

Data visualization with ggplot2

AUTHOR

Rogelio Castro

##Welcome

```
library(ggplot2)

# Basic plot
p1 <- ggplot(cars) +
  aes(x = speed, y = dist) +
  geom_point() +
  geom_smooth()

# Plot with custom labels
p2 <- ggplot(cars) +
  aes(x = speed, y = dist) +
  geom_point() +
  labs(
    title = "Speed and Stopping Distances of Cars",
    x = "Speed (MPH)",
    y = "Stopping Distance (ft)",
    subtitle = "Your informative subtitle text here",
    caption = "Dataset: 'cars'"
  ) +
  geom_smooth(method = "lm", se = FALSE) +
  theme_bw()

# Read data
url <- "https://bioboot.github.io/bimm143_S20/class-material/up_down_expression.txt"
genes <- read.delim(url)

# Summary statistics
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

```
ncol(genes)
```

```
[1] 4
```

```
nrow(genes)
```

```
[1] 5196
```

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

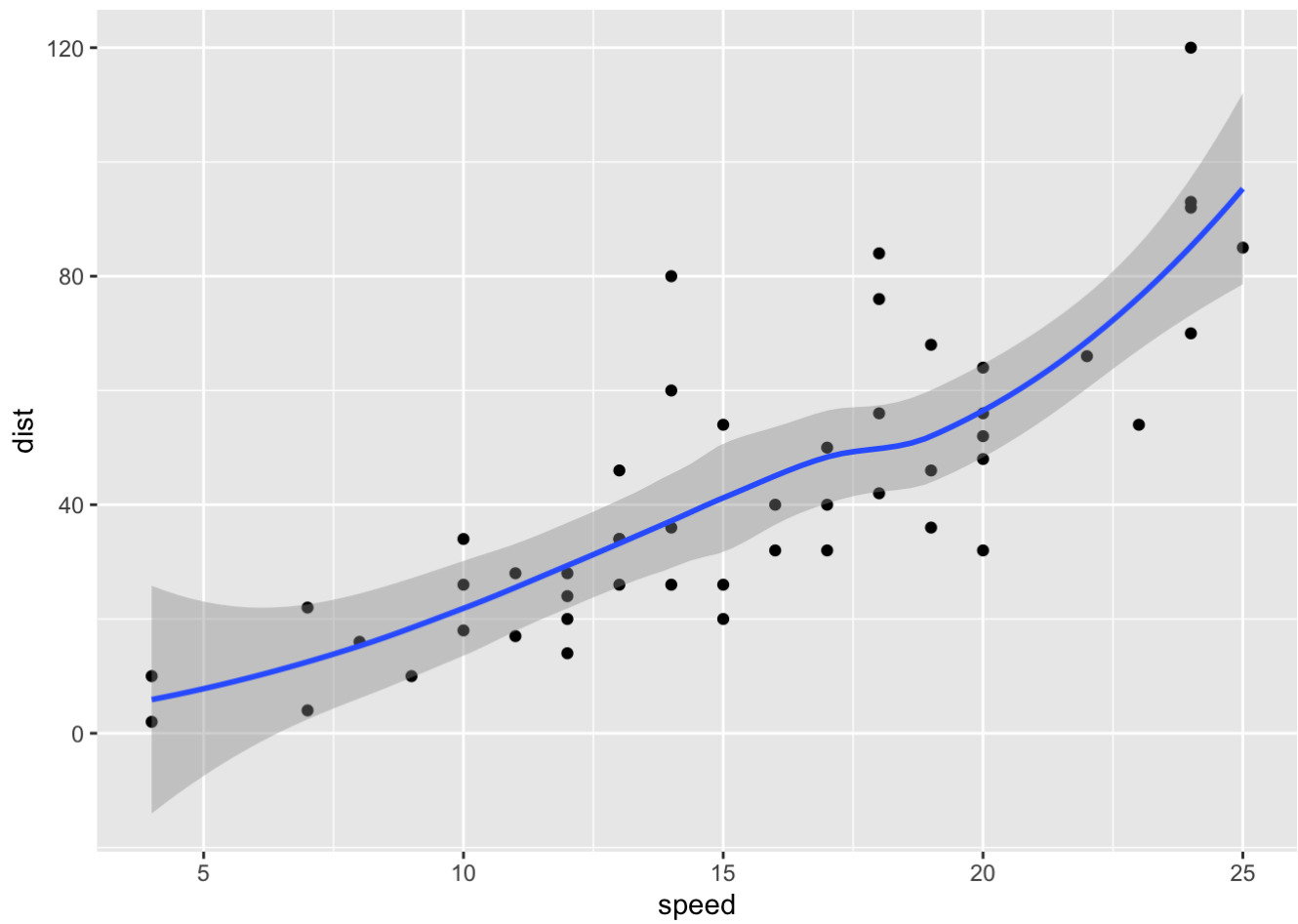
down	unchanging	up
1.39	96.17	2.44

