Dashboard of COVID-19 cases and corresponding travel behavior pattern in Salt Lake County

Process book

Moongi Choi, Nancy Lyu, Xinyan Yan

Github Link: https://github.com/MC1316663/dataviscourse-projectCovid
Web site link: : https://mc1316663.github.io/dataviscourse-projectCovid/Project.html

Table of Content

Overview and Motivation

The outbreak of COVID-19 has greatly changed our daily life and working style. For example, in the early stage of a pandemic, most organizations, such as companies, restaurants and schools, moved their meetings, services and courses online. People started working from home and tried to avoid going out as much as possible. Besides, the severity of the epidemic varies in different periods, which further affects people's life and work behaviors.

Inspired by this, this project aims to investigate the evolution of people's travel behavior in Salt Lake County during the COVID-19 epidemic (2020-2021) from a visual perspective. Given the available data and the goal of this project we aim to inspect three kinds of people' behaviors for each day: how many time people stay at home, how many people working in person, and how many time people spent in outdoor activities. To be specific, we study the pattern of such behaviors, the relationship between the behaviors and the Covid cases, and more importantly, the spatial positions of each area in Salt Lake county of by visualizing the map.

Related Work

Our inspiration for the design of visualization and interaction is mainly from the class discussion and the assignments.

Visualization design. As discussed in class that position is an effective channel in describing the quantitative attributes and that multiple visualizations facilitate the comparison analysis. We designed the evolution of the three kinds of patterns and Covid cases with bar charts for each behavior, where they share the x-axis indicating time. Inspired by the concept of 'Overlay', we designed an overlay on the map consisting of multiple block areas, where each is colored by a specific attribute.

Interaction design. We mainly used several interaction techniques taught in class,
(1)'Overview + Detail'. We first show the overall behavior pattern, then users can select a specific area and time period to conduct further analysis.
(2)'Filtering'. We allow users to filter a specific period of areas.
(3)'Brushing'. The filtering function is achieved mainly by brushing the time axis.
(4)'Linking'. We link all of these views together, for instance, every time users select a time period, we update the value in the map view; every time users select an area block, we update the information in the right view.

Questions

Our current questions includes:

- Show the evolution of various behaviors of people in Salt Lake County during the pandemic.
- Reveal the relationship between people's behaviors and the number of Covid cases.
- Investigate the detailed information of a specific area in Salt Lake City.

Data

We used three kinds of data for this project: 1) 'social distancing matrix' from SafeGraph (2020.01.06 to 2021.02.03), 2) 'census block group' spatial data (.json or .shp file) 3) Total count of COVID-19 cases, recovered population and casualties in Salt Lake County. Social distancing matrix is the data that records individual phone device counts at home, work and others (other behaviors). Spatial unit and time unit of the data is CBG (Census Block Group) and hourly based (some are minute based but we will convert those to hourly based data) separately.

Data	Description	Source
1. SD (Social Distancing) Matrix	Smart device counts in Salt Lake county by time. The data divide people's travel behavior by census block group unit with home, work (full, part time), other behaviors.	Safe Graph (Metadata and schema) (https://docs.safegraph.com/docs/social-distancing-metrics) Example map: (https://carto.com/spatial-data-catalog/browser/dataset/sg_social_dist_667d8e8e/map)
2. Census block group (GEOJSON)	Spatial data (Polygon) illustrating Salt Lake County by census block group unit	Utah GIS (https://opendata.gis.utah.gov/datasets/utah-census-block-groups-2020/explore?location=39.472883%2C-111.547240%2C-1.00)
3. Total Count of covid-19 cases	Total Covid-19 cases, death rate, recovery rate from 2020.01.06 to 2021.02.03 in Salt Lake County	Salt Lake County Health Department (https://slco.org/health/COVID-19/data/)

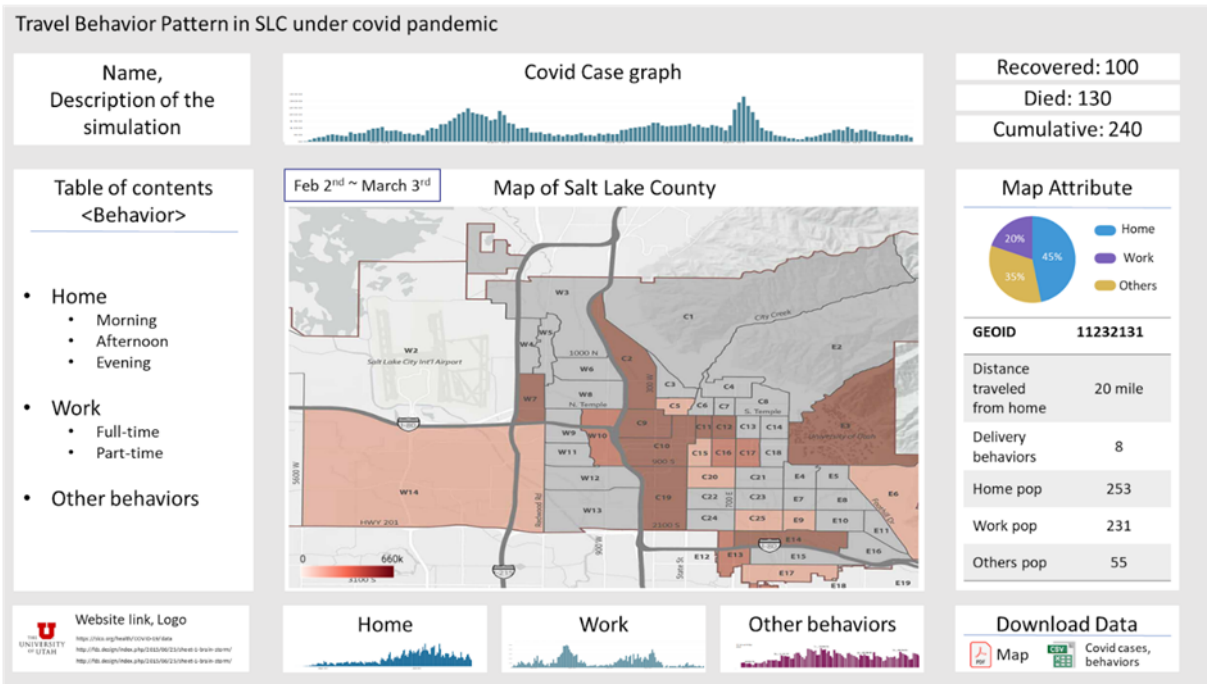
Data processing

As the spatial data and the Covid data of salt lake county is available and can be used directly, our data processing mainly focus on the ‘Social distancing matrix’ data. We will first filter out all data related to the Salt Lake Countym and then only reserve attributes that we will use in this system. With the filtered data, we then check the completeness and formart of the data. We will annotate the missing day, and for the missing attributeon a particular day, we will update it with the average or the most common value of other available rows. To visualize the number of people staying at home, working full time or part time acorss the entire time span, we need to select all relevant data along with the time attribute.

