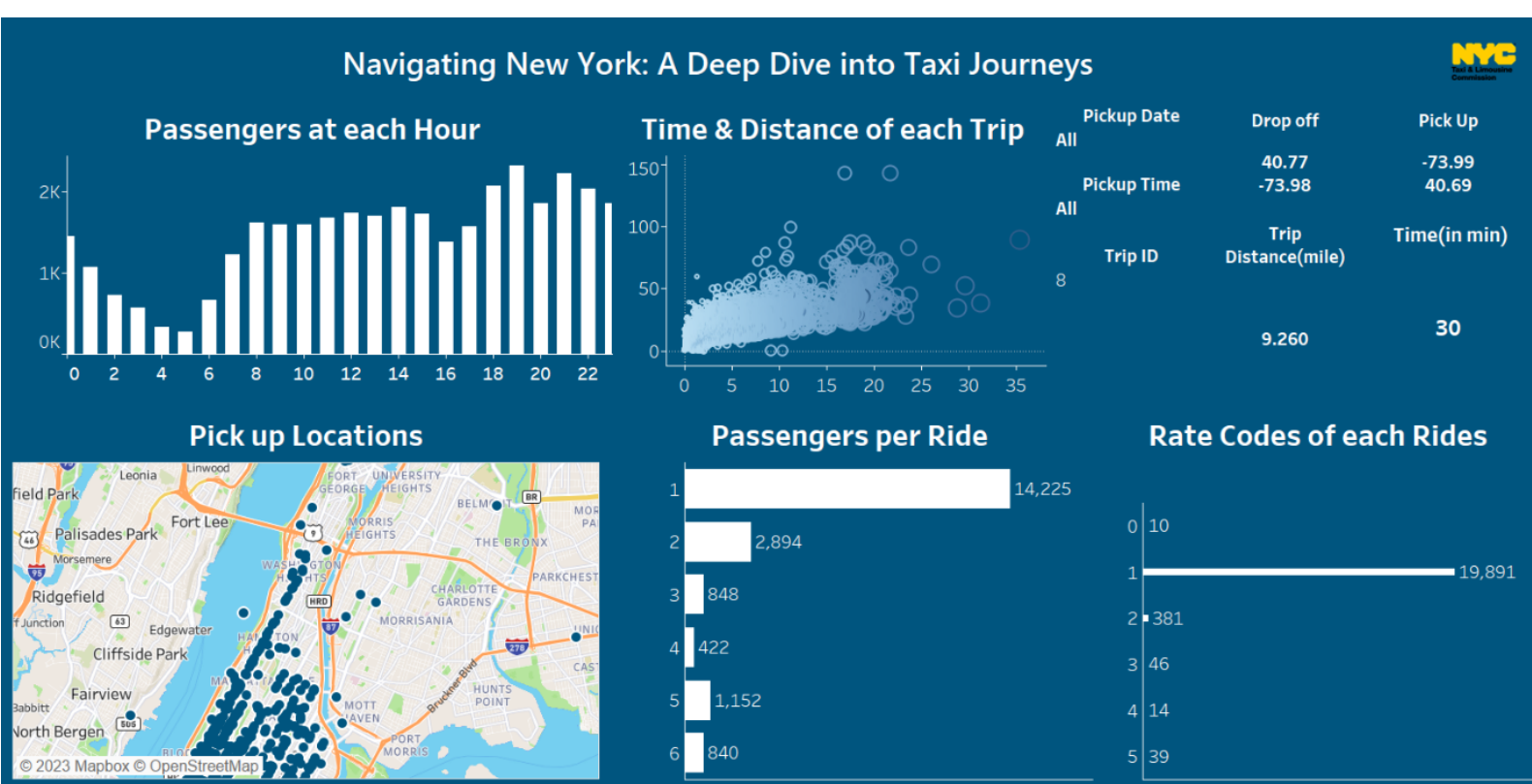# PROJECT: 2
Navigating New York: A Deep Dive into Taxi Journeys

Chandrakanth Naidu Gonuguntla

We used Tableau Software for this project, the data we used for this project is real-time data collected from the NYC Taxi Commission website.

