Stat 565

Homework 1

Due in class Jan 14th

Q1

Google maintains records of the popularity of search terms over time at http://www.google.com /trends You can examine daily data by selecting a time period of less than 90 days, otherwise you'll see weekly or monthly data (e.g. searches for "hangover" in the last 90 days). The numbers reported are relative search volume. Your task is to find:

- one search term dominated by a weekly or annual seasonal pattern
- one search term dominated by an increasing trend
- one search term that would be poorly described by a combination of trend and seasonality

In each case, you should download the relevant data as a .csv file (click on the … on the trends page), read the data into R and produce a plot of the data.

Q2

I've downloaded much higher resolution data from the Corvallis Municipal Airport weather station. I am giving you ten years of about thrice hourly (I think, you should probably check) observations. Download the data from: http://stat565.cwick.co.nz/data/corv_sub.rds. You can get it into R with:

```
corv <- readRDS("corv_sub.rds")
head(corv)
```

Check out the structure, I've been nice and already converted the `date` and `datetime` to R date time classes.

Before I move to a new place I like to look up what the climate is like there. But I normally end up looking at a graph that shows me the average min, mean and max temperature by month, and maybe total precipitation, and that doesn't really help me decide how my day to day life will be affected by the weather. I'm much more interested in things like:

- How often will I be biking to work in below freezing temperatures?
- How many days a year will I see the sun?
- How many days do I need a raincoat, assuming I'm only outside during my morning and evening commute?

Your task is to create a climate metric that **you** would find useful, and by using the data provided

and `dplyr`, calculate it and present it visually. (And since this is a statistics class you might give a thought to how you would quantify and present the variability in your metric).

Challenge Problem

I will occasionally add a problem to a homework that is more challenging. These problems are not worth any credit, but provide opportunity for you to explore more on your own.

These problems are not necessarily harder, but they will tend not to have a single solution, and there is no guarantee they are completely solvable (i.e. I may not have even attempted them).

Douglas J. Keenan has posed a challenge to detect trends in time series. He has generated 1000 time series some of which have a trend and will award $100,000 to the first person to correctly classify 900 of the series.

The challenge can be found at: http://www.informath.org/Contest1000.htm

Challenge task: Perform some EDA on the 1000 series. Do you believe his claim that he added trends of +1 or -1 degree C / century? About what proportion do you think he added trends to?