Process document: An American Day

Ben McMorran and Francisco Sanchez

Overview and motivation

Provide an overview of the project goals and the motivation for it. Consider that this will be read by people who did not see your project proposal.

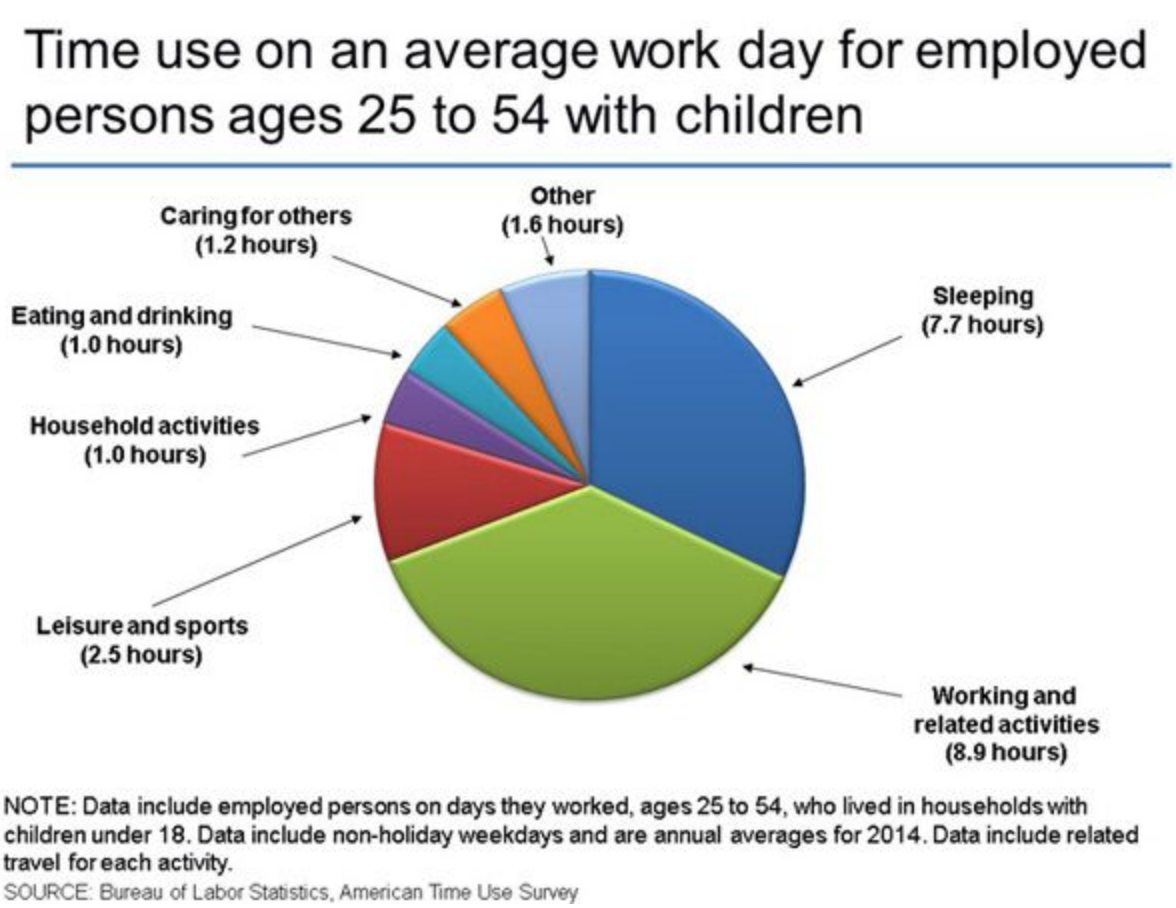
How does the average American spend their day? Using data collected by the Bureau of Labor Statistics (BLS) in the American Time Use Survey (ATUS), our project set out to explore the the daily habits that we all share. ATUS is a continuously run survey that asks respondents aged 15 and older to recall every activity they did in the last 24-hour period. Over 3,000 samples are collected each month, and this information is used in a wide variety of applications such as demographic research, quantifying the amount of non-paid time that people work, and measuring time spent on elder and child care.

Compared to these economic applications, our project is directed towards a much more casual, exploratory audience. We seek to create an interactive, fun, exploratory tool for users quickly gain a high level understanding of the ATUS data, see how they compare to the average, and look for insights in time use differences among demographics. While the ATUS data obviously works for comparing daily routines, the interview date combined with age of the respondent allows us to to explore other timescales as well, such as a week, year, or even a lifetime. Our visualization should make it easy to jump between these timescales during exploration.

Related work

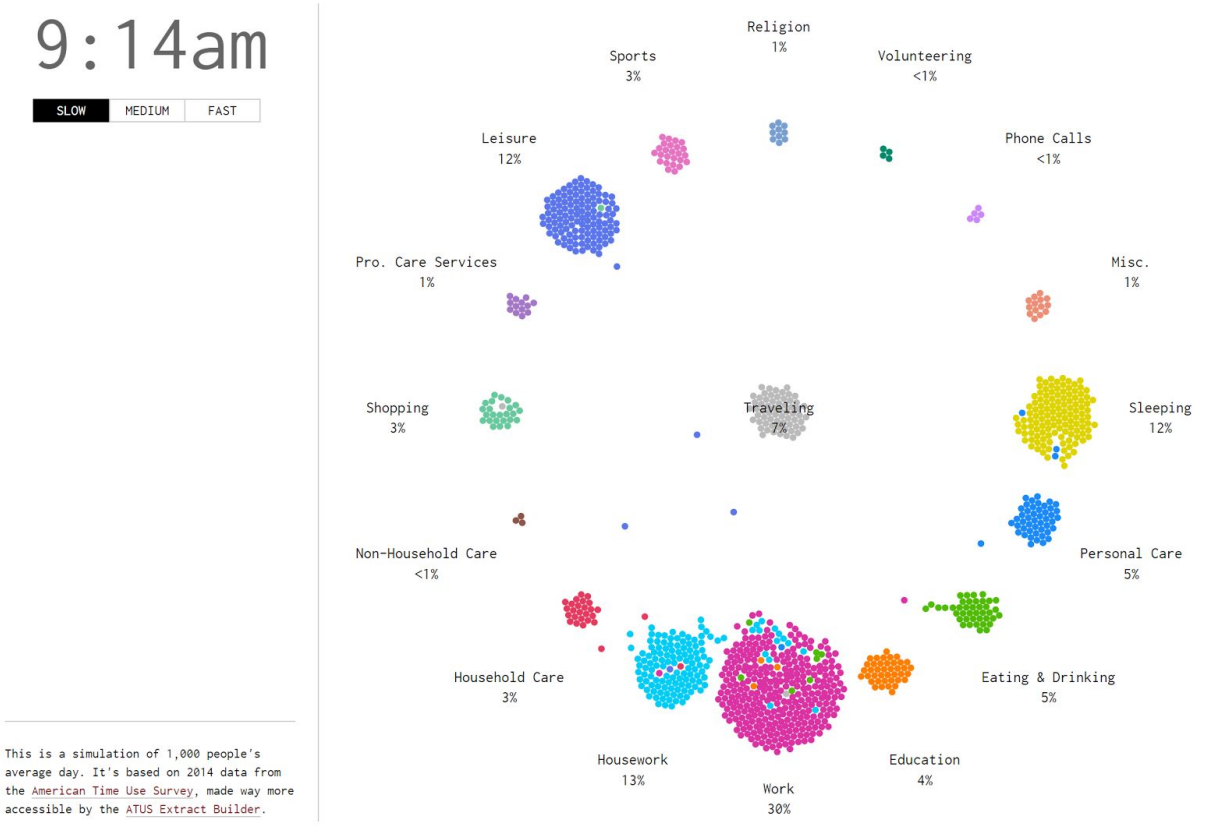
Anything that inspired you, such as a paper, a web site, visualizations we discussed in class, etc.

We are far from the first to visualize the ATUS data, and we drew inspiration from several sources. First, the BLS itself provides some simple (albeit non-interactive) visualizations of the data in the PDFs they publish on their website.



While these simple visualizations work to some degree for conveying very high level statistics, they are poor from a visualization perspective. Pie charts use an angular encoding, which is known to be perceptually inferior to position encoding. BLS also provides bar charts for some statistics, but here they suffer from being non-interactive and trapped inside a PDF. However, the BLS information can still provide useful guidance for finding insights in the data, and our tool should be able to visualize these same insights.

Another visualization comes from FlowingData. This visualization focuses much more on individuals that simple averages like the BLS. Each dot in the visualization represents a person moving about the activities of their day. Transitions are simulated by a Markov process trained on the original ATUS data.



The individualized aspect of this visualization adds a lot of visual interest, but it's difficult to gain deep insight. Time is mapped to actual time through animation, so it is difficult to compare relative percentages of individuals in different activities at different times of day. One thing this visualization does well is show the sequence of activities that an individual participates in throughout their day. It's easy for this to wash out in averages, but the sequence may provide important insights.

FlowingData has several other visualizations of the ATUS data focusing on different aspects like the most common activities broken down by age and sex and different times of day, when Americans leave for work by county, and histograms of how many hours Americans spend on different activities. Each of these visualizations do an excellent job of telling a focused, compelling story about the data while allowing the user to interact with the data to see how they compare.

Questions

What questions are you trying to answer? How did these questions evolve over the course of the project? What new questions did you consider in the course of your analysis?

Our visualization is designed primarily as a tool to explore the answers to two main questions:

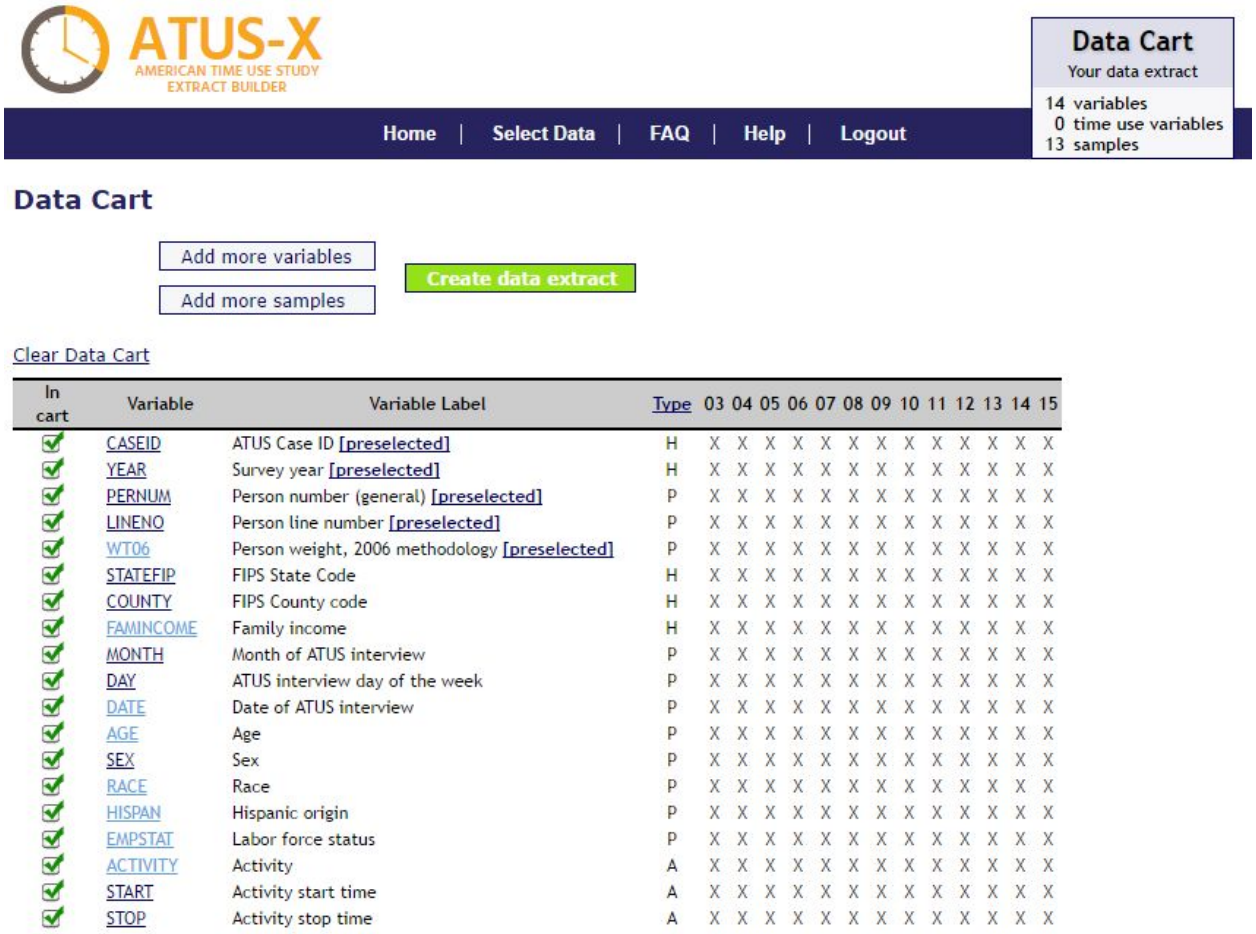
- Which activities do Americans do most commonly over different time periods (day, week, year, and lifetime)?
- Are there significant differences in activity prevalence between different demographics (split on age, sex, race, employment, and income)?

Originally we were also interested in exploring how activities vary by location (FIPS codes are available in the data), but due to time constraints we were not able to explore this dimension. Otherwise our questions remained mostly intact from conception in the proposal through to the implementation.

Data

Source, scraping method, cleanup, etc.

As discussed above, our data originally came from the BLS. However, the raw BLS data is spread across several files and has slight methodology changes year-to-year, making it difficult to interpret. Instead we used the ATUS Extract Builder (ATUS-X). This tool allows users to select the fields of interest from the ATUS data and then download a subset of the original data. It is possible to select time use variables as well and have the system automatically determine average hours per day for these activities, but doing so produces a rectangular extract which removes the time of day information that we need for visualization. We instead opted for a hierarchal extract (shown below). While more difficult to process, it provides much richer information on individual activities.



In cart	Variable	Variable Label	Type	03	04	05	06	07	08	09	10	11	12	13	14	15
☑	CASEID	ATUS Case ID [preselected]	H	X	X	X	X	X	X	X	X	X	X	X	X	X
☑	YEAR	Survey year [preselected]	H	X	X	X	X	X	X	X	X	X	X	X	X	X
☑	PERNUM	Person number (general) [preselected]	P	X	X	X	X	X	X	X	X	X	X	X	X	X
☑	LINENO	Person line number [preselected]	P	X	X	X	X	X	X	X	X	X	X	X	X	X
☑	WT06	Person weight, 2006 methodology [preselected]	P	X	X	X	X	X	X	X	X	X	X	X	X	X
☑	STATEFIP	FIPS State Code	H	X	X	X	X	X	X	X	X	X	X	X	X	X
☑	COUNTY	FIPS County code	H	X	X	X	X	X	X	X	X	X	X	X	X	X
☑	FAMINCOME	Family income	H	X	X	X	X	X	X	X	X	X	X	X	X	X
☑	MONTH	Month of ATUS interview	P	X	X	X	X	X	X	X	X	X	X	X	X	X
☑	DAY	ATUS interview day of the week	P	X	X	X	X	X	X	X	X	X	X	X	X	X
☑	DATE	Date of ATUS interview	P	X	X	X	X	X	X	X	X	X	X	X	X	X
☑	AGE	Age	P	X	X	X	X	X	X	X	X	X	X	X	X	X
☑	SEX	Sex	P	X	X	X	X	X	X	X	X	X	X	X	X	X
☑	RACE	Race	P	X	X	X	X	X	X	X	X	X	X	X	X	X
☑	HISPAN	Hispanic origin	P	X	X	X	X	X	X	X	X	X	X	X	X	X
☑	EMPSTAT	Labor force status	P	X	X	X	X	X	X	X	X	X	X	X	X	X
☑	ACTIVITY	Activity	A	X	X	X	X	X	X	X	X	X	X	X	X	X
☑	START	Activity start time	A	X	X	X	X	X	X	X	X	X	X	X	X	X
☑	STOP	Activity stop time	A	X	X	X	X	X	X	X	X	X	X	X	X	X

The raw data is a plain text file that looks like this:

```
120030100013280020030606037013
220030100013280020030100100106020030103815546 2.67215800006001011001000 2
320030100013280020031301240 4:00:0005:00:00
42003010001328002003
320030100013280020030102010 5:00:0005:30:00
42003010001328002003
320030100013280020030101010 5:30:0015:30:00
42003010001328002003
```

The first character encodes the record type, and subsequent character sequences encode a subset of the attributes shown above. We wrote a Python script that, given a configuration JSON file and the raw data file, creates four CSV files with aggregated activities and demographics. One CSV file is produced for each timescale of interest (day, week, year, and lifetime). Both the demographics and the activities of interest are set in the configuration file.

```
Minute,all:work,all:calls,all:social_relax_leisure,all:household_activities
…
0,0.01629,0.00029,0.01119,0.00498,…
1,0.01627,0.00029,0.01124,0.00500,…
2,0.01628,0.00028,0.01131,0.00489,…
```

These CSV files are suitable for final preprocessing in d3 before being displayed in the in visualization. This preprocessing extracts the demographic and activity name from each column, converts the minute/day of week/day of year/age to an actual date, and computes stacks to plot a stacked area chart.

Exploratory data analysis

What visualizations did you use to initially look at your data? What insights did you gain? How did these insights inform your design?

Because we were planning to create an exploratory analysis tool, we did less exploratory analysis in our initial planning that other projects may have. Instead we focused primarily on exploring how others had visualized the data (see related work). ATUS is a very rich data set, so we had to figure out which aspects in particular we wanted to highlight to avoid overwhelming the user.

None of the visualizations we found provided a lifetime view of the ATUS data, which we felt could provide useful insight into how people's habits change over time. To validate the feasibility of such a view, we used ATUS-X to quickly create an extract of respondent age versus time spent on education, leisure, and work. We used Excel to create a line graph of this data and found that clear, interesting trends were visible.

We also tried developing tools to find the medoid, or most central, respondent in a given demographic, with the thought that these values would represent a “typical” day. However, computing the medoid is an N^2 problem, so computing it for the entire data set is infeasible.

Finally, we looked at the distribution of activities to guide our choice of how to bin activities. Sleep, for example, is counted by the BLS under the category of personal care, but it is common enough that we felt it warranted its own bin. Other activities, like government services and calls, are so infrequent that they can be combined into a miscellaneous bin. Top level BLS categories were used for initial prototype design, but these custom designed bins were used in the final visualization.

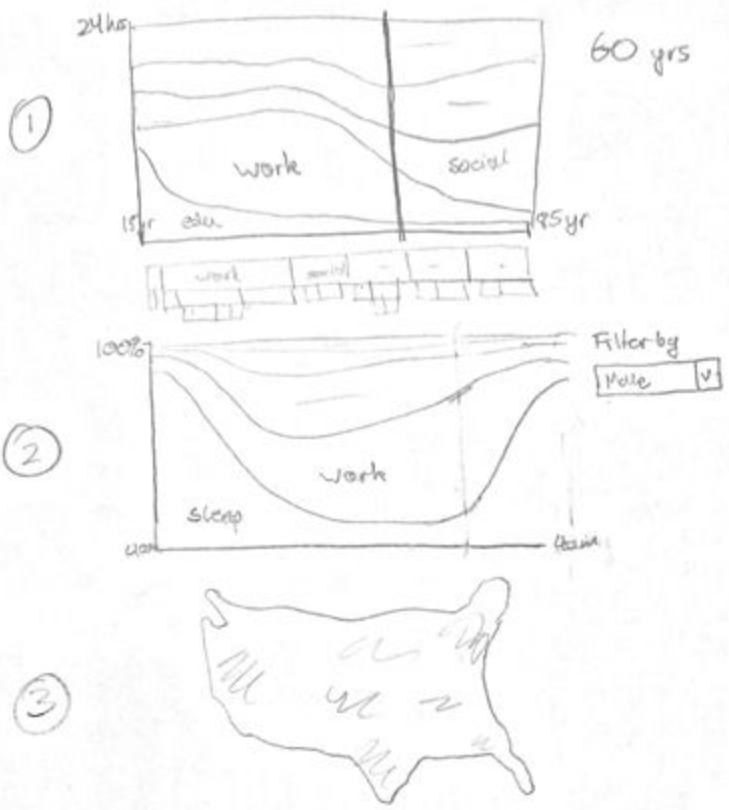
Design evolution

What are the different visualizations you considered? Justify the design decisions you made using the perceptual and design principles you learned in the course. Did you deviate from your proposal?

