Airline_fatalitiesComparison

Caroline Oliver

4/3/2018

Load in the dataset from csv file

NOTE: Make sure to replace the path variable with your own path to the file

```
library(readr)
airline_safety <- read_csv("/Users/carolineoliver/airline-safety.csv")
```

```
## Parsed with column specification:
## cols(
##   airline = col_character(),
##   avail_seat_km_per_week = col_double(),
##   incidents_85_99 = col_integer(),
##   fatal_accidents_85_99 = col_integer(),
##   fatalities_85_99 = col_integer(),
##   incidents_00_14 = col_integer(),
##   fatal_accidents_00_14 = col_integer(),
##   fatalities_00_14 = col_integer()
## )
```

```
# REPLACE LINE ABOVE WITH YOUR PATH: airline_safety <- read_csv("path_to_csv_file_here")
```

Get number of ASK based on trillions (like in the study)

```
airline_safety$ASKperTrillion = airline_safety[, "avail_seat_km_per_week"]/1000000000000
```

Get number of ASK based on billions to compare charts

```
airline_safety$ASKperBillion = airline_safety[, "avail_seat_km_per_week"]/1000000000
```

add columns to dataset for 85 to 99 - INCIDENTS (per trillion and per billion)

```
airline_safety$Incidents_85_99_trillion =
  airline_safety$incidents_85_99 / airline_safety$ASKperTrillion
airline_safety$Incidents_85_99_billion =
  airline_safety$incidents_85_99 / airline_safety$ASKperBillion
```

add columns to dataset for 00 to 14 - INCIDENTS (per trillion and per billion)

```
airline_safety$Incidents_00_14_trillion =
  airline_safety$incidents_00_14 / airline_safety$ASKperTrillion
```

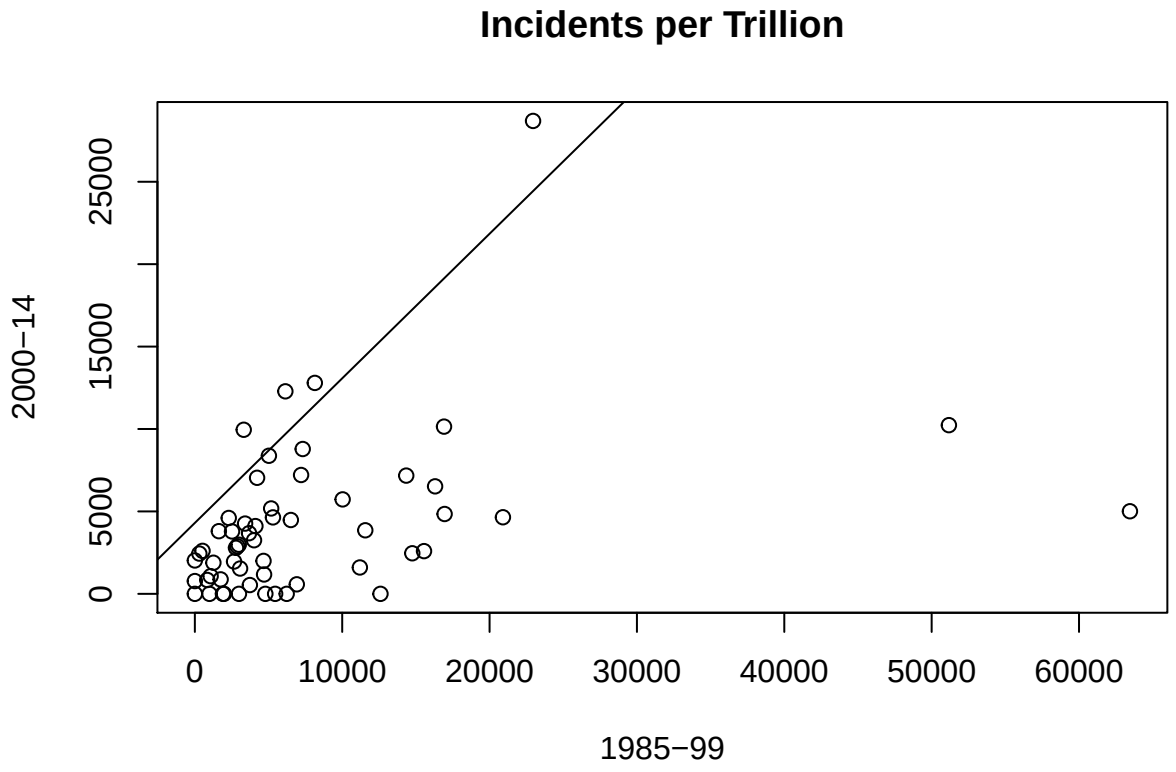
```
airline_safety$Incidents_00_14_billion =
  airline_safety$incidents_00_14 / airline_safety$ASKperBillion
```