When interacting with the system, we need to dynamically aggregate over a subset of data or retrieve the data of a particular day. To facilitate this, we need to store all data into a list, where each element represents one day, in this way, we can quickly retrieve the subset of the data according to the time.

Design evolution

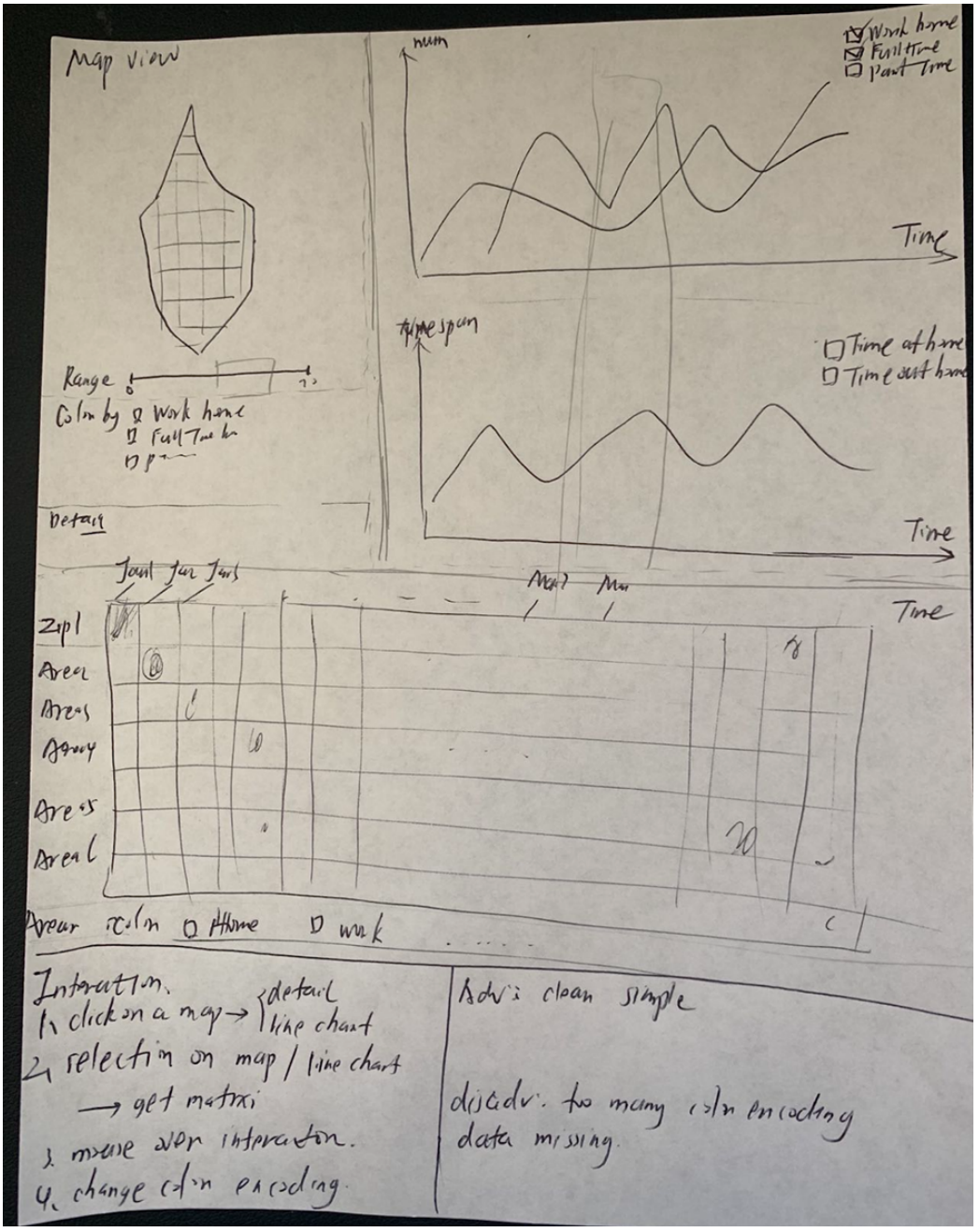
Prototype 1



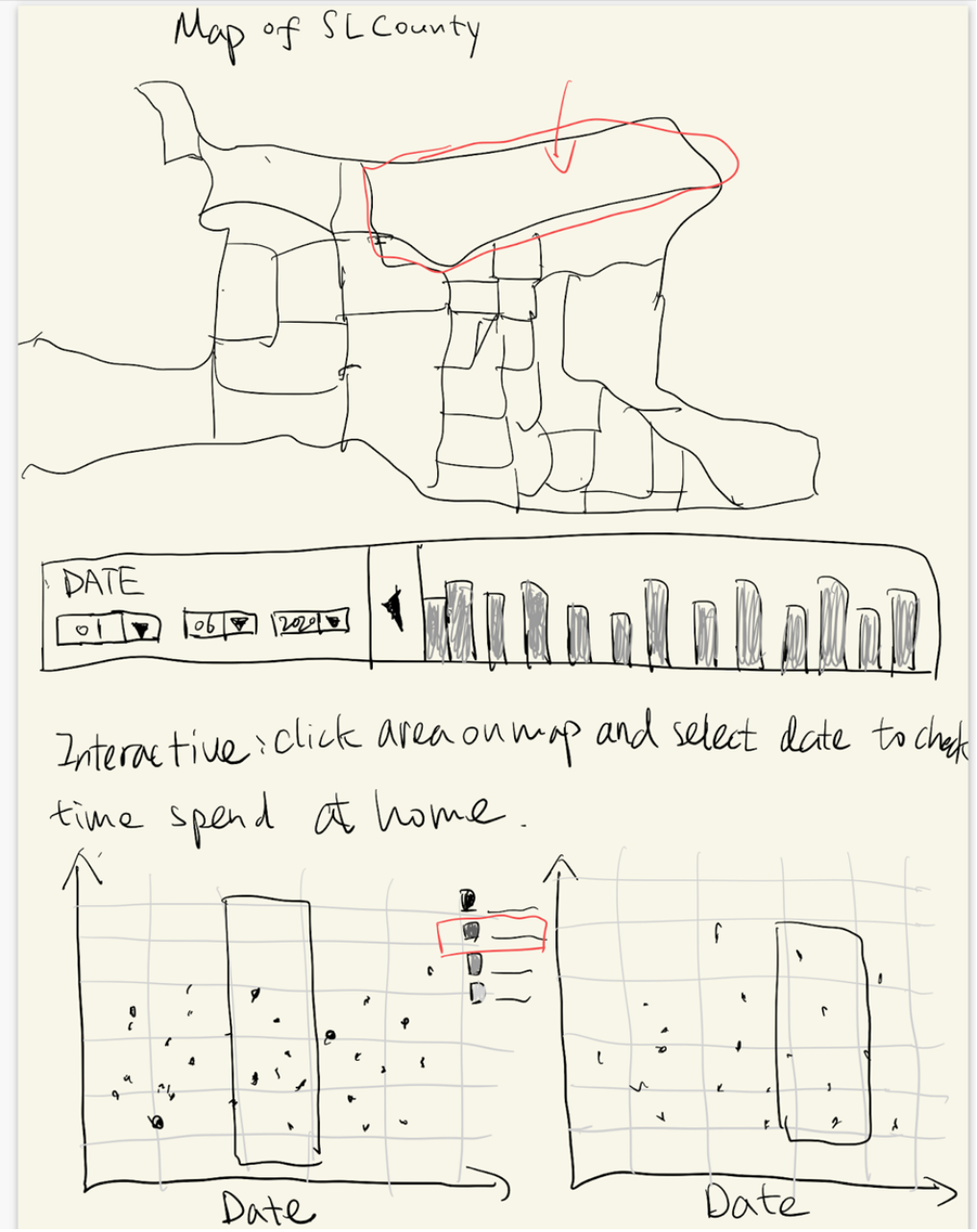
The purpose of this visualization design is to draw a web dashboard showing travel behavior pattern change in Salt Lake County by the chosen period under covid-19 pandemic. There are several elements which shows spatial and temporal filtered dataset such as how many people got infected and recovered, bar graph of covid-19 cases, bar graph of population who are at home and work, choropleth geographic map classified by each behavior's counts, pie-chart and detail attribute of one polygon (census block group) in time period chosen.

