Week 8: Process Book & Project Proposal

Authors: Sophie Sun, Verena Lin, and Eric Lin

Consumer Spending Telltales

Abstract

The lipstick index was inspired by a rise of cosmetic sales in the midst of the economic recession of the 2000s. Since then, economists have found characteristics in consumer spending that may predict future economic trends. Our motivation in pursuing this project is to illuminate consumer spending patterns that may predict the state of the economy. By analyzing certain product sales and overall industry health, we hope to evaluate whether these consumer spending trends can predict the financial market and the economic cycle (primarily via the national GDP). Our goals include mapping consumer behavior with future economic states, as well as proposing the implications that spending may have on policy making. We plan on using consumer expenditure surveys, as well as US Bureau of Labor Statistics datasets in order to map consumer patterns to periods of economic recessions and growth.

Week 9: Team Agreement & Detailed Project Plan

Team Agreement

1. Preparation
    a. We will read the assignment, come up with at least two broad ideas, and be prepared to listen to others' ideas prior to meetings
2. Communication
    a. We will communicate primarily through our text group chat and expect responses within three hours, with the exception of the 10:00 pm to 7:00 am hours (local hours)
    b. We will remind or bump members who aren't responding
    c. We will keep the rest of the team in the loop for when we are "going dark," or not responding due to other assignments or other life events
    d. We will get to the root of the problem and depending on circumstances, redistribute work if certain situations or surprises arise
3. Responsibility
    a. We will structure our work by brainstorming and discussing ideas together

while fleshing out the top ideas individually. We will then reconvene during our weekly meetings and discuss these ideas more in detail
    b. We will assign weekly individual to-dos
    c. We will assign a point person for each milestone - this point person will make final edits and submit the assignment for the milestone by the deadline
    d. We will organize each project milestone as follows:
        i. The team brainstorms and works on the milestone together in Google Docs after class or on the agreed upon date.
        ii. Individual to-dos should be completed by the set deadline established during the meetings
4. Logistics
    a. We will meet every 3:00 pm EST to discuss the milestone and lecture assignment. We will meet via this recurring Zoom link
5. Team Dynamic
    a. We will assume best intentions, and if conflicts arise, we will be respectful of the concern that is voiced and attempt to solve the conflict. If the team cannot solve the conflict, we will bring it up to our TF
    b. We will be open to expressing these concerns, and the team will proactively address the conflict immediately
    c. We will ensure that everyone has the opportunity to speak their ideas and voice their opinions by noticing who has spoken and who has not - for those who have not, we will ask them about their thoughts before moving on

We will use all of the above to fulfill the following goals:

1. Continue to learn about data visualization
2. Explore different types of data visualization by incorporating at least one advanced visualization
3. Make new friends
4. Have a good time!

Signatures: Sophie Sun, Verena Lin, Eric Lin
Date: 10/28/2020

Detailed Project Plan

Basic Information:

Project Title: Consumer Spending Telltales
Names: Sophie Sun, Verena Lin, and Eric Lin
Email Addresses: sophiesun@college.harvard.edu, tlin@college.harvard.edu, eric_lin@college.harvard.edu
Team Name: Sunny Lins

Background and Motivation:

Our project is inspired by economic indicators like the leading lipstick index, hot waitress index, and skirt length (hemline) theory. They all serve to correlate consumption patterns with the economic cycle. For example, the leading lipstick indicator proposed by Estee Lauder's chairman Leonard Lauder predicts that sales of less expensive indulgences like lipstick increase when consumers feel less confident about the economic future. Similarly, George Taylor of Wharton Business School proposed that skirt hemlines are higher when the economy is performing better. Through data visualization, we seek to present correlations similar to the one described by the lipstick index. We will start off by confirming or invalidating the lipstick index. Afterward, we will investigate whether such patterns can be observed through the consumption of other goods.

Related Work:

The lipstick index was initially proposed by Estee Lauder's Leonard Lauder. He did so based on his observations of lipstick sales. Recent research motivated by the ongoing pandemic has discounted the lipstick index -- pointing instead to products like face mask and moisturizer. There has been no official research confirming or invalidating any of these observational conclusions, so we seek to do so through this project.

Interestingly, fashion photographer Bill Cunningham wrote in his book *Fashion Climbing* that "timing is one of the most important ingredients of design. It's exactly like the Wall Street stock market." The comparison he draws between the runway and the stock market reflects our motivation for this project.

