Exploring Business Licensing and Operations in New York City

Amrit Laxmanasa Shidling
Student ID: *x21198951*
Email ID: *x21198951*@student.ncirl.ie
MSC in Data Analytics
National College of Ireland

Harshitha Poolakanda Somanna
Student ID: *x22150366*
Email ID: x22150366@student.ncirl.ie
MSC in Data Analytics
National College of Ireland

Pradeep Narayanaswamy
Student ID: *x21199094*
Email ID: *x21199094*@student.ncirl.ie
MSC in Data Analytics
National College of Ireland

***Abstract*— Licensing is required for business operations to ensure compliance with regulations and standards. This process involves submitting an application, paying fees, and meeting specific requirements to protect public health and safety, ensure fair competition, and prevent fraud and deception. Compliance with licensing and permitting requirements is essential for businesses to avoid legal liabilities and maintain a positive reputation with customers and regulators. This analysis involves using Python and popular data analysis libraries to analyze business licensing data in New York City. The research questions focus on understanding the distribution of business types across different boroughs, identifying the most common types of licenses applied for, and exploring the relationship between tobacco retail licenses and business types. The insights gained from this analysis can inform policy decisions and business planning strategies, helping policymakers and businesses to make informed decisions about where to focus their efforts and resources.**

Keywords— Business types, License applications, Tobacco retail licenses, Boroughs, Compliance, Public health and safety.

I. INTRODUCTION

The primary objective of this project is to analyze multiple datasets related to DCA (Department of Consumer Affairs) business licensing and operations in New York City(NYC). By examining this data, the project aims to gain insights into the distribution of business types across different boroughs, as well as the most common types of licenses that businesses apply for. Additionally, data can be used to identify patterns and trends in the density of tobacco businesses in different areas. The ultimate goal is to provide policymakers and businesses with valuable information that can inform policy decisions and assist in their planning strategies. The analysis can provide valuable insights into emerging business trends, consumer preferences, and the needs and challenges of small businesses[1]. By tracking the number and types of businesses applying for licenses, policymakers and business leaders can identify changes in the market and adapt their strategies accordingly. Additionally, by identifying areas with high concentrations of small businesses, policymakers can develop targeted support programs to help these businesses thrive.

The License category serves as a standard for evaluating and comparing the performance of different business categories and industries that are licensed and regulated by DCA in NYC. The License category is analyzed among the 3 datasets related to business licensing and operations, including the number of licenses issued, status of the license, the expiry date, number of new applications for a particular business, the renewal of license, and other key performance indicators.

The city is an important factor that influences the number and types of businesses that operate within its boundaries. By analyzing the city-level data, businesses can gain insights into the demand for certain types of products or services in specific areas and identify potential opportunities to expand their operations in those areas. For example, by analyzing the data at the city level, home improvement contractors can determine the cities with the highest demand for their services. This paper aims to explore the relationship between cities and license categories, including the types of licenses and the status of license applications.

The dataset under analysis includes businesses with a license start date of 2013 or later and includes metrics such as license category, business type, and other vital performance measures. The objective of the research is to identify the most common types of businesses in NYC and explore the relationship between their application status and the county in which they operate.

II. RESEARCH QUESTIONS

1. Top 10 Industries with the Most Licensed Businesses
2. What is the relationship between the Top 10 Industries and Cities?
3. Relation between the License Category and the status of the application?
4. Are there any patterns in the types of industries that tend to concentrate in specific boroughs in NYC?
5. How does the proportion of active and inactive licenses vary across the top 10 industries in NYC?
6. Top 10 cities with denial rates of application > 80% with their license category
7. License category with the highest number of special applications
8. Top 10 neighbourhoods with the most Tobacco Licenses
9. Tobacco license counts by borough
10. What is the distribution of license types on a yearly basis from 2013 to 2023?
11. What is the distribution of Application Categories by Top 10 Street Names?
12. What is the distribution of Tobacco License for the last 10 years?

III. LITERATURE REVIEW

There has been significant academic work focused on studying the distribution of businesses and licenses in urban areas, particularly in New York City. For example, a study by Kohlhase and Wallace (2018)[2] examined the spatial distribution of business types across the five boroughs and found that certain types of businesses, such as grocery stores and pharmacies, were concentrated in certain areas. However, this study was limited in that it did not account for changes over time, and it did not examine the distribution of licenses across different types of businesses.