We approached the design process by first brainstorming ideas individually before coming together to discuss our different approaches. This allowed us to have a broader basis on which to implement our final design. We had various different ways to visualize comparative data, through bar charts, line graphs and choropleth, as well as time disturbed data through stacked area charts, scatter plots and stacked bar charts. Additionally, we came up with some ways to represent hierarchical data through sunburst charts and treemaps, and some ways to summarize the entire data set through matrixes and common activity timeline bars.

The three designs we came up with to represent the dataset can best be summarized as an age focused visualization, customization focused visualization, and a comparative focused visualization. The age focused visualization utilized a stack area chart with age as an x-axis from 15 to 85 years of age, and had the option to show a sunburst chart as well as another stacked area chart based on a 24-hour period for any age you hover over. The 24-hour period stacked area chart uses percentage of respondents doing a given activity for the areas, and the 24-hour time scale for the x-axis. Additionally, by clicking an activity from the 24-hour period stacked area chart, you would bring up a choropleth of the united states, that shows the distribution of respondents doing the selected activity based by state location. This design does a good job of conveying the continuous nature of age, and has the opportunity to show the full hierarchy of activities, however it is not very good for comparing across various classes, and only allows comparisons via stacked visualizations.

Layout

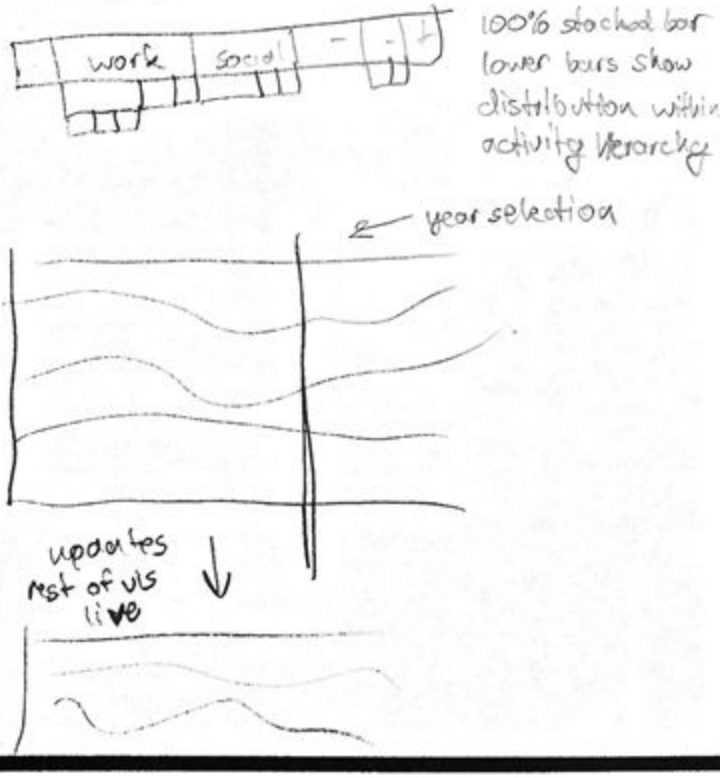


Title:
Auther: Ben McMorran
Date: 10/31/16
Sheet: 2
Task:

Operations

Slide across (1) to select age. Hierarchy distribution updates, Dropdown in (2) to filter day distribution to certain class. Click activity in (2) to show distribution in (3). Slide across (2) to show certain hour.

Focus

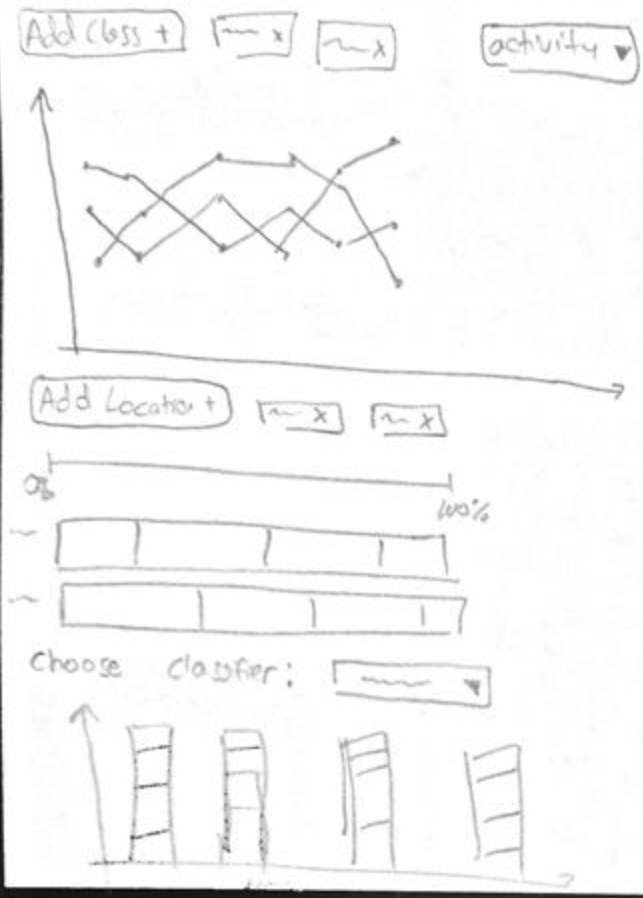


Discussion

- hard to compare across classes
+ conveys continuous nature of age
+ full hierarchy of activities
- all comparisons between activities are stacked. Unaligned worse than aligned.

The customization focused visualization utilized various drop down and additive buttons to select various different configurations for the graphs to represent the data. This visualization utilized line graphs to show how time affects various activities, stacked horizontal and vertical bar charts to show how classes affect the percentage of time spent doing each activity. The benefit of this visualization is that it allows a great deal of user interaction and customization so the user can view whatever they feel is most important, and given the compact design can show all visualizations without having to scroll down. The drawbacks of this visualization is that certain insights may be lost because it relies almost entirely on user interaction. Dropdowns have the potential to obscure information.

Layout



Title:

Auther: Francisco Sanchez

Date: 10/31/15

Sheet: 3

Task:

Operations

Add Class + Button allows uses to select a class to add to the scatterplot

Activity allows user to select activity for comparison.

Add Location + similar button to add class

classifier allows user to track a certain subgroup through the years

Focus

Each set of visualizers features an interactive set of classifiers to compare classes and characteristics to one another.

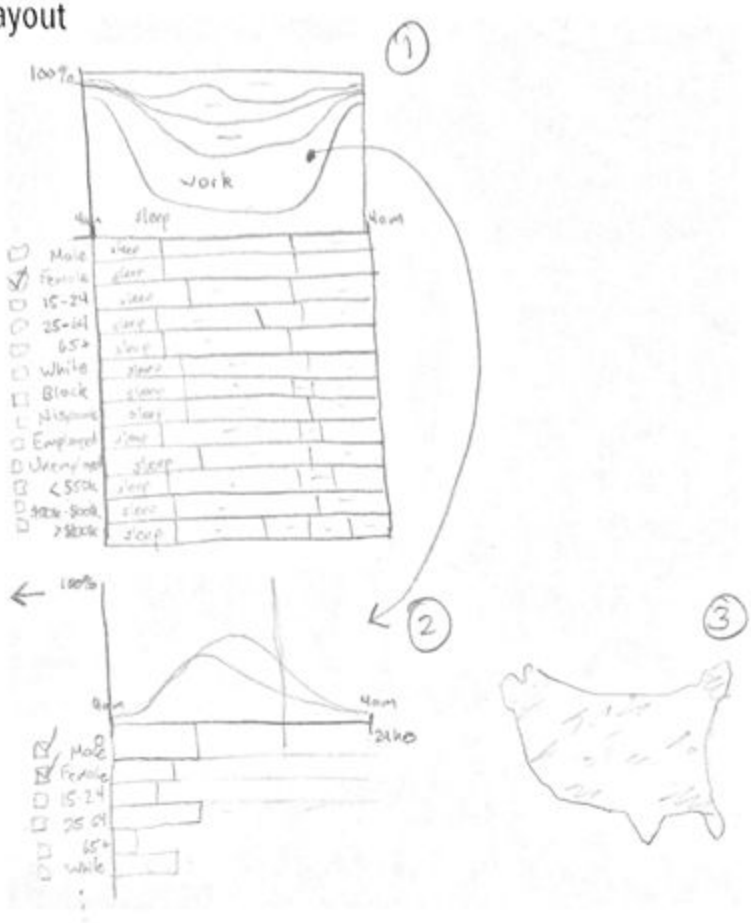
Discussion

+ allows deep user customization
+ Can show many comparison on a single page
- some insights may be lost.

The comparative focused visualization utilized a combination of a stacked area chart and timeline bars to show what percentage of respondents are doing each activity at each hour of the day, as well as breaking that up into each of the different classes as well as buckets of age groups. Additionally, by clicking an activity on the stack area chart, this visualization brings up a line graph of the distribution of respondents doing that activity throughout an average day.

Underneath this line graph is a bar chart that shows how many hours on average each class spends doing that activity in a 24-hour period, with the ability to select any class to add it to line graph. Finally, there is a vertical hover cursor that displays the time distribution of each state on a US map. This visualization is good to make comparisons between classes because they are on a shared axis, and it reduces area for the visuals by using a stacked area chart and as a result is fairly compact. Some cons with this visualization is that common activities may be very similar across classes, the map is only for display, and that the age is a ratio but is only represented in discrete bins.

Layout



Title:

Auther: Ben McMorran

Date: Oct 31 2016

Sheet: 1

Task:

Operations

Click on activity in (1), animate to selected activity

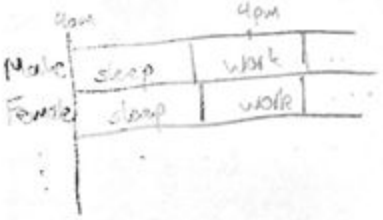
(2). Filter in (1) w/ checkboxes

Compare in (2) w/ checkboxes

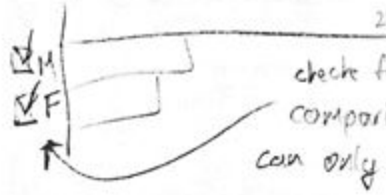
Slide cursor over (2) to animate (3) (if only 1 or zero comparisons selected)

Back arrow to return to (1)

Focus



shows most common activity at time of day for class



shows # of hrs/day on activity

check for comparison

can only compare within type (age, sex, e

shaded for avg day normally if 1 comparison checked then shading is filtered

shows hour of day as cursor slides across (2)

Discussion

- most common activity may be similar enough that bars under (1) aren't interesting
+ comparison between classes on shared axis
- map only for display, not interactive
+ stacked area reduces to single area for better readability
- age is ratio, but represented in discrete (large) bins
+ compact comparison of all classes

Our final design sketch attempts to incorporate the best elements of all three earlier designs. The main view uses a stacked area bar chart over a 24-hour period (by default) to quickly convey the distribution of activities over the course of a day. Unlike earlier designs, the time scale of this graph can be changed using a slider at the top to cover a week, year, or lifetime. Time has a deeply hierarchical nature, and adding this slider lets the user easily explore time at all these scales. Although the stacked nature of a stacked area graph means that activities cannot be compared along a common baseline, we felt that this tradeoff was acceptable given how well the stacked area bar chart conveys the part-to-whole relationship. The number of categories should be small enough (< 15) that easily distinguishable colors can be used for each activity type and activities will not be lost.

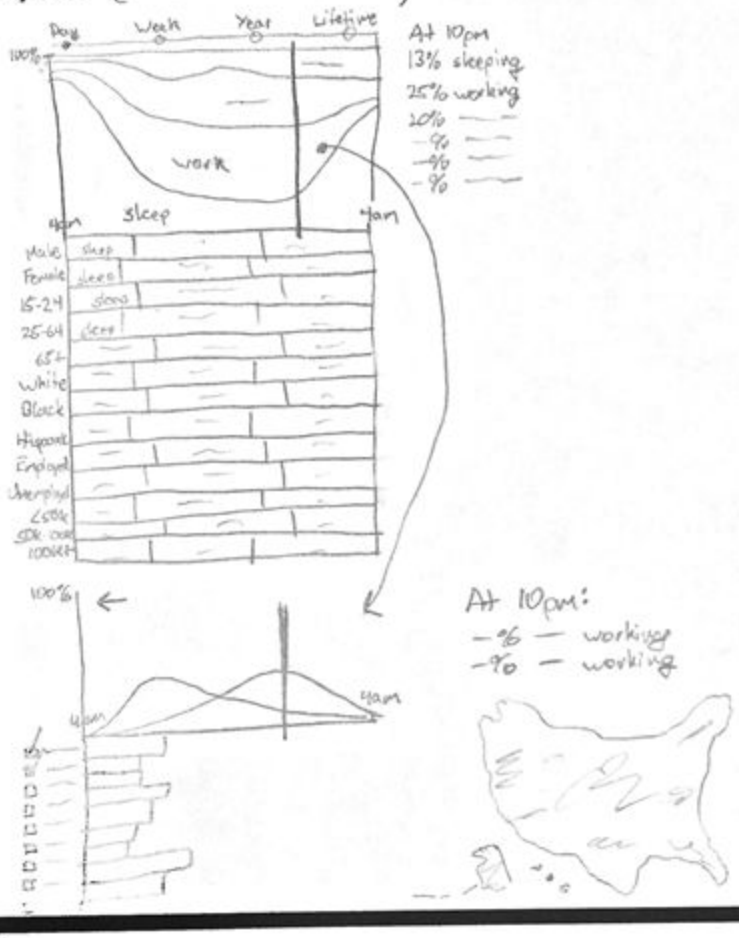
Below the stacked area graph is a list of demographics by age, sex, race, employment status, and income. Next to each item is a quick visual summary of the most common activity of the given demographic at a given time of day, using the same horizontal axis as the stacked area bar chart above. Clicking on one of these bars will filter to stacked area bar chart to show only results from that demographic. Clicking again will remove the filter. The horizontal summaries for each demographic, while certainly not comprehensive, provide a way to quickly compare all classes without extra interaction. More detailed comparison is available at the individual activity level.

Hovering over the stacked area graph will display a vertical cursor at a certain time of day. A sidebar displays an exact percentage breakdown for activity distribution at that time of day. If the user is not hovering over the chart, the sidebar simply displays averages for the time period.

Clicking on an area in the stacked area graph will filter the graph to display only that activity, turning the stacked areas into a simple line graph. This facilitates better comparison of activity distribution over the time period by using a common baseline. Additionally, the summary bars for each demographic at the bottom are replaced by a horizontal bar whose length represents the number of hours per day that the demographic spends on the selected activity. Again this design facilitates direct comparison between demographics in the most perceptually accurate way by using a length encoding on a common baseline.

If additional demographics are selected a new line is added to the line graph specifically for that demographic. The color of the line matches the color of the horizontal bar, creating an implicit legend. This allows the user to compare certain demographics for a given activity across the time period on a common baseline.

Layout (similar to sheet 1)



Title:

Auther: Ben & Francisco

Date: 10/31/16

Sheet: 4

Task:

Operations

Hover over main chart to get % breakdown @ point in time.

Change timescale at top.

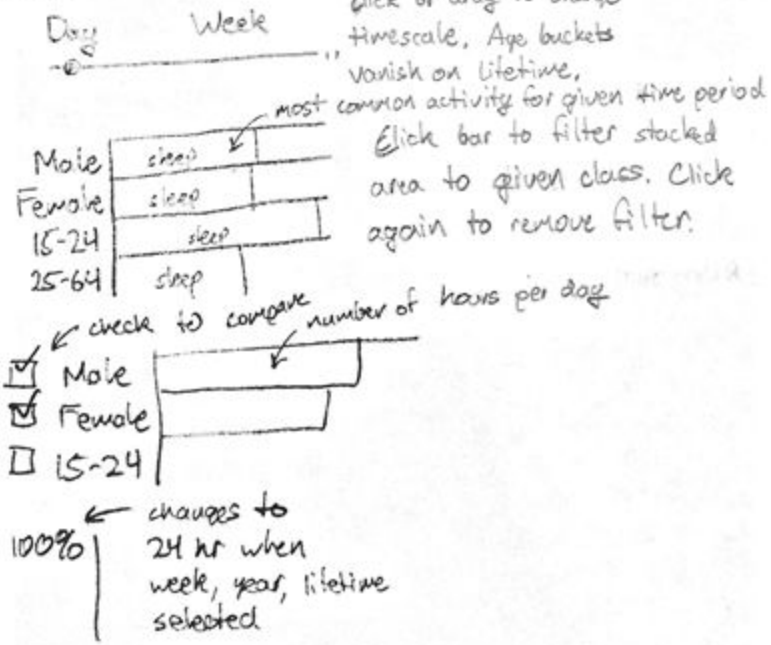
Click bar to filter by class.

Click activity to remove all other activities

Check classes to compare

Click over single activity to update % breakdown and map

Focus



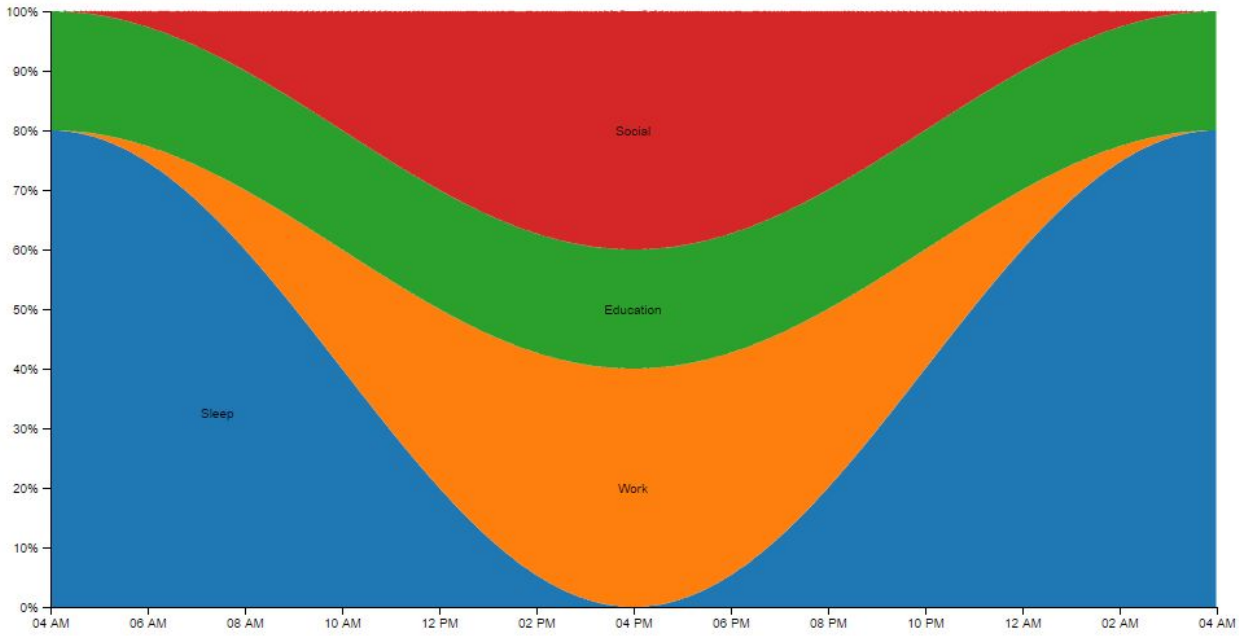
Detail

Class-level summaries may look silly at anything other than Day. Check w/ real data.

