**Annual Bicycle Count**

City of Portland

12.3.2018

**─**

Prepared by:

Arihan Jalan. Audrey Tay. Sidney Hutchison.

Instructors:

Mayuree Binjolkar. Zhiyong Cui

# Table of Contents

[**Table of Contents**](#_ziamv9tnln5n) **1**

[**Introduction**](#_ud4rqu44p4c0) **2**

[**Assumptions**](#_u7y1xryjzqh1) **2**

[**Description of Database**](#_ghjpqr63mzeq) **3**

[Relational Schema](#_ec2bo6c605p6) 3

[**User Manual**](#_x9lwck755gh1) **4**

[Database Attributes and Relations](#_thjsltdvj9vi) 4

[Location Table](#_cn6v4upicuku) 4

[Survey](#_1cg3r2w42y2u) 4

[Counts](#_jlptaljih8et) 5

[Shiny Application User Manual](#_5lgq3r9zmgo2) 6

[First Tab: Home Page](#_qgt5cwc3wffs) 6

[Second Tab: Data Input: Counts, Location](#_liaedk5oxjrk) 6

[Third Tab: Data Input: Survey](#_trklbb7p4vcg) 6

[Fourth Tab: Static Visualization](#_dxxzuuulgvfp) 6

[Fifth Tab: Interactive Visualization](#_9wqpg3auneyw) 7

[**Appendix**](#_6lpied4pd94) **9**

[Team Member Contribution](#_pvdwyo4vuvbd) 9

[First tab: Shiny App Home Page](#_he5gu2de5srg) 9

[Second tab: Bicycle Counts Form for counts, locations table](#_t8p8nwkjh1kc) 9

[Example Data Collection form used by volunteers](#_5c187hldfisd) 10

[Third tab: Survey Data Input Form](#_fb9ugz34a0lb) 10

[Shiny Dashboard (user entering data)](#_ffspq4n8hk5s) 10

[Shiny Dashboard (user submitted data; UI confirmation)](#_8haeomfxh05w) 11

[SQL Query against DBMS within R to verify user submitted data is uploaded to SQL](#_63bluy1ujezh) 11

[Fourth Tab: Static Visualization created within Excel](#_cvey7bsebcci) 12

[Fifth tab: Dynamic Visualization created using Shiny and Leaflet Library](#_5hpw2qu80ioz) 13

[SQL Scripts for table creation](#_h4vetqw53450) 14

[Locations table SQL script](#_gfvp8gwcgs3r) 14

[Counts table SQL Script](#_aohgb04cjxx) 14

[Survey Table SQL Script](#_a78qcxq292rz) 15

[SQL code for testing insertions into Survey table](#_xqbie2wyp34n) 15

# Introduction

The City of Portland wants to know information about cyclists in the city. To do this, they have volunteers collect counts at specific intersections during specific time periods (7:00 AM - 9:00 AM and 4:00 PM - 6:00 PM) using the Portland Bicycle Count form included in the Appendix of this report. The volunteers where the cyclists are heading, their gender, whether or not they’re wearing a helmet, and information about the weather conditions. Additionally, volunteers may conduct in-person surveys with cyclists as they choose and record data about the cyclist’s trip origin, destination, and purpose, as well as the frequency of their bicycle use.

We have created an app where volunteers can input the survey data that is updated in the SQL database, and a proof of concept submission form for the counts (not connected to database). Additionally, users of the app will be able to see and interact with the data visually to assess how cycle use is distributed spatially while controlling for traffic class. This report will detail the assumptions and process that went into creating the database and the app.

