location, destination location, start time, vehicle weight class) with unique keys indicating the trip ID. The second set contained the locations and time stamps for each GPS probe classified by their appropriate trip ID.

Geographic Data

As a part of our visualization we need to visualize the map of state Utah and its counties. We have prepared the required data. The details are given in the following section.

Data Processing

Trajectory Data

The processing of this data took roughly six months for the Marković research group to perform. It included looking for trips that INRIX had incorrectly classified and mapping approximate GPS probe locations to the appropriate roadway. The data cleanup and preparation were extensive and would take too long to detail in a proposal. The data we have from the research group consists of a simple list of trips with their starting counties, ending counties, and departure times. We will just need to find the counts for each origin/destination combination during departure time intervals. This is a simple process. We just create a list of all possible origin-destination combination (841 total) and count them for our time interval or intervals.

Geographic Data

Our geographic data includes the map of state Utah and its counties. We have followed the following steps to prepare it:

- We downloaded the shape file of state Utah from this.
- We generated the GeoJSON file from this.

The map data its ready now, and we have also uploaded the data on GitHub repository.

Exploratory Data Analysis and Design Evolution

Sketches of all our designs can be found in the appendix. We've decided to use three graphics, each having their own purpose. First, we'll have a Circos plot to give a general overview of all trip counts within the state because it's very good at highlighting the most significant areas. Another major advantage of using a Circos plot is its memorability factor because most people are not familiar with Circos plots. Circos plots have a major flaw though in that they make it very difficult to see important details. Especially when there are multiple origins and destinations, the bands will overlap making seeing fine details practically impossible.

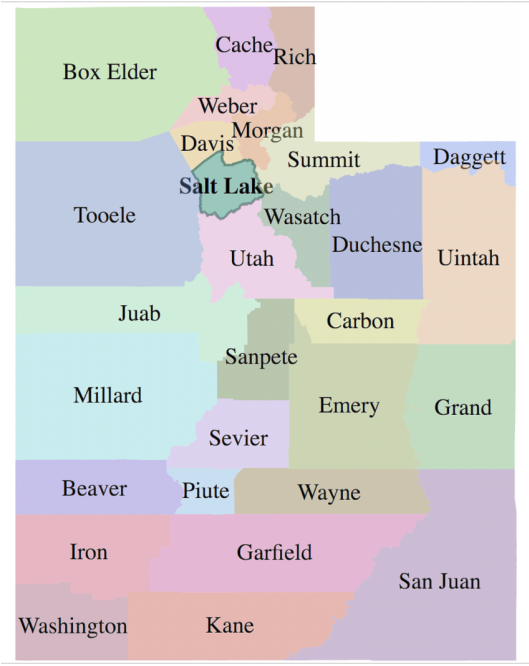
To supplement this difficulty, we've used bar charts that vary based on user input. When the user hovers their mouse over an element of the Circos plot, it generates bar charts for trips entering the selected area and exiting the area; one selection of bars for trips entering, and one adjacent selection of bars for trips exiting. Bar charts make it easy for the user to compare differences because of all potential indicators, length is perceived most accurately by humans.

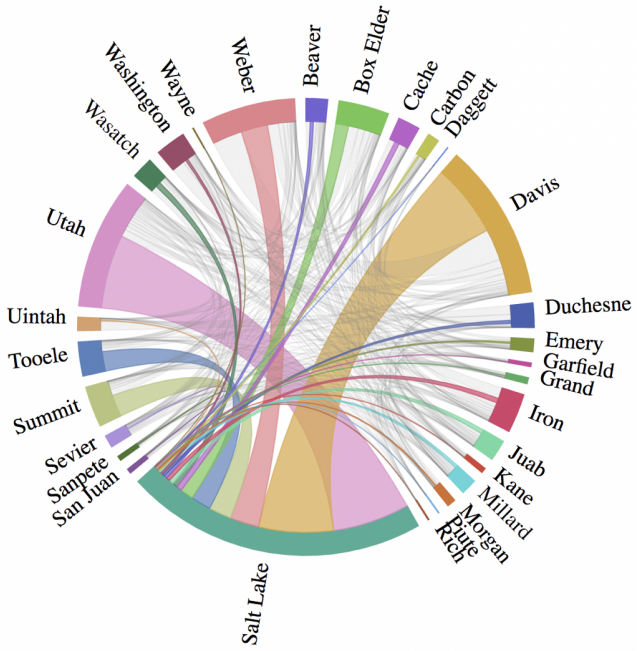
To tie these two elements together, we'll have a map. The map will operate similarly to the Circos plot where hovering over a region will update the bar charts. We'll also tie the Circos plot and the map so when the user hovers over one, the other one shows the same as though it had been selected. For more details see the Sketch of Chosen Design in Appendix I.

