CSMC 6950 Project Proposal

I will be using the historical weather data for the town of Sexsmith, because quite frankly that just sounds hilarious, and for the year of 2001. I will be looking at the mean and standard deviation of the temperature of each month to determine if the std is a function of the mean. I will consider converting the temperature to kelvin to make presenting and discussing the results easier. An interesting point of order is whether the temperature distribution is skewed or not, it is probably not going to be perfectly symmetric about a given axis, but it would be interesting to see to what extent the data is skewed. For this I will have to experiment with where I take the axis of symmetry which should be a relatively straight forward task once I get a plot of the temperature distribution (see below). I would also like to look at the temperature distribution per month. Perhaps violin plots for each month would be useful.

Another thing to look at is the following: first the probability of there being precipitation on any day given a particular month of the year $P(P|M)$, next the probability of there being precipitation on any day given a month and temperature $P(P|(M,T)$ (where P is precipitation, M is month and T is temperature). I could do some linear regression using likelihood estimation or perhaps even Bayesian linear regression would be used here but I can't make any promises on the latter as I have only read so far into it a few months ago but it would be pretty cool to pull off. At any rate some basic analysis on these two metrics can be expected in the project.

Quick note: when I say temperature, I am talking about mean temperature not min or max.