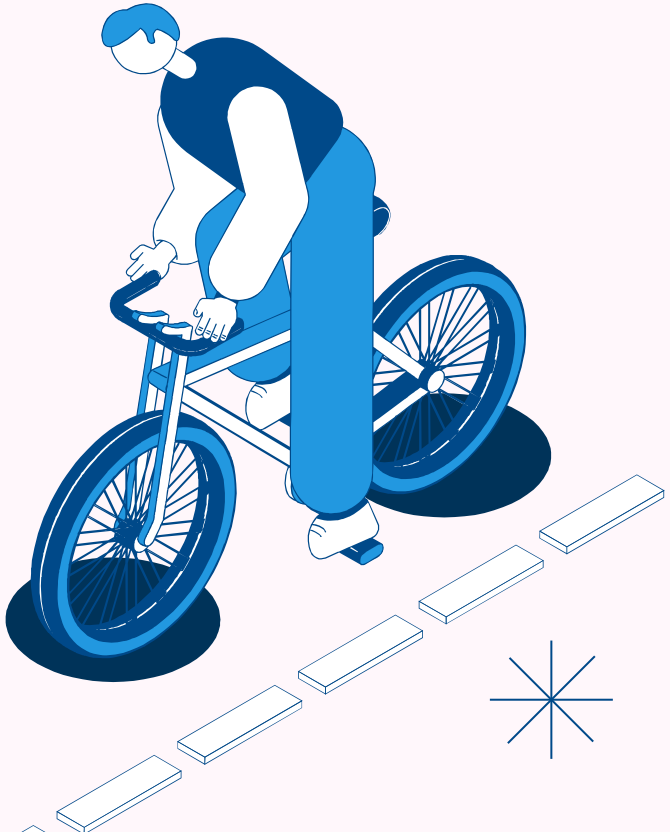
Project 3
Divvying in the Rain!



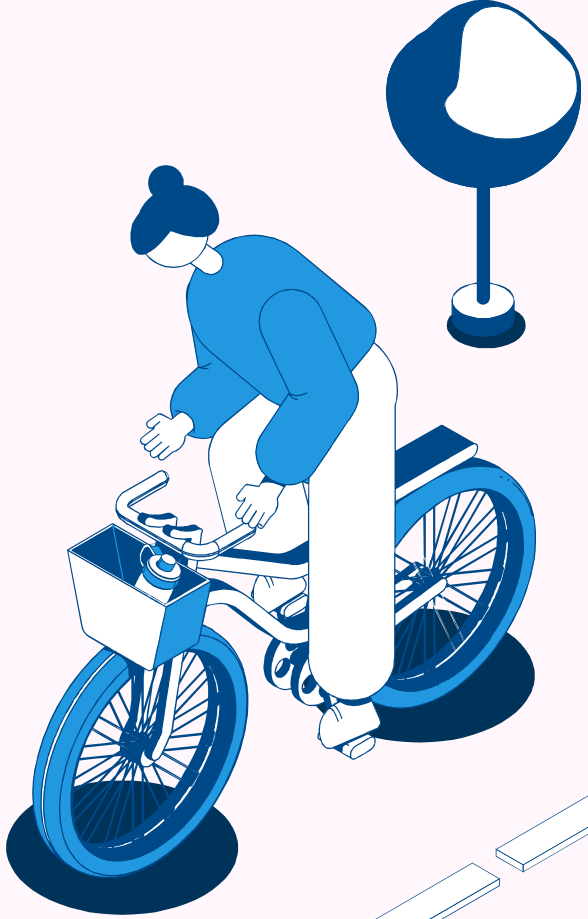
Group 1
Jason Pealy, Andrew Hawthorne, Ben Richardson, Aileen Alvarez, & Alicia Hlavac

TABLE OF CONTENTS

01

Project Overview

Objective & Research Questions



Objective

Using Divvy trip history data and OpenWeather API data, from January to December of 2022, our goal was to identify how many Divvy members keep riding, rain or shine.

Research Questions

- What are the top 10 start stations, top 10 end stations, and top 10 routes?
- Which season(s) have the most and/or least Divvy Rides?
- How does the total count of Divvy Rides vary by month?
- How does precipitation affect the average daily rides?

