Gabrielius Kudirka

CS 6630: Visualization for Data Science

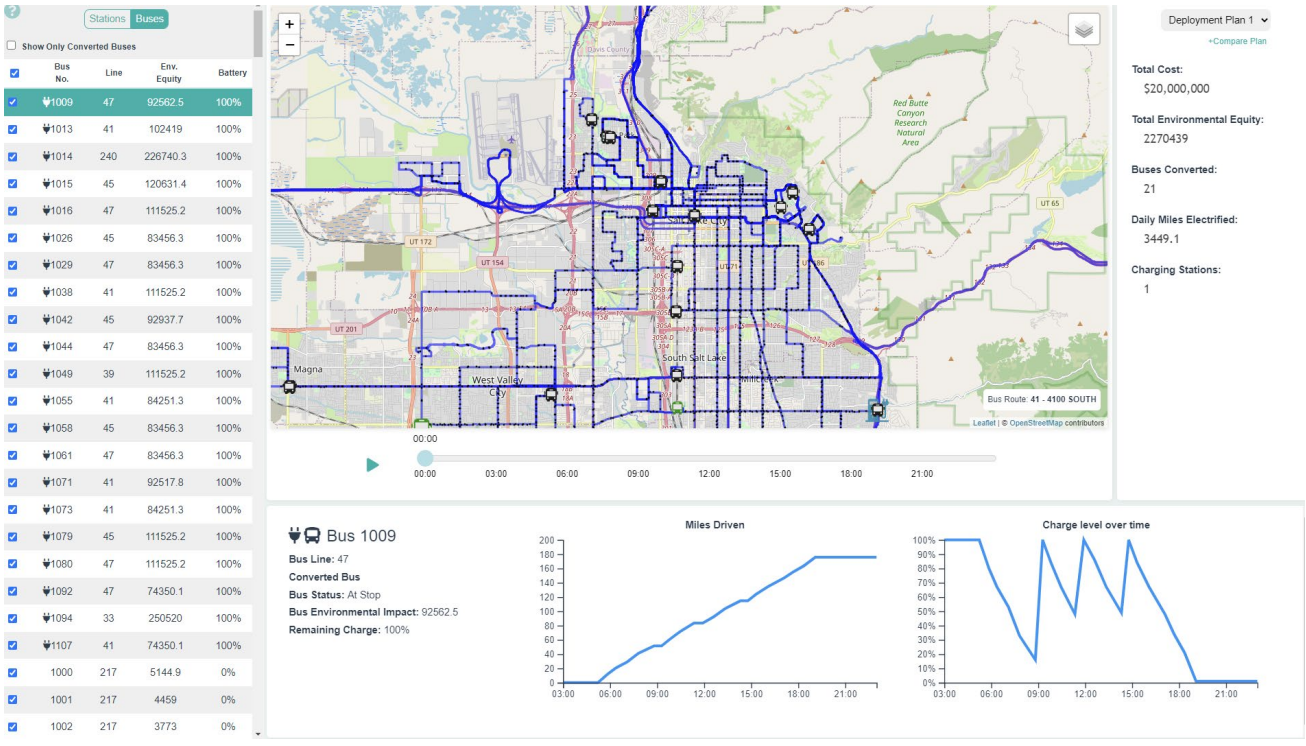
Project Process Book

Overview & Motivation

This project is an extension of my research on the Seed2Soil project headed by Dr. Bei Wang Phillips and Dr. Cathy Liu. The aim of this project is to provide a visualization tool for performance analysis of service fleet vehicles at the University of Utah. This tool will visualize telematic and operational data from University of Utah vehicles to allow for analysis of the performance of fleet operation. I hope that that this tool will allow for the development of strategic solutions for more cost effective and energy efficient fleet service management.

Related Work

In the fall of 2021 I took Dr. Bei Wang Phillips' Advanced Data Visualization class where I built a visualization tool for the deployment of electric busses in Salt Lake City. For this visualization I visualized the locations of busses under this plan according to their routes and the charging stations associated with them. I also added an extra panel to visualize data about selected busses and charging stations. After I completed that class, I continued to work on the visualization with Dr. Bei Wang Phillips and Dr. Cathy Liu for a paper on the efficient deployment of electric busses. This visualization was a starting point for this project and I am using its overall layout and application structure for this project. This is a screenshot of that visualization:



Questions

For the Seed2Soil project we are working closely with the administration to tailor our research towards their needs. The primary questions that they hope to answer with our project and visualization are:

How can we optimize the utilization of vehicles between departments to consolidate our fleet?

Can we reduce the number of trips taken by vehicles by consolidating similar routes?

Can we reduce vehicle idle time?

Can we increase driving efficiency by analyzing driver behavior?

By answering these questions, we hope that we will be able to reduce costs and emissions from fleet vehicles at the University of Utah.

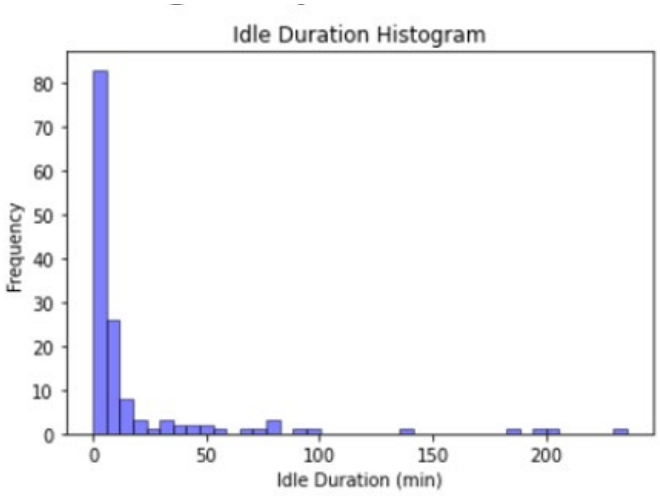
Data

The data that I have been using for the project so far has been from a platform called GeoTab which the University of Utah currently uses to house data on their fleet. On this platform are two main pieces of data that we are analyzing, descriptions of the fleet vehicles, and descriptions of the trips that these vehicles make. The data on the fleet vehicles themselves is in csv format and includes the vehicle's ID, make, model, year, class, department, and VIN. The trips data is also in csv format where each row corresponds to a stop made on the trip. Each stop includes the vehicle that made the stop, the department of the vehicle, the time the stop started, the time the stop ends, the duration of the stop, the duration of idle time during the stop, the distance traveled since the last stop, and the address of the stop location.

