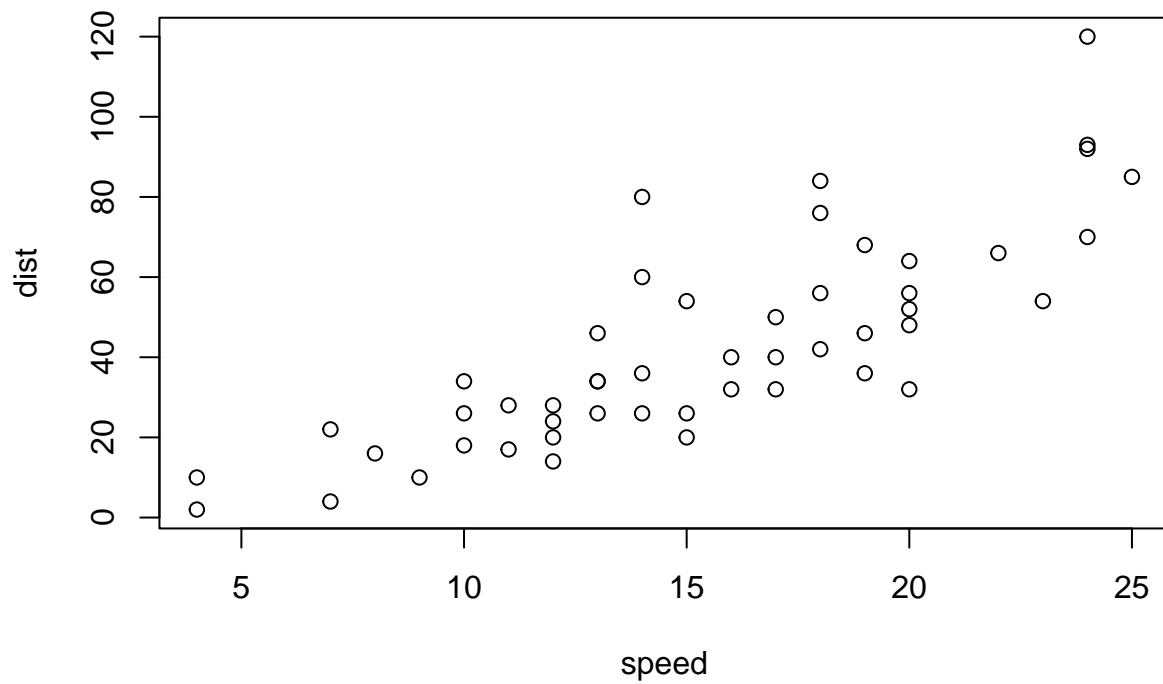


week05.R

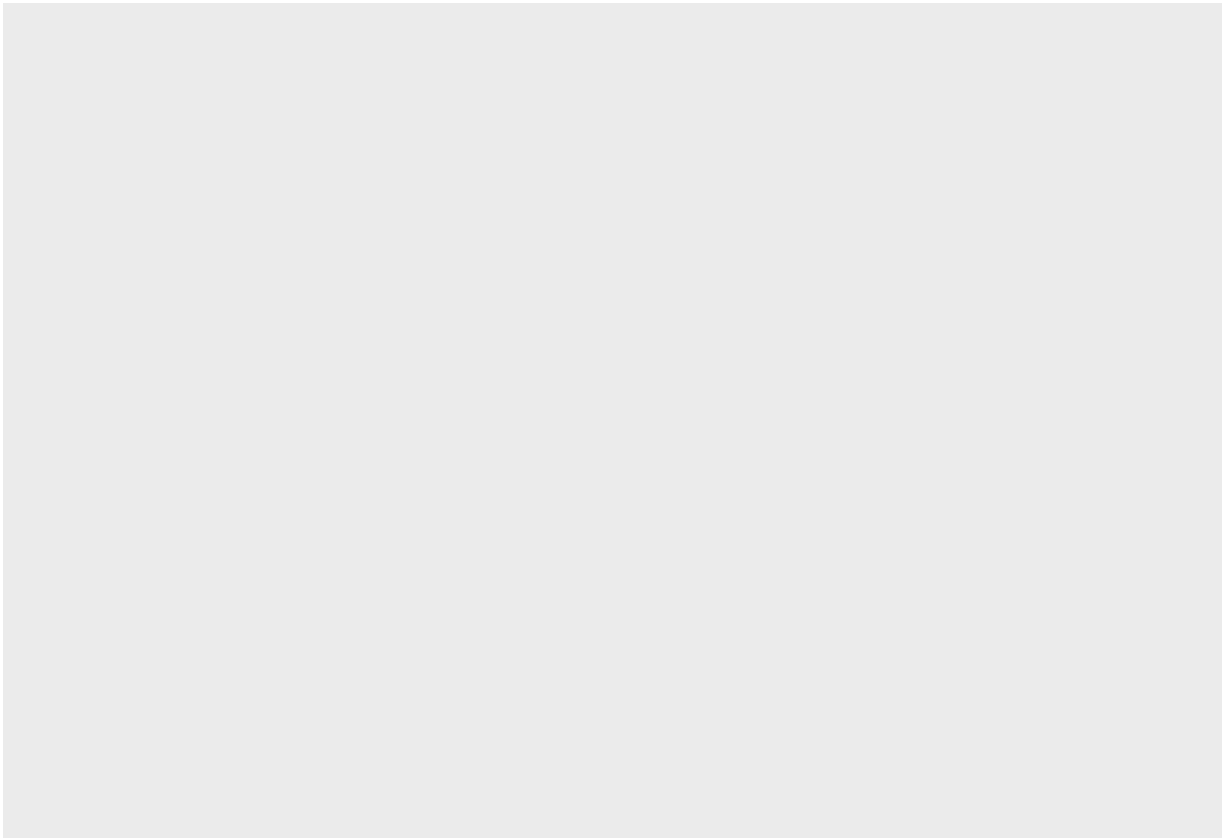
LaneyP

2022-02-01

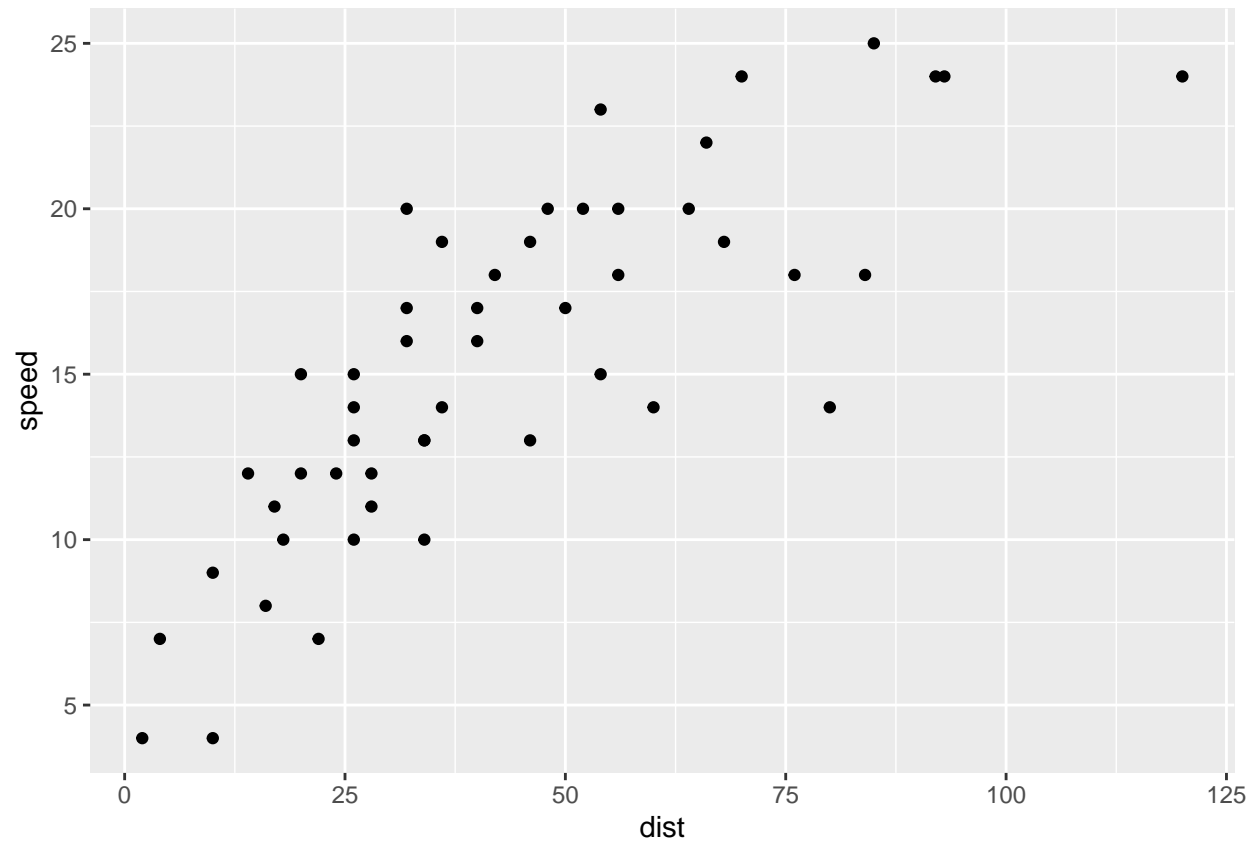
```
# Class 05 Data Visualization  
  
# This is the "base" R plot  
plot(cars)  
  
# We are going to get a new plotting package called ggplot2  
# ggplot(cars)  
  
# install.packages("ggplot2")  
# Now we need to call/load the package  
library(ggplot2)
```



