

Week2_MelbWalk

N Kaso

09/09/2020

Week 2 Melbourne Walk Exercise

Read the data

```
walkers <- read_csv("~/Documents/data_science_decision_making/melb_walk.csv")
```

```
## Parsed with column specification:
## cols(
##   Sensor = col_character(),
##   Date_Time = col_datetime(format = ""),
##   Date = col_date(format = ""),
##   Time = col_double(),
##   Count = col_double()
## )
```

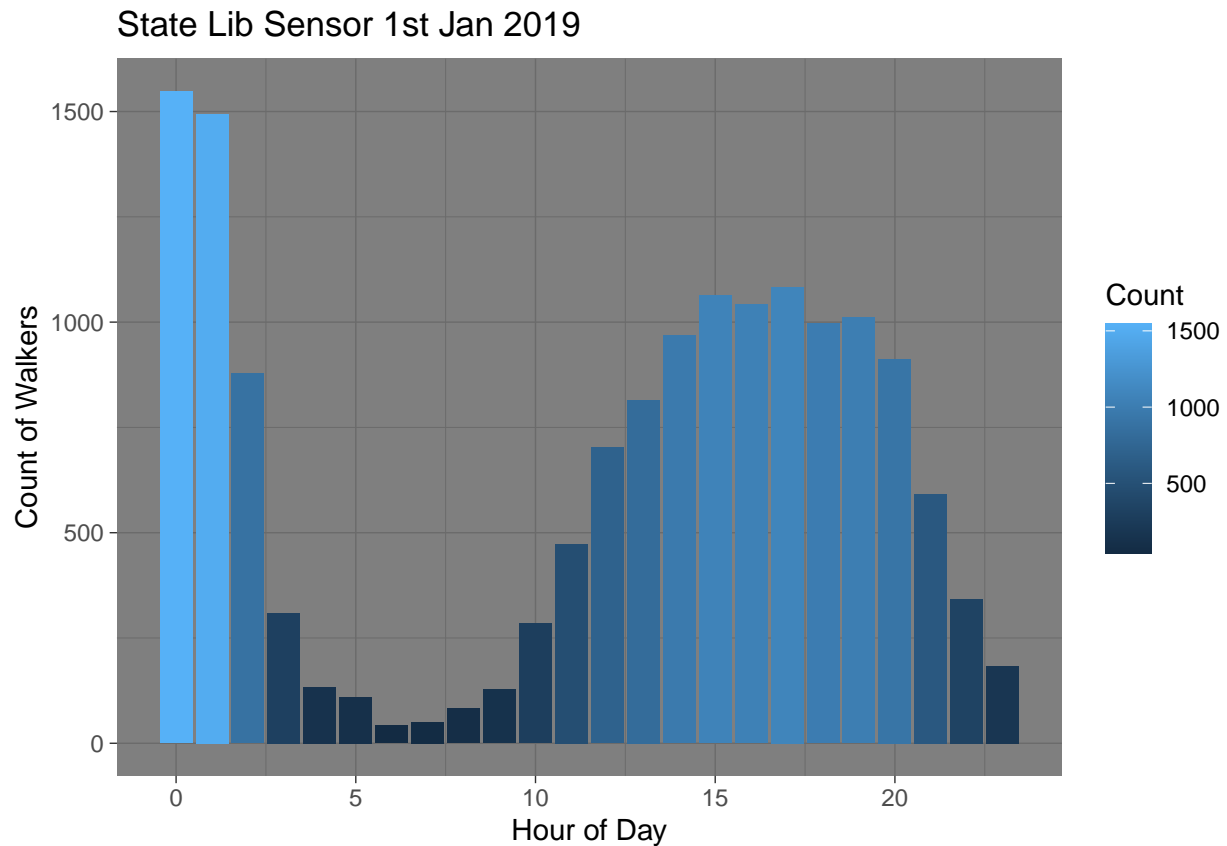
Filter for State Library on 1st Jan 2019

```
state_lib <- filter(walkers, Date == "2019-01-01", Sensor == "State Library")
state_lib
```

```
## # A tibble: 24 x 5
##   Sensor      Date_Time      Date      Time Count
##   <chr>      <dtm>      <date>    <dbl> <dbl>
## 1 State Library 2018-12-31 13:00:00 2019-01-01     0  1548
## 2 State Library 2018-12-31 14:00:00 2019-01-01     1  1494
## 3 State Library 2018-12-31 15:00:00 2019-01-01     2   878
## 4 State Library 2018-12-31 16:00:00 2019-01-01     3   309
## 5 State Library 2018-12-31 17:00:00 2019-01-01     4   133
## 6 State Library 2018-12-31 18:00:00 2019-01-01     5   110
## 7 State Library 2018-12-31 19:00:00 2019-01-01     6    42
## 8 State Library 2018-12-31 20:00:00 2019-01-01     7    50
## 9 State Library 2018-12-31 21:00:00 2019-01-01     8    83
## 10 State Library 2018-12-31 22:00:00 2019-01-01     9   128
## # ... with 14 more rows
```