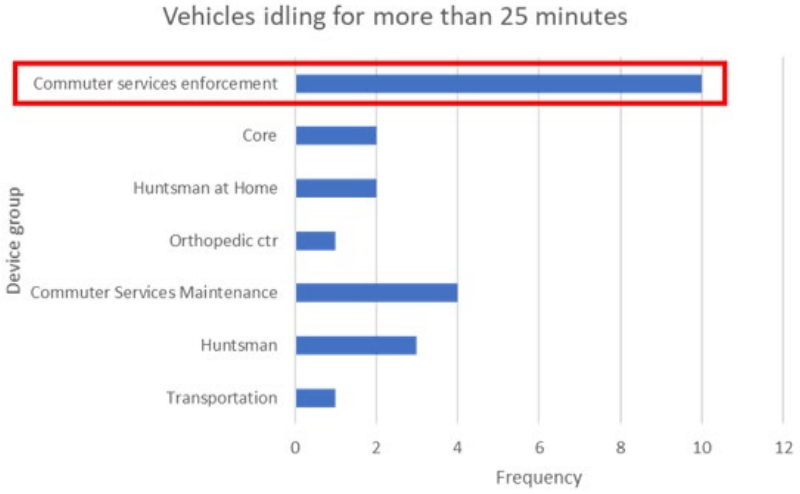
For the data processing I created three 4 distinct objects: departments, vehicles, trips, and stops. Each object contains the information related to that type of object. Each department is associated with the vehicles in that department, each vehicle is associated with the trips that vehicle made and the department it belongs to, each trip is associated with the individual stops on the trip, and the stops are associated with the vehicle and the trip that it was a part of. The greatest challenge of the data processing stage was that each stop only included the address, not the actual coordinates of the stop which are necessary for plotting on a map. To overcome this I used the ArcGIS API to go through each address and convert it into readable format. I then converted this coordinate into a GeoJson point feature to be plotted. I performed all of the initial data processing using python and then saved the arrays of objects to JSON files which I then read in my application.

Exploratory Data Analysis:

Before completing the data extraction I described above for the application, I created some preliminary visualizations to explore what we want to include in the overall visualization. Firstly, I looked into the idling time since that was a point of interest for the University of Utah and came up with the following visualizations:

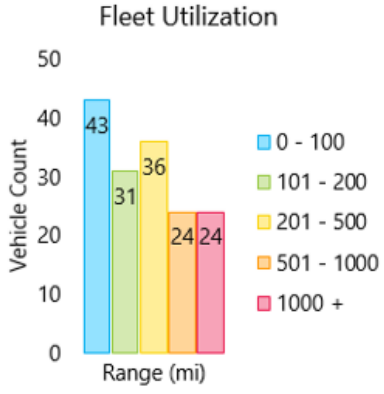


This histogram displays the frequency of idle durations across stops made. From this we can see that most of the time, vehicles are not idling although there are some major outliers. From this I figured that it would be important to include more fine grain analysis of these outlying idle times.

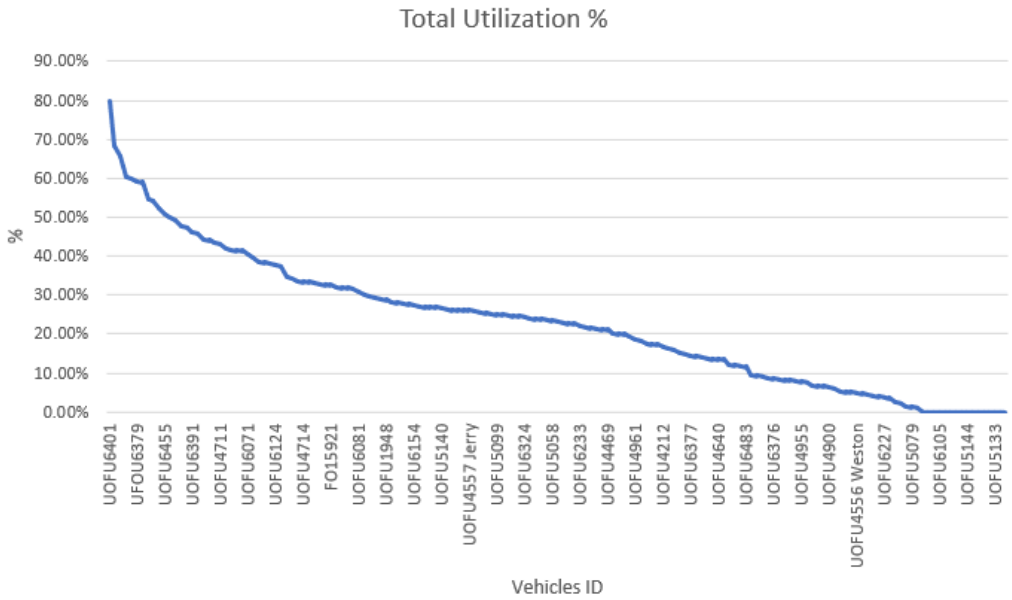


This visualization displays the frequency of long idle times across departments. In order to help the university discover where they can improve their fleet efficiency, we want to include analysis of idle time by department in the application.