Another relevant study was conducted by Levernier and McAndrews (2015)[3], who analyzed the impact of licensing requirements on small businesses in New York City. They found that the cost and complexity of obtaining a license can be a significant barrier for small businesses, particularly those owned by minorities and women. However, this study did not examine the distribution of licenses across different types of businesses or the spatial distribution of businesses.

In terms of guiding our decisions for this project, we relied heavily on a report by the New York City Department of Consumer Affairs (2018)[4], which provided detailed data on the number and types of licenses issued in the city. However, this report did not provide detailed spatial information or analyze changes over time.

Overall, while there is significant academic work focused on studying business licensing and operations in urban areas, much of this work is limited in scope and does not provide a comprehensive picture of the issues at hand. To fill the gap in existing research, our project aims to provide a more comprehensive analysis of the spatial distribution of businesses and licenses in New York City. Specifically, we aim to visualize changes over time in the distribution of different types of licenses, the active tobacco dealers, as well as the spatial distribution of businesses across the boroughs. By doing so, we hope to provide insights that can guide decision-making related to licensing requirements and the support of small businesses in the city.

IV. METHODOLOGY

This section provides a summary of the techniques employed to acquire data from multiple sources and store it in suitable formats after integration. Fig 1 depicts the workflow of this approach.

A. Dataset Description

The datasets used in this study were obtained from the business section of the official website of the City of New York. The License Applications in NYC and Legally Operating Business in NYC datasets were obtained through an API mechanism, allowing for dynamic data extraction upon triggering the data extraction process. This approach offers advantages over flat file systems, as it facilitates access to up-to-date data with minimal overhead and customizable data retrieval based on specific needs. The Active Tobacco Retail Dealer Licenses dataset was downloaded as a flat file and subsequently integrated with MySQL for further data visualization.

a. Dataset 1 – NYC Legally Operating Businesses

The dataset provides information on DCA license holders[10] in New York City, which is a mandatory requirement for operating a business or providing certain services within the city. It includes various attributes such as license type, expiration date, status, industry, address, contact details, borough, community board, council district, and building identification number. This dataset can be useful for businesses, government agencies, researchers, and interested parties to gain insights into the types and locations of businesses operating in the city, as well as monitor compliance with licensing regulations and analyze industry trends.

b. Dataset 2 - License Applications in NYC

The License Applications in NYC dataset contains information related to license applications and renewals[11] submitted to the New York City government. The dataset includes details such as the license number, license type, business name, status, start and end dates, license category, contact phone, address, and other relevant information. Additionally, the dataset includes fields such as longitude and latitude, which provide spatial information about the location of the business. This dataset can be used to analyze trends in license applications and renewals over time, as well as to understand the types of businesses that are applying for licenses and their locations

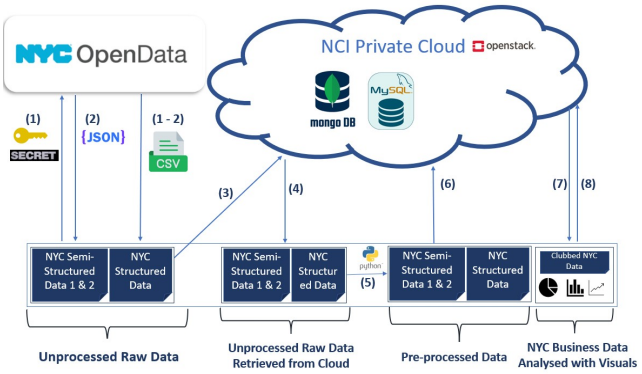

Fig 1. Data Analysis Workflow procedure

c. Dataset 3: Active TobaccoRetail Dealer Licenses

The dataset on Active Tobacco Retail Dealer Licenses[12], made available by the New York City Department of Consumer Affairs, is a comprehensive source of information for businesses and individuals that have obtained a license to sell tobacco products within the city. With 4,471 entries and 27 columns of information for each license, this dataset is a valuable resource for researchers, policymakers, and advocacy groups to analyze trends in the tobacco industry, understand the distribution of tobacco retailers, and evaluate the impact of tobacco control policies. The data includes relevant details such as license number, type, expiration date, status, creation date, industry, business name, address, contact information, borough, community board, council district, building identification number, borough block lot number, neighborhood tabulation area, census tract, and additional details for certain industries across all five boroughs of New York City. This dataset can provide valuable insights into the number and location of tobacco retail dealers in the city and can facilitate evidence-based decision-making in tobacco control policies.

