Ggplot2 tutorial - Command lines Part 1

Ariane Ducellier

10/10/2023

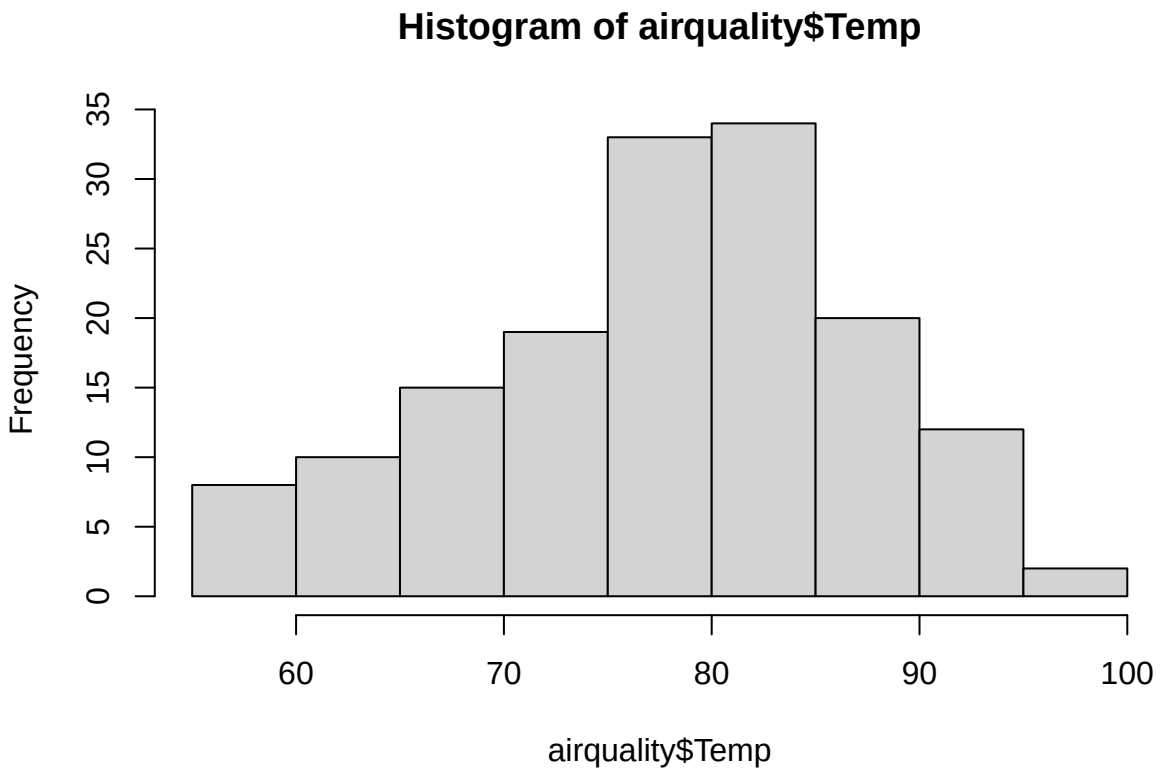
Load R packages

```
library(ggplot2)
library(Lock5Data)
```

Part 1 - Basic Plotting in ggplot2

Histograms

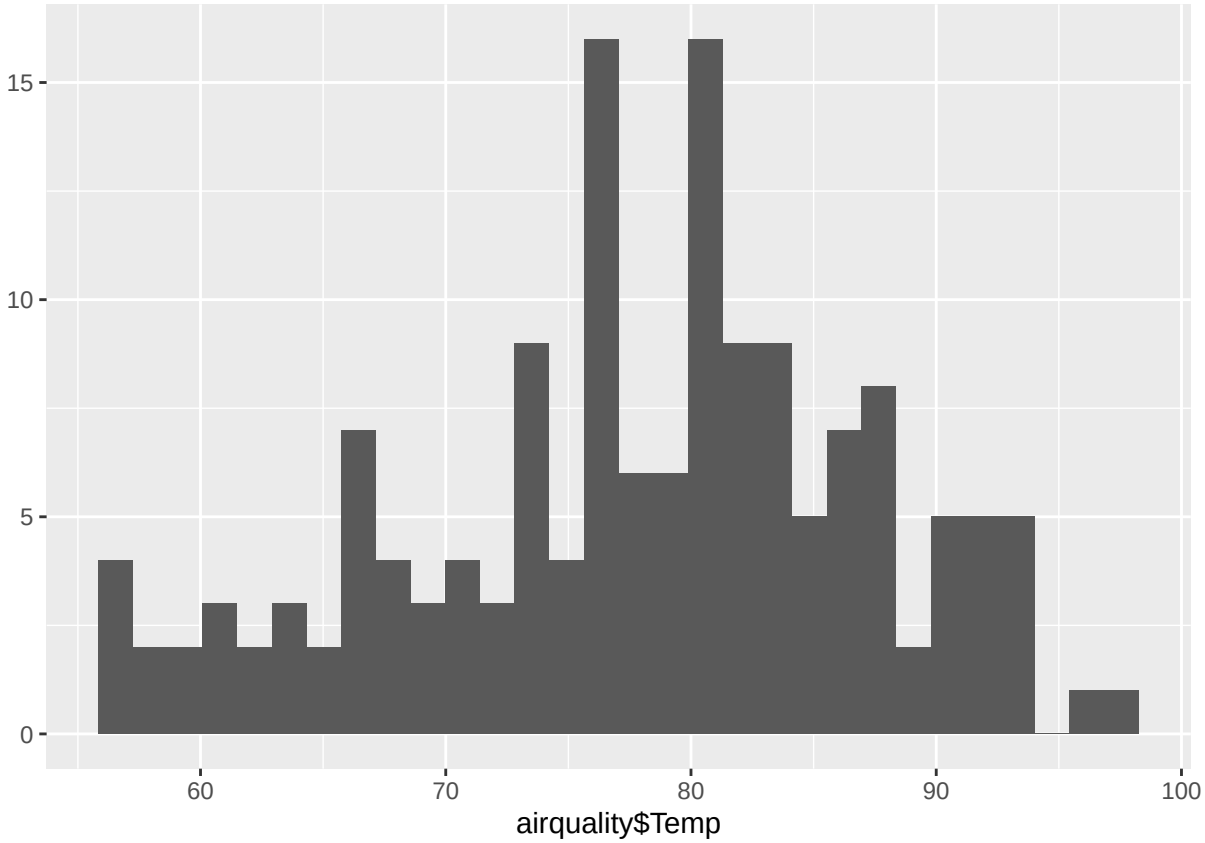
```
hist(airquality$Temp)
```




```
qplot(airquality$Temp)
```

