Writeup For DataFest

Jonas

13/06/2020

####The Statistically Analysis regarding the number of Occupants in GTA shelters during the COVID-19 lockdown

#*OverView*

Jonas

hello

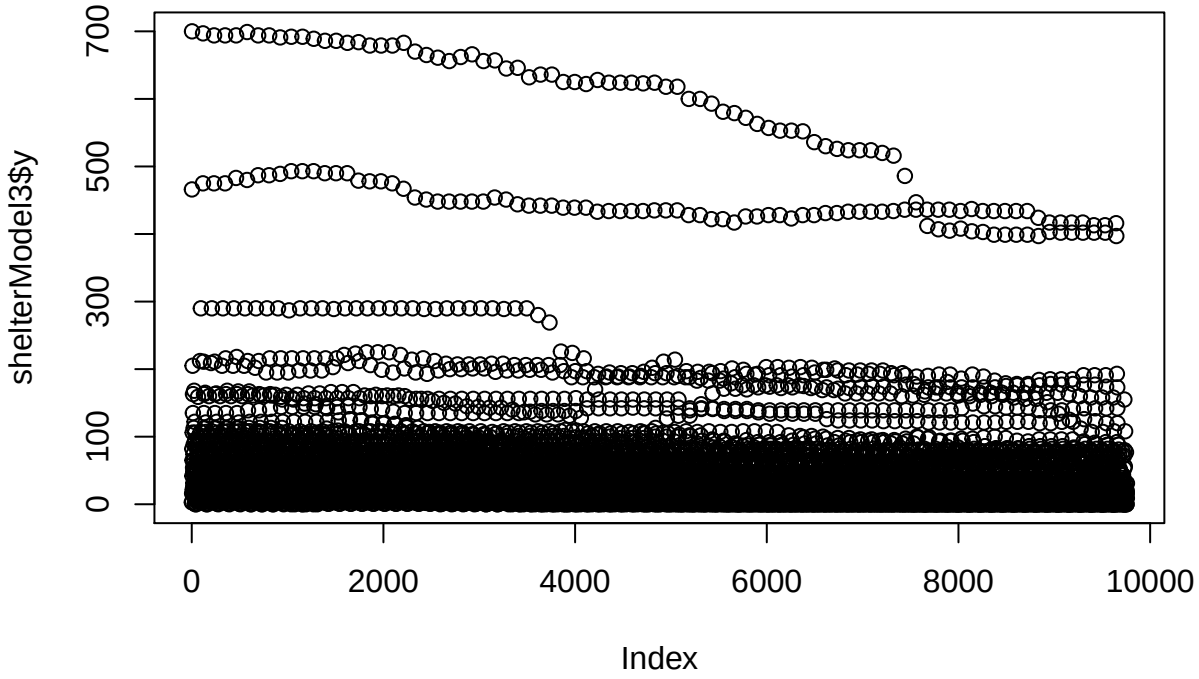
For our project we decided to focus on the issue regarding whether there was a change in the number of people using the GTA shelters between March 17, 2020 and June 8, 2020.

To help us analyze this issue we first gathered the daily data about the shelters listed in the GTA, that are publicaly avaibale here:"https://open.toronto.ca/dataset/daily-shelter-occupancy/". An appropiate possion regression model waas made to fit the data. And from the model we were able to determine that there was either an uniform or decrease in the number of people that stayed in the GTA shelters during the time period.

#*DescriptiveSummaryStatisticsOfOccupancy*

Sector	*Mean*	*StandardDeviation*	*Q*1	*Median*	*Q*3	*Min*	*Max*	*IQR*
Total	47.04	74.72	10.00	28.00	56.00	0.00	700.00	5
Men	39.16	42.15	12.00	28.00	58.00	0.00	290.00	5
Women	27.40	28.65	4.00	21.00	38.00	0.00	166.00	5
Co-ed	39.52	36.44	10.00	33.00	70.00	0.00	178.00	5
Families	119.40	150.99	25.00	72.00	143.0	0.00	700.00	5
Youth	24.44	18.57	11.00	22.00	32.00	0.00	95.00	5

####*SummarizingStatisticalTestOutcomes* :



For the purpose of our project we assumed a signifigance level $alpha$=0.05. From the plot it can be seen that the shelters had a decrease in the number of occupants using their shelter, or there were some slight dips but followed a unfiform distribution.

#TRYING STUFF OUT:

Therefore in conclusion we can see that that there is a singificant drop in the number of occupants in the shelters as COVID-19 progressed. Additionally from the plot(shelterModel3$y) it can be seen that there is a downwards trend. Finally when computing the AIC model it can be seen that the same predictors are used, hence confirming that the predictors are statistically significant.

This is where I show the model I use as well as the results from the model. It is also where I talk about the signifigance of the p-values and why they were used to justify their use for the model (maybe get rid of that part?), comparing t-tests ANOVA summarizing correlation and Regression Analyses

POINT 331 MAY BE A LEVERAGE POINT (OR MAYBE AN INFLUENTIAL POINT)