There are two that should be automatically filtered and updated according to the user's choice. Covid-19 cases and social distancing matrix. There are two things users can choose, the period of what they want to see, and which behavior (home - morning, afternoon, evening, work - part time, full time, other behaviors) that will be drawn. From their choice, the graph, map, attribute table, and downloadable data will be changed. This is the filter process we need to employ for each data.

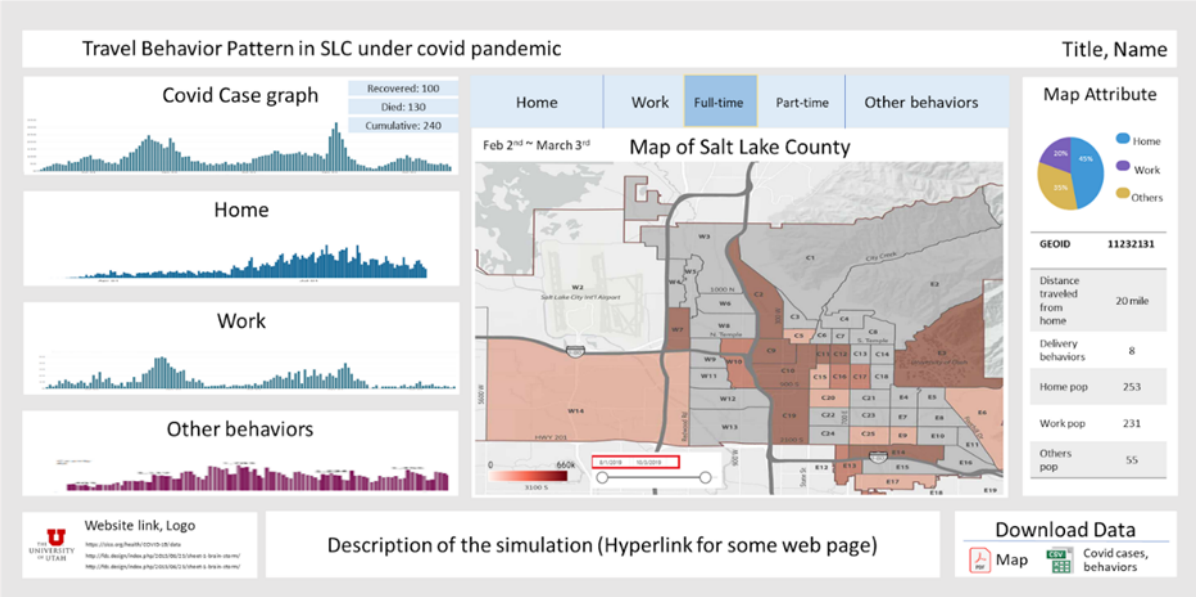
Prototype 2


Map view
num
Work home
Full Time
Part Time
Time
Range
Colm by
Work home
Full Time
Part Time
timespan
Time at home
Time not home
Time
Detail
Time
Zipl
Area
Area
Area
Area
Area
Area
Area
Area colm
Home
Work
Interaction.
1. click on a map → detail / line chart
2. selection on map / line chart → get matrix
3. mouse over interaction.
4. change color encoding
Adv: clean simple
disadv: too many color encoding
data missing.