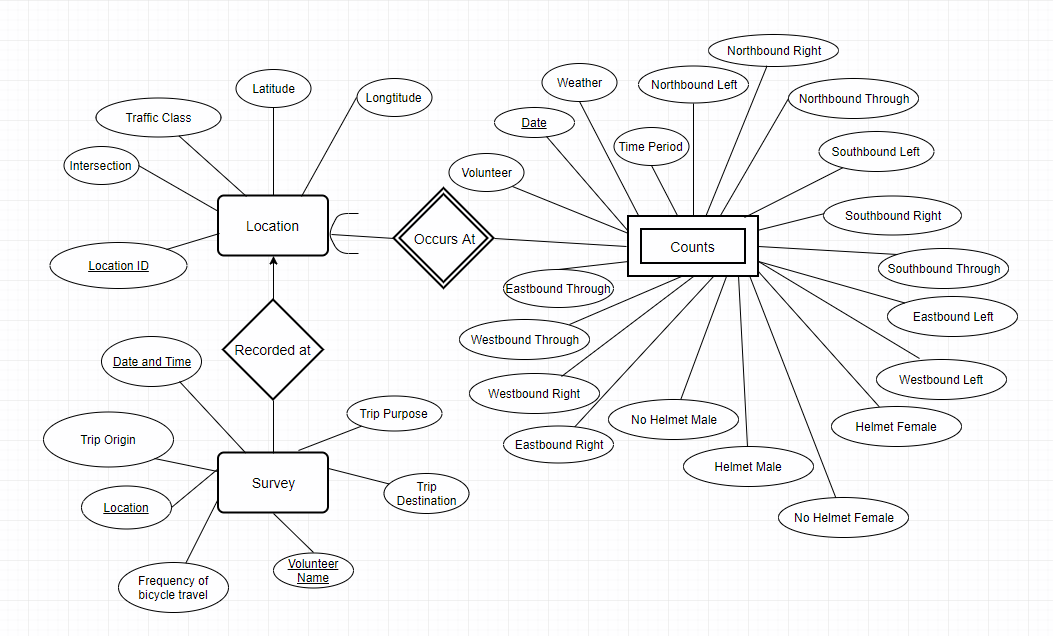
# 

# 

# Assumptions

* Data will not be collected at the same location more than once each day.
* All locations share the same weather conditions at a given time.
* The app is to be used by program staff only, not the general public. So everyone using the app has a general understanding of the annual bicycle count program, and the type of data being recorded (e.g.- numbers, strings, date formatting, etc.)
* Traffic Class refers to the number of non-cycle vehicle traffic at an intersection.
* Traffic Class, LocationID, Latitude, and Longitude are autocomputed within SQL based on input from the user (e.g.- trafficClass is assigned based on traffic data).
* Only volunteers who have previously recorded data will enter new survey data.
* No two volunteers have the same name.

# Description of Database



**Picture 1. E/R Diagram**

## 

## 

## Relational Schema

Location(LocationID, Intersection, TrafficClass, Latitude, Longitude)

Survey(Location, DateandTime, VolunteerName, TripPurpose, TripOrigin, TripDestination, Frequencyofbicycletravel, Location.LocationID)

Counts(Location.LocationID, Date, Volunteer, Weather, TimePeriod, NorthboundLeft, NorthboundRight, NorthboundThrough, SouthboundLeft, SouthboundRight, SouthboundThrough, WestboundLeft, WestboundRight, WestboundThrough, EastboundLeft, EastboundRight, EastboundThrough, HelmetMale, HelmetFemale, NoHelmetMale, NoHelmetFemale)

# User Manual

## 

## Database Attributes and Relations

For the entity sets, the following attributes were chosen. This includes all of the attributes that were given in the dataset. In addition to the dataset, we also included another entity called Survey, which included attributes for future survey records.

### Location Table

|  |  |
| --- | --- |
| **Attribute** | **Explanation** |
| Location ID | The location identification number assigned to the specific intersection |
| Intersection | Name of the intersection |
| Traffic Class | The typical level of traffic an intersection services (high, medium, or low) |
| Latitude | Latitude value for the specific intersection |
| Longitude | Longitude value for the specific intersection |

### Survey

|  |  |
| --- | --- |
| **Attribute** | **Explanation** |
| Location | The name of the intersection at which the volunteer collected information from the bicyclist |
| Date and Time | The numerical date on which the data was collected (YYYY-MM-DD) and Time |
| Volunteer Name | The name of the volunteer(s) that collected the data |
| Trip Purpose | The purpose of the trip as reported by the bicyclist to the volunteer |
| Trip Origin | The location that the trip started as reported by the bicyclist to the volunteer |
| Trip Destination | The location that the trip would end as reported by the bicyclist to the volunteer |
| Frequency of bicycle travel | How often the bicyclist travels as reported by the bicyclist to the volunteer |

