# Note in the Pocket Exploratory Data Analysis Report

Alvin Sheng, Jimmy Hickey, Naomi Giertych, Yinqiao Wang, Jake Koerner, Hyoshin Kim

**Abstract**

This report contains an exploratory data analysis of the requests filled by Note in the Pocket from 2019 to the present.

## Contents

# 1 Count plots for requests given to Note in the Pocket

## 1.1 Clothing categories

In this section and the last section, we count the number of individuals that request clothing of a given gender and size:

- An individual that requests clothing labeled as gt, gb, gu, gs, gc is considered as "female small."
- An individual that requests clothing labeled as bt, bb, bu, bs, bc is considered as "male small."
- An individual that requests clothing labeled as wt, wb, wu, ws, wc is considered as "female large."
- An individual that requests clothing labeled as mt, mb, mu, ms, mc is considered as "male large."

## 1.2 Plot of Items by Date

Note that the red vertical line marks the COVID starting date, the green areas the spring semester and the brown areas the fall semester for Wake County public schools. The blue curves are the fitted curves with respect to the number of requests for the female and male items.

- Based on the time plots for each clothing item (female small, male small, female large, male large), there is a recurring pattern each year. A small peak in spring, and a larger peak in fall for the number of requests, reflecting the start of the school semester.
- There seems to be a smaller number of requests during the 2020 spring semester after COVID starts.
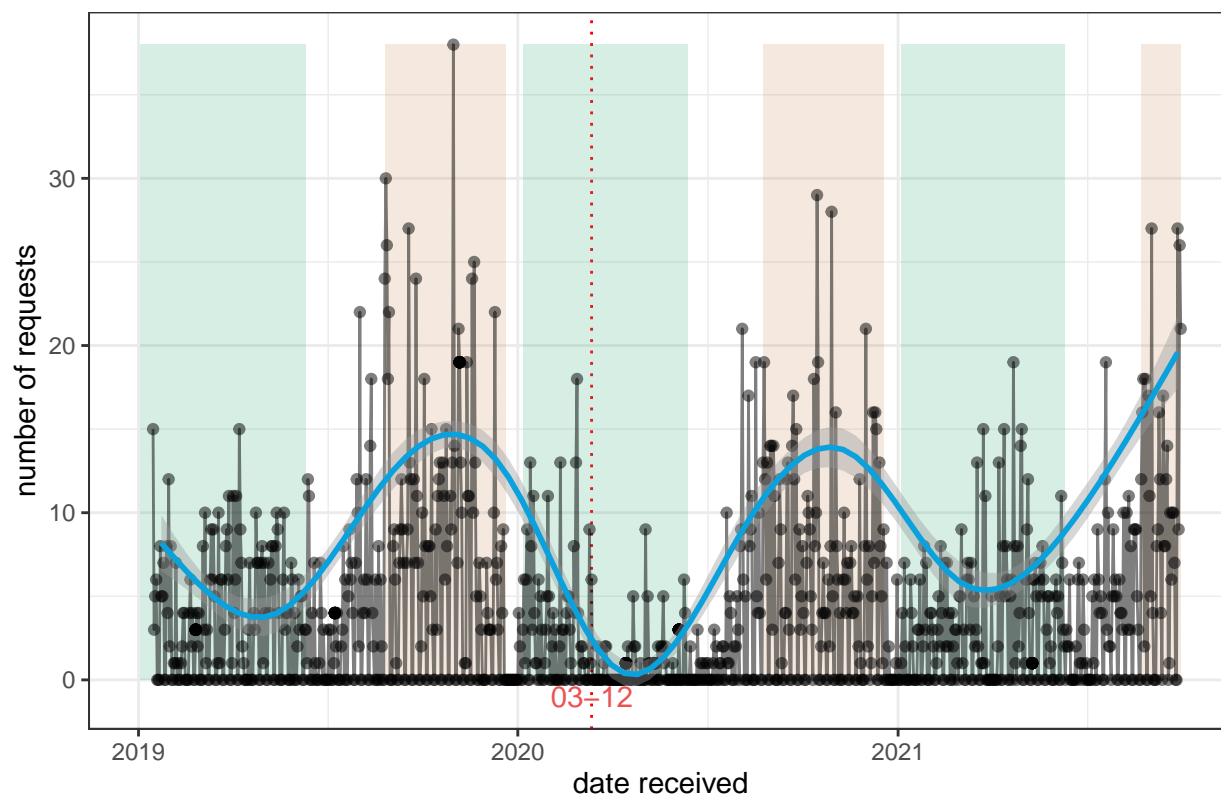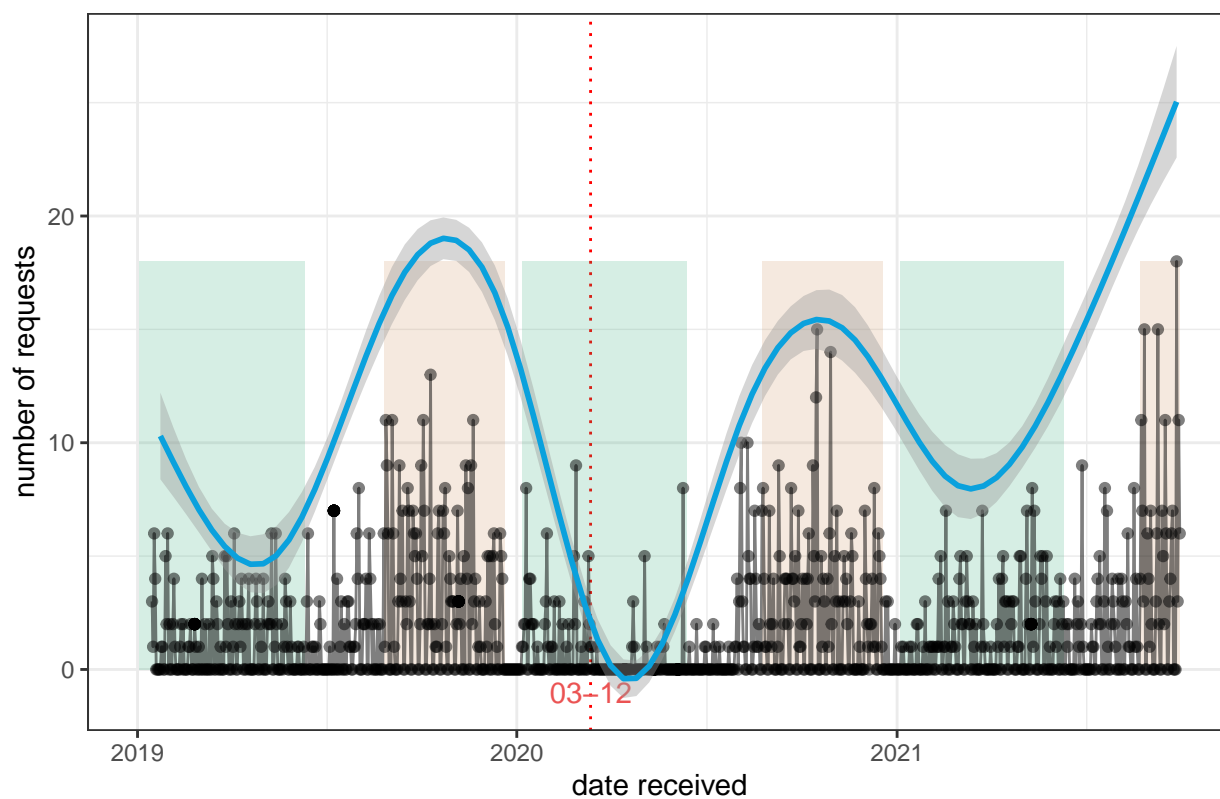
Total number of requests
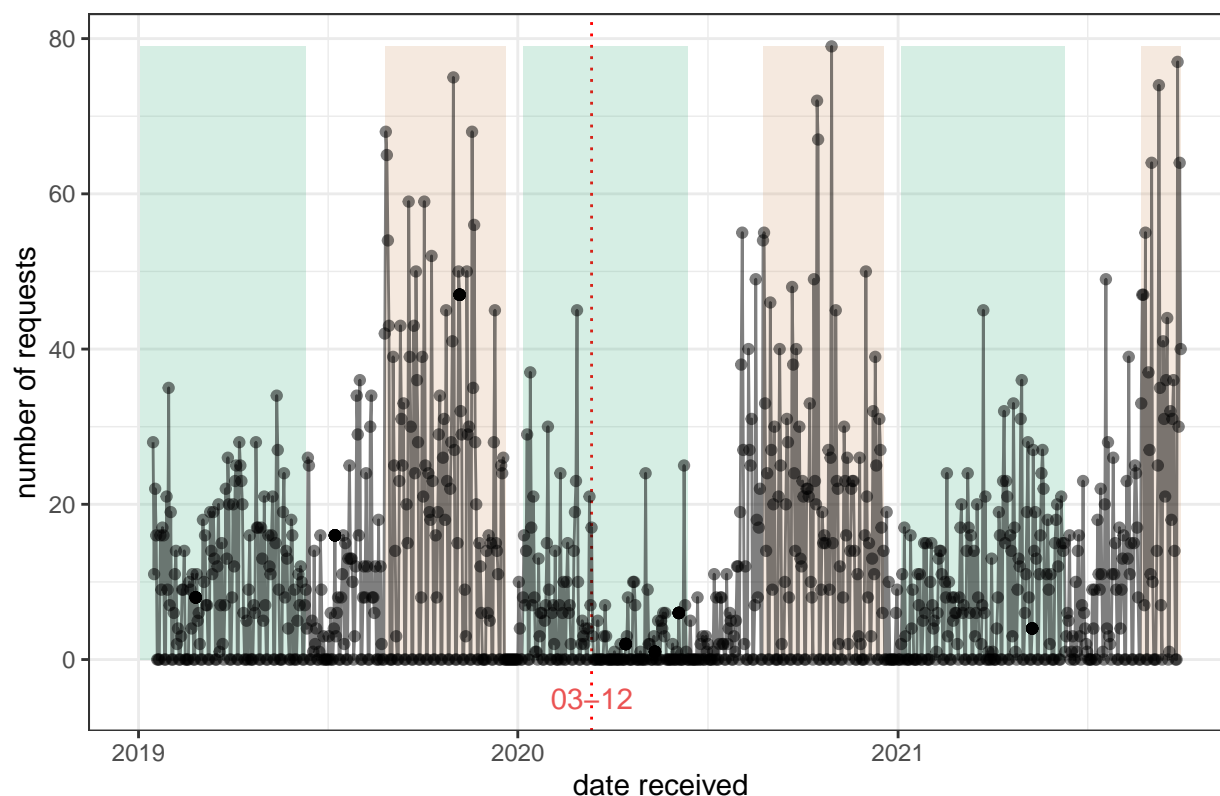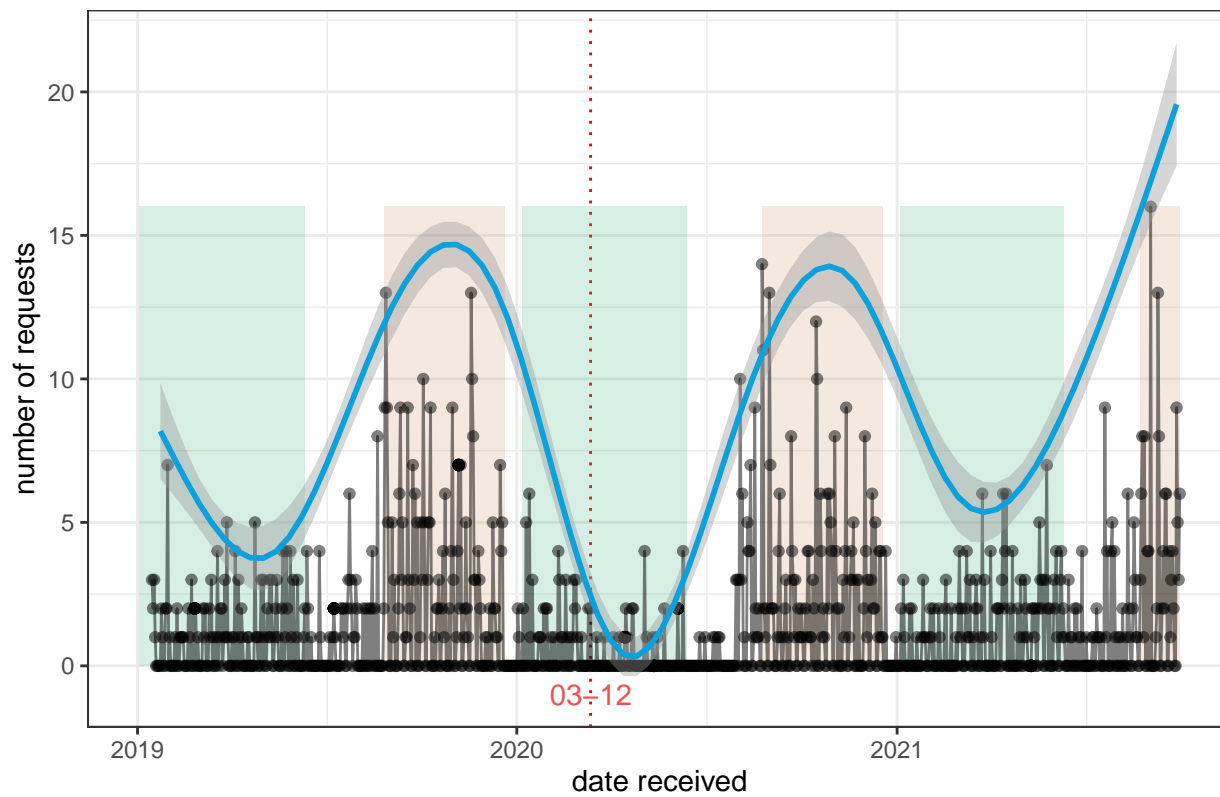
## Female Small
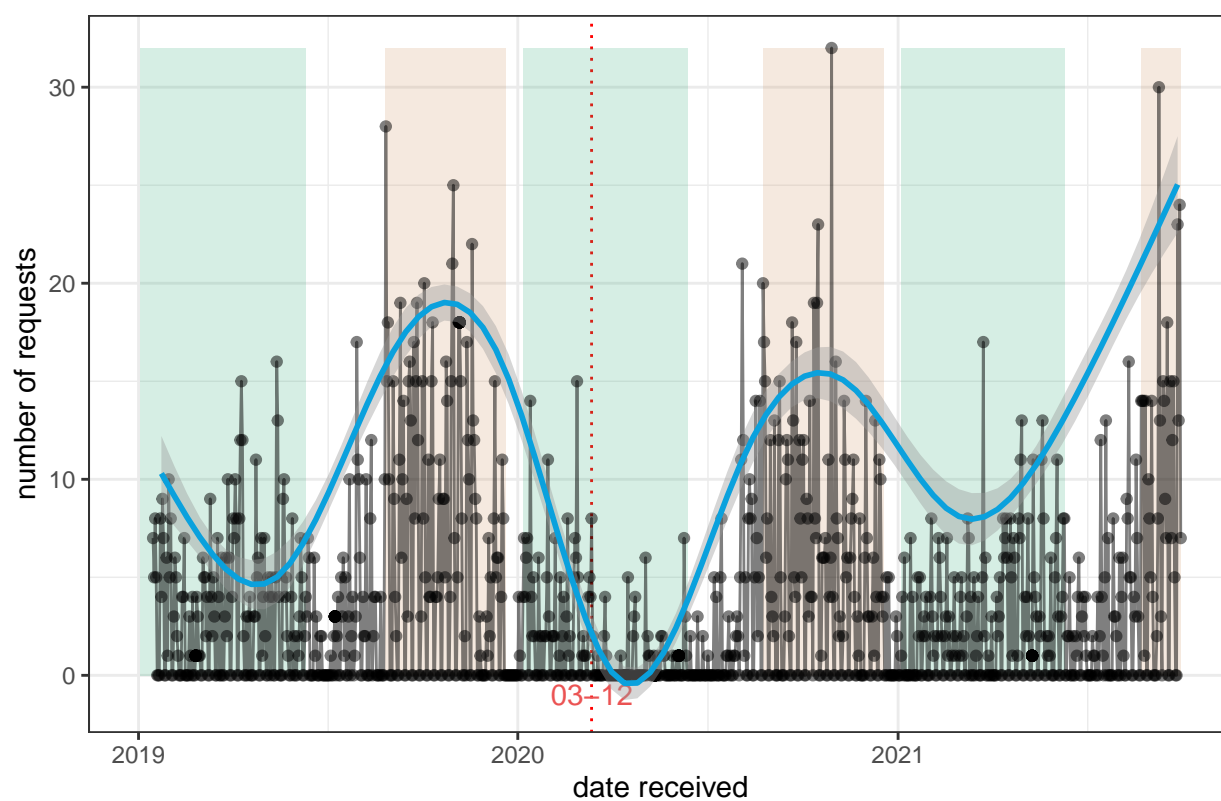


## Male Small

Female Large

Male Large

## 1.3    Plot of Total Number of Requests and Student Status

To see if student status has an impact on the the number of requests, let us look at the plot of the total number of requests by student status. The count plots for Student and Non-student follow a similar trend as before, so we can conclude that looking at just the total number of requests will be sufficient.