Make the plot for 'per trillion' - INCIDENTS

the data for the columns needed is originally stored as a list

must be converted to numeric vector to be used in plot function

```
tril_85_99 = as.numeric(as.character(unlist(airline_safety$Incidents_85_99_trillion)))
tril_00_14 = as.numeric(as.character(unlist(airline_safety$Incidents_00_14_trillion)))
plot(tril_85_99, tril_00_14, xlab="1985-99", ylab="2000-14", main = "Incidents per Trillion")
abline(lm(tril_85_99 ~ tril_00_14))
```



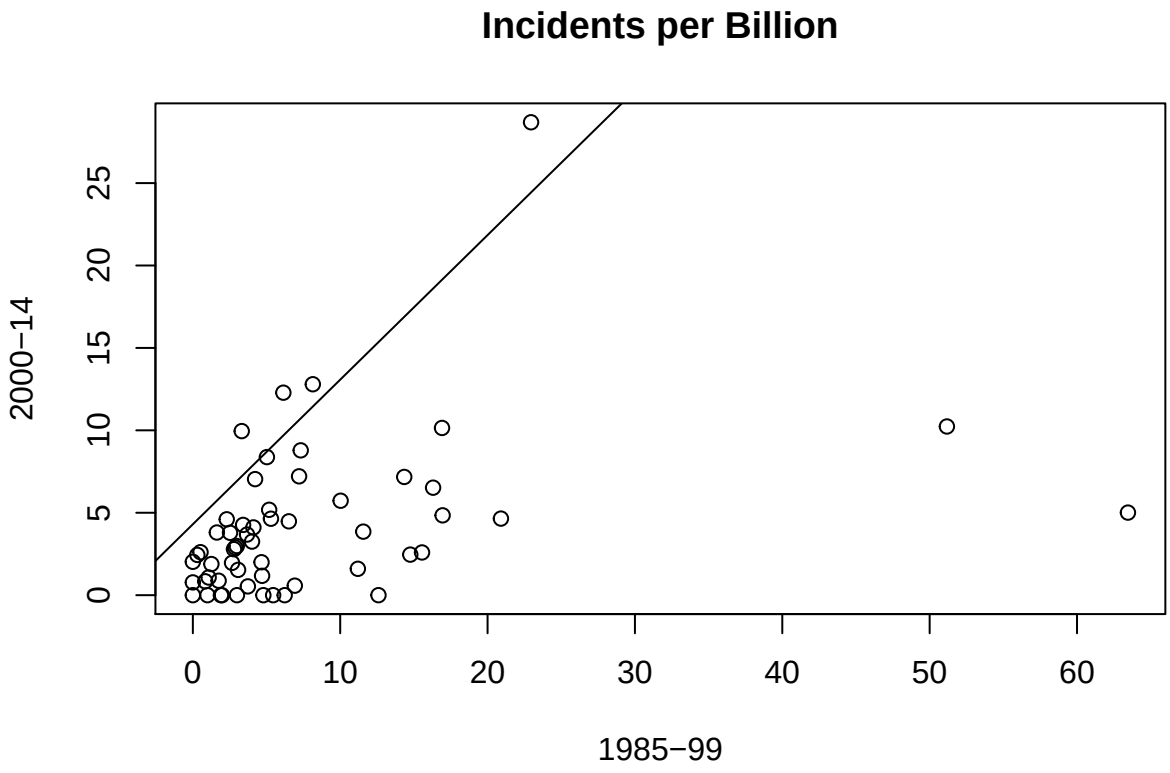
Make the plot for 'per billion' - INCIDENTS

the data for the columns needed is originally stored as a list

must be converted to numeric vector to be used in plot function

```
bil_85_99 = as.numeric(as.character(unlist(airline_safety$Incidents_85_99_billion)))
bil_00_14 = as.numeric(as.character(unlist(airline_safety$Incidents_00_14_billion)))
```

```r
plot(bil_85_99, bil_00_14, xlab="1985-99", ylab="2000-14", main = "Incidents per Billion")
abline(lm(bil_85_99 ~ bil_00_14))
```



NOTES SO FAR ON GRAPHS ABOVE VS STUDY

- per Trillions does not yeild same axis/numbers as the study does

- per Billions is closer to the numbers yeilded on the study but still not exact

Like the FatalitiesComparison graphs, the points on the plots themselves have the same pattern as the ones produced in the study.