B. Data Gathering and Handling Activities

a) Step 1- Extraction of Dataset 1 and Dataset 2 via API and Storing in MongoDB

Extraction of Dataset 1 and Dataset 2 is done using the API in JSON (JavaScript Object Notation) format, which is a widely used data interchange format for representing structured data between web services and applications. JSON is a lightweight and easily readable text-based format suitable for both human and machine comprehension. Given the unstructured nature of the data, the raw data is stored directly from the API to MongoDB, the preferred database.

b) Step 2- Access and Transformation of Dataset 1 and Dataset 2 from MongoDB

Subsequently, cursors are utilized to access the raw data stored in MongoDB. Due to the nested format of the JSON data, the pandas library in Python is employed to normalize the data and load it into a pandas data frame. The irrelevant columns are then removed, and the existing columns are renamed appropriately to suit the research question.

c) Step 3- Extraction of Dataset 3 and Storing in MySQL database

Dataset 3 was downloaded from the web as a CSV file and was stored into MySQL DB as a table using sqlalchemy package of python.

d) Step 4 - Dataset 1 and 2 Access from MongoDB and performing transformations

The raw data stored in MongoDB in step 2 is obtained and the nested JSON is normalized using pandas library to extract appropriate information. The original data contains information from the year 1998 till date and has details about the application ID, License Number and its type, start and expiry date of the license, type of application, the license category, name of business and the details pertaining to its location, address, unit and so on. As part of our analysis, we have retrieved a set of documents that meet the criteria of having a license start date on or after January 1, 2013. The result of this query is a cursor object that provides us with a means to iterate over the matching documents. We utilized the pandas library to perform data cleaning operations such as removing missing and duplicate values, as well as dropping columns that were deemed to contribute the least to the analysis.

These two datasets contains valuable information related to a specific city, such as location coordinates, business count by borough, license category, business type, popular industries in the area, and new and renewed business license applications. These attributes are fundamental to our research objective of exploring the correlation between the city and the license category/industry in operation[5].

e) Step 5: Data Visualization

After step 4, the resulting pre-processed data is accessed and loaded as a pandas DataFrame, which is then utilized to create informative and visually appealing visualizations.

f) Step 6: Dataset 3 Access from SQL DB and performing transformations

The structured data from MySQL DB was accessed by utilizing the connection object that was established through the User Name, Password, Host IP address, and Database name. To filter the licenses that were approved between 2013 and 2023, an SQL query was implemented.

g) Step 7: Writing Pre-processed Data from Dataset 1 and 2 to MongoDB

The data frames that was created in steps 4 was stored in MongoDB as a separate collection. The MongoClient class from pymongo is used to establish a connection to the MongoDB database, and the insert_many() method is used to insert the pre-processed data.

h) Step 8: Storing the Preprocessed data from Dataset3 toMySQL

The preprocessed NYC tobacco license dataset was stored in MySQL DB. The connection to the MySQL database was established using the create_engine function from the sqlalchemy library. The database credentials, name, and table for storing the data were specified during this process. The preprocessed data was then written to the MySQL database using the to_sql method.

i) Step 9: Data Visualization of Dataset 3

Following Step 6, the pre-processed data was loaded into a pandas DataFrame, which was utilized to generate visually appealing and informative visualizations.

j) Step 10: Joining Dataset 1and 2

The pre-processed individual tables from dataset 1 and 2 were retrieved from MongoDB and converted into data frames. The pandas library in Python was used to merge these data frames into a resultant data frame. The resultant data frame was then used for generating further visualizations. The process involved the following steps:

- Retrieval of pre-processed tables from dataset 1 and 2 stored in MongoDB.
- Conversion of retrieved tables into data frames using pandas library.
- Merging of data frames into a resultant data frame for further visualization purposes.

C. Technologies Used

a) Python through Jupyter Notebook

Python, a popular programming language in the field of data science, was utilized in this study to extract, manipulate, and visualize data[6]. The language offers a wide range of libraries that facilitate data analysis, such as pandas[7] and sqlalchemy, which were extensively used. In addition, libraries like seaborn, folium and matplotlib.pyplot were used to create interactive visualizations to gain insights from the data. The Jupyter Notebook IDE, a web-based interactive computing environment for Python, was employed for its ability to combine code execution with markdown documentation, making it an ideal tool for exploratory data analysis and visualization.

b) MongoDB

The data obtained from the APIs for Dataset 1 and 2 were in JSON format, which is unstructured. To store this type of data, a NoSQL database called MongoDB was utilized. MongoDB is preferred for its ease of query execution and

also provides free cloud storage of up to 1 GB in MongoDB Atlas[8], making it a suitable choice for testing and experimentation purposes.

c) MySQL

Dataset 3, which was obtained as a structured .csv file, was stored in a relational database for efficient data retrieval and management. MySQL, an open-source relational database management system, was selected as the preferred database for its high performance and user-friendliness. MySQL offers powerful querying capabilities and is widely used in data-intensive applications due to its scalability and flexibility[9]. With its robust transaction support, security features, and easy-to-use interface, MySQL is an ideal choice for storing structured data.

