Has Rush Hour Traffic Drastically Changed in Response to the COVID-19 Pandemic?

Introduction:

Due to the ongoing COVID-19 pandemic, everyday life has been significantly impacted in most of the world, including our own local community, the Northwest section of Santa Clara County. As avid cyclists, a common feature of our own personal lives has always been dealing with heavy traffic on weekdays during rush hour, the time during which the most people commute and congestion is therefore at its highest. On our numerous bike rides, we previously witnessed traffic jams on highways during rush hour that back up for what appears to be miles on end. Since the majority of jobs only involve working during weekdays, the amount of traffic on weekends during the same times as when rush hour would normally occur is usually significantly less, barring occasions such as holidays or vacation. As such, we typically observe rush hour traffic on weekdays to be much heavier than on weekends. However, the current social distancing and stay-in-place measures have significantly reduced the amount of traffic during rush hour as many people are now remaining at home and away from the road. Both of us have been frequently biking around the local area to keep healthy and stay positive throughout these trying times. On our separate rides, we have both observed drastic reductions in traffic during rush hour on both weekdays and weekends, to the extent where we perceive no difference in rush hour traffic volume between weekdays and weekends. Though we are aware that the stay-in-place measures have significantly altered traffic flow, we felt skeptical that these changes in traffic have been so drastic as to actually equalize the difference in rush hour traffic between

weekdays and weekends. As such, we decided to perform a sanity check and determine if rush hour traffic was still greater on weekdays than on weekends in our area using statistical methods.

Statistical Question:

Is there still less afternoon rush hour traffic in the Northwest area of Santa Clara County on weekends than on weekdays despite the ongoing situation with COVID-19?

$$H_0: \mu_{weekends} - \mu_{weekdays} = 0 \qquad H_a: \mu_{weekends} - \mu_{weekdays} > 0$$

$$\alpha = 0.05$$

Where… $\mu_{weekends}$ = the population mean time for 40 cars (in seconds) to pass by on highways in the Northwest area of Santa Clara County during rush hour on weekends (Saturday or Sunday)

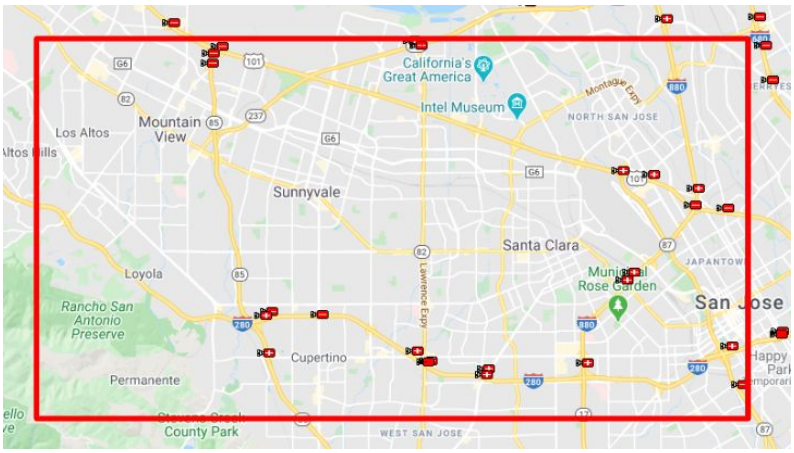
$\mu_{weekdays}$ = the population mean time (in seconds) for 40 cars to pass by on highways in the Northwest area of Santa Clara County during rush hour on weekdays (Monday through Friday).

Methodology:

We conducted an observational study that collected data on the traffic in our local highways during rush hour in order to answer our statistical question. Highways were used for our study because unlike other roads, cars travelling through them are not affected by stop signs or traffic lights, which could impact the quality of our data. Moreover, traffic on highways is usually less localized compared to other types of roads, meaning it is more indicative of the general level of traffic in a given area. These conditions made highways the most suitable road for study in our research. We chose highways in the Northwest section of Santa Clara County

because we sought to confirm our suspicions that there was a significant change in traffic in our local community. It is noteworthy to mention that shelter-in-place orders in this county have been some of the strictest in the nation, and as such this area likely has had some of the greatest changes in traffic behavior. The California Department of Transportation provides live camera feeds for various locations throughout the state highway system. The feeds provide an unobstructed view of a section on the highway. An example image is provided below.

