2-2022

Slices of the Big Apple: A Visual Explanation and Analysis of the New York City Budget

Joanne Ramadani
The Graduate Center, City University of New York

SLICES OF THE BIG APPLE: A VISUAL EXPLANATION AND

ANALYSIS OF THE NEW YORK CITY BUDGET

by

JOANNE RAMADANI

A capstone project submitted to the Graduate Faculty in Data Analysis and Visualization in

partial fulfillment of the requirements for the degree of Master of Science, The City University

of New York

2021

Slices of the Big Apple:

A Visual Explanation and Analysis of the New York City Budget

by

Joanne Ramadani

This manuscript has been read and accepted for the Graduate Faculty in Data
Analysis and Visualization in satisfaction of the capstone project requirement for
the degree of Master of Science.

_____   _____

Date   Aucher Serr

                                                 Thesis Advisor

_____   _____

Date   Matthew Gold

                                                 Executive Officer

THE CITY UNIVERSITY OF NEW YORK

ABSTRACT

Slices of the Big Apple:

A Visual Explanation and Analysis of the New York City Budget

by

Joanne Ramadani

Advisor: Aucher Serr

As a component of government, budgets are fundamental not only to improving the quality of a shared society, but also to understanding what our government officials consider to be their priorities. However, most budgets can be difficult to understand, using terms that are not familiar to people who have not studied finance or economics. To that end, Slices of the Big Apple is an interactive, centralized narrative website that uses visualizations at its core in order to: 1) facilitate a holistic understanding of the New York City government budget for NYC residents; and 2) conduct a five-year analysis of Community Board budget allocation to demonstrate one way to analyze the budget data. Using five years (Fiscal Years 2017-2021) of budget data from the New York City Comptroller's Office, Slices of the Big Apple takes a two-fold approach to building literacy in our government finances by first visualizing the components of the budget and its process, and then leading the audience through the data for one slice of the budget: community board funding. By choosing an interactive narrative approach to visualizing budget data, this project offers an alternative to the classic "budget data dashboard" that makes such an important part of local government accessible to all.

ACKNOWLEDGMENTS

So many people supported me through the creation of this enormous project, and without them, getting to a point of completion would have been much more difficult than it was.

First, this project would not have been possible without the guidance and support of my advisor, Aucher Serr—thank you so much for all of your wise words, helpful tips, feedback, and constant cheerleading. It's been one of the highlights of my academic career to have been your student and I truly appreciate your patience and wisdom throughout not only this project, but all the projects I had the opportunity to create in your classes.

My thanks also to the professors I've had during this degree program, who gave me the foundation that I needed to build this project, and to Matt Gold and Jason Nielsen for providing guidance and structure, and for answering questions that I had not only during this capstone project, but also during my time in the Data Analysis and Visualization program. I've learned so much during my time at the Graduate Center, and that knowledge has been an incredible gift.

Thank you to my mother, who has been a rock when things got tough (which was frequently), who was one of the first to actually set eyes on the unfinished project (and gave me some great feedback), and who listened through every roadblock, breakthrough, and stalemate— you're one of the most awesome people I know, and I'm so grateful to have you as my mom.

My gratitude also to my beautiful, amazing communities of writers and artists who provided encouragement, support, a co-working space, and reassurance that the project was not as bad as I thought it was in the early stages, with particular thanks to Krissey, Grace, Tom, Lindsey, and Vol, who sprinted with me for hours even though I was actually coding instead of writing prose. I promise I'll stop using the "!sprint join edit" command sometime soon for things that aren't editing!

TABLE OF CONTENTS

LIST OF FIGURES

DIGITAL MANIFEST

I.   Capstone Whitepaper (PDF)

II.   WARC files

      a. Project Website (archived version of https://jramadani.github.io/nycbudget-slices/)

III. Code and other deliverables

      a. Zip file containing the contents of the GitHub repository at time of deposit.

          (https://github.com/jramadani/nycbudget-slices)

TECHNICAL SPECIFICATIONS

Slices of the Big Apple can be experienced in a live format on the web at

https://jramadani.github.io/nycbudget-slices/, while the code repository for the project is located on

GitHub here: https://github.com/jramadani/nycbudget-slices.

The project was built using HTML5, CSS3, and JavaScript ES6, relying on the JavaScript

packages d3.js (version 7.1.1) , scrollama.js (version 2.2.3), and Node.js for visualization,

scrollytelling, and package management/server initialization respectively. Express.js (version

4.17.2) was also used in conjunction with Node to create a server file that will allow users to

initialize the project locally, if files are downloaded to their machine.

The repository contains, among other files, one main HTML file (index.html), one main

CSS file (style.css), and one main JavaScript file (index.js). In addition to these files, there is a

"node_modules" subfolder which contains all dependencies for this project, which are listed with

their correct versions in the package.json file in the root folder.

To launch this project on a local machine, users must first make sure they have installed

Node.js globally. The version of Node used in development of this project is 14.11.0, and on a full

install of Node, npm will also be installed, allowing full access to package management. The

version of npm used in this project is 7.21.1. After making sure that all necessary dependencies

have been installed, navigate to the project folder on the command line (PowerShell for Windows

users). Once in the root of the project folder, run "node server.js" to launch the local server created

for this project.

In a web browser, enter the URL "localhost:5500". If port 5500 is already occupied, modify

line 6 of server.js to an unoccupied port number.

NARRATIVE

Background

During the Black Lives Matters protests in 2020, the New York City (NYC) budget came to the forefront of public conversation when citizens began to question why the New York Police Department (NYPD) receives so much funding, leading to a diversion of $1 billion from the NYPD to other agencies.[1] Despite this renewed spotlight on the budget, the resources explaining what the budget is and what it does are scattered across the city's various agency websites, leaving residents to jump from site to site, attempting to piece together the fundamentals of how the budget works.