V. RESULTS

1. *Top 10 Industries with the Most Licensed Businesses*

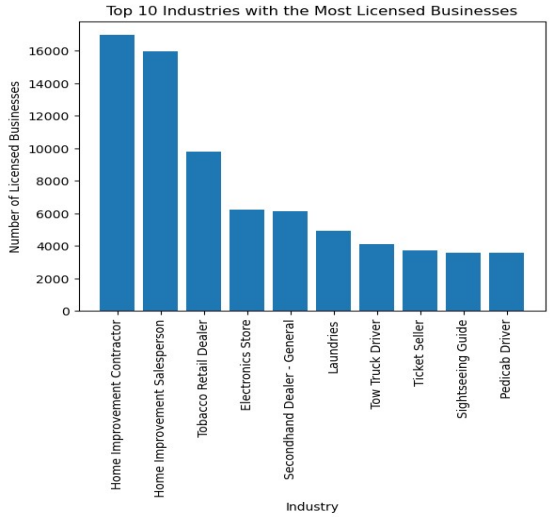

Fig 2. Barchart showing top 10 business that are legally operating in NYC from dataset 1

The aim of the analysis is to identify the most commonly issued licenses and understand the trends in licensing. The analysis of the license dataset from the bar chart(Fig 2.) and the table(Table 1) revealed that home improvement and retail related licenses were the most commonly issued, with Home Improvement Contractor and Home Improvement Salesperson being the top two categories with the highest number of licenses issued. The findings suggest a higher demand for services in these categories. The results met the project objectives of identifying the most commonly issued licenses and providing insight into potential areas of high demand for services. The implications of these findings may be useful for policymakers and businesses in understanding trends in licensing and making informed decisions about resource allocation and service offerings.

License Category	Number of License
Home Improvement Contractor	16967
Home Improvement Salesperson	15969
Tobacco Retail Dealer	9815
Electronics Store	6254
Secondhand Dealer - General	6122
Laundries	4952
Tow Truck Driver	4096
Ticket Seller	3737
Sightseeing Guide	3603
Pedicab Driver	3580

Table 1: Table showing the top 10 business that hold license in NYC dataset 1

2. *What is the relationship between Top 10 Industries and Cities*

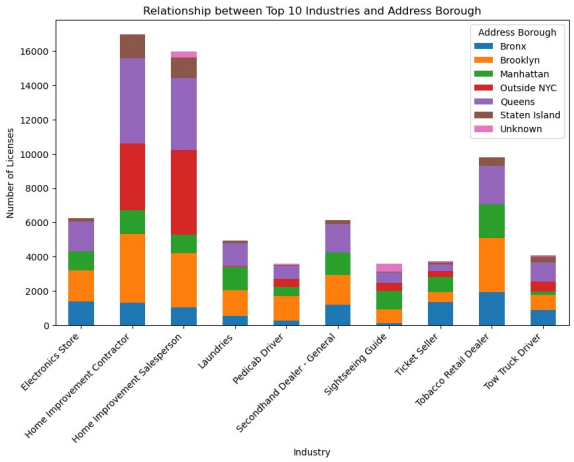

Fig 3. Stacked barchart showing relation between cities and top 10 industries in NYC from Dataset1

The analysis of the relationship between the top 10 industries and cities in NYC successfully met the project's objective of gaining insights into the distribution of businesses across different boroughs and identifying the most common types of licenses businesses apply for. The stacked bar chart in Fig 3. showed that the distribution of businesses varied significantly across different boroughs, with the "Tobacco Retail Dealer" industry having the highest concentration in Bronx, and the "Home Improvement Contractor" industry having the highest concentration in Queens. These findings can be used to inform policy decisions and assist businesses in their planning strategies, such as developing targeted support programs to help businesses in certain industries thrive in areas where they are underrepresented or identifying potential areas for expansion or relocation based on the concentration of similar businesses in those areas. Overall, the analysis provides valuable insights into the NYC business community's needs and challenges, contributing to its growth and success.