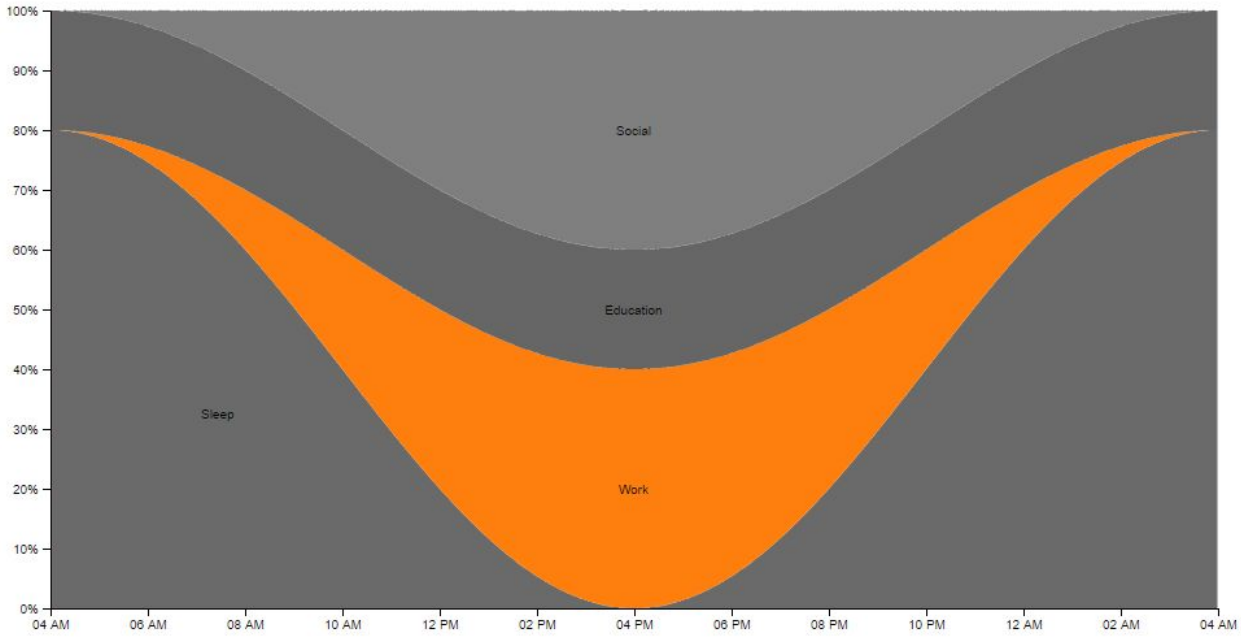
Need real-time interactive updates for slides to be effective

Preprocessing to reduce respondent level data to summaries

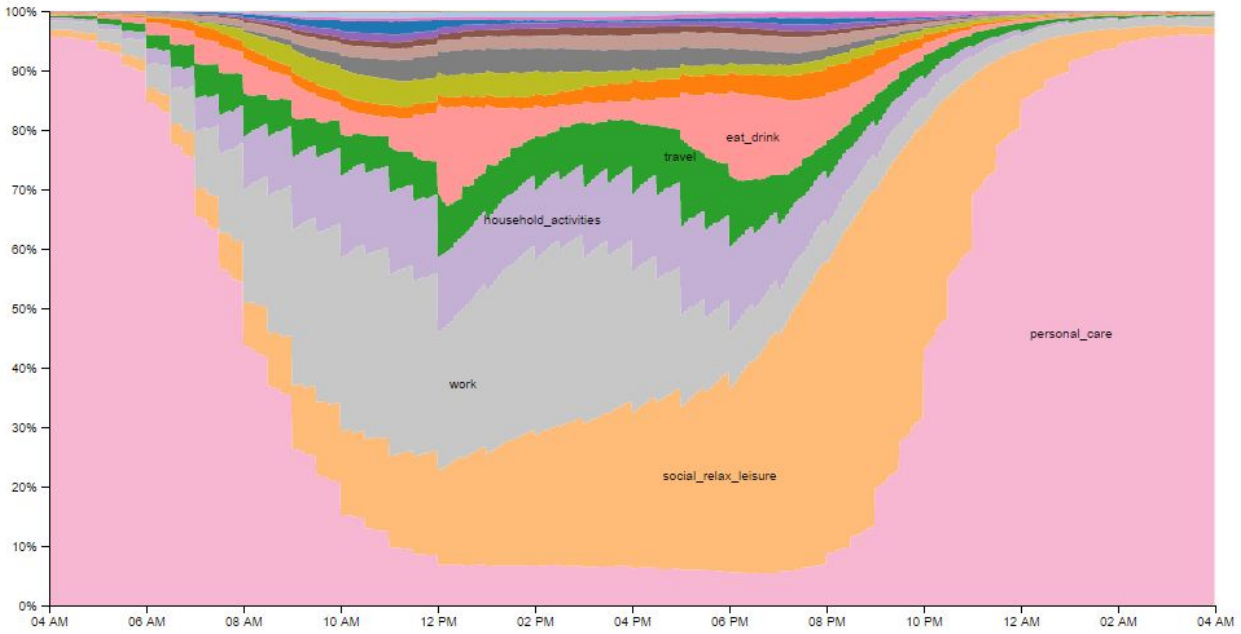
As seen above, most of our design ideas centered around a stacked area chart to provide a high level overview of the data, so we first developed a stacked area chart using synthetic data to ensure we understood how to build one in d3. We also experimented with approaches for automatically placing labels on the graph. In this screenshot labels are placed by finding the centroid of the largest area.



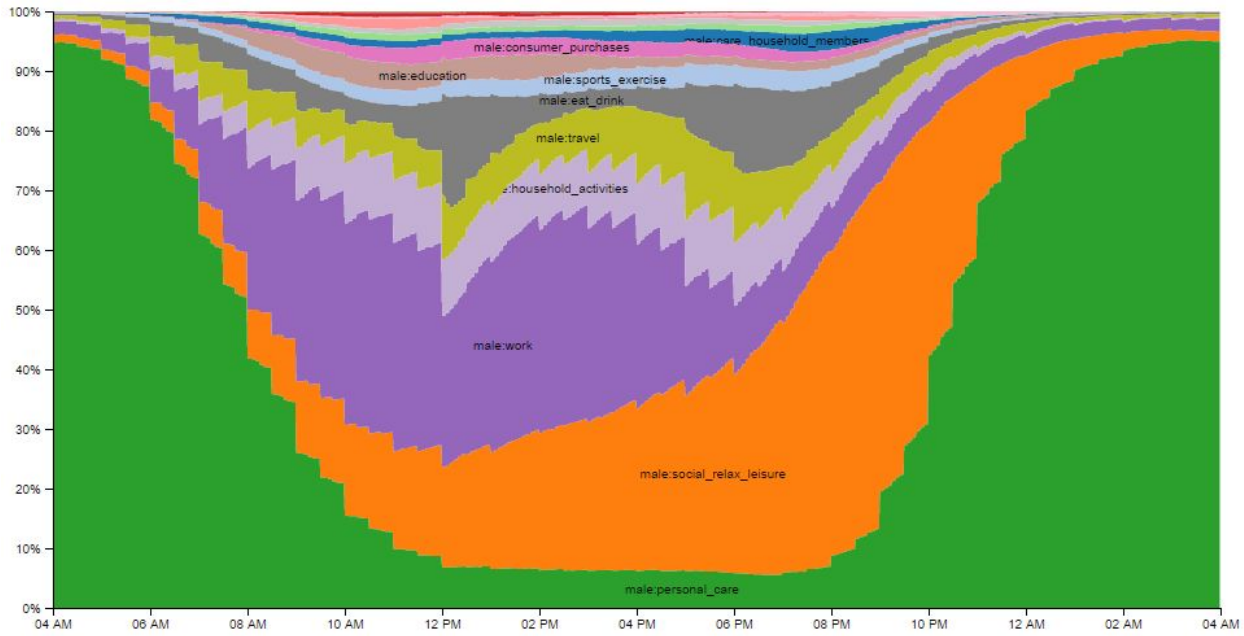
We also started experimenting with interaction techniques. This screenshot demonstrates desaturating the colors for every other activity when hovering over one activity. This technique makes use of preattentive visual pop out to make tracking the activity of interest easier.



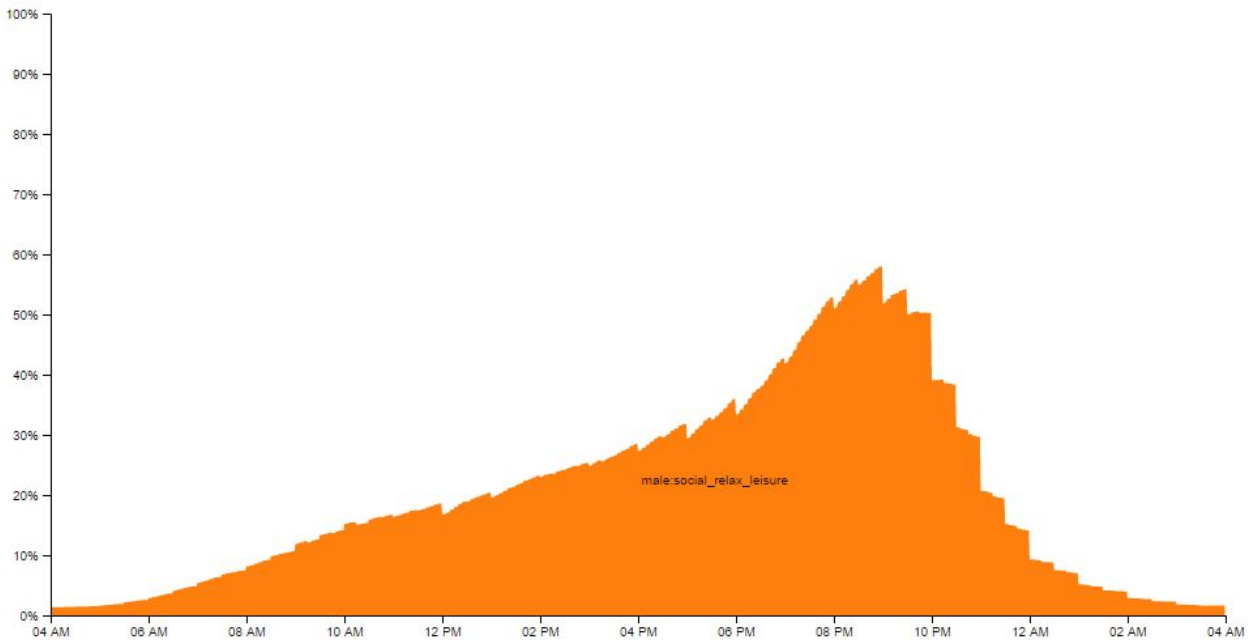
Confident in our ability to create the chart, we next tried incorporating real data into the visualization. This was a good test of our entire data processing pipeline. Already some interesting insights begin to emerge, like how people tend to get off work on the hour. There are also some problems with label placement beginning to show up. Finally, there are many thin, infrequent activities at the top of the visualization. Changing the binning of activity codes and providing alternative ways of visualizing these infrequent activities may help.



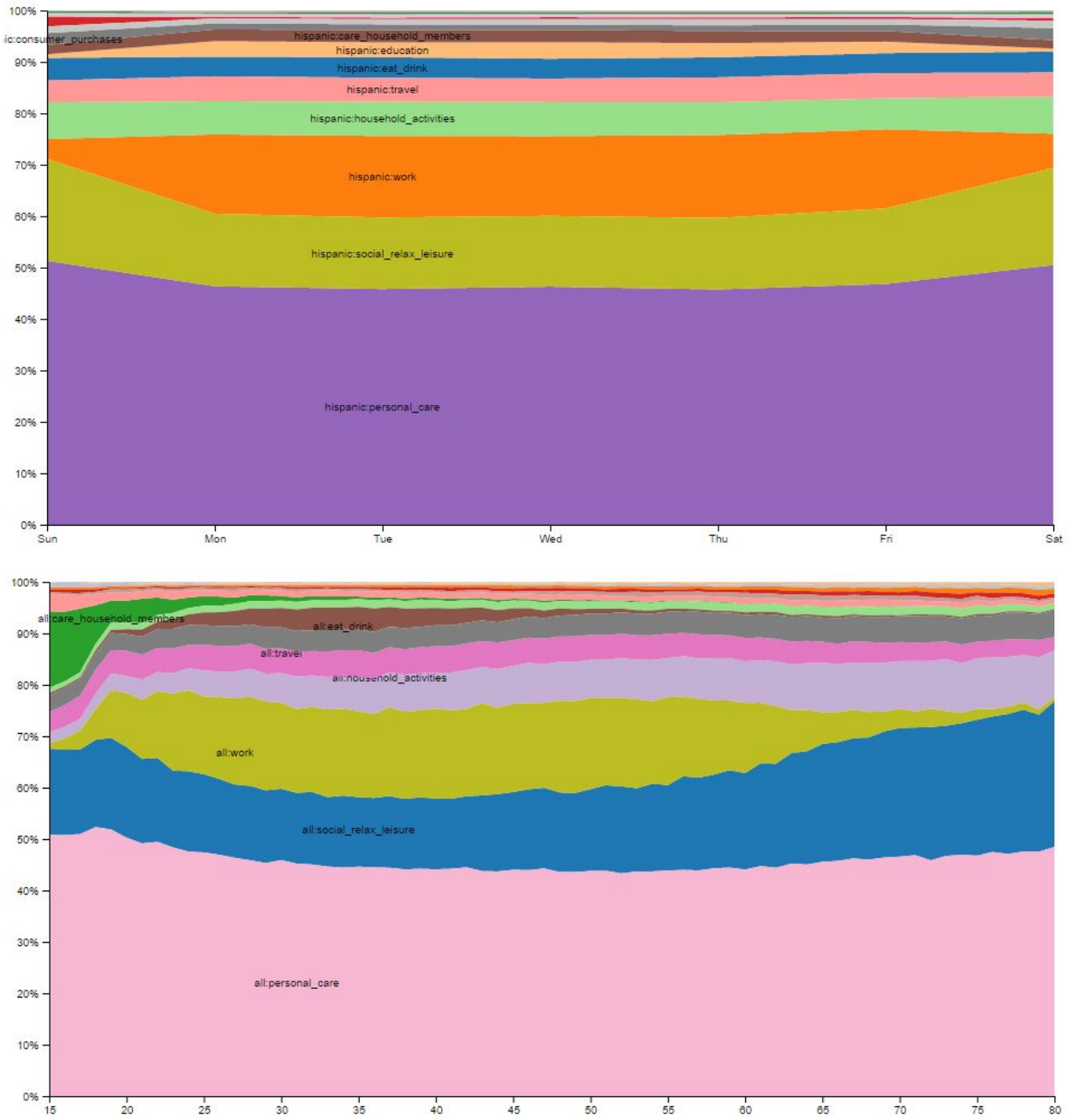
We also tried successfully visualizing the daily activity breakdown for a specific demographic (in this case male). Notice the slightly increased amount of work and decreased amount of household activities when compared to the average.



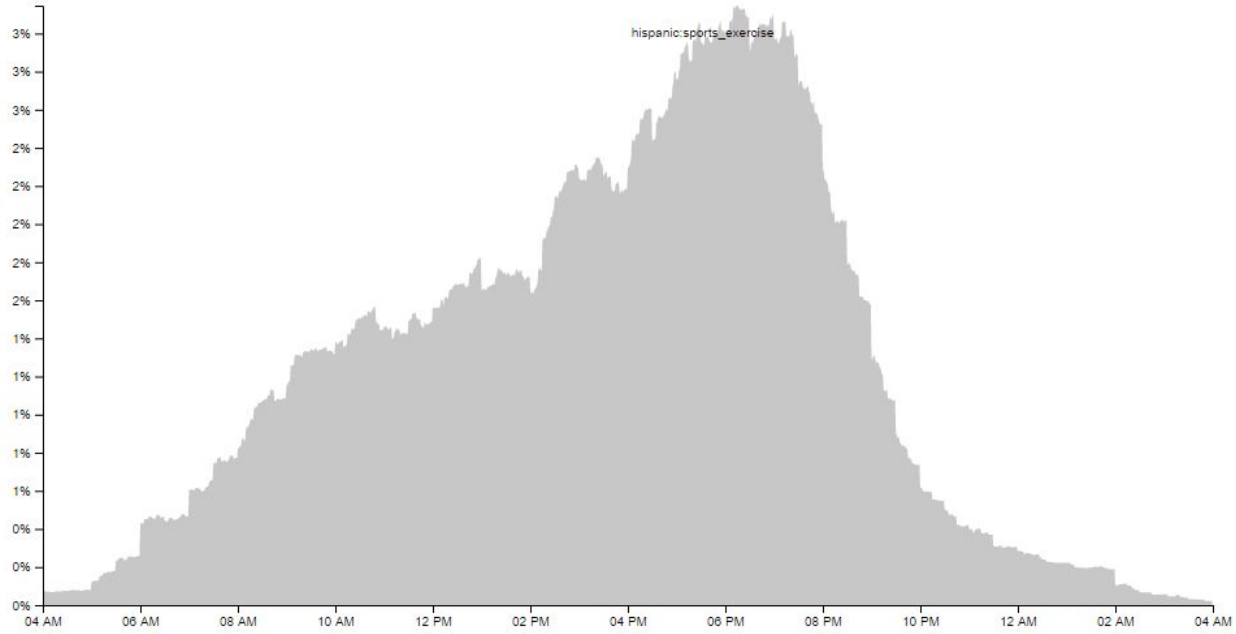
With a high-level summary visualization complete, we next turned to the visualization of individual activities. Stacked area charts do a good job of compactly conveying relative distributions over time, and they use a perceptually accurate length encoding. However, the visualization is less well suited for visualizing an individual activity over time because each time slice does not share a common baseline. We developed an animated transition that fades out all other activities, then slides the selected activity down to a common baseline when it is clicked. This new visualization provides better perceptual accuracy and the transition helps the user maintain context while the visualization is transformed. Here you can easily see how leisure activities increase in the evening as people come home from work.



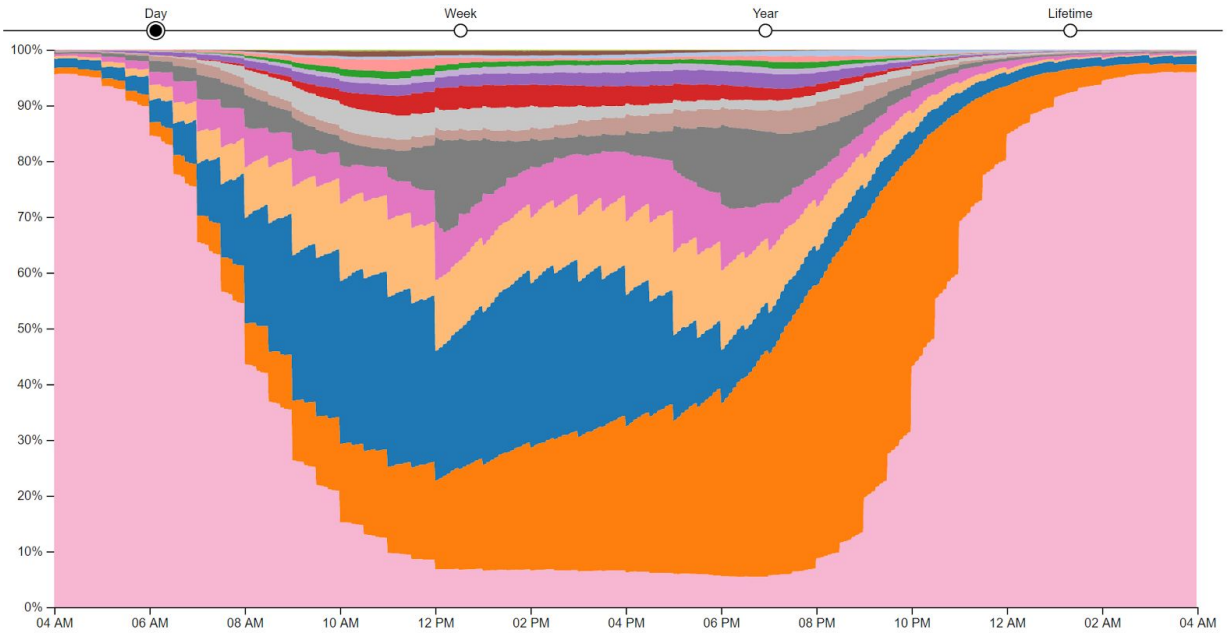
We also started developing the code necessary to support additional timescales. Week views and lifetime views are shown below.



