1. Thinking about it logically, if you look at bike trips vs temperature it is most likely going to be a parabola. In that case, the best model would most likely be something like a second degree polynomial and an inverse parabola since as it gets warmer there will be more bike trips and as it gets colder there will be fewer.

Chart, line chart

Description automatically generated

* 1. The graph for N = 2 looks to be less accurate than I anticipated at first. However, it does represent the trend I predicted; as temperatures increase, the number of rides will increase up to a certain point before they begin dipping down.

Chart, scatter chart

Description automatically generated

* 1. I graphed N = 3 for the sake of seeing it, but it does not work since an x^3 graph will always have the tail at the left end that makes it inaccurate.

Chart, scatter chart

Description automatically generated

* 1. N = 4 looks to be the more accurate, but as we increase the degree, we run the risk of overfitting. This can also be said of the N = 5 graph, and if we look at the R^2, it seems to be the most “accurate” of the options.

Chart, scatter chart

Description automatically generated Chart, scatter chart

Description automatically generated

For this, I would pick N = 4. While N = 5 looks to be more fitted, the coefficient on it is miniscule. The N = 4 fit looks to have more of a relevant coefficient. In addition, it follows my initial statement that it should be an inverse parabola.

1. MSE

A picture containing text

Description automatically generated

This took me a while to code, and even when I finished the MSE was ridiculously large here, resulting in 179267273139 or 15840853955 depending on the dataset used to calculate it. With an average of 25201 trips, this meant very high variance. If we square root it, we get 423399 or 125860, which makes a lot more sense.

1. Visualization
   1. This graph clearly shows trends between the time of the year and trips made; as we near colder months like November rides will decrease, and as we near warmer months like May rides will increase.
   2. This can be represented in different ways, including the N = 4 graph I had picked in part 1, but the scatter plot below more clearly shows the strong correlation between warmer temperatures and number of rides.

Chart, scatter chart

Description automatically generated

* 1. I assume that two graphs were meant by this; one for months and one for weekdays. For these, I used histograms representing frequency. On one axis, we have the month or whether it was a weekday or not, and on the y-axis we have the number of times a ride was made in that month or day.