After doing some preliminary analysis of idle times, I also began exploring how we can visualize fleet utilization in the application.



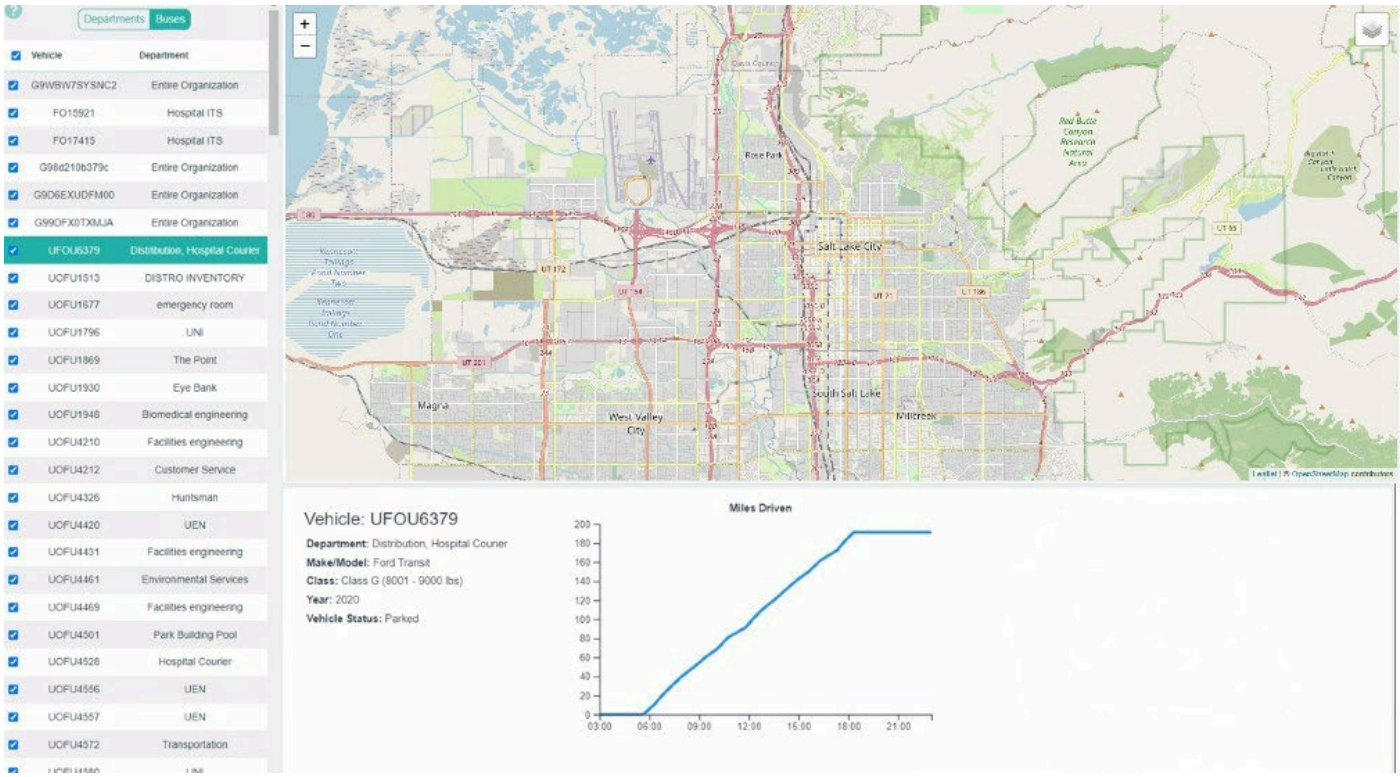
This visualization shows the frequency of high and low usage among university vehicles. From this we can see that the most vehicles are not used very much with 43 traveling between 0 and 100 miles in a one-month span. From I figured it would be important to provide comparisons of usage across vehicles in the application.



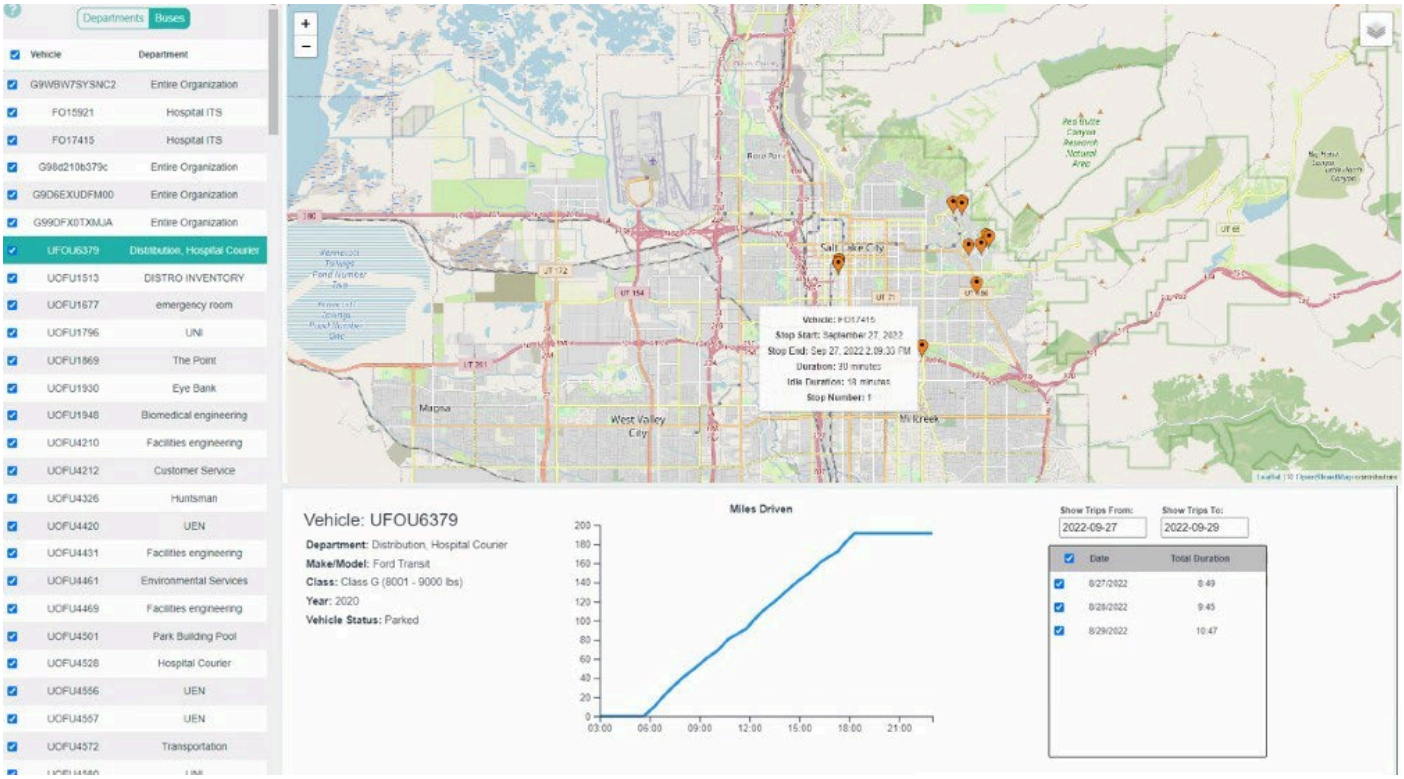
This visualization displays the usage of individual university vehicles, from most utilized to least utilized. I calculated utilization by the number of days with a trip divided by the total number of days in a one month span. I believe that visualizing a utilization metric like this will be important in the final application.