```
# Basic gene expression plot
p3 <- ggplot(genes) +
  aes(x = Condition1, y = Condition2, col = State) +
  geom_point()

# Gene expression plot with custom colors and labels
p4 <- ggplot(genes) +
  aes(x = Condition1, y = Condition2, col = State) +
  geom_point() +
  scale_colour_manual(values = c("blue", "gray", "red")) +
  labs(
    title = "Gene Expression Changes Upon Drug Treatment",
    x = "Control (no drug)",
    y = "Drug Treatment"
  )

# Display plots
p1
```

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

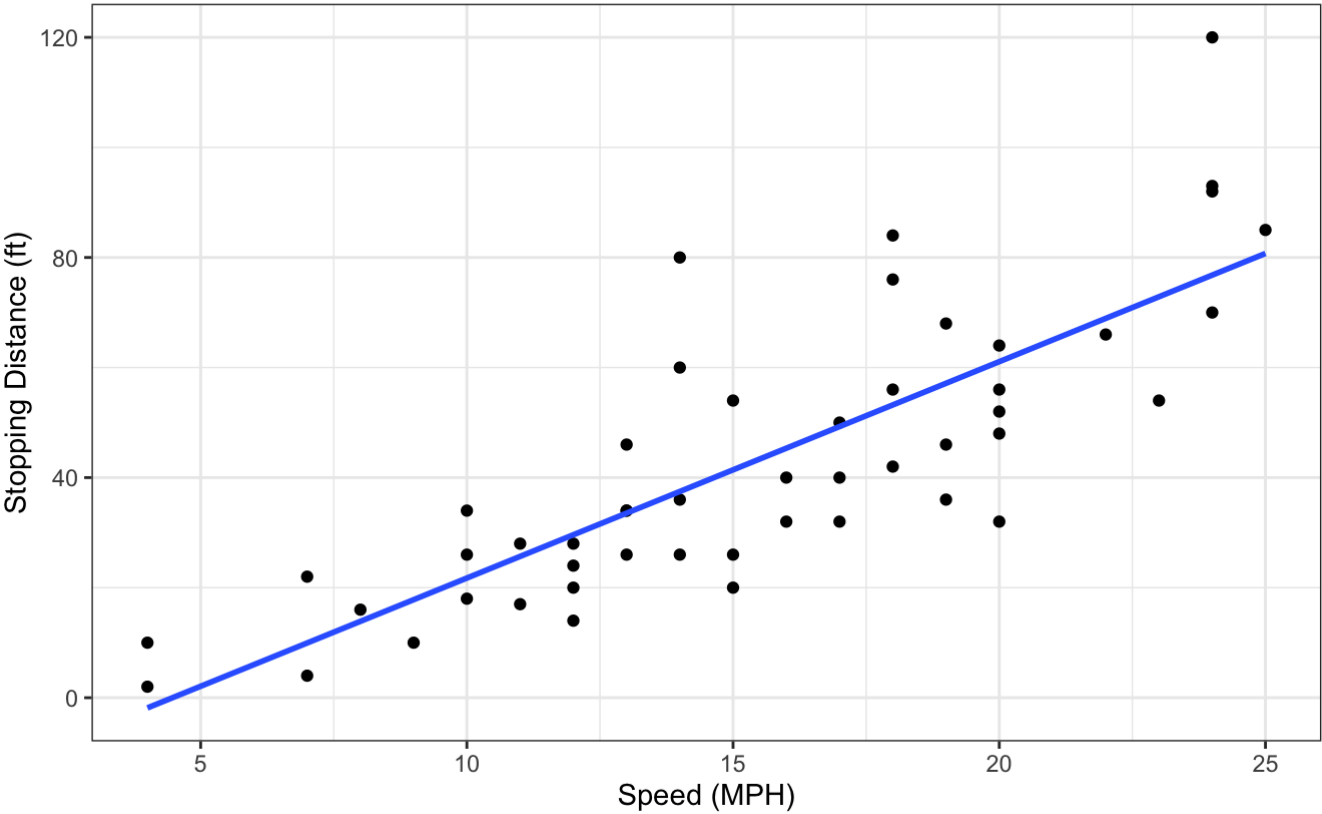


p2

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

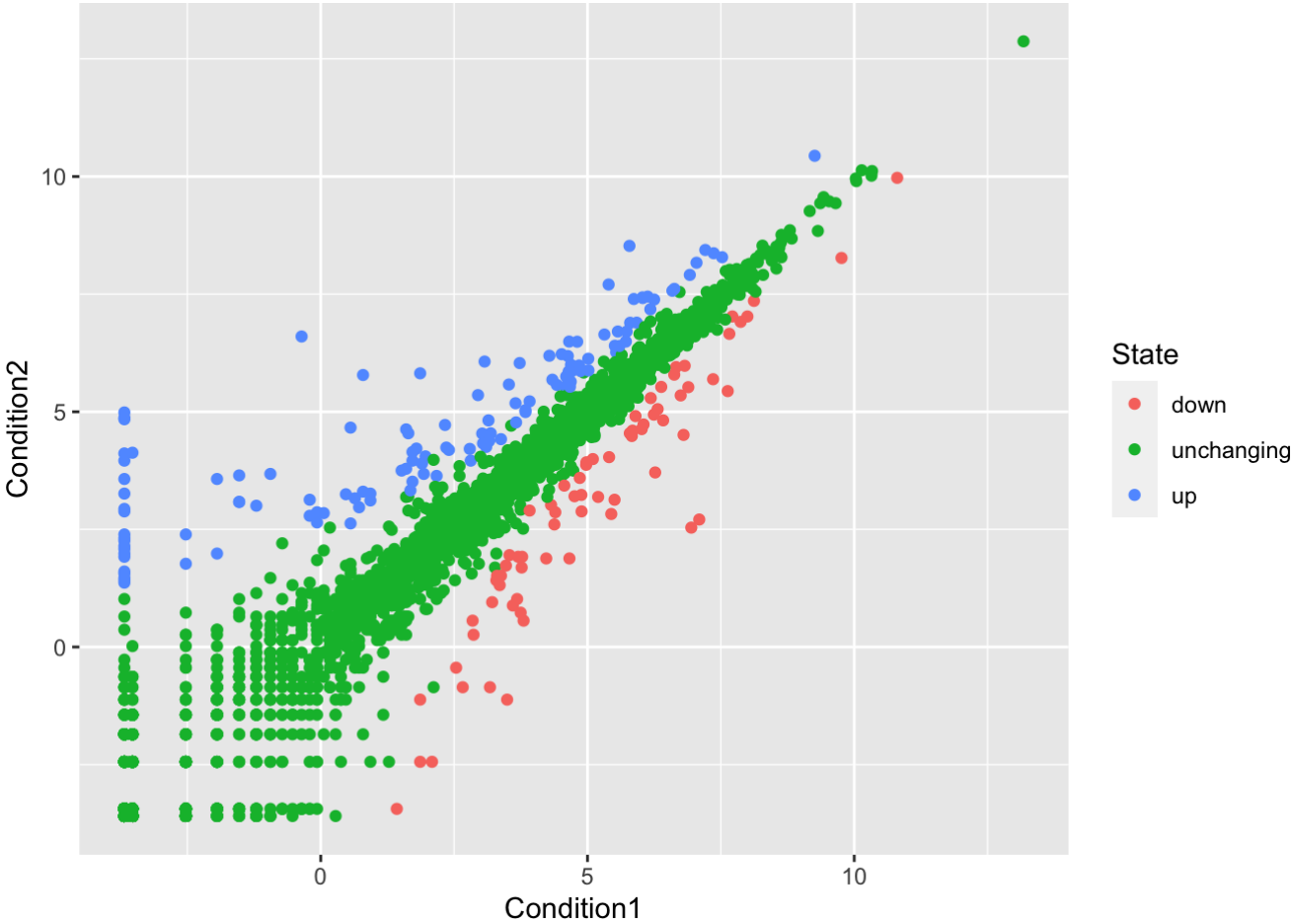
Speed and Stopping Distances of Cars

Your informative subtitle text here



Dataset: 'cars'

p3



p4

Gene Expression Changes Upon Drug Treatment

