

# Class05.R

jaimebruce

2022-02-01

```
# Class 05 Data Visualization
```

```
# This is the "base" R plot
```

```
plot(cars)
```

```
# We are going to get a new plotting package called ggplot2
```

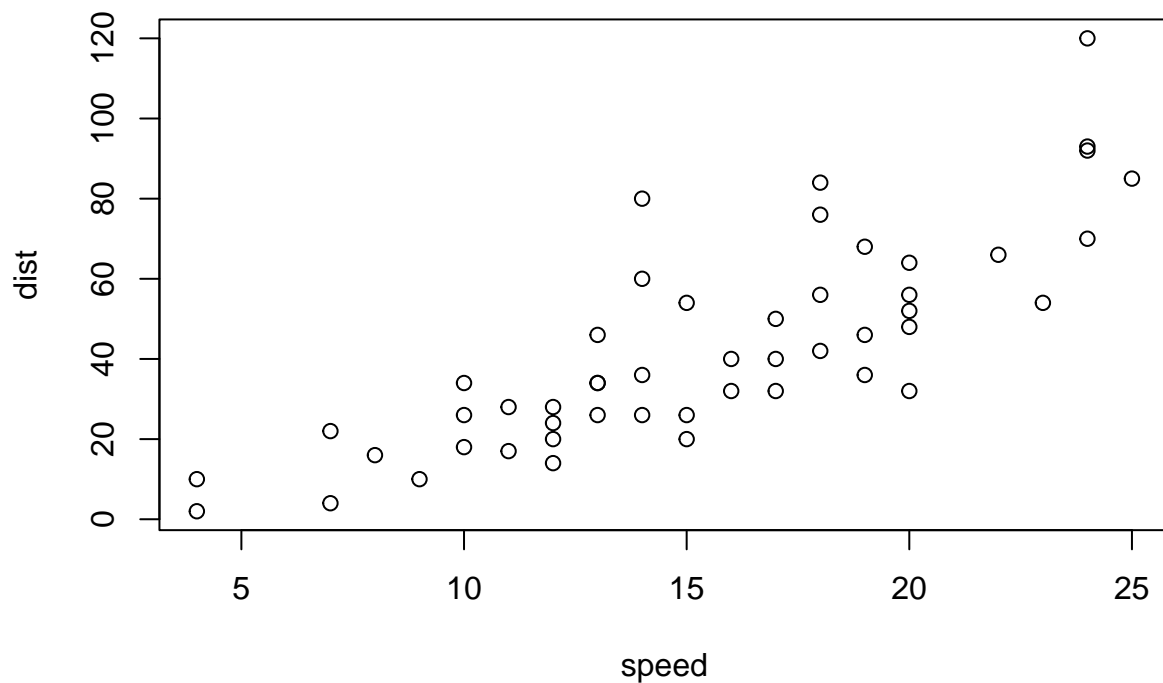
```
# ggplot(cars)
```

```
# install.packages("ggplot2")
```