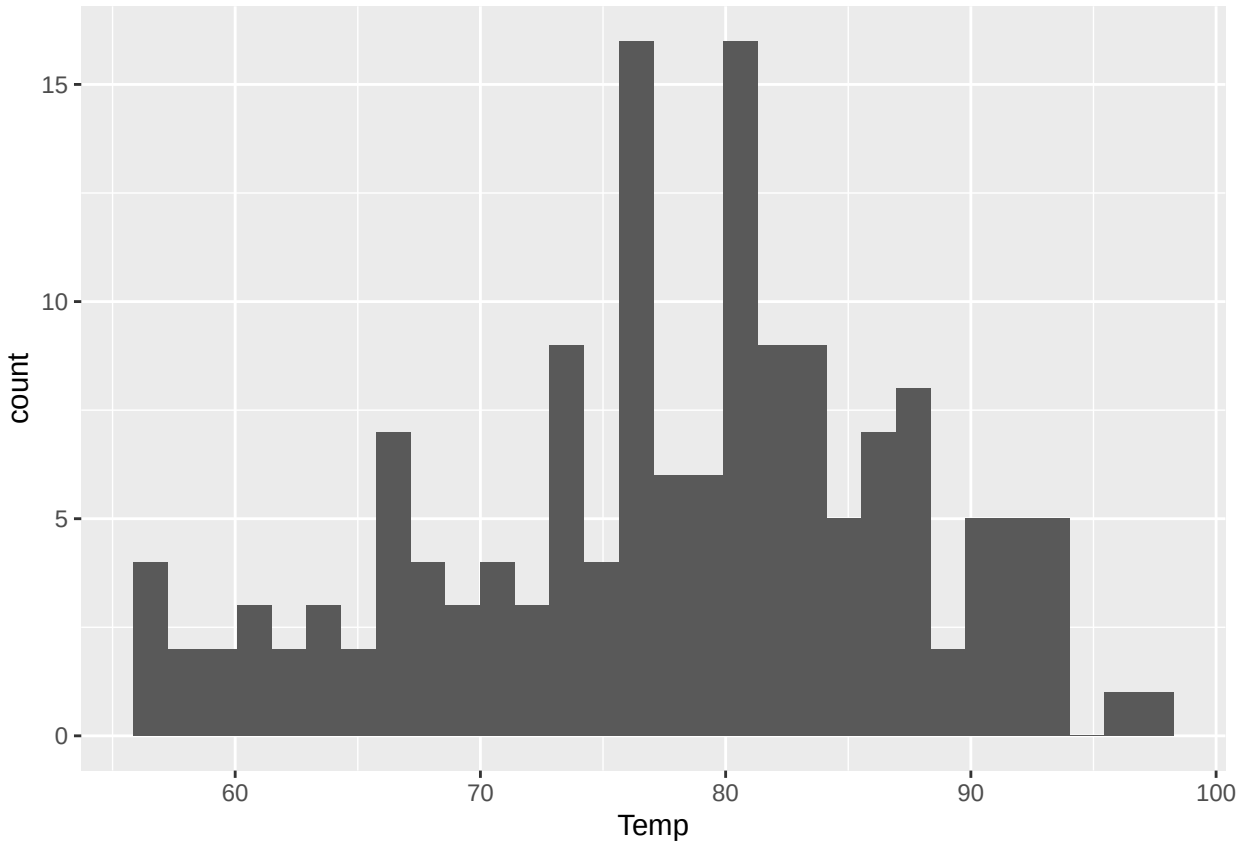
```
## Warning: `qplot()` was deprecated in ggplot2 3.4.0.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



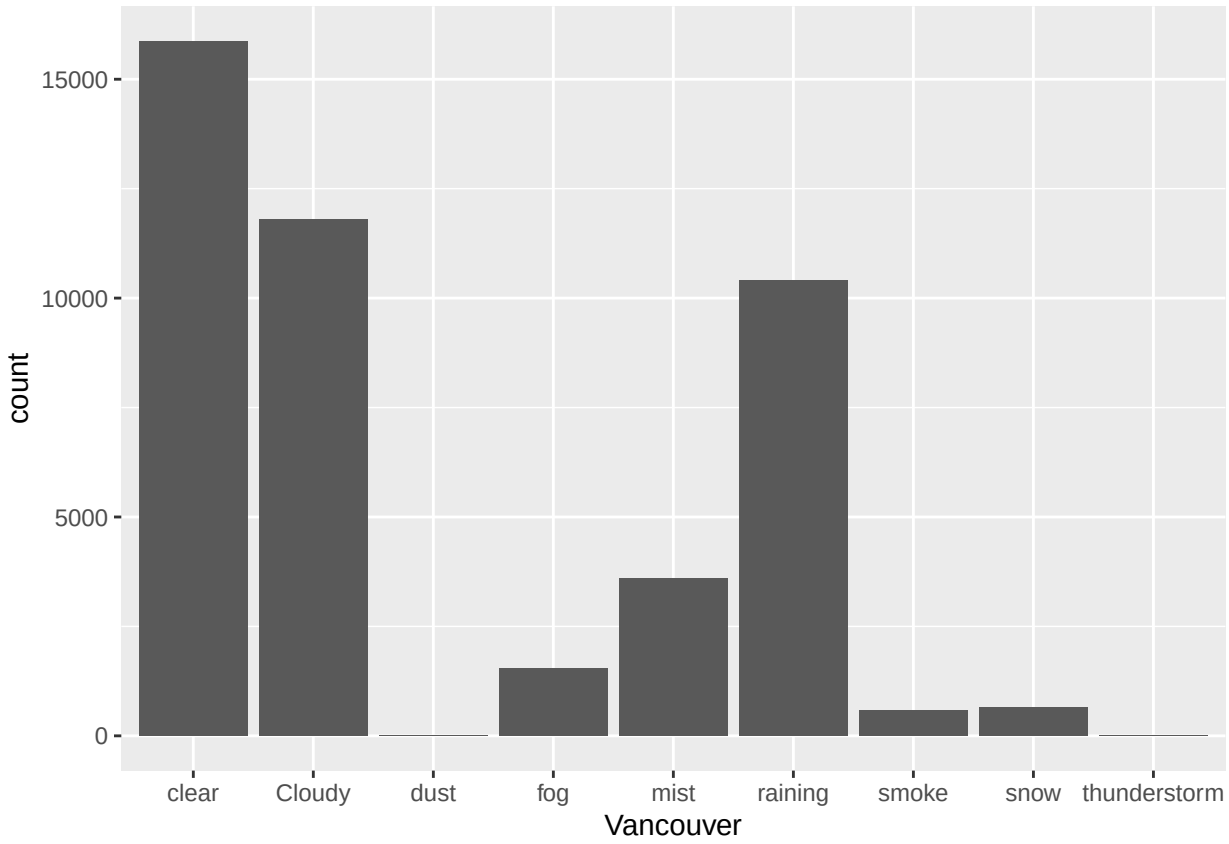