### 

### Counts

|  |  |
| --- | --- |
| **Attribute** | **Explanation** |
| Date | The numerical date on which the data was collected (YYYY-MM-DD) |
| Volunteer | The name of the volunteer(s) that collected the data |
| Weather | A combination of the average temperature and the weather conditions recorded |
| Time Period | The specific two-hour time interval in which the data was collected; either 7:00-9:00 AM or 4:00-6:00 PM |
| Northbound Left | A count of the number of bicycles traveling Northbound that turned left at that intersection |
| Northbound Right | A count of the number of bicycles traveling Northbound that turned right at that intersection |
| Northbound Through | A count of the number of bicycles traveling Northbound that went straight through that intersection |
| Southbound Left | A count of the number of bicycles traveling Southbound that turned left at that intersection |
| Southbound Right | A count of the number of bicycles traveling Southbound that turned right at that intersection |
| Southbound Through | A count of the number of bicycles traveling Southbound that went straight through that intersection |
| Westbound Left | A count of the number of bicycles traveling Westbound that turned left at that intersection |
| Westbound Right | A count of the number of bicycles traveling Westbound that turned right at that intersection |
| Westbound Through | A count of the number of bicycles traveling Westbound that went straight through that intersection |
| Eastbound Left | A count of the number of bicycles traveling Eastbound that turned left at that intersection |
| Eastbound Right | A count of the number of bicycles traveling Eastbound that turned right at that intersection |
| Eastbound Through | A count of the number of bicycles traveling Westbound that went straight through that intersection |
| Helmet Male | A count of the number of bicycle travelers that came to the specified intersection at the specified date/time that both wore a helmet and were identified as male |
| Helmet Female | A count of the number of bicycle travelers that came to the specified intersection at the specified date/time that both wore a helmet and were identified as female |
| No Helmet Male | A count of the number of bicycle travelers that came to the specified intersection at the specified date/time that both did not wear a helmet and were identified as male |
| No Helmet Female | A count of the number of bicycle travelers that came to the specified intersection at the specified date/time that both did not wear a helmet and were identified as female |

## Shiny Application User Manual

Note- screenshots for the content for each tab in the Shiny app is provided in the appendix

### First Tab: Home Page

Introduction of our application.

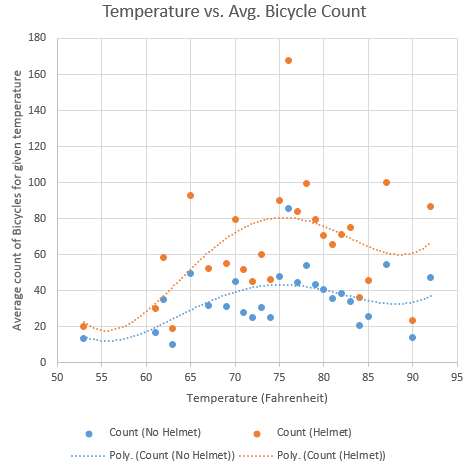
### Second Tab: Data Input: Counts, Location

Data entry interface for the counts information collected by volunteers.

### Third Tab: Data Input: Survey

Data entry interface for the survey information collected by volunteers.

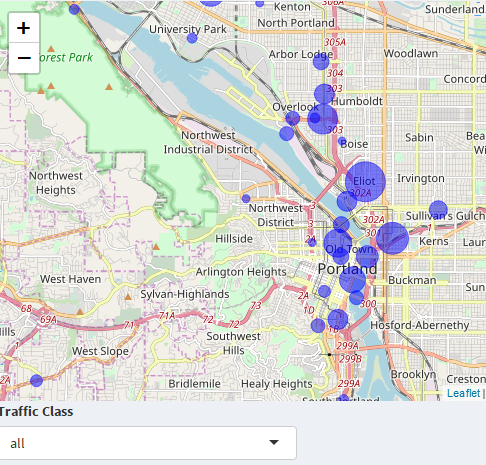
### Fourth Tab: Static Visualization



The static graph used in this tab shows the Temperature vs. Average Bicycle Count data for bicyclists. The scatterplot visualizes the average count of cyclists for a given temperature; and the line of best helps show how they’re correlated. The number of bicycles at a given location is computed by taking a group average of “HelmetMale”, “HelmetFemale” into “Helmet”, and “NoHelmetMale”, “NoHelmetFemale” into “No Helmet” for a given temperature.

