PLAN 390 Homework 2 Write Up

Link to github repository: https://github.com/akunna1/PLAN\_390\_Homework\_2

Using the Fire Department Dataset:

Answers to questions:

1. How long does it take Wake County Fire to respond to incidents, on average (i.e. the

time between when an incident is dispatched and when firefighters arrive on the

scene)? (hint: you can subtract lubridate date columns from each other). [1 point]

* 318.7927 seconds (approximately 5.3 minutes) is the average response time for Wake County Fire to respond to incidents.
* I calculated that from arrival\_time – dispatch\_time = response\_time\_secs
* Average (response\_time\_secs). I had to exclude all the NA values here using na.rm = TRUE

2. Does this response time vary by station? What stations have the highest and lowest

average response times? [0.5 points]

* Each station has a unique average response time ranging between 200 to 500 seconds
* Station 29 has the highest average response time, which is 495.7640 secs
* Station 13 has the lowest average response time, which is 223.0000 secs

Showing the maximum average response time:

Table

Description automatically generated

Showing the minimum average response time:

Table

Description automatically generated

3. Have Wake County Fire’s response times been going up or down over time? What might

be the reason for these changes? [0.5 points]

* Yes, it has been going up and down over time as seen in the graph below.
* The changes in response time could be due to the availability of resources and the time of the day. For example, if Wake County Fire is called multiple times a day to the same area or neighborhood, response times in other areas may suffer. The time of day has an effect on response times. Graveyard shift response times (12am – 6am) are generally the slowest because there are generally fewer firefighters on duty ready to respond to emergencies during those times.

Chart, histogram

Description automatically generated

4. At what times of day are fire calls most likely to occur? [1 point]

* 22h (10pm)
* I got that by looking at the count for dispatch\_hour

5. The dataset contains all types of fire department calls, other than emergency medical

services (which are removed to protect privacy). The codes for the different incident

types can be found on page 3-22 of the National Fire Incident Reporting System

Complete Reference Guide. How many calls to Wake County Fire are recorded in this

dataset, and how many of them are actual fires? [0.5 points]

* 229,047 calls to Wake County Fire are recorded in this dataset
* 17,231 calls to Wake County Fire are actual fires

6. It is reasonable that firefighters might respond more quickly to some types of incidents

than others (e.g., a building fire, code 111 might be higher priority than a cat stuck in a

tree, code 542). Using the reference guide linked above to determine appropriate

incident codes, evaluate the average response time to actual fires. Is this response time

faster than the average response time for all incidents? [0.5 points]

* It takes 311.1936 seconds (approximately 5.2 minutes) for Wake County Fire to respond to actual fires. It is faster than the average response time for all incidents, which was 318.7927 seconds (5.3 mins)

7. Repeat the analysis for questions 2-4 for actual fires, rather than all incidents. [2 points]

2) Does this response time vary by station? What stations have the highest and lowest

average response times?

* Each station has a unique average response time ranging between 200 to 600 seconds
* Station 29 has the highest average response time, which is 586.3713 secs
* Station 13 has the lowest average response time, which is 232.7666 secs

Showing the maximum average response time:

Table

Description automatically generated

Showing the minimum average response time:

Table

Description automatically generated

3) Have Wake County Fire’s response times been going up or down over time? What might

be the reason for these changes?

* Yes, it has been going up and down over time as seen in the graph below.
* The changes in response time could be due to the availability of resources and the time of the day. For example, if Wake County Fire is called multiple times a day to the same area or neighborhood, response times in other areas may suffer. The time of day has an effect on response times. Graveyard shift response times (12am – 6am) are generally the slowest because there are generally fewer firefighters on duty ready to respond to emergencies during those times.

Chart

Description automatically generated

4) At what times of day are fire calls most likely to occur?

* 21hr (9 pm)
* I got that by looking at the count for dispatch\_hour

New columns I created:

**arrival\_time:** includes arrival date and time. It was added to the fire\_incidents\_data table.

**dispatch\_time:** includes dispatch date and time. It was added to the fire\_incidents\_data table.

**cleared\_time:** includes cleared date and time. It was added to the fire\_incidents\_data table.

**response\_time\_secs**: It was added to the fire\_incidents\_data table. It was created from arrival\_time – dispatch\_time.

**dispatch\_hour:** Itwas created from dispatch date and time; it includes only hours. It was added to clean\_fire\_incidents\_data

Data frames used and created:

* **fire\_incidents\_data:** the original data. I added 4 extra columns to it i.e. arrival\_time, dispatch\_time, cleared\_time (was not used) response\_time\_secs,
* **clean\_fire\_incidents\_data:** similar to the fire\_incidents\_data dataset but includes only response\_time\_secs that are greater than 0 and are not “NA” values.
* **station\_summary:** was created using summarize and group by station. It was used to find which station had the highest and lowest average response time in question 2.
* **dispatch\_hour\_count:** shows count for dispatch hours
* **actual\_fire\_data:** includes only actual fire i.e. where the numbers in incident\_type column are less than 200 (all 100s values).
* **actual\_fire\_station\_summary**
* **actual\_fire\_dispatch\_hour\_count**

Small calculations made using columns:

* **avg\_response\_time:** the average of response\_time\_secs in the fire\_incidents\_data dataset
* **actual\_fire\_avg\_response\_time**

Notes about missing data for each data set:

* **fire\_incidents\_data** contains missing data which is represented by “NA”. It has 229,047 entries.
* To calculate **avg\_response\_time** for question 1, I removed all the “NA” values.
* To create the **clean\_fire\_incidents\_data**, I had to remove all **response\_time\_secs** (i.e., a column in the data set) that were less than 0 and were “NA” values (i.e. missing values). It has 184,382 entries.
* To create **station\_summary**, I removed all the “NA” values (i.e., missing values)
* To create **actual\_fire\_data**, I removed all the “NA” values (i.e., missing values). It contains 17,231 entries.