```
ggplot(airquality, aes(x=Temp)) + geom_histogram()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

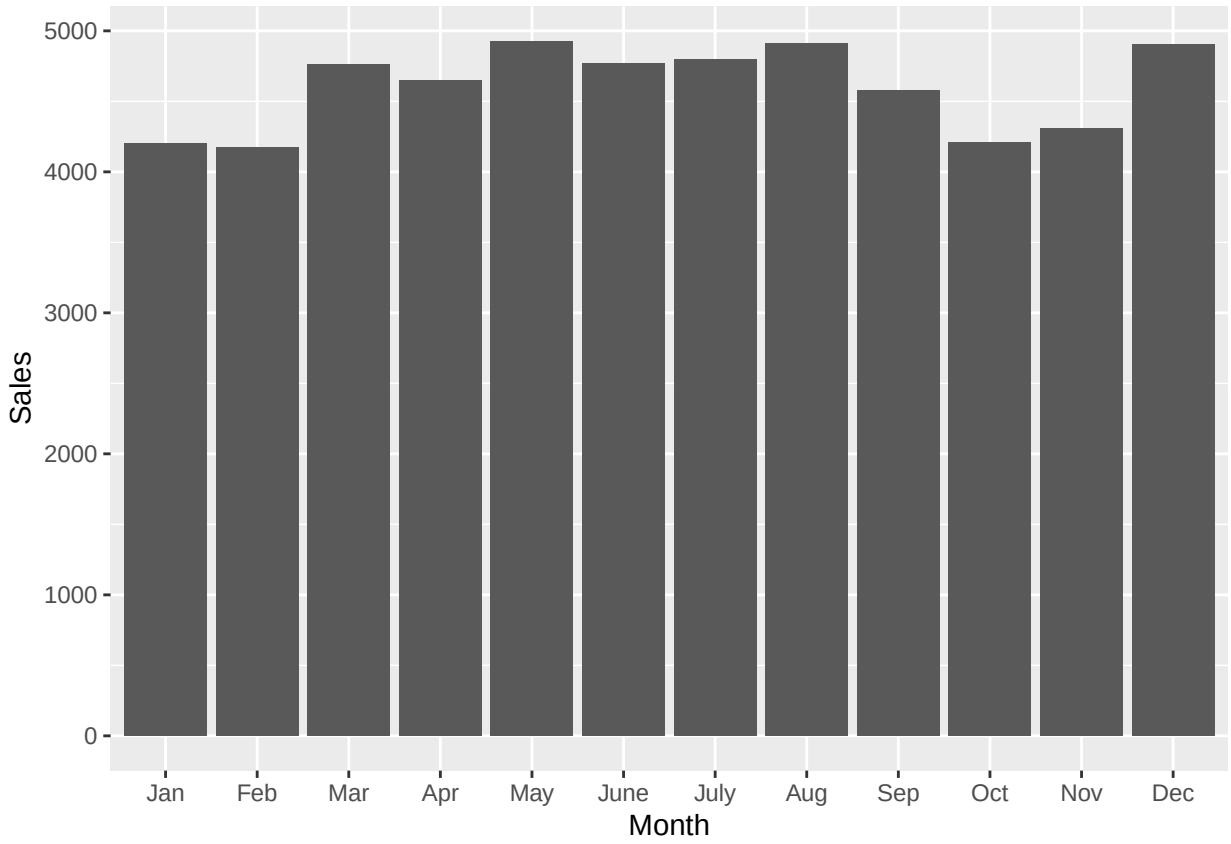


Bar plots

```
df_desc <- read.csv("../data/historical-hourly-weather-data/weather_description.csv")
ggplot(df_desc, aes(x=Vancouver)) + geom_bar()
```