Design Evolution & Implementation:

The starting point of the design for this visualization was the previous project for which I displayed a screenshot in the related work section. In that design, the right panel displays the different deployment plans for electric busses. This was pretty specific to that project so I first decided to cut that out and expand the map section. Next, after processing the data I added all of the vehicles to the list on the left side and made them selectable such that on selecting, information about the vehicle is displayed on the bottom panel. Here is how the application looks in this stage. As of this point the visualization chart in the bottom panel is just a placeholder and is not showing any useful information. This stage did not affect any actual visualization in the application, however this addition increases the interactivity by allowing users to select vehicles they want more detail on.



In the next stage I expanded the visualizations for vehicles. A main focus point for the people working at the U was to be able to analyze easily analyze where vehicles are moving, so I started off with geo-spatial visualizations. First, I wanted to display where the vehicles were on the map by placing markers at each stop location for each of a vehicle's trips. This resulted in a significant amount of clutter on the map so I also included a time varying component so that you can visualize stops that a vehicle made within a specific period of time. To vary the window of time you wish to view, you can update the from and to date select boxes in the bottom panel. I also wanted the user to be able to view specific trips that a vehicle has made, so on the bottom panel, I included a list of all of the trips made by the vehicle within the specified time range. When you uncheck a trip in this list, the stops for that trip will not be displayed on the map. When you check the checkbox on the top of the list, all of the trips for that vehicle will once again be checked and all of the stops for all of the trips will be displayed on the map. Also, to increase the interactivity of the visualization, I added a tooltip for each of the markers which displays the vehicle associated with the stop, the start and stop time of the stop, the idle duration of the stop, and the stop number on the trip. This is a view of the application in this stage:

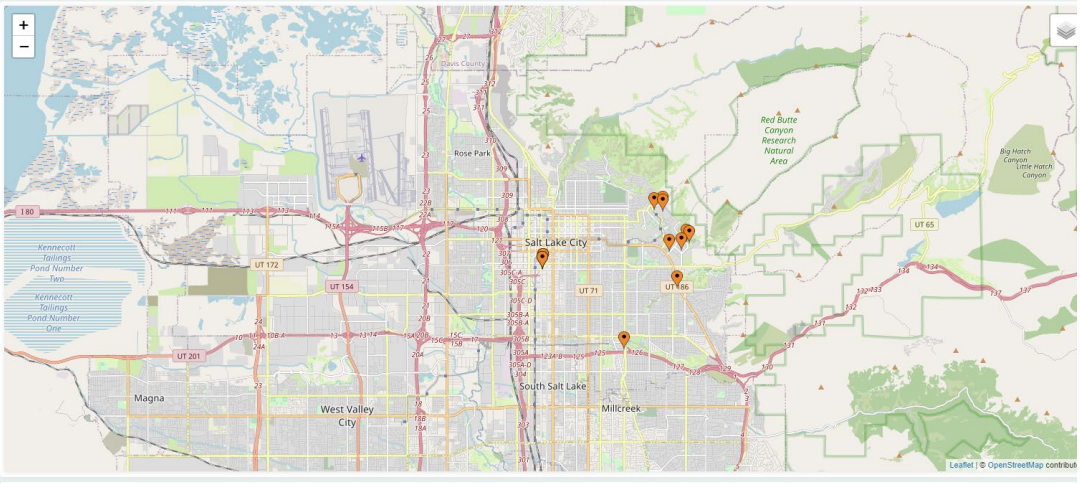


In this next stage I wanted to add two important pieces to the visualization, a view for visualizing department wide statistics, and changing the placeholder chart to visualize real information. For the visualization for a single vehicle, I decided to visualize its driving time across the selected dates. To do this I added a bar chart where the dates in the selected range are on the x axis and the driving duration on a particular day is on the y axis. In the future I will add more selectable visualizations for other pieces of the data like idle duration but felt that this was the most important so I added it first. Also, on the vehicle information panel, I made the department name of the vehicle to be a button where upon clicking, the department view of the vehicle's department will be displayed.