In a general sense, budgets are plans covering future income and spending, but they are also a reflection of a society's priorities.[2] Much like politics, they are foundational to deciding what the quality of daily life will be like for residents, and this is no different in New York City. The budget for NYC is supposed to account for the needs of approximately 8 million citizens, and because of this, the budget is suitably large on a yearly basis—covering roughly 150 agencies and over 300 departments, the budget is an amalgam of seven different components that do not necessarily all add up to one cohesive whole. Rather, those different components describe different parts of the budget, sometimes overlapping and sometimes extending beyond the range of the current fiscal year.

Part of NYC's unique budget is also the existence of participatory budgeting, which is positioned as a way for citizens to get more directly involved in the budget process. According to Ebdon et al, "there is some evidence that the governmental environment is related to the

---

[1] "New York City Passes Budget with $1 Billion Cut to NYPD amid City Hall Protest," ABC7 New York, June 30, 2020, https://abc7ny.com/defund-police-protesters-black-lives-matter-mayor-bill-de-blasio/6284680/.

[2] R. Mark Musell and Ryan Yeung, *Understanding Government Budgets: a Guide to Practices in the Public Service,* (Milton: Routledge, 2020), 2.

involvement of citizens in the budgeting process,"[3] so it might seem to be the case that participatory budgeting is a good thing overall. In actuality, this shows itself to be little more than a sandbox provided to residents by legislators—according to an article published in the journal *Administration & Society*, the participatory budgeting process in NYC "earmarks" or sets aside a certain amount from the budget in order to generate goodwill between council members and their constituents, leading to funding more projects at lower average amounts than in districts without participatory budgeting.[4]

All of this information was synthesized into the project, which walks the audience through these points in a setting that attempts to make this seemingly opaque and immense topic a relaxed and approachable experience.

The Project

Slices of the Big Apple is an interactive narrative that uses data visualization and animation in conjunction with data storytelling to: 1) holistically explain the New York City budget system to an everyday resident of the city; and 2) attempt to find an effective way to visualize the enormity of budgets in that sphere to the benefit of those involved in the actual process.

I am deeply interested in making inscrutable topics accessible to the average person, particularly if those topics relate to political science, which was one of my previous academic areas of focus. The intersection of political science and data visualization tends to be where it can be game-changing to very clearly explain information that has a direct impact on a person's daily life,

---

[3] Carol Ebden and Aimee L. Franklin, "Citizen Participation in Budgeting Theory," *Public Administration Review* 66, no. 3 (2006): 438, accessed July 23, 2021, http://www.jstor.org/stable/3843923.

[4] "Does Participatory Budgeting Alter Public Spending? Evidence From New York City", *Administration & Society*, October 2020, https://advance-lexis-com.ezproxy.gc.cuny.edu/api/document?collection=news&id=urn:contentItem:6123-8R81-JBMY-H041-00000-00&context=1516831.

and public fund allocation in particular is something that can feel very daunting to the average person, but has an immediate and enormous impact on their daily lives. When conceptualizing the project, I wanted to see if I could push the boundaries of what other scholars had accomplished in attempting to visualize budgets, like Leuschner's approach to visualizing Berlin's budget data over time[5] (which utilized several of the same chart types that I employed throughout the narrative), but using modern best practices and animation.

This project is primarily the answer to a question I asked myself when trying to understand the budget myself: "What if the city budget's parts could be illustrated as interactive data stories instead of simple bar charts or static articles?"

Development Process

The main goals of this project were to facilitate a holistic, clear understanding of the NYC budget, raise awareness of the potential of citizen participation in allocations that affect them, and to find an effective way to visualize budgets using data storytelling. In light of these goals, my process for development was fourfold: 1) gather research; 2) conduct analysis/data cleaning and processing; 3) design a clear and effective user experience; 4) bring that design into a web experience that accomplishes the goals I set out to achieve.

This project's research phase was intensive, since budget data requires an immense amount of context that is not readily given by the Comptroller's office. As one of the project's main goals was to clarify budget practices for New Yorkers with little-to-no knowledge of budget practices (especially those unwilling to read across several site pages), the first thing I wanted to understand was where the budget's boundaries were. This information, covered in both the

---

[5] Wladimir Silantiy Leuschner, "Berlin Viz: systematic approach of visualizing governmental budget data over time," (2015), 23-33.

background of this paper as well as the site's narrative, directly contributed to the animated

visualization of the "core" of the budget (Figure 1). Using resources from the Office of

Management and Budget[6] [7] [8] and the Independent Budget Office,[9] both of which are directly

connected to the city, I constructed both the narrative and the explanatory visualizations. The

second thing I wanted to clarify was the difference between Adopted and Modified funds, which

were explained on the Comptroller's data site, but not in a way that a layperson could understand.[10]

This was vital to the creation of several of the 5-year community board budget analysis

visualizations, as there were different use cases for either. When doing a historical look-back at

budget data, the Adopted vs. Modified becomes a matter of expectation vs. reality, which was

something that couldn't be communicated in the Comptroller's codebook. With more research, I

learned how the Office of Management and Budget is involved in the budget process[11] and gained

an understanding of the budget's balancing requirement, which ended up contributing greatly to the

narrative.

       The data cleaning and processing phase took place across a number of tools, namely

Python in a Jupyter notebook, Tableau, QGIS, and Excel. Most of the data was processed in the

Jupyter notebook using pandas and numPy, but Excel was utilized for a subset of the datasets

(including using VLOOKUP to match the geoJSON community district numbers with their

respective community board for use in a choropleth), Tableau was utilized for testing data shapes,

[6] "New York City Budget Cycle," NYC Office of Management and Budget, accessed July 23, 2021, https://www1.nyc.gov/site/omb/about/new-york-city-budget-cycle.page.

[7] "Frequently Asked Questions," NYC Office of Management and Budget, accessed July 23, 2021, https://www1.nyc.gov/site/omb/faq/frequently-asked-questions.page.

[8] "Description of all Publications - OMB," NYC Office of Management and Budget, accessed July 23, 2021, https://www1.nyc.gov/site/omb/publications/description-of-all-publications.page.

[9] "The Road to Adopting New York City's Budget," accessed July 23, 2021, NYC Independent Budget Office, https://www.ibo.nyc.ny.us/IBORoadmap.pdf.