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



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

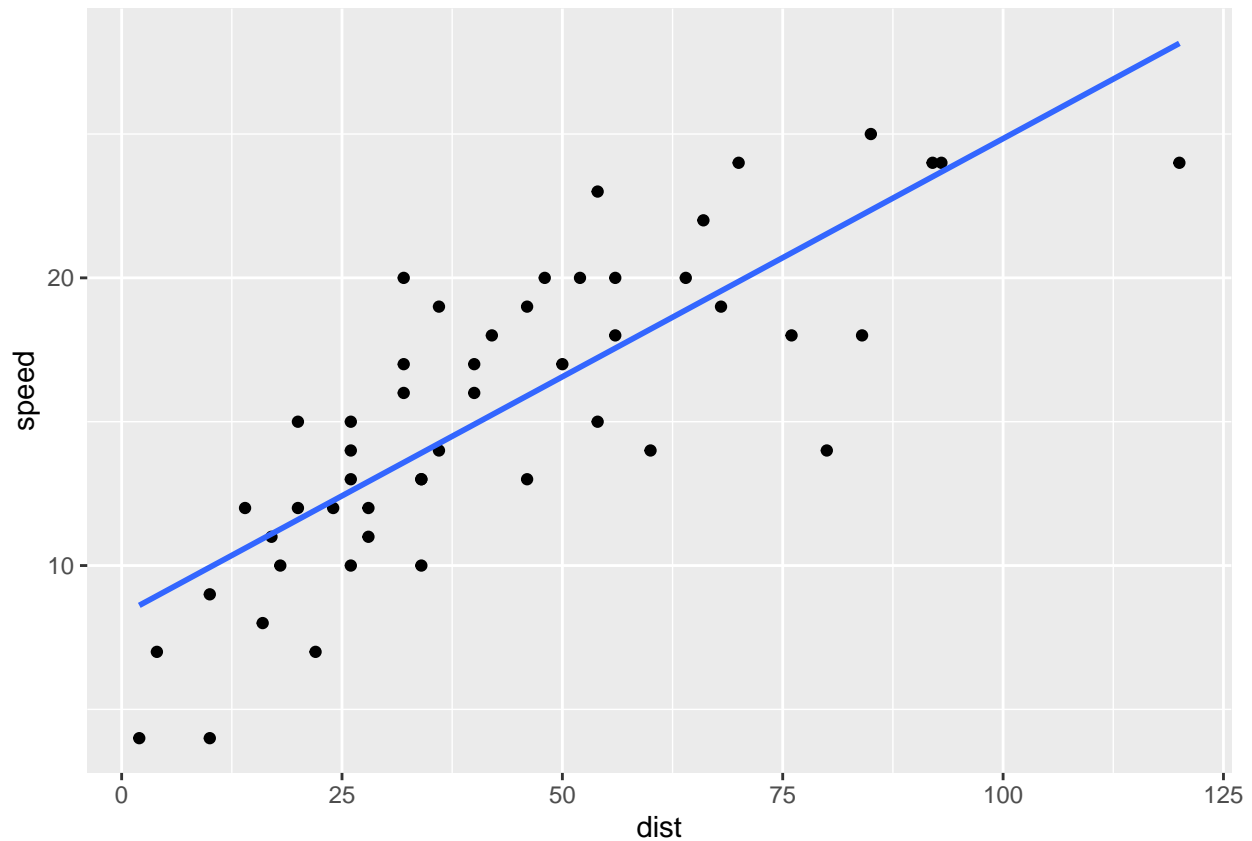


```
# One last thing. Let's add a line to the data
```

```
# I want a linear model
```

```
gg + geom_smooth(method="lm", se=FALSE)
```