Temperature and the number of cyclists are positively correlated- as temperatures increase- there are more cyclists; there is a maxima at around 75 °F. The graph also shows that generally, there are more cyclists with helmets than those without.

### Fifth Tab: Interactive Visualization



The number of bicycles at a given location is computed by averaging the “HelmetMale”, “HelmetFemale”, “NoHelmetMale”, “NoHelmetFemale” for a given intersection. Locations that did not have geolocation data are omitted. If data is marked with the incorrect geolocation information (different intersections having same geolocation, or latitude, longitude data)- only the first instance of the same latitude, longitude pair is kept.

The map above visualizes the average bicycle count for a given intersection in Portland.The blue circles represent the bicycle counts at the specified intersections. The size of a blue circle is larger if the bicycle count is larger at that location. The map allows the user to see the bicycle count data of an intersection through the lens of the vehicular traffic class of those intersections. For example, the user can choose to look at all intersections simultaneously or they can choose to only look at the intersections with a particular traffic class: high, low, or medium.

The map shows the city of Portland, the largest city in the state of Oregon, has many large circles which represent a high number of recorded cyclists- the circle sizes reduce as we move away from the city center. This could probably be explained by there being a higher population density and more commuted activity within the city- which could mean there are more people that cycle, or people need to travel more- hence there are more cyclists (e.g.- attending work, going to events, etc.)