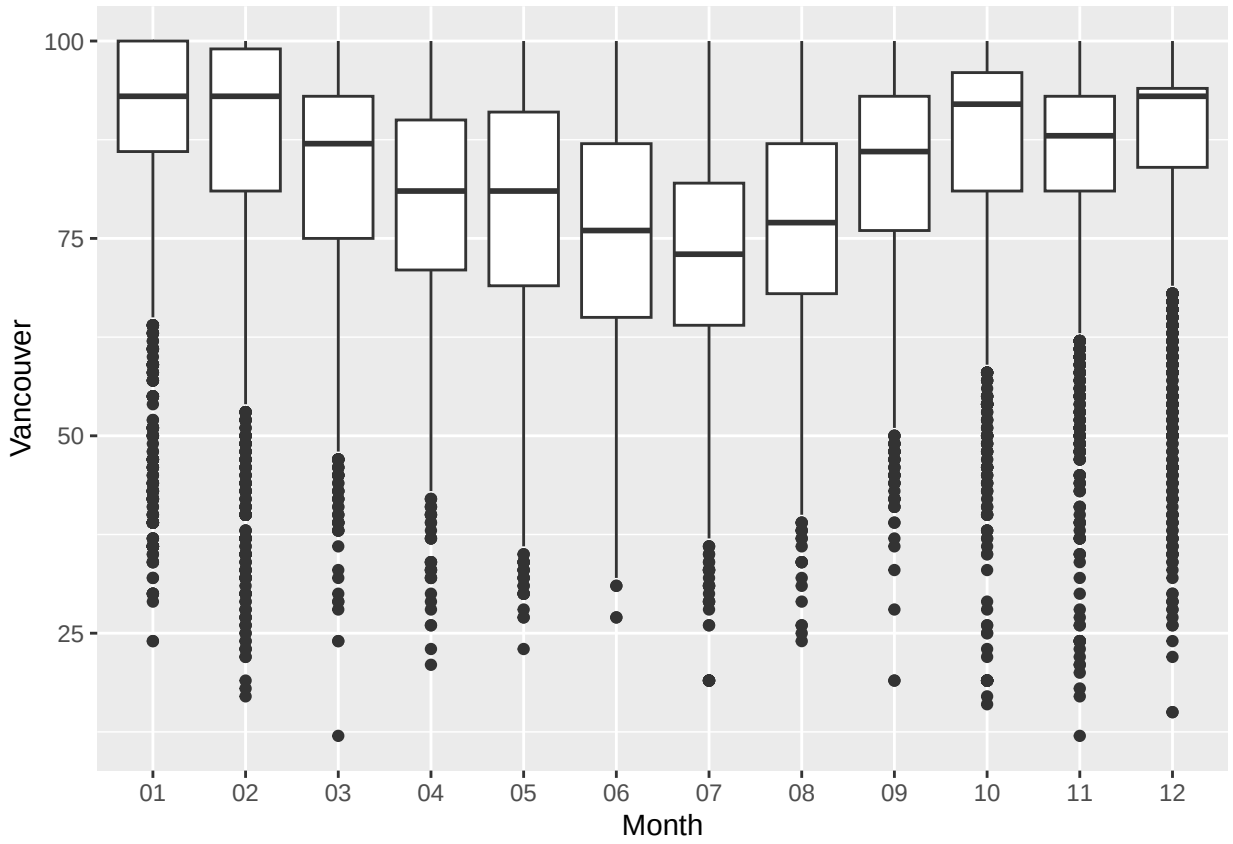

```
df <- na.omit(RetailSales)
months_of_the_year <- c("Jan", "Feb", "Mar", "Apr", "May", "June",
                        "July", "Aug", "Sep", "Oct", "Nov", "Dec")
ggplot(df) +
geom_bar(aes(x=factor(Month, months_of_the_year), y=Sales), stat="identity") +
xlab("Month")
```



Box plots

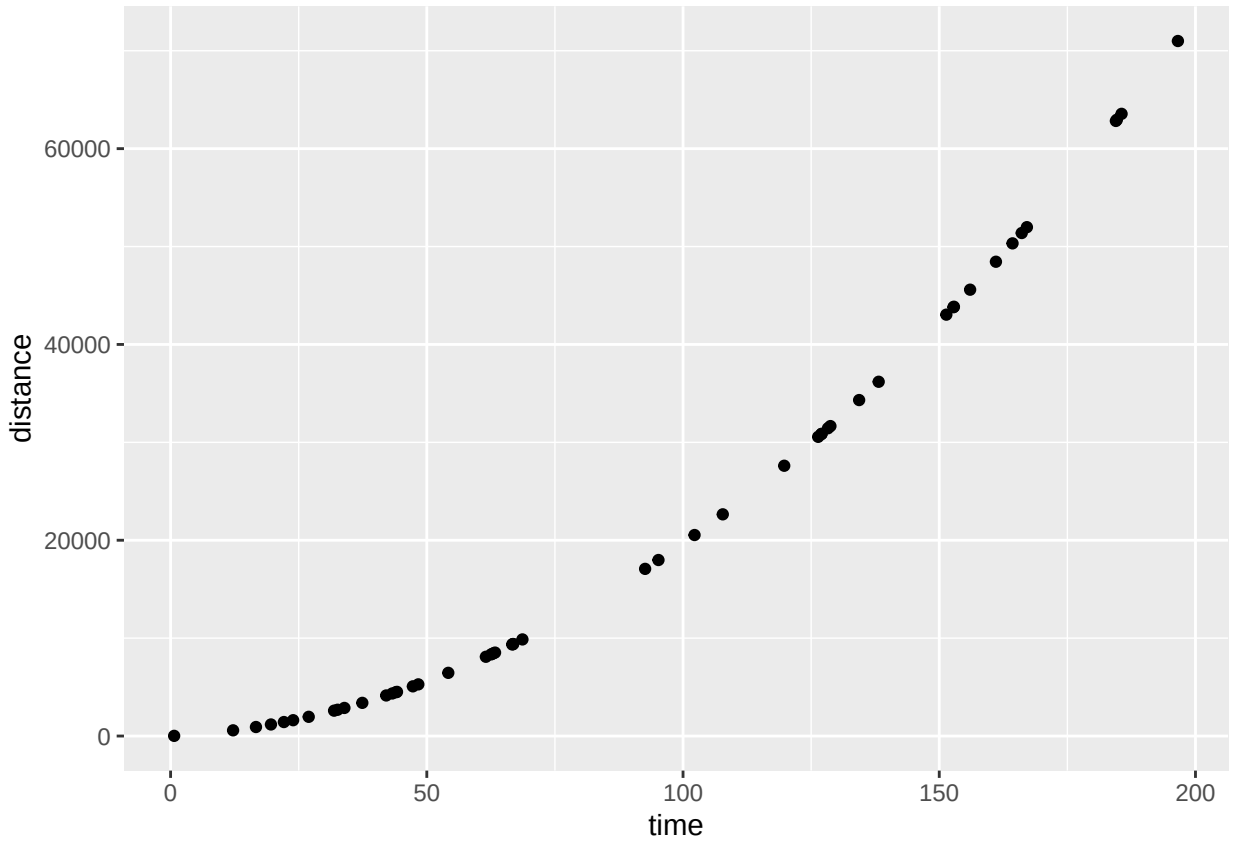
```
df_hum <- read.csv("../data/historical-hourly-weather-data/humidity.csv")
df_hum$datetime <- as.character(df_hum$datetime)
df_hum$Month <- substr(df_hum$datetime, 6, 7)
ggplot(df_hum, aes(x=Month, y=Vancouver)) +
geom_boxplot()
```

```
## Warning: Removed 1826 rows containing non-finite values (`stat_boxplot()`).
```