Some other related work and aesthetic visualizations include these income visualizations and consumer visualizations.

Audience and Questions:

We hope to inform economists and policymakers with our visualizations. By displaying consumer spending on certain products, as well as their expenses, and correlating them with times of economic recessions and booms, we will be able to illustrate what consumer behaviors look like during certain economic cycles. Therefore, the primary questions we will attempt to answer with our data story include the following:

1. What types of products are most strongly correlated with future economic recessions?
2. What types of products are negatively correlated with future economic recessions?
3. What types of products are positively correlated with future economic recessions?
4. Does consumer behavior differ among people of different income brackets?
5. How long in advance do spikes or dips in these purchases occur before an economic recession?
6. How did consumer behavior change during COVID? Was there a shift towards certain sectors of the economy?
7. What sectors of the economy are performing well and which are still lagging behind?

Our overarching goal from this project is to evaluate how consumer spending behavior can foreshadow economic recessions.

Data:

We will use data collected by the US Bureau of Labor Statistics to find and show patterns within consumer expenditure. We are choosing this source because it is extremely reliable and consistent (collected year after year). It is commonly used for economic visualizations published in the New York Times (such as this Voronoi diagram) and other credible sources.
We will also consider utilizing other sources from government agencies or economic institutions. Here is our full list:

- US BEA (Monthly expenditures; includes 2020!)
- US Bureau of Labor Statistics (Consumer Expenditure Survey)
- UK consumer expenditure
- US Census Bureau Expenses and Expenditures
- Collection of data sources, including the World Bank
- List of other potential sources

Data Cleanup:

Each of the economic datasets have up to 100k+ features and variables. As such, one core challenge we face is to navigate the dataset in a way where we can narrow down and focus on certain variables without losing the entire picture. Since these are government datasets published routinely every year / quarter, we do anticipate to spend some time cleaning the data but trust in their overall reliability.

Week 10: Data, Sketches, Decide & Storyboard

Meeting with Yalong

- Paper to read about visualizations and storytelling: http://vis.stanford.edu/files/2010-Narrative-InfoVis.pdf
- Example of excellent storytelling: http://www.r2d3.us/visual-intro-to-machine-learning-part-1/ http://mbtaviz.github.io/
- Examples of visualization with temporal data:
    - https://www.cs.middlebury.edu/~candrews/showcase/infovis_techniques_s16/themeriver/themeriver.html
    - https://vcg.informatik.uni-rostock.de/~ct/timeviz/timeviz.html

Data

The dataset we are planning to use for our project can be found at this Google drive link here. This dataset was found on the US Census website that breaks down consumer spending by month - it has done so since 1992. We took the adjusted sales data (which "reflect corrections to data," a fuller explanation can be found here) for 32 categories including the total retail and food services sales, motor vehicle and parts data, furniture and home furnishings stores, food and beverage stores, grocery stores, health and personal care stores, clothing stores, department stores, and more, all broken down by month. This dataset includes consumer spending data starting in January 2007 and ending in August 2020.

We also want to explore GDP trends. Here is a dataset for GDP per state. The data is broken down into a separate csv file per state from 2005 to 2020. We then merge the csv's together into a nationwide dataset over the past 15 years. There is also quarterly data by state of consumer expenditures (similar to the above paragraph). The data is segmented into different sectors of the economy and both absolute and relative (percentage of the total U.S.) values are given.

Sketches

See below, under Storytelling.

Deciding

Rough draft of a general outline for how our final website may look like:

1. GDP comparison of 2008 and 2020
2. Interactive component
3. Transition to online learning — Zoom
    a. In parallel, stock market crash 2008
4. How long it took for recessions to reach lowest point / how sharp was decline
    a. WHO global crisis
    b. Lehman Brothers
5. Consumer expenditure
    a. Compare sectors
6. Narrow down into one sector
7. Conclusion: public policy

Storytelling

Storyboard link here.