Prototype 3


Map of SL County
DATE
01
06
2022
Interactive: click area on map and select date to check
time spend at home.
Date
Date

Final Design

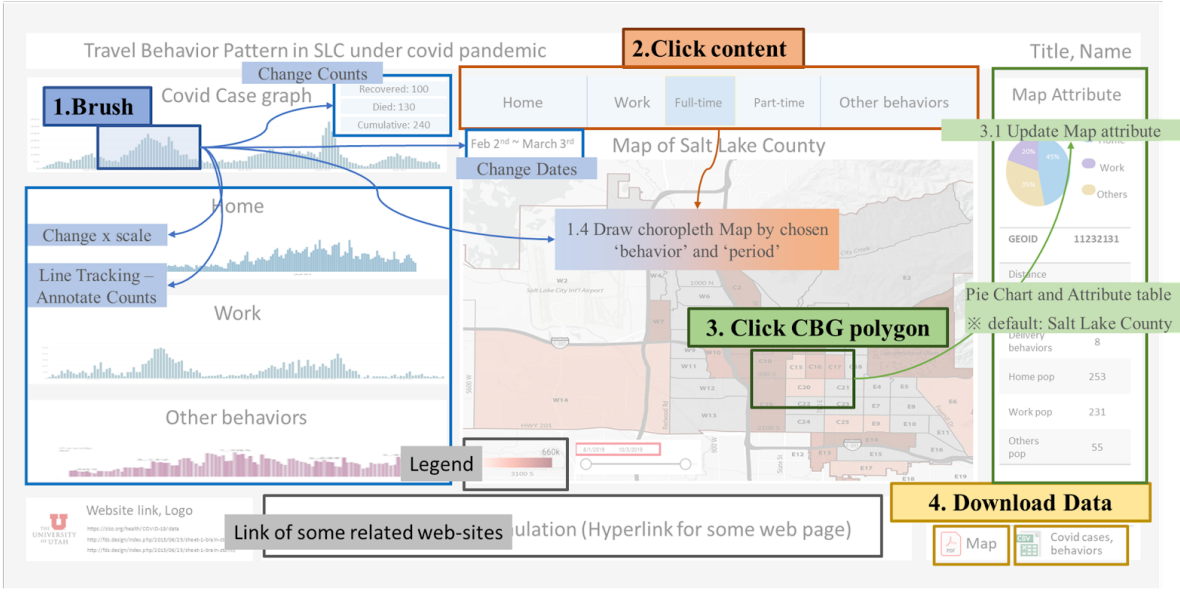


We changed the layout more concisely and efficiently for users to understand the dashboard. And combine our ideas of design and interaction functions by brainstorming. We place the main time graph to the left side to see the change of each data pattern by time more clearly. We shrink the size of the map but at the same time, we put the table of contents (home, work, others) and range in the map layout section so that people can see map-related elements in the same area. On the right panel, we make the attribute table and pie chart smaller than the initial design so we can secure more space for the graphs.

Implementation

Interaction

Interaction



There are 4 main interactions in this dashboard and those have each several functions being connected to other elements. Each interaction, function and filter data plan are described in table below

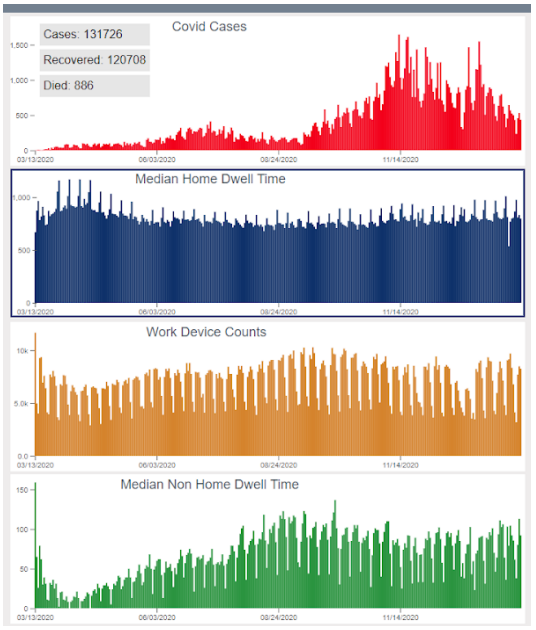
Table 3.Interactions and functions description of Final design

Interactions	Functions	Data Filter
1. Brush	1.1 Change counts	**Covid-19 case data** Filter 3 columns (Recovered, Died, Cumulative cases) selected by period from brush
	1.2 Change x scale of 3 graphs in the left panel	**SD (Social Distancing) data** Filter Home, Work (sum of part- and full-time work), and other behaviors selected by period from brush
	1.3 Line tracking function for each graph in the left panel	
	1.4 Change Dates in map	**Covid-19 case data** Filter the period from brush
2. Click content	1.5 / 2.1 Draw choropleth map	**SD data, Map polygon data (GeoJson)** Filter SD data by selected content (home, work, work-full time, work- part time, others) and period selected by brush
3. Click CBG polygon	3.1 Update Map attribute	**SD data**

		Filter SD data (mean or sum) by selected polygon and the period selected by brush to make pie chart and attribute table
4. Download Data	5.1 Download Map and csv	**SD data, Covid-19 case data** Filter data from period from brush Captured map from web map

