

London Crimes

Suz

2019 10 14

Answer the following questions. –Each answer should include a plot or graph. – Each answer should provide some information about error (confidence).

```
library('tidyverse')
```

```
## -- Attaching packages
```

```
----- tidyverse 1.2.1 --
```

```
## √ ggplot2 3.2.1      √ purrr  0.3.2
## √ tibble  2.1.3      √ dplyr  0.8.3
## √ tidyr   0.8.3      √ stringr 1.4.0
## √ readr   1.3.1      √ forcats 0.4.0
```

```
## -- Conflicts
```

```
----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
```

```
# List all files ending with csv in directory
csv_files = list.files(path = 'dt', pattern = "csv$", full.names = TRUE)
# Read each csv file into a list
csv_list = lapply(csv_files, read.csv, header=TRUE)
# Concatenate the data in each csv file into one combined data frame
london_street = do.call(rbind, csv_list)

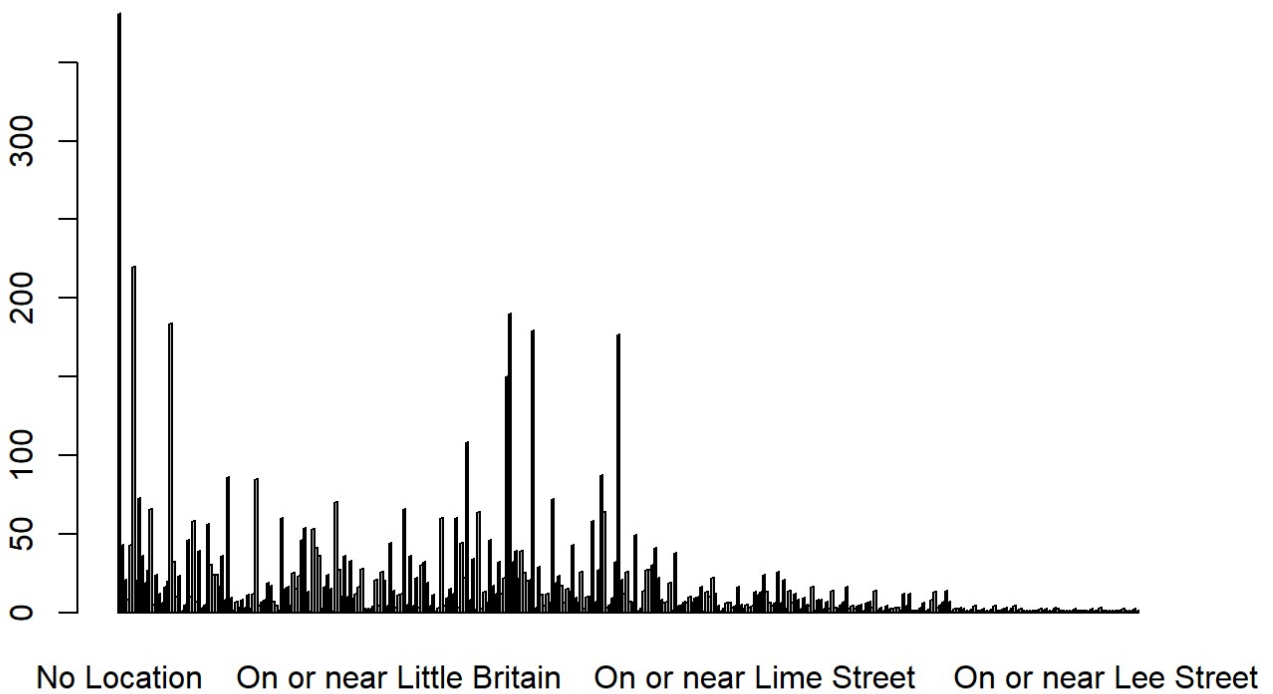
#london_street[london_street$Location=='On or near Aldersgate Street',]
```

- What is the probability that any crime will happen on or near Aldgate?

```
aldersgate <- london_street[grep("Aldersgate",london_street$Location),]
nrow(aldersgate)/nrow(london_street)
```

```
## [1] 0.003163056
```

```
t<-table(london_street$Location)
barplot(t)
```