Implementation

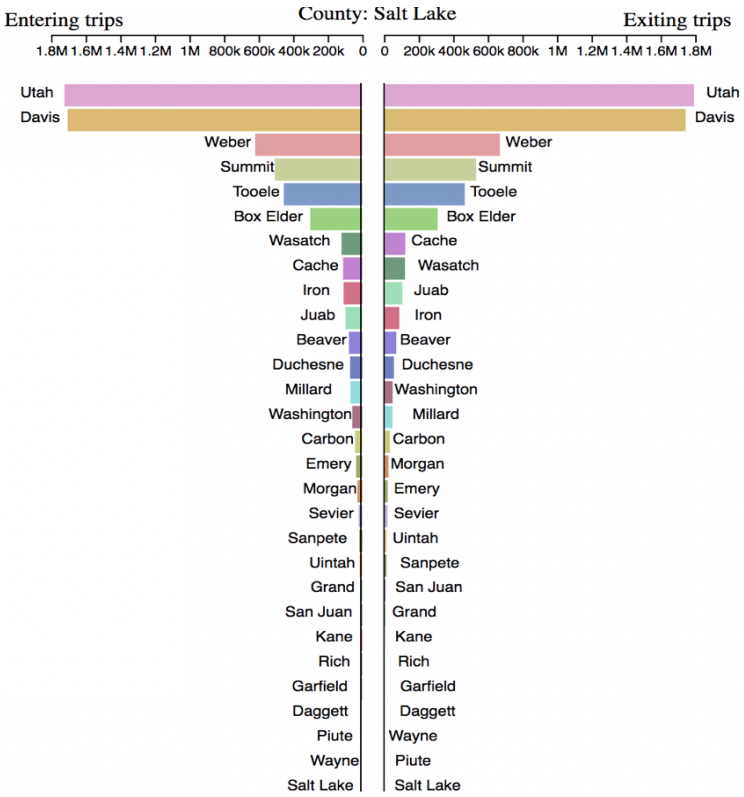
Our first implementation can be found in Appendix II. And the final implementation is given in Appendix III. We've used three elements: a chord diagram, a map, and a two bar charts. The chord diagram provides the general overview of all transit between two regions. The bar chart adds to this by separating inflow and outflow for each region and shows intensity of inflow and outflow by length of the bar making

visualizing the amount of flow much more strait forward than trying to understand the overlapping lines. We've also included a map in our design so that the viewer can easily identify regions. The visuals are color coded so that the user can easily connect each of the three components to each other. The initial state of these three visualized elements is shown for Salt Lake county. We've highlighted the selected county on both map and Circos plot:





The bar charts show the entering and exiting trips for the selected county. Each of the bar charts are sorted separately based on the number of entering trips and exiting trips. Hovering on the bar charts shows the exact number of trips. Since all the three visualizations are interactive, the user can explore and make updates by clicking on the map, Circos plot, and also bar charts.



Evaluation

As we mentioned earlier, we had a clear goal for this project. Based on our visualization, we made a clear visualization for entering and exiting trips of different regions for the state of Utah. The map identifies the area for the user. The chord diagram gives a general overview of the trips between areas. And more importantly, the user can find the exact information on the bar charts.

