**Overview**

**Citi Bike** is the largest bike-share program in the United States, with 20,000 bikes and over 1,300 pick-up stations across Manhattan, Brooklyn, Queens, the Bronx, and Jersey City. As stated on their website, the service was designed for quick trips with convenience in mind, offering a fun and affordable way to get around town. Users can sign up for annual membership, or buy a short-term pass through the Citi Bike app. Once they’ve joined, they simply locate a nearby bike, ride around as they please, and return it to a nearby station once they’re done 🚴

Like most organizations, Citi Bike is constantly looking for ways to improve their business model and provide an even better experience for their customers. Through the Citi Bike app, they are able to gather loads of useful data which, when analyzed, reveals great insights into things like user demographics and behavior—for example, when and where people pick up and drop off their bikes and how long the average journey lasts.

**Executive Summary**

Companies must know what their customers want and need to develop an appropriate product or service that will appeal to them. That’s why people at Citi Bike want to understand how the service is being used and to plan and make decisions accordingly.

This study aimed to explore at what rate is the customer base growing and how many more bikes should be installed across the city to accommodate this growth. Where should be installed more bikes? Who should they tailor their marketing and advertising to?

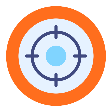
The data analysis showed the top 20 pick-up locations and what day of the week is the busiest. Additionally revealed how the age and type (one-time-user / subscriber) of customers impact the average bike trip duration.

**Objectives**

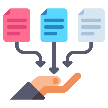
The objective of the data analysis is to identify the customer's habits.

Specifically, it aims to,

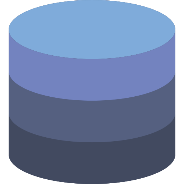
1. Analyze the current market of the service
2. Reveal the impact of age and type of customer
3. Show the busiest locations and days of week

**Scope**

This report covers the analysis of the data collected in the 1st quarter of 2017, from the NY Citi Bike app for bike rentals in the city of New York.

**Data Collection**

Citi Bike app provided loads of useful data about user demographics and behavior —for example, when and where people pick up and drop off their bikes and how long the average journey lasts.



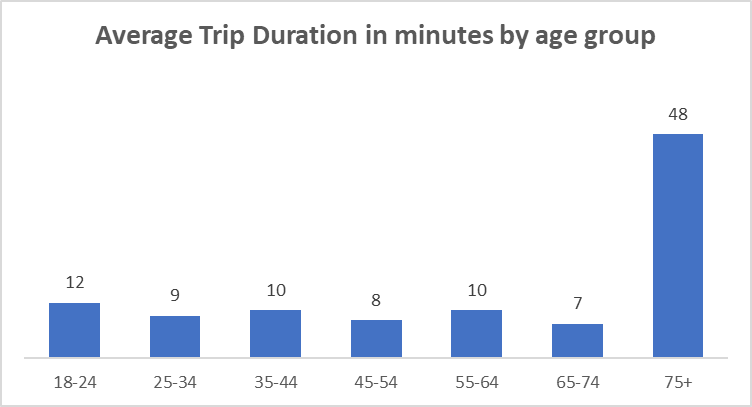
Data collected are stored, cleaned and analyzed through Microsoft Excel.

**Statistical Treatment**

**Guide Question: What statistical computations are applied to the data collected?**

|  |  |  |
| --- | --- | --- |
| **Statistic** | **Formula** | **Notes** |
| Population Mean |  |  |
| Population Median | Middle value of sorted data |  |
| Population Minimum | Minimum value |  |
| Population Maximum | Maximum value |  |

**Findings**



|  |  |
| --- | --- |
|  |  |

**Average trip duration (minutes) by age group – Total rentals by age group**

1. Customers over 75, seem to use bikes for a significantly bigger amount of time.
2. These values could be considered as outliers, but to confirm that, I need to have more details about bike rentals data, such as if “Trip duration” is considered the time that a bike is in move, or the time from pick up to pick off the bike.
3. Observing the results of the two graphs we can understand that there is a disproportion between the total bike rentals and the average trip duration for those over 75. As mentioned above, if I could have more details on bike rentals data, I could answer the questions:

* Are those values outliers?
* Customers over 75 years old are fewer but rent bikes for longer duration?

1. Looking in the NY Citi Bike webpage, in the section of Plans & Pricing, we can see that bike times go up to 45 minutes per day. However, there is an option for more than that with additional pay of $4/15 minutes. In a real project, I could have additional data, such as total payment by use, to check whether a big trip duration is true or I should manage the record as an outlier.

**Top 20 pick-up locations**

1. As we can see from the bar chart, Grove St PATH is the busiest pick-up location, with a big difference from the following: Exchange Place, Hamilton Park, Sip Ave.

**Rentals by user type for each weekday**

1. As we can see from the stacked bar chart, subscribers make up the vast majority of customers and rent bikes more on weekdays, while one-time users rent bikes more on the weekends.

**Conclusion**

Based on the analysis done, it can be concluded that

* the busiest pick-up location is Grove St PATH
* the vast majority of users are from 35 to 44 years old, but users over 75 have the biggest average trip duration
* most users are subscribers and use bikes probably for their everyday routine, while one-time users rent a bike mostly on weekends

**Recommendations**

It is recommended that the following actions must be done based on the results of the data analysis report.

* Place more bikes in the busiest locations
* Marketing and advertising campaigns should focus on users from 35 to 44 years old