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.

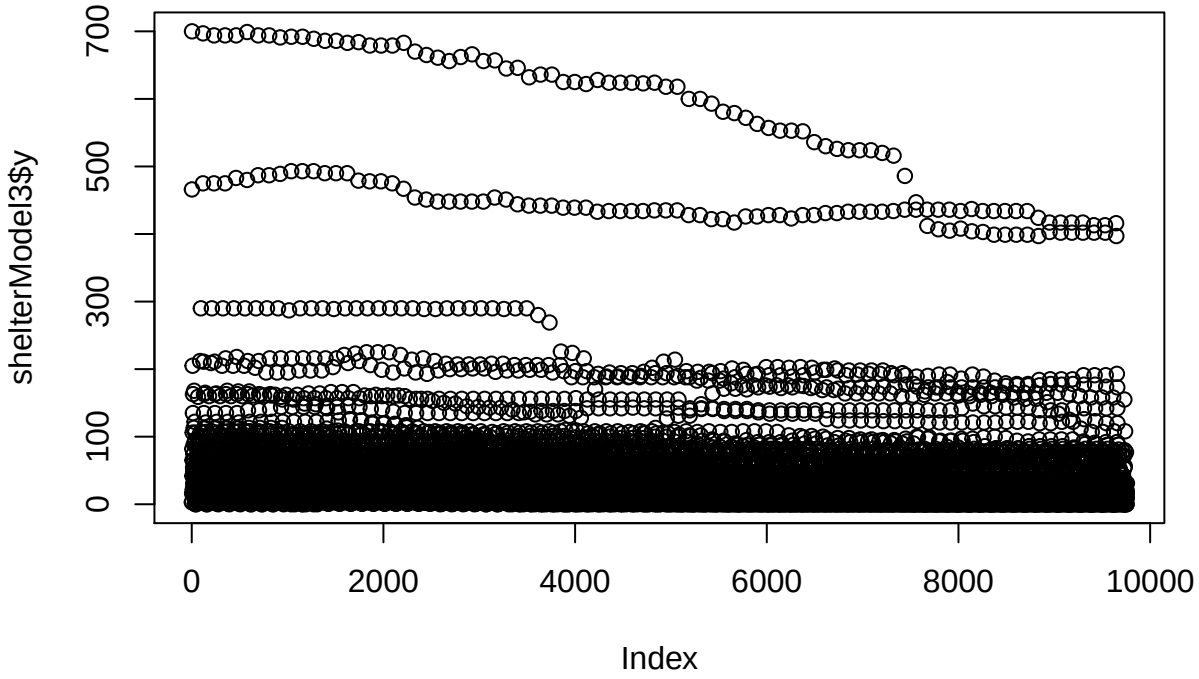
Portion = total number of occupants in a given sector / total number of occupants

#*DescriptiveSummaryStatisticsOfOccupancy*

Sector	*Mean*	σ	$Q1$	*Median*	$Q3$	*Min*	*Max*	*Proportion*
Total	47.04	74.72	10.00	28.00	56.00	0.00	700.00	1.00
Men	39.16	42.15	12.00	28.00	58.00	0.00	290.00	0.061
Women	27.40	28.65	4.00	21.00	38.00	0.00	166.00	0.12
Co-ed	39.52	36.44	10.00	33.00	70.00	0.00	178.00	0.034
Families	119.40	150.99	25.00	72.00	143.0	0.00	700.00	0.061
Youth	24.44	18.57	11.00	22.00	32.00	0.00	95.00	0.18

####*HypothesisTesting* : H_0: From March 17, 2020 till June 8, 2020 there is no significant change in the number of occupants residing in GTA shelters. H_1: From March 17, 2020 till June 8, 2020 there is a significant change in the number of occupants residing in GTA shelters. $\alpha = 0.05$

####*SummarizingStatisticalTestOutcomes* :



For the purpose of our project we assumed a signifigance level $alpha$=0.05. From the gg-plot we noticed that there was a trend where shelters either approximately maintained a constant number of occupants in the shelters as time progressed, or had either a slight or significant decrease in the number of occupants in certain shelters or had certain dips in the number of reported occupants in shelters but was able to rally and proceed to approxiamtely have a constant number of occupants. In particular we were able to identify drastics drops in reported number of occupants in certain shelters, once a significant number number of reported COVID-19 cases were reported in an alternative shelter. Additionally notice that when the likelihood Ratio of Test Of Nested Models occurs, a p-value smaller thena 0.05 is obtained indicating that our model is statistically significant

Something else of other signifigance is that the point 331 is an influential point.