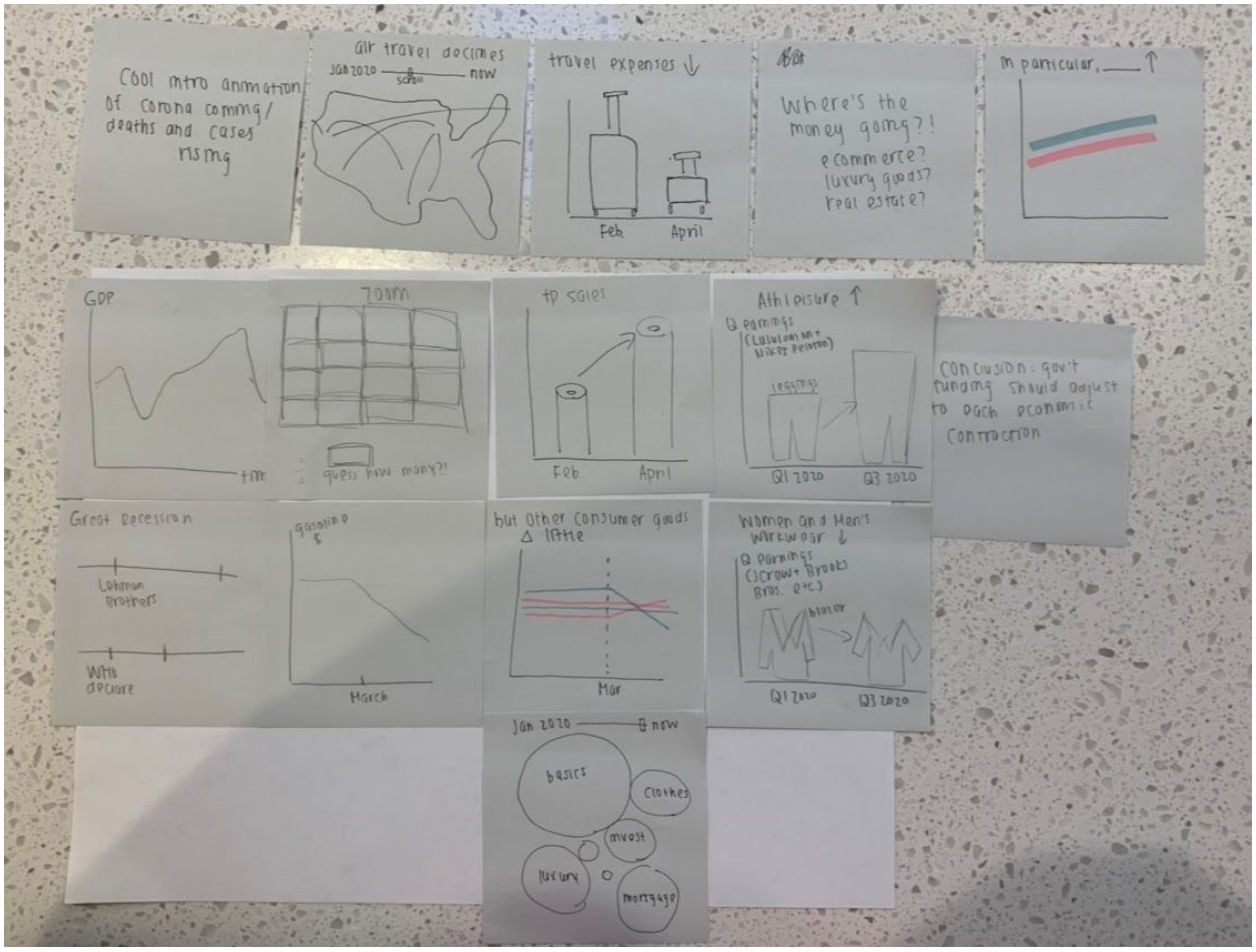
We will first begin by analyzing the data from the recent COVID-19 recession. In doing so, we will be able to capture the audience's attention by displaying the data that is most relevant to their lives now. We will set up a quick background with regards to the fact that consumer behavior changes when the economy in the US changes, specifically when the US faces a recession. We will analyze specific products that consumers purchased using the monthly dataset and exploit trends in the data. The story here is that consumer spending increases and decreases based on 1. The product that they purchase and 2. The state of the economy.

From there, we will look back into history by analyzing the data from the 2007-2009 US financial crisis. We will analyze general consumer spending behavior and attempt to use the data points from this historical recession to predict and analyze the COVID-19 recession.

The difference between the aforementioned COVID-19 and 2007-2009 financial crises is that we will analyze more specific trends of consumer spending behavior, such as athleisure, when exploring the COVID-19 data on consumer behavior. The story here is that although consumer spending changes according to the state of the economy, the *different* events that caused the current state of the economy leads to differences in consumer spending changes.