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

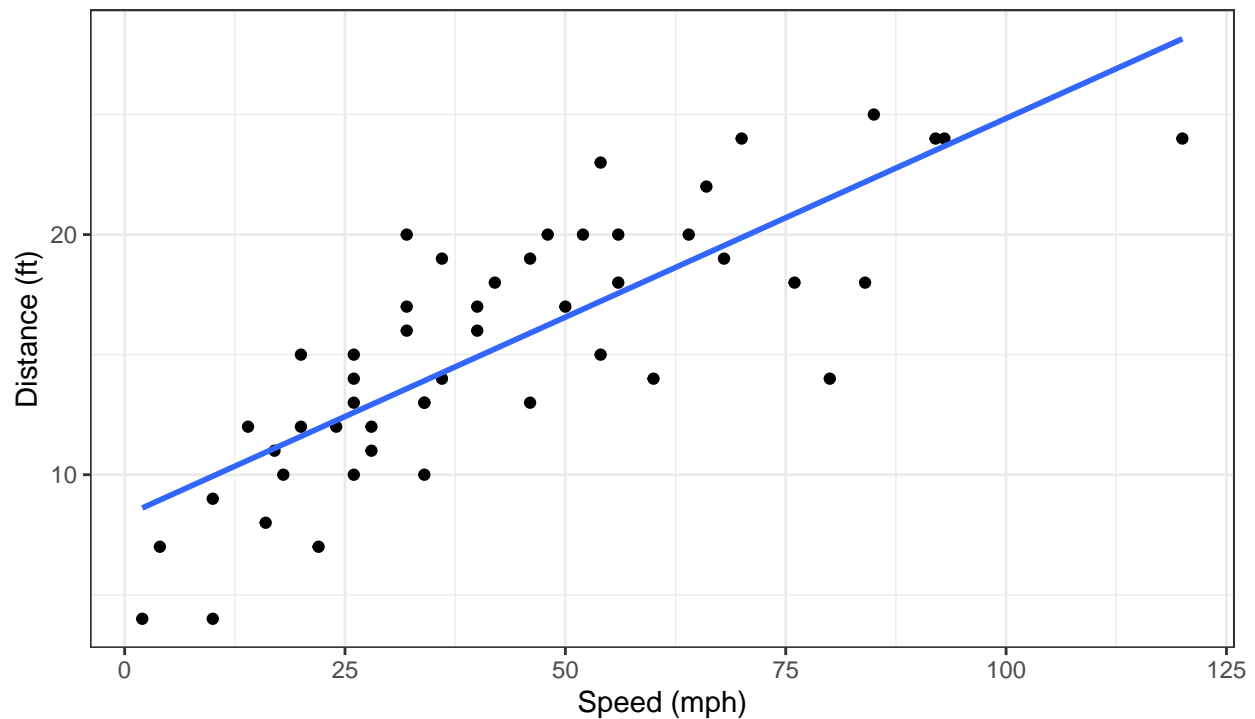


```
# Add labs() and theme_bw()
gg + labs(title="Speed and Stopping Distances of Cars",
          x="Speed (mph)", y="Distance (ft)",
          subtitle = "The stopping distance of cars increase as speed increases",
          caption="Dataset: 'cars'") +
  geom_smooth(method="lm", se=FALSE) +
  theme_bw()
```