# Appendix

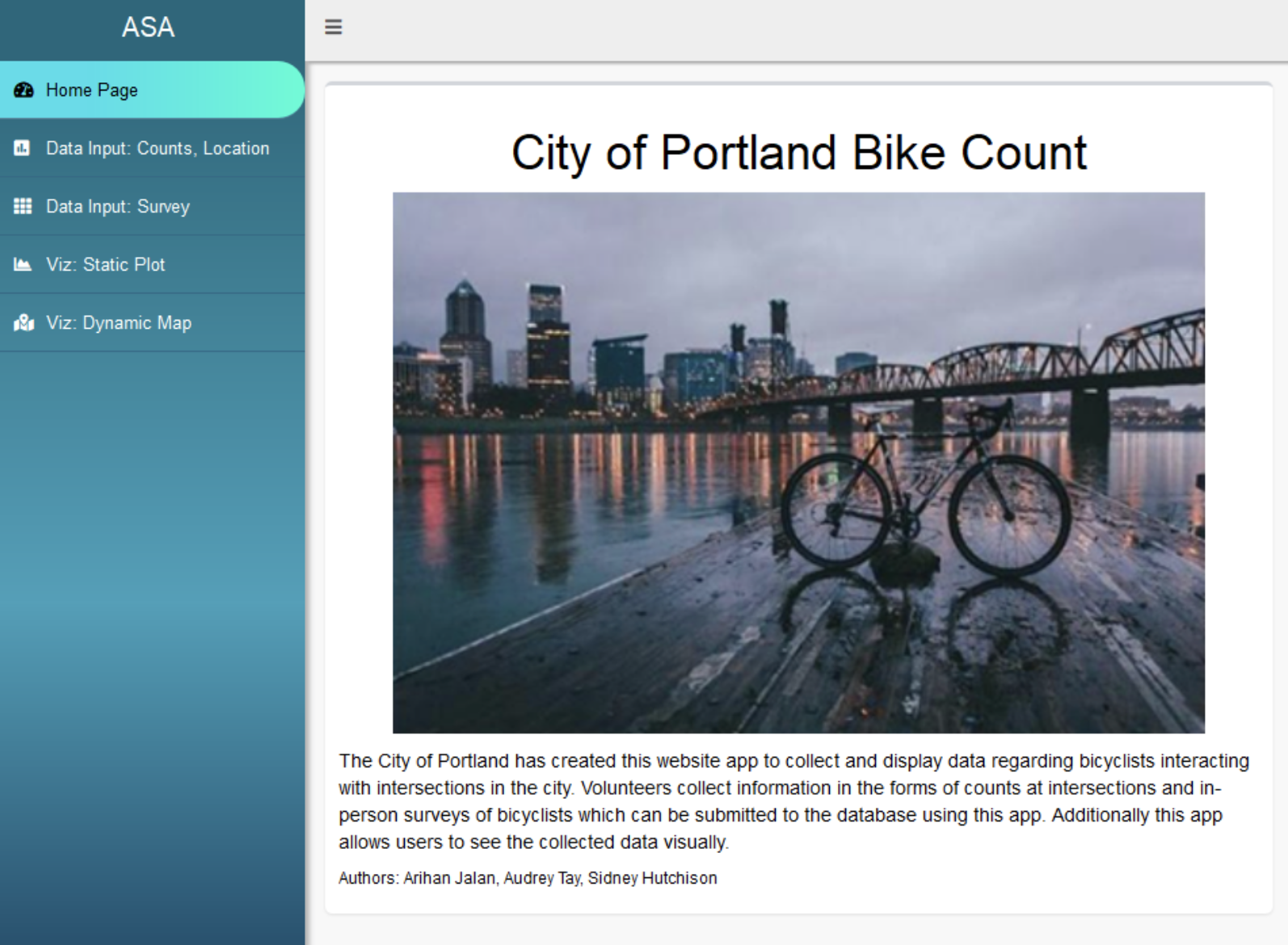
## Team Member Contribution

**Sidney 33.3%**

**Audrey 22.3%**

**Arihan 44.3%**

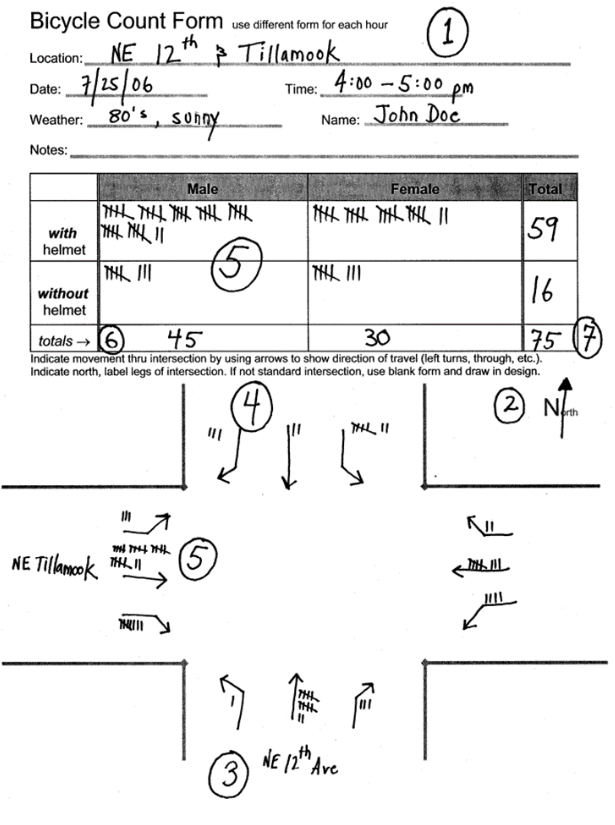
