Starbucks Stores Analysis

Vansh Rajesh Jain, Sharvari Kalgutkar, Kshitij Parab, and Ankit Tripathi

University of Southern California,

Abstract. This research paper presents a data visualization project focused on drawing valuable insights about Starbucks. The dashboard aids strategic decision-making for Starbucks Executives and competitor analysts. By examining factors such as population and median income, we provide insights to pinpoint optimal locations for new stores, aligning with population trends and store density. The interactive visualizations were created using d3.js and mapbox with a user friendly website design created using bootstrap, CSS and HTML. Charts like Chloropleth Map, Bar Charts, Dot Maps, Proportional Symbol maps and Scatterplots are used to turn data into actionable insights for strategic decision-making.

Keywords: Data, Starbucks, D3.js, MapBox

1 Introduction

With Starbucks closing 7 Starbucks stores in October 2023 in San Francisco, brought into light Starbucks data centric approach in analysing store portfolios to identify locations for opening new stores or closing stores that are not bringing in enough profits and meeting customers changing need. Our visualization aims to analyse and aid strategic decision making to help understand what factors influence number of stores that Starbucks owns in a region.

Motivation: We have developed visualization that aid business analysis in order to pinpoint optimal locations for new stores based on income, population trends, and existing store density.

User studies: Starbucks Executives, Competitor Analysts

1. Starbucks Executives: Starbucks Executives can use the analysis for strategic expansion. Analysis will help them to tailor marketing and product offerings by understanding income distribution and correlating it with store locations, analyzing coffee market trends, and maintain a competitive edge.
2. Competitor Analysts: Competitor stores can use the dashboard to understand the concentration of Starbucks stores. Coffee shops can identify areas where competitors could gain market share.

   The organisation of the paper is as follows. Section 2 of the paper highlights the Related Work. Section 3 walks through the Dataset used and the cleaning performed. Section 4 talks about the Design Approach and Visualisations. Section 5 describes the Technical stack used to develop the dashboard. Section 6 outlines the conclusions that can be drawn from the dashboard. Section

seven lists down the contributions of each team member. Section 8 lists down the references that were used during the research, design and development phase of the dashboard.

2 Related Work

- Strategic Analysis of Starbucks Corporation: [1] A paper demonstrating the Strategic Analysis of Starbucks Corporation by using various scores like VRIO Analysis, Value Chain.
- Los Angeles County Starbucks Market Analysis: [2] A page demonstrating different visualization for Starbucks Analysis in the Los Angeles County by using maps like Choropleth, Dot Map on different topics. We took inspiration for our visualizations. We plotted La stores using interactive Mapbox. Instead of just plotting LA stores dot map, we also plotted normal coffee stores to understand their distribution relative to each other. We took inspiration from the starbucks location in relation to households with income under poverty line and did a similar graph for california
- Complete Data Set of Coffee Shops in Los Angeles County :[3] We took inspiration from this chart to plot coffee houses in LA. Cleaned data for removing dot(coffee houses outside LA)

3 Dataset and Data Cleaning

The analysis utilises data from data.world[4], which consists of four CSV files focused on Starbucks. The four csv files are as follows:

- Starbucks in California City Statistics
- Starbucks in California County Statistics
- Starbucks in California
- Starbucks World Statistics

The dataset contains details about starbucks across the world in different countries, cities in California, counties in California including Los Angeles county. It details the store count in each country, city, or county, accompanied by their respective latitude and longitude coordinates.For California county and city levels, median household income and median age is included for each location.

The data.world csv files regarding Starbucks stores in California required some cleaning and processing. For the LA County graph, data regarding LA county was extracted from the Starbucks in California county level data-set using python and a new smaller csv file pertaining to coffee shops in LA County was created. Median income was in string format example median income $40,000 was converted to 40000 using python spit function and then converted to integers. The latitude longitude locations were present in POINT(latitude, longitude)

format which was parsed using python. Continents were added to the world Starbucks dataset to provide additional insights about continent.

The data for normal coffee shops in Los Angeles county was obtained from a github repository that extracted coffee shops information like ratings, address, alias etc from Yelp[3]. The data contained data points outside Los Angeles County and was cleaned using geoContains() to identify if the data point lies inside the geojson boundary of Los Angeles County.

4 Approach & Visualizations

Our approach involves a global analysis of Starbucks stores, followed by a specific focus on California's cities and counties, taking into account factors such as population and median income. Additionally, our study includes a comparative analysis of market dominance, examining Starbucks and its competitors in the Los Angeles County.

The study design is as follows:

- **Home Page** - The homepage serves as an intuitive interface, offering users a quick overview of the analysis. It provides convenient links to access other graphs. Furthermore, the page incorporates seamless animations and hover effects, making it visually appealing.
- **Find A Store** - This page offers a comprehensive view of all stores in Los Angeles, enabling users to visually explore the stores within its vicinity. This functionality is implemented through the use of MapBox.
- **Country Level Performance** - Each page features a top navigation bar to ensure smooth and intuitive navigation. The color scheme across all pages aligns consistently with the Starbucks logo. On this specific page, a bar chart is employed to display the top 5 and bottom 5 countries based on the number of stores. In instances where the window size restricts the full visibility of the graph, it seamlessly transitions into a lollipop chart for optimal user experience.
- **Population Density and Starbucks** - This graph shows the relation between the number of stores and the population of a country, The graph shows an increasing trend.