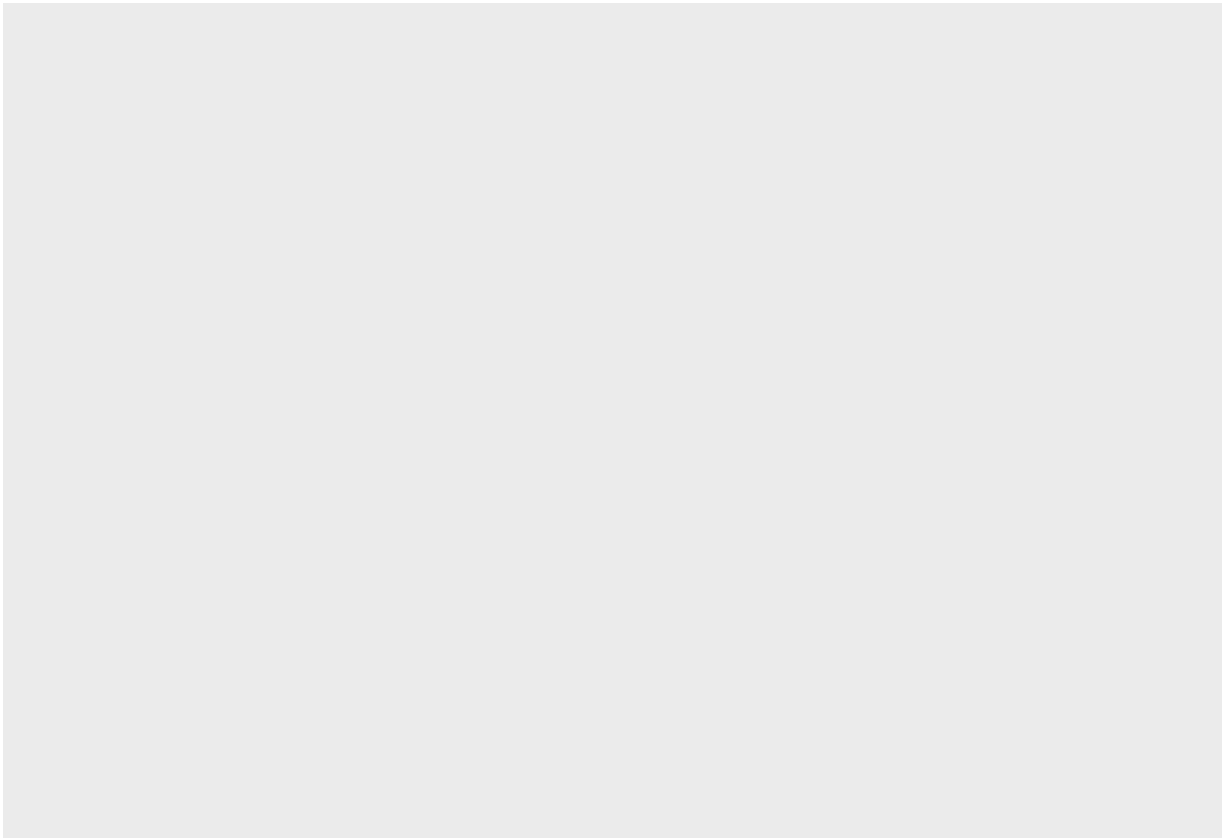
```
# Now we need to call/load the package
```

```
library(ggplot2)
```

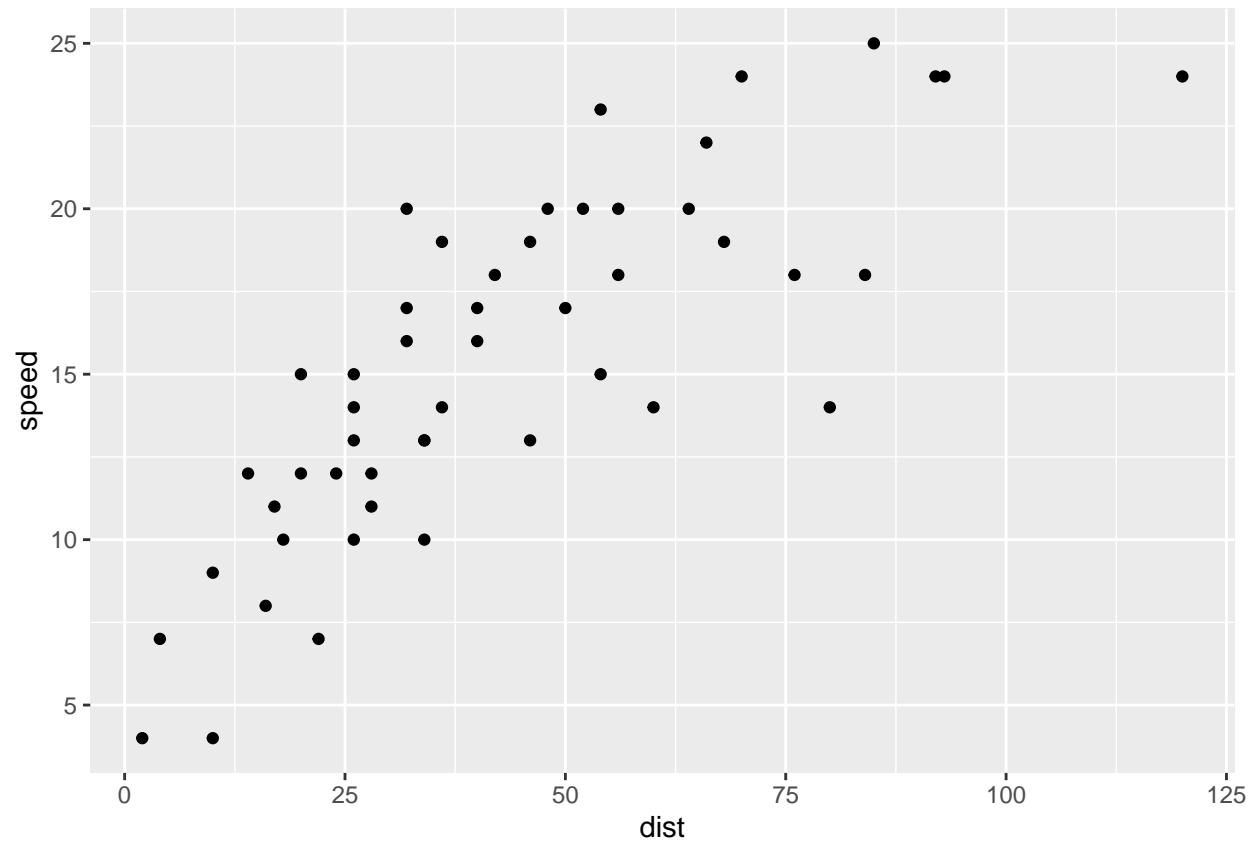
```
## Warning in register(): Can't find generic 'scale_type' in package ggplot2 to  
## register S3 method.
```



```
# This sets up the plot  
ggplot(cars)
```