## First tab: Shiny App Home Page



## Second tab: Bicycle Counts Form for counts, locations table

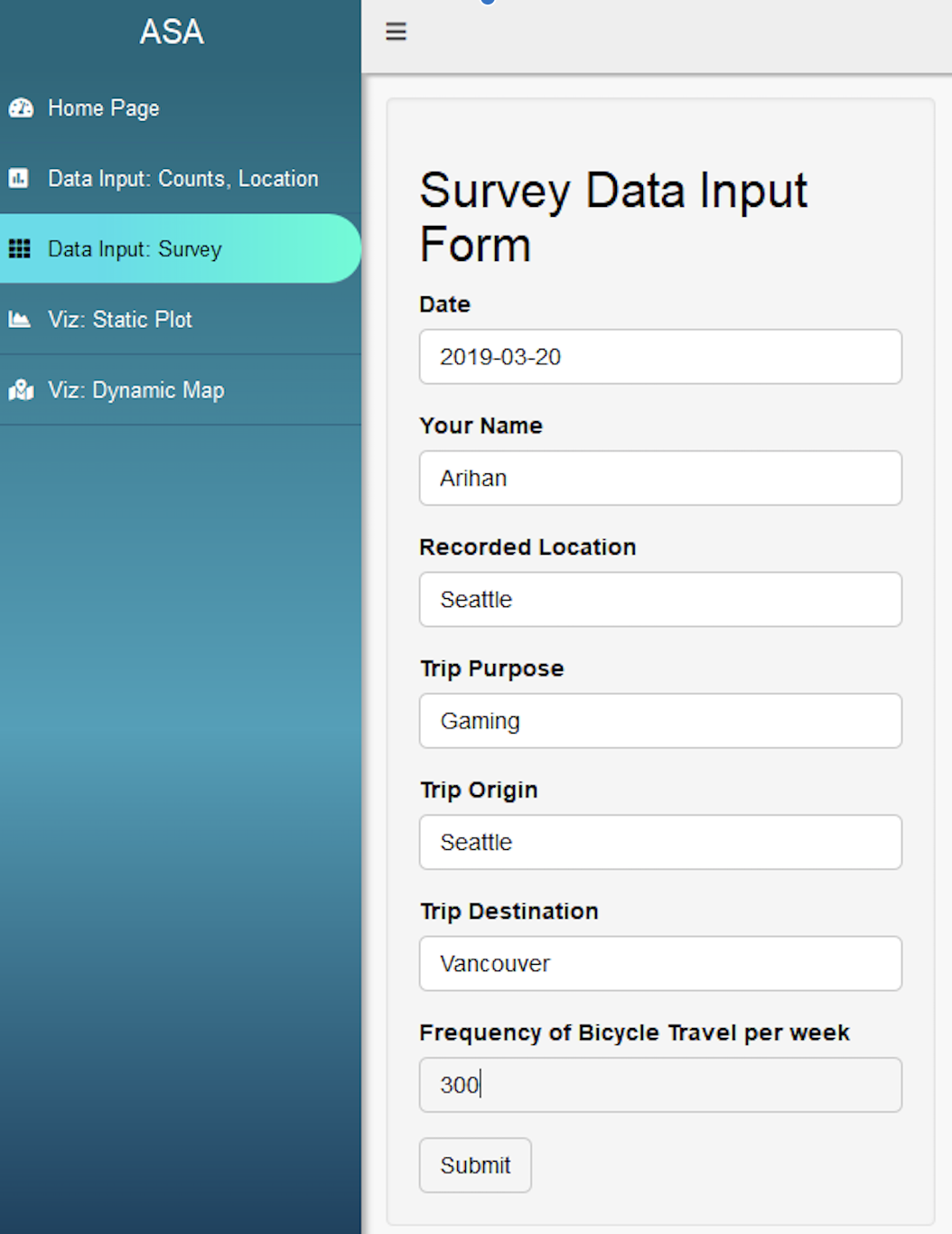


### Example Data Collection form used