02

Preparing & Processing the Data

Data Sources , Data Collections, & System Architecture

Data Sources

Daily data from January 1st, 2022 to December 31st, 2022

Open Weather API

- Date
- Precipitation
- Morning Temperature
- Afternoon Temperature
- Evening Temperature
- *Significant Precipitation**
- *Average Daily Temperature***

**Significant precipitation is greater than 0.1 inches*
***Average Daily Temp calculated using the morning, afternoon, and evening temperature*

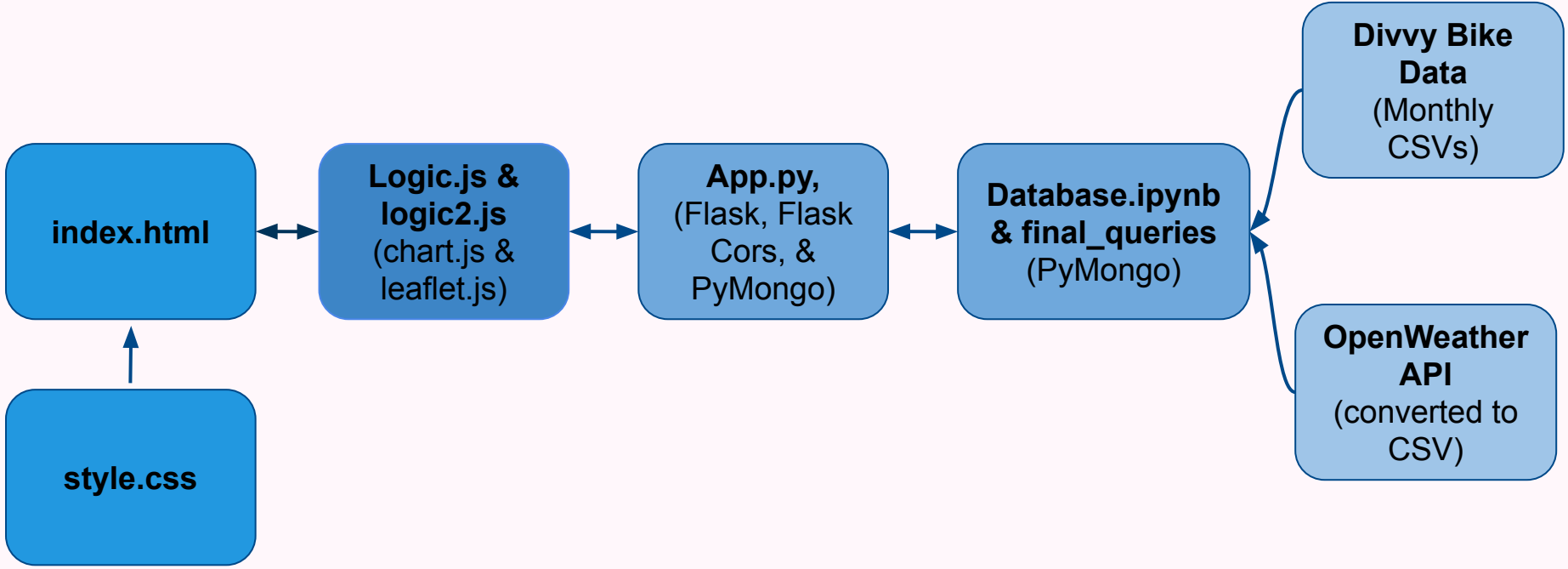
Divvy Trip History Data

- Ride ID
- Started At
- Start Date
- Start Time
- Start Station Name
- Start Latitude
- Start Longitude
- Ended at
- Ended Date
- Ended Time
- End Station Name
- End Latitude
- End Longitude

Data Collections (MongoDB)

divvy_ridedata	Monthly Divvy CSV Files combined to yearly file (2022)
weather_daily	Daily Chicago weather (2022)
divvy_ridedata_merged	Weather_daily merged to divvy_ridedata
WithLatLong	Divvy Rides without zeros or empty strings for Latitude and Longitude
withStationName	Divvy Rides with no empty strings for station names
divvy_rides_by_season	Total Count of Divvy Rides by Season
Top10StartStations	Top 10 Start Station Names with Total Count of Divvy Rides
Top10EndStations	Top 10 End Station Names with Total Count of Divvy Rides
RouteDistance	Calculated from the Pythagorean Theorem using Start & End Latitude and Longitude
Top10Routes	Top 10 most popular routes with Total Count of Divvy Rides
Divvy_rides_by_month	Total Count of Divvy Rides per Month