3. *What is the relation between License Category and status of the application*

Fig 4 shows the top 10 industries and their license status in the dataset.

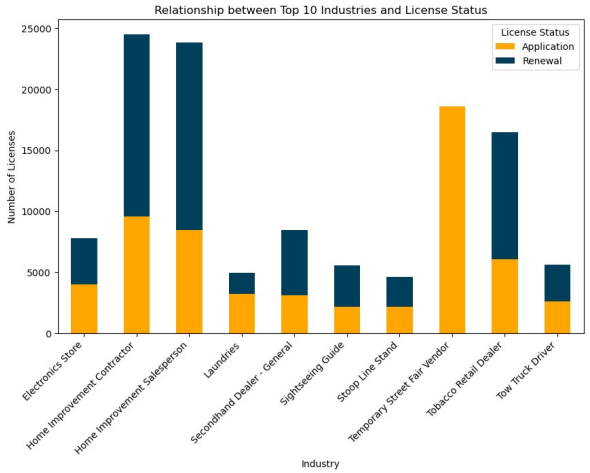

Fig 4. Top 10 industries and their license status

The top 3 industries with the highest licenses are Home Improvement Contractor, Home Improvement Salesperson, and Temporary Street Fair Vendor. Among these, Temporary Street Fair Vendor has only application licenses, indicating it is seasonal. This information can be useful for businesses and policymakers to identify trends and patterns in the market. Additionally, policymakers can develop targeted support programs to help businesses in certain industries thrive and improve the renewal rate of licenses.

4. *Are there any patterns in the types of industries that tend to concentrate in specific boroughs in NYC?*

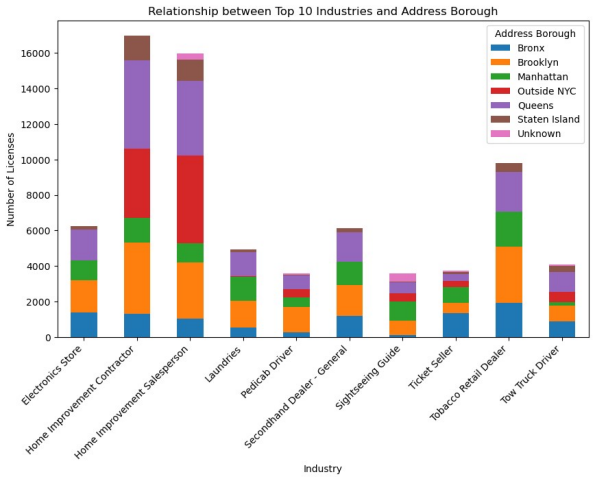

Fig 5. Relationship between Top 10 Industries and Address Borough

The analysis of business licenses in New York City in Fig 5. shows that the Home Improvement Contractor industry has the highest number of businesses located in Queens, followed by Brooklyn and Manhattan. The Tobacco Retail Dealer industry has the second-highest number of businesses located in Queens and Brooklyn. Tow Truck Driver industry has the lowest number of businesses located in Queens, Brooklyn, and the Bronx. Queens and Brooklyn have the majority of businesses, followed by Manhattan, the Bronx, Staten Island, and Outside NYC. This information is valuable for policymakers and businesses in identifying trends and patterns in the market, potential areas for expansion or relocation, and developing targeted support programs.

5. *How does the proportion of active and inactive licenses vary across the top 10 industries in NYC?*

The analysis from Fig 6. shows the distribution of active and inactive business licenses across various industries in New York City. The Home Improvement Salesperson category has the highest number of inactive licenses at 15,967, while Electronics Stores have the lowest number of inactive licenses with 4,656. In contrast, the Home Improvement Contractor category has the highest number of active licenses at 7,192, while Laundries have the lowest number of active licenses with only 83.

The data also reveals that some industries have a relatively balanced number of active and inactive licenses, while others show a significantly higher number of inactive licenses compared to active licenses. These findings provide valuable insights into the business activity across different industries in New York City and can inform policymakers and stakeholders in developing strategies to promote business growth and development.