We define the Northwest area of Santa Clara County that comprises our local area to be the region within the red box on the map below.



The red camera icons indicate a live camera feed location. The area we chose to examine had 24 locations that are shown on the map above. We decided to use these provided camera feeds in place of live observation in order to follow the ongoing stay-in-place orders. We took a simple

random sample of 5 locations to observe by assigning each camera feed a number from 1 to 24. We then used a random number generator to select 5 numbers without replacement in this range. The numbers were 18, 3, 20, 12, and 6, which corresponded to the following locations:

- Hwy 280 at Saratoga Ave
- Hwy 880 at The Alameda
- Hwy 101 at Moffett Blvd
- Hwy 85 at Stevens Creek
- Hwy 680 at King Road

Our observation period was over the course of two weeks. We decided to collect data on Mondays for weekdays and on Saturdays for weekends, as we could not find any basis for believing that there would be significant variation in traffic behavior between the various weekdays, or between Saturday and Sunday. We chose the afternoon rush hour period to be between 5:00 PM and 6:00 PM. For each day, we took three data points from each of the four locations, for a total of 15 data points. Over the 2 week period, this translated to 30 data points for weekdays and 30 data points for weekends, for a combined total of 60 data points. Each data point was randomly assigned a time for data collection by generating random numbers between 0 and 59. For instance, if 12 was drawn, the data for that point would be collected at 5:12 PM.

Our method for measuring the flow of traffic was based on timing how long it would take for forty vehicles to pass a specific point on the highway, all lanes included. A highway with greater traffic flow would have a lower time, while one with lower traffic would have a higher time. We felt this measurement would be indicative of the general volume of traffic at a given time. In order to eliminate as many external variables, we made sure to check prior to data

collection that the no lanes of the highway were closed during measurement time, that the weather was pleasant with good visibility, and that there was no ongoing construction or any nearby accidents that would have interrupted traffic flow. Fortunately, all of these conditions were satisfied during the time intervals we had randomly selected and we were able to collect data as planned. We recorded our results in a table that will be shown in the Data Display section of our report.

Once we recorded 30 data points for weekdays and 30 data points for weekends, we took the mean time for weekdays and weekends, respectively, to determine the sample means that would be used to conduct our hypothesis test. We did not receive any guidance from an adult for any aspect of our observational study or for the writing of our report.

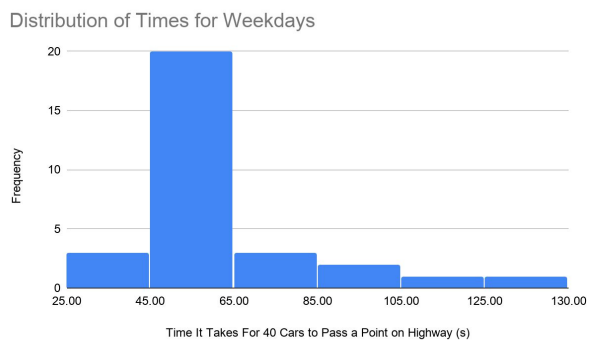
Data Display:

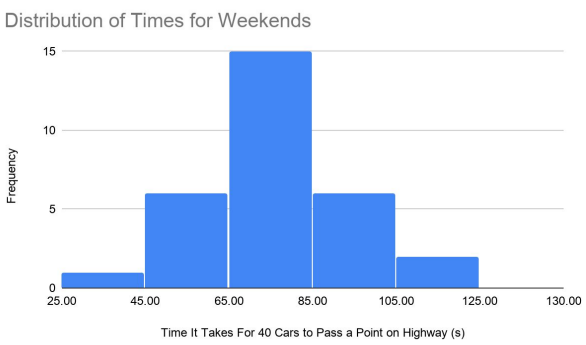
The raw data from our observational study is displayed below:

Week 1 (4/13-4/19)				
Highway	Weekday Time	Timestamp	Weekend Time	Timestamp
280 @ Saratoga Ave (SB)	49.83	5:58 PM	76.52	5:29 PM
	54.36	5:49 PM	73.16	5:18 PM
	29.46	5:26 PM	68.35	5:55 PM
880 @ Alameda (SB)	57.96	5:29 PM	79.88	5:26 PM
	47.7	5:16 PM	73.76	5:32 PM
	64.49	5:05 PM	73.1	5:54 PM
101 @ Moffett Blvd (SB)	38.37	5:18 PM	54.32	5:43 PM
	45.21	5:56 PM	62.18	5:47 PM
	46.43	5:30 PM	67.32	5:23 PM
85 @ Stevens Creek (SB)	67.97	5:10 PM	75.4	5:45 PM
	108.1	5:22 PM	81.4	5:07 PM

	126.39	5:53 PM	123.21	5:59 PM
680 & King Road (SB)	46.72	5:44 PM	70.32	5:44 PM
	52.59	5:42 PM	72.46	5:31 PM
	49.87	5:06 PM	81.48	5:57 PM
Week 2 (4/20-4/26)				
Highway	Weekday Time	Timestamp	Weekend Time	Timestamp
280 @ Saratoga Ave (SB)	47.73	5:03 PM	96.56	5:48 PM
	50.23	5:36 PM	62.41	5:12 PM
	46.98	5:20 PM	37.25	5:32 PM
880 @ Alameda (SB)	57.74	5:13 PM	74.09	5:25 PM
	59.4	5:51 PM	82.46	5:33 PM
	88.8	5:50 PM	87.62	5:41 PM
101 @ Moffett Blvd (SB)	36.96	5:04 PM	56.26	5:15 PM
	49,62	5:16 PM	60.81	5:06 PM
	57.54	5:54 PM	49.25	5:46 PM
85 @ Stevens Creek (SB)	76.85	5:39 PM	91.54	5:01 PM
	100.65	5:02 PM	107.83	5:27 PM
	73.28	5:23 PM	85.59	5:59 PM
680 & King Road (SB)	61.82	5:40 PM	87.21	5:37 PM
	57.39	5:11 PM	75.38	5:09 PM
	64.35	5:38 PM	90.46	5:37 PM

After collecting the data on the traffic volume of highways, we graphed the times for weekdays and weekends into two separate histograms displayed below:





From the histograms above, it is clear that the distribution of traffic times for weekdays is skewed to the right towards the greater values, while the distribution of traffic times for weekends is approximately symmetrical and appears approximately normal. The median of the weekday distribution appears to be smaller than the median of the weekend distribution, and the weekday distribution appears to have a greater range than the weekend distribution. Neither the weekday distribution nor the weekend distribution appears to have any gaps, and there are no discernable outliers. From our inspection of the data visualizations, it appears that the times for weekdays tend to be less than the times for weekends based on the location of the centers of the distributions. However, there seems to be more variation among the times for weekdays than among the times for weekends. To better understand our data, we also computed 1-variable statistics for both distributions, which are displayed in the table below:

Day	N	Mean	Std	Min	Q1	Median	Q3	Max
Weekends	30	75.919	17.091	37.25	67.32	74.735	85.59	123.21
Weekdays	30	60.493	21.385	29.46	47.7	55.875	64.49	126.39

Clearly, we can see that our observations regarding the visualizations are confirmed through the 1-variable statistics. The times for weekends tend to be greater than those for weekdays, as demonstrated by mean ($75.919 > 60.493$) and median ($74.735 > 55.875$) values. Greater times means less traffic because there are less cars on the road, which increases the amount of time taken for 40 cars to cross a given point on the highway. Already, there seems to be evidence in support of our suspicions against our original perceptions, which suggested that the amount of traffic during rush hour is the same for weekdays and weekends.

