Temperature and Crime: A Relationship

Kevin Arnold

DSC410-T301 Predictive Analytics

Professor Fadi Alsaleem

November 14, 2023

The topic which I choose for this term long project centers around the theory that the weather and, specifically, the outdoor temperature affects the number of crimes that occur. The idea is that the colder and more brutal the weather is outdoors, the less number of crimes occur. To test this theory, I utilized data from my hometown of Omaha, Nebraska, and sourced it from the Omaha Police Department's website.[1]

The data is represented by the reported incident date and time, the RB Number (which is a unique identifier assigned to each incident), the statute and/or ordinance broken, the physical location, the police district in which the incident occurred, and the latitude and longitude of the physical location.

After performing basic data wrangling procedures and exploratory data analysis, a linear regression model was chosen due to its simplicity to attempt to prove that the temperature and number of incidents have a direct, linear correlation. To implement this model, scikit learn's ordinary least squares Linear Regression[2] was chose to solve for this regression problem.

Once the model was fine-tuned using many of the methods learned within this class, a Dash app front-end was created to allow a user to input a hypothetical temperature (in degrees Fahrenheit) into the app and have it return the predicted number of crimes that would occur. While the initial layout is basic, there are additional opportunities to make it more user friendly and allowing for additional inputs like the precinct in which the user desires to make the estimate. Additional tuning of the model is also possible by adding features in the future that factor in further location and meteorological details such as "is there precipitation? Are any major events occurring the location? And so on…"

[1] https://police.cityofomaha.org/crime-information/incident-data-download
[2] https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html

While both my model and the front-end Dash app are extremely basic, I'm excited to work on and iterate through different versions of this project. Ideally, I will get it to a point where I can actively measure my model's performance against real-time, daily data to see how well it does in predicting criminal events and analyzing what other factors my contribute that I need to consider within the model. I would also like to test different regression methods beyond linear regression and use packages other than sci-kit learn to both scale out my knowledge and improve my model. I still have quite a bit to learn but this course was a great start on an end-to-end project such as this.