Homework-1

Tadesse Zemicheal

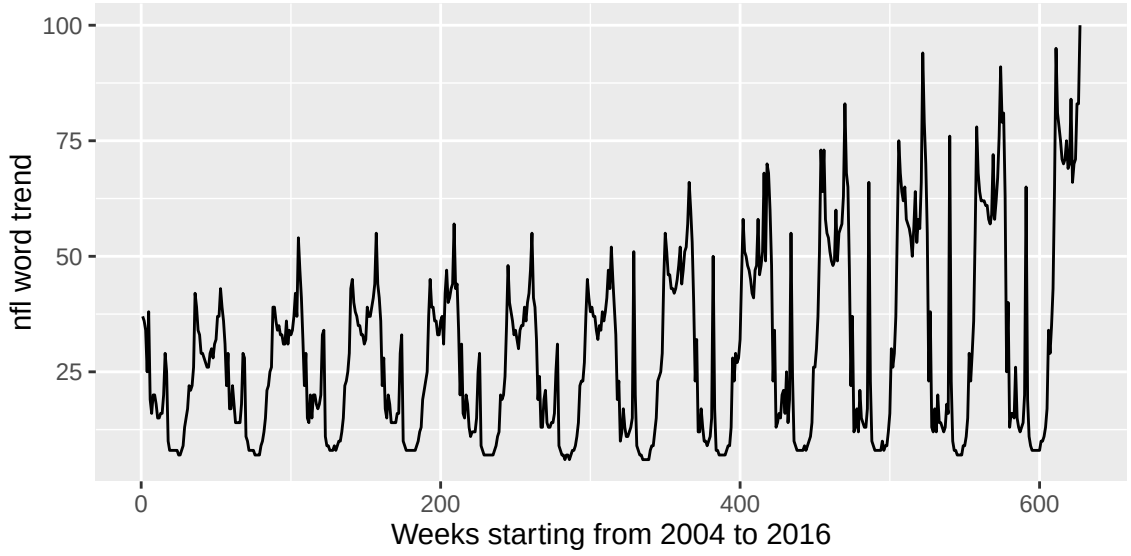
January 13, 2016

Trends in Google search

Search term dominated by a weekly or annual seasonal pattern

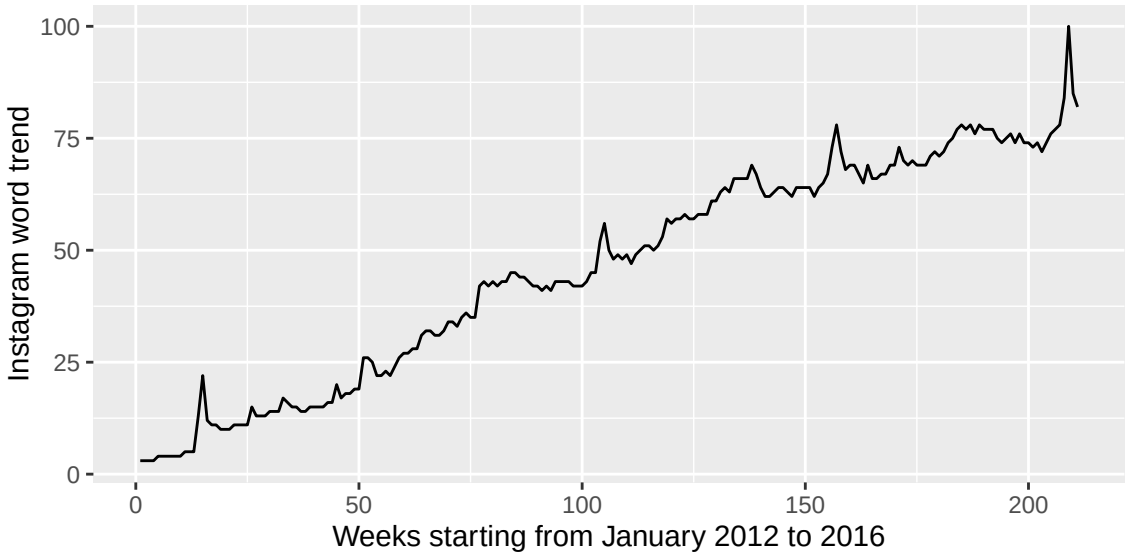
Google maintains records of the popularity of search terms over time at http://www.google.com/trends You can examine daily data by selecting a time period of less than 90 days, otherwise you'll see weekly or monthly data (e.g. searches for "hangover" in the last 90 days). The numbers reported are relative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