Plot a bar chart

```
state_lib_plot <- ggplot(state_lib, aes(x = Time, y = Count, fill = Count)) +  
  geom_bar(stat = "Identity") +  
  labs(x = "Hour of Day", y = "Count of Walkers", title = "State Lib Sensor 1st Jan 2019")  
  theme_dark()  
state_lib_plot
```



Filter again for one sensor over all days

```
state_lib <- filter(walkers, Sensor == "State Library")  
state_lib
```

A tibble: 744 x 5

##	Sensor	Date_Time	Date	Time	Count
##	<chr>	<dtm>	<date>	<dbl>	<dbl>
##	1	State Library	2018-12-31 13:00:00	2019-01-01	0 1548
##	2	State Library	2018-12-31 14:00:00	2019-01-01	1 1494
##	3	State Library	2018-12-31 15:00:00	2019-01-01	2 878
##	4	State Library	2018-12-31 16:00:00	2019-01-01	3 309
##	5	State Library	2018-12-31 17:00:00	2019-01-01	4 133
##	6	State Library	2018-12-31 18:00:00	2019-01-01	5 110
##	7	State Library	2018-12-31 19:00:00	2019-01-01	6 42
##	8	State Library	2018-12-31 20:00:00	2019-01-01	7 50
##	9	State Library	2018-12-31 21:00:00	2019-01-01	8 83

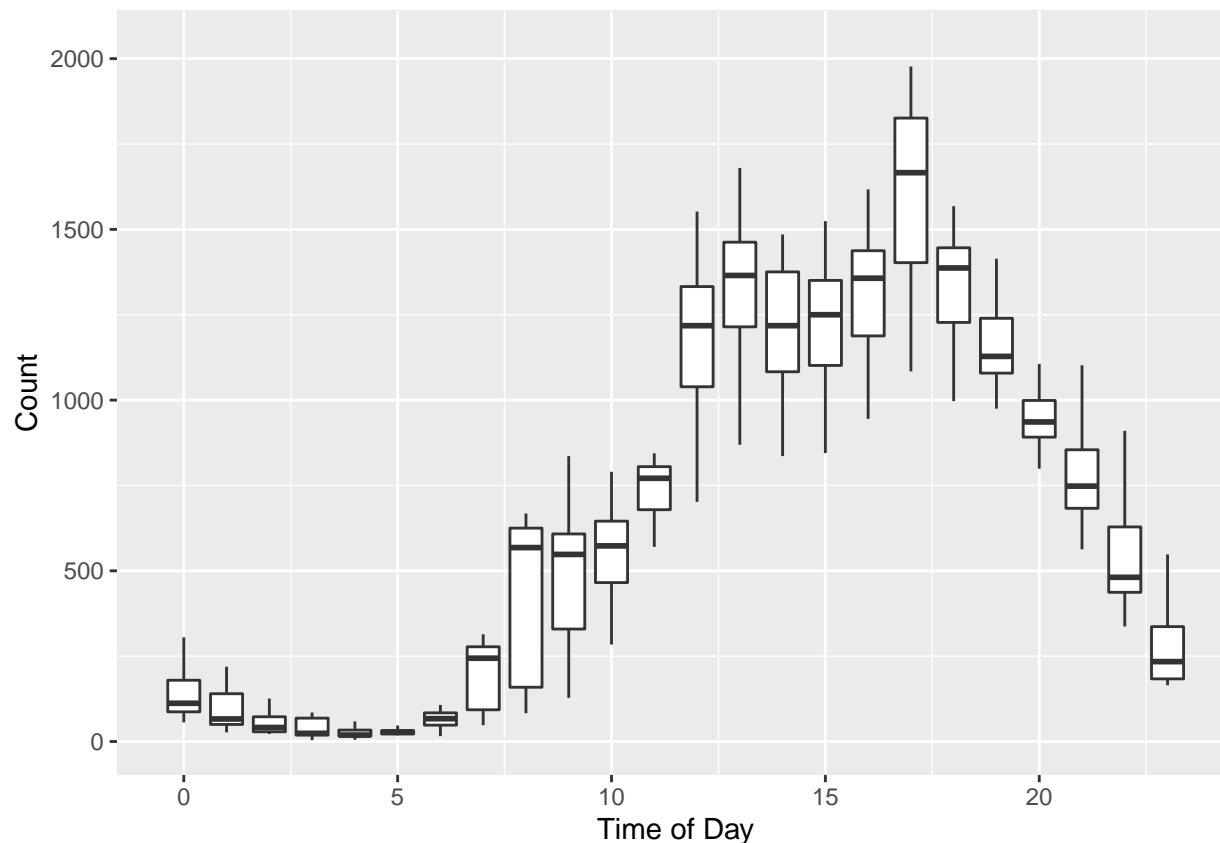
```
## 10 State Library 2018-12-31 22:00:00 2019-01-01      9    128
## # ... with 734 more rows
```