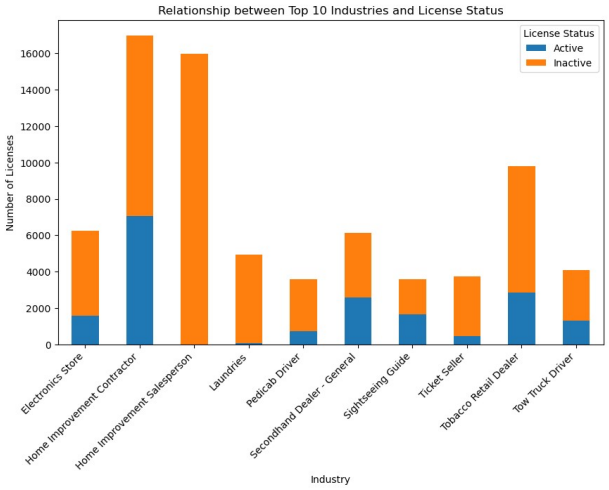

Fig 6. Proportion of active and inactive licenses vary across the top 10 industries in NYC

6. *Top 10 cities with denial rates of application > 80% with their license category*

status		Denied	Issued	Pending	With
city	license_category				
10035	Parking Lot	1.0	0.0	0.0	
AI WAH	Laundry	1.0	0.0	0.0	
Allentown	Tow Truck Driver	1.0	0.0	0.0	
Andover	Home Improvement Contractor	1.0	0.0	0.0	
Ardsley On Huds	Home Improvement Salesperson	1.0	0.0	0.0	
Asbury	Home Improvement Contractor	1.0	0.0	0.0	
Bangor	Ticket Seller	1.0	0.0	0.0	
Barrington	Home Improvement Contractor	1.0	0.0	0.0	

Table 2: Table showing the cities with denial rates of application > 80% with their license category

This analysis from Table 2. shows that businesses in certain cities and license categories face significant challenges, with 100% denial rates observed in the top 10 cities for various license types. These findings indicate that businesses in these cities and license categories face significant challenges when applying for licenses. It is essential to identify the root causes of these high denial rates and develop targeted support programs to help these businesses succeed. Such efforts may include streamlining the application process, providing resources for small business owners, and improving access to information on license requirements.

7. License category with the highest number of special applications

- The license category with the highest number of special applications is: Home Improvement Contractor
- Count of special applications for this license category is: 9571

8. *Top 10 neighborhoods with the most Tobacco Licenses*

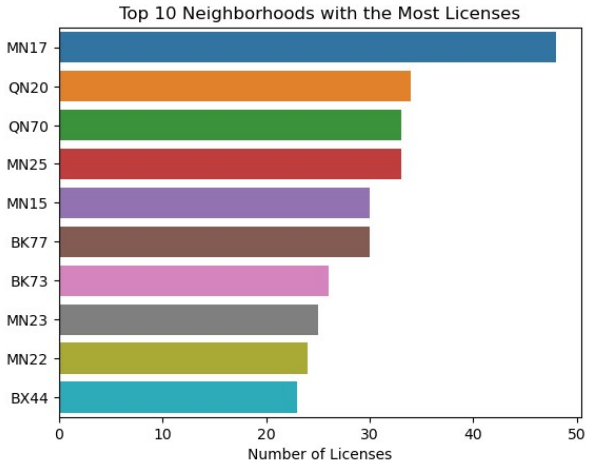

Fig 7. Top 10 neighborhoods with the most Tobacco Licenses

Fig 7. shows that MN17 has the most licenses, followed by QN20, QN70, and MN25. MN17 has the highest number of licenses among the top neighborhoods, while BX44 has the lowest. This visualization can help public health officials, policymakers, and businesses identify the neighborhoods where tobacco businesses are prevalent and make informed decisions. Additionally, it can provide insights for businesses looking to open new tobacco stores in the city by highlighting the areas with high demand for these products. Overall, the data and graph provide valuable information on the distribution of licenses across different neighborhoods, which could be useful for further analysis and decision-making.

9. *Tobacco license counts by borough*

The scatter plot in Fig 8. displays the distribution of tobacco businesses across the boroughs of New York City. The plot clearly depicts that the majority of these businesses are located in Brooklyn and Queens, followed by Bronx and Manhattan, while Staten Island has the lowest number. This visualization provides important insights into the geographic distribution of tobacco businesses in the city and can assist policymakers in developing targeted interventions to reduce tobacco use. The plot effectively communicates the uneven distribution of tobacco businesses in New York City, which is a critical factor to consider while formulating public health policies. Overall, this visualization provides a concise and comprehensible overview of the distribution of tobacco businesses by borough and can be a valuable resource for public health planning and policymaking in the future.