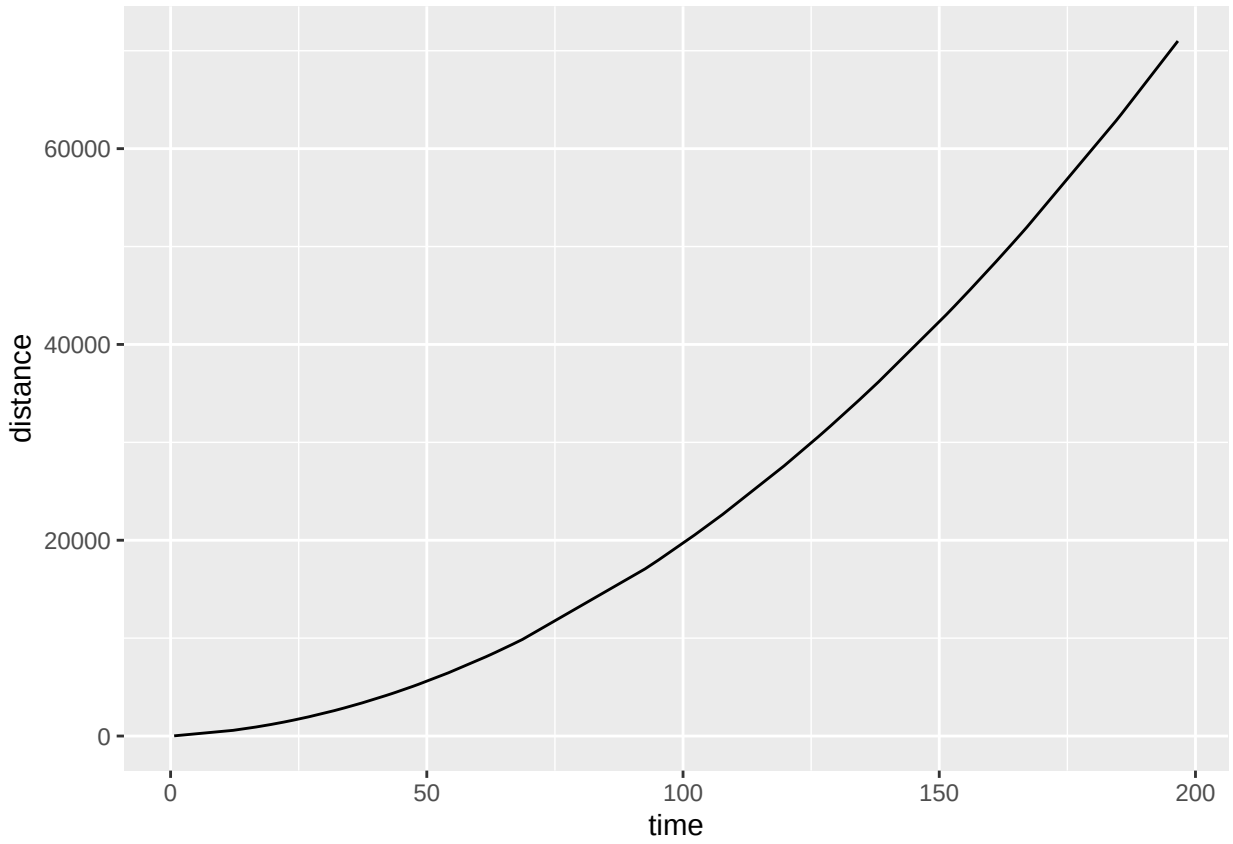


Scatter plots and line plots

```
a = 3.4
v0 = 27
time <- runif(50, min=0, max=200)
distance <- sapply(time, function(x) v0 * x + 0.5 * a * x^2)
df <- data.frame(time,distance)
ggplot(df, aes(x=time, y=distance)) + geom_point()
```




```
ggplot(df, aes(x=time, y=distance)) + geom_line()
```