Plot a box-plot

```
state_lib_boxplot <- ggplot(state_lib, aes(x = Sensor, y = Count)) +  
  geom_boxplot()  
state_lib_boxplot
```



```
state_lib_boxsbs <- ggplot(state_lib, aes(x = Time, y = Count, group = Time)) +  
  geom_boxplot(outlier.alpha = 0) +  
  xlab("Time of Day")  
state_lib_boxsbs
```

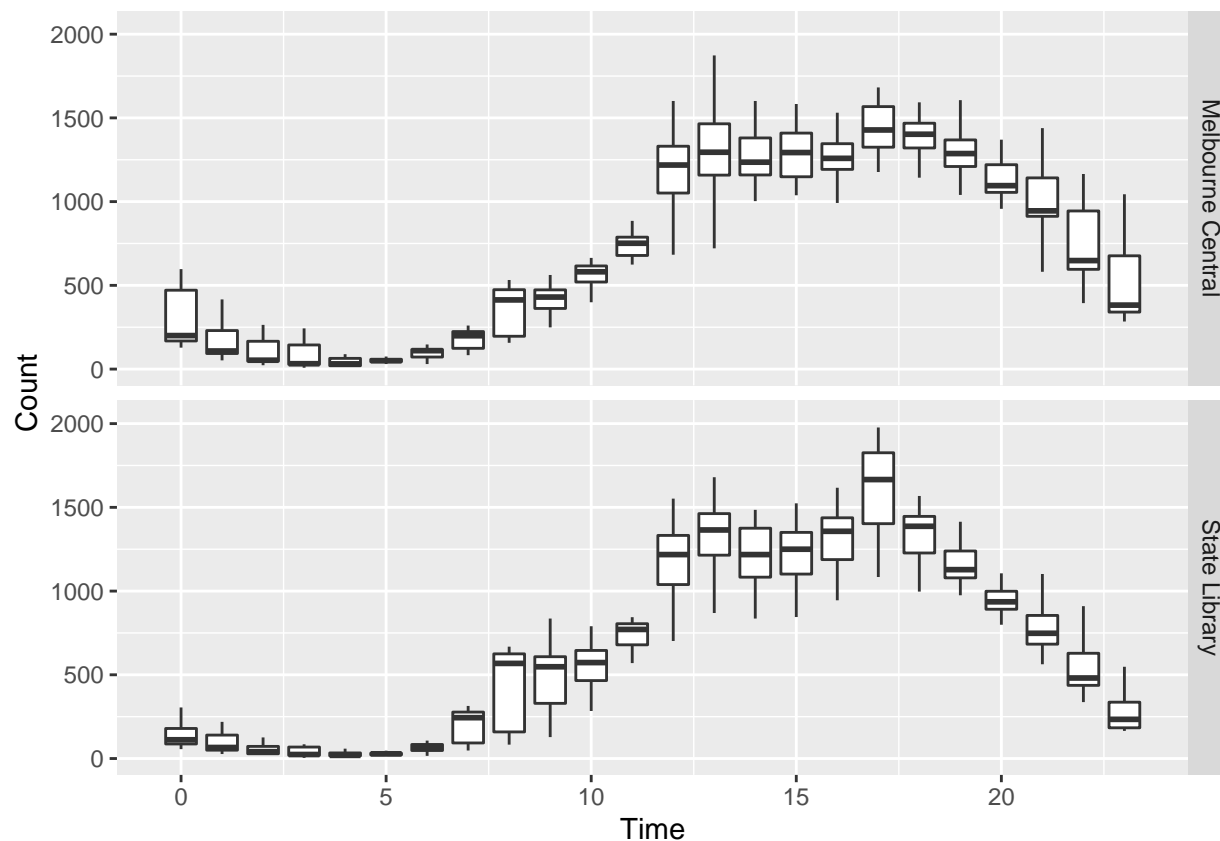


```
### Filter again for two sensors
```

```
walkers_mc_sl <- filter(walkers, Sensor %in% c("Melbourne Central", "State Library"), !is.na(Count), Count > 0)
walkers_mc_sl
```

```
## # A tibble: 1,359 x 5
##   Sensor      Date_Time      Date      Time Count
##   <chr>      <dtm>      <date>    <dbl> <dbl>
## 1 State Library 2018-12-31 13:00:00 2019-01-01     0 1548
## 2 State Library 2018-12-31 14:00:00 2019-01-01     1 1494
## 3 State Library 2018-12-31 15:00:00 2019-01-01     2  878
## 4 State Library 2018-12-31 16:00:00 2019-01-01     3  309
## 5 State Library 2018-12-31 17:00:00 2019-01-01     4  133
## 6 State Library 2018-12-31 18:00:00 2019-01-01     5   110
## 7 State Library 2018-12-31 19:00:00 2019-01-01     6    42
## 8 State Library 2018-12-31 20:00:00 2019-01-01     7    50
## 9 State Library 2018-12-31 21:00:00 2019-01-01     8    83
## 10 State Library 2018-12-31 22:00:00 2019-01-01     9   128
## # ... with 1,349 more rows
```