by volunteers

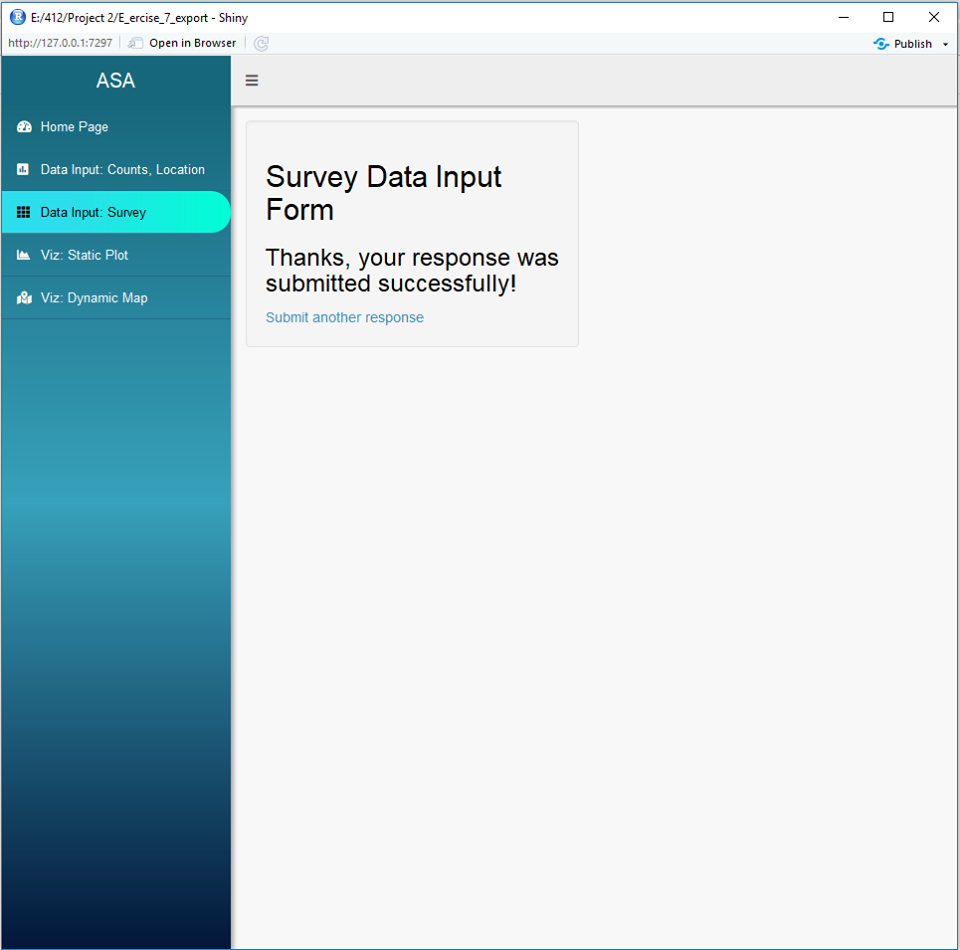


## Third tab: Survey Data Input Form

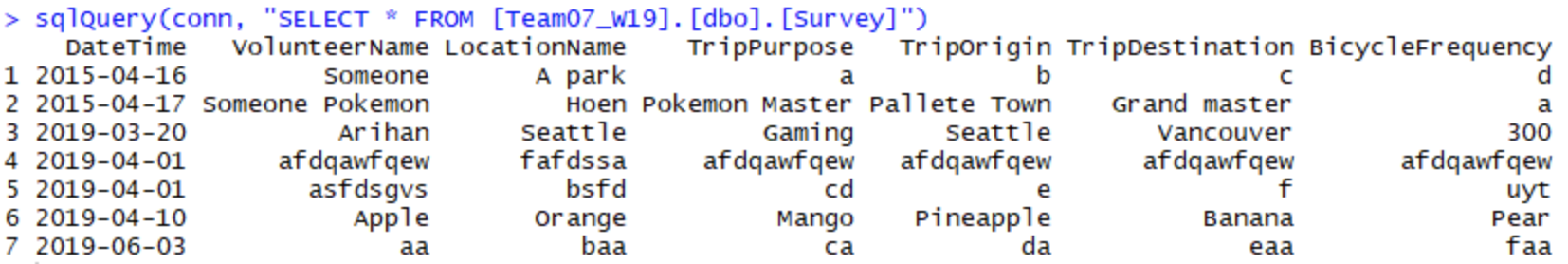
### Shiny Dashboard (user entering data)



