This was a project for a course I took from Google on Business Intelligence. This is one of the final projects you can select from.

The project began during the Foundations of Business Intelligence course. I was presented with notes from a fictional company, Cyclistic. This is assuming I was just hired as a new member of the business intelligence team at Cyclistic.

I attended a meeting and made notes. (actually provided during the course). During the meeting, I found out that Cyclistic was creating a business plan for next year and wanted to identify customer demand at different station locations.

I went through the notes and considered who was on the team, what they were trying to accomplish, what data I had access to, and what I needed to do to aid in that goal. With that in mind, I considered what dashboard I would need and what visualizations would help illustrate and explore the related data. At this point I did not begin on the dashboard, just gave it some initial thoughts before I even looked at the data.

I then accessed the various data sources and brought all of that data into a single table for later visualization use. This table is where I will manipulate the data in any way necessary (splitting fields, correcting obvious errors, etc.) before moving to any visualizations. In this process I used BigQuery to access public tables, uploaded a zipcode table provided to me, extracted necessary data from the various tables using multiple joins and grouping, exported the results to a .csv for later use in analysis.

After moving data into the single table, I began to look directly at the data to see what needed to be done as far as manipulations, and to see what kind of insights I think I will be able to pull from this data. I also verified that I have the information that I already know I'll need to answer the questions I'm already aware of.

I loaded the data into Tableau and started to create some visualizations. Early on I realized there were some entries with trip durations that exceeded a full day, so in Tableau I added a filter on the data source, removing entries with trip durations over one day in length.

The first visualization I created for Cyclistic was a line chart showing the total number of trips over time. To give a detailed breakdown that wasn't as unwieldly as daily totals, I showed totals broken down every week. It's more granular than monthly, which I think lost some helpful details. Because one of the things the team at Cyclistic wanted to be able to see is the difference between walk-up and subscriber trips, I added that in as a tooltip. While this will also appear on other charts, it seemed logical to show that for the day-to-day counts as well.

They also wanted a visualization showing the number of trips from each location per month. I created a table showing this information, setting the background color based on the number of trips to make it easier to discern at a glance which months and which locations have the largest number of trips. I did this by creating calculated fields to control the shade Tableau assigns as well as one to set the cell fill size (this has to do with how Tableau handles fills and doesn't have an option to fill a cell completely, but can fill based on a number). I also added trip totals by user type in tooltips for this visualization.

One thing they wanted to be able to look at was the total trip minutes shown by destination. Here I broke it down by borough and zip code to give the most detailed data we reasonably could. I used a bar code, allowing me to show them directly the difference between subscriber trips and walk-up trips. I also give the average trip duration in the tooltip.

The next visualization shows, in a more obvious and detailed manor, the average arrival time per destination. Here I show the average minutes for trips ending at the borough, neighborhood, zipcode, and I detail the difference between walk-up and subscriber customers. I used different bar colors to make it more obvious at a glance, as well.

And next I created a bar chart that shows trips ending at each location and broken down by clear days vs. days with precipitation. To do this, I created calculated fields to help sum the days with any precipitation vs. days without any. I also modified the tooltip to give the total trips per day without accounting for the weather.

Lastly, I created a visualization showing the average trip duration to answer the questions they had about whether subscribers or walk-up customers spend longer on each trip, and I broke this down by neighborhood and zipcode. It made sense to use a bar chart to display this. I also included the total number of trips by subscribers and walk-up customers in the tooltips on this chart.

With these visualizations, I created a dashboard for the Cyclistic team that allows them to gather insights. Within that dashboard we can, for example, see that the number of trips increased drastically from 2019 to 2020. We can see which neighborhoods are the most popular starting locations and which ones are barely used. We can also see that the same locations that are popular trip starting locations are also popular ending locations for trips. Weather is major factor for how many trips leaving from each location, with about half the number of trips taken on rainy days as their are on clear days. And finally, we can see that walk-up customers take trips about 50% longer than subscribers, but that subscribers take far more trips overall.

I added an insights tab, pointing out some of what I found, and in the process I decided to add one more visualization, Trip Totals Year to Year Comparison, which shows the number of trips each month with a different color of line for each year. This new visualization better displays the increased number of trips in 2020 whn compared to the same time period in 2019 in an obvious and easy to see way, but also reduces some of the day-to-day fluctuations due to weather and other temporary factors.

These are all important insights that the Cyclistic team can use in their decision-making process and when making future plans. If I were to expand this dashboard in the future, one thing I could try to add would be a visualization that makes it easy to see the congestion at each location. But this dashboard in its current form gives the Cyclistic team the information they identified as most important to them at this time.