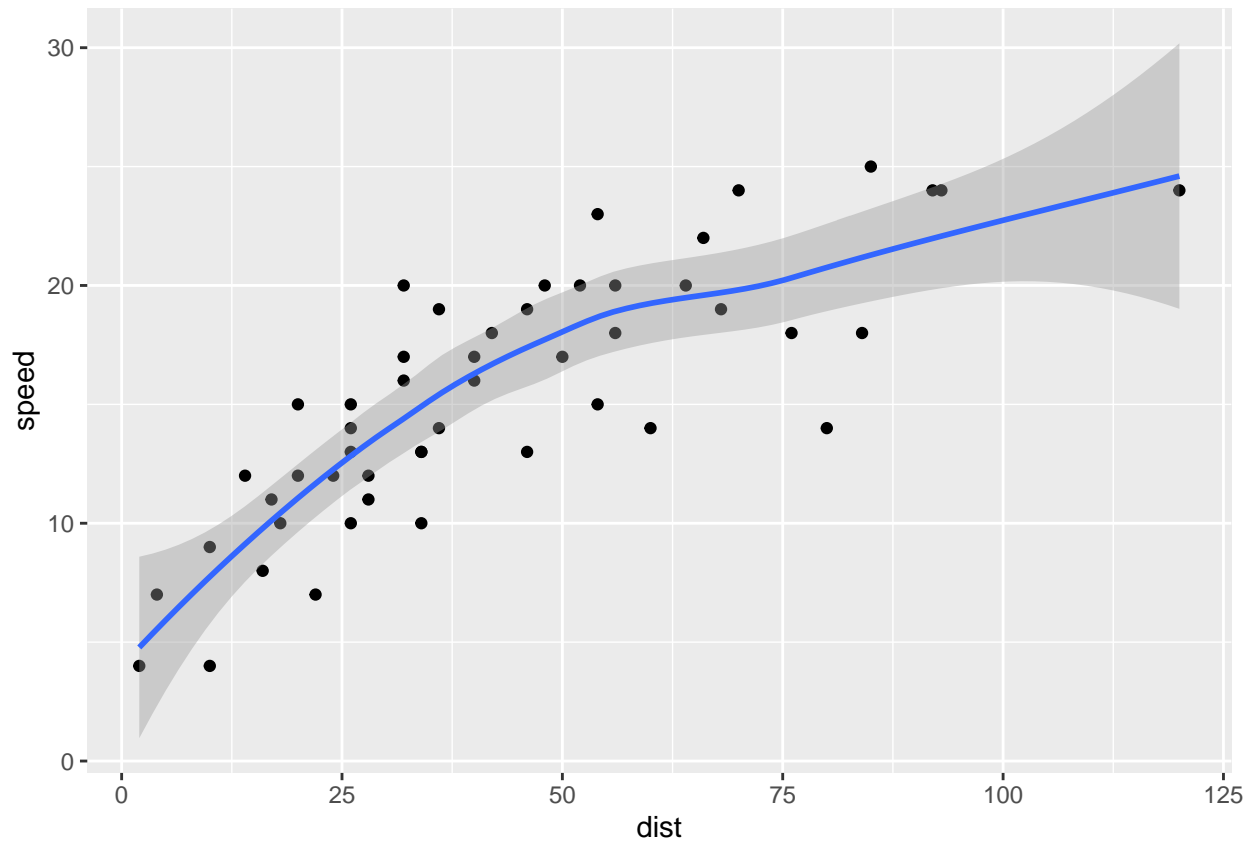
```
gg <- ggplot(cars) + aes(x=dist, y=speed) + geom_point()  
gg
```



```
# One las thing. Let's add a line to the data
```

```
gg + geom_smooth()
```