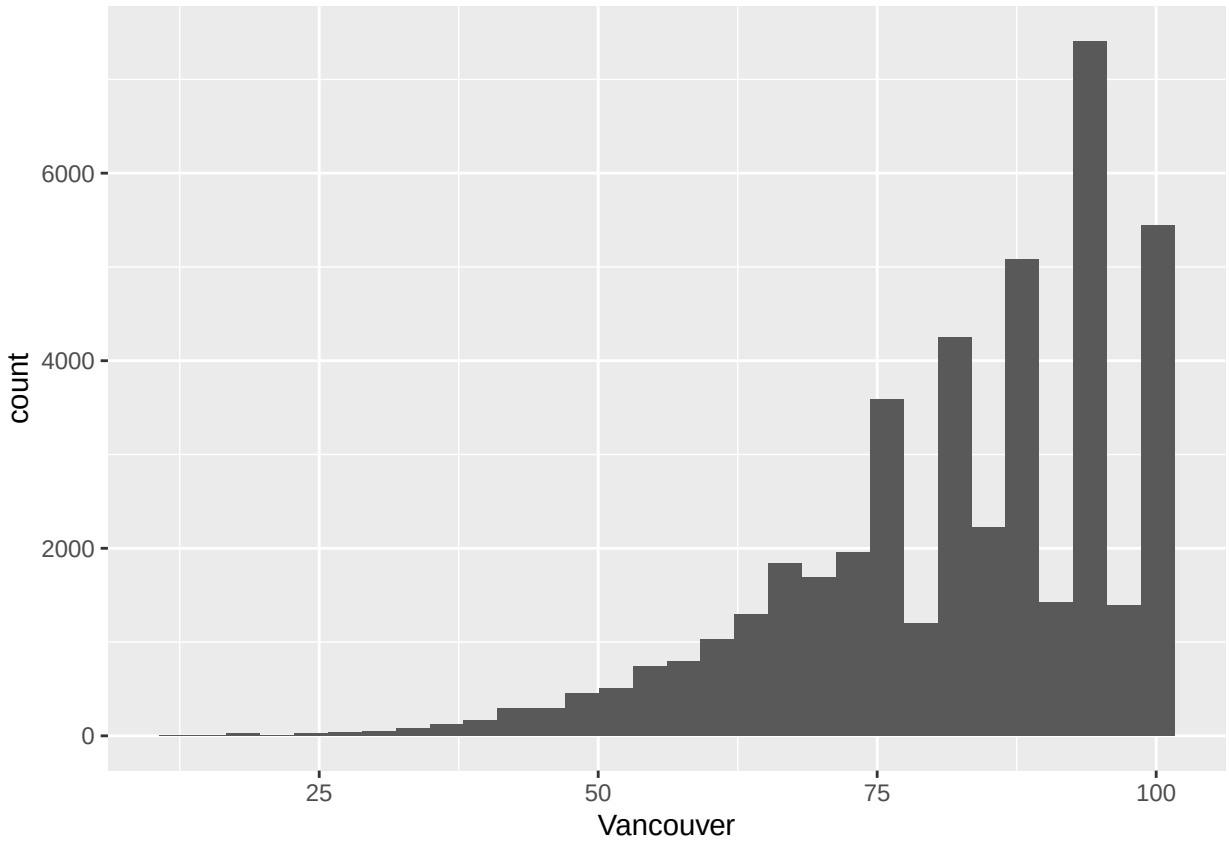


Changing histogram defaults and adding aesthetics

```
df_hum <- read.csv("../data/historical-hourly-weather-data/humidity.csv")
ggplot(df_hum, aes(x=Vancouver)) + geom_histogram()
```