One thing we noticed while developing the transitions between the individual activity view and complete view was that the resulting aligned baseline chart was often nearly unreadable for infrequent activities, like sports and exercise. To fix this, we adjusted the transition to align the baseline, then rescale the y axis to put the peak of the selected activity at the top of the chart. This two-step transition helps the user maintain context throughout the entire transition while improving the readability of the final chart.

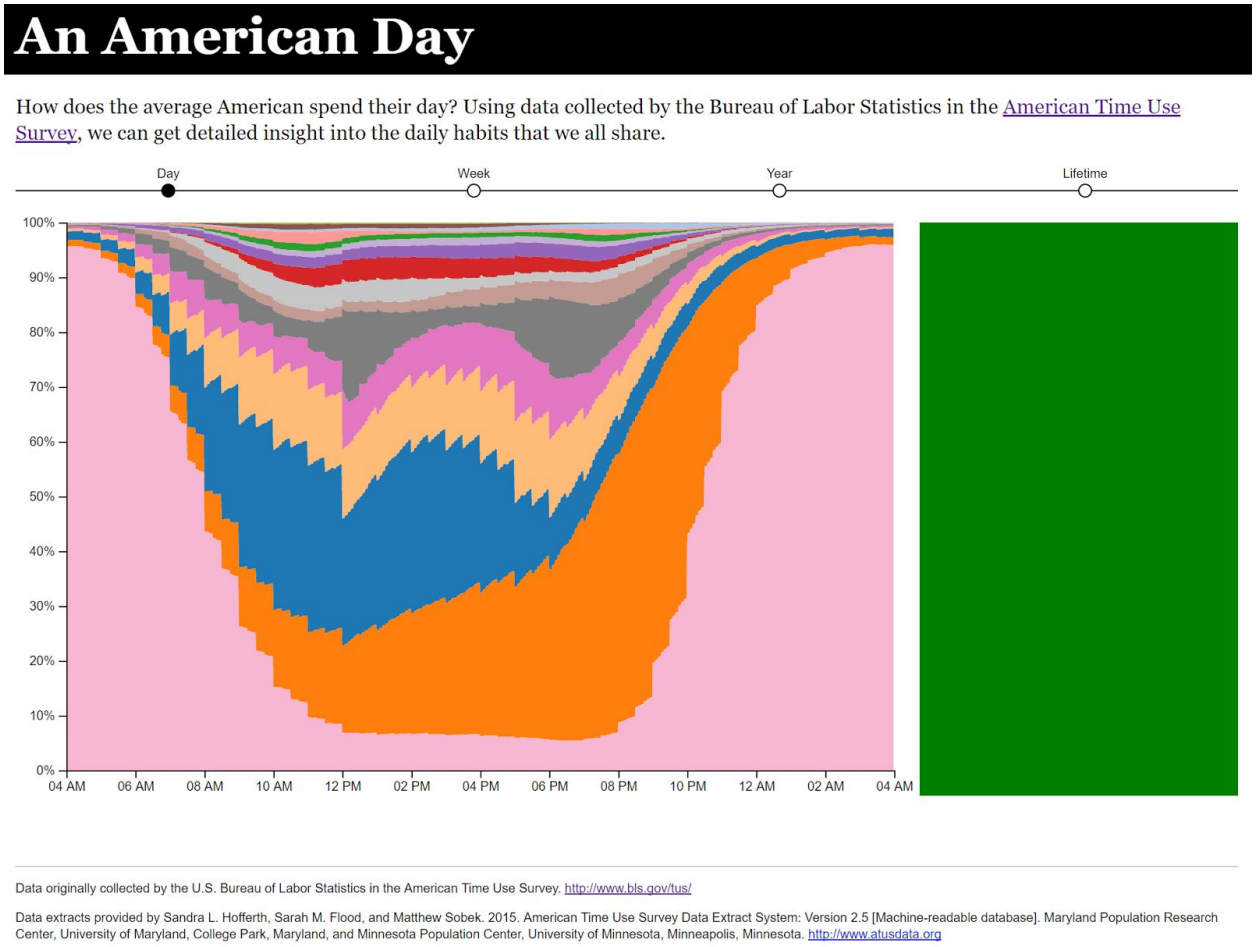


At this point we felt confident in the base view of our visualization, and started adding additional views to provide additional insight and context. The first thing we did was add a slider at the top of the visualization to select the timescale of interest. Transitions between timescales are animated. Placing this element at the top is appropriate because the timescale provides the context for every other element of the visualization.



Next we began drafting how the final website layout would look. The screenshot below shows a header, some draft introductory text, the timescale slider, stacked area chart, and a blocked out

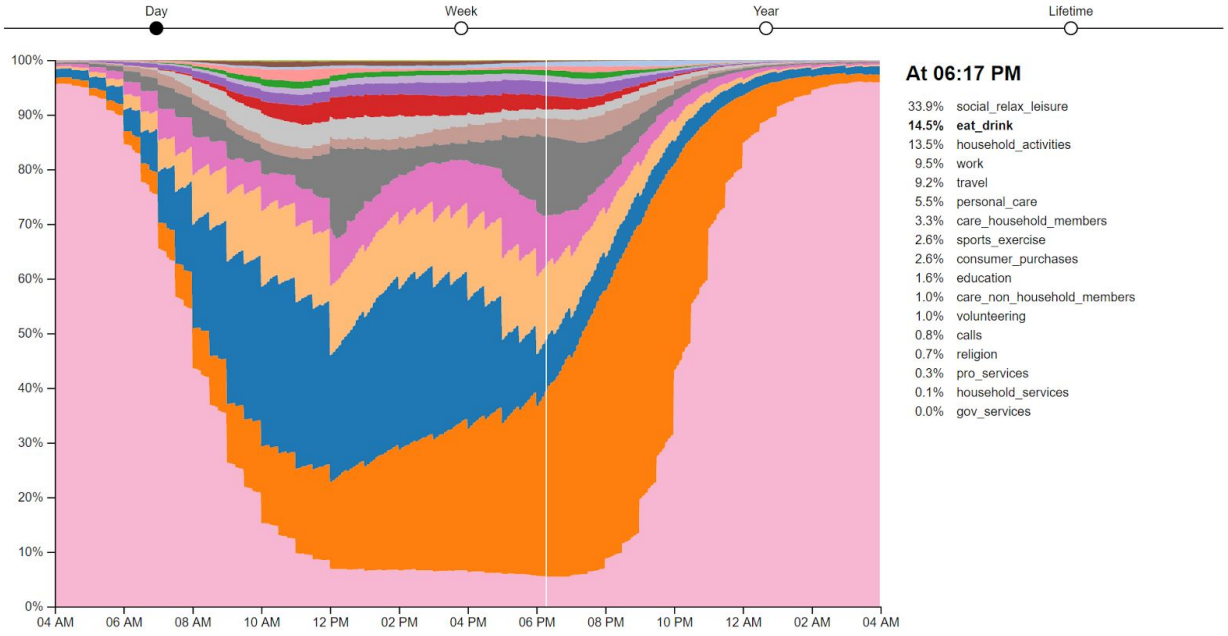
area for displaying detailed information when the user hovers over the area chart. Area labels are absent in this view because we are still determining the best way to place them.



The sidebar view shows exact percentage breakdowns for each activity either on average if the user is not hovering over the stacked area chart, or at the specific instant they are hovering over. A thin vertical cursor helps the user understand which moment in time they are currently selecting. The list of activities in the sidebar is automatically sorted in order from most frequent to least frequent. The interaction technique is designed to only provide detailed information when the user requests it to avoid overwhelming them. The activity that the user is hovering over is displayed in bold in the sidebar. This provides some context, but direct activity labeling is still needed.

An American Day

How does the average American spend their day? Using data collected by the Bureau of Labor Statistics in the American Time Use Survey, we can get detailed insight into the daily habits that we all share.



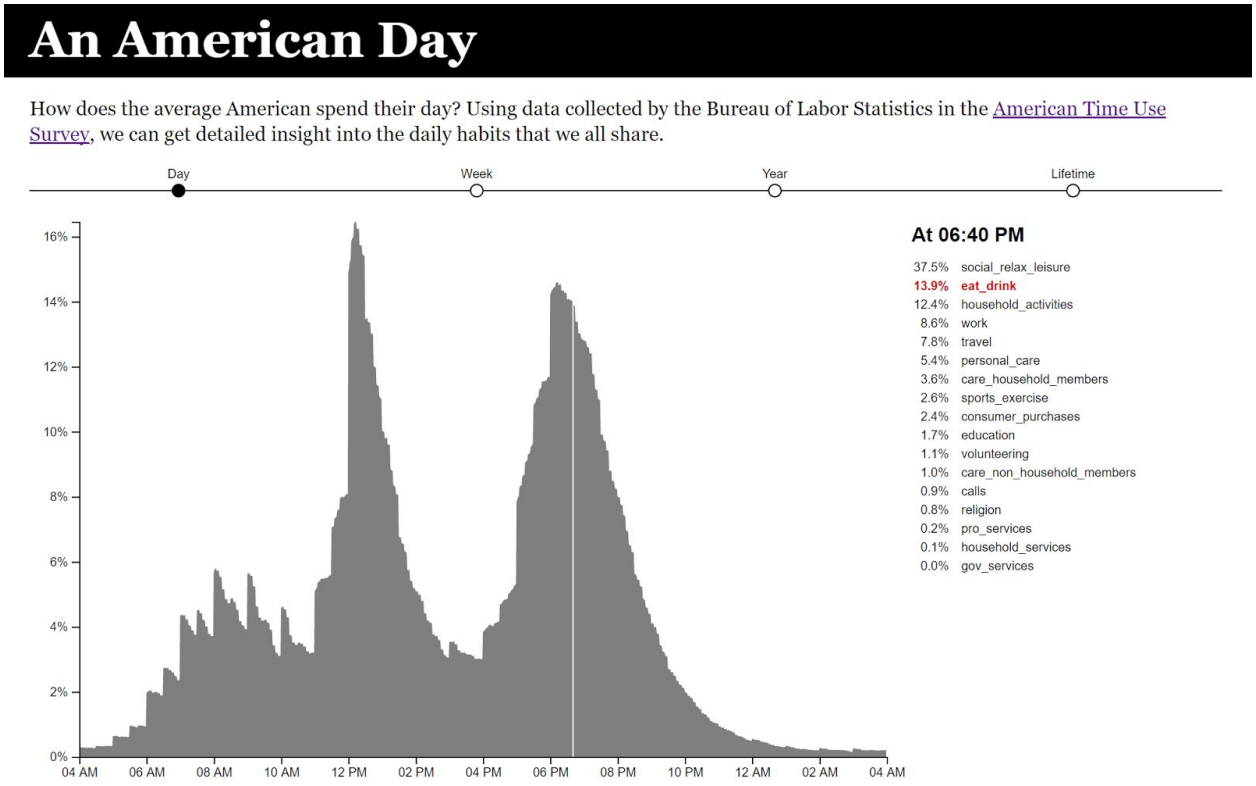
Data visualization by Ben McMorran and Francisco Sanchez.

Data originally collected by the U.S. Bureau of Labor Statistics in the American Time Use Survey. http://www.bls.gov/tus/

Data extracts provided by Sandra L. Hofferth, Sarah M. Flood, and Matthew Sobek. 2015. American Time Use Survey Data Extract System: Version 2.5 [Machine-readable database]. Maryland Population Research Center, University of Maryland, College Park, Maryland, and Minnesota Population Center, University of Minnesota, Minneapolis, Minnesota. http://www.atusdata.org

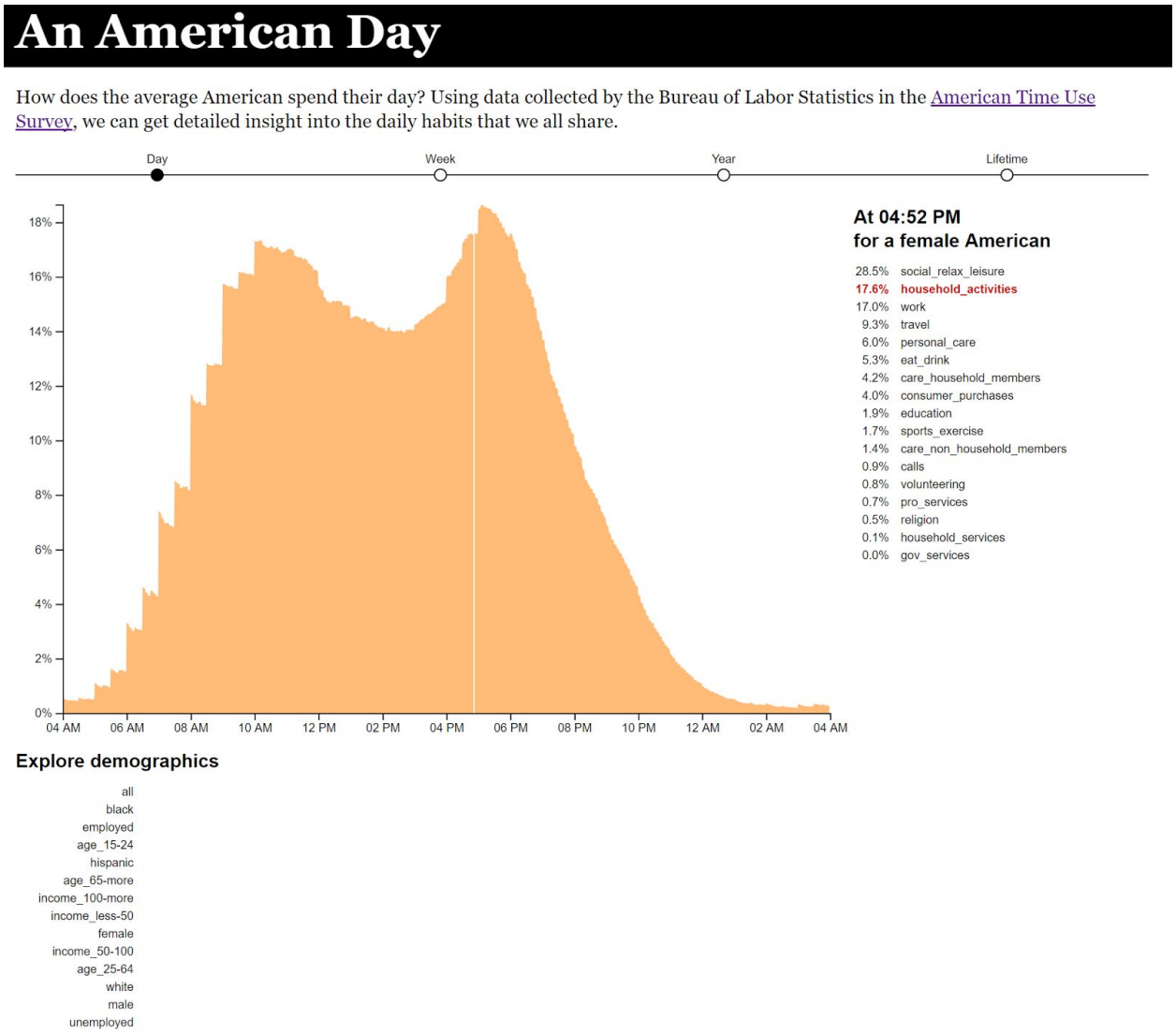
Favicon made by Freepik from www.flaticon.com. Used under CC 3.0 BY.

The sidebar displays the same information when an individual activity is selected. This helps provide global context even when viewing detailed information.



We next began building demographic information into the visualization. The list of demographics in the lower section of the visualization are clickable, and this action will trigger an animated transition in the area chart from the current demographic to the selected demographic. Animating this transition is important because it highlights the relative differences between demographics. Initially this animation both rescaled the y axis and animated the shape of the area in the same transition, but animation does a poor job of showing absolute differences in activity prevalence between demographics. Instead we settled on a two-step transition that first animates the shape on a fixed axis, then rescales the y axis to normalize the graph again.

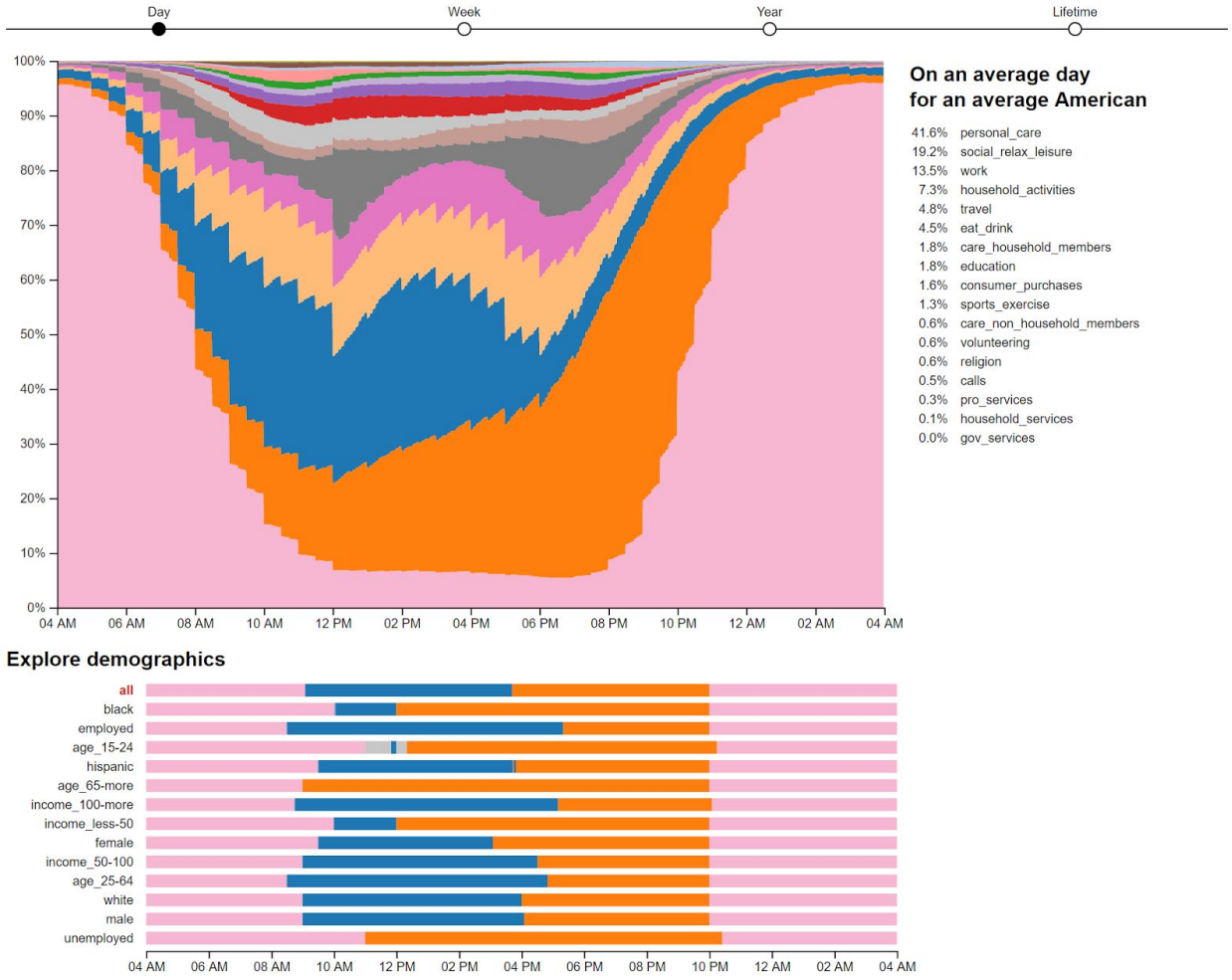
The sidebar text also updates to remind the user of the demographic they are currently exploring, in this case for females.