DASHBOARD:

Sample data:

This data describes the usage of taxis in New York City in May 2013.

	Trip_ID	medallion	hack_licens	vendor_ic	rate_code	passenger_	trip_dista	pickup_lo	pickup_la	dropoff_l	dropoff_l	pickup_tin	pickup_date	dropoff_ti	dropoff_dat	trip_time
0	3B1A31779	AED0496C9	VTS	1	1	1.34	-73.9823	40.77282	-73.9862	40.75874	0	01-05-2013	0	01-05-2013	8
1	A38D671A1	47AC05DE1	VTS	1	5	2.5	-73.9775	40.74531	-73.9556	40.77421	9	05-05-2013	9	05-05-2013	8
2	7B201D72A	F2020C3F1&	VTS	1	1	1.59	-73.9622	40.77911	-73.9507	40.79422	12	05-05-2013	12	sni	7
3	AB4CE77E2	3B4701521:	VTS	1	1	1.72	-74.0055	40.74593	-73.9905	40.76638	12	05-05-2013	12	05-05-2013	4
4	198CE3EF5!	BF3544AD0	VTS	1	5	1.9	-73.9938	40.72091	-74.0026	40.73966	2	05-05-2013	3	05-05-2013	10
5	6262E716C!	38072A275!	VTS	1	5	3.04	-73.959	40.78099	-73.9903	40.75178	11	05-05-2013	12	05-05-2013	11
6	BE0D66D2F	4F1AC06CD	VTS	1	1	1.8	-73.9754	40.77704	-73.96	40.77052	11	05-05-2013	11	05-05-2013	9
7	B82032932!	3BB26E7CB	VTS	1	6	3.15	-73.7898	40.64656	-73.9997	40.72702	9	05-05-2013	10	05-05-2013	15
8	C0ABDE211	0B70B06FD	VTS	1	5	9.26	-73.9894	40.68881	-73.9831	40.76969	10	05-05-2013	10	05-05-2013	30
9	AB4F8FB61!	E107CC6BB	VTS	1	1	2.05	-73.9642	40.75728	-73.984	40.7344	10	05-05-2013	11	05-05-2013	7
10	068BBBB50	CFE3E6148:	VTS	1	1	0.4	-73.9548	40.77767	-73.9548	40.77767	12	05-05-2013	12	05-05-2013	3
11	E438441E8!	5FFC09022&	VTS	2	1	16.87	-73.9801	40.75868	-73.7891	40.64166	12	05-05-2013	12	05-05-2013	27
12	383E0C5CD	4D5DD72C(	VTS	1	6	1.99	-73.9695	40.78463	-73.9733	40.76091	10	05-05-2013	10	05-05-2013	8
13	0CAC8E999	DF618A0B8	VTS	2	1	17.39	-73.7913	40.64565	-73.9726	40.75588	8	05-05-2013	9	05-05-2013	34
14	9B5BE818E	4D85F9862(	VTS	1	3	8.35	-73.9539	40.78777	-73.864	40.76948	13	05-05-2013	13	05-05-2013	19
15	C41A7FA97	1DC96DC34	VTS	1	6	0.39	-73.9859	40.73253	-73.9911	40.73502	10	05-05-2013	10	05-05-2013	4
16	BBD41D16E	160C474EC!	VTS	1	1	2.13	-73.9822	40.74504	-74.0021	40.71931	9	05-05-2013	10	05-05-2013	23
17	AE6D16985	287758B2F!	VTS	1	1	2.83	-73.9631	40.80268	-73.9632	40.77503	12	05-05-2013	13	05-05-2013	15
18	9383E6D0E	455ADB884	VTS	1	1	0.85	-73.9626	40.76708	-73.9718	40.76586	13	05-05-2013	13	05-05-2013	6
19	0587ABA72	FE9CF687B!	VTS	1	6	1.19	-73.9807	40.77734	-73.97	40.79435	12	05-05-2013	12	05-05-2013	6
20	CF6864240	F522B35C1,	VTS	1	1	0.92	-73.9813	40.76852	-73.9793	40.7776	13	05-05-2013	13	05-05-2013	5