### Shiny Dashboard (user submitted data; UI confirmation)



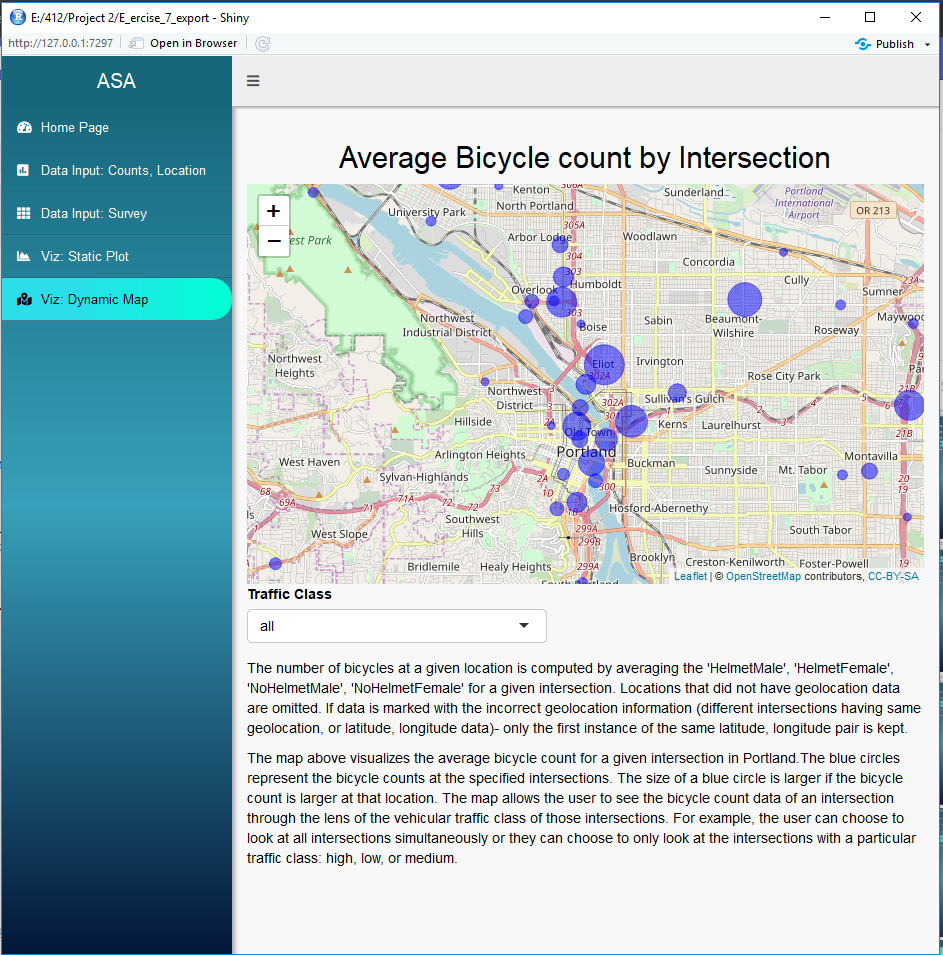
### SQL Query against DBMS within R to verify user submitted data is uploaded to SQL



## Fourth Tab: Static Visualization created within Excel



## Fifth tab: Dynamic Visualization created using Shiny and Leaflet Library



## SQL Scripts for table creation

### Locations table SQL script

CREATE TABLE [dbo].[Locations] (

[LocationID] int NOT NULL,

[Intersection] nvarchar(255) NOT NULL,

[TrafficClass] nvarchar(255),

[Latitude] float,

[Longitude] float,

PRIMARY KEY ([LocationID])

)

### Counts table SQL Script

CREATE TABLE [dbo].[Counts] (

[LocationID] int NOT NULL FOREIGN KEY REFERENCES dbo.Locations(LocationID),

[Volunteer] nvarchar(255),

[Date] datetime NOT NULL,

[Weather] nvarchar(255),

[TimePeriod] nvarchar(255),

[NorthBoundLeft] float,

[NorthBoundRight] float,

[NorthBoundThrough] float,

[SouthBoundLeft] float,

[SouthBoundRight] float,

[SouthBoundThrough] float,

[EastBoundLeft] float,

[EastBoundRight] float,

[EastBoundThrough] float,

[WestBoundLeft] float,

[WestBoundRight] float,

[WestBoundThrough] float,

[HelmetMale] float,

[HelmetFemale] float,

[NoHelmetMale] float,

[NoHelmetFemale] float,

PRIMARY KEY ([LocationID], [Date])

)

### Survey Table SQL Script

CREATE TABLE [dbo].[Survey] (

[DateTime] datetime NOT NULL,

[VolunteerName] nvarchar(255) NOT NULL,

[LocationName] nvarchar(255) NOT NULL,

[TripPurpose] nvarchar(255),

[TripOrigin] nvarchar(255),

[TripDestination] nvarchar(255),

[BicycleFrequency] nvarchar(255),

PRIMARY KEY ([DateTime], [VolunteerName], [LocationName])

)

### SQL code for testing insertions into Survey table

INSERT INTO dbo.Survey

VALUES ('20150416', 'Someone', 'A park', 'a', 'b', 'c', 'd');