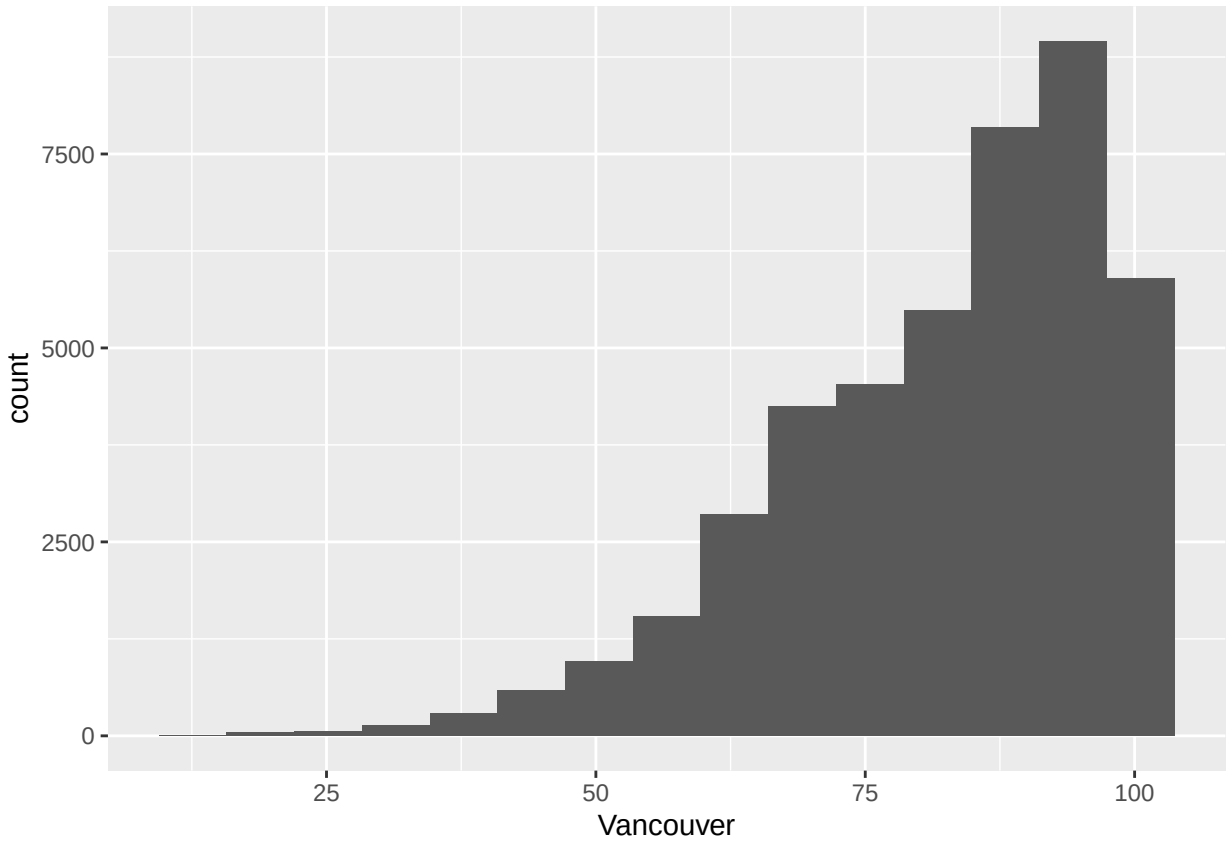
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
## Warning: Removed 1826 rows containing non-finite values (`stat_bin()`).
```



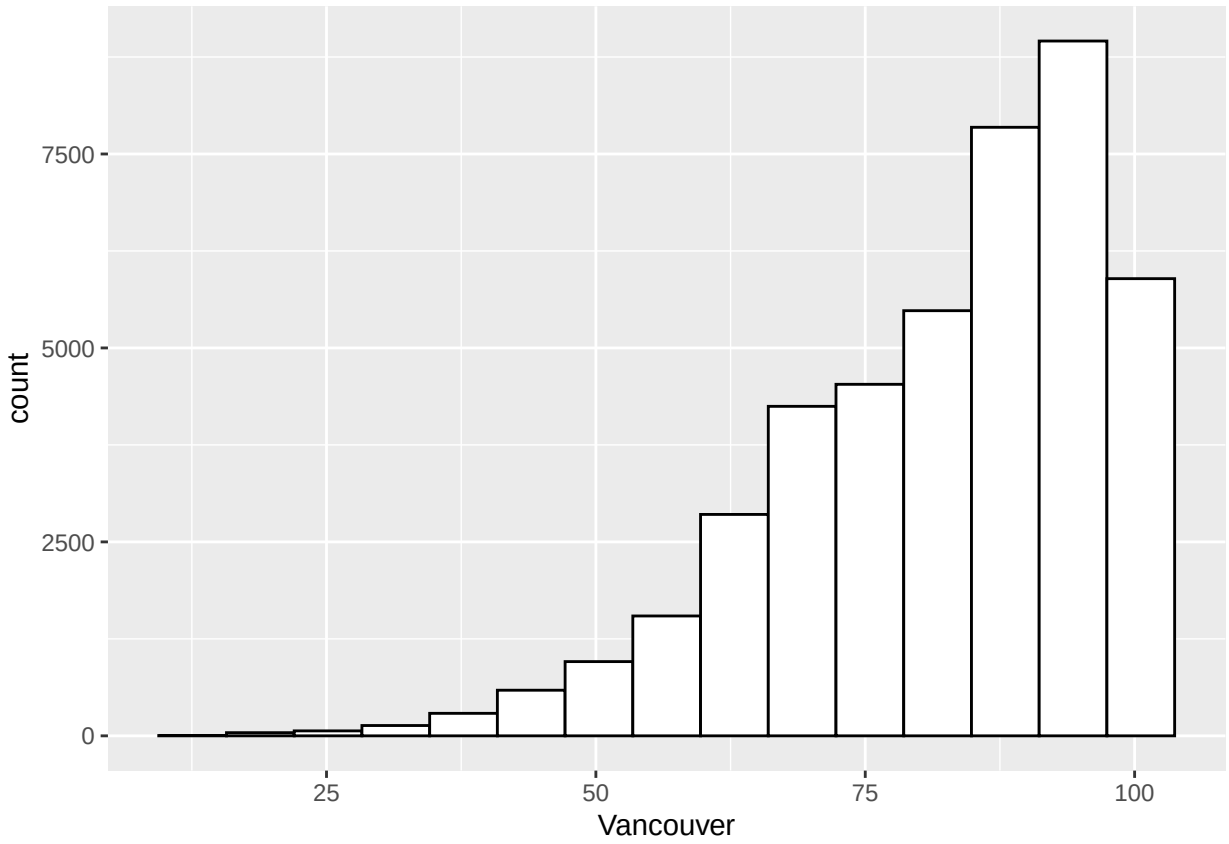

```
ggplot(df_hum, aes(x=Vancouver)) + geom_histogram(bins=15)
```

```
## Warning: Removed 1826 rows containing non-finite values (`stat_bin()`).
```



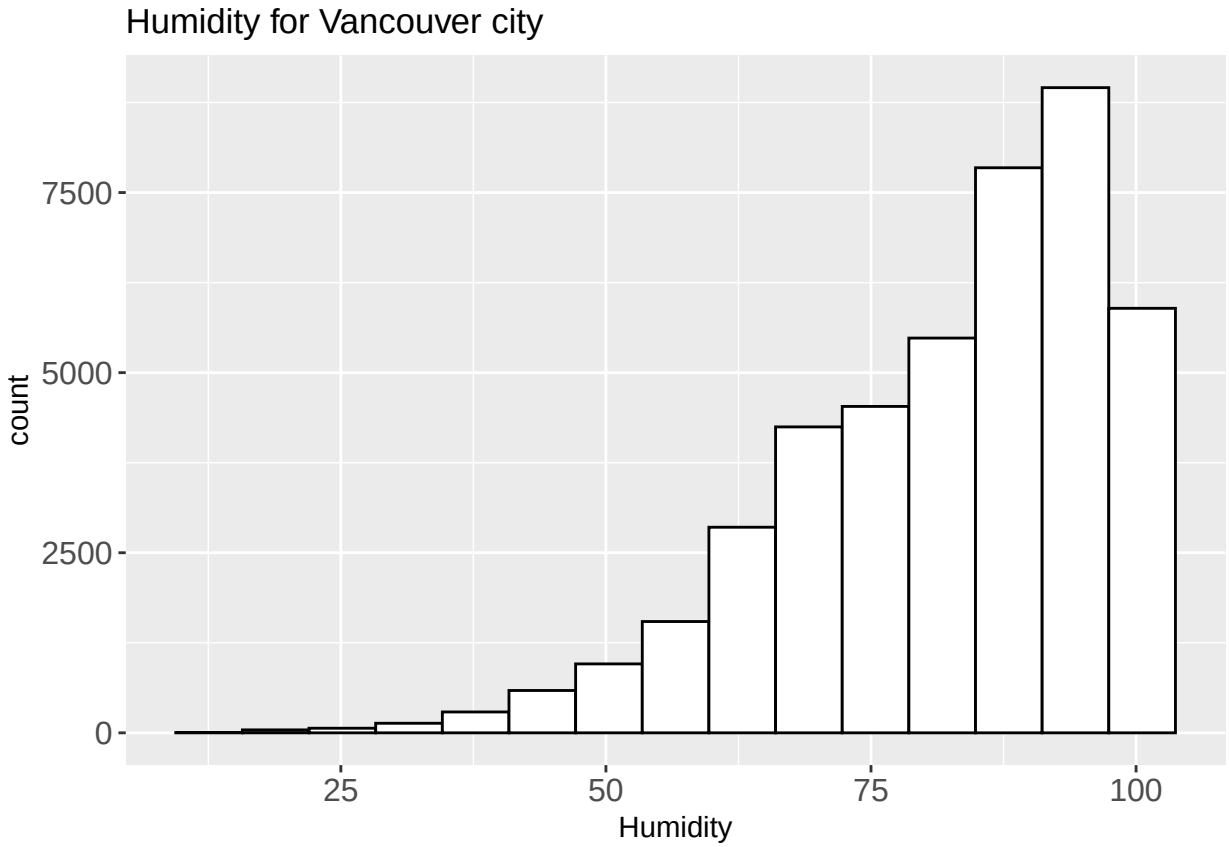

