# Data Sources Description

## Overview

For this analysis, I'm using historical trip data from Divvy, Chicago's bike-share service (the fictional Cyclistic in our case study). As specified in the case study instructions, I've downloaded the following datasets:

1. **Divvy\_Trips\_2019\_Q1.csv** - First quarter data from 2019
2. **Divvy\_Trips\_2020\_Q1.csv** - First quarter data from 2020

## Source

The data was downloaded from the official Divvy trip data repository hosted on Amazon S3:

* <https://divvy-tripdata.s3.amazonaws.com/Divvy_Trips_2019_Q1.zip>
* <https://divvy-tripdata.s3.amazonaws.com/Divvy_Trips_2020_Q1.zip>

## Data Description

According to the Divvy website, each trip record is anonymized and includes:

* Trip start day and time
* Trip end day and time
* Trip start station
* Trip end station
* Rider type (Member, Single Ride, and Day Pass)

The data has been processed by Divvy to remove trips taken by staff for service and inspection purposes, as well as any trips that were below 60 seconds in length (potentially false starts or users trying to re-dock a bike to ensure it was secure).

## Data Organization

The data is organized in CSV format with the following characteristics:

* Divvy\_Trips\_2019\_Q1.csv: 50,528,553 bytes
* Divvy\_Trips\_2020\_Q1.csv: 71,273,013 bytes

## Data Credibility Assessment (ROCCC)

* **Reliable**: The data comes from the official Divvy system, which tracks all bike trips.
* **Original**: The data is from the primary source (Divvy/Lyft Bikes and Scooters, LLC).
* **Comprehensive**: The data includes all trips taken by Divvy users during the specified time periods.
* **Current**: The data represents historical information from 2019 Q1 and 2020 Q1 as required by the case study.
* **Cited**: The data is made available by Motivate International Inc. under license as mentioned in the case study.

## Privacy and Security Considerations

As noted in the case study, data-privacy issues prohibit the use of riders' personally identifiable information. The data does not include:

* Personal details of riders
* Credit card numbers
* Individual user identifiers that could connect pass purchases across multiple trips

This ensures compliance with privacy regulations while still providing sufficient information for analysis of riding patterns between member types.