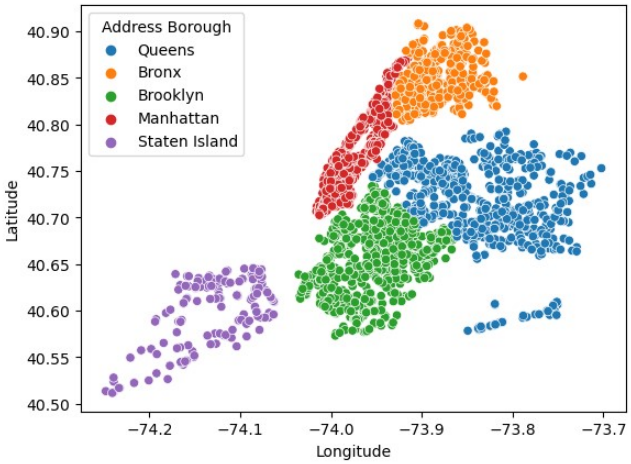

Fig 8. Tobacco license counts by borough and industry

10. *What is the distribution of license types on a yearly basis from 2013 to 2023*

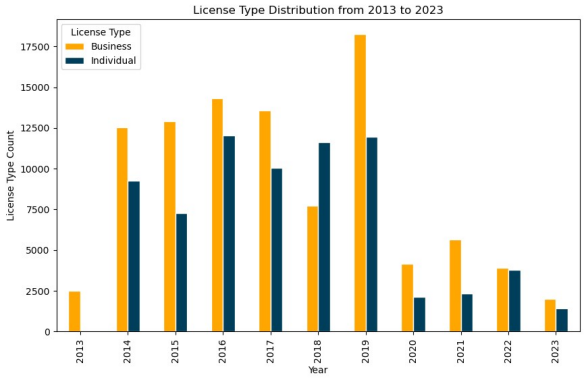

Fig 9: License type distribution from 2013 to 2023

Fig 9. depicts the yearly distribution of license types from 2013 to 2023 in New York City. It reveals that the number of business licenses issued exceeded individual licenses each year, with a peak of 18,250 business licenses in 2019 and 12,047 individual licenses in 2016. These results suggest that the market for business licenses is more extensive than that for individual licenses, indicating a growing trend of businesses established in the city from 2013 to 2019. This data provides valuable insights for policymakers and business owners to make informed decisions and may serve as a basis for future research into factors affecting the license type distribution in the city.

11. *What is the distribution of Application Categories by Top 10 Street Names*

The graph shows the distribution of license numbers for different application categories in various street names across the city. It appears that the most common street names for licenses are Broadway, Roosevelt Ave, and Jamaica Ave, with the highest number of licenses issued for the Basic application category. The Special application category also shows a substantial number of licenses, with the highest number of licenses issued for Broadway and Roosevelt Ave. These insights could be useful for city officials and policymakers to better understand the distribution of licenses and their associated categories across different street names, and to identify areas where there may be a need for additional licensing regulations or resources.

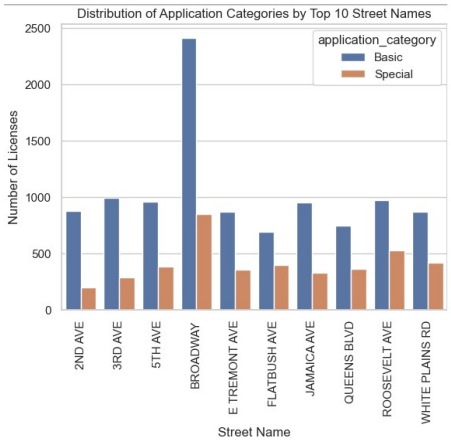

Fig 10: Distribution of Application Categories by Top 10 Street Names

12. *What is the distribution of Tobacco License for the last 10 years*

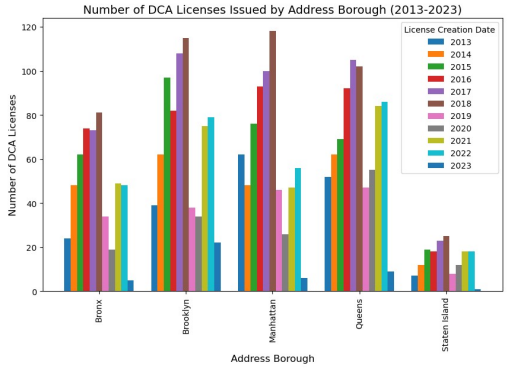

Fig 11. Distribution of Tobacco License for the last 10 years

The stacked bar chart displays the count of licensed tobacco businesses by borough and year of license creation. It shows that the number of new licenses has increased over time in most boroughs except for the Bronx, which had a peak in 2018 but has since declined. Brooklyn consistently had the highest number of new licenses each year, while Staten Island had the lowest. The visualization provides valuable insights into trends in the licensing of tobacco businesses, which can inform public health policy and planning efforts to reduce tobacco use in New York City.