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

Speed and Stopping Distances of Cars

The stopping distance of cars increase as speed increases



Dataset: 'cars'

```
# 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
```

```
# How many genes are in data set
nrow(genes)
```

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

```
# How many columns are there
colnames(genes)
```

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

```
ncol(genes)
```

```
## [1] 4
```

```
# How many "up" regulated genes are there
```

```
table(genes$State)
```

```
##
```

```
##      down  unchanging      up
```

```
##      72      4997      127
```

```
# What fraction of total genes is up-regulated in this data set
```

```
round( table(genes$State)/nrow(genes) * 100, 2 )
```

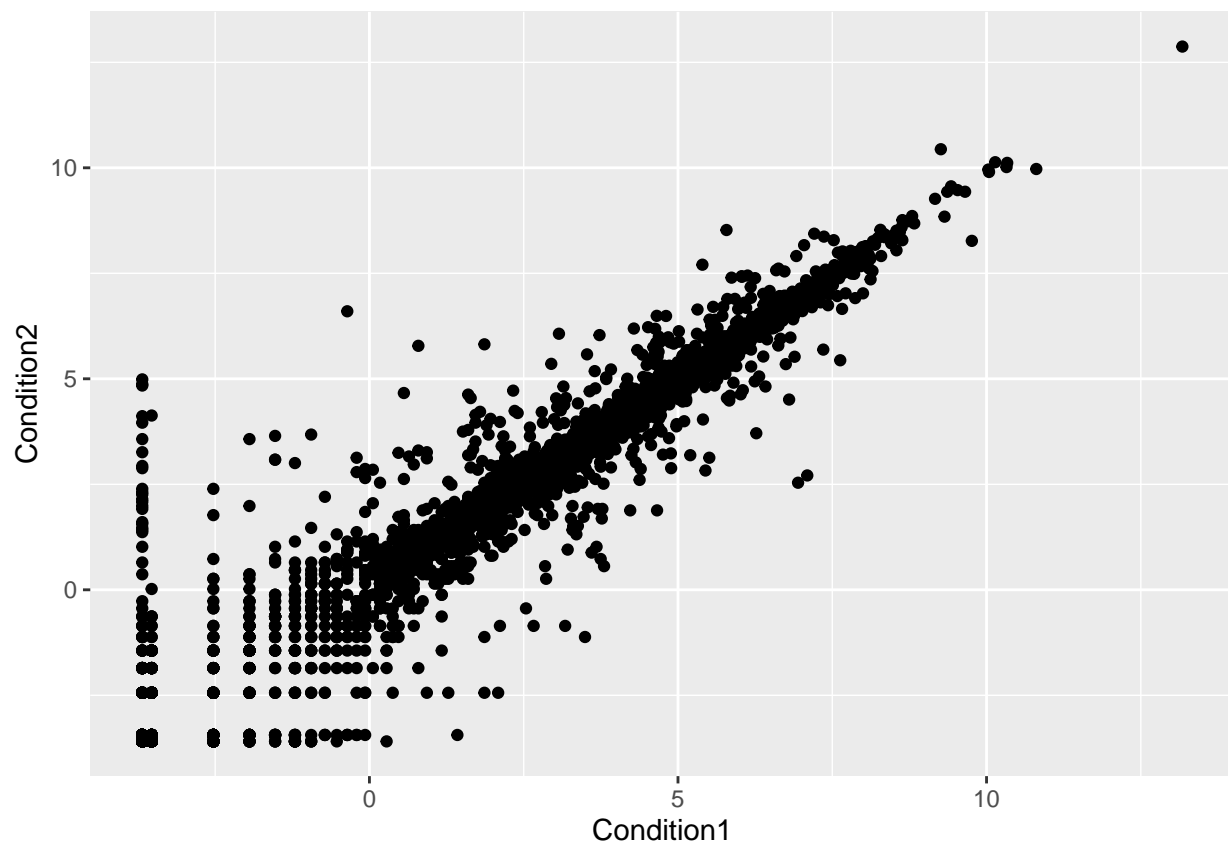
```
##
```

```
##      down  unchanging      up
```

```
##      1.39      96.17      2.44
```

