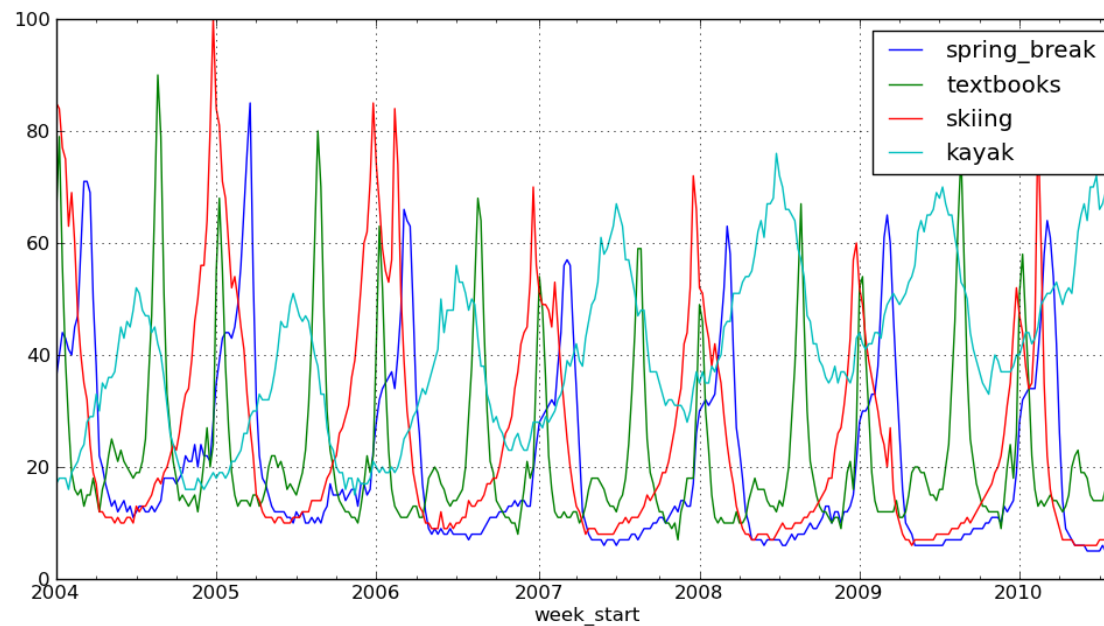# Homework 1

Google Trends is pretty awesome, except that on the site you cannot do more than overlay plots. Here we'll play with search term data downloaded from Google and draw our own conclusions. We will be using pandas and numpy to explore the data.

1. Use trends.csv file and pandas.read_csv() to import the data and reproduce this plot:

2. Determine in which week of each year (for all five search trends including "global warming") that search reached its peak and its minimum. Are there any trends you can spot with any of the terms?

3. Which term has the largest scatter about its median value? Which term has the smallest scatter?

4. Determine the time lag, in units of weeks, that maximizes the cross-correlation between "skiing" and "spring break". Do this also for "skiing" and "global warming"

5. Download trend data on two terms of your choosing and redo the questions above
http://www.google.com/insights/search/#

Hints:

1. You might start your script with:

```
import pandas as pd
import numpy as np
df=pd.read_csv('trends.csv',index_col=0,parse_dates=True)
```

2. numpy has tools for cross-correlation

```
result = np.correlate(df.spring_break,df.spring_break, mode='full')
plot(arange(result.size) - result.size/2,result)
```