Data Analysis:

We conducted a two-sample t-test for a difference between two means to determine whether we had convincing statistical evidence to believe that there was still less traffic during weekends than during weekdays. A t-test was used instead of a z-test because we do not know the population standard deviations. The hypotheses are defined in the statistical question. Once again, it is important to note that a greater mean time means smaller traffic flow, while a smaller mean time means greater traffic flow.

Next, we verify the conditions for statistical inference. Since the data was collected using random sampling, the random condition has been satisfied. The sizes of both the weekday and weekend samples are 30, which is sufficiently large to meet the conditions for the central limit theorem. As such, we can assume the sampling distribution of the sample means for weekend and weekday times is approximately normal, thus satisfying the normality condition. Finally, we believe the independence condition has been met because our data is a metric of traffic flow, and we believe that our 30 observations of 40 cars comprise less than 10% of the total rush hour traffic on highways for both weekdays and weekends in the Northwest area of Santa Clara County, as the county population is approximately 1.928 million people[1] and the volume of afternoon rush hour traffic associated with such a population should be large enough to prevent a major violation of the independence condition.

We test our hypotheses using the data collected from our observational study. The sample mean $\bar{x}_{weekends}$ is 75.919 and the sample standard deviation $S_{weekends}$ is 17.091, while the sample mean $\bar{x}_{weekdays}$ is 60.493 seconds and the sample standard deviation $S_{weekdays}$ is 21.385. Since the size of both samples is 30, the degrees of freedom is df = 30 - 1 = 29.

The test statistic is calculated from the following formula:

$$t = \frac{(x_1 - x_2)}{\sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}}}$$

Here, 1 is the weekend sample and 2 is the weekday sample. The calculated t-statistic is 3.086. Given that df = 29, we determine the p-value to be p = P(t > 3.086) = 0.00222.

Conclusion:

Since 0.00222 < 0.05, we believe that there is convincing statistical evidence to reject the null hypothesis and accept the alternative hypothesis that the population mean time for 40 cars to pass by on highways during rush hour on weekends is greater than the population mean time for 40 cars to pass by on highways during rush hour on weekdays. As such, it can be concluded that even in the current situation with the stay-in-place measures throughout Santa Clara County, there is still less rush hour traffic on weekends than on weekdays as usual.

Reflection:

Overall, we feel that the observational study went well. Our findings provided evidence against our initial perceptions while cycling that there was no difference between afternoon rush hour traffic on weekdays and on weekends. Instead, our results concluded that there is still less afternoon rush hour traffic on weekends than on weekdays. However, we did face some limitations with our study. First, our focus on highways may not be entirely indicative of the afternoon rush hour traffic that we observed while cycling on local roads. Our data also only covered a relatively small area of the San Francisco Bay Area, and it is possible we may have arrived at different conclusions if we examined other areas in the study.

Ultimately, we were able to truly understand the importance of statistics in the real world. What we may perceive or believe may not necessarily be the case and can be disproven by analyzing data through statistical methods and making conclusions based on the results. Though the COVID-19 outbreak has disrupted many aspects of our everyday lives, perhaps it is somewhat comforting to know that a normal nuisance like rush hour traffic still remains relatively the same in its behavior. Next steps to expand upon this initial study include collecting more data in other regions and on local roads in order to determine whether our conclusions hold in areas besides our nearby highways. We could also examine changes in traffic behavior as the county transitions into its various phases of reopening. We never thought that a few interesting observations and one quick conversation between us about our bike rides would lead down this rabbit hole of statistical discovery. In the future, we will continue to apply our statistical knowledge to uncover other insights in the world around us.

Works Cited:

1. "U.S. Census Bureau QuickFacts: Santa Clara County, California." Census Bureau QuickFacts, www.census.gov/quickfacts/santaclaracountycalifornia.