This has public policy implications, especially in how the government should react to different economic contractions. Depending on what's driving the contraction, different sectors of the economy are positively vs. negatively impacted. For example, some sectors (i.e. food) were affected in 2008 but not during the pandemic. Thus, in

crafting programs like the Paycheck Protection Program (PPP), the government should focus their funding on suffering industries, rather than make the funding available to all.



Note: top row additional drawings/notes that could come in use depending on data analysis

Week 11: Prototype V1

- Name of students that worked on prototype V1 submission: Sophie Sun, Eric Lin, and Verena Lin.
- Data scraping and cleaning complete: consumer spending dataset, GDP per state dataset, consumer expenditure dataset.
- Storytelling:
    - We will first begin by analyzing the data from the recent COVID-19 recession. In doing so, we will be able to capture the audience's attention by displaying the data that is most relevant to their lives now. We will set up a quick background with regards to the fact that consumer behavior changes when the economy in the US changes, specifically when the US faces a recession. We will analyze specific products that consumers

purchased using the monthly dataset and exploit trends in the data. The story here is that consumer spending increases and decreases based on 1. The product that they purchase and 2. The state of the economy.
    - From there, we will look back into history by analyzing the data from the 2007-2009 US financial crisis. We will analyze general consumer spending behavior and attempt to use the data points from this historical recession to predict and analyze the COVID-19 recession.
    - The difference between the aforementioned COVID-19 and 2007-2009 financial crises is that we will analyze more specific trends of consumer spending behavior, such as athleisure, when exploring the COVID-19 data on consumer behavior. The story here is that although consumer spending changes according to the state of the economy, the *different* events that caused the current state of the economy leads to differences in consumer spending changes.
    - This has public policy implications, especially in how the government should react to different economic contractions. Depending on what's driving the contraction, different sectors of the economy are positively vs. negatively impacted. For example, some sectors (i.e. food) were affected in 2008 but not during the pandemic. Thus, in crafting programs like the Paycheck Protection Program (PPP), the government should focus their funding on suffering industries, rather than make the funding available to all.
- Each of the following can be found in this zip file.
    - At least one D3 visualization already partly implemented, and drafts for 2 more visualizations
    - Rough webpage design and structure has to be done and implemented
    - The first design of an innovative view
    - Interactions (e.g., filtering, brushing, etc.) have to be designed

Week 12: Prototype V2

Week 12: Prototype V2
We expect you to be 95% done with the implementation of your data story. It should be ready to be tested by a random person the following week.

Specific requirements:

- Please submit only code (and have an up to date process book)!
- All views must be complete and working by this point!
- We will evaluate the progress you made from Prototype V1 to Prototype V2

- You will have two weeks until the final submission. The last week is meant for fine-tuning, incorporating feedback, documenting, creating the screencast, etc.

If the dataset is too large to upload, please contact your project TF so that they can either copy your dataset directly or use dropbox, etc.

Finally, you will fill out this self and peer evaluation form to reflect on how well you and your teammates performed in the final project so far. **Every team member** needs to fill out this form. Only the teaching team will see your data. This form may affect the grades of team members who did not pull their weight. We will ask you to fill in the same form again at the end of the final project.

This is a team assignment. Please submit only once! The submission should be done by the assigned person in your team (usually the team leader) who is doing the submissions.

Please check the final project overview for more information about the final project.
Please submit your a **before 11:59 pm EDT on Sunday, November 22**.

Week 13: Test

Week 13: Test
You will conduct a think-aloud study with a random person from another team that we will assign to you. You will document the results of the study and use them to make changes to your prototype.

Week 14: Wrap Up & Submission

Week 14: Wrap Up & Submission
You will finalize your web-based data story and submit your final project at the end of the week. You will host your final project as a public website using GitHub pages or any other web hosting service of your choice (we strongly encourage you to use GitHub). You need to make sure your public website is up and running and include the link of your website in your process book.

In addition to the process book, data, and code of your project you will also create a **two-minute screen-cast with narration** showing a walkthrough of your data story. You can use any screencast tool of your choice. Please make sure that the sound quality of your video

is good - it may be worthwhile to invest in an external USB microphone. Please use a standard video file format with a standard video codec.