Total Number of Requests and Student Status
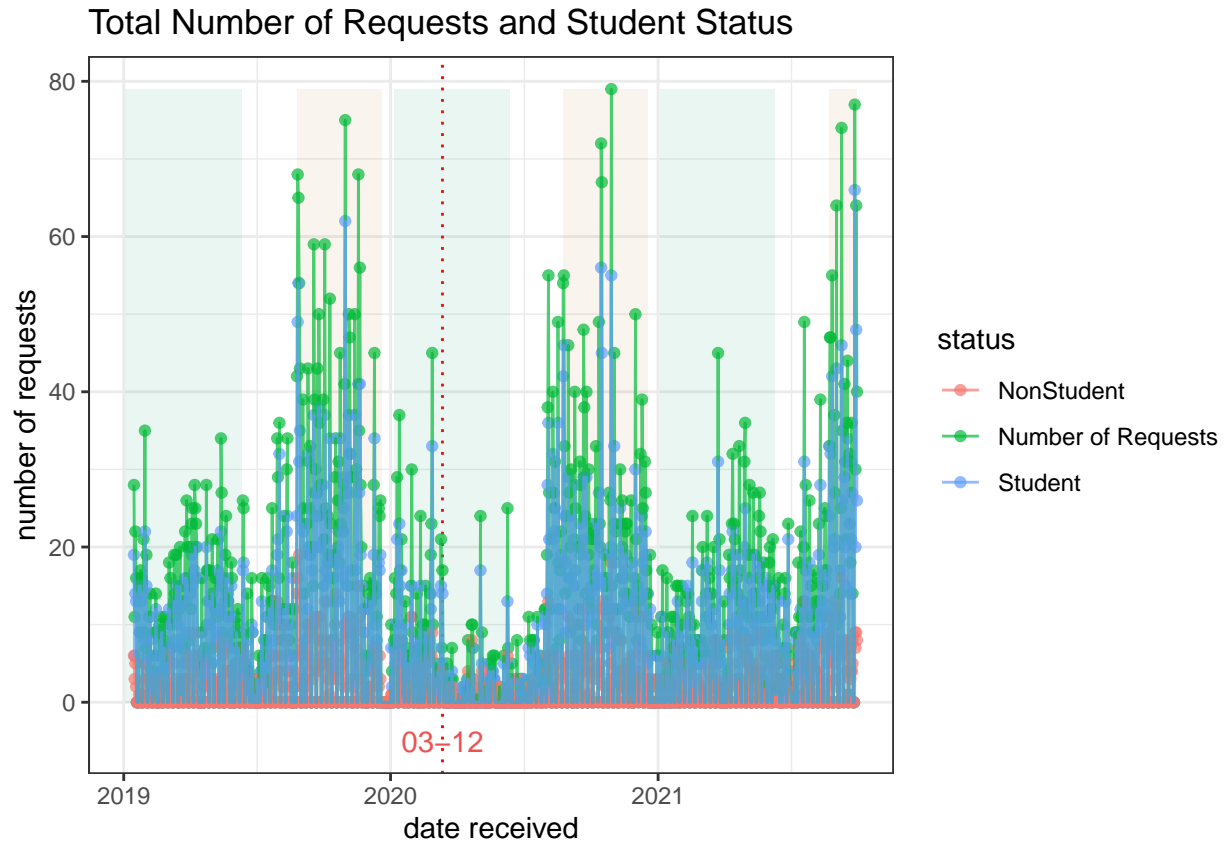
# 2   Request Lag Plots

Understanding the trends in the **lag time** is the first step towards improving order turnaround speed. The lag time is the number of days between an order being received and fulfilled. For example, if an order was received on March 2 and fulfilled on March 10, the lag time would be 8 days. To get view of how the lag time has changed over time, we plot lag time on the y-axis against the date the orders were received on the x-axis. If an order was received on March 2 2020 and filled on March 10 2020 that point would show up at March 2 2020 on the x-axis with a y-value of 8. We also stratified by gender to get a sense of if the lag time distributions was different for males and females. Because points can over lap, e.x. two orders were received on March 2 and both were filled in 8 days, we also adjusted the transparency of the points. A darker point means that there are actually many orders stacked at the same value. The vertical line at March 15, 2020 is to give a general date for when COVID may have started to affect operations.
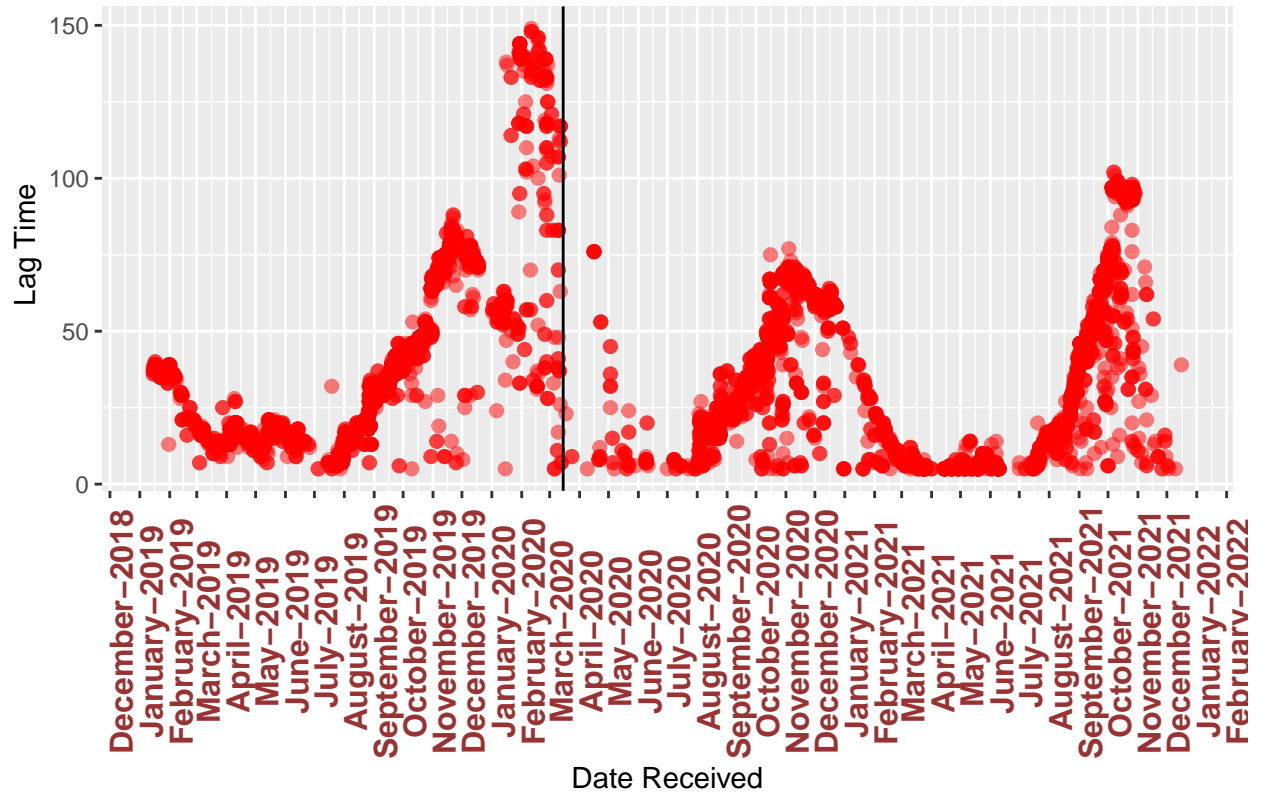