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

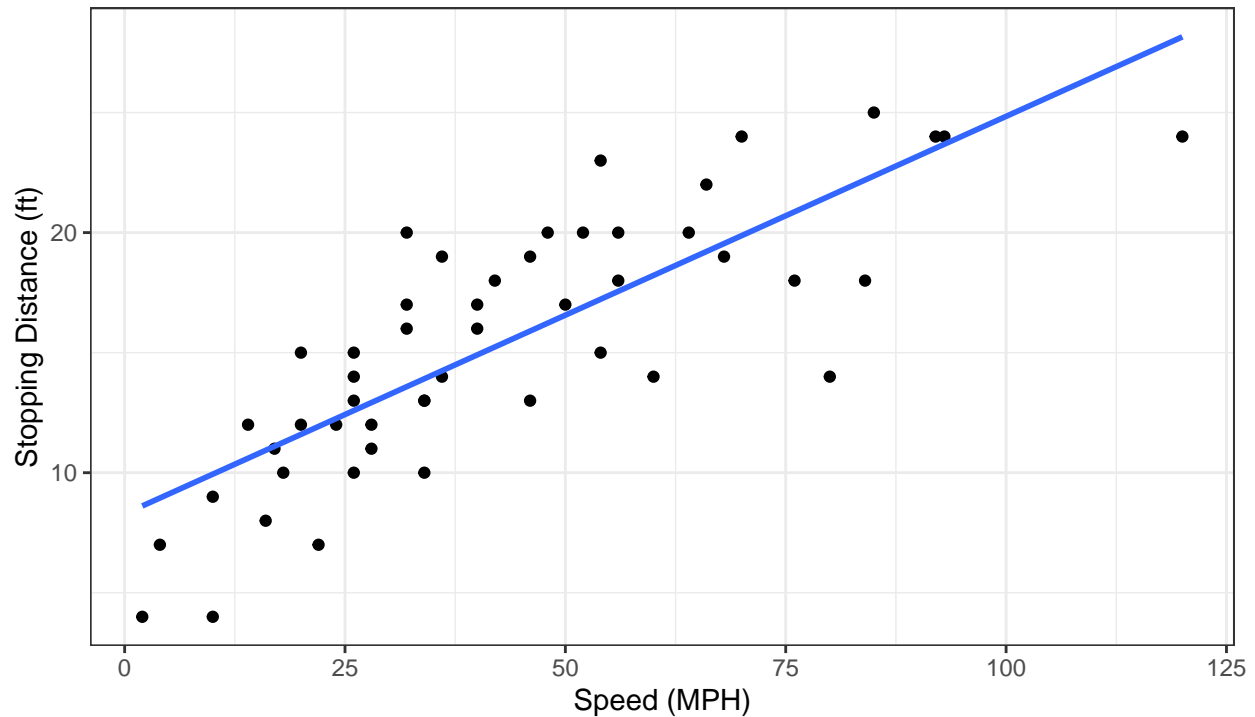


```
# I want a linear model
gg + labs(title="Speed and Stopping Distances of Cars",
  x="Speed (MPH)",
  y="Stopping Distance (ft)",
  subtitle = "Car stopping distance",
  caption="Dataset: 'cars'") +
  geom_smooth(method="lm", se=FALSE) + theme_bw()
```

```
## 'geom_smooth()' using formula 'y ~ x'
```

