Basic Loading and Transformation

Jawaid Hakim

2022-09-05

```
library(dplyr)
library(plotly)
```

Loading data in Data Frame

For this assignments a number of data sets are available at https://data.fivethirtyeight.com/.

Based on my interest in politics, I selected https://fivethirtyeight.com/features/the-big-lie-voting-laws/, which contains data on legatislative measures for major electoral change.

Data for this data set is publickly available as a google spreadsheet. I downloaded the google spreadsheet as a csv file. Both the R file and data set are available on github:

```
ds <- read.table(file="The Big Lie's Long Shadow - Sheet1.csv", header=TRUE, sep=",")
```

A glimpse shows the data set has 579 rows and 5 columns:

```
summary(ds)
```

```
##     State               Bill           Introducing.Party    Category
##  Length:579          Length:579         Length:579          Length:579
##  Class :character    Class :character   Class :character    Class :character
##  Mode  :character    Mode  :character   Mode  :character     Mode  :character
##     Status
##  Length:579
##  Class :character
##  Mode  :character
```

Status of the proposed legislative measure is an important value. It tells us whether the proposed legislative measure passed or not, became laws or not. Let's convert this column data type (Mode) from character to factor:

```
ds <- ds %>% mutate(Status = as.factor(Status))
summary(ds)
```

```
##     State               Bill           Introducing.Party    Category
##  Length:579          Length:579         Length:579          Length:579
##  Class :character    Class :character   Class :character    Class :character
##  Mode  :character    Mode  :character   Mode  :character     Mode  :character
##
```

```
## 
## 
## 
##                                                               Status
##  Legislature adjourned without passing              :283
##  Referred to committee                              : 96
##  Signed into law                                    : 50
##  Passed committee, legislature adjourned without passing: 41
##  Passed Senate, House adjourned without passing     : 15
##  Passed House, Senate adjourned without passing     : 13
##  (Other)                                            : 81
```

Let's take a look at the complete summary of Status. There are **30** unique nominal values (a lot). Measures traverse a long journey from being introduced to passage: they may (may not) pass the House or the Senate, may (may not) be signed into law:

```
summary(ds$Status)
```

```
##                                              Became AB 6970
##                                                           1
##                                              Became AB 7478
##                                                           1
##          Became HB 1498, HB 1499, HB 1501, and HB 1502
##                                                           1
##                                              Became HB 2905
##                                                           1
##                                              Became SB 1819
##                                                           1
##                                              Became SB 2545
##                                                           1
##                                               Became SB 264
##                                                           1
##                                               Died in House
##                                                           1
##                                           Failed to cross over
##                                                           3
##                       Legislature adjourned without passing
##                                                         283
##                                             Never introduced
##                                                           1
##                                             Passed committee
##                                                           7
## Passed committee, legislature adjourned without passing
##                                                          41
##                                                 Passed House
##                                                           5
##                               Passed House, died in Senate
##                                                           5
##          Passed House, Senate adjourned without passing
##                                                          13
##                                               Passed Senate
##                                                          11
##                             Passed Senate & House, then died
```

2

```
##                                                               3
##                          Passed Senate, died in House
##                                                               6
##           Passed Senate, House adjourned without passing
##                                                              15
##           Passed Senate; House adjourned without passing
##                                                               1
##                                   Referred to committee
##                                                              96
##                                   Rejected by committee
##                                                               1
##                                      Rejected by House
##                                                               3
##                                     Rejected by Senate
##                                                               3
##                                          Signed into law
##                                                              50
##                                  Veto-overriden into law
##                                                               3
##                                                  Vetoed
##                                                              13
##                     Voting restriction removed from bill
##                                                               4
##                                               Withdrawn
##                                                               4
```

A quick plot of Status values shows that for a majority of measures either the 'Legislature adjourned without passing' or the measure was 'Referred to committee':

```
plot_ly(ds, x=ds$Status)
```

The nominal value **Signed into law** identifies the legislative initiatives that actually survived the long legislative process and became laws. To make it easy to identify measure that became law, we do the following:

1. Add a new logical column **Signed Into Law**
2. Each entry in this column will be TRUE (FALSE) if the corresponding Status is (is not) 'Signed into law'

```
ds <- ds %>% mutate('Signed Into Law' = Status == 'Signed into law')
```

Final data set:

```
summary(ds)
```

```
##      State               Bill            Introducing.Party    Category
##  Length:579         Length:579         Length:579         Length:579
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
```

```
##
##                                                      Status      Signed Into Law
##  Legislature adjourned without passing                 :283    Mode :logical
##  Referred to committee                                 : 96    FALSE:529
##  Signed into law                                       : 50    TRUE :50
##  Passed committee, legislature adjourned without passing: 41
##  Passed Senate, House adjourned without passing        : 15
##  Passed House, Senate adjourned without passing        : 13
##  (Other)                                               : 81
```