I added the department view because the administrators at the U want to be able to analyze the performance of department wide statistics. To do this I first added the functionality to the top left toggle so that when it is switched, the list on the left panel displays all of the departments rather than vehicles. Upon selecting a department, the view on the bottom panel will display information on that specific department, rather than of a vehicle. For the visualization on the department view I wanted to show the usage of the entire department's vehicles. I thought it would be important to know which vehicles were being used so rather than a normal bar chart like in the vehicle information panel, I added a stacked bar chart where each section in the bar is color coded to represent the usage of a specific vehicle in the department. Moving forward, I intend to add a legend for this visualization as well as a tooltip where upon hovering over a specific piece of the bar, you can see which vehicle that piece of the bar is representing. I also wanted to make it so that from this view you could view and select the vehicles in that department. I implemented this similarly to how I implemented the trips list on the vehicle panel view, and by selecting a vehicle on this list, the information for that vehicle will be displayed on the bottom panel.

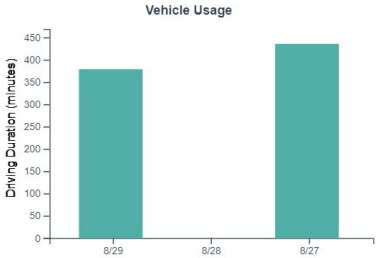
Departments | Vehicles

Vehicle	Department
G9WBW7SYSNC2	Entire Organization
FO15921	Hospital ITS
FO17415	Hospital ITS
G98d210b379c	Entire Organization
G9D6EXUDFM00	Entire Organization
G99DFX0TXMJA	Entire Organization
UFOU6379	Distribution, Hospital Courier
UOFU1513	DISTRO INVENTORY
UOFU1677	emergency room
UOFU1796	UNI
UOFU1869	The Point
UOFU1930	Eye Bank
UOFU1948	Biomedical engineering
UOFU4210	Facilities engineering
UOFU4212	Customer Service
UOFU4326	Huntsman
UOFU4420	UEN
UOFU4431	Facilities engineering
UOFU4461	Environmental Services
UOFU4469	Facilities engineering
UOFU4501	Park Building Pool
UOFU4528	Hospital Courier
UOFU4556	UEN
UOFU4557	UEN
UOFU4572	Transportation
UOFU4580	UNI



Vehicle: FO17415

Department: Hospital ITS
Make/Model: Ford Transit
Class: Class G (8001 - 9000 lbs)
Year: 2016
Vehicle Status: Parked

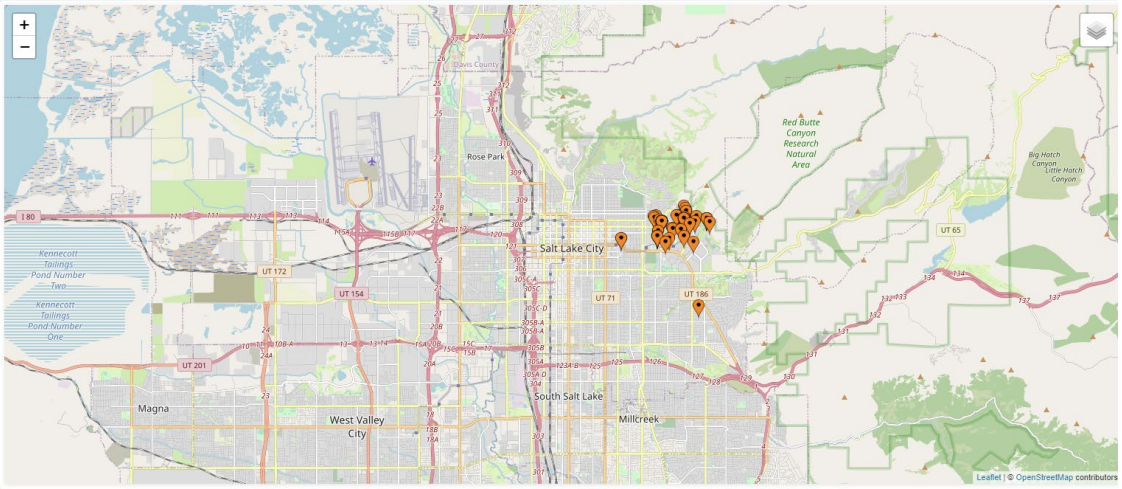


Show Trips From: 2022-09-27
Show Trips To: 2022-09-29

Date	Total Duration
8/27/2022	7:49
8/29/2022	6:52

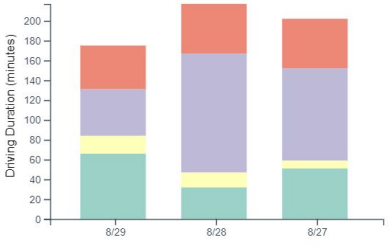
Departments | Vehicles

Department	Num Vehicles
Academic District	1
Air Med	2
Auxiliary Services	1
Biomedical engineering	3
CPD Pool	2
Cardio	2
Carpenter shop	5
Commuter Services	1
Commuter Services Maintenance	5
Commuter services enforcement	11
Construction	4
Core	7
Customer Service	1
DISTRO INVENTORY	2
Distribution, Hospital Courier	1
Electronics	1
Emergency Services	2
Engineering District	3
Entire Organization	6
Environmental Services	1
Eye Bank	1
Facilities engineering	11
Geology	1
Grounds	1