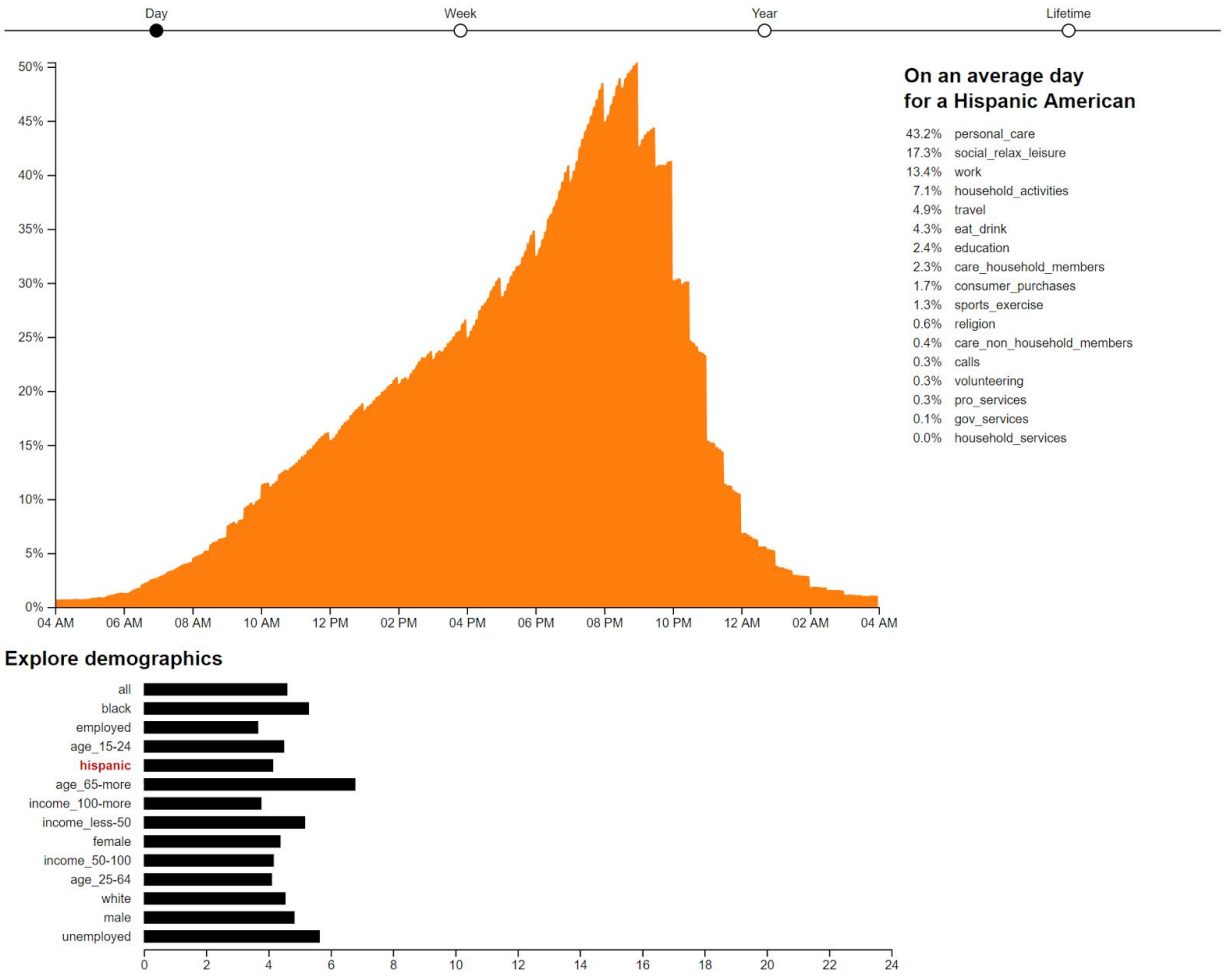
The screenshot below shows a first attempt at providing a summary visualization for each demographic in the demographics section. This visualizations shows the most common activity at each time of day for each demographic. While visually interesting, we are concerned that this visualization is potentially misleading because no one person in the dataset has a day like this. For example, it is unlikely that the average black American that works works only 2 hours per day, but it would be easy to draw such a conclusion from the visualization. Additionally, labels are needed to avoid depending entirely on color for activity encoding.

An American Day

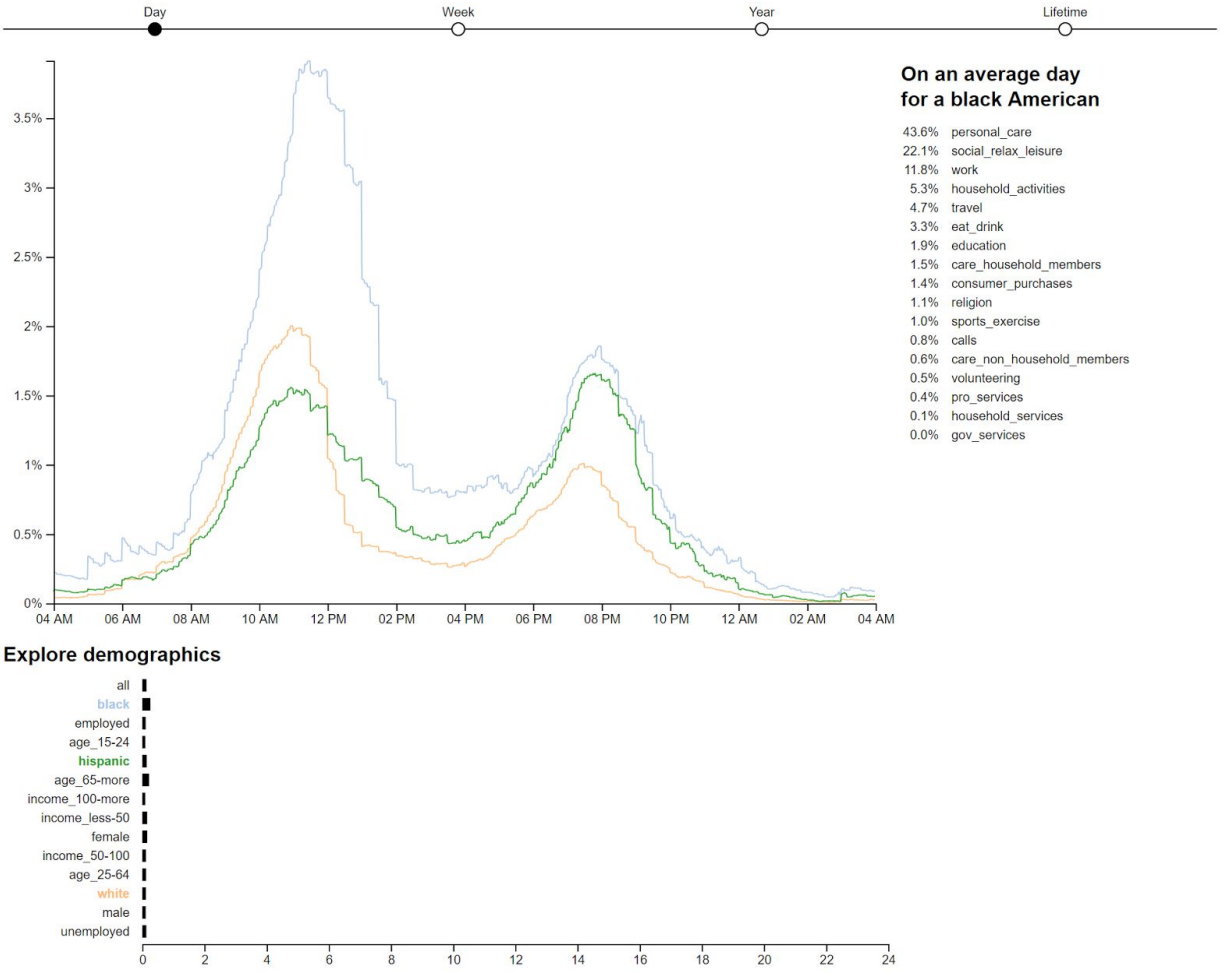
How does the average American spend their day? Using data collected by the Bureau of Labor Statistics in the American Time Use Survey, we can get detailed insight into the daily habits that we all share.



The demographic view also adjusts to display average hours per day in each demographic when an individual activity is selected. This visualization provides context for the selected activity can suggest other demographics for the user to explore. In this example, we can see that older people tend to do leisure for significantly more hours per day than other demographics.

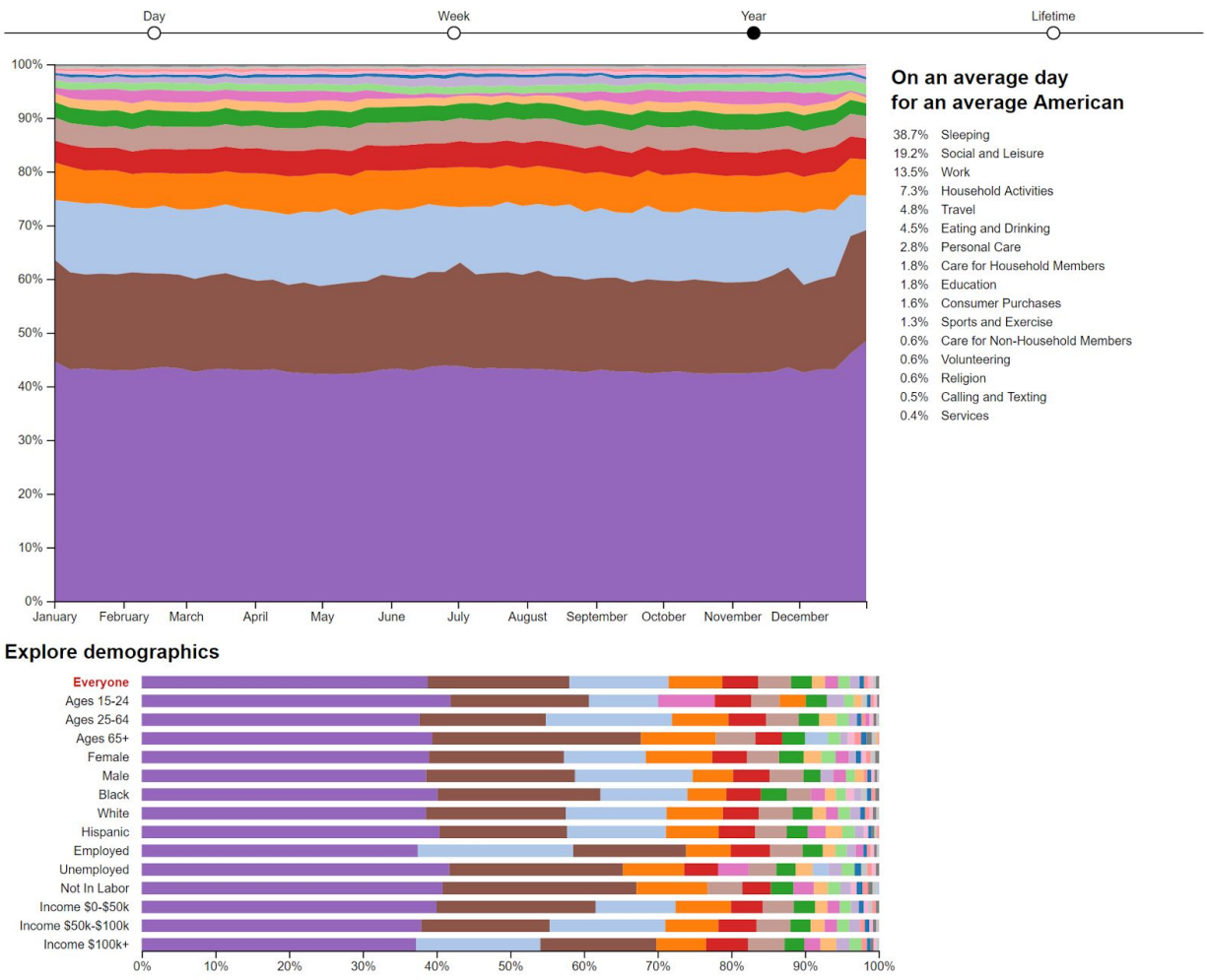


The visualization techniques developed so far are effective at presenting information on individual activities and demographics, but they make it difficult to compare activity prevalence across demographics. There are often interesting insights to be found in these comparisons, like the difference in housework between genders or differences in sleep patterns across age groups. To support these comparisons, we added the ability to select multiple demographics when an individual activity is selected and show each demographic as a line in the chart. The color of the line matches the color of the demographic below.



Here we are visualizing differences in religion between black, white, and Hispanic Americans. Notice how black Americans participate in religious activities more in general, but Hispanic Americans have relatively more activity in the evening. This may be due to a strong Protestant influence in black communities and a Catholic influence in Hispanic communities.

Two final changes we made to the visualization were averaging the year view by week instead of day, and changing the demographic view to simply show average percent of time in each activity. Averaging by week helps to smooth the significant noise we were observing in the daily data, and also smooths out the BLS reporting gaps on July 4 and Christmas Day. The new demographic view provides a more consistent summary summary of each demographic with less potential to mislead the user.



Implementation

Describe the intent and functionality of the interactive visualizations you implemented. Provide clear and well-referenced images showing the key design and interaction elements.

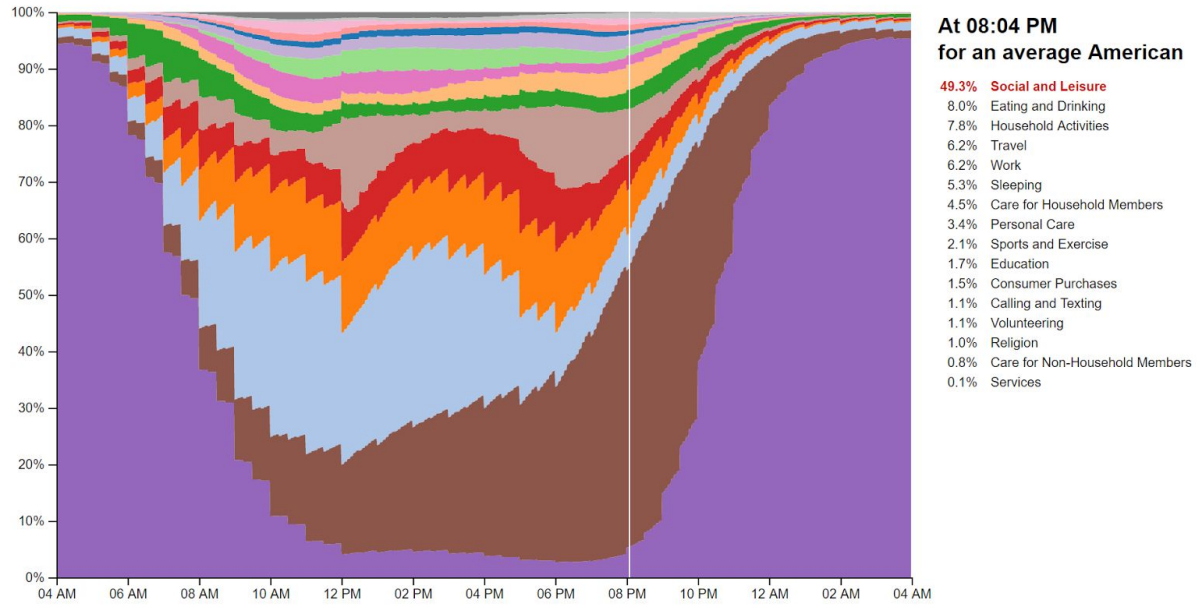
Our design supports five main interactions:

1. Change timescale
2. Hover for details
3. Select single activity
4. Compare demographics for single activity
5. Change demographic for all activities

Change timescale

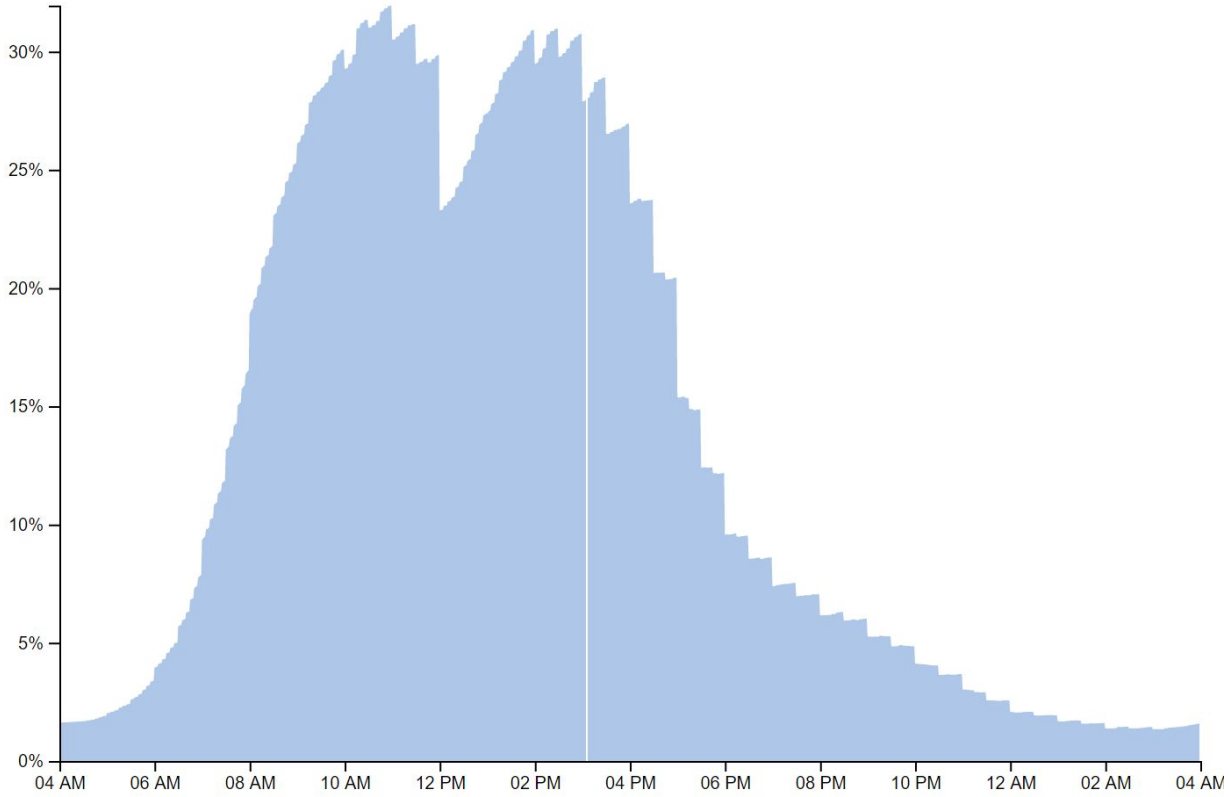
The slider at the top of the visualization provides an easy way for the user to change between different timescales. Subtle animated transitions provide feedback on hover and click. Changing timescales also animates the x axis of the stacked area chart.