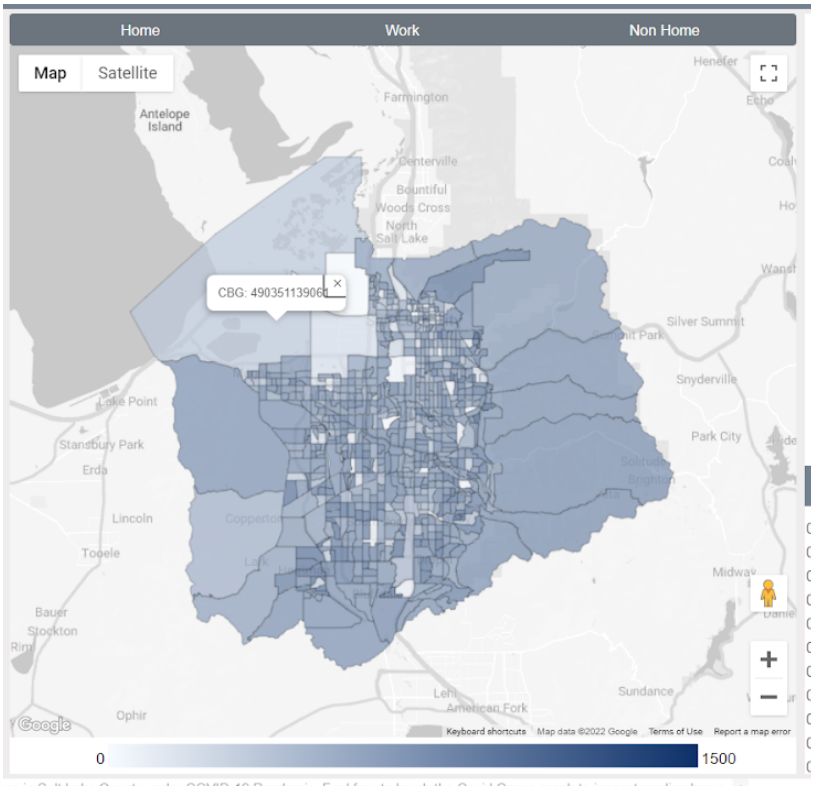
Each design

1) Bar Charts



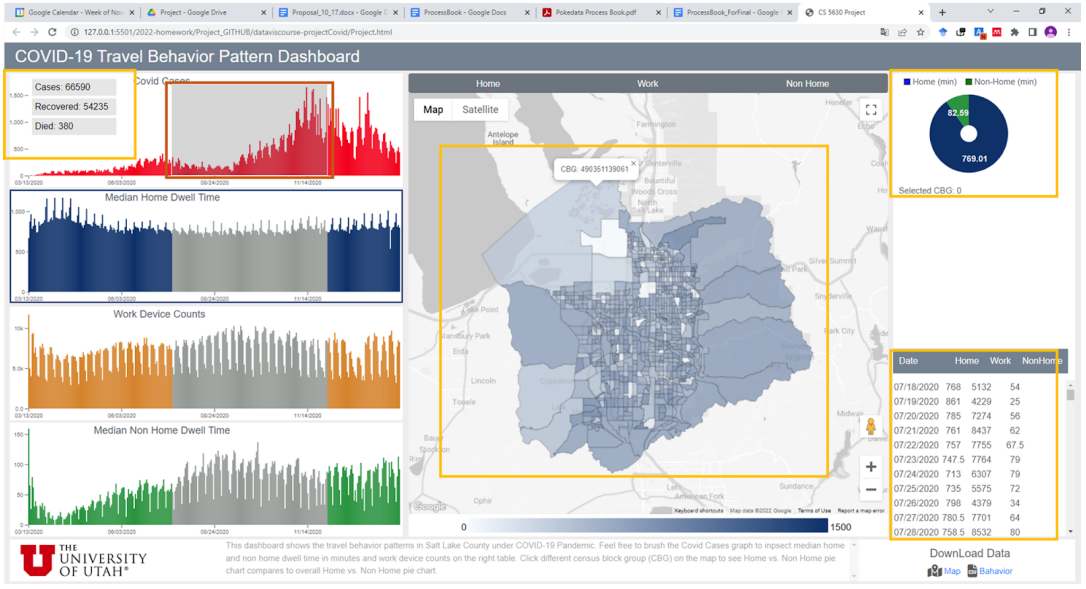
- We have four bar charts in the dashboard that users can have interactive actions and check travel behavior patterns.
- Each one is Covid-19 cases, Median Home dwell time, Work device counts, and median non-home dwell time.
- From bruising within Covid-19 cases graphs, we can select the certain time period that we want to see the changing pattern of the map, pie chart, and table

2) Map



- In the center of the dashboard, we have an interactive map to zoom in and out.
- This is a choropleth map that interacts with brushing period, and button on top (home, work, non-home)
- Colors are automatically changed by brushing and button
- Tooltip of CBG ID when user's mouse is over each CBG.

3) Brush

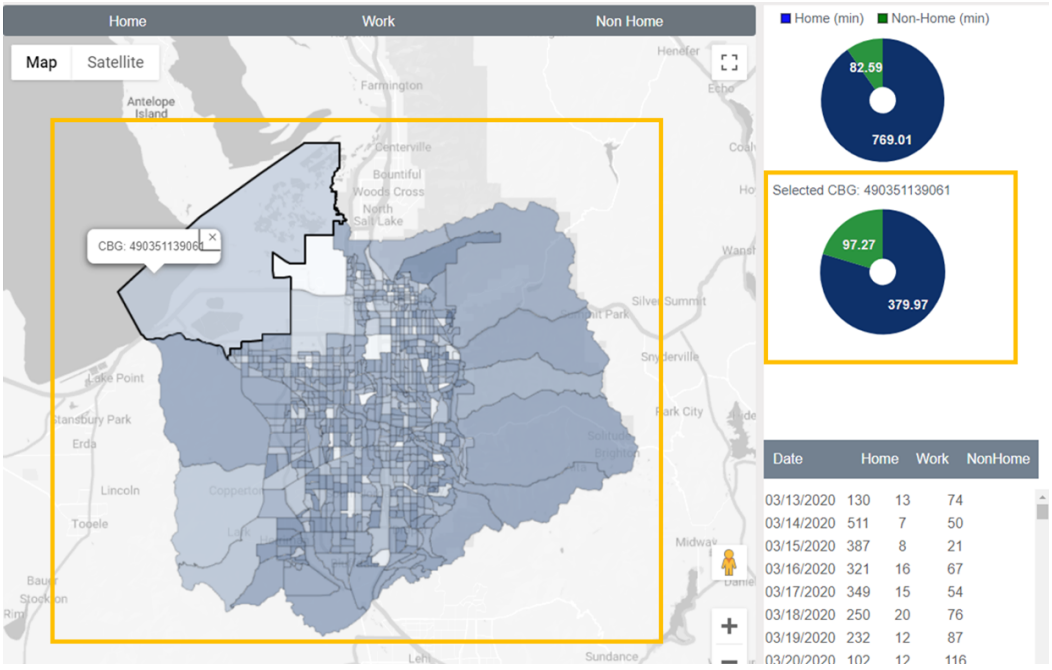


- When the user uses the brush function and selects the time extent, it changes several visualizations in each section. First, change counts information such as Recovered cases, casualties from COVID-19 within the period selected. Second, Highlight the x scale of 3 graphs in the left panel. Third, update choropleth maps according to the selected period. fourth, change the pie chart and table on the right panel.