## Speed and Stopping Distances of Cars

Car stopping distance



Dataset: 'cars'

```
# RNASeq experiment dataset
```

```
# Read the data into R
```

```
url <- "https://bioboot.github.io/bimm143_S20/class-material/up_down_expression.txt"
```

```
genes <- read.delim(url)
```

```
head(genes)
```

```
##      Gene Condition1 Condition2      State
## 1   A4GNT -3.6808610 -3.4401355  unchanging
## 2   AAAS  4.5479580  4.3864126  unchanging
## 3  AASDH  3.7190695  3.4787276  unchanging
## 4  AATF  5.0784720  5.0151916  unchanging
## 5  AATK  0.4711421  0.5598642  unchanging
## 6 AB015752.4 -3.6808610 -3.5921390 unchanging
```

```
# > Q. How many genes in this dataset?
```

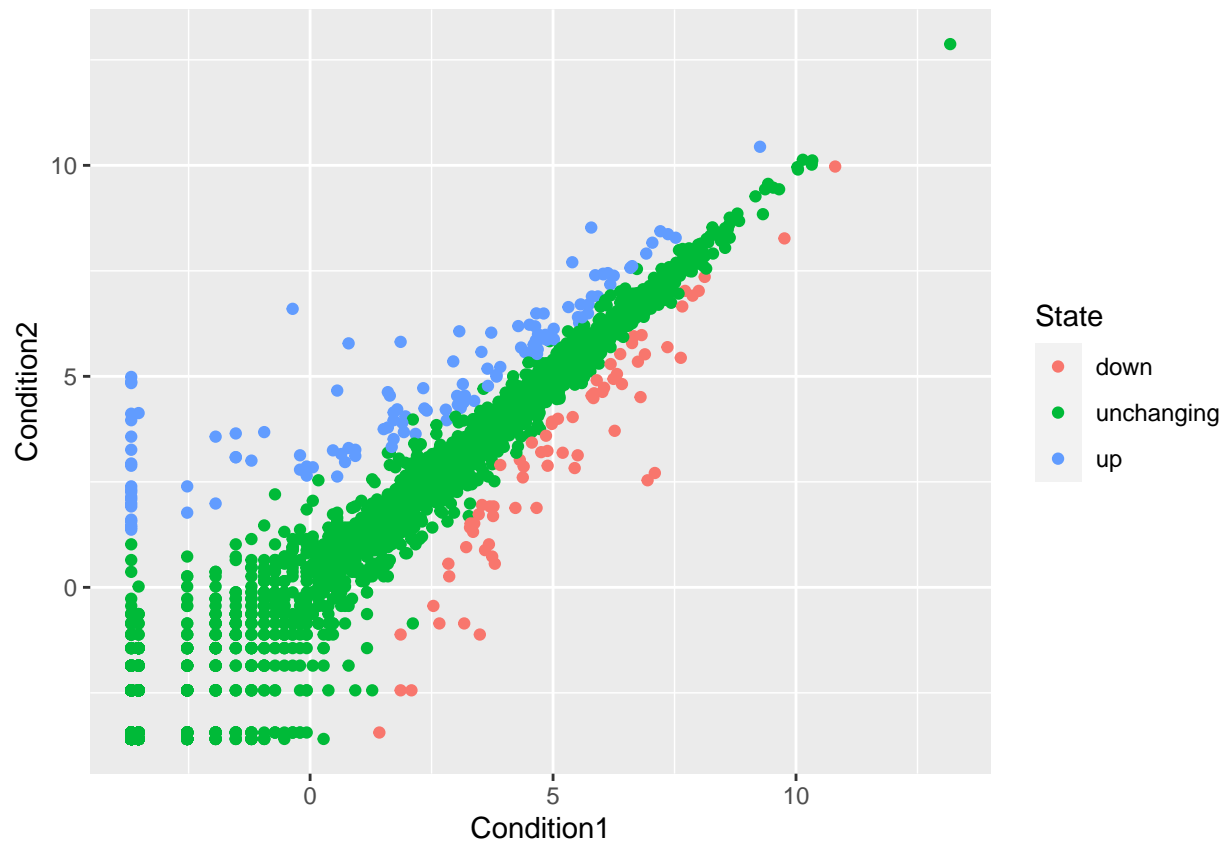
```
nrow(genes)
```

```
## [1] 5196
```

```
colnames(genes)
```

```
## [1] "Gene"      "Condition1" "Condition2" "State"
```

```
# I want to plot this result
ggplot(genes) +
  aes(x=Condition1, y=Condition2, col=State) +
  geom_point()
```



```
# Q. How many genes are "up" regulated in this
table( genes$State )
```