[10] "Checkbook NYC," New York City Office of the Comptroller, accessed September 27, 2021, https://www.checkbooknyc.com/site-overview.

[11] David J. Levy, "New York, New York: Lessons Learned from the Biggest Local Government," *Journal of Policy Analysis and Management* 24, no. 2 (2005): 425-27, accessed July 23, 2021, http://www.jstor.org/stable/3326219.

and QGIS was solely used to reproject the geoJSON obtained from the NYC Department of City Planning's data portal, Bytes of the Big Apple. The budget data was obtained from the Comptroller's data portal, Checkbook NYC, as five separate datasets that were concatenated in the Jupyter notebook and then processed into five of the ten final datasets.

The following dataset explanations are all a result of data processing done in that Jupyter notebook, resulting from the budget data. Adoptedtotal.csv was a result of grouping the data by the feature "Year" and using the built-in sum function to sum by the feature "Adopted." Topten-mod-total.csv also used the base budget dataset, grouping the data by "Year" and "Agency" and then performing a sum on the "Modified" feature. This was then pivoted, sorted by the highest values per year, and sliced down to the top ten, then later reduced to only five entities in the editing process. Commboards.csv is the most comprehensive Community Board-centered dataset and the basis for many of the visualizations in the second half of the explainer, created by doing a string search for the term "Community Board" in the "Agency" feature. From that subset, commsmod.csv was created using a for loop to search through every row in the "Agency" feature for a borough match, pushing the result to an empty array, then assigning the array to a new feature in the dataset titled "Borough." The resulting dataset was also used to create commboard-totals.csv, which takes the new feature, groups the data by both "Year" and "Borough," and sums the feature "Modified." After conducting this groupby, the data was pivoted, using the years as columns and the boroughs as rows, with the Modified numbers as the values of the dataset.

One of the datasets was partially constructed in Excel, but modified using Python— the dataset in question, tenmil-withpercent.csv, was retrieved from the Independent Budget Office's brochure-based explanation of the budget to be reimagined as a visual component. The data was reconstructed as "parts of a whole," and additional calculations were conducted using the same brochure for some of the units to obtain their totals. The percentage column was created in

the Jupyter notebook, but the visualizations required distillation of the numbers into their unit price and total price to visualize the total amount required for those categories. Finally, all of the calendar datasets (dates21.csv, datespartial-half1.csv, and datespartial-half2.csv) were created in Excel using autofill for simplicity and to avoid using JavaScript to save computation power.

For a five-year analysis, I had to focus on one topic so that the analysis itself wouldn't be overwhelming—when I examined the data, I noticed that community boards were given an allocation, and from prior experience attending a community board meeting, I was very interested in diving deeper into those fund allocations. Community boards themselves function as yet another level of government between residents and city government and play a vital role in the budget process. As I conducted preliminary descriptive analysis, I began to wonder if, on the whole, there were any patterns in which boards got more funding, which led to two questions: 1) "Is there a relationship between a board's location and how much it gets in funding?"; and 2) "Is there a relationship between a board office getting more funds and the difference between their adopted vs. modified budget for the fiscal year?" In order to answer these questions, I conducted a series of tests, including an ANOVA test and a Pearson's correlation matrix.

The ANOVA test turned out to be more useful than the correlation matrix, and I put the second question to the side to keep the scope of the project from becoming too large. Setting aside boroughs that did not have an adequate sample size from which to draw, I conducted the test on Brooklyn, Manhattan, and Queens community boards, which had normal distributions and a good sample size to draw from. The f-statistic was ~74.73, and the p-value was far beneath the standard alpha threshold of 0.01, so the statistics indicated that, in this case, the null hypothesis should be rejected and that the answer to the first question was that there is a relationship between location and funding. I decided not to display the results of the test in the narrative mainly to keep it accessible for an everyday audience, who might not be engaged by the results. This process did,

however, influence the spontaneous creation of a jitterplot (Figure 2) inspired by one of the preliminary descriptive analysis visualizations, which made the results of this test very clear.

Turning to narrative and design, I conducted both processes in concert. The idea behind this was to make the narrative and visualization design as seamless as possible. From a design perspective, one of my main goals was to make the site beautiful and engaging with the intent to ease understanding. Large paragraphs can be difficult to parse, so something I wanted to do very early on was highlight important key terms and phrases that would draw the audience's attention and help them synthesize the information to the point of understanding. I centered the design theme around the city's nickname, "The Big Apple," and used gala apples as the centerpiece for the site's color palette. After selecting the colors, I tested them using Susie Lu and Elijah Meeks' VizPalette[12] for accessibility and clarity in all forms of visualization (dots, lines, bars, etc). In terms of typography, I wanted web fonts that were fun, but also reduced fatigue as much as possible, allowing for ease of readability on all levels. The decision to use Capriola for headings and Asap for the body text was to remove distraction as much as possible—Capriola has so much character that I was worried it would steal the show from the visuals, so Asap, which is also a sans serif font that has much in common with Capriola without having as much showmanship, felt like a natural choice (Figure 3). The roundness of the typefaces was also very intentional, drawing attention back to the theme of the "Big Apple."

That design theme is similarly invoked throughout the narrative, using an apple as an overarching metaphor for the budget. With that theme in mind, I did multiple iterations of storyboarding for each part, highlighting key points of the high-level overview section of the narrative, and then constructing a whittled-down version of my initial concept for the five-year

---

[12] Susie Lu and Elijah Meeks. "Viz Palette." https://projects.susielu.com/viz-palette

analysis. In this section, I was surprised to find that the COVID-19 pandemic *had* had a noticeable impact on the budget, and made a note of that in the narrative—a potential explanation for the decrease in overall funding may lie in the fact that there was a massive amount of flight from the city during the early days of the pandemic, which is interesting when considering that the budget is required to be balanced within $5 million.[13] According to Jordan, there is a prevailing theory that citizens can express their discontent with government policy when they "exit the market by moving out of the jurisdiction."[14] In this case, the reduced overall revenue from the lack of ability to tax those residents who had fled the city would naturally lead to a reduction in expenditure allocation. Moreover, according to Gordon et al, "most states and localities started 2020 in very favorable fiscal condition,"[15] and NYC's 2020 Fiscal Year adopted allocation reflected that as well. This created a great moment in the narrative where I was able to create a small insight before drilling down into the next level of the analysis. From that insight, shifting the focus in the narrative onto the community boards gave me the opportunity to explore the concept from a geographic level while also creating a fun interactive moment that asked the audience to identify themselves within the position of the narrative, giving it stakes.