Request fulfillment lag time
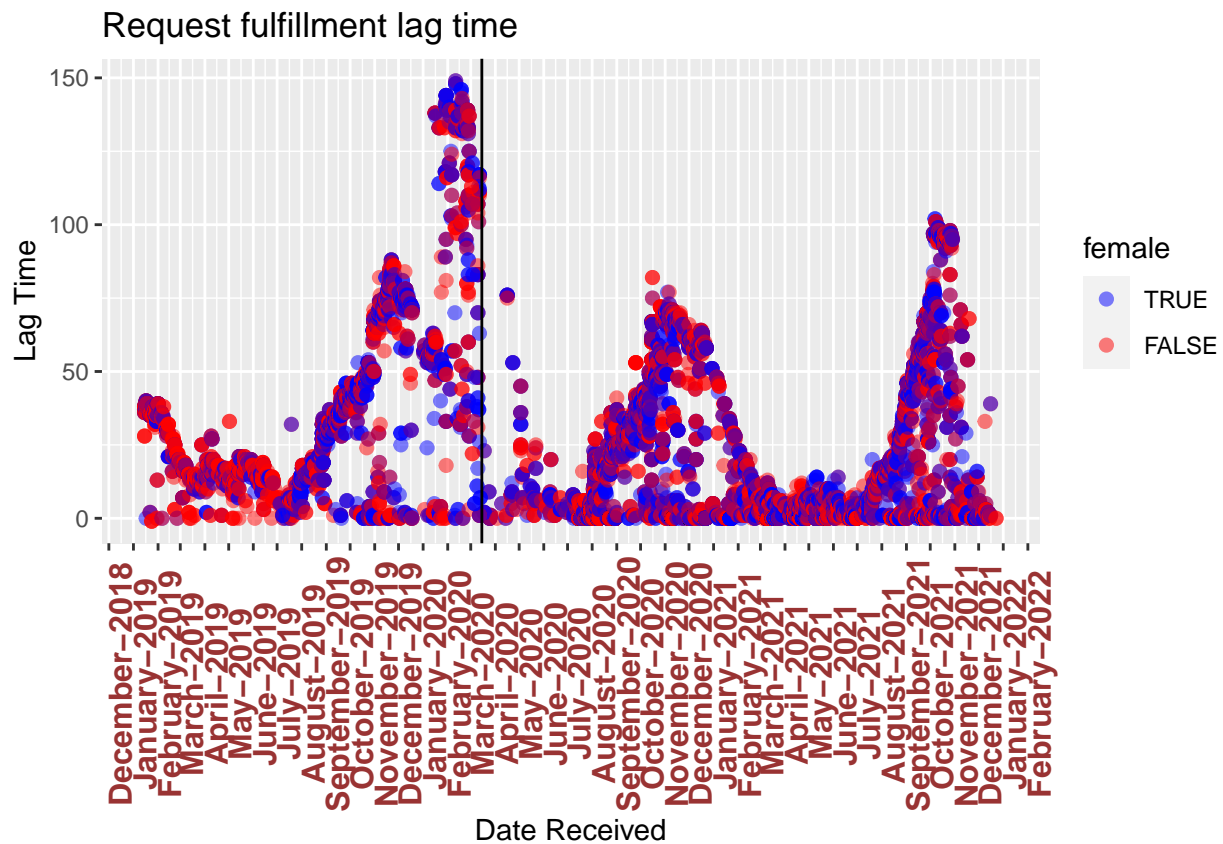
Next we filter out any orders that take less than 5 days to fulfill. This helps remove emergency orders that needed immediate attention and focus on more standard orders.
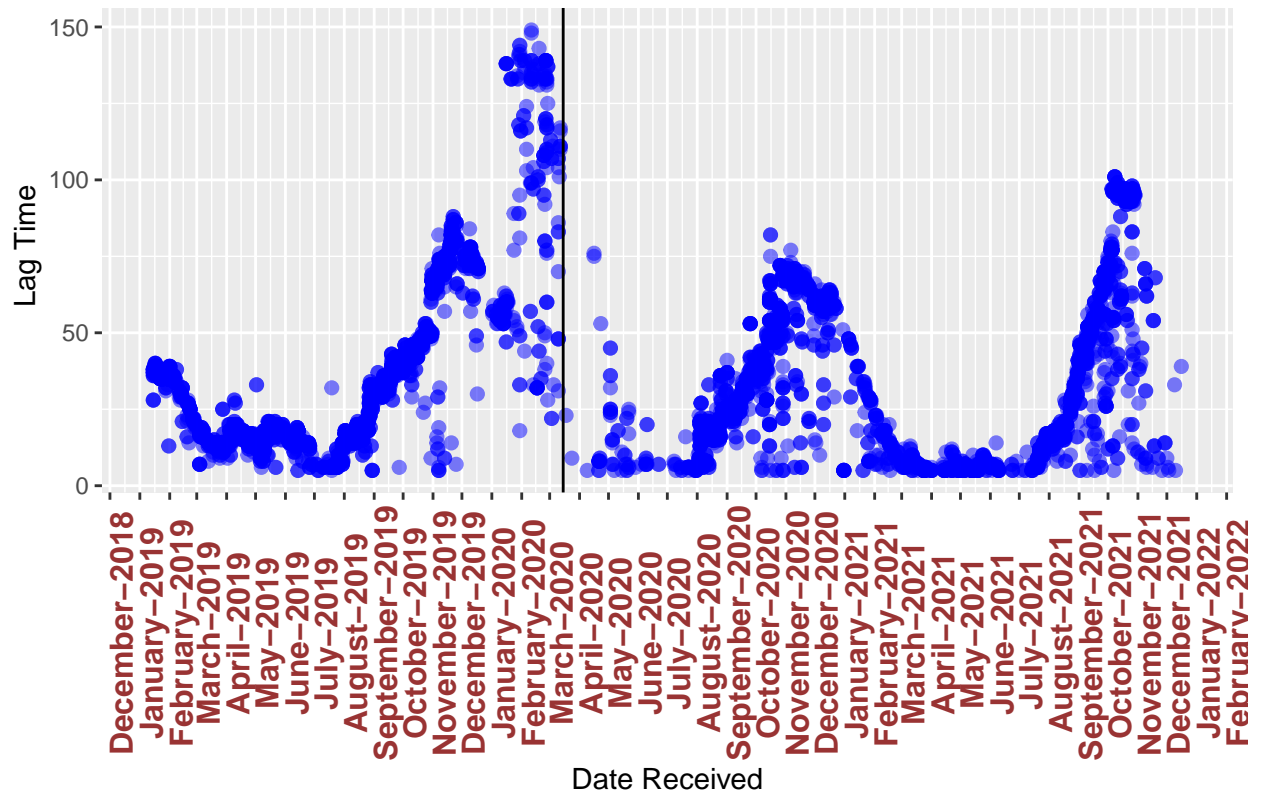
# Request fulfillment lag time > 5 days
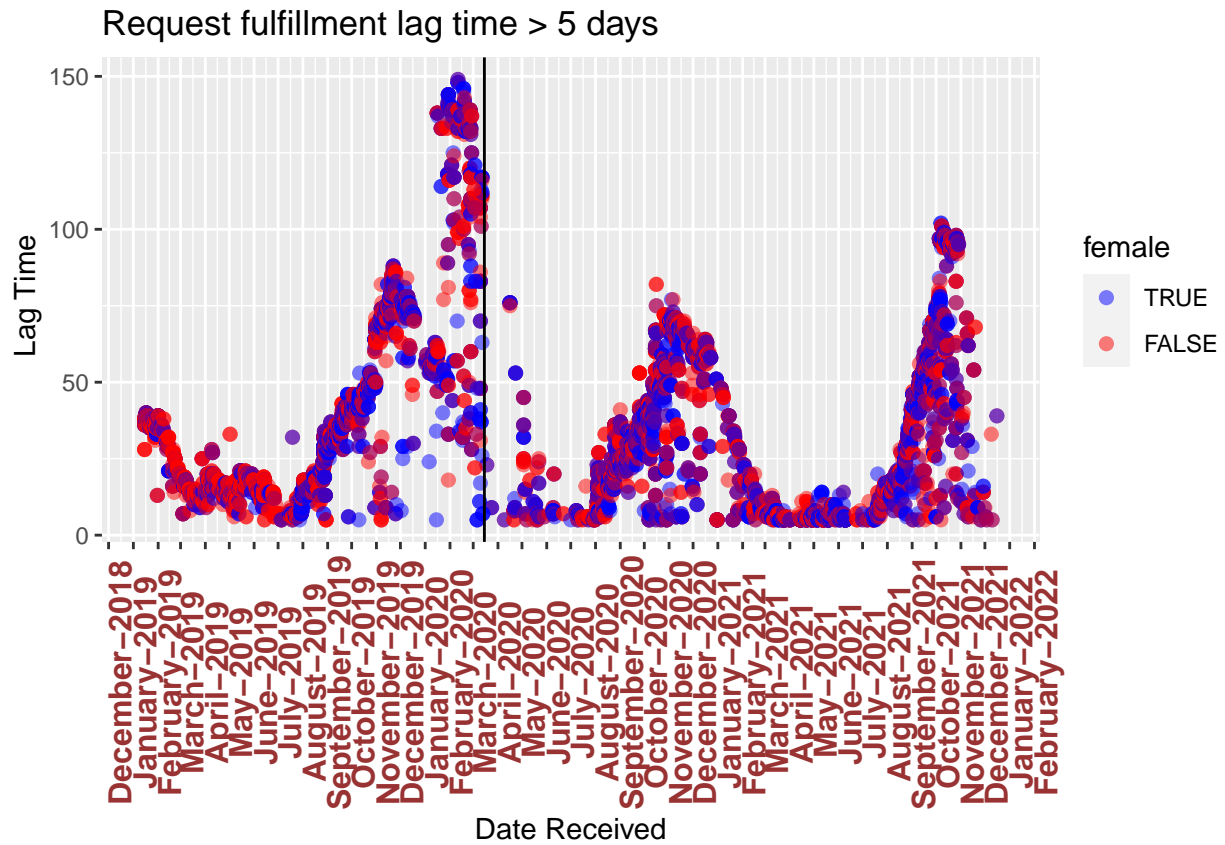


To reduce clutter, we also plot the requests for each gender separately.

