Introduction

Hello, welcome to my application documentation for Project 1 of CS 424 - Visual Analytics. This project's main purpose is to teach how to use the R programming language and how to visualize it using ggplot; a library that has built in functions in order to plot data and gives the programmer many different options on how to manipulate and visualize data. The library ggplot allows us to manipulate the raw data that we're receiving in various forms like scatter plots and line graphs and even box-and-whisker plots. For the purposes of this project, my application will use ggplot to plot the data in the forms of bar plots and tables. The other libraries that I will be using are Shiny and Shiny Dashboard. These libraries enable an R program to create a webpage that can display all the plots created by ggplot and it lets users be able to dynamically change the input data and have the plots change accordingly based on what the user does. For example, if someone only wanted to see a specific range of data for only one year, all they would have to do is click on a dropdown menu and click the year that they want and it would display the data that they want only in that specific time period.

The data that I'll be visualizing is data on riders on the Chicago L over the last 20 years. This would be from 2001 to 2021. For 60% of the grade we'll need to download the data from the Chicago Data Portal's website and clean the data into a usable format. The plots that are needed for this section are a bar chart for the total entries at UIC-Halsted for each year,
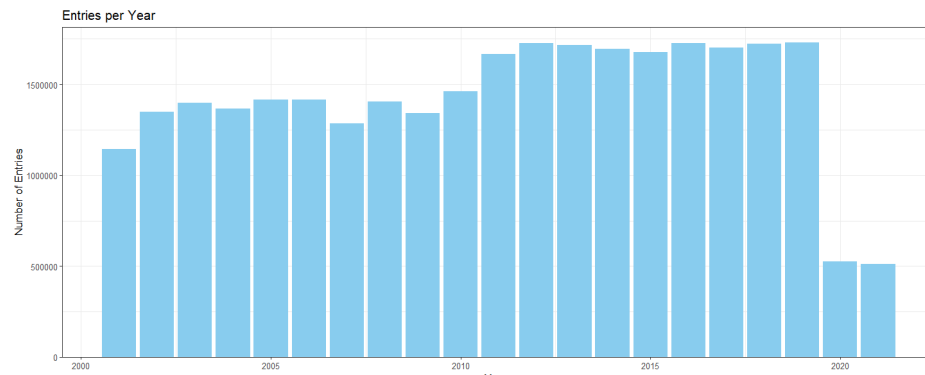
Figure 1: Entries per Year

each day for 2021 at UIC-Halsted, each Month for 2021 at UIC-Halsted, and each Day of the week for 2021 at UIC-Halsted.
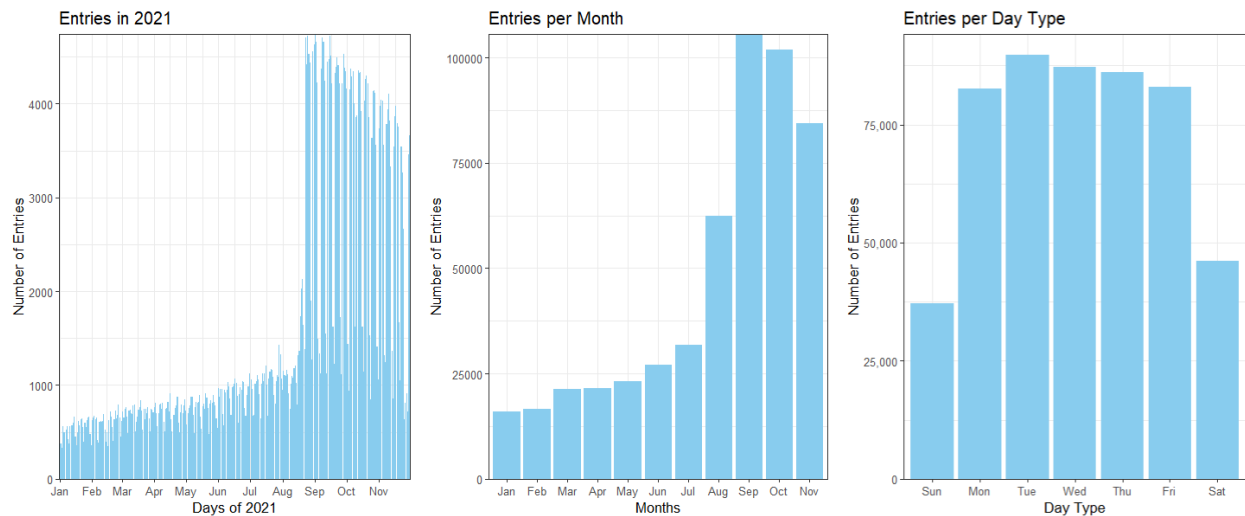

Figure 2: Entries for each day, month, and week day

There should also be a menu to choose any of the years from 2001-2021 and have all the charts update for the chosen year. There should also be a table for each of the graphs that are plotted as well.

Years	Entries		V1	V2		V1	V2
Search:			Search:			Search:	
2001	1143302		January	2459879		Sun	37224
2002	1350077		February	2834072		Mon	82535
2003	1398751		March	2679422		Tue	89695
2004	1364987		April	2802109		Wed	87213
2005	1414985		May	1739560		Thu	86059
2006	1417008		June	1702280		Fri	82932
2007	1286209		July	1761178		Sat	46141
2008	1406969		August	2165666			
2009	1340464		September	3393969			
2010	1461571		October	3656629			

Showing 1 to 7 of 7 entries

Previous 1 Next

Showing 1 to 10 of 21 entries

Previous 1 2 3 Next

Showing 1 to 10 of 12 entries

Previous 1 2 Next

Figure 3: Charts for Yearly data, Monthly data, Day data

Finally, there should be an about page.

About

Jack Martin created this app for Project 1 of UIC's CS 424 - Visual Analytics.

This data is from the Chicago Data Portal. More specifically, the 'CTA - Ridership - L Station Entries - Daily total'. The main components on why we are given this project is to teach us and give better familiarity with both the R language and Shiny and Shiny dashboard. We were tasked with analyzing and plotting Entries over specific stations over 2001-2021 and over each Day of the Week and Month.

Figure 4: About Page

Personally, I wasn't able to get to the part past this point but for the rest of the points there would need to be a side by side comparison between Halsted and O'Hare station data with the capabilities where the user can choose which year gets compared. For example, monthly data from 2002 for Halsted and daily data from O'Hare 2005 can be

compared against each other. There should also be another station for the last 20% of the grade that will be added to the same comparison options for the previous 20%.