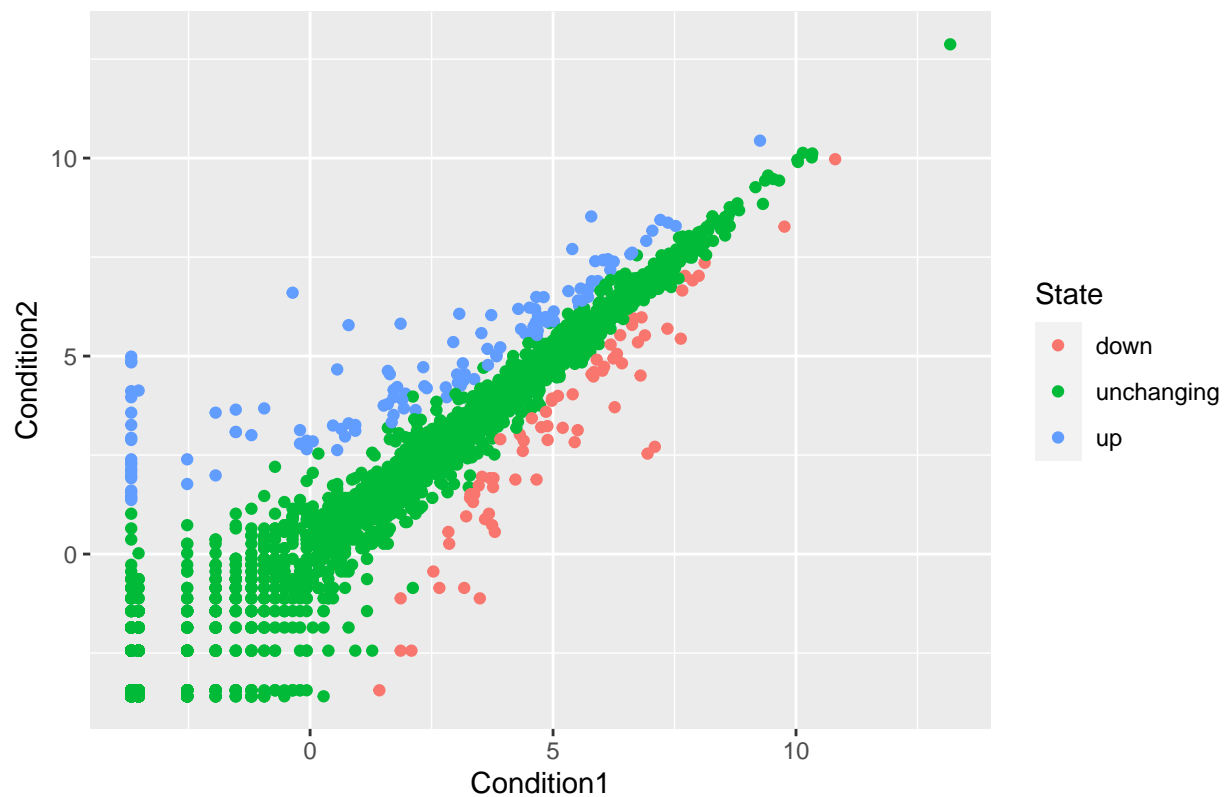
```
##
##      down  unchanging      up
##      72      4997      127
```

```
# Q. Using your values above and 2 significant figures. What fraction of total genes is up-regulated in
round( table(genes$State)/nrow(genes) * 100, 2 )
```

```
##
##      down  unchanging      up
##      1.39      96.17      2.44
```

```
# Add some color to denote up/down values
ggplot(genes) +
  aes(x=Condition1, y=Condition2, col=State) +
  geom_point() +
  labs(title="Some gene expression data")
```

Some gene expression data



```
p <- ggplot(genes) +
  aes(x=Condition1, y=Condition2, col=State) +
  geom_point()
p + scale_colour_manual(values=c("blue","gray","red")) +
  labs(title="Gene Expression Changes Upon Drug Treatment",
       x="Control (no drug) ",
       y="Drug Treatment")
```

Gene Expression Changes Upon Drug Treatment