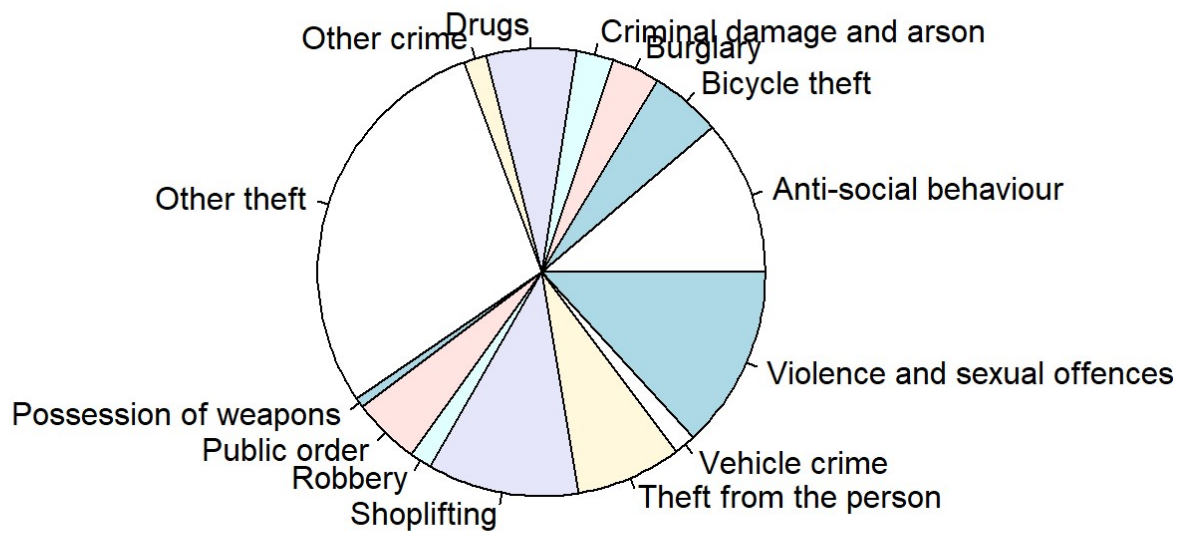


- What percentage of crime in London is 'violence and sexual offences'?

```
crime<-london_street[grepl('Violence and sexual', london_street$Crime.type),]  
nrow(crime)/nrow(london_street)
```

```
## [1] 0.1317413
```

```
t2<-table(london_street$Crime.type)  
pie(t2)
```

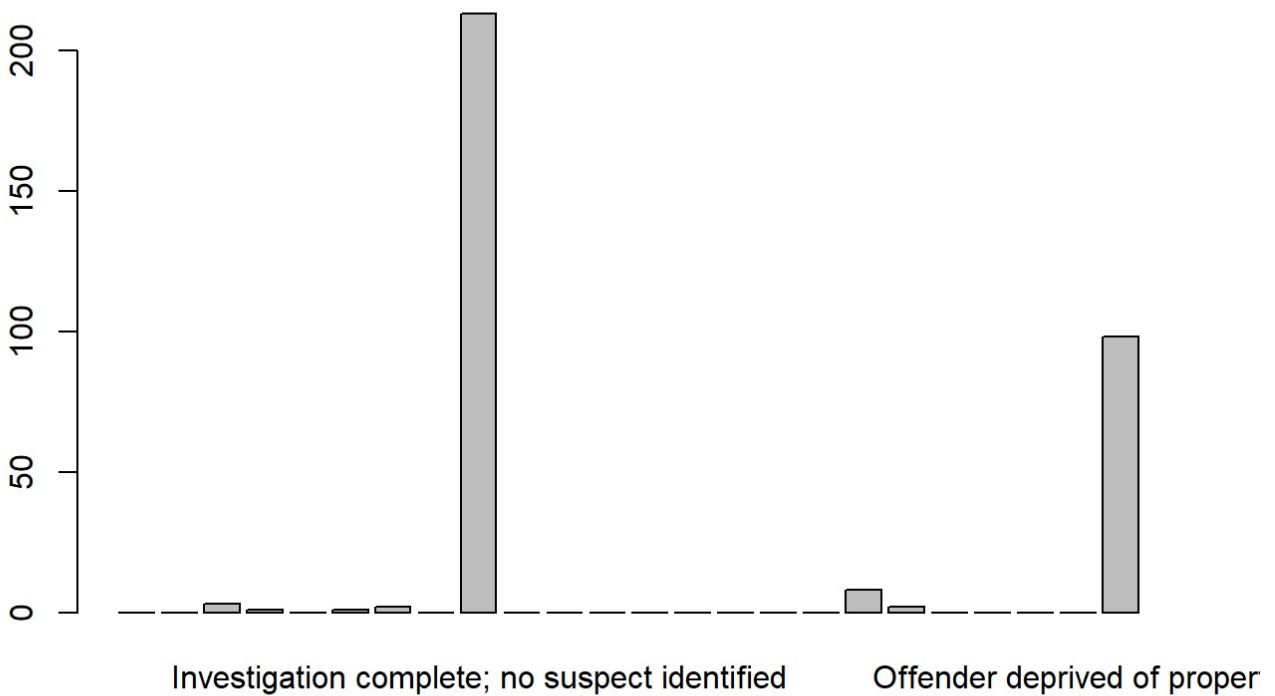


- What is the probability that no suspect will be identified in a bicycle theft?

```
bike <- london_street[grepl('Bicycle', london_street$Crime.type),]  
bike_nosuspect <- bike[grepl('no suspect', bike$Last.outcome.category),]  
nrow(bike_nosuspect)/nrow(bike)
```

```
## [1] 0.6493902
```

```
t3<-table(bike$Last.outcome.category)  
barplot(t3)
```



- What is the probability of 'robbery' on Carthusian Street?

Error: In london street data, there's no robbery on Carthusian Street. In london outcomes, there's no crime type column.

```
csv_files2 = list.files(path = 'dt/outcomes', pattern = "csv$", full.names = TRUE)
csv_list2 = lapply(csv_files2, read.csv, header=TRUE)
london_outcomes = do.call(rbind, csv_list2)

cart_street <- london_street[grepl('Carthusian', london_street$Location),]
cart_outcomes <- london_outcomes[grepl('CARTHUSIAN', london_outcomes$Location),]

cart_rob_street <- cart_street[grepl('Robbery', cart_street$Crime.type),]
nrow(cart_rob_street)
```

```
## [1] 0
```