```
ggplot(df_hum, aes(x=Vancouver)) + geom_histogram(bins=15, fill="white", color=1)
```

```
## Warning: Removed 1826 rows containing non-finite values (`stat_bin()`).
```




```
ggplot(df_hum, aes(x=Vancouver)) +
geom_histogram(bins=15, fill="white", color=1) +
ggtitle("Humidity for Vancouver city") +
xlab("Humidity") +
theme(axis.text.x=element_text(size=12), axis.text.y=element_text(size=12))
```

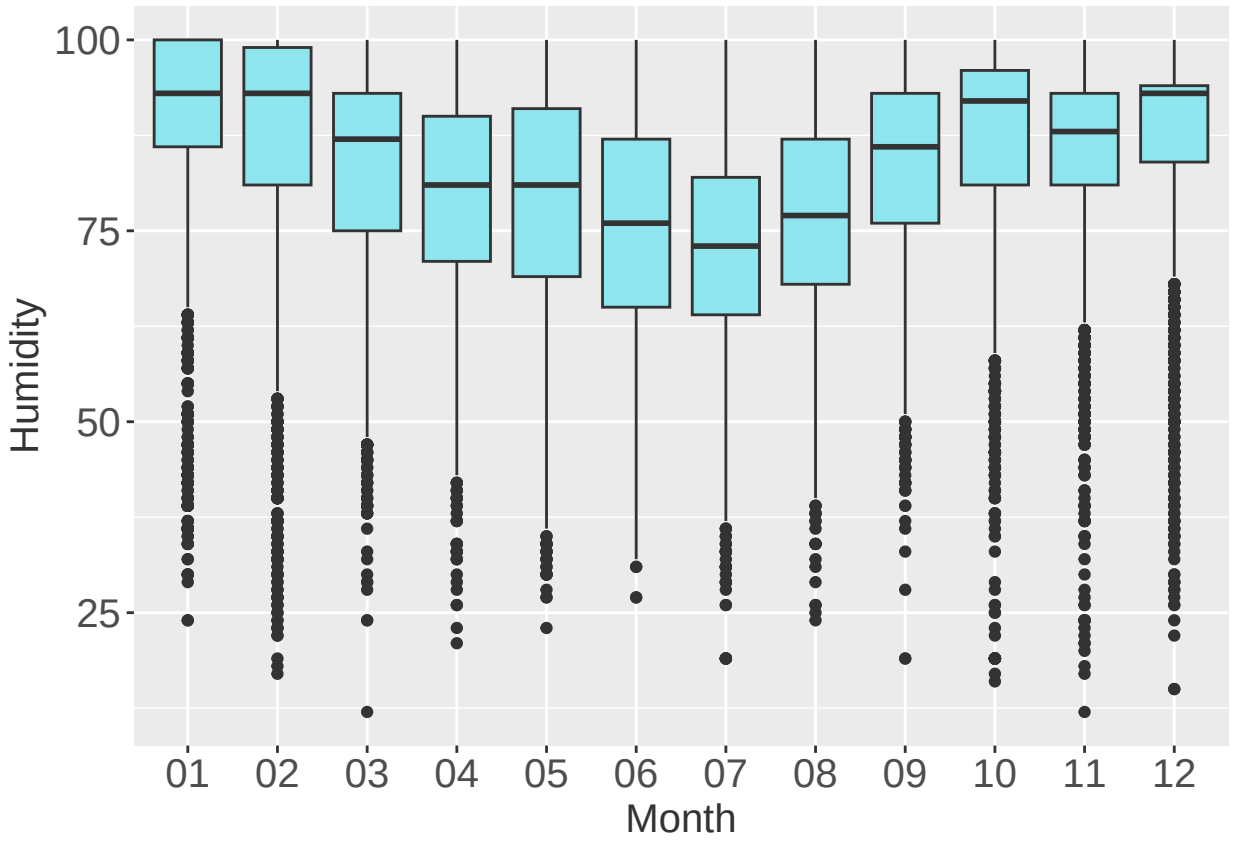
```
## Warning: Removed 1826 rows containing non-finite values (`stat_bin()`).
```



Changing boxplot defaults and adding aesthetics

```
df_hum <- read.csv("../data/historical-hourly-weather-data/humidity.csv")
df_hum$datetime <- as.character(df_hum$datetime)
df_hum$Month <- substr(df_hum$datetime, 6, 7)
ggplot(df_hum, aes(x=Month, y=Vancouver)) +
geom_boxplot(color="gray20", fill="cadetblue2") +
ylab("Humidity") +
theme(axis.text.x=element_text(size=15),
axis.text.y=element_text(size=15),
axis.title.x=element_text(size=15, color="gray20"),
axis.title.y=element_text(size=15, color="gray20"))
```

```
## Warning: Removed 1826 rows containing non-finite values (`stat_boxplot()`).
```


100
75
50
25
Humidity
01
02
03
04
05
06
07
08
09
10
11
12
Month