Carpenter shop

Number of Vehicles: 5

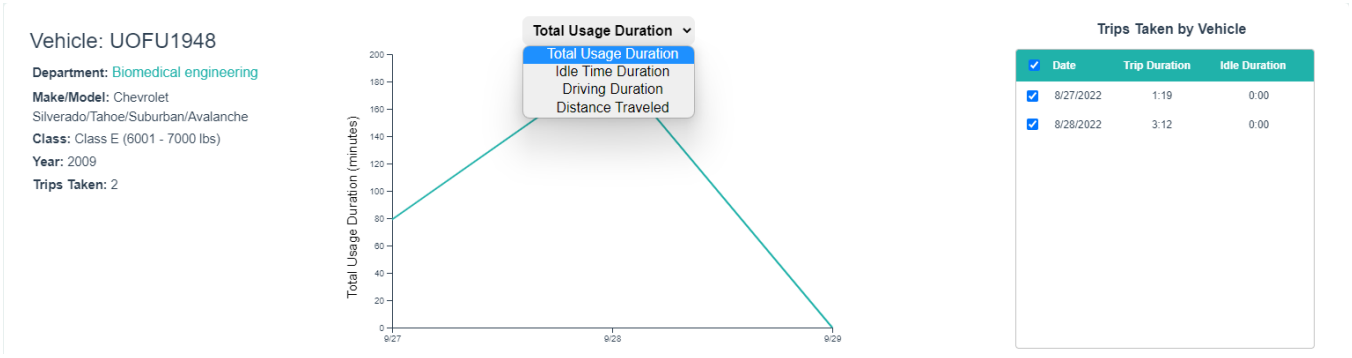


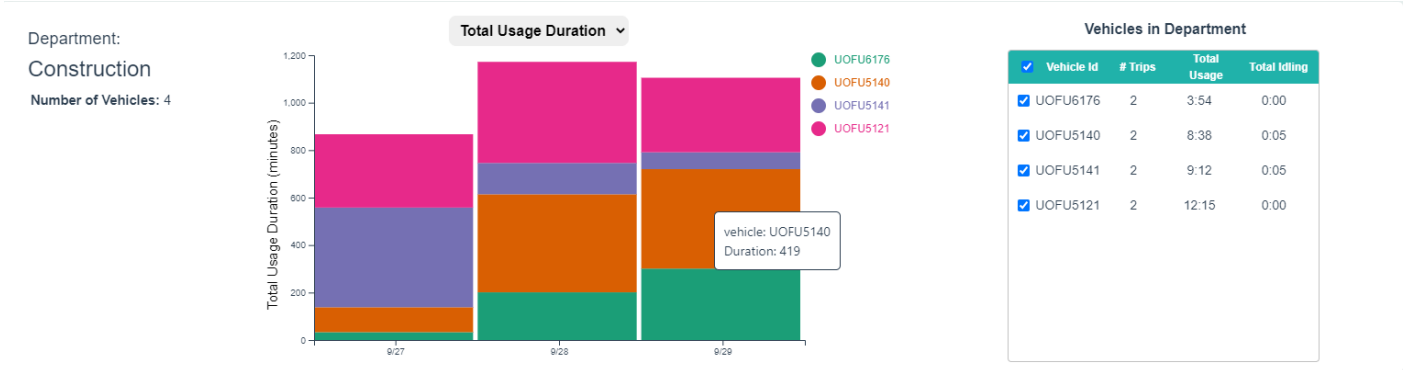
Show Trips From: 2022-09-27
Show Trips To: 2022-09-29

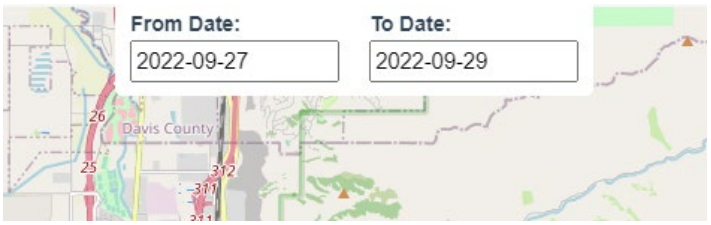
Vehicle Id	Year
UOFU5119	2017
UOFU5120	2017
UOFU6171	2019
UOFU6172	2019
UOFU6311	2020

In the next stage, I first finished the work on the stacked bar chart on the department panel by adding a legend and a tooltip for each section of the stacked bar chart. Next, I wanted to be able to visualize the distance traveled, the idle duration, and the driving duration of vehicles and departments on the bottom panel. To implement this, I decided that rather than adding multiple small charts that would be difficult to read, I added a select box on the bottom panel that allowed users to select the data they wish to visualize. For the department view, I stuck with a stacked bar chart for each of the visualizations since I felt that this gave a good overview of the whole department as well as the ability to see statistics on the individual vehicles in a department.