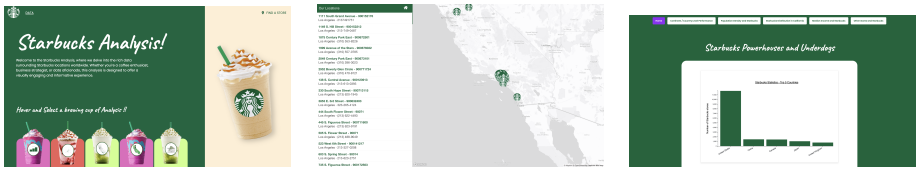

Fig. 1: Home Page, Find A Store - Mapbox, Country Level Performance

- **Starbucks in California** - This represents a county-level analysis of California utilizing a Proportional Symbol Map. The graphical representation features bubbles, where the size corresponds to the number of stores in each county.
- **Median Income and Starbucks** - This analysis contrasts the number of stores with the median income of counties in California.
- **Other Stores and Starbucks** - This graphical representation provides a clear insight into Starbucks' dominance compared to other local coffee shops within Los Angeles County. The data visually emphasizes the extent to which Starbucks prevails in the local coffee market landscape, offering a snapshot of its comparative market share and competitive standing in the region.

Fig. 2: Population and Starbucks, Number of Stores in California, Median Income and Starbucks, Starbucks and Other Local Coffee Shop in Los Angeles

5 Technical Stack

We utilized the following technologies to create this study:

- **HTML**: Used for structuring the website and creating graphs.
- **CSS**: Employed to design the website, ensuring an attractive and intuitive user experience for quick comprehension.
- **Javascript**: JavaScript was used to make the website dynamic, enabling functions like data reading, button and tab functionalities, dynamic graph creation, and tooltip implementation for enhanced interactivity.
- **D3.js**: D3.js is at the core of our graphs, providing the essential framework for creating and manipulating SVG elements.
- **Python**: Python was used for data processing and data cleaning.

6 Conclusion

The dashboard allows users to perform Starbucks stores analysis. This dashboard can help Starbucks executives to make informed business decisions, whether it's expanding operations to meet high demand or strategically closing stores to minimize losses. The dashboard is also useful for competitor coffee shops to identify areas where they could gain market share.

Through our findings, we discovered that Starbucks tends to increase the number of stores in a country as the population increases. Starbucks stores in California are in high numbers in areas with a high working population, such as San Mateo, Santa Clara, and Los Angeles. This indicates a focus on areas with a significant working population, suggesting that a major portion of Starbucks sales is driven by customers in these areas.Starbucks stores are consistently located in areas where the median income exceeds $70,000. This underscores the company's preference for areas with a higher income demographic, aligning with the assumption that areas with greater median incomes exhibit increased coffee consumption and potentially higher sales for Starbucks.

Future work could include addition of income data for each Starbucks store to analyse relation between population, median income etc. and store sales.

7 Contributions

- Vansh Rajesh Jain - Data cleaning and processing, Research and Visualization ideation, Website design , development and integration. Made Scatterplot for World population and number of Starbucks, Dot Map and Chloropleth for median income and number of stores in California, Paper writing, made Presentation.
- Sharvari Kalgutkar - Research and Visualization ideation, Bar Charts for Top and Bottom countries according to number of stores, Dot Map for Los Angeles Coffee shops and Starbucks stores, Website design, Report, Paper, made Presentation, ReadMe, video for demo.
- Kshitij Parab - Research similar work and Visualization ideation, Data cleaning and pro- cessing, Developed map that shows Los Angeles Starbucks store's locations using Mapbox, final paper writing, presented and made the presentation and wrote Readme file.
- Ankit Tripathi - Research and Visualization ideation, Data cleaning and processing, Worked on Choropleth Dot Map for number of stores in California , made Presentation, Paper writing.

8 The References Section

References

1. Nithin Geereddy:- Strategic Analysis Of Starbucks Corporation: https://scholar.harvard.edu/files/nithingeereddy/files/starbucks_case_analysis.pdf
2. Lily Maddox: – Los Angeles County Starbucks Market Analysis: https://storymaps.arcgis.com/stories/e236e40cc259480d936c480ad982a2ad
3. Vu Duong: – Complete Data Set of Coffee Shops in Los Angeles County: https://vu-d.gitbook.io/journey/data-analytics/complete-data-set-on-coffee-shop-footprints-in-los-angeles-county
4. Alice c: – Starbucks: https://data.world/alice-c/starbucks
5. GitHub - Devmode-on/Starbucks-Landing-Page: Animated Starbucks Landing Page Using HTML CSS." GitHub: https://github.com/devmode-on/Starbucks-Landing-Page