## Request fulfillment for females



## Request fulfillment for males

The request distributions look similar for both males and females. There is a clear spike in request fulfillment time around the onset of COVID, when Note in the Pocket had to temporarily close and volunteering was not allowed

Finally, we stack all of the data up by month. This more easily visualizes trends that occur throughout a year.
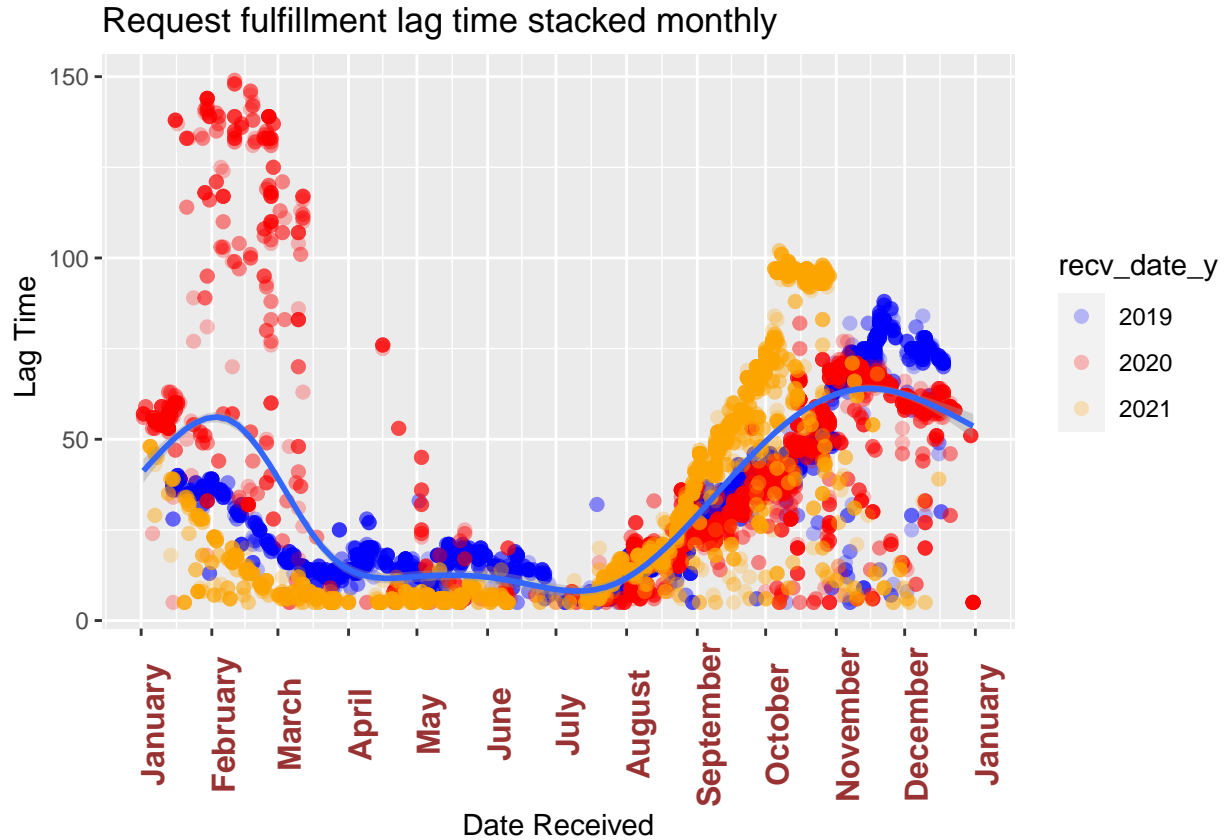


As expected, the summer months have fewer request and also less lag with requests ramping back up at the beginning of the academic year. We can clearly see the effect of the onset of COVID at the beginning of 2020 compared to the rest of the years.
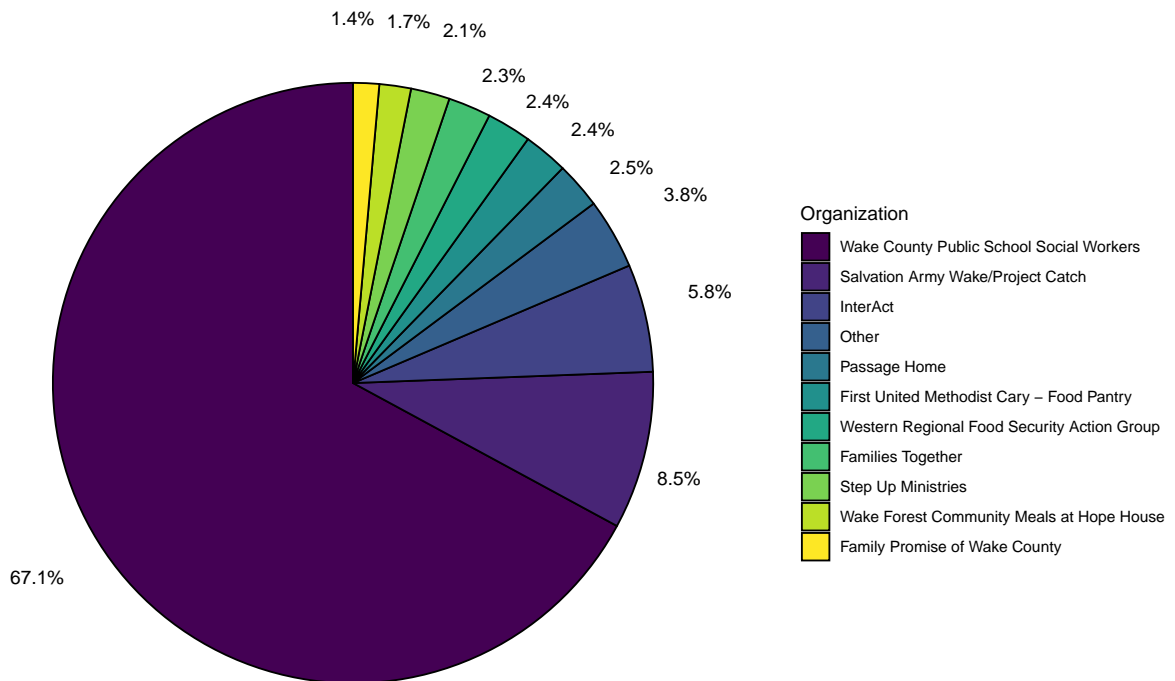
# 3   Map of Wake County Zip Codes