4) Click contents

- When the user clicks the button of home, work, non-home activity, it will filter the data by selected content and automatically update the map.

5) Click CBG polygon

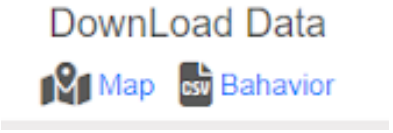


- If the user clicks each CBG block, the right panel shows the detailed information of the selected CBG and pie chart showing the home and non-home time of that CBG (So use can compare the home and non-home time at the CBG they clicked to the average home and non-home time at SLC.

6) Table

- When we brush the period or click the CBG in the map, the table keeps changing its value corresponding to the actions. For example, if we brush and click CBG at the same time, the table shows the information in that CBG in the selected period.

7) **Download Data**



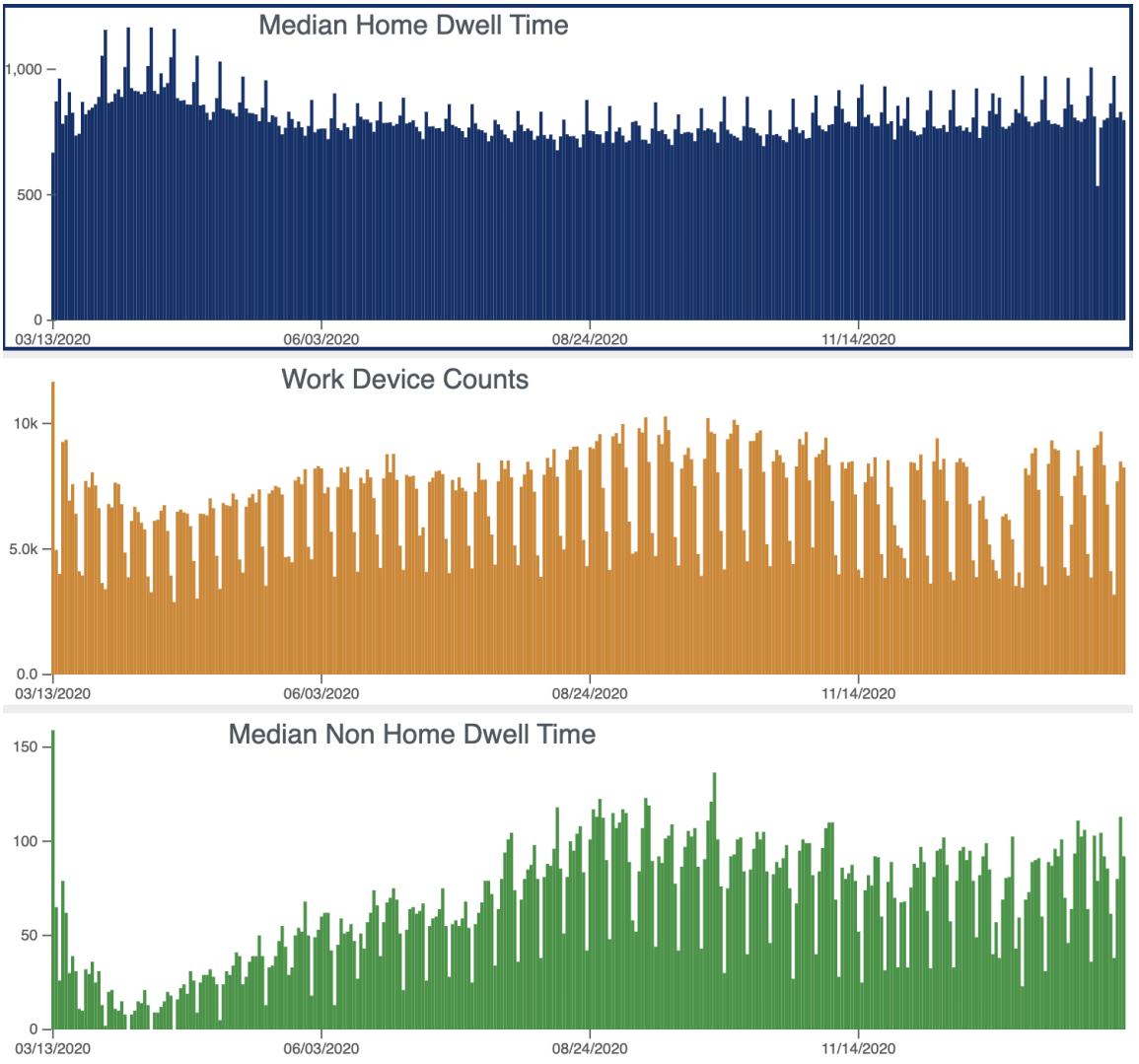
- User can download the origin data to use their own research

Evaluation

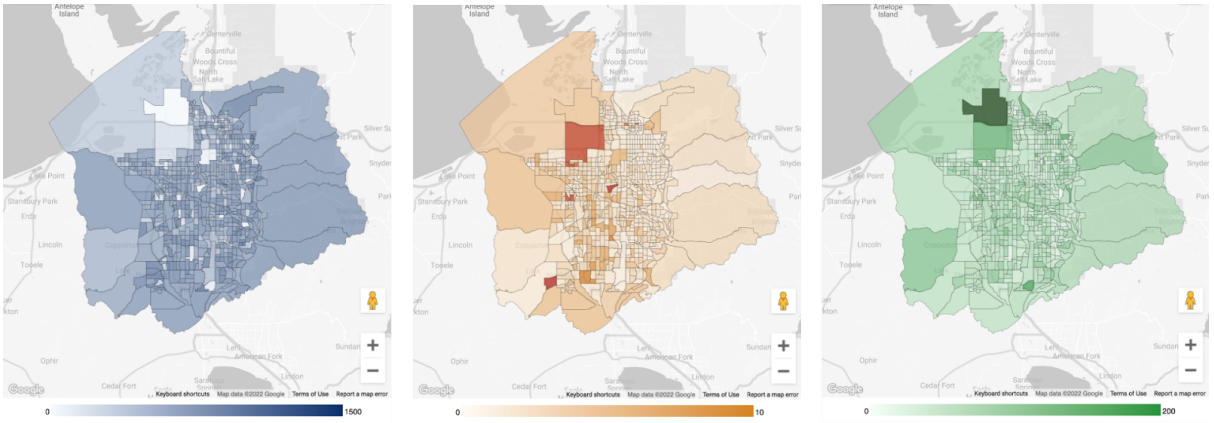
Answeing Questions

Overall, through the visualization and data, we learnt the evolution of various activities under Covid pandemic, we will elaborate these insights and findings through answering question we proposed.