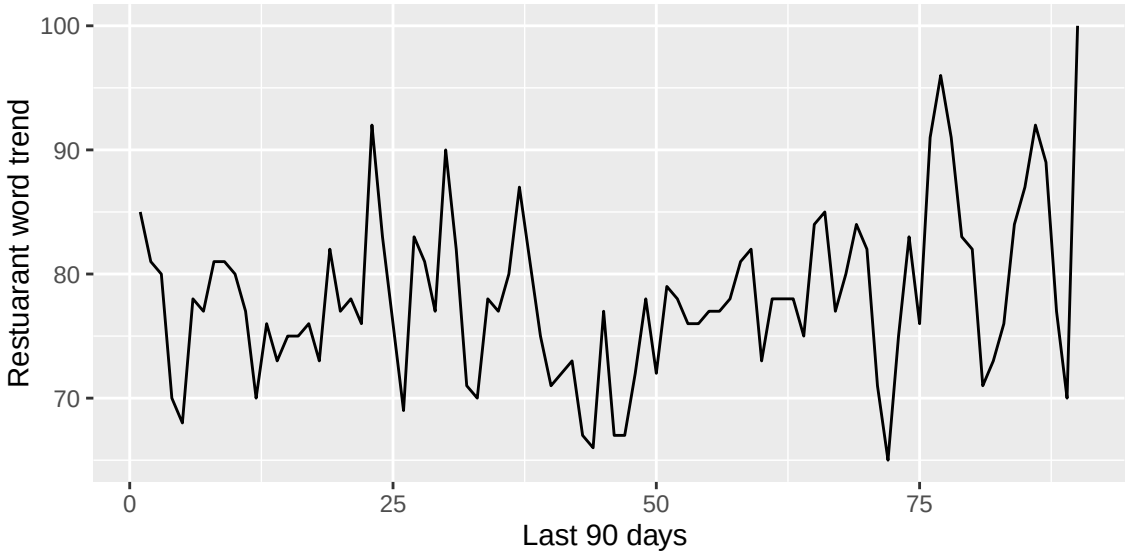


Search term dominated by an increasing trend

Search term "instagram" from period of January 2012 to January 2016



Search term that would be poorly described by a combination of trend and seasonality