Because there are zip code maps in the following section, I include a map of the zip codes in Wake County, for reference (https://www.cccarto.com/nc/wake_zipcodes/).
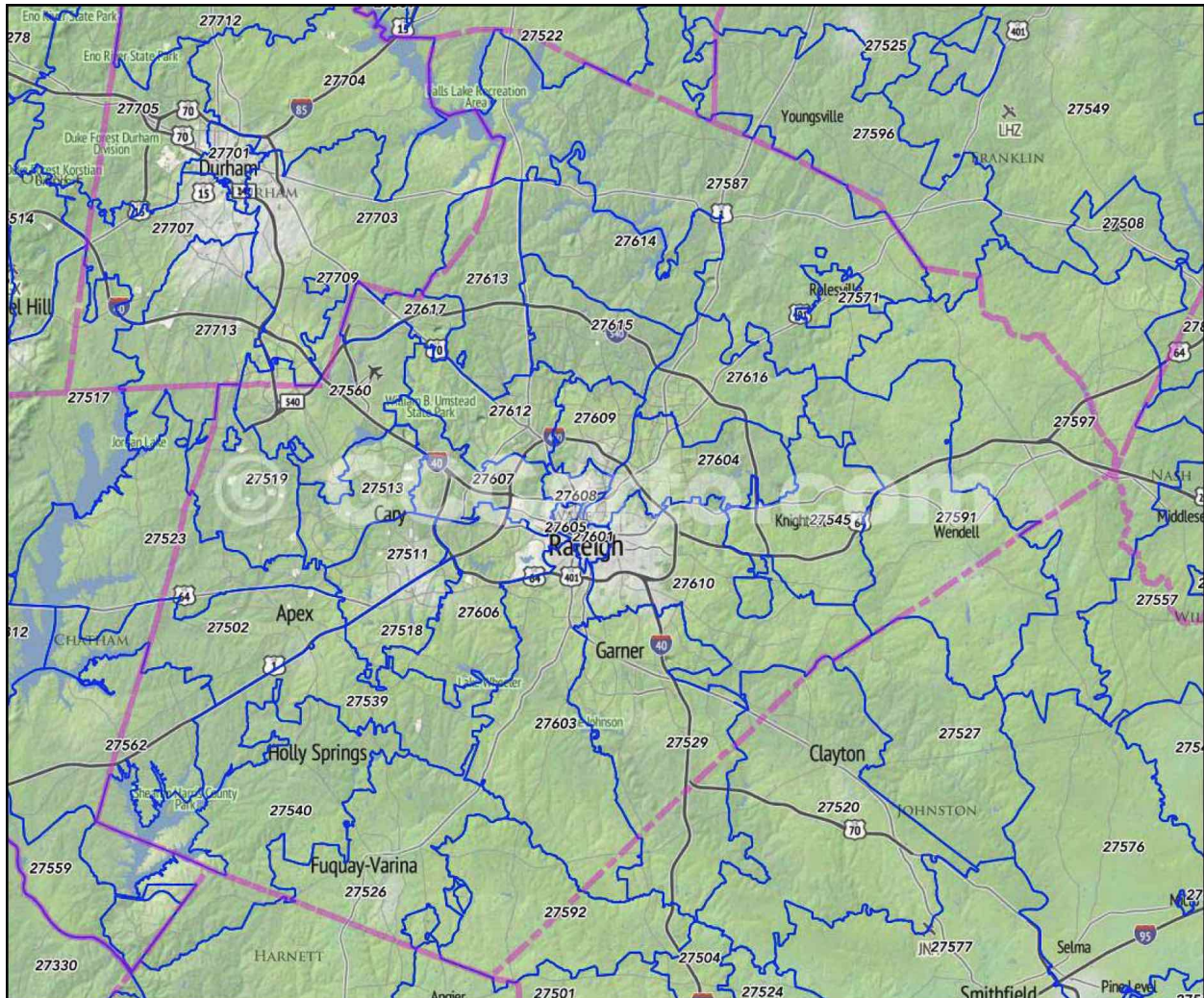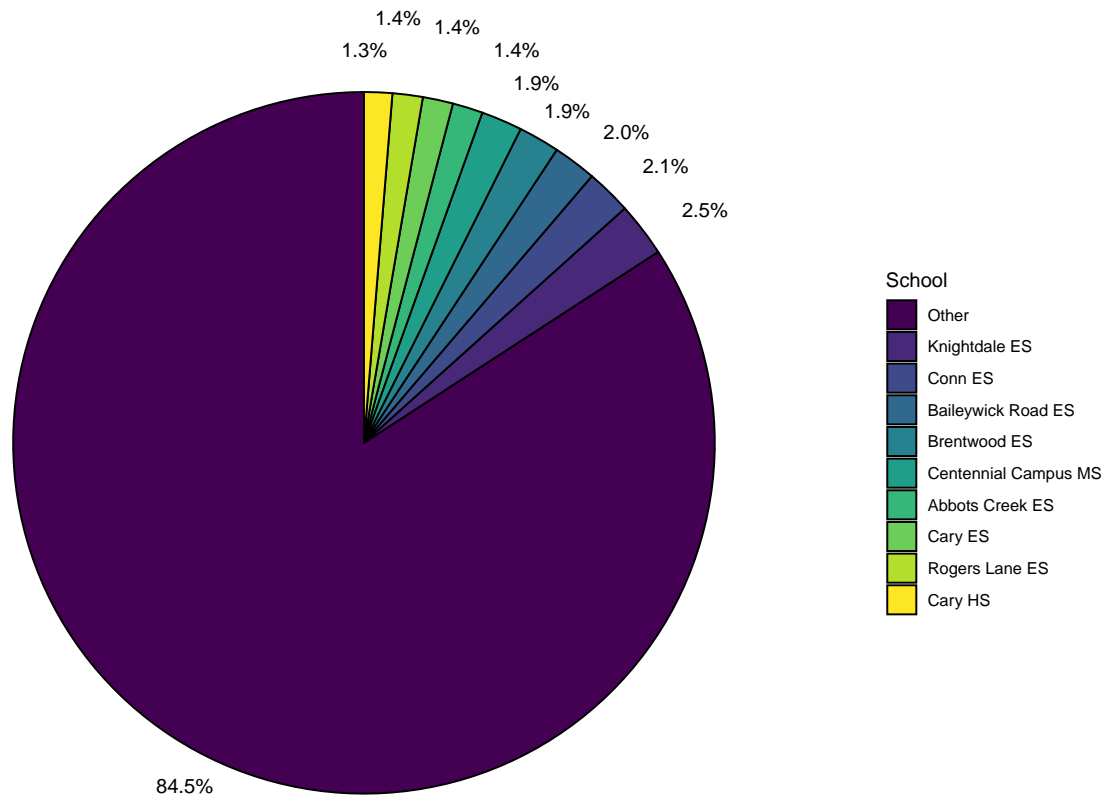
# 4 Exploratory data analysis for the categorical variables

We selected three categorical variables to explore: Organization, School, and City. We first calculated the number of individuals served by each organization/school/city. Then, based on the total number of individuals served, we calculated the proportion of service by each organization/school/city. We can find the top 10 organizations/schools/cities that mostly collaborated with the Note in the Pocket. Meanwhile, we mapped all the zipcode information except for missing value and outside North Carolina's zipcode. Combining the information from the pie chart, it is not surprising to see the Raleigh, Cary, Garner areas are the top three cities for the proportion. More service may be provided to the individuals in the Apex area, considering the influence of distance.
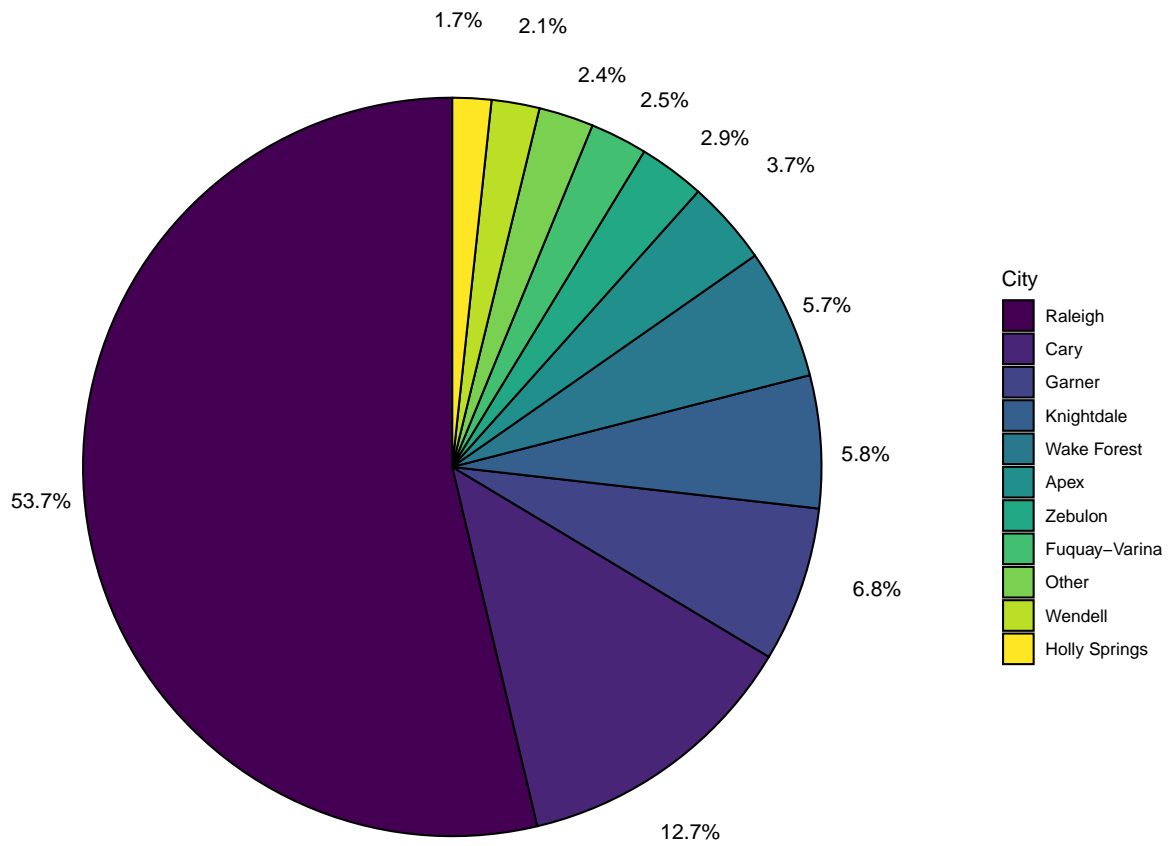
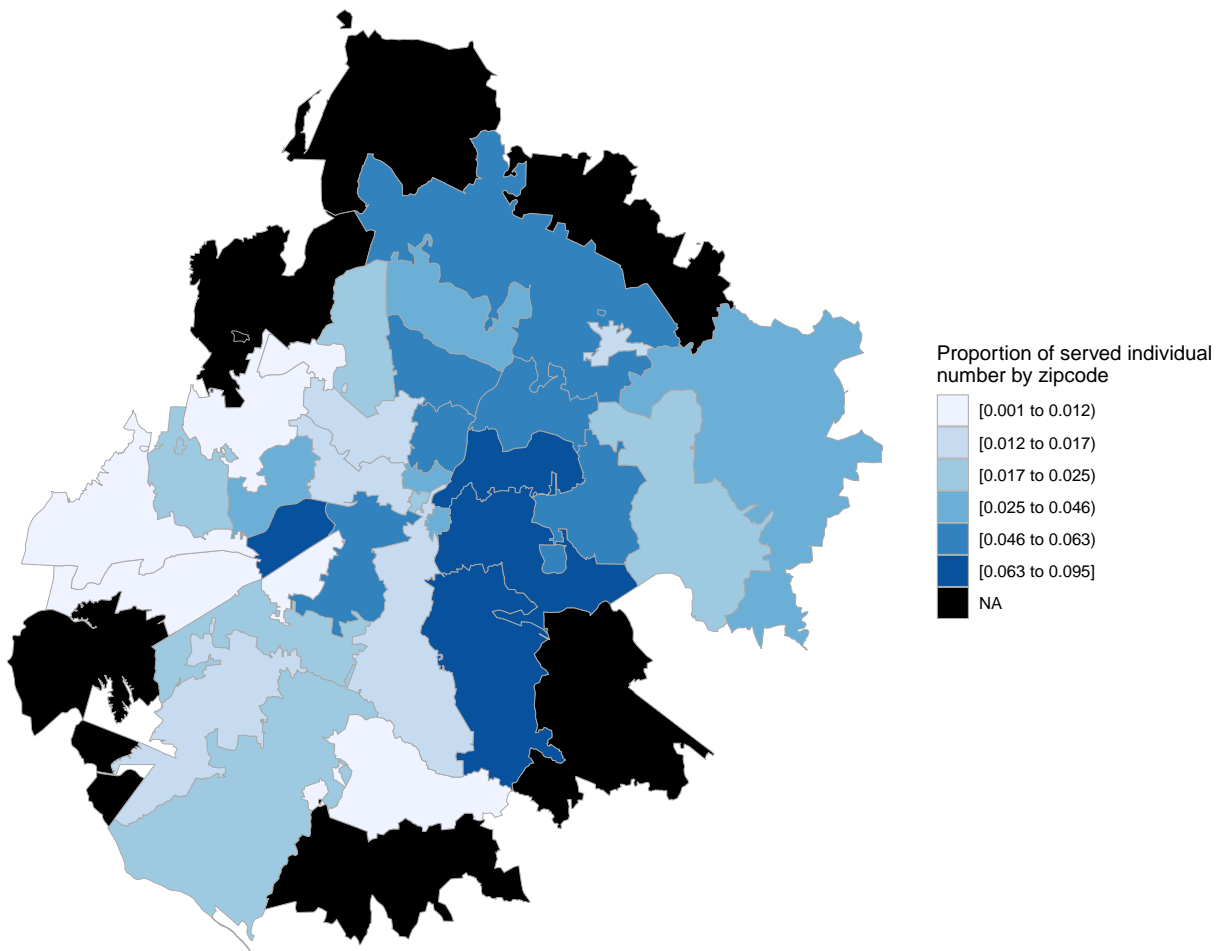Proportion of served individual number by organization