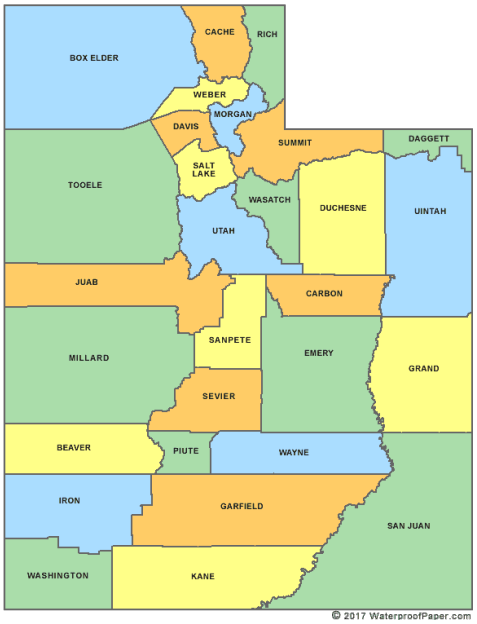
The user can much learn about the data through the interactivity. The chord diagram shows the most significant area immediately. The bar charts are sorted separately based on the number entering and exiting trips. Therefore, the user can recognize the counties with the most travels from/towards the selected county.

This project was restricted to visualizing the trips between regions for one time slot. We have the potential to extend this project for more time slots. In that case, there are more potentials to learn patterns from the data.

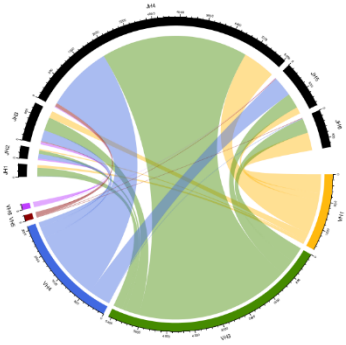
Appendix I – Designs

Sketch of Chosen Design

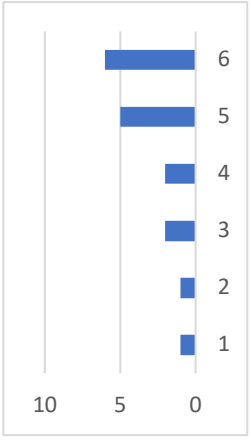
Map of all counties in Utah, hovering over specific county highlights in Circos Plot and sets bar chart to location



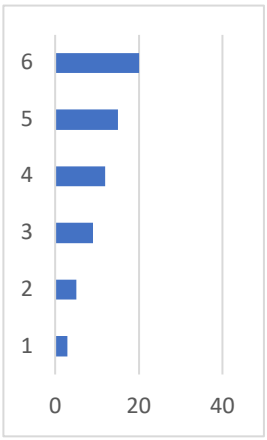
Circos Plot of All Trips in Utah Split by County – Hovering over out circle selects location and highlights place on map map and sets bar chart to location