For the visualizations, I created basic designs that were iterated on multiple times, based on both the data and the narrative. Previous attempts by scholars to visualize other government budgets tended to have success with treemaps,[16] so I knew that using a treemap (Figure 4) to show parts of a whole was going to come into play, which manifested while I was

[13] Levy, 425.

[14] Meagan M. Jordan, "Punctuations and Agendas: A New Look at Local Government Budget Expenditures," *Journal of Policy Analysis and Management* 22, no. 3 (2003): 348, accessed July 23, 2021, http://www.jstor.org/stable/3326031.

[15] Tracy Gordon, Lucy Dadayan, and Kim Rueben, "STATE AND LOCAL GOVERNMENT FINANCES IN THE COVID-19 ERA," *National Tax Journal* 73, no. 3 (2020): 737.

[16] Francisco Cifuentes-Silva, Daniel Fernández-Álvarez, and Jose E. Labra-Gayo 2020, "National Budget as Linked Open Data: New Tools for Supporting the Sustainability of Public Finances," *Sustainability* 12, no. 11: 4551, 7-8, https://doi.org/10.3390

storyboarding the second part of the narrative, and I already had the idea to create a map (Figure 5) using the skills I learned in a previous map-focused course, so from there, it was a matter of asking what the best representation was for the other points in the process. I decided on a horizontal bar chart (Figure 6) for the initial step in part two in order to start from a high-level overview, then zoom down into the community boards by using a heatmap table (Figures 7A, 7B) as an intermediary step. I completed conceptualization for the visualizations in reverse, since the second part of the narrative was more data-driven than the first, and I wanted to make sure everything made sense before stepping back and creating the visuals for the explanations.

Many of the visualizations for the first part were not driven by large amounts of data, but were visual representations of concepts, save for the section showing how big $10 million is in the scope of the entire Adopted budget for one year. That section in particular was one of the most difficult for me to both conceptualize and code, as I was reimagining a previously-conducted concept from the Independent Budget Office[17] that was in line with the second goal of this project—attempting to see if there was a way to show this concept in an even more comprehensible way. In the end, I decided to use animation to represent this concept in a way that a static representation wouldn't accomplish (Figures 8A, 8B).

When I entered the code phase, my process became even more integrated—with the structure of my narrative and visualizations down, I was constantly iterating on both as I coded, adding or removing parts as necessary. I contemplated using a framework like Vue.js to set the project up, but felt that it would add a level of complication that would not be helpful to the process. Instead, I opted for a very basic setup of one HTML file, one CSS file, and one JavaScript file, which would work in tandem to create the site. In addition, debugging would be less

---

[17] "Understanding New York City's Budget: A Guide," NYC Independent Budget Office, July 2021, https://www.ibo.nyc.ny.us/iboreports/understandingthebudget.pdf.

complicated. My coding process generally began with pen and paper, re-sketching the intended visual and then pseudocoding the logic necessary in order to bring the visual to life. For some of the visuals, I used an Observable HQ notebook to test out code in an asynchronous environment. I had used scrollama once before, in a project for a previous class, but I had used it with Vue.js and wasn't sure how I would structure it in this project. In the end, I set it up in the initialization function, which executes once all of the data is loaded in.

Throughout the coding phase, I was constantly learning and becoming familiar with scrollama, altering the logic of the code on an as-needed basis. One of the principles I tried to implement was the idea of keeping the code DRY (don't repeat yourself) and using functions where possible to both keep the code organized and efficient. This meant that, several times through the process, I would notice a block of code that could be made DRY and edited when necessary.

Connections to Focus Area and Coursework

Throughout the project, which is primarily visualization-based, I was constantly leveraging prior experiences from previous courses. The skills and languages I learned in my Data Mining class with Professor Ashwin Satyanarayana and Working With Data class with Professor Timothy Shortell (Python in particular, using numPy and pandas) were crucial to parsing and analyzing the 800,000 total rows that resulted from combining five years of budget data, especially when it came to using Python to filter, pivot, and transform the data I needed and calculating year-over-year percentage changes for the community boards. Additionally, the skills I learned in my Data Analysis Methods class with Professor Howard Everson were crucial to answering the pivotal question in part two of the narrative—without that foundation in statistical analysis, I would not

have been able to confirm that there was a relationship between the location of a community board and its increased funding.

My course work was also based in best practices for data visualization, particularly considering how people understand information, points of data feminism (including data ethics), and accessibility standards, all gathered through my Fundamentals of Data Visualization class taught by Professor Michelle McSweeney and Data, Culture, and Society class taught by Professor Katherine Behar. Both classes approached data and visualization with care, and I attempted to use those same principles in assembling the narrative of the project. Part of my course work also entailed taking electives like my course on Geographic Information Systems taught by Professor Shipeng Sun, which was instrumental to representing the community boards as a choropleth over a period of time, and a digital humanities course taught by Professor Bret Maney, which influenced my design and narrative-constructing sensibilities to effectively deliver the humanistic narrative experience that I landed on with this project.

The visualization studio courses I took with Professor Aucher Serr on creating web visualizations were the most vital to the project's success—many of the examples we investigated in both the Interactive Data Visualization and Advanced Interactive Data Visualization studio classes structurally inspired and laid the groundwork for this project. These classes taught me most of what I know about front-end development of visualizations, particularly in constructing my data visualization process and introducing me to resources like the visualization-building package d3.js and the package scrollama.js. These classes also showed me the relevance and power of web-based visualization in terms of making information accessible to a wider audience, especially now, when web-based journalism is using these technologies and techniques. Throughout the project, I was constantly going back to concepts taught in these classes, and they were fundamental to the core of the project.