Hover for details



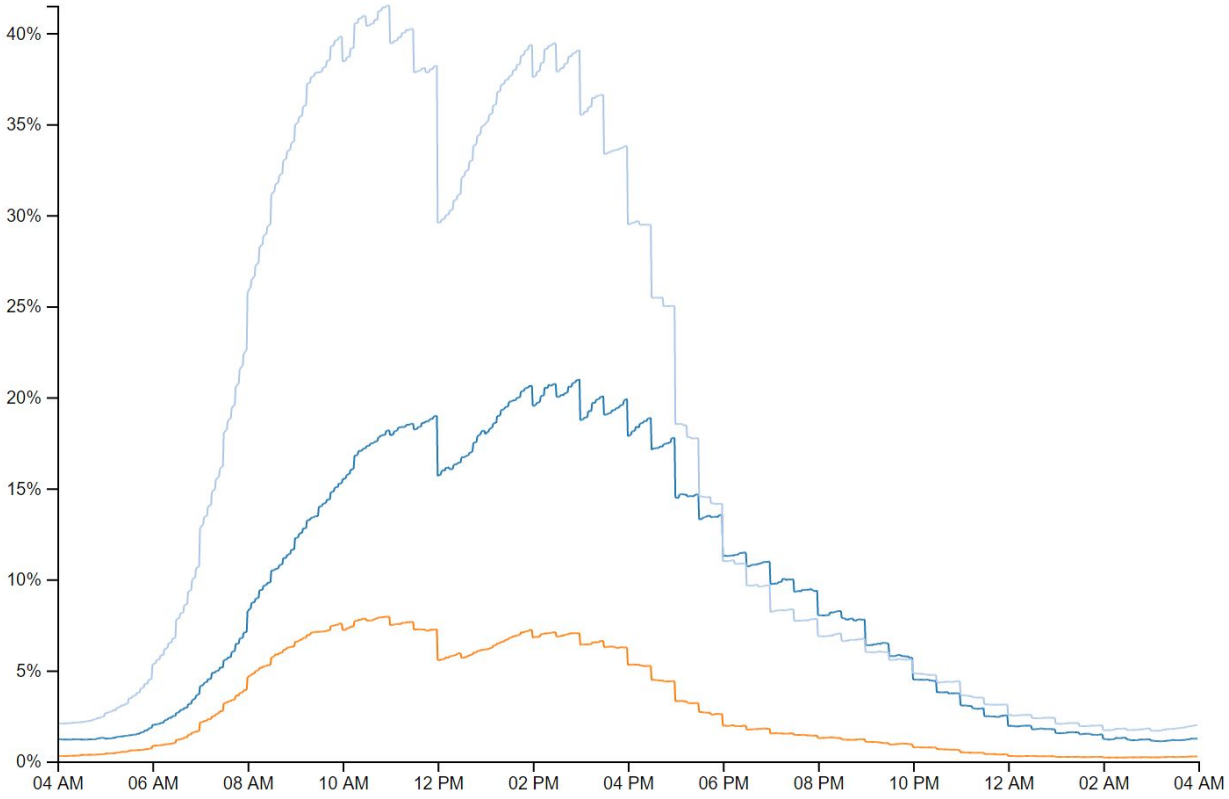
Hovering over the stacked area chart displays a thin vertical cursor. The sidebar updates to provide detailed information about the distribution of activities at that point in time. Additionally, the name of the activity the user is hovering over is highlighted in red. This hover interaction only provides additional information on demand to avoid overwhelming the user when they first encounter the visualization.

Select single activity

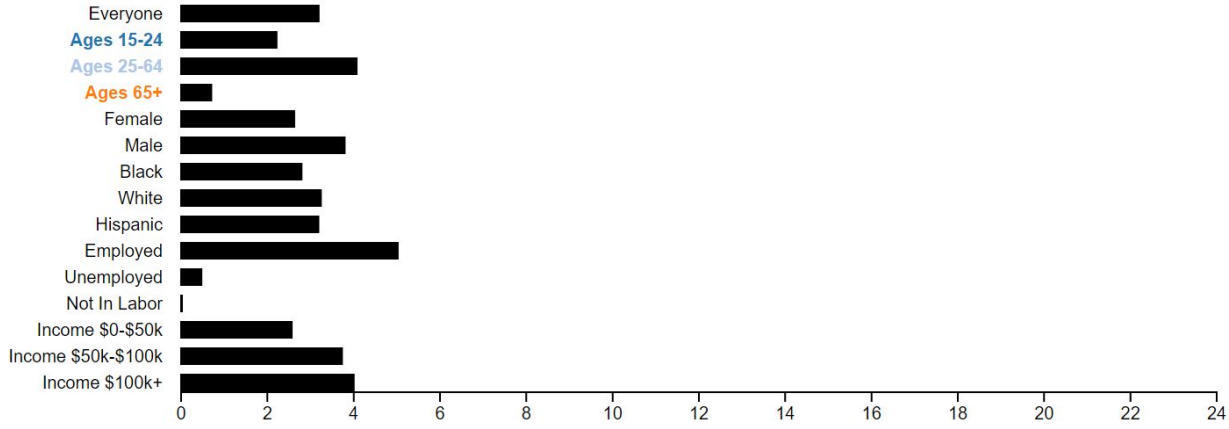


Clicking on an activity in the stacked area chart will trigger an animated transition to show only that activity along a common baseline. Hovering for details still works like normal. The animation helps to preserve context while transitioning to a common baseline. Clicking on the activity again will animate back to the full stacked area chart.

Compare demographics for single activity

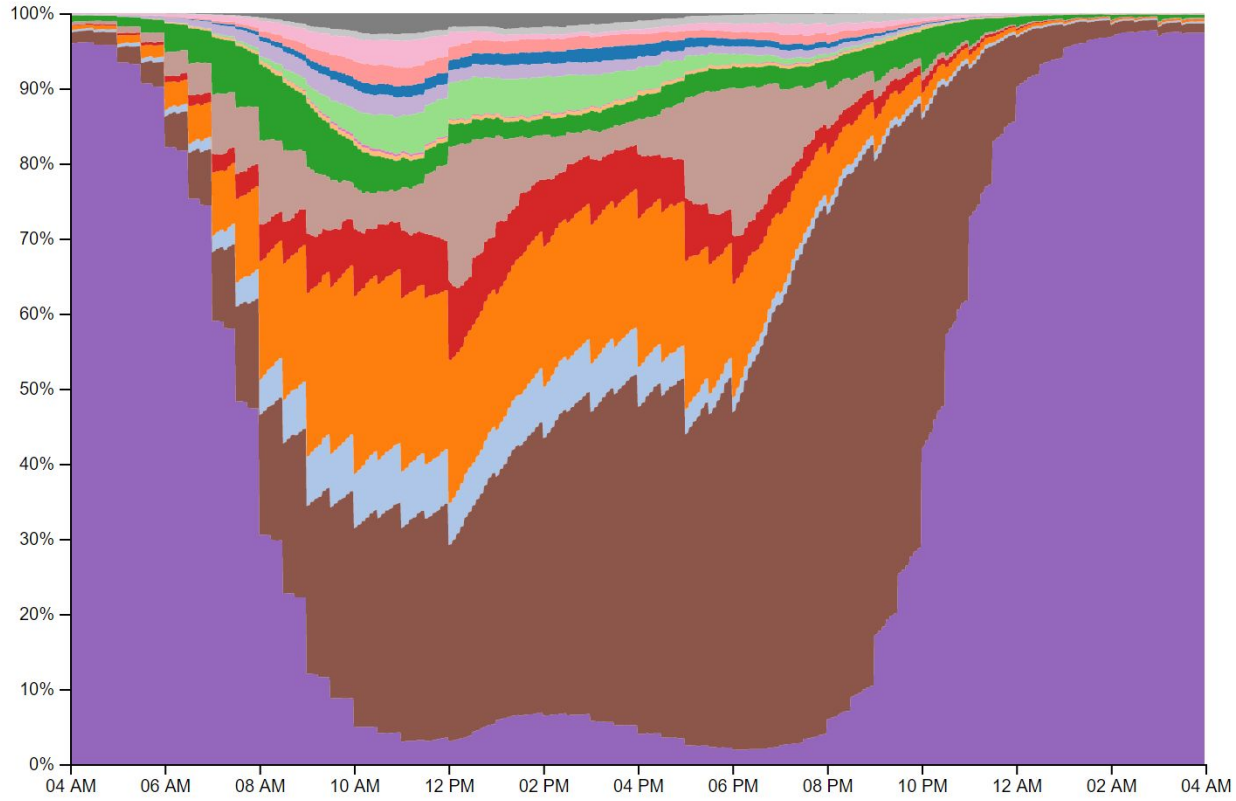


Explore demographics

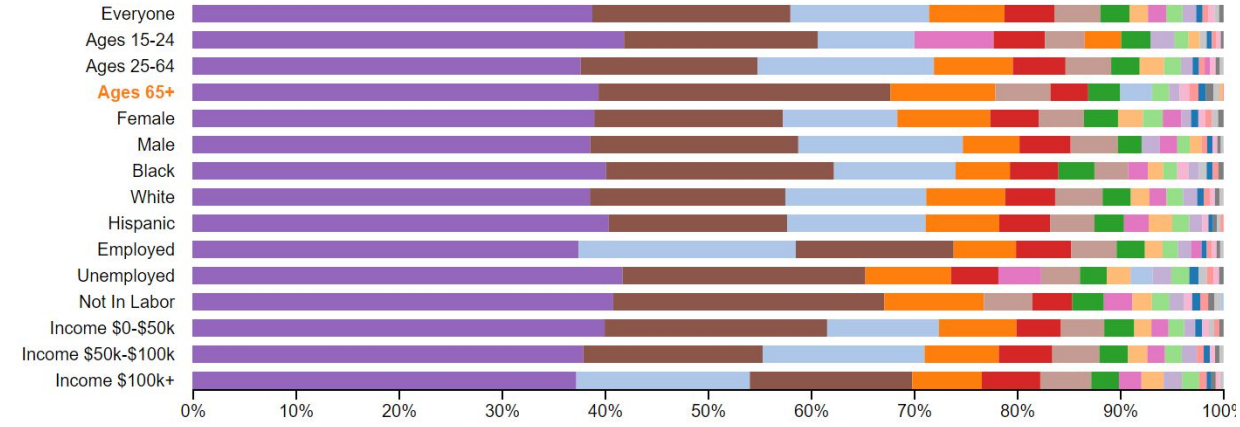


With a single activity selected, the user can click additional demographics in the demographic view below to compare them. The area chart will transition to a simple line chart. The color of each line maps to the color of the demographic in the view below. The line chart provides a way to directly compare demographics on a common baseline. Clicking on a demographic again will deselect it.

Change demographic for all activities



Explore demographics

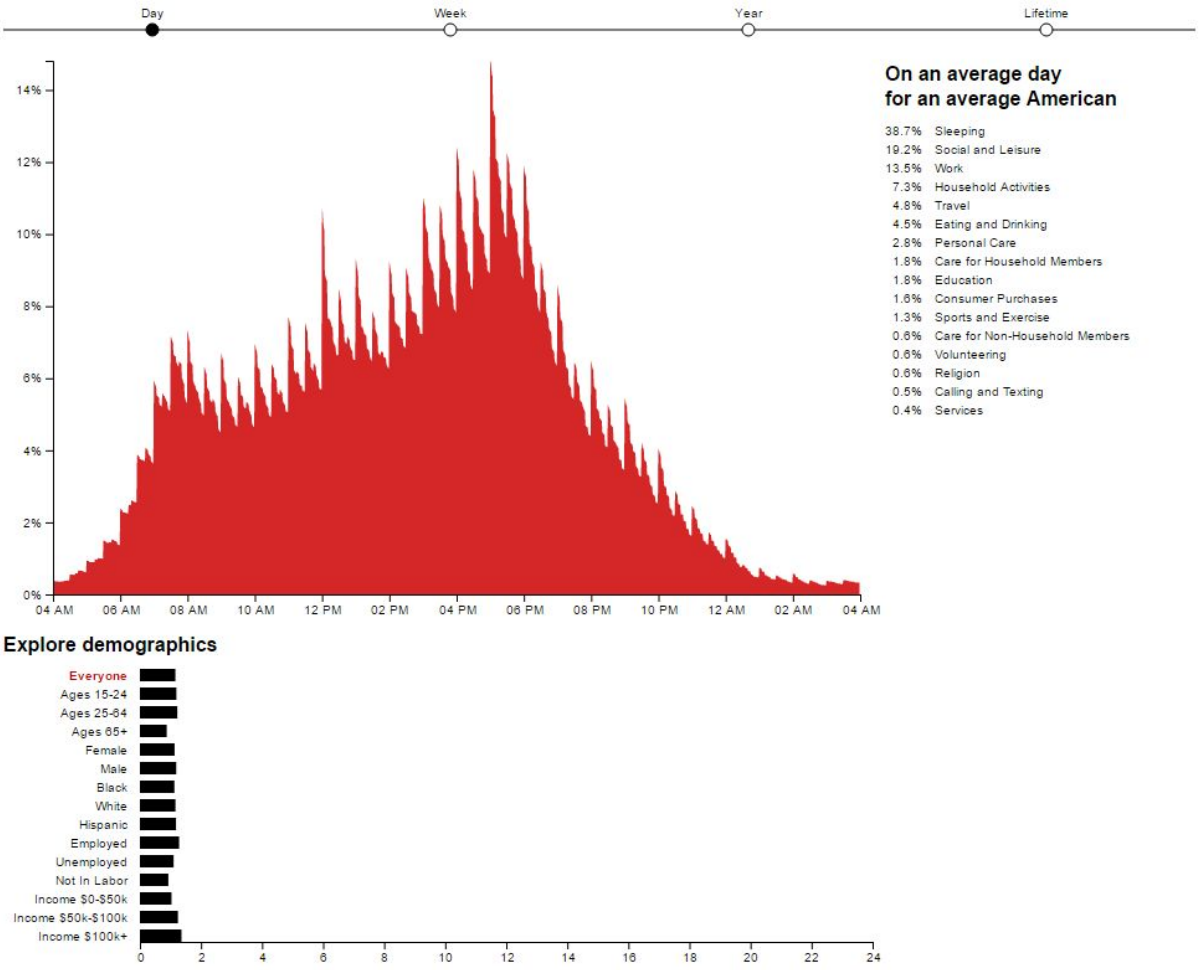


Clicking on a demographic when no activity is selected will animate a transition to that demographic in the stacked area chart. The sidebar will also update to display information from the selected demographic on hover. Again, animating this transition helps the user maintain context and makes differences between demographics more apparent.

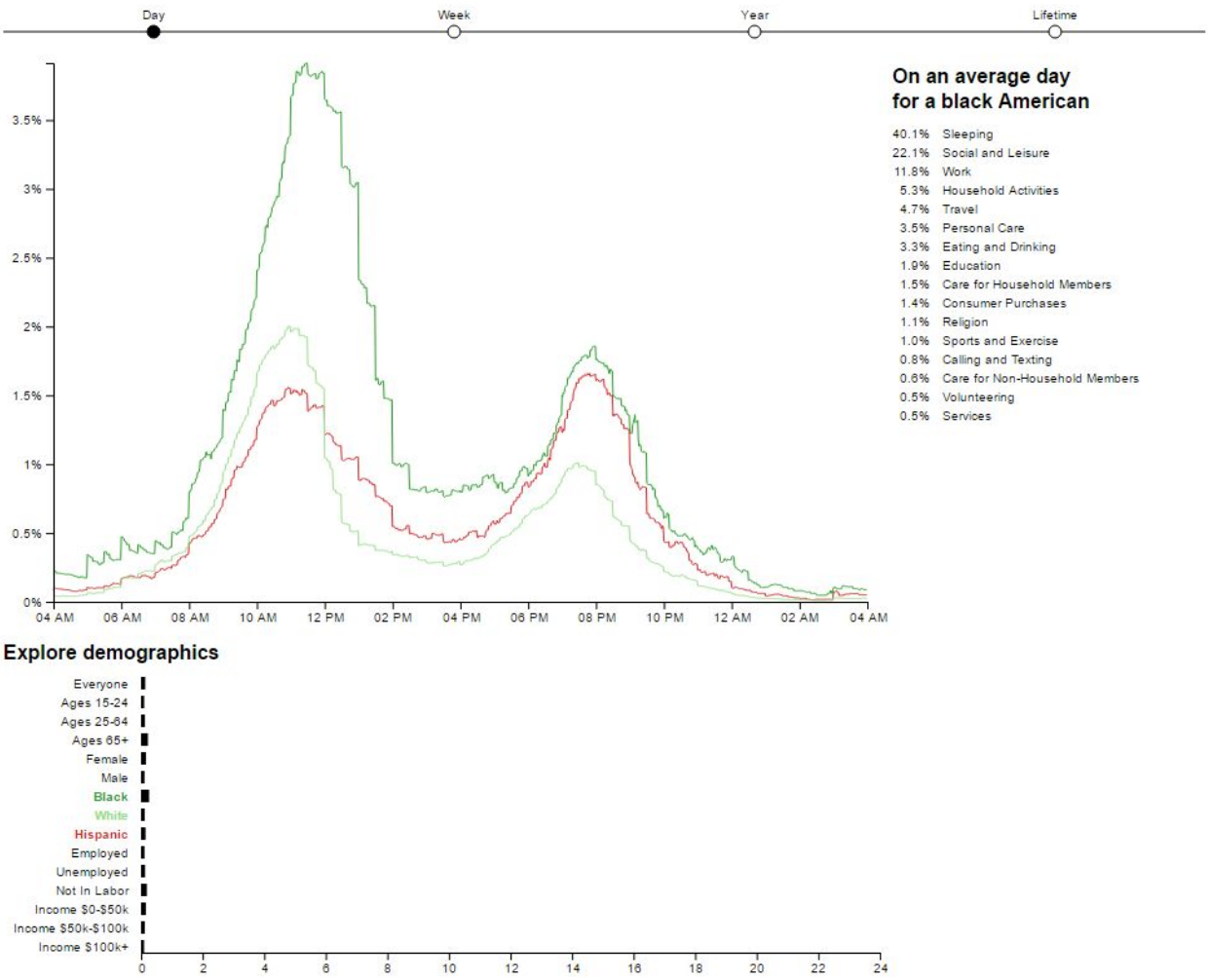
Evaluation

What did you learn about the data by using your visualizations? How did you answer your questions? How well does your visualization work, and how could you further improve it?

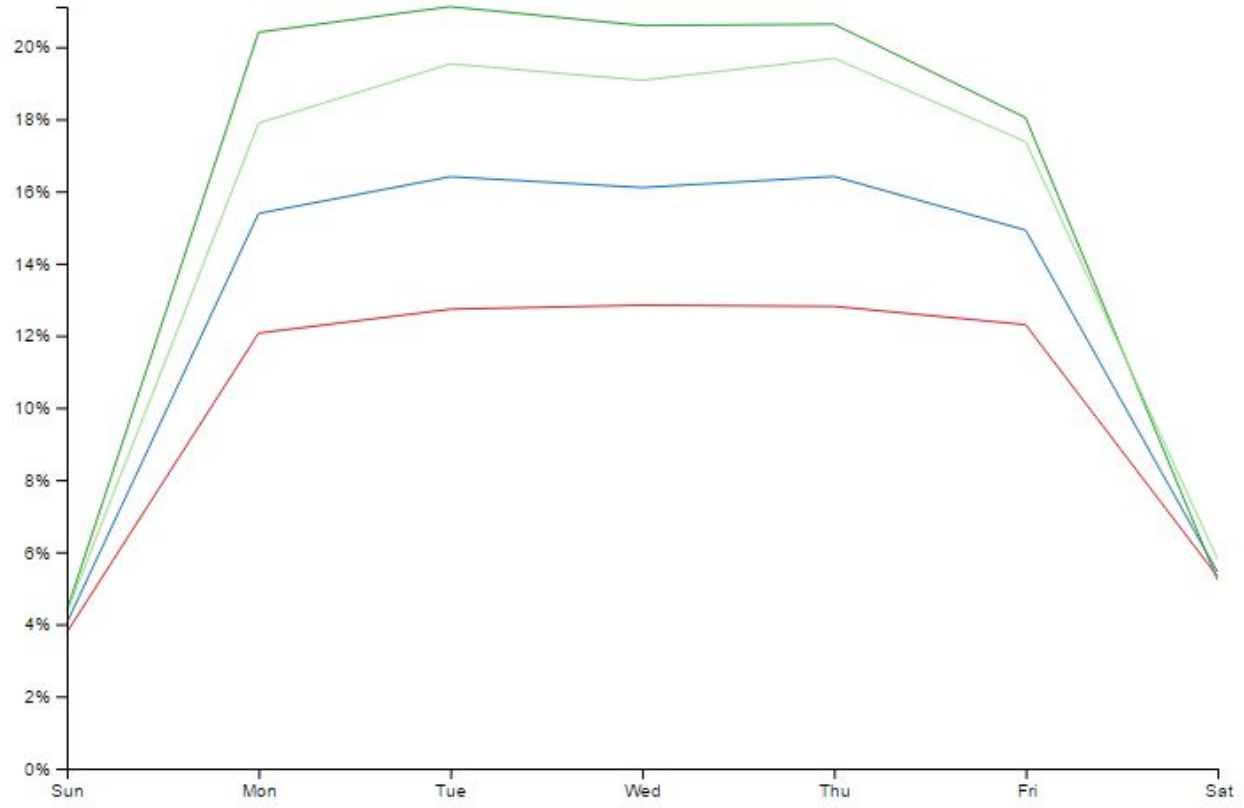
After exploring our data, we were able to answer our two questions to a good degree. For instance, through the primary day view, we were able to notice trends in travel, such as how most Americans reported leaving at the hour or half hour.