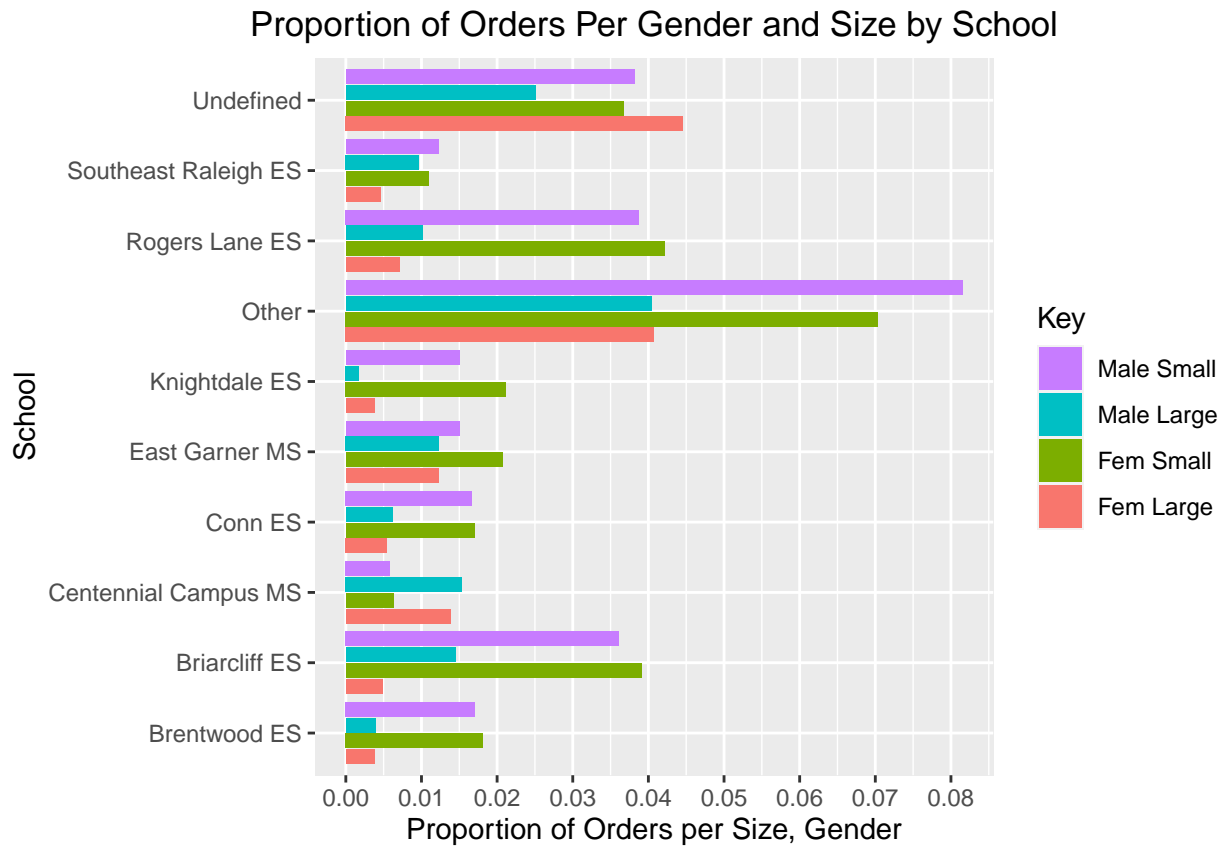Proportion of served individual number by school



School
- Other
- Knightdale ES
- Conn ES
- Baileywick Road ES
- Brentwood ES
- Centennial Campus MS
- Abbots Creek ES
- Cary ES
- Rogers Lane ES
- Cary HS

Proportion of served individual number by city

Note in the Pocket_Request order zipcode information map



Proportion of served individual
number by zipcode

- [0.001 to 0.012)
- [0.012 to 0.017)
- [0.017 to 0.025)
- [0.025 to 0.046)
- [0.046 to 0.063)
- [0.063 to 0.095]
- NA

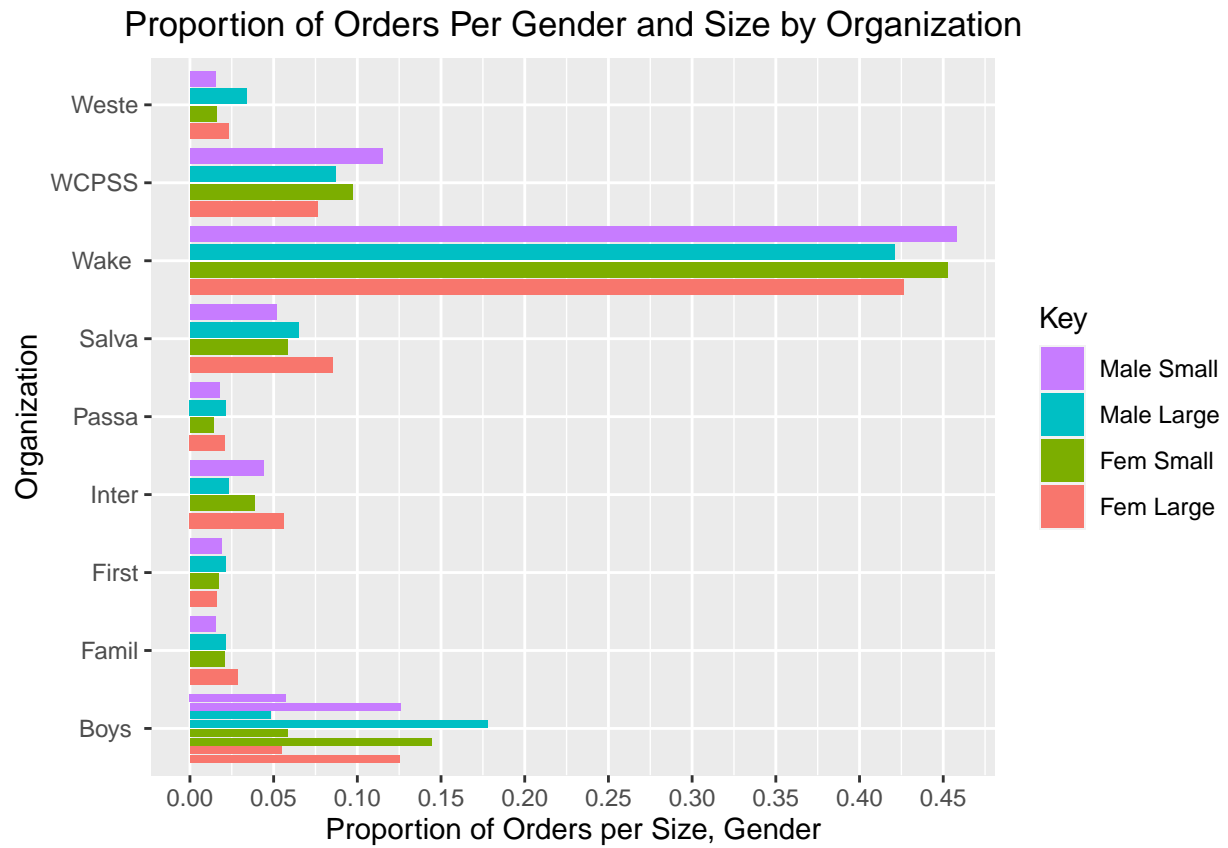**Proportion of Orders for Gender, Size by School:**