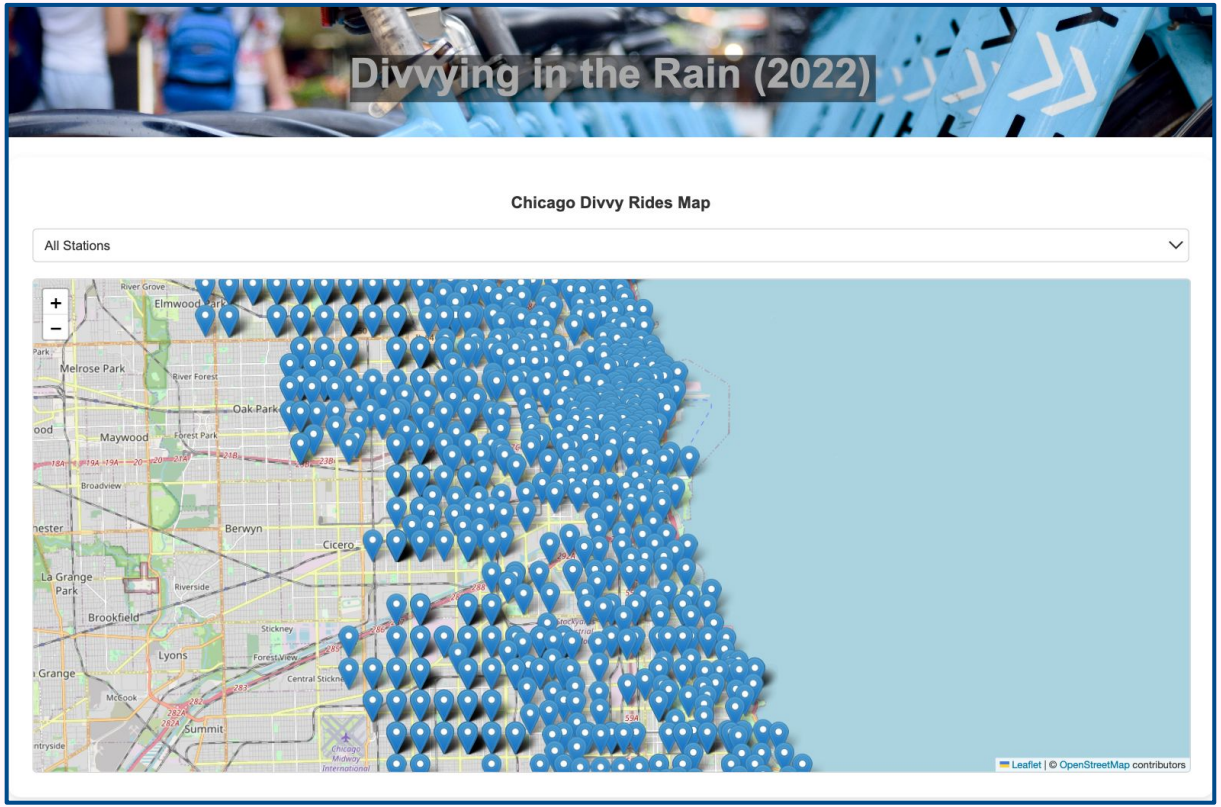
System Architecture



03

Interactive Visualizations

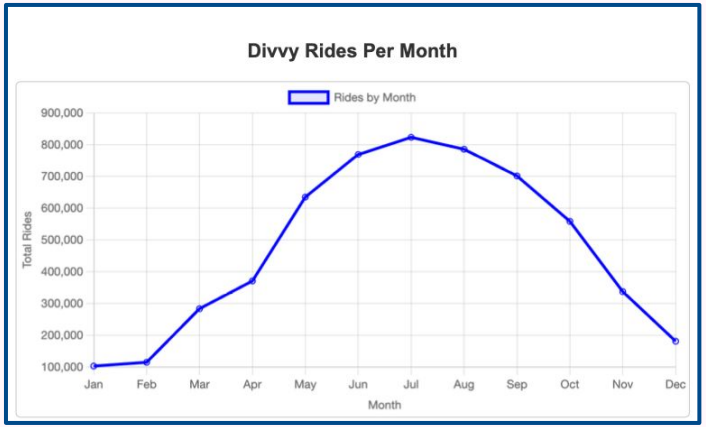
Dashboard: Chicago Map

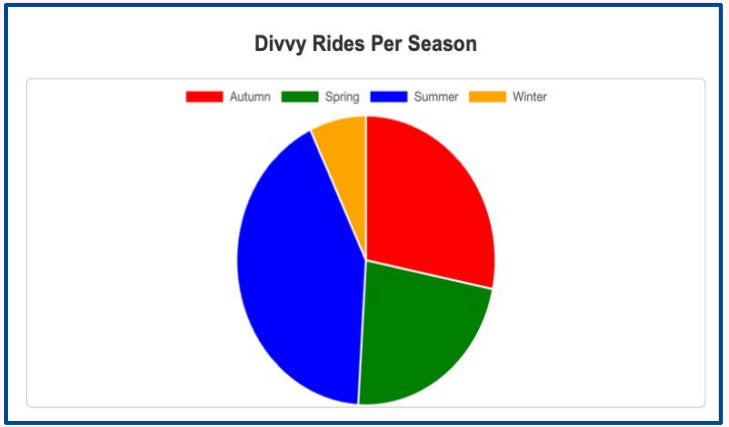


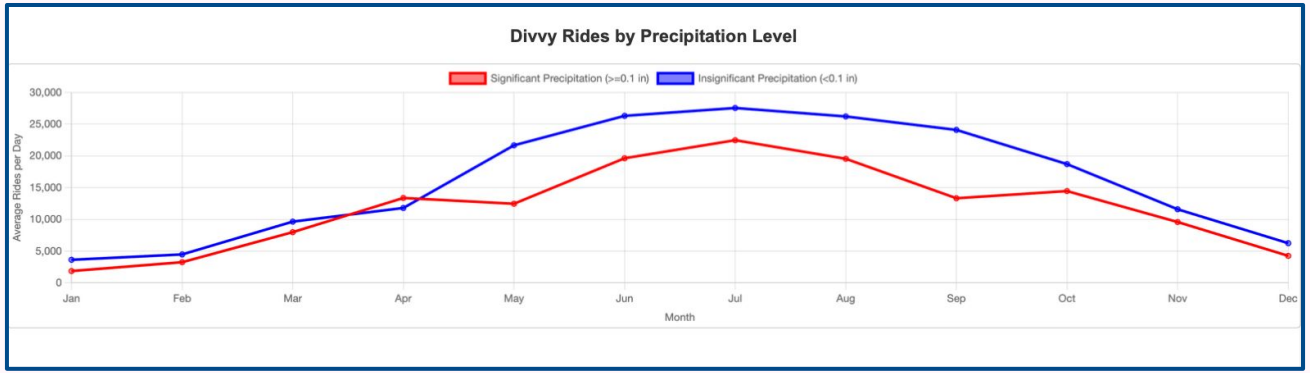
The drop down function allows us to narrow down to the:

- Top 10 Start Stations
- Top 10 End Stations
- Top 10 Routes

Dashboard: Breakdown of Divvy Rides







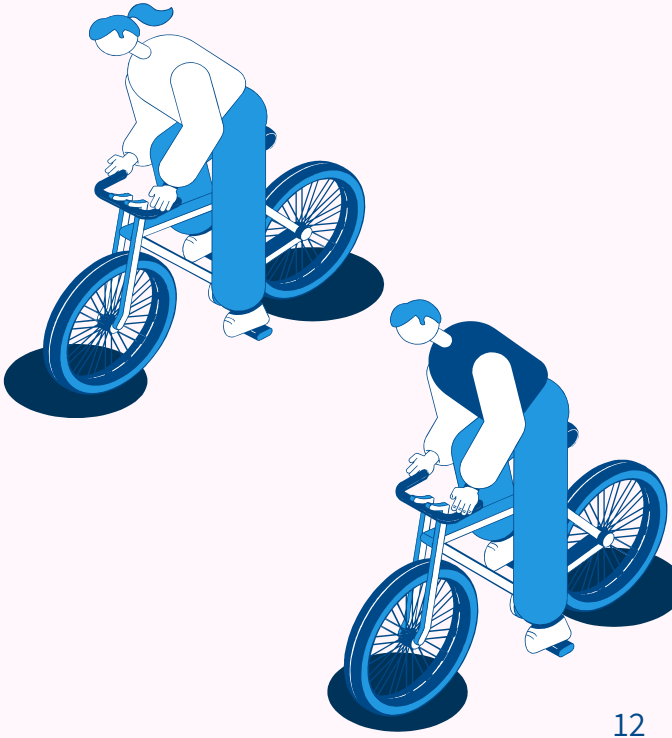
04 Takeaways

From the Data:

- The most popular Divvy Stations are within Downtown Chicago
- Riders tend to prefer to bike in the summertime
    - Total Count of Divvy rides peak in July
- Excluding April, there are fewer rides year round when there is significant precipitation present

From the Project:

- Increased awareness of the process of using various applications to display findings to end user
- Performing significant data merges on MongoDB vs. SQL



Next Steps:

- Look further into electric bike ride data - only used for overall counts as many did not have start and/or end stations (can be left anywhere)
- See if there are any significant differences between Divvy members and casual riders
- Investigate why many “top routes” start and end at the same station, and see if the riders of these routes tend to be more casual than members (for example, tourists might pick up and return a bike to the same station)

Q&A