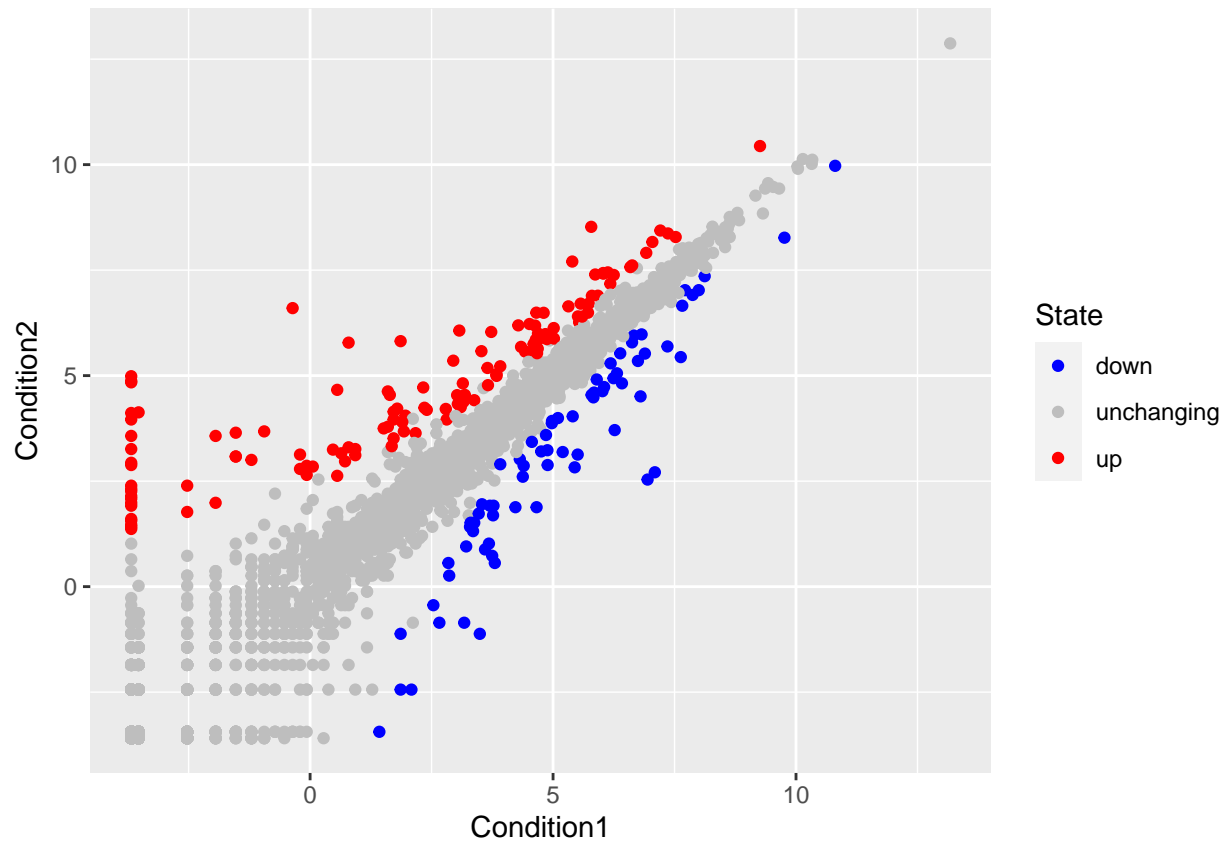
```
# I want to plot this result
```

```
ggplot(genes) +  
  aes(x=Condition1, y=Condition2) +  
  geom_point()
```



```
# Add some color to denote up/down values
p <- ggplot(genes) +
  aes(x=Condition1, y=Condition2, color=State) +
  geom_point()

# Add better color
p + scale_colour_manual( values=c("blue","gray","red") )
```



```
# Adding plot annotations
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