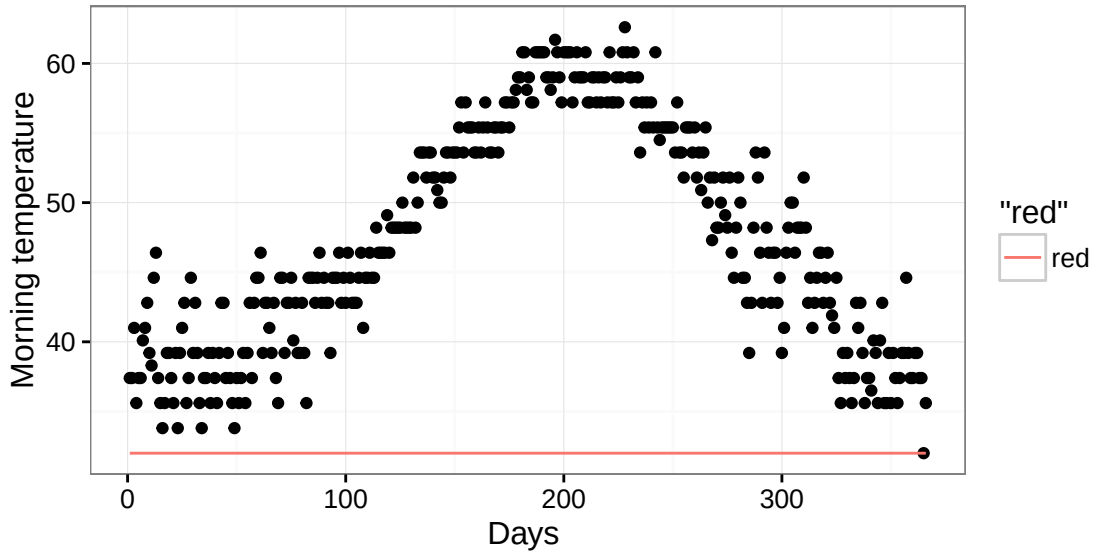


Q2

Corvallis Weather data #

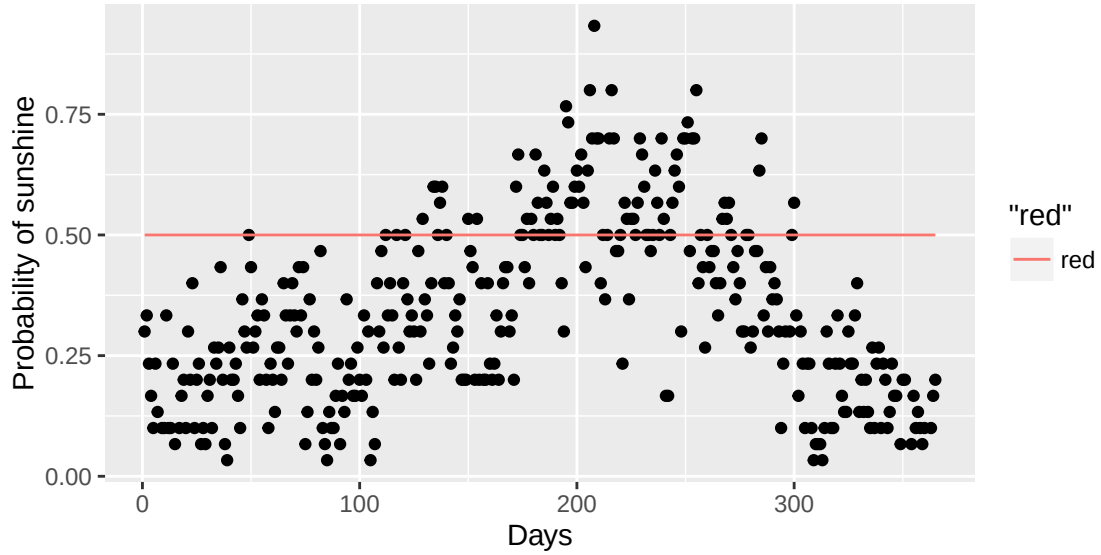
How often I will bike to work below freezing temperature

In this plot, I averaged morning temperature from 6AM to 9 AM over the 10 years data. Then to find the number of possible freezing days, I filtered the average temperature below freezing (32 F) temperature. The graph shows, on average freezing morning temperature rarely happens in Corvallis, with few possibility on late Decem-



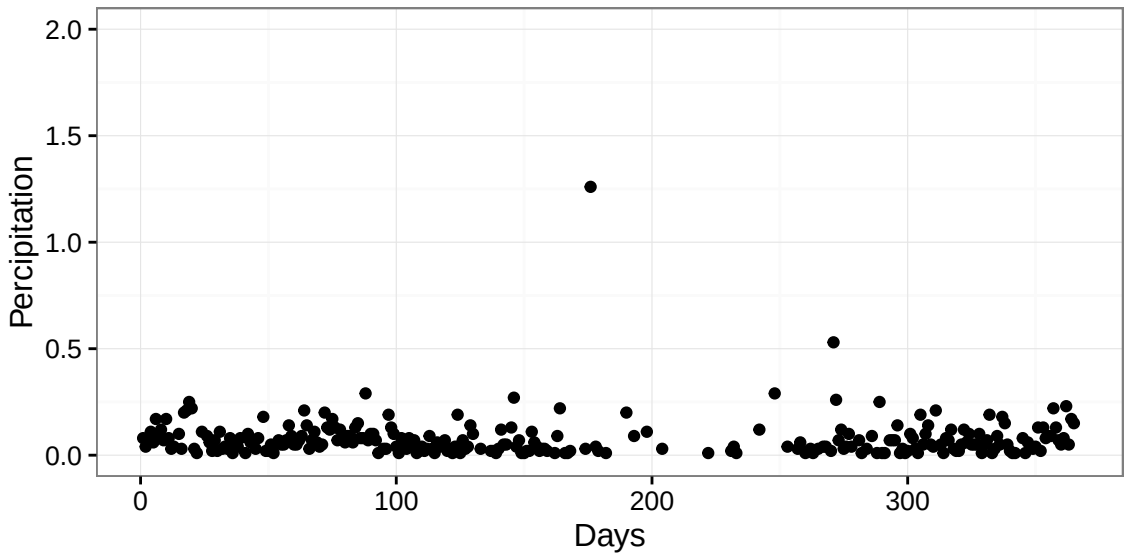
ber.

