Stage_II_Report

Stage 2 was about understanding the data and using poisson distribution to look at the probability and what could happen. For task 1, I did not calculate the mode because there are not a lot of values re-occurring so there is not point on getting the mode. Just like we did on the member task, I have calculated the needed data for the state and 45 other states. Every data needed like mean_cases, mean_deaths, median_cases, median_deaths are inside a main table for every state. The data can be compared after that. When normalized, Iowa has the highest mean_cases and massachussets has the highest mean_deaths. California is second with the mean_cases but as the data is normalized, Iowa has the highest data here. I have plotted the cases and deaths for the state of Washington after that.

For task 2, We had to graphically represent the data. I have 12 histograms which show the cases and deaths of 6 states. One of Washington and five others that I chose. The histograms are normalized and skewed. The total_cases are different in each case. Some are in decimal and some are in whole numbers which tells us how the numbers of cases differs with respect to the population of that state. There is other data like skewness, variance, kurtosis, standard deviation, median on top of each histograms to give more details about the data. Standard deviation ranges from 0-2 expect for cases of Washington where it is 6.21 which shows the difference in the cases and its timeline. For the correlation, I wanted to find the relation between the population and beds in a state and how it affects the overall cases and deaths. I have posed questions about the effects of number of beds per person in the notebook.

Overall, I think this project showed us a little bit about predicting what might happen with poisson distribution and we were much familiar with implying it.