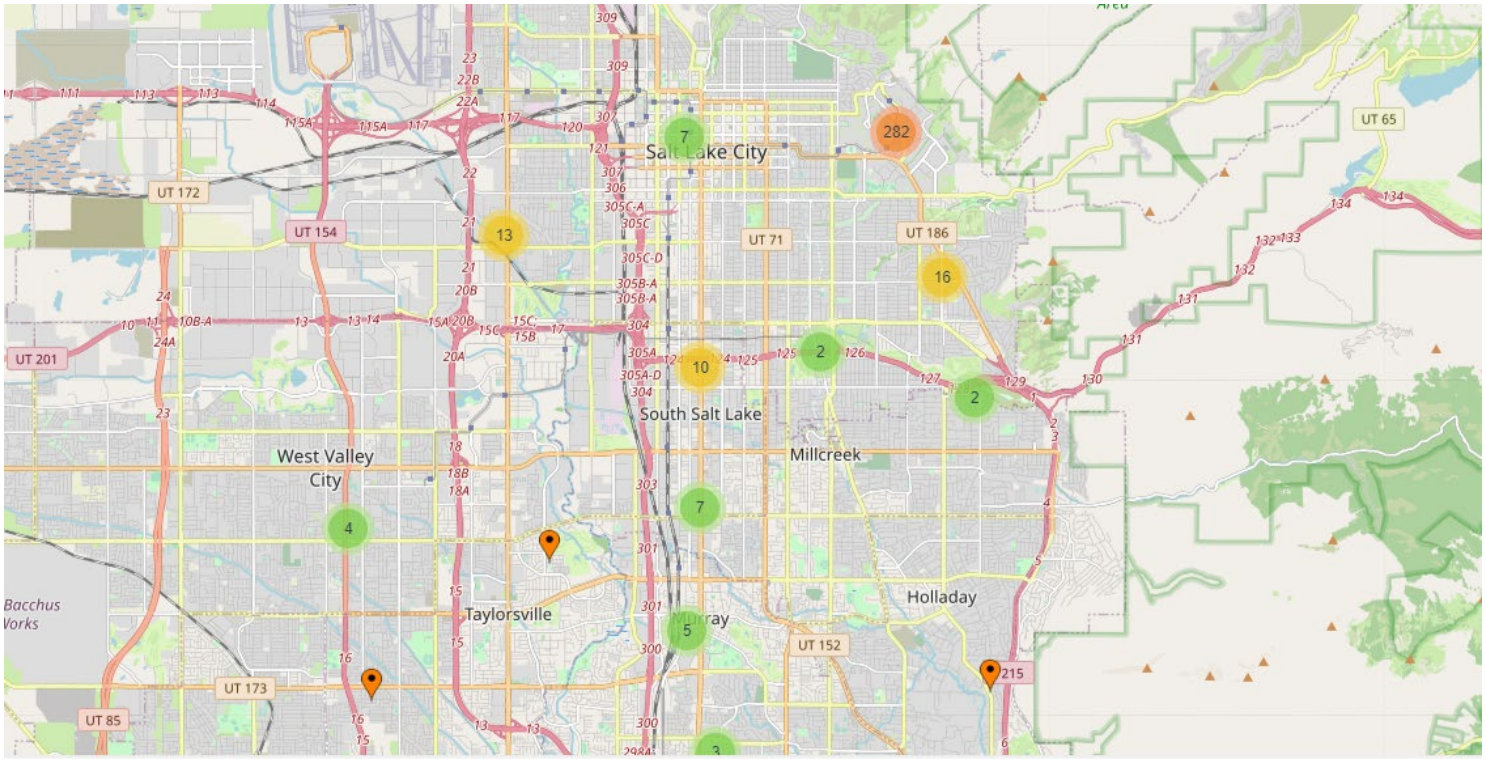
For the vehicle panel view, before adding the other visualizations, I decided that a line chart would work better than a bar graph. This is because when visualizing across a larger date range (at this point there were only 3 days of data in the application) a line graph would be better for making out trends over time. Once I switched from a bar graph to a line graph, I added a select box, similar to the one in the department view, that allows a user to switch between which data they wish to visualize. Lastly, in this stage I felt like since the time select is a central part of the application, I moved it to its own spot at the top of the map instead of on the bottom panels.