Trip Count ***Entering*** Selected Location

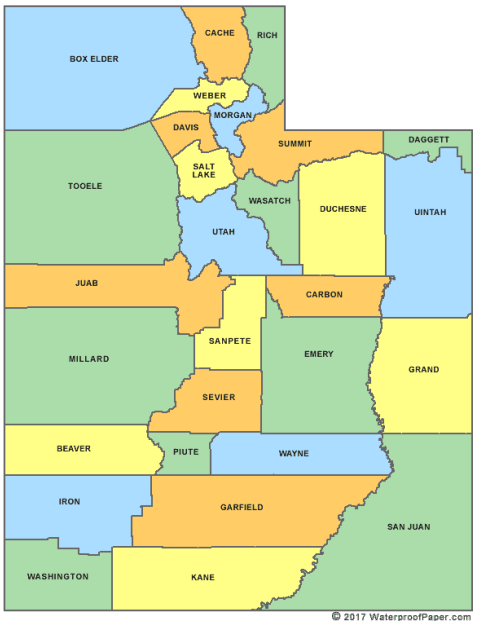


Trip Count ***Leaving*** Selected Location

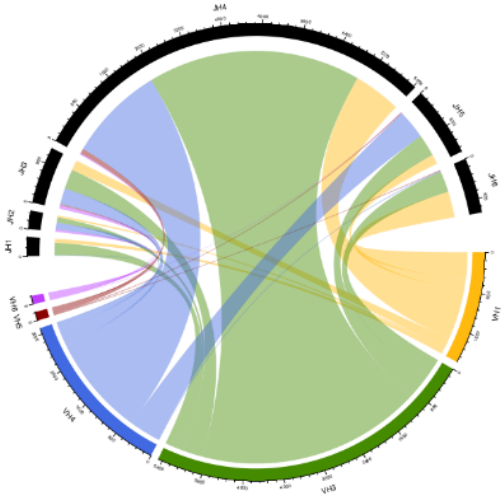


Design 1

Map of all counties in Utah, hovering over specific county highlights in Circos Plot

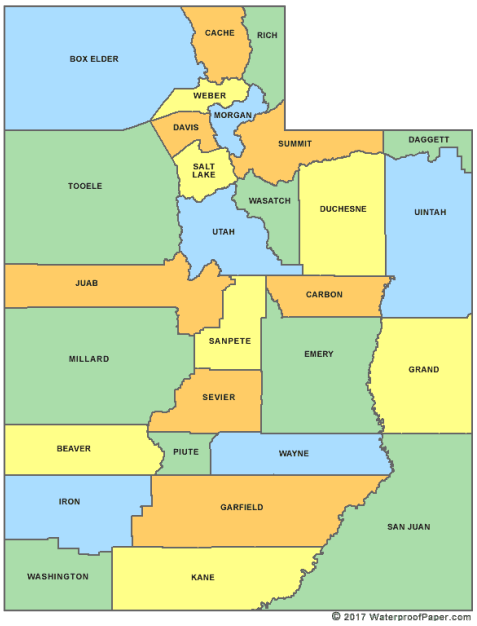


Circos Plot of All Trips in Utah Split by County – Hovering over out circle selects location and highlights place on map

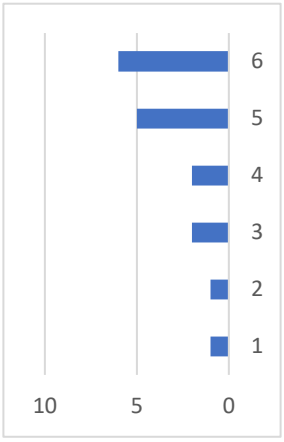


Design 2

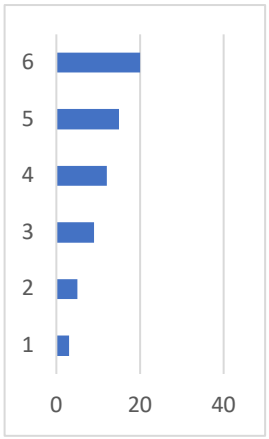
Map of all counties in Utah, hovering over specific county sets bar chart to location