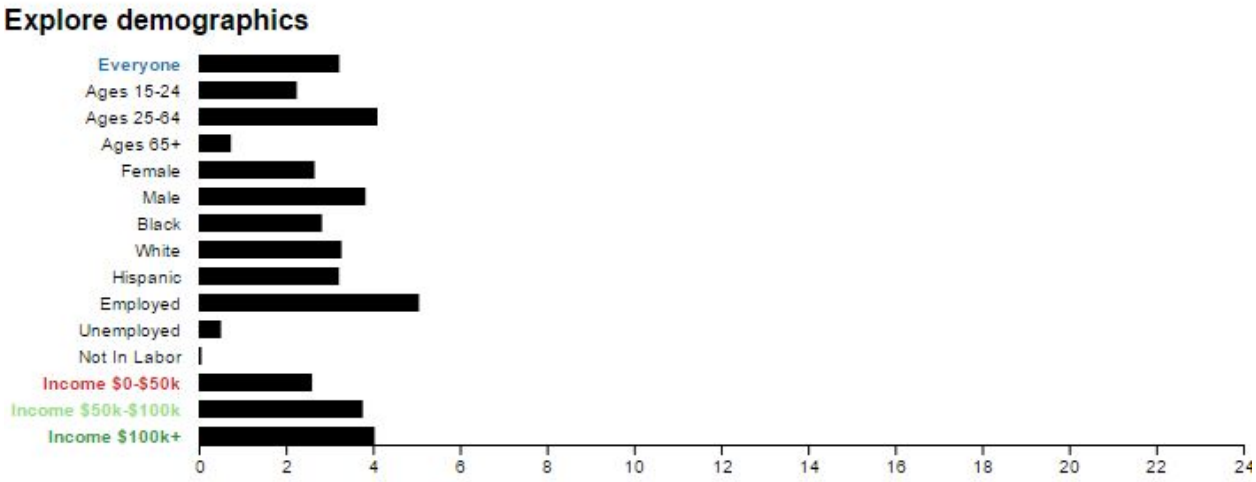


Additionally, day view we noticed interesting differences on what times and what percentage of different demographics practice religion. The black demographic had a higher overall percentage of Americans participating in the religion activity than the white demographic, and the hispanic demographic had a noticeable shift in the time when the majority of Americans participate in religion, from the earlier morning to the later evening service.

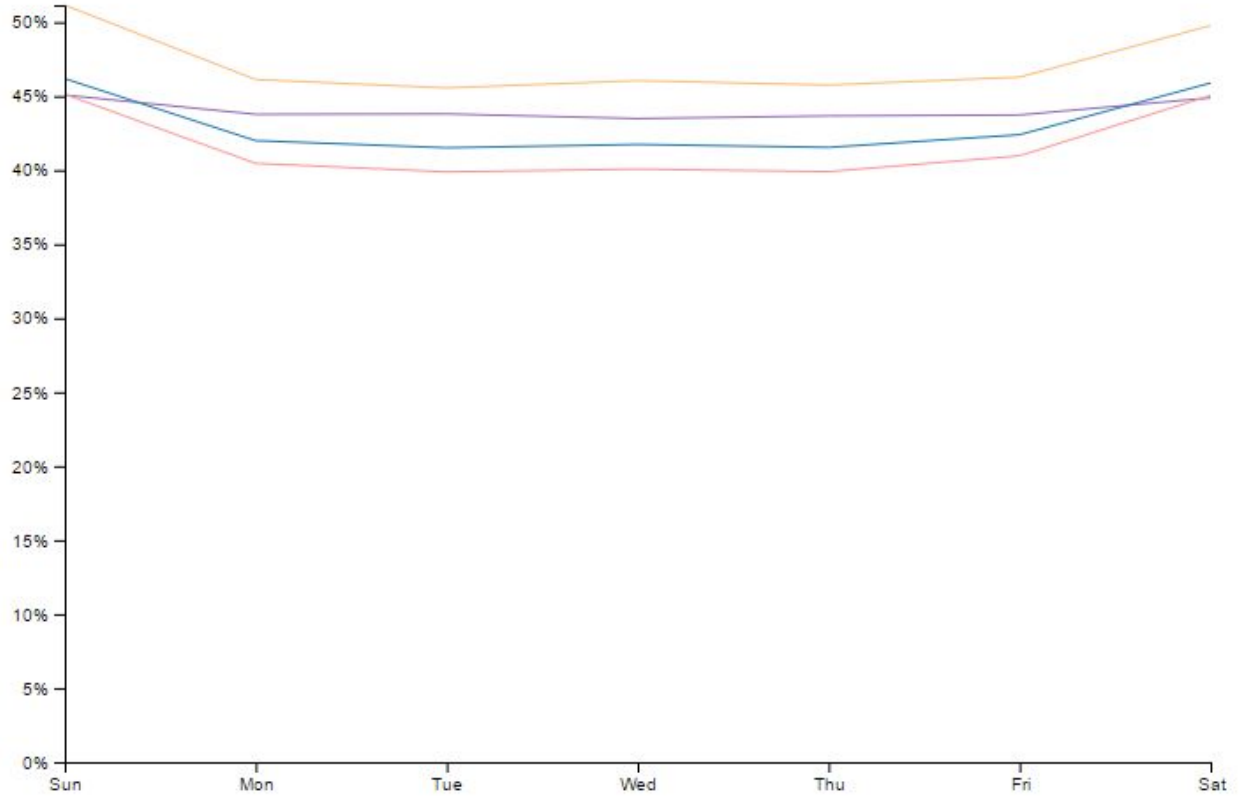


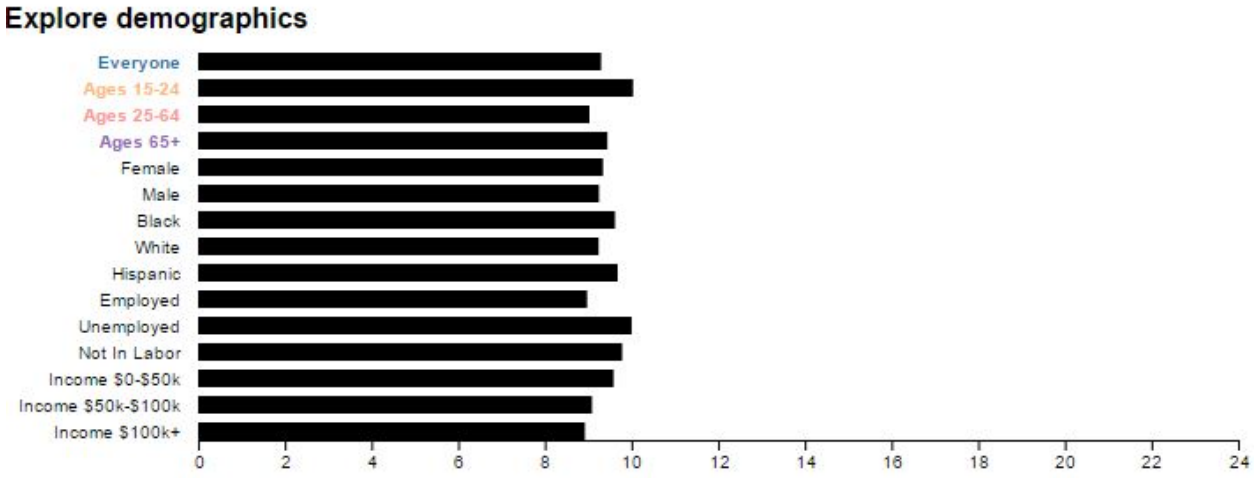
From the week view, we were able to learn that the percentage of each weekday that an average American spent working was influenced by their income status. For example, the income from $0 to $50,000 demographic spent the least percentage of day working throughout the week with an average of about 12% of each weekday. Contrastingly, the entire dataset had values of about 14% per weekday, the income from $50,000 to $100,000 demographic had values of about 18% weekday, and the income greater than $100,000 demographic had values of up to 20% per weekday. The largest difference represents about a 66% increase in amount of day spent working. Additionally, from the week view, you can see how with the entire dataset and the demographics for any income greater than $50,000, there is a trend to spend less time working on Friday. This makes sense, since most people tend to want to leave early from work on Friday in order to enjoy their weekend better.



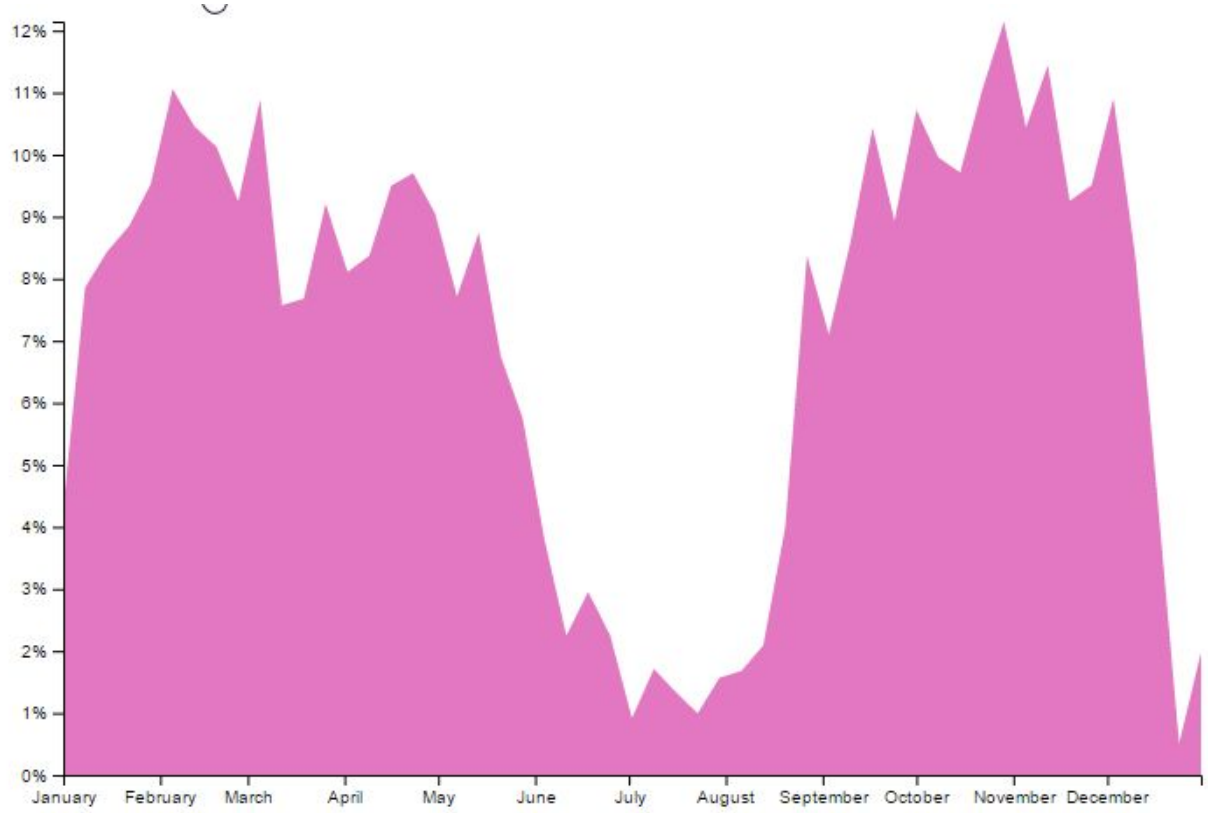


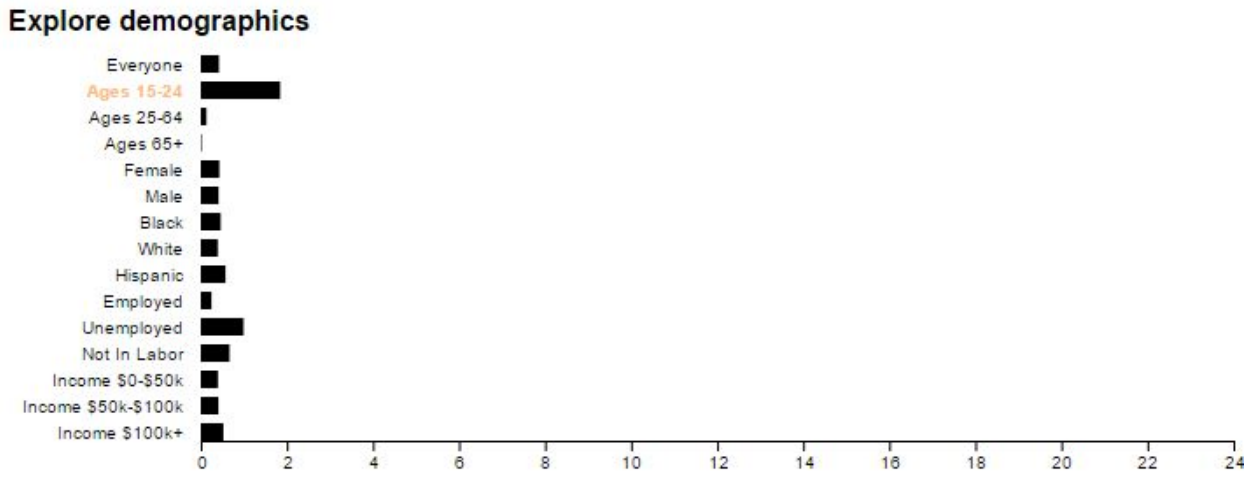
Further, from the week view, we noticed that sleep has a couple of interesting patterns. The first of which is that the age demographic of 15 to 24 has higher average percentage of day spent sleeping than both the overall population and other age demographics. Additionally, we observed that, the Americans on average spend longer sleeping on weekends however, with the age 65+ demographic, this is not nearly as apparent. Part of this may be because 65+ year olds have a lot more free time and less obligations that they are able to sleep however long they want every day.



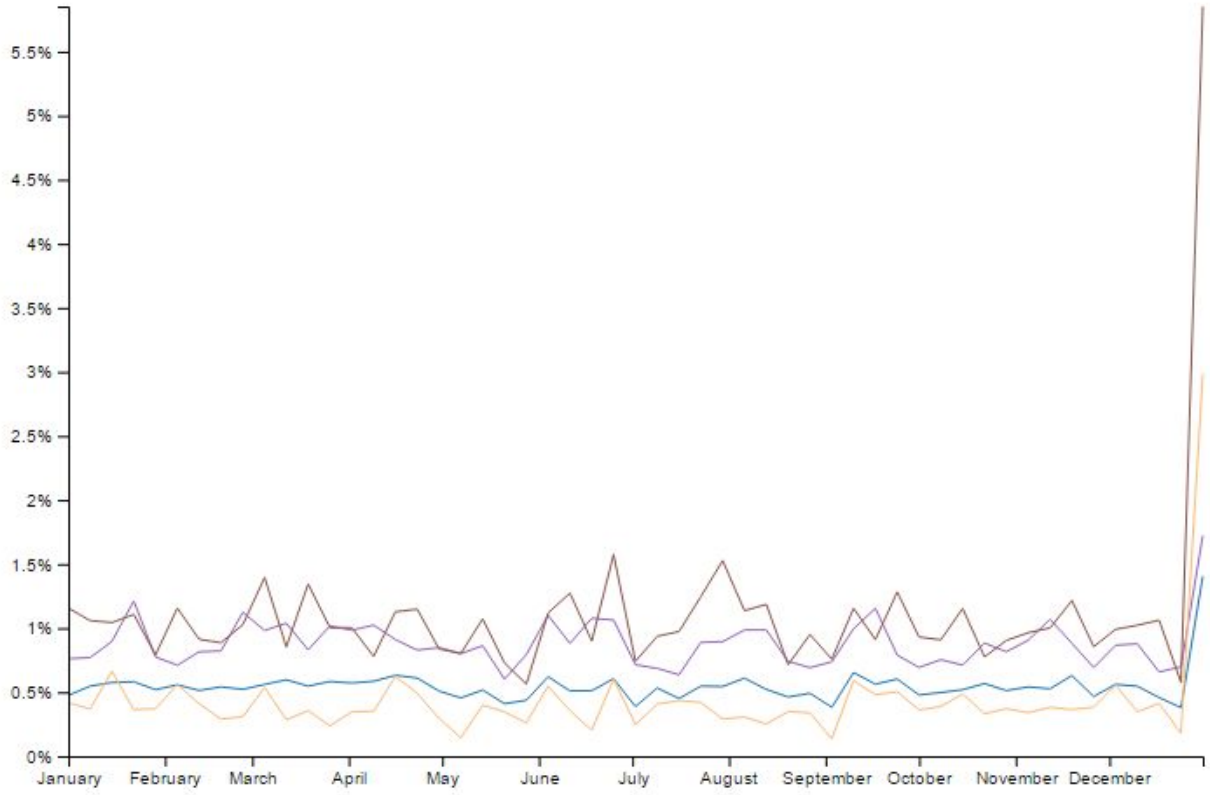


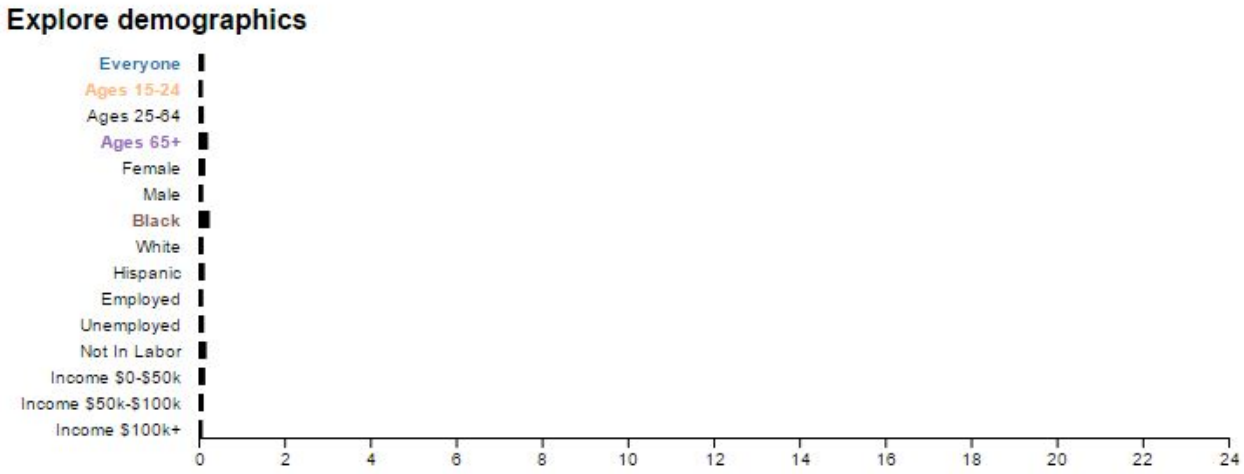
Through the year view, we reaffirmed the intuition that the amount of time spent on education decreases during summer break and other holidays, since students are no longer in school. This drop in time spent on education per day is especially apparent when looking at the data through the age 15 to 24 demographic.