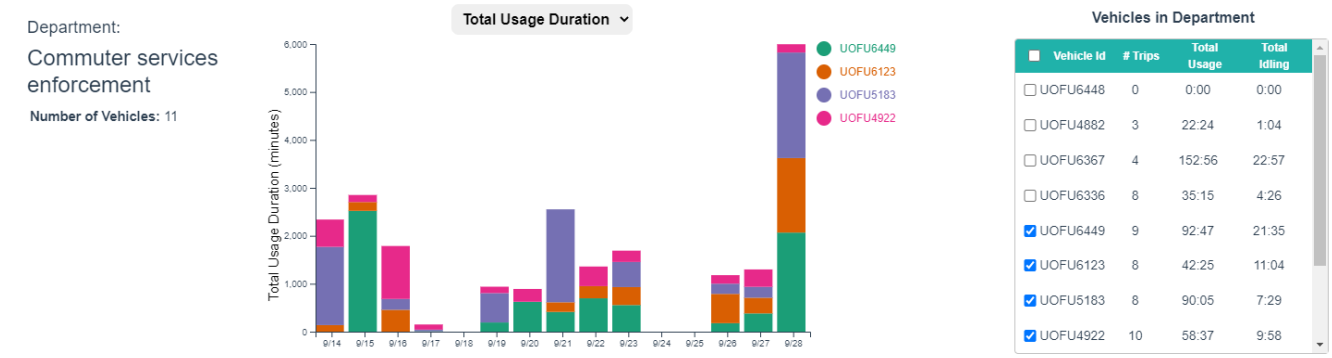




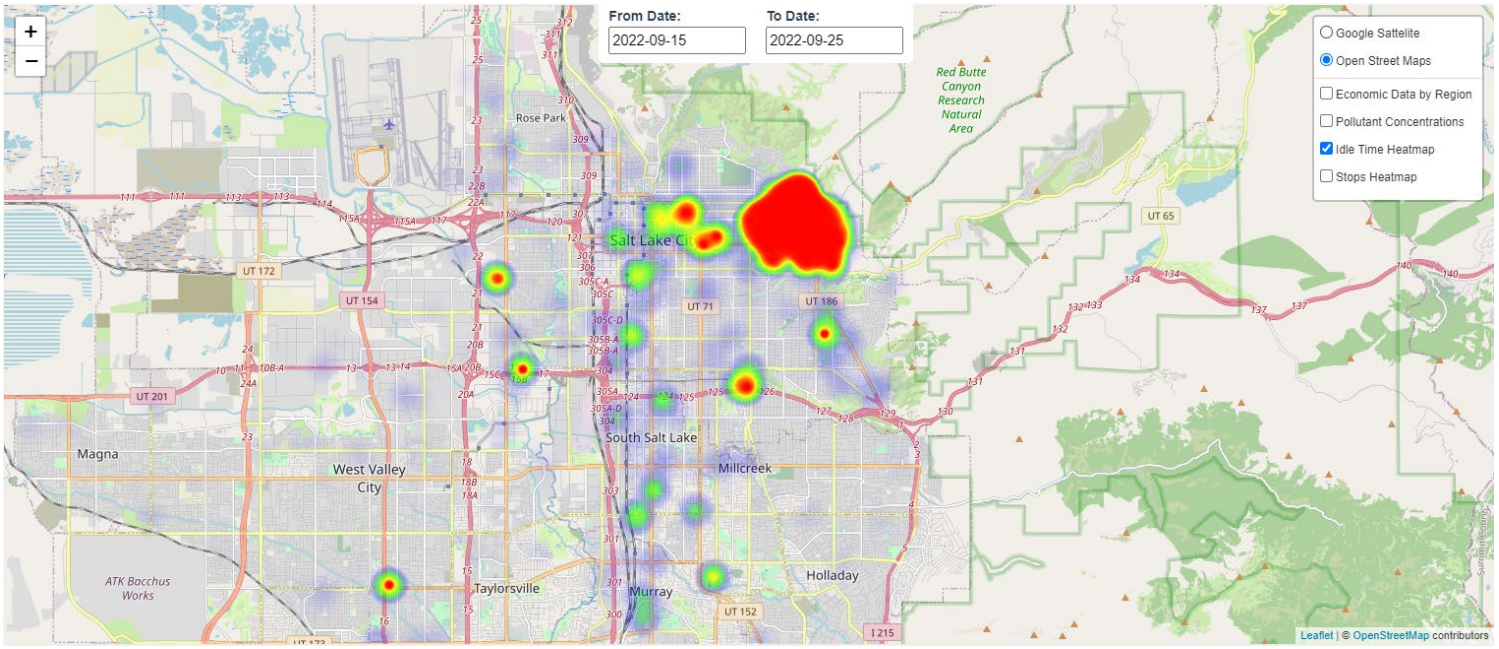
The first thing I did in this next stage was add two months' worth of data into the application instead of the three days that were there before. Doing so brought about two main challenges, the first was that when the selected time range was large, the x axis on the bottom panel visualizations would have too many ticks and became jumbled. The second was that on the map, there would be too many markers that would often overlap and cluttered the map. To fix the first issue, I had to limit the number of ticks that were on the x axis based on the range of dates selected. For the map clutter issue, I decided rather than just showing all of the markers on the map, I clustered nearby points together, such that each cluster is shown on the map. Each cluster is color coded from green to red where clusters with fewer points are green and clusters with more points are red and the number of points in the cluster are displayed on the marker. Upon zooming in on a cluster, the individual points in the cluster are shown and like before, by hovering over the actual marker you can view information on the stop.



After adding more data to the application, I also noticed that some of the departments have quite a few vehicles (the maximum was 11) and for these departments there were too many vehicles to make out which is which on the bar chart. To combat this I added the ability for the user to choose which vehicles would be visualized on the stacked bar chart by checking and unchecking vehicles on the vehicle list on the bottom panel.



As I mentioned before, the people we are working with at the U wanted to focus on geo-spatial visualizations so I then wanted to add visualizations to view where overall stops and idling are happening the most for the whole fleet. To accomplish this I decided to added two heatmap overlays, selectable in the upper right corner of the map. The first is titled ‘idle time heatmap’ and displays a heatmap on the map for idle time, the second is titled ‘stops heatmap’ and shows a heatmap displaying where stops are most frequent.



Lastly, I worked fairly hard on trying to implement a grid heatmap visualization on a panel that would be toggleable in place of the map. On this grid vehicles would be represented by the rows and columns and each cell would be color coded based on the similarity of two given vehicles behavior, calculated by a score developed by the other members of our research project. This visualization is intended to help decision makers consolidate work into fewer vehicles. Unfortunately, I got COVID during the week before this project is due and because of that was unable to complete this part of the visualization in time. This visualization will be completed as we continue the research project.

Evaluation:

This project has helped me learn quite a bit about the vehicles that are managed by the University of Utah. The main thing I learned, and I hope the users at the U will learn using my tool, is that there are quite a few vehicles on campus that are severely underutilized. I would estimate that at least 10 vehicles could be removed from our fleet to reduce costs if managed more efficiently, which would significantly reduce cost. Especially since currently, the U leases out a portion of it's vehicles from external sources. I also found that idle time was not that great for the most part across the fleet, however a few outlier vehicles idled significantly, and most of the idling happened on campus.

In my proposal, some of the questions I wanted to answer and some of the 'optional features' that I wanted to include were dependent upon access to fine graine trajectory data of the vehicles. Unfortunately for this project and the research that we are trying to accomplish, this trajectory data was never made available to us and therefore I was unable to create visualizations for it. Hopefully this will become possible next semester.

Overall, I am happy with the visualization application I created and I believe that it works well for its intended purpose. A central focus for me on the project was interactivity, I wanted to provide a tool that would allow for in depth exploration of data rather than presenting my own analysis. The tool that I have created has a significant degree of interactivity and allows users to quickly and easily interact with the data. While I was not able to complete all of the visualizations that I had hoped to be able to include, I feel as though what I was able to complete allows users to understand the data very well and answers most of the questions that I set in my proposal. I put most of my work into the geo-spatial visualizations, as guided by the administrators at the U we are working with and am happy with what I was able to produce there. I had strongly hoped that I would be able to include a heatmap visualization as I feel like that will be very helpful for decision making but due to COVID, I was not able to complete it.

Moving forward I intend on completing the heatmap visualization I previously mentioned as well as more fleet wide visualizations. I feel like while I provide a good fine grain view of the data, I do not provide much in the way of viewing the fleet efficiency as a whole which will be useful. I intend on including these visualizations on a toggleable panel which would take the place of the map view.