Trip Count ***Entering*** Selected Location

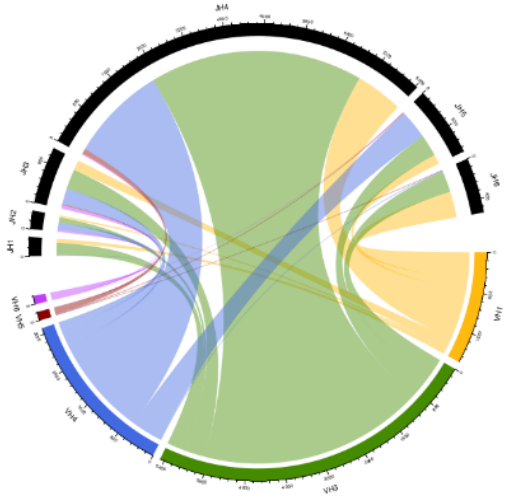


Trip Count ***Leaving*** Selected Location



Design 3

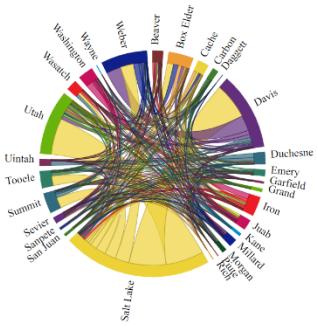
Circos Plot of All Trips in Utah Split by County – Hovering over out circle puts boarders of the square of the heatmap for specific location

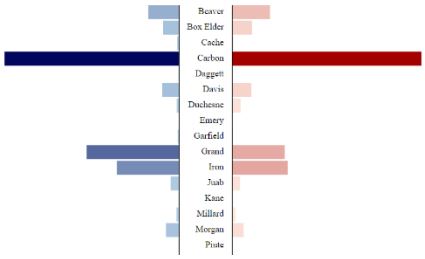


Heatmap of all trips in Utah split by County – hovering over a row or column with the county name updates the Circos plot.

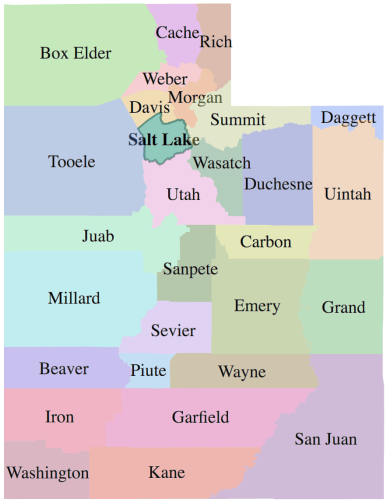
	Utah	SL	Weber
Utah	50	10	5
SL	9	50	6
Weber	20	30	30

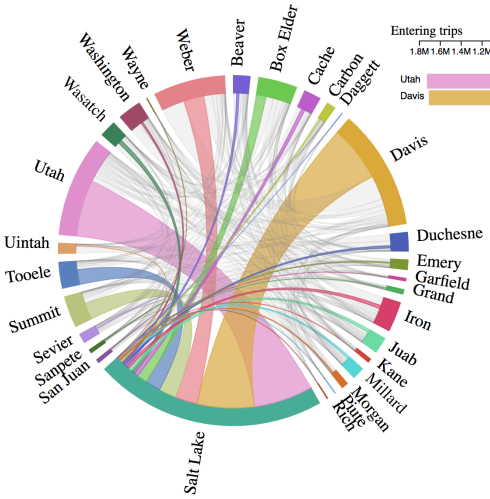
Appendix II – Initial Implementation

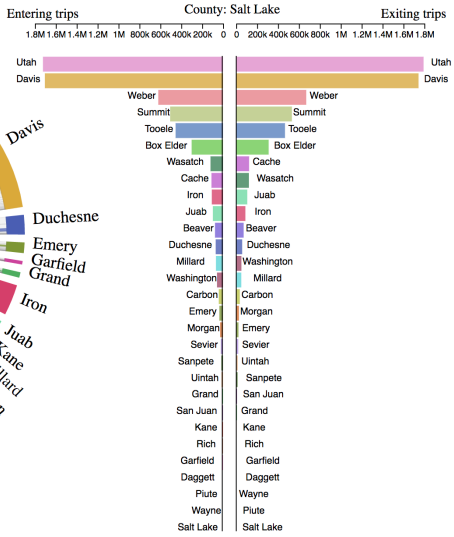




Appendix III – Final Implementation







Appendix IV – Ciros Plot and Map Sources

https://geology.com/county-map/utah.shtml

https://www.biostars.org/p/298547/

http://census.ire.org/data/bulkdata.html

https://mapshaper.org