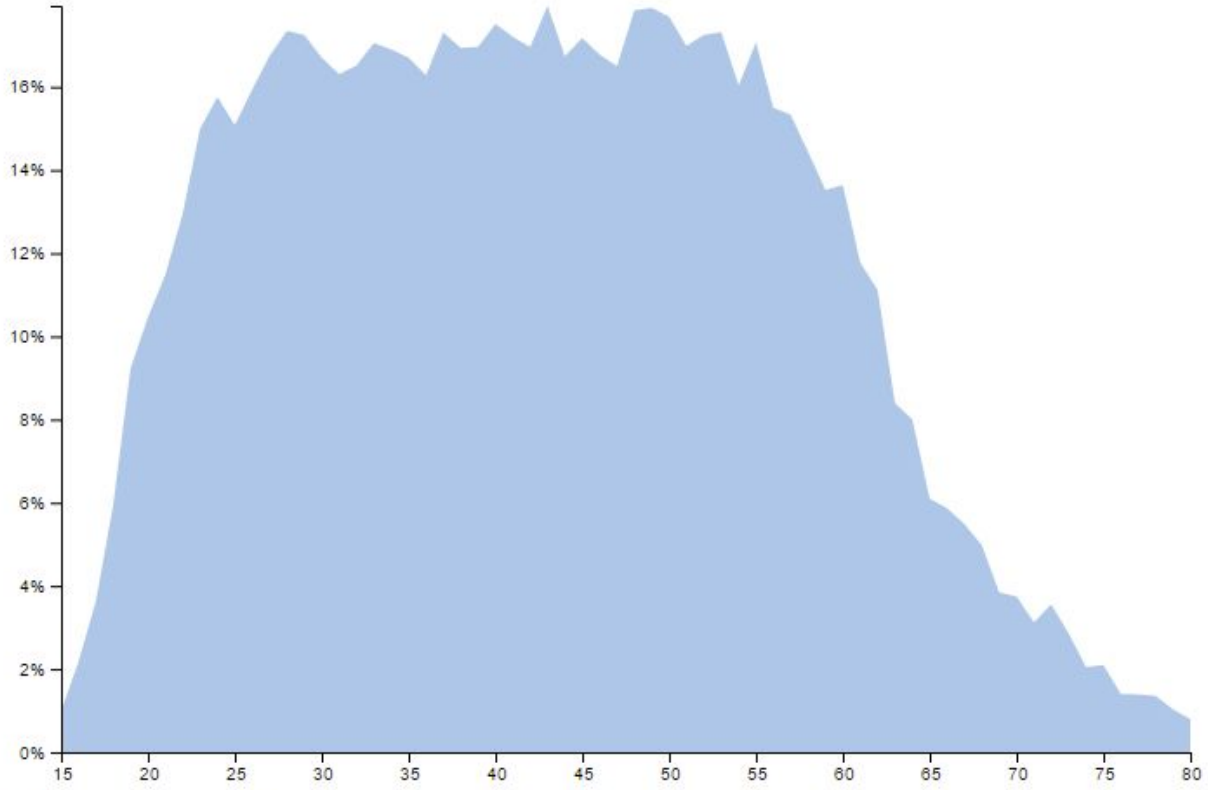


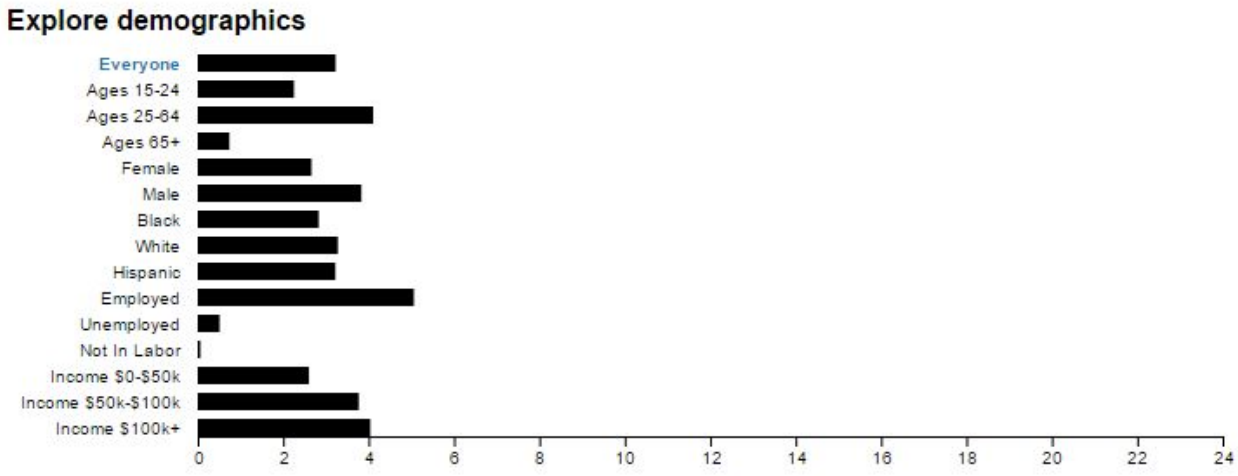
Additionally, in the year view, you can see a large spike in percentage of day spent on religion during December, likely due in large part to Christmas Mass. This spike is even larger for the age 15 to 24 demographic, since those americans are likely back home for the holidays and go to religious services with their families. Contrasting, for the age 65+ and the black demographics the overall yearly participation in religion is higher than the general american dataset. This might be due to their excess free time, and higher religious responsibilities respectively.



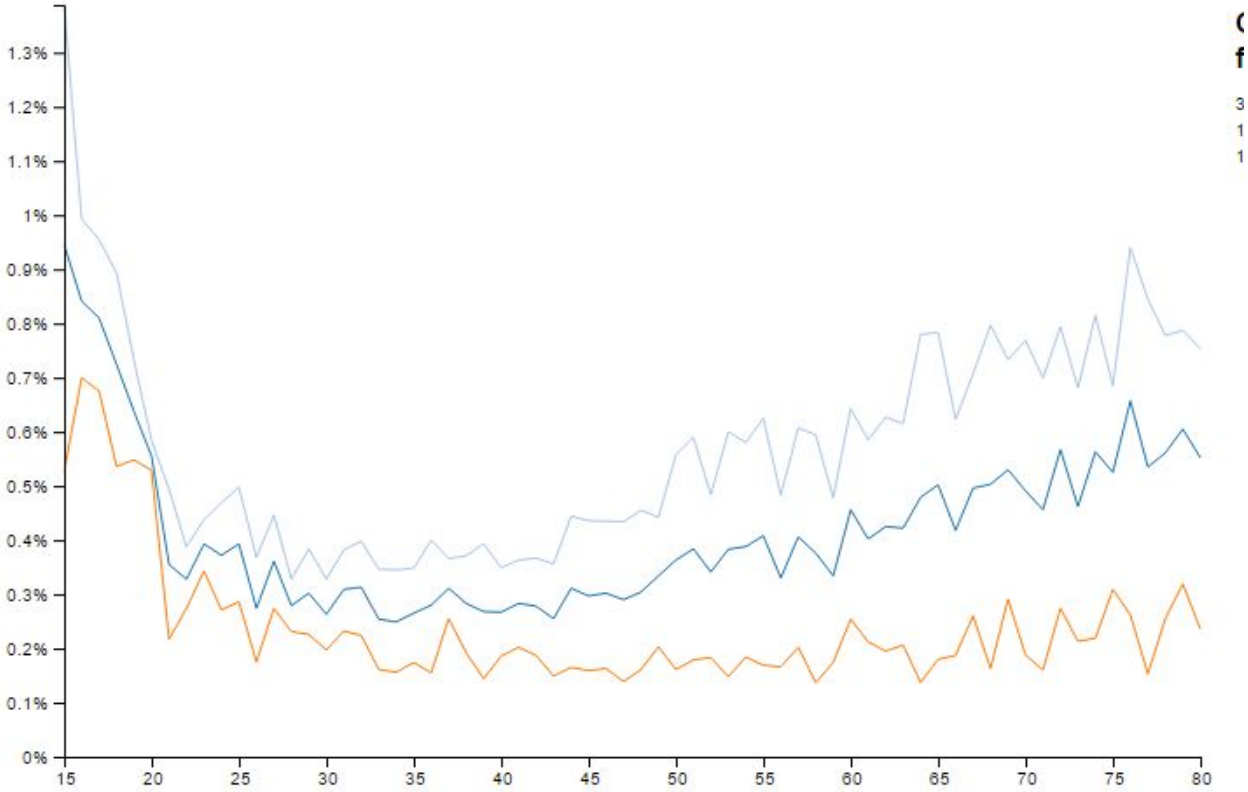


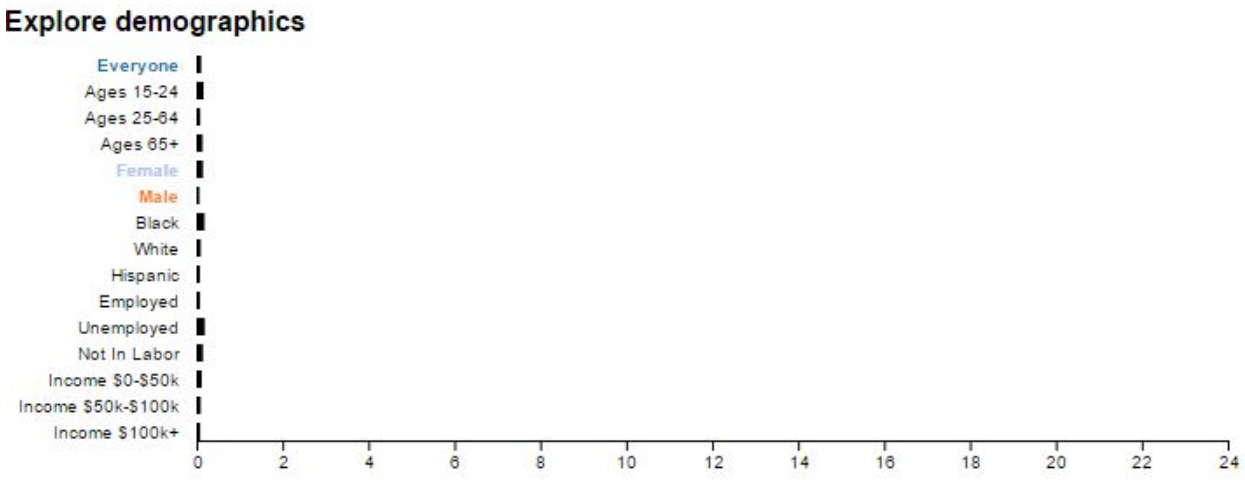
Finally, in the lifetime view, some observations we made were that the percentage of an average day spent working increases rapidly during the early the early years, and begins to drop off at around age 55.



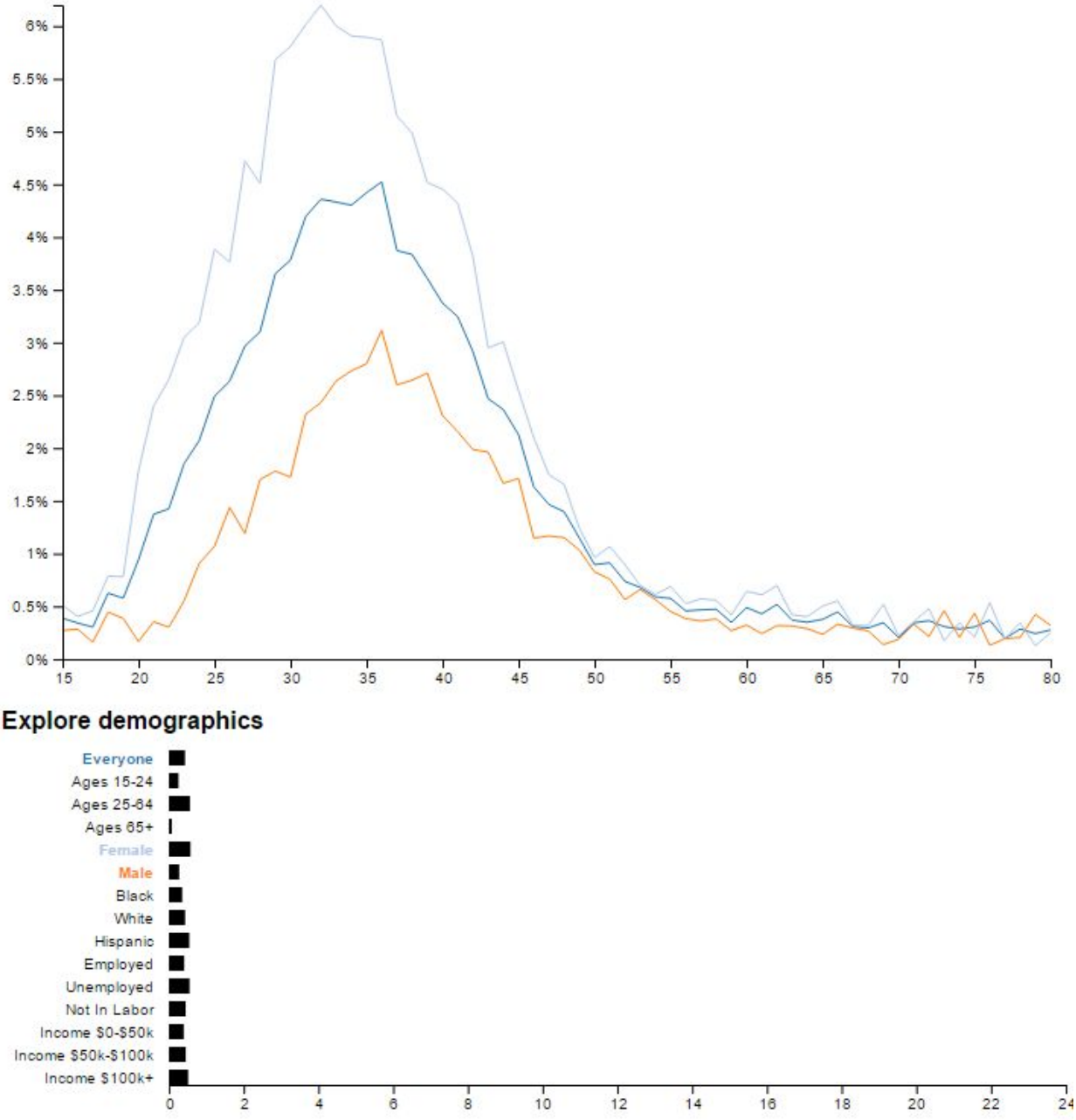


Further, the percentage of a day spent calling and texting has a huge spike in the age 15 to 20 range, which drops off quickly and then slowly rises again as age increases. This is likely due to the prevalence of texting within younger adults, and the amount of time older relatives spend calling their families.

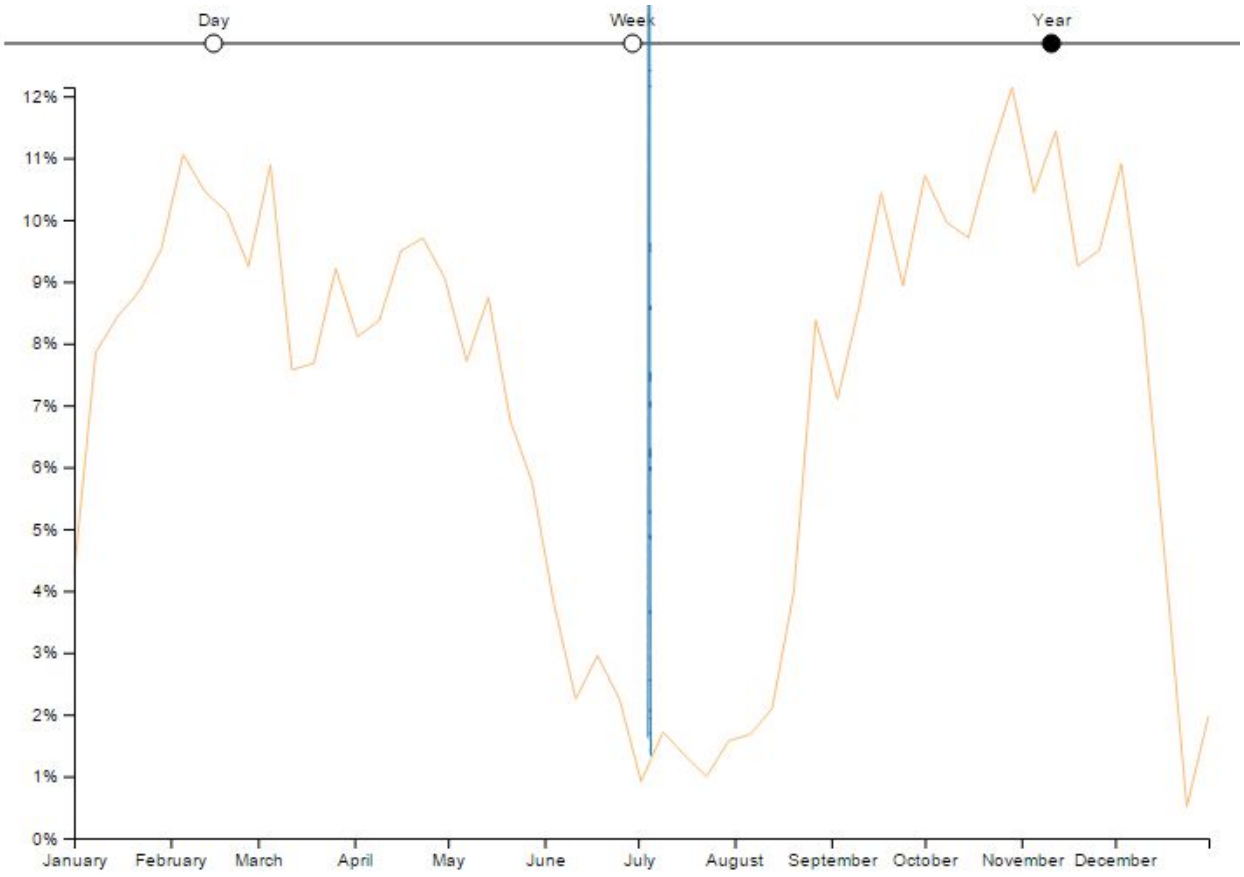




Finally, in lifetime view we noticed that there is a noticeable spike in the percentage of a day spent caring for household members from the age 25 to 45 range. This spike is larger for the female demographic and smaller for the male demographic. Both of these make sense because around age 25 is when most couples have children and start taking care of them, and around age 45 is when most of the children leave for college. Additionally, in American gender and social norms usually have the mother taking care of children, and as a result, the spike is larger for the female demographic.



Our visualization works fairly well however there are some improvements we could make, to make it more robust overall. For instance, in its current state, when clicking an activity on the stacked area chart, the demographics section lacks a good transition from a stacked horizontal bar chart to a single bar chart, as well as color change. Furthermore, there is a bug in our transitions that can be triggered by selecting an activity, selecting multiple demographics to compare and then changing the timescale without deselecting multiple demographics. When this is done, the next time a user selects an activity and begins clicking multiple demographics, the last selected demographic in the stack's line becomes either completely horizontal or vertical and does not represent the data it is supposed to.



Another improvement we could make on our visualization is to add a color identifier to the sidebar so that you can tell what each activity is on the stack area chart without having to hover over the area itself. This would be especially useful for the smaller activities that are at the top of the area chart. Additionally, having another section underneath the sidebar and adjacent to the demographics section would be useful to display summary information on the activity selected or useful things the user can look at within the graph. Overall, the visualization we created is fairly complete in its given state but could be improved to make it more user friendly and responsive.