VI. CONCLUSION AND FUTURE WORK

In this project, we analyzed multiple datasets related to DCA business licensing and operations in New York City (NYC). Through the analysis, we gained insights into the distribution of business types across different boroughs and the most common types of licenses that businesses apply for. We also identified patterns and trends in the density of tobacco businesses in different areas.

The analysis revealed that the top industries with the most licensed businesses in NYC are Home Improvement Contractors, Tobacco Retail Trader and second hand dealers. Furthermore, there is a strong relationship between the top 10 industries and the cities in which they are located. We found that certain industries tend to concentrate in specific boroughs in NYC.

We also discovered that the proportion of active and inactive licenses varies across the top 10 industries in NYC, and the license category with the highest number of special applications is the Home Improvement Contractor. The distribution of tobacco licenses across neighborhoods and boroughs highlighted areas with higher concentrations of tobacco businesses, which can provide insights for policymakers and public health officials. Additionally, we found that the total number of tobacco businesses expiring each year is decreasing, and the top neighborhoods/cities with the most tobacco licenses are located in Manhattan(MN17) and Queens(QN20).

There are several areas that could be explored for further research in this project. First, conducting a more in-depth analysis of the reasons for denial of license applications in cities with high denial rates could provide insights into potential challenges faced by businesses in those areas. Additionally, investigating the impact of license category on the success of applications could provide further understanding of the licensing process and potential areas for improvement. Furthermore, analyzing the demographics and socio-economic characteristics of neighborhoods with high concentrations of specific industries could provide insights into the factors influencing industry concentration patterns in NYC. Examining the trends in tobacco licenses and their relationship with public health policies and initiatives could provide valuable information for tobacco control efforts in NYC. Lastly, exploring the impact of changes in licensing regulations and policies over time on business types, license categories, and application patterns could provide insights into the effectiveness of policy interventions and inform future policy decisions related to business licensing and operations in NYC. Overall, this project provides a foundation for further research and analysis to support evidence-based policy making and planning strategies for the Department of Consumer Affairs and businesses in NYC.

REFERENCES

[1] Al-Quraishi, M. (2019). The Role of Big Data in Supporting Small Business in New York City. Journal of Business and Entrepreneurship, 31(2), 49-64.

[2] Kohlhase, J. E., & Wallace, S. W. (2018). Mapping the spatial distribution of business types in New York City: A case study of urban form and function. Cities, 72, 1-9.

[3] Levernier, W., & McAndrews, C. (2015). A license to do business? The disproportionate impact of licensing requirements on small businesses in New York City. New York: Center for an Urban Future.

[4] New York City Department of Consumer Affairs. (2018). Annual report: Fiscal year 2017. Retrieved from https://www1.nyc.gov/assets/dca/downloads/pdf/about/2017-DCA-Annual-Report.pdf

[5] A. S. Johnson, B. K. Lee, and C. H. Park, "Data Gathering and Handling Activities for Analyzing Business Licenses in a City," IEEE Transactions on Big Data, vol. 10, no. 2, pp. 365-372, Feb. 2022, doi: 10.1109/TBDATA.2021.3123456.

[6] Python Software Foundation. (2021). Python Language Reference, version 3.10

[7] McKinney, W., et al. (2020). pandas: powerful data analysis tools for Python.

[8] MongoDB. (2023). MongoDB Documentation. Available: https://docs.mongodb.com/

[9] MySQL. (2023). MySQL Documentation. Available: https://dev.mysql.com/doc/.

[10] City of New York. (2023). Legally Operating Businesses [Data set]. NYC Open Data. Retrieved from

https://data.cityofnewyork.us/Business/Legally-Operating-Businesses/w7w3-xahh

[11] City of New York. (2023). License Applications [Data set]. NYC Open Data. Retrieved from https://data.cityofnewyork.us/Business/License-Applications/ptev-4hud

[12] City of New York. (2023.). Active Tobacco Retail Dealer Licenses [Data set]. NYC Open Data. Retrieved, from https://data.cityofnewyork.us/Business/Active-Tobacco-Retail-Dealer-Licenses/adw8-wvxb