How many days will I see sunshine? ### In this exploratory analysis, I took a sunny sunshine say at noon. Then I counted all appearance of **clear** sky at noon overall all 10 year data. Then my hypothesis is if in more than half of the years a **clear** sky seen, then I concluded there is high probability the day is sunshine.



###

How many days will I need a raincoat during commiting. ### Suppose the commuting time is from 7 AM to 9 AM in morning and 5 PM to 7 PM in the evening. Then we can take filter the daily average precipitation for the specified time period to explore the possibility of rainy days. I assume in order to bring raincoat, it should rain above some tolerance precipitation (say 0.05) . Then this can give how many days I can bring my raincoat.



##

Appendix

```
knitr::opts_chunk$set(echo=FALSE, message = FALSE,
  warning = FALSE, results = "hide", fig.height = 3, fig.width = 6)
library('ggplot2')
library('lubridate')
library('dplyr',quietly =T)
nfl<-read.csv('NFL-trend-2004-2016.csv',header=T,comment.char = '#')
nfl$week.no <- 1:nrow(nfl)
qplot(data=nfl,week.no,nfl,geom="line") +xlab("Weeks starting from 2004 to 2016")+
    ylab("nfl word trend")
instagram<-read.csv('istagram-increase-trend.csv',header=T)
instagram$week.no <-1:nrow(instagram)
qplot(data=instagram,week.no,Instagram,geom="line") +xlab("Weeks starting from January 2012 to 2016")+
    ylab("Instagram word trend")
restuarant<-read.csv('restuarant-past-90days.csv',header=T)
restuarant$restuarant.days <-1:nrow(restuarant)
qplot(data=restuarant,restuarant.days,restuarant,geom="line") +xlab("Last 90 days ")+
    ylab("Restuarant word trend")
#load data
cor <-readRDS('corv_sub.rds')
#str(cor) #check the structure
corv <- mutate(cor,year = year(date),
                            month = month(date),
                            yday = yday(date))
#Average daily morning temperature from 6 AM to 10 AM
avg.daily.temp <- corv %>%
    filter(hour %in% c(6,7,8)) %>%
    group_by(yday) %>%
    summarize(min.temp=median(temp,na.rm = T))
#plot the mean morning temperature over all data
#Below red line indicates freezing temperature
qplot(data=avg.daily.temp,yday,min.temp) +
    geom_line(aes(y=32,color='red')) +
    xlab("Days ") +ylab("Morning temperature")+theme_bw()
```

```r
cat("Number of days with freezing temperature",nrow(avg.daily.temp%>%filter(min.temp<=32)))
#Take a typical time say at noon
#What is the average probability that we will see clear sunshine at noon over all years
sunshine.at.noon <- corv %>%
    filter(hour==12 & conditions=='Clear') %>%
    group_by(yday) %>%
    summarize(prob.sunshine=length(yday)/30)  #30 because we have 3 reading per year and over 10 years
qplot(data=sunshine.at.noon,yday,prob.sunshine,geom="point") + geom_line(aes(y=0.5,color='red'))+
    xlab("Days") +
    ylab("Probability of sunshine")
sunshine.days <-nrow(sunshine.at.noon %>% filter(prob.sunshine>0.5))
cat("Based on average noon weather probability, the number of sunshine days are ",sunshine.days)
morningEveningRain <- corv %>% filter(hour %in% c(7,8,17,18)) %>% group_by(yday) %>%
    summarize(max.rain=max(precip,na.rm = T))
qplot(data=morningEveningRain,yday,max.rain,geom='point') + ylim(0,2)+
    ylab('Percipitation') +xlab('Days')+theme_bw()
#Number of rain that need coat.
#Assume we can tolerate rain of 0.07
tolerance <- 0.05 #tolerance of percipitation
number.of.days<-na.omit(morningEveningRain) %>% filter(max.rain>tolerance) %>%nrow
cat(number.of.days,"days will rain\n")
```