## Proportion of Orders Per Gender and Size by School
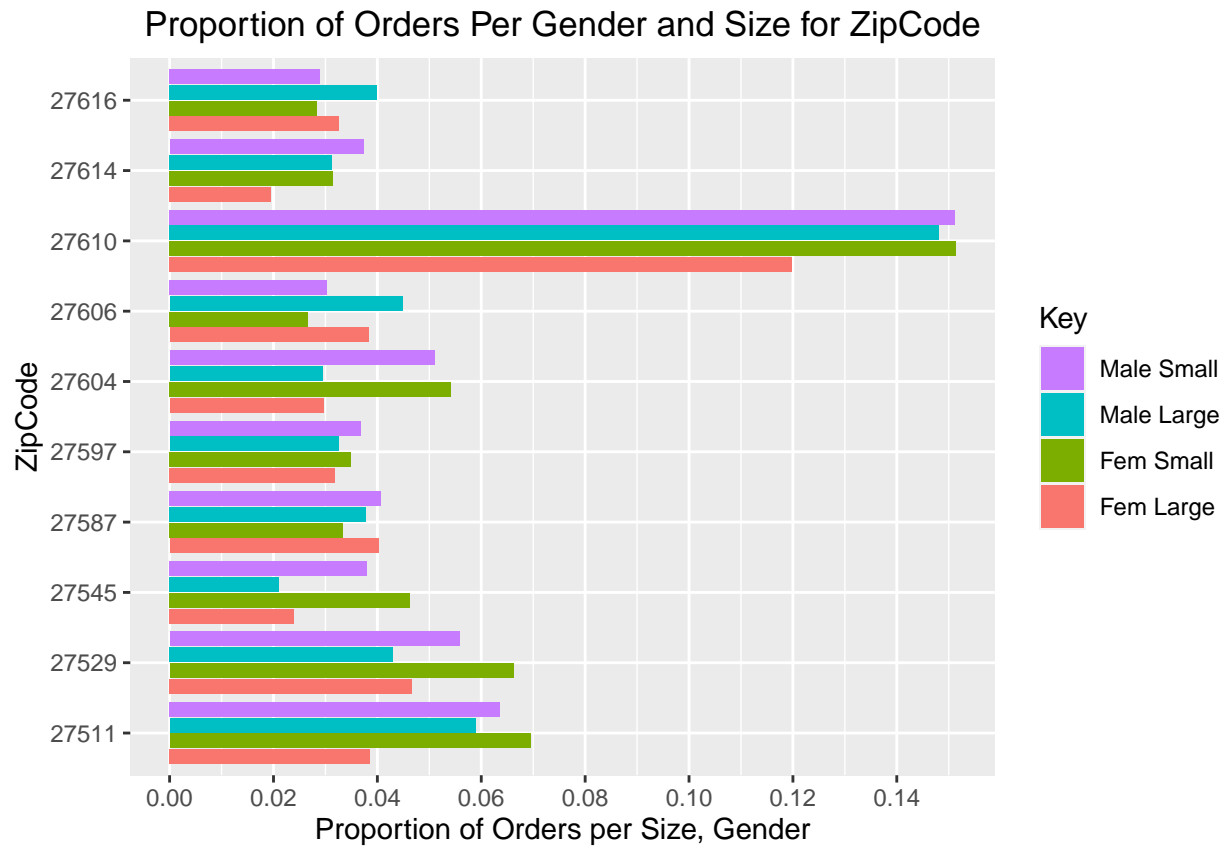


**Description:** The above graph plots the ratio of orders for a gender,size per school to the total number of orders for that gender,size over all schools.

# Proportion of Orders Per Gender and Size by Organization



**Description:** The above graph plots the ratio of orders for a gender,size per organization to the total number of orders for that gender,size over all organizations.

**Proportion of Orders for Gender, Size by ZipCode:**

## Proportion of Orders Per Gender and Size for ZipCode



**Description:** The above graph plots the ratio of orders for a gender,size per zip code to the total number of orders for that gender,size over all zip codes.

## Mean PCI vs Total Orders by Zip
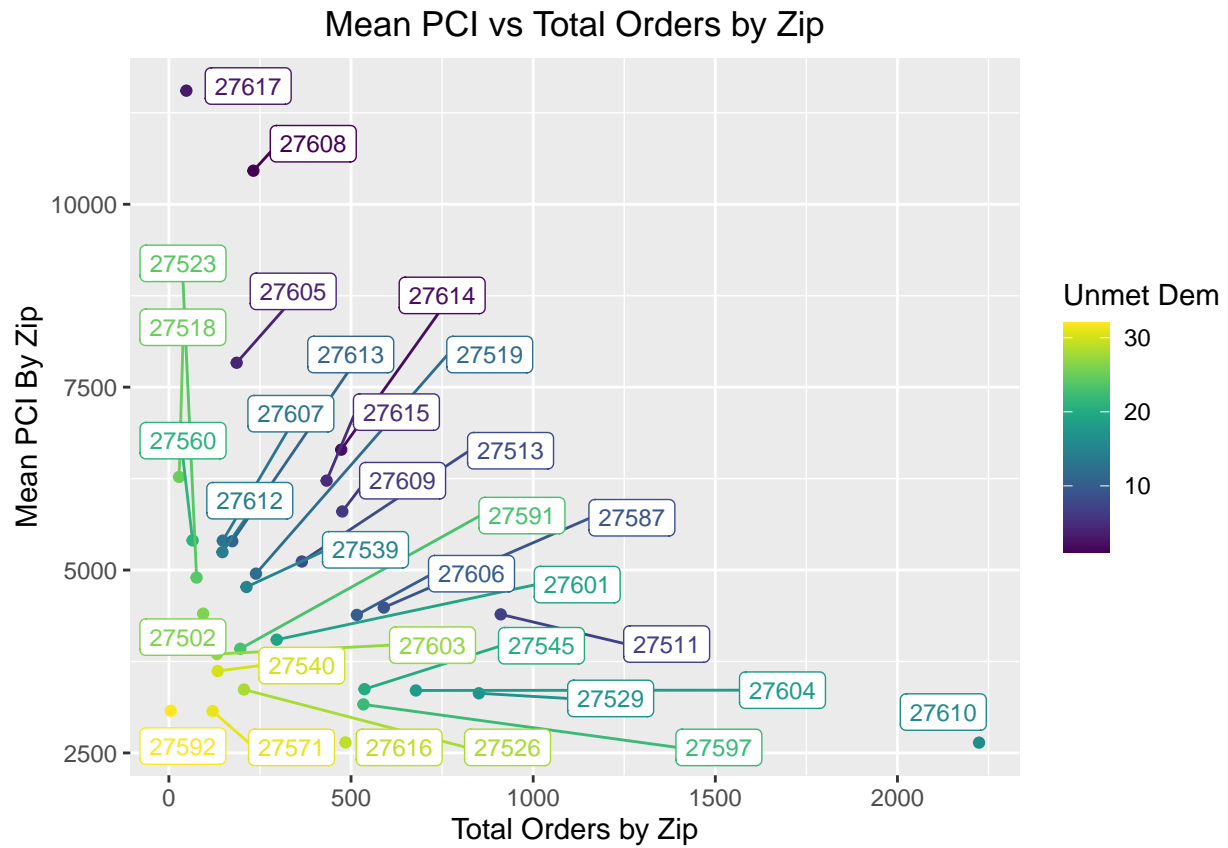


**Description:** The above graph plots mean PCI for a zip code versus total orders for that zip code. Coloring is based on a metric to measure unmet demand in that zip code, i.e. where mean PCI and number of orders are both low. More yellow/large Unmet Dem indicates greater unmet demand, whereas more purple/smaller Unmet Dem indicates less demand to be met.