

# Data 650 Midterm

*Charles Redmond*

*October 17, 2017*

## Set-Up

Download the Kaggle data set found here (<https://www.kaggle.com/rtatman/us-candy-production-by-month/data>). The data is daily United States candy production as a percentage of 2012 production. Read this data into R.

You will need dplyr, ggplot2, stringr, and lubridate.

## Problem 1

Look at the first plot in the dashboard found here (<https://bendazz.github.io/grad-sol/>). It is a time series plot indicating the candy production for every day in the data set. Reproduce this plot.

## Problem 2

Look at the second plot in the dashboard found here (<https://bendazz.github.io/grad-sol/>). It is a time series plot for only the days in 1990. Reproduce this plot, being sure to have each month labeled as shown.

Hint: "%b"

## Problem 3

Look at the third plot in the dashboard found here (<https://bendazz.github.io/grad-sol/>). It is a bar plot comparing the average candy production for the 1980's, 1990's, and 2000's (2000-2009). Reproduce this plot as it is shown.

Hint:

```
str_sub("1980-01-01",3,3)
```

```
## [1] "8"
```

Hint:

```
paste("8","0's",sep='')
```

```
## [1] "80's"
```

Hint:

```
standing<-c("junior","freshman","sophomore")
standing<-factor(standing,levels=c("freshman","sophomore","junior"))
levels(standing)
```

```
## [1] "freshman" "sophomore" "junior"
```

## Submission

Make a dashboard out of the three plots.

Use this link (<https://www.dropbox.com/request/GXoc7m8z3FIJsrz7dGeV>) to submit your dashboard Rmd file.