Question 1: Show the evolution of various behaviors of people in Salt Lake County during the pandemic.



- For the ‘Home Dwell time’, we can tell that at the start of the pandemic, the time increase greatly, which means that people tend to stay at home as much as possible. But after 2-3 month, the number starts to decrease and keeps steady.
- For the ‘Work device count’, at first, there are no many devices, but as times goes, the devices gradually increase, which has two peaks in Jun and September of 2020.
- For the ‘Non Home Dwell time’, we can notice that such behaviors decreases dramatically, then increase gradually which aligns with the ‘Home Dwell time’. Also, acrond 11/14, 2020, there was another local minimum may indicating that the pandemic was severe again.
- We can also notice a pattern from the three charts that every five days, there is a new low/high peak, indicating that the difference between week days and weekends.

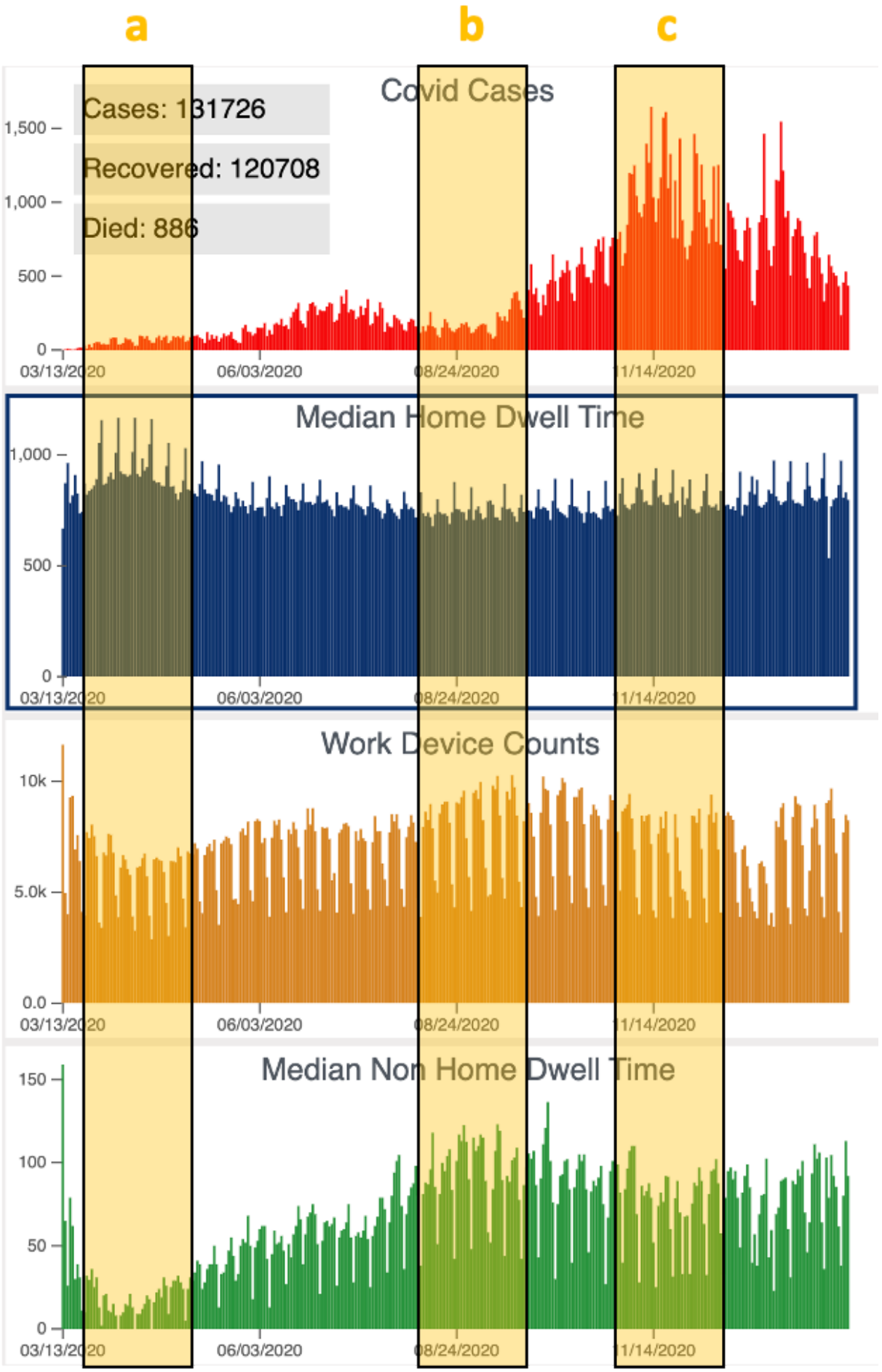


The above figure shows the three kinds of activities from a high level, where each area is colored by the average value of this block across the entire timespan. We can tell that the the values vary on different areas, and areas with low value of ‘Home dwell time’ tyopically have high values of ‘Work device count’ and ‘Non home dwell time’, which is also reasonble.