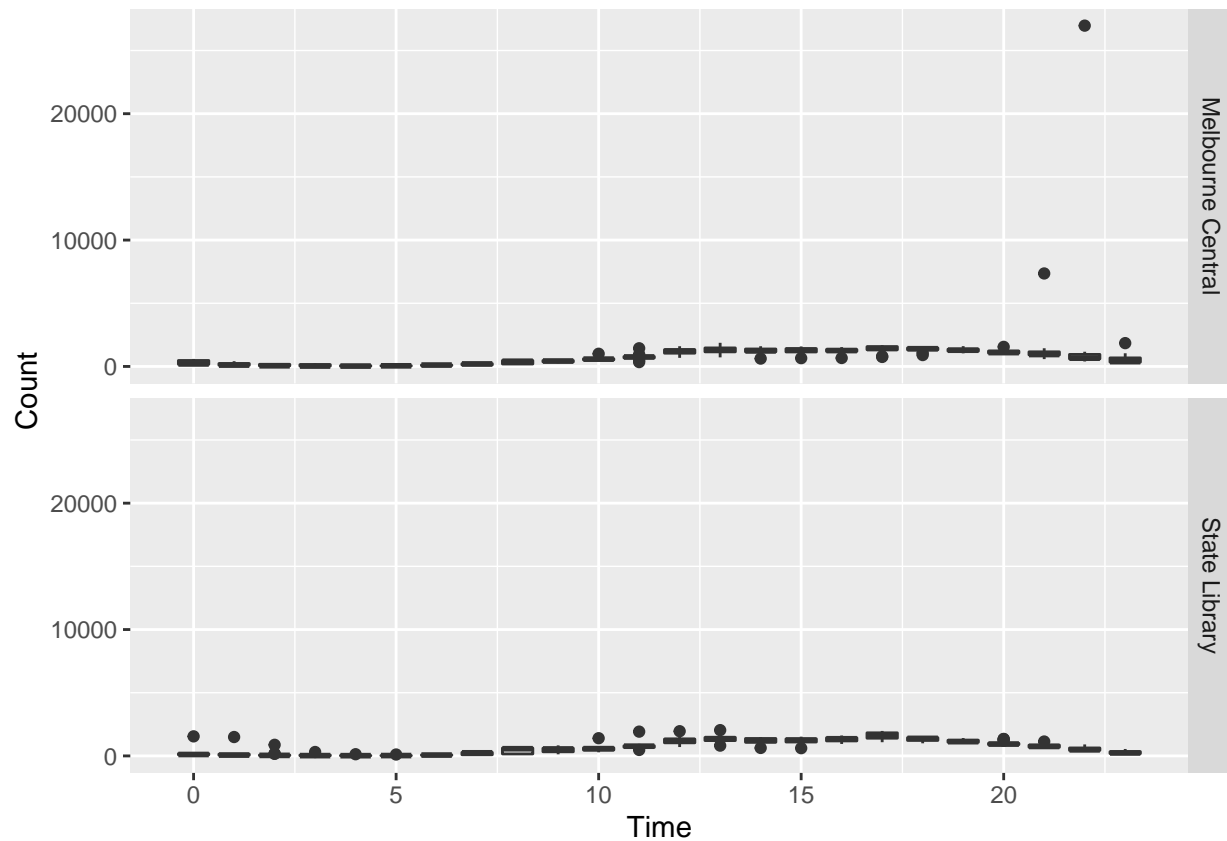
```
state_lib_boxsbs2 <- ggplot(walkers_mc_sl, aes(x = Time, y = Count, group = Time)) +
  geom_boxplot(outlier.alpha = 0) +
  facet_grid(Sensor~.)
state_lib_boxsbs2
```



```
walkers_mc_sl <- filter(walkers, Sensor %in% c("Melbourne Central", "State Library"), !is.na(Count))
walkers_mc_sl
```

```
## # A tibble: 1,361 x 5
##   Sensor      Date_Time      Date      Time Count
##   <chr>      <dtm>      <date>    <dbl> <dbl>
## 1 State Library 2018-12-31 13:00:00 2019-01-01     0 1548
## 2 State Library 2018-12-31 14:00:00 2019-01-01     1 1494
## 3 State Library 2018-12-31 15:00:00 2019-01-01     2  878
## 4 State Library 2018-12-31 16:00:00 2019-01-01     3  309
## 5 State Library 2018-12-31 17:00:00 2019-01-01     4  133
## 6 State Library 2018-12-31 18:00:00 2019-01-01     5  110
## 7 State Library 2018-12-31 19:00:00 2019-01-01     6   42
## 8 State Library 2018-12-31 20:00:00 2019-01-01     7   50
## 9 State Library 2018-12-31 21:00:00 2019-01-01     8   83
## 10 State Library 2018-12-31 22:00:00 2019-01-01     9  128
## # ... with 1,351 more rows
```

```
state_lib_boxsbs2 <- ggplot(walkers_mc_sl, aes(x = Time, y = Count, group = Time)) +
  geom_boxplot() +
  facet_grid(Sensor~.)
state_lib_boxsbs2
```



```
# Load packages
library(tidyverse)
library(lubridate)

##
## Attaching package: 'lubridate'

## The following objects are masked from 'package:base':
##
##   date, intersect, setdiff, union

library(rwalkr)

# Use melb_walk() to extract data
ped <- melb_walk(from=dmy("01072018"),
                 to=dmy("31072018"))

# Filter the data
ped <- ped %>%
  filter(Sensor %in% c("Melbourne Central", "Flinders Street Station Underpass"))

# Plot box plots
ggplot(ped, aes(x=Sensor, y=Count)) +
  geom_boxplot()
```