Evaluation

This project was an enormous undertaking, and there were plenty of challenges along the way. In terms of the initial vision of the project, I accomplished the primary goals of what I set out to do, but reduced some of the more ambitious elements in the pursuit of those goals. Some of the challenges were in the data side of things, but most were in the coding process, which was dependent on other packages.

In my initial plan for this project, I cast a somewhat wide net in terms of the audience and the five-year analysis section. Though fundamentally, the goal was to analyze the budget data over five years, I soon found that there were so many agencies that attempting to visualize all 150+ agencies would get overwhelming for any audience looking for a clear explanation of the budget. My advisor, Professor Serr, was extremely helpful in guiding me towards a thought process that would create a more manageable scope, and in the end, what resulted was better for the project than what I originally had in mind.

Additionally, at the start, I identified three audiences for the project: NYC residents, NYC officials who are interested in alternative methods of budget visualization, and activists based in the city who need a way to reference and advocate for changes in budget spending. While the first two audiences remained central to the project, the third audience would have needed more of an exploratory tool that would allow them to find what they need. Due to my need to reduce the scope of the project in the interest of time and ease of understanding, I removed this particular intended audience from the scope of the completed project.

In every other matter, though, the project mostly aligns with the initial vision—I successfully centralized a narrative that facilitates a holistic understanding of the NYC budget and conducts a five-year analysis. Though that analysis is not necessarily of the consequences of budget

cuts, which would have taken me far beyond the scope of the initial project, it does conduct a five-year analysis of an important part of New Yorkers' daily lives, raising awareness of this tool that residents have that can improve their quality of life.

Regarding challenges and setbacks, the data was initially troublesome due to the conflicting ways that the Comptroller's data codebook and the Independent Budget Office's explainers each defined "adopted" versus "modified." This was a major part of the project, because getting the intent of these numbers correct was crucial. Through the data analysis process and cross-referencing numbers between the IBO and the Comptroller's office, I was able to understand the difference and implement a clear explanation in the project so that future developers could more easily understand the codebook.

In the sphere of coding, there were several minor challenges that simply involved pseudocoding in an iterative manner until I understood what the obstacle was. However, one major challenge faced during the late development stages was the update of the scrollama content delivery network from version 2.2.3 to version 3.0, which completely broke my scrollytelling capabilities. This setback gave me the opportunity to rethink the way I had set up my repository, leading me to figure out that GitHub pages requires an empty ".nojekyll" file in the root repository in order to access the node_modules subfolder when using npm to manage packages. From there, I streamlined my other code packages and ended up with a far more manageable setup that won't break (as CDNs are vulnerable in precisely the way described here).

This project succeeded in providing an engaging and high-level overview of the NYC budget, in addition to raising awareness about Community Board existence and office funding—these functions have a deep impact on residents' daily lives, but because of the lack of visibility, residents don't usually even know that the boards exist, let alone that they exist to improve quality of life. Where this project could use some improvement is in terms of a wholly WCAG-compliant

accessible and mobile web experience, which I plan to rectify in the near future—as one person attempting to take this project on, I was limited in terms of having enough time to make sure that the visualizations worked and were wholly accessible at the same time. Though I did my best to keep content hierarchical and links underlined, I will revisit the project in the future to make sure that as many people can access it as possible. Despite this, my initial goals were accomplished, making this project an overall success.

Continuation of the Project

As mentioned, due to time constraints, I wasn't able to optimize the site for mobile or ensure that members of my audiences using screen readers can access the visualizations, but I plan to do so in the future, as I want the project to be as accessible to as many members of my intended audiences as possible. I also plan to make the site completely WCAG-compliant in terms of accessibility, adding descriptions to the SVGs/divs and textures to visualizations that only use color to encode values. Currently, the site is partially accessible by being structured hierarchically, having underlined links and large and legible fonts, using lists where items are listed, and using the word "interact" instead of "click" to be more inclusive for keyboard and mobile users. However, more can still be done, especially in terms of the visualization.

In the text of the site, I also outline another future endeavor: "This project will eventually have another sandbox component that allows you to fully dig into [each Community Board's spending allocations]." In the process of analyzing data, I obtained some very interesting results regarding the dollar difference between each Community Board's adopted vs. modified funding that would be a large project in its own right, and due to scoping, I set that and other large sandbox-type exploratory visualizations aside; however, I intend to create a second part to this

narrative that will function as a separate sandbox in the future that will bring my intended third audience—activists—back into the picture.

Additionally, on the coding side, I intend to refactor the main JavaScript file into a module and components—my development process required that I work in a single file in order to have the most comprehensive view of the code, especially with the use of scrollama, but I would like the project to be able to be parsed by other developers and be of more use to others attempting to visualize budget data in the future.

