San Fransisco Larcency Crimes

Tadros Salama

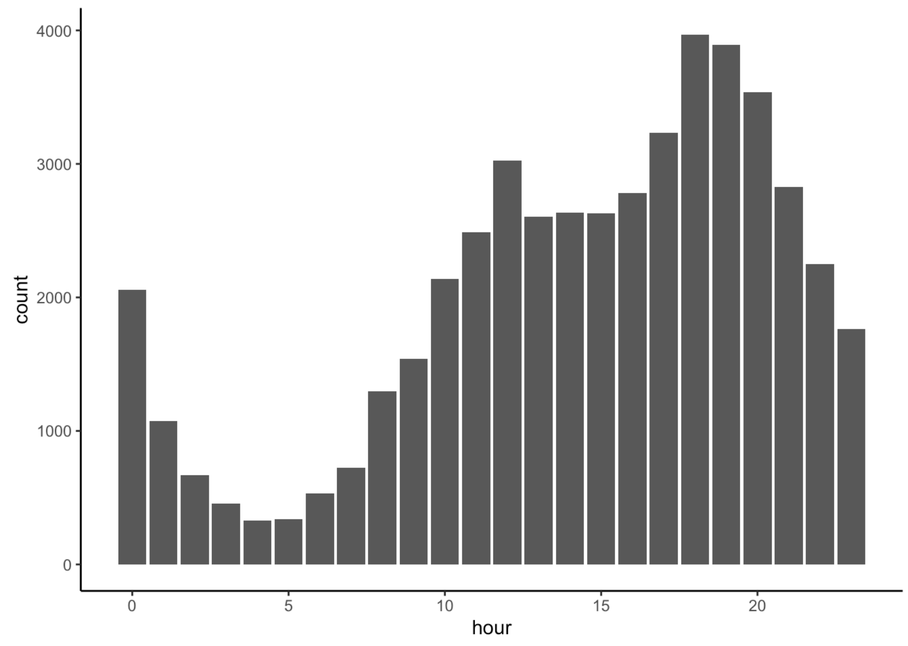
12/5/2020

Amount of larcency crimes in San Fransisco during the whole year of 2018.

```
sfl <- tibble(
date = date(t),
month = month(t, label = T), weekday = wday(t, label = T), hour = hour(t)
)
sfl
```

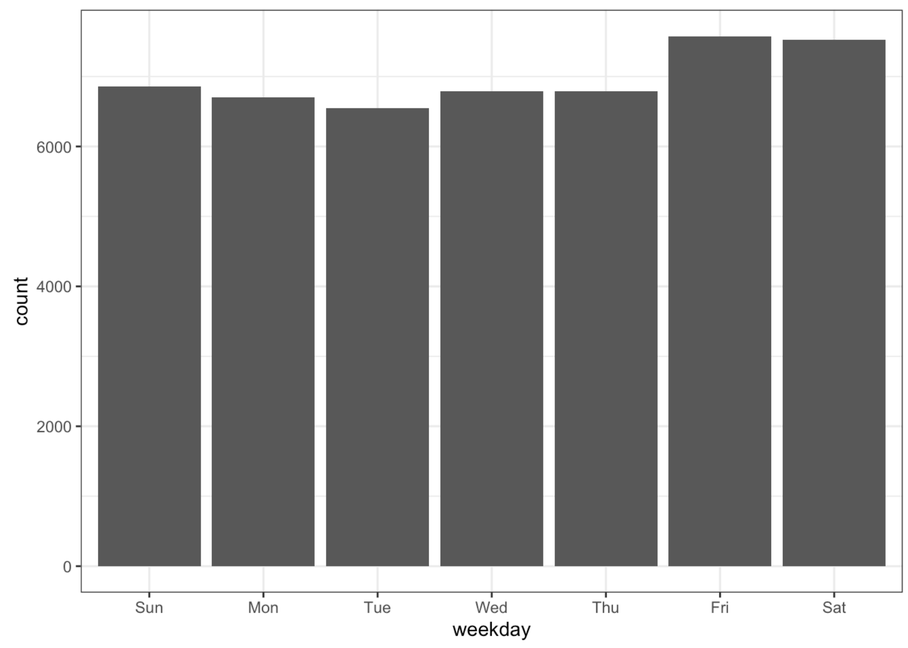
```
## # A tibble: 48,789 x 4
##    date       month weekday  hour
##    <date>     <ord> <ord>   <int>
##  1 2018-10-05 Oct   Fri        16
##  2 2018-10-05 Oct   Fri        15
##  3 2018-06-01 Jun   Fri         0
##  4 2018-10-06 Oct   Sat         4
##  5 2018-10-06 Oct   Sat         8
##  6 2018-10-06 Oct   Sat        10
##  7 2018-10-06 Oct   Sat        16
##  8 2018-10-06 Oct   Sat        15
##  9 2018-10-06 Oct   Sat        18
## 10 2018-10-06 Oct   Sat        16
## # … with 48,779 more rows
```

```
ggplot(sfl, aes(hour)) + geom_bar() + theme_classic()
```



