Developing Data Products - Week 3 Assignment

```
knitr::opts_chunk$set(echo = TRUE, cache = TRUE)
```

Introduction

- This project was created as part of the Developing Data Products course of the Coursera Data Science Specialisation.
- The goal of the project is to create a web page presentation using R Markdown that features a plot created with Plotly, and to host the resulting web page on either GitHub Pages, RPubs, or NeoCities.
- The interactive plot on the next slide represents the number of road accidents in Great Britain from 2005 to 2015, grouped by severity (slight, serious, or fatal).

```
rm(list=ls())
library(plotly)
library(data.table)
library(tidyr)
library(lubridate)
library(zoo)
```

```
accidents0514 <- fread("data/Accidents0514.csv", header = TRUE, sep = ",")
```

```
##
Read 0.0% of 1640597 rows
Read 9.1% of 1640597 rows
Read 17.7% of 1640597 rows
Read 26.2% of 1640597 rows
Read 28.6% of 1640597 rows
Read 37.8% of 1640597 rows
Read 43.3% of 1640597 rows
```

```
Read 53.0% of 1640597 rows
Read 62.8% of 1640597 rows
Read 65.8% of 1640597 rows
Read 78.0% of 1640597 rows
Read 87.2% of 1640597 rows
Read 97.5% of 1640597 rows
Read 1640597 rows and 32 (of 32) columns from 0.210 GB file in 00:00:17
```

```
accidents0514 <- accidents0514 %>%
    select(Accident_Severity, Date)

accidents15 <- fread("data/Accidents_2015.csv", header = TRUE, sep = ",")
accidents15 <- accidents15 %>%
    select(Accident_Severity, Date)

# concatenate data tables and free up environment
accidents <- rbind(accidents0514, accidents15)
rm(list = c("accidents0514", "accidents15"))
```

```
# convert severity to factor and add labels
accidents$Accident_Severity <- factor(accidents$Accident_Severity, levels = 1
:3, labels = c("Fatal", "Serious", "Slight"))


# convert date strings to Date objects
accidents$Date <- dmy(accidents$Date)


# group data by date and severity, get count, one row per date
accident_count <- accidents %>%
    group_by(Date, Accident_Severity) %>%
    summarise(count = n()) %>%
    spread(key = Accident_Severity, value = count) %>%
    as.data.frame()
```

```
# create a smoother for each severity to visualise general trends
loess_slight <- loess(Slight ~ as.numeric(Date), data = accident_count)
loess_serious <- loess(Serious ~ as.numeric(Date), data = accident_count)
loess_fatal <- loess(Fatal ~ as.numeric(Date), data = accident_count)
```

Road accidents in Great Britain (2005-2015)

```
# plot data
plot_ly(accident_count) %>%
    add_trace(x = ~Date, y = ~Slight, type="scatter", mode = "markers",
              name = "slight", legendgroup = "slight",
              marker = list(color = "#52A9BD")) %>%
    add_trace(x = ~Date, y = ~Serious, type="scatter", mode = "markers",
              name = "serious", legendgroup = "serious",
              marker = list(color = "#FFF16B")) %>%
    add_trace(x = ~Date, y = ~Fatal, type="scatter", mode = "markers",
              name = "fatal", legendgroup = "fatal",
              marker = list(color = "#F5677D")) %>%
    add_trace(x = as.Date(loess_slight$x), y = fitted(loess_slight),
              type="scatter", mode = "lines",
              line = list(color = '#1A7A90'),
              name = "slight Loess smoother", legendgroup = "slight",
              hoverinfo = 'none', showlegend = FALSE) %>%
    add_trace(x = as.Date(loess_serious$x), y = fitted(loess_serious),
              type="scatter", mode = "lines",
              line = list(color = '#E9D625'),
              name = "serious Loess smoother", legendgroup = "serious",
              hoverinfo = 'none', showlegend = FALSE) %>%
    add_lines(x = as.Date(loess_fatal$x), y = fitted(loess_fatal),
              type="scatter", mode = "lines",
              line = list(color = '#DC2340'),
              name = "fatal Loess smoother", legendgroup = "fatal",
              hoverinfo = 'none', showlegend = FALSE) %>%
```

```
    layout(xaxis = list(title = "Year"),
           yaxis = list(title = "Number of Accidents")
    )
```

```
## Warning: Ignoring 39 observations
```