Overall, I truly think this project could be a useful tool for those seeking to understand the NYC budget and get involved in their communities, and I intend to make it as complete as possible so that it can make a meaningful change in the way we live as New Yorkers.
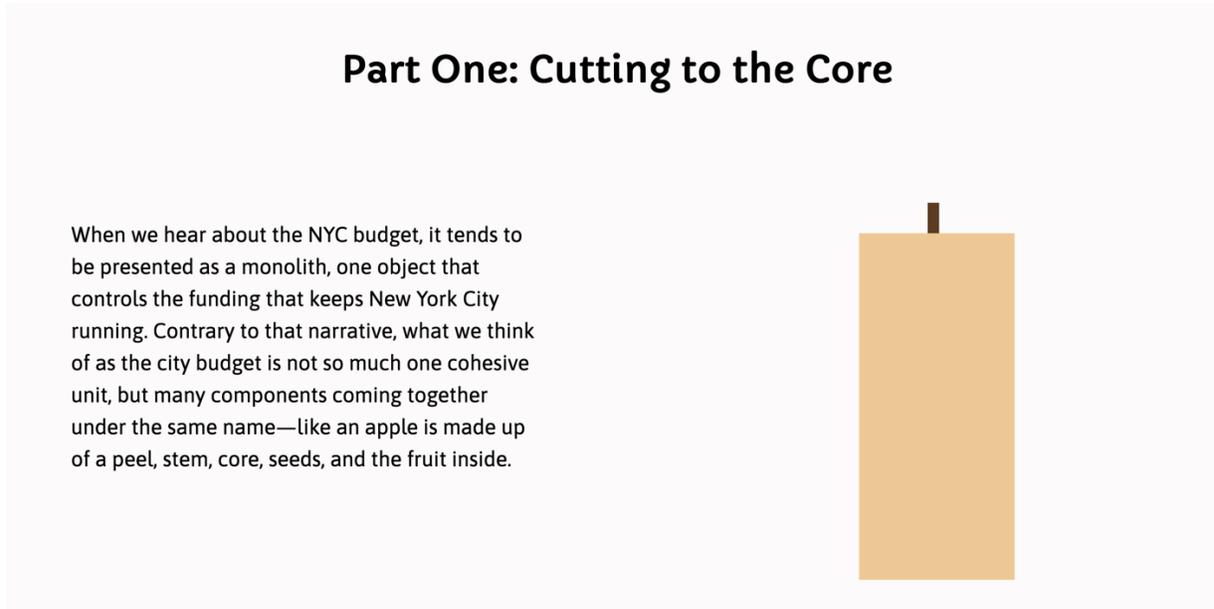
Figure 1: A screenshot of the "core" visual which shows a breakdown of the aspects of the city budget.

Figure 2: A screenshot of the "jitterplot" which shows the distribution of values for all Community Boards by borough over a period of five years.
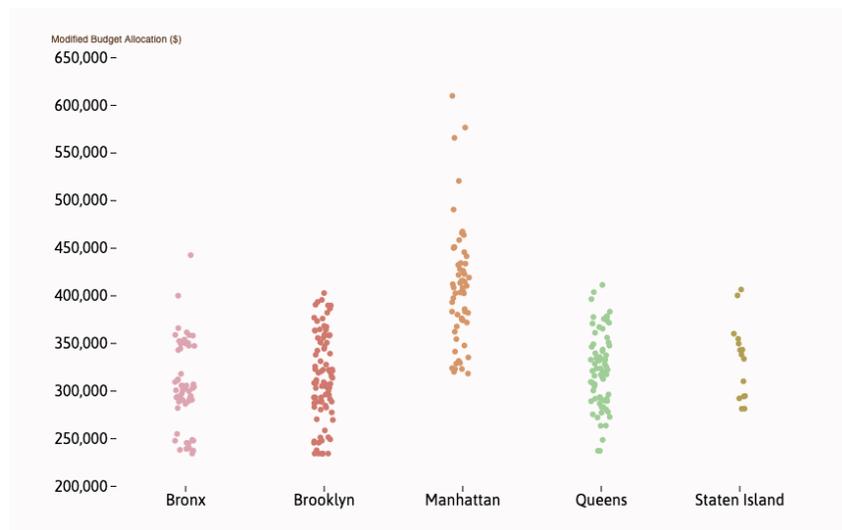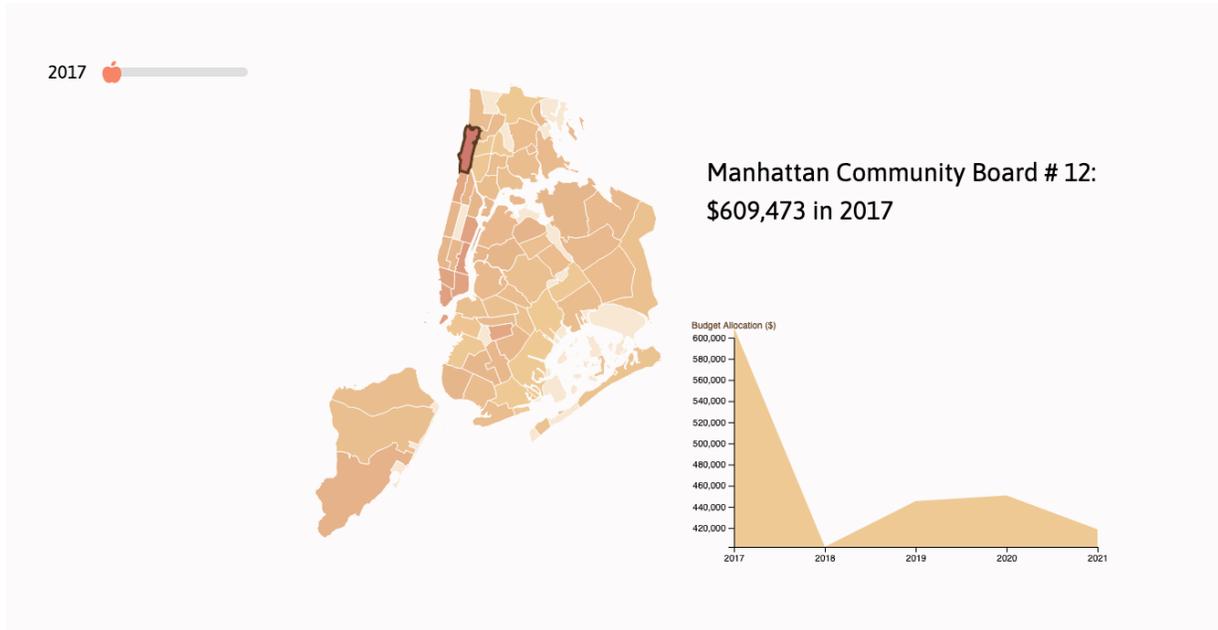
Figure 3: A screenshot of the first part of the narrative, displaying the effect of the typography in context.

Figure 4: A screenshot of one of the completed treemaps showing a breakdown of allocated funding for one community board (Manhattan Community Board # 12) in 2021.

Figure 5: A screenshot of the map interface within the narrative, with range slider, choropleth, detail pane, and supplementary area chart.

Figure 6: A screenshot of the horizontal bar chart in the second part of the narrative showing total adopted budget funds for each year.