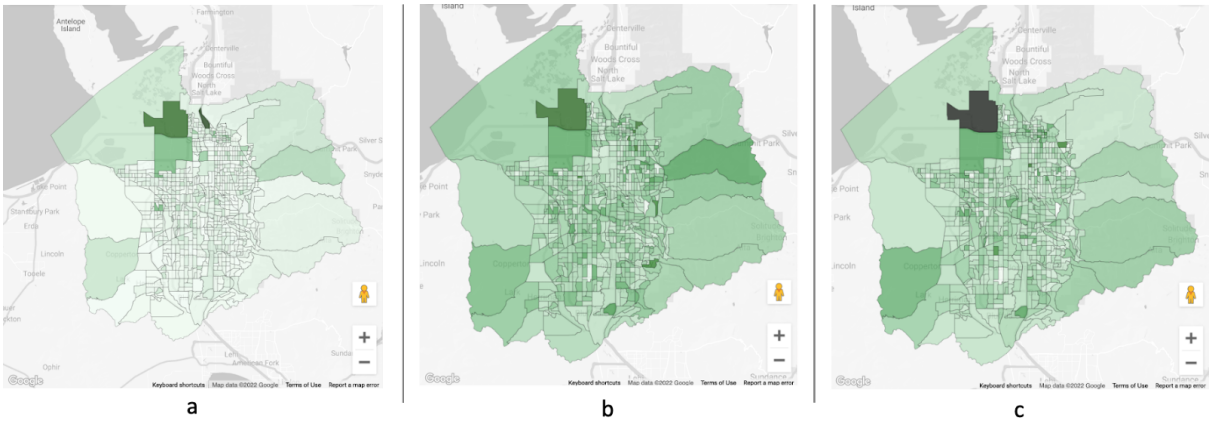
Question 2: Reveal the relationship between people’s behaviors and the number of Covid cases.

As shown in Figure below, we marked three time periods ‘a’, ‘b’ ‘c’, representing the start of the pandemic, the middle of the pandemic, and another peak of the pandemic, respectively.

At the start of the pandemic, a lot of people started to stay at home more and decreased the time being outside in case being affected. And there were no many work devices.
In the middle of the pandemic, people starts to go outside more and the values keep steady, and the number of work device increased. This results may represent that people already get used to the pandemic and found ways about being survived during pandemic.
In the another peak of the pandemic, we don’t notice the change of ‘Home dwell time’, but there is a small decrease of ‘Non home dwell time’, which may indicate people will take more actions during severe periods but the pandemic will not effect people’s life too much.

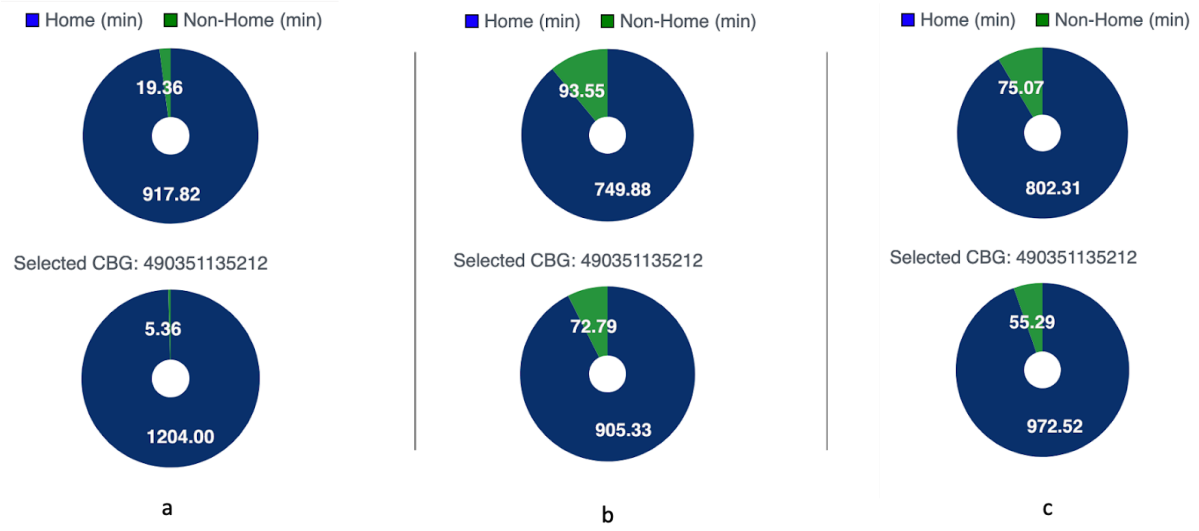

a
b
c
Covid Cases
Cases: 131726
Recovered: 120708
Died: 886
1,500
1,000
500
0
03/13/2020
06/03/2020
08/24/2020
11/14/2020
Median Home Dwell Time
1,000
500
0
03/13/2020
06/03/2020
08/24/2020
11/14/2020
Work Device Counts
10k
5.0k
0.0
03/13/2020
06/03/2020
08/24/2020
11/14/2020
Median Non Home Dwell Time
150
100
50
0
03/13/2020
06/03/2020
08/24/2020
11/14/2020

We further compared the map of ‘Non Home Dwell Time’ at the three time periods, respectively. As shown below. For the middle and buttom areas, At first, many people don’t have much outdoor time (a), but the time increases gradually as shown in (b). Even in severe time, the time being outside didn’t decrease too much (c).



a b c

Question 3: Investigate the detailed information of a specific area in Salt Lake City.



After clicking on a specific area, we can look at the data of this area shown in the table of the right panel, also, we show the pie charts fo this area representing the proportion of time staying at home and time staying non-home.

As shown in above figure, we select an area (CBG ID: 490351135212), and display its states at time periods a, b, and c, which are consistent with the above time period. The above one represents the average time composition across the Salt Lake County, and the buttom one represents the selected CBG. We can notice that this area always has a lower value of non-home time than average for the three time periods. In particularly, during period a, most activities took place at home; during time period b, people started to go outside more; however, during time period c, proportion of being outsides decreased mildly.

Improvement

Overall, our visialization can answers all of our proposed questions. However, there still has lots of improvement space.

- Show more detailed information of the selected area. Currently, we only show the three types of activities of the selected area, but showing more information will give users more includes. For example, as we use the average value, showing the number of this residents in this area will reflect the reliability of the data. Also, showing the type of this area, such residential area or bussiness area, is also meaningful in reasoning.
- Compare multiple areas. We only support to select one area and look at its information. Enabling comparison among multiple areas would help detailed inspection of the data. For example, by comparing two areas' activities over time, we can get new insights about what factors may cause such difference.
- Improve the data precision and visualize the uncertainty inside of data. As the data is collected by devices, there are a lot of factors that will effect the precision of the data. Thus, showing these data in a certain way may introduce bias, in the further, we can improve this by visualizoing the uncertainty of the information.