Data types:

Categorical	Quantitative
1. Rate code 2. Passenger Count	1. Pick up locations 2. Trip Distance 3. Drop off location 4. Trip time

Tasks

Discover- 1. Peak hours of taxi usage throughout the day.
2. The average number of passengers per taxi ride.
3. The distribution of different rate codes across taxi rides.

Identify- 1. Geographical distribution of taxi pickup points using the scatter plot on the map.
2. Relationship between trip duration and distance traveled.

Users

Taxi Companies like Uber and Lyft and also Industrial Automation companies like VTC.

Variables encoded (marks, channels)

Marks: points, lines
Channels: horizontal and vertical position, color luminance, shape.

Idioms

1. Bar chart
2. Scatter Plot
3. Geographic maps

Explanation:

The dashboard is disorganized due to the alphabet and labels, which is why I neglected to mention any labels or units. Therefore, I took it out and provided a detailed explanation in the explanation below. Clear explanation for each component of the dashboard based on the graphs:

- **Bar Plot for Passengers at Each Hour:**
  - **X-AXIS:** Pickup Time in Hours (0-24)
  - **Y-AXIS:** Count in Numbers
  - **Explanation:** The vertical bar plot represents the distribution of taxi passengers throughout the day, segmented by each hour. This visualization aims to identify peak hours of taxi usage, helping to optimize resource allocation and understand the temporal patterns of demand.

- **Map Showing Pickup Points:**
  - **X-AXIS:** Longitudes in degrees
  - **Y-AXIS:** Latitudes in degrees
  - **Explanation:** The map with scatter plot points visualizes the geographical distribution of taxi pickup points in New York City. Each point on the map represents a pickup location, allowing for a spatial understanding of where taxi services are in high demand. This can be valuable for identifying popular pickup areas or potential areas for service improvement.

- **Horizontal Bar Chart for Passengers per Ride:**
  - **X-AXIS:** Passengers Count in Numbers
  - **Y-AXIS:** Count or Sum in Numbers
  - **Explanation:** The horizontal bar chart illustrates the average number of passengers per taxi ride. This allows for a quick comparison of ride occupancy, helping to identify trends or anomalies. For instance, are there specific times or areas where taxis tend to have more passengers?

- **Horizontal Bar Plot for Rate Codes:**
  - **X-AXIS:** Rate Code in Numbers
  - **Y-AXIS:** Count or Sum in Numbers
  - **Explanation:** The horizontal stacked (or grouped) bar plot visualizes the distribution of different rate codes across taxi rides. This chart helps to understand the variety of rate codes used and their relative frequencies. It could reveal insights into the types of rides (e.g., standard fares, fixed-rate rides) and their popularity.

- **Scatter Plot Showing Time and Distance:**
  - **X-AXIS:** Trip Distance in miles.
  - **Y-AXIS:** Trip Time in minutes.
  - **Explanation:** The scatter plot displays the relationship between the duration of taxi trips and the corresponding distances traveled. Each point on the scatter plot represents an individual trip, providing insights into the efficiency of taxi rides. For instance, it can reveal if there's a correlation between longer distances and longer trip times.

Interactive Filters:

Explanation: The interactive filters (pickup date (Days in May), pickup time (0 -24 hours), and trip ID(Unique ID for each taxi ride)) provide users with the flexibility to focus on specific subsets of the data. Users can dynamically adjust the date, and time, or explore individual trips using the trip ID filter. This interactive feature enhances the dashboard's utility, allowing users to perform detailed explorations based on their specific interests or queries.