Figures 7A, 7B: Screenshots of the resulting heatmap tables in part two. Figure 7A (top) shows the top 5 agencies by total amount of funding awarded per year, while Figure 7B (bottom) shows the same for each borough of the city.

Budget Agency	2017	2018	2019	2020	2021
Department of Education	23.5B	25.1B	27.1B	28.1B	29.2B
Department of Social Services	9.71B	9.92B	10.2B	10.6B	10.2B
Pension Contributions	9.40B	9.63B	9.95B	9.79B	9.47B
Miscellaneous	9.27B	9.26B	10.2B	10.0B	11.5B
Debt Service	5.89B	6.70B	6.38B	6.58B	8.20B

Community Board Borough	2017	2018	2019	2020	2021
Bronx	3.30M	3.35M	3.99M	4.08M	3.63M
Brooklyn	5.15M	5.18M	6.22M	6.29M	5.47M
Manhattan	4.72M	4.47M	5.21M	5.30M	4.80M
Queens	4.09M	4.13M	4.94M	5.03M	4.39M
Staten Island	917k	918k	1.08M	1.10M	961k

Figures 8A, 8B: Screenshots of the visual representation of funding in context of the total budget amount for the year 2021. Figure 8A (left) shows $10 million in that total context, while Figure 8B (right) shows the approximate amount necessary to fund 13 services for the year in that context.
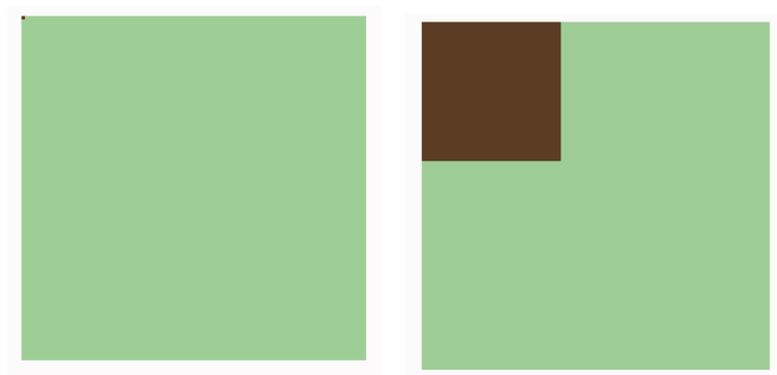
Listed in order of appearance in the file index.js

cd-boundaries-albers.geojson – community district boundaries

Features	Polygons and properties of each community district
Name	Name of file
Type	Type of data

adoptedtotal.csv – aggregated initial funding for all agencies per fiscal year

Year	Fiscal year of budget
Adopted	Total amount initially awarded to all agencies

commboards.csv – a subset of the budget data centered around Community Boards

Accrued Expense	An accounting expense recognized by the City before it is paid for
Adopted	Initially allocated funds (before any modifications during the year)
Agency	City agencies and departments; in this dataset, lists Community Boards
CBnum	Column created to match Community Board with its relevant district
Budget Code	Internal code used by the Comptroller's office for classification
Budget Name	Classification name for allocated funds
Cash Expense	Liquid funds, which can be an expense or revenue.
Department	High-level description of where the allocated funds will be used
Encumbered	Amounts committed to pay for goods and services in contract but that have not been received yet
Expense Category	Granular description of what the allocated funds will be used for
Modified	Budget allocations reflecting changes made to the Adopted budget after the start of the fiscal year
Post Adjustments	Events after the date of the financial statements which result in adjustments to the financial statements when they are next issued
Pre-Encumbered	Amount expected to spend, but for which there is no legal obligation to spend
Year	Fiscal year of budget

tenmil-withpercent.csv – describes services that can be funded with $10 million

Division	Category of selected service
Part	Amount of the service that $10 million can fund
Total	Total amount for the year
Unit	Description of service in prose
Percent	Percent of service that can be funded with $10 million for the year (calculated by dividing "part" by "total" and multiplying by 100)
Unitprice	Unit price of service (calculated by dividing 10 million by "part" for each item)
Totalprice	Total price of individual service (calculated by multiplying "unitprice" by "total" for each item)

dates21.csv – a list of all dates in 2021

Date	Dates for the calendar year 2021

dates-partial-half1.csv – dates for FY 2021 in the second half of 2021

Date	Dates for Fiscal Year 2021/Fiscal Year 2022 planning (July 1 – December 31, 2021)

dates-partial-half2.csv - dates for FY 2021 in the first half of 2022

Date	Dates for Fiscal Year 2021/Fiscal Year 2022 planning (January 1 – June 30, 2022)

topten-mod-total.csv – top ten agencies and departments by Modified funding, FY 2017-2021

Agency	City agencies and departments
2017	FY 2017 aggregated amounts (Modified allocation values)
2018	FY 2018 aggregated amounts (Modified allocation values)
2019	FY 2019 aggregated amounts (Modified allocation values)
2020	FY 2020 aggregated amounts (Modified allocation values)
2021	FY 2021 aggregated amounts (Modified allocation values)

commboard-totals.csv – aggregated community boards by borough with Modified funding as values, FY 2017-2021

Borough	Boroughs of New York City
2017	FY 2017 aggregated amounts (Modified allocation values)
2018	FY 2018 aggregated amounts (Modified allocation values)
2019	FY 2019 aggregated amounts (Modified allocation values)
2020	FY 2020 aggregated amounts (Modified allocation values)
2021	FY 2021 aggregated amounts (Modified allocation values)

commsmod.csv – budget funds aggregated by individual community board with borough attached

Year	Fiscal year of budget
Agency	City agencies and departments (community boards only in this particular dataset)
Adopted	Initially allocated funds (before any modifications during the year), aggregated by Community Board
Modified	Budget allocations reflecting changes made to the Adopted budget after the start of the fiscal year, aggregated by Community Board
Borough	Boroughs of New York City