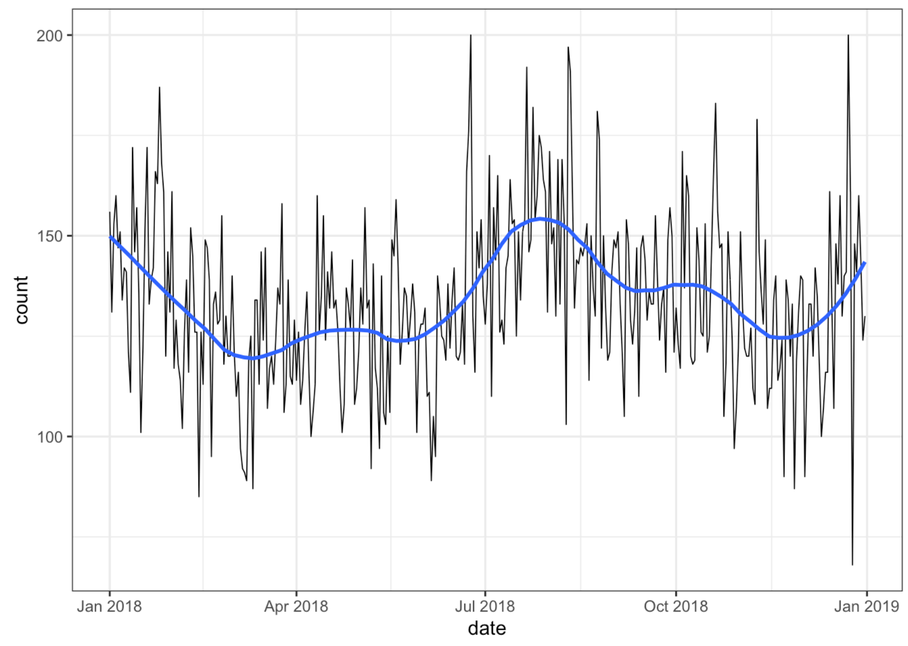
(d)

```
ggplot(sfl, aes(weekday)) + geom_bar() + theme_bw()
```




```
sfl_daily <- sfl %>% group_by(date) %>% summarize(count = n())
 ggplot(sfl_daily, aes(x=date, y=count)) + geom_line(size=0.3) +
geom_smooth(span = 0.3, se = FALSE) + theme_bw()
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```



Bar plot **(c)** of the hourly larceny thefts, shows a concentration of thefts between 5:00pm - 8:00pm. Since that is the typical period people leave work, a majority of these thefts could be targeting commuters.

```
dec_sfl <- sfl %>% group_by(date, month) %>% summarise(count = n()) %>%
  filter(month == 'Dec')
```

```
## `summarise()` has grouped output by 'date'. You can override using the `.groups` argument.
```

```
ggplot(dec_sfl, aes(x=date, y=count)) + geom_line(size=0.5) +
geom_smooth(span = 0.5, se = FALSE) + theme_bw()
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```