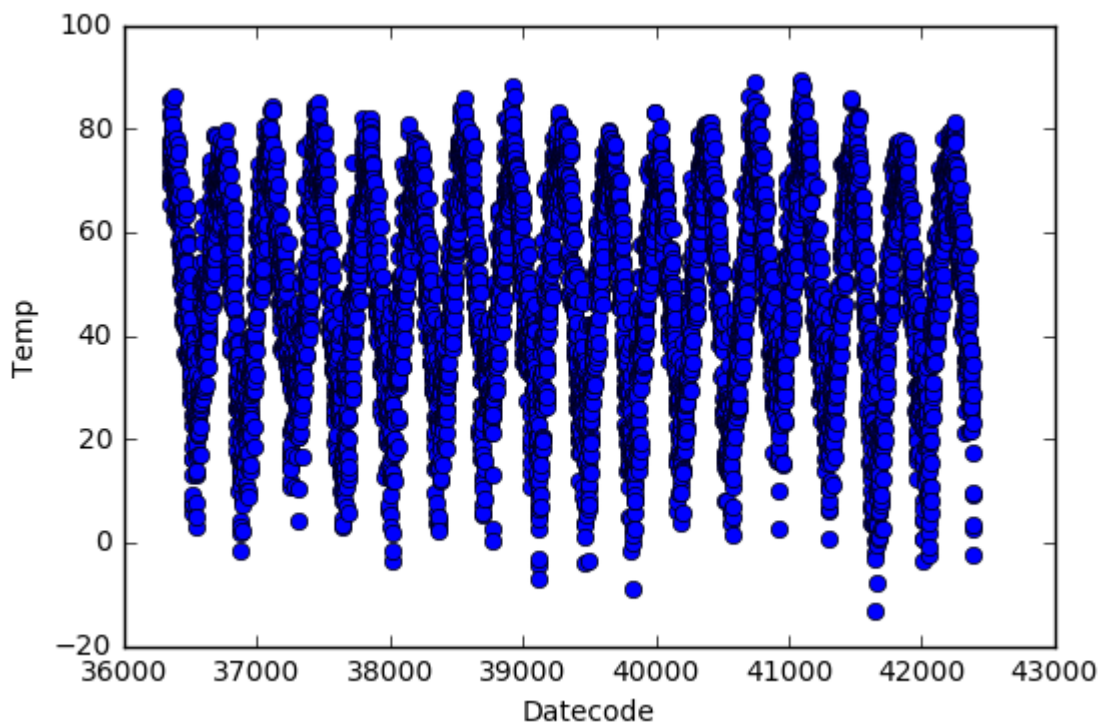Jakob Horner
COSC 4931 – Data Science
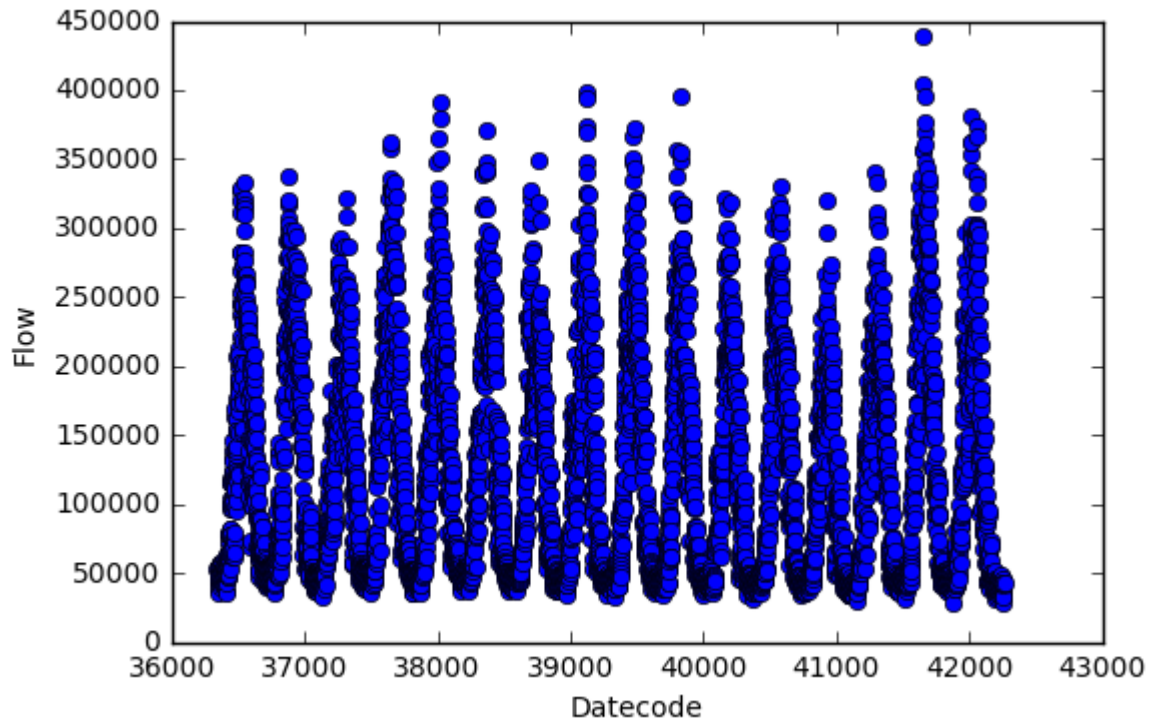March 9, 2017

Data Science Project

In this project, I am using data given by GasDay. This data is a confidential combination of many of GasDay's different data sets. Since this is data is a combination of data sets, some of the data might be more averaged out, however, we can still recognize the patterns and some outstanding data points.

If we analyze the full data set of data points, we can see patterns that roughly repeat once a year.

This shows clear patterns throughout the years. We see the temperature going up during summer times and down during winter times. This shows us that our data is following realistic patterns we can understand.

Similarly, we can see patterns for the flow of gas. We see opposite swings, when temperature is lower, flow is high.

However, this large overall dataset isn't as helpful to use averages and other similar descriptions since it is pattern based rather than mean based.

One method attempted is to separate this data by month.