Week3_Assignment

Anish Raj

10/11/2019

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 2017 to 2018, grouped by severity (slight, serious, or fatal).
  - A Loess smoother line has been added to highlight the overall evolution of the number of accidents.

```
rm(list=ls())
library(plotly)
```

```
## Loading required package: ggplot2
```

```
## Warning: package 'ggplot2' was built under R version 3.5.3
```

```
##
## Attaching package: 'plotly'
```

```
## The following object is masked from 'package:ggplot2':
##
##     last_plot
```

```
## The following object is masked from 'package:stats':
##
##     filter
```

```
## The following object is masked from 'package:graphics':
##
##     layout
```

```
library(data.table)
```

```
## Warning: package 'data.table' was built under R version 3.5.3
```

```
library(tidyr)
```

```
## Warning: package 'tidyr' was built under R version 3.5.3
```

```r
library(lubridate)
```

```
## Warning: package 'lubridate' was built under R version 3.5.3
```

```
##
## Attaching package: 'lubridate'
```

```
## The following objects are masked from 'package:data.table':
##
##     hour, isoweek, mday, minute, month, quarter, second, wday,
##     week, yday, year
```

```
## The following object is masked from 'package:base':
##
##     date
```

```r
library(zoo)
```

```
## Warning: package 'zoo' was built under R version 3.5.2
```

```
##
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```r
# The source data sets are not included in this repository.
# To reproduce this presentation, you will first need to download the two
# following zipped data sets:
# - All STATS19 data (accident, casualties and vehicle tables) for 2005 to
#   2014", from
#   https://data.gov.uk/dataset/road-accidents-safety-data/resource/8ecee6ac-33fd-4f5b-8973-e900cc65d24
# - Road Safety - Accidents 2015, from
#   https://data.gov.uk/dataset/road-accidents-safety-data/resource/ceb00cff-443d-4d43-b17a-ee13437e956
# Then extract the `Accidents0514.csv` and `Accidents_2015.csv` files from
# the zip files in a subdirectory named `data`.
# read data for 2005-2014 and 2015 as data tables and keep only severity and
# date columns

accidents2017 <- fread("C:/Users/samsung/Desktop/Coursera/Module 9/Week3/dftRoadSafetyData_Accidents_20

accidents2017 <- accidents2017 %>%
  select(Accident_Severity, Date)

accidents2018 <- fread("C:/Users/samsung/Desktop/Coursera/Module 9/Week3/dftRoadSafetyData_Accidents_20

accidents2018 <- accidents2018 %>%
  select(Accident_Severity, Date)
```

concatenate data tables and free up environment

```
accidents <- rbind(accidents2017, accidents2018)
rm(list = c("accidents2017", "accidents2018"))
```

```
# 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()) %>%

    ##get count one row per date
    spread(key = Accident_Severity, value = count) %>%
    as.data.frame()
```

```
## Warning: Calling `n()` without importing or prefixing it is deprecated, use `dplyr::n()`.
## This warning is displayed once per session.
```

```
# 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 (2017-2018)

```
# 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 = "date"),
    yaxis = list(title = "number of accidents")
  )
```

```
## This version of Shiny is designed to work with 'htmlwidgets' >= 1.5.
##     Please upgrade via install.packages('htmlwidgets').
```

```
## Warning: Ignoring 14 observations
```