APPENDIX C: GLOSSARY OF FUNCTIONS

Listed in order of appearance in the file index.js

init(): initializes the scroller, calls all functions

core(): creates the budget parts of whole visual

coreUpdate(): handles the budget parts of whole visual using a switch statement and helper function

updater(itemclass, text): selects each div by class and adds descriptive text

fiscYear(caldata, placement, color): creates the calendar, taking calendar dates (array), an HTML id of the visual's container (string), and a hex code (string) as input

pathMonth(t): generates the month separating lines for the calendars

involvement(): creates the force simulation for the entities involved

ticked(): updates position for involved()

tenMil(): creates the visual that represents $10 million in the context of a single year's total allocation

tenMilCategories(): creates the percentage bars in the $10 million section

calUpdate(): updates the calendar with conditional and switch statements using helper functions calColorUpdate() and calColorRange()

calColorUpdate(selecting, highlight, datecompare): recolors single rects in the calendar, taking an HTML id of the visual's container (string), an HTML id of supplementary text to be highlighted (string), and date (string in "MM/DD/YYYY" format) as input

calColorRange(selecting, highlight, first, last): recolors a continuous range of rects in the calendar, taking an HTML id of the visual's container (string), an HTML id of supplementary text to be highlighted (string), starting date (string in "MM/DD/YYYY" format) and ending date (same format as starting date) as input

involveUpdate(): updates involvement() to fade the circles in based on class using a switch statement

hbc(data, placement): creates a horizontal bar chart with 0 width, taking data (array of objects) and an HTML id of the visual's container (string) as input

hbcUpdater(placement): updates the horizontal bar chart created by hbc() using a switch statement in conjunction with hbcReuse(), takes an HTML id of the visual's container (string) as input

hbcReuse(placer): dependent on being placed within hbcUpdater(), selects each individual bar and brings to its correct width based on case in switch statement, takes HTML class of the bar (string) as input

heatTable(placement, columns, data): creates a heatmap table, taking HTML id of the visual's container (string), table columns (array), and data (array) as input

comparative(dropdown, location, details, data, default_selection): selects a dropdown, populates it with data, performs a data transformation, and then calls treemap(), taking HTML id of the select element/dropdown (string), HTML id of the treemap's container (string), HTML id of the descriptive tooltip (string), data (array) and a default selection (string) as input

treemap(wrappeddata, element, item, reusable): creates a treemap, taking data (array), HTML id of the visual's container (string), HTML id of the descriptive tooltip (string), and the HTML id of the visual's container again if the treemap is meant to be called multiple times (string) as input

getContrastYIQ(hex color): calculates RGB contrast for an element's background and returns "black" or "white" based on the amount of contrast that is optimal, takes a hexadecimal code (string in "FFFFFF" format, no #) as input

geomap(): creates Community District map and populates range slider

draw(): updates information by Community District map, populating with matching details

supplementalTrend(data): creates area graph for Community District map, taking data (array) as input

jitterplot(): creates jittered scatterplot

compareCB(): takes the dropdown values in the state and creates a comparative statement as detail information

BIBLIOGRAPHY

"Checkbook NYC." New York City Office of the Comptroller. Accessed September 27, 2021. https://www.checkbooknyc.com/site-overview.

Cifuentes-Silva, Francisco, Daniel Fernández-Álvarez, and Jose E. Labra-Gayo 2020. "National Budget as Linked Open Data: New Tools for Supporting the Sustainability of Public Finances" *Sustainability* 12, no. 11: 4551. https://doi.org/10.3390

"Description of all Publications - OMB." NYC Office of Management and Budget. Accessed July 23, 2021.https://www1.nyc.gov/site/omb/publications/description-of-all-publications.page.

"Does Participatory Budgeting Alter Public Spending? Evidence From New York City". Administration & Society. October 2020. https://advance-lexis-com.ezproxy.gc.cuny.edu/api/document?collection=news&id=urn:contentItem:6123-8R81-JBMY-H041-00000-00&context=1516831.

Ebdon, Carol, and Aimee L. Franklin. "Citizen Participation in Budgeting Theory." *Public Administration Review* 66, no. 3 (2006): 437-47. Accessed July 23, 2021. http://www.jstor.org/stable/3843923.

"Frequently Asked Questions." NYC Office of Management and Budget. Accessed July 23, 2021.https://www1.nyc.gov/site/omb/faq/frequently-asked-questions.page.

Gordon, Tracy, Lucy Dadayan, and Kim Rueben. "STATE AND LOCAL GOVERNMENT FINANCES IN THE COVID-19 ERA." National Tax Journal 73, no. 3 (2020): 733–758.

Jordan, Meagan M. "Punctuations and Agendas: A New Look at Local Government Budget Expenditures." *Journal of Policy Analysis and Management* 22, no. 3 (2003): 345-60. Accessed July 23, 2021. http://www.jstor.org/stable/3326031.

Leuschner, Wladimir Silantiy. "Berlin Viz: systematic approach of visualizing governmental budget data over time." (2015).

Levy, David J. "New York, New York: Lessons Learned from the Biggest Local Government." *Journal of Policy Analysis and Management* 24, no. 2 (2005): 424-28. Accessed July 23, 2021. http://www.jstor.org/stable/3326219.

Lu, Susie, and Elijah Meeks. "Viz Palette." https://projects.susielu.com/viz-palette

Musell, R. Mark, and Ryan Yeung. *Understanding Government Budgets: a Guide to Practices in the Public Service*. Milton: Routledge, 2020.

"New York City Budget Cycle." NYC Office of Management and Budget. Accessed July 23, 2021. https://www1.nyc.gov/site/omb/about/new-york-city-budget-cycle.page.

"New York City Passes Budget with $1 Billion Cut to NYPD amid City Hall Protest." ABC7 New York. June 30, 2020. https://abc7ny.com/defund-police-protesters-black-lives-matter-mayor-bill-de-blasio/6284680/.

"The Road to Adopting New York City's Budget." Accessed July 23, 2021. NYC Independent Budget Office. https://www.ibo.nyc.ny.us/IBORoadmap.pdf.

"Understanding New York City's Budget: A Guide." NYC Independent Budget Office. July 2021. https://www.ibo.nyc.ny.us/iboreports/understandingthebudget.pdf.