class05.R

ziyuanhan

2021-10-13

```r
#Class 05: Data visualization

#load the package
library(ggplot2)
ggplot(cars)
```

```r
head(cars)
```

```
##   speed dist
## 1     4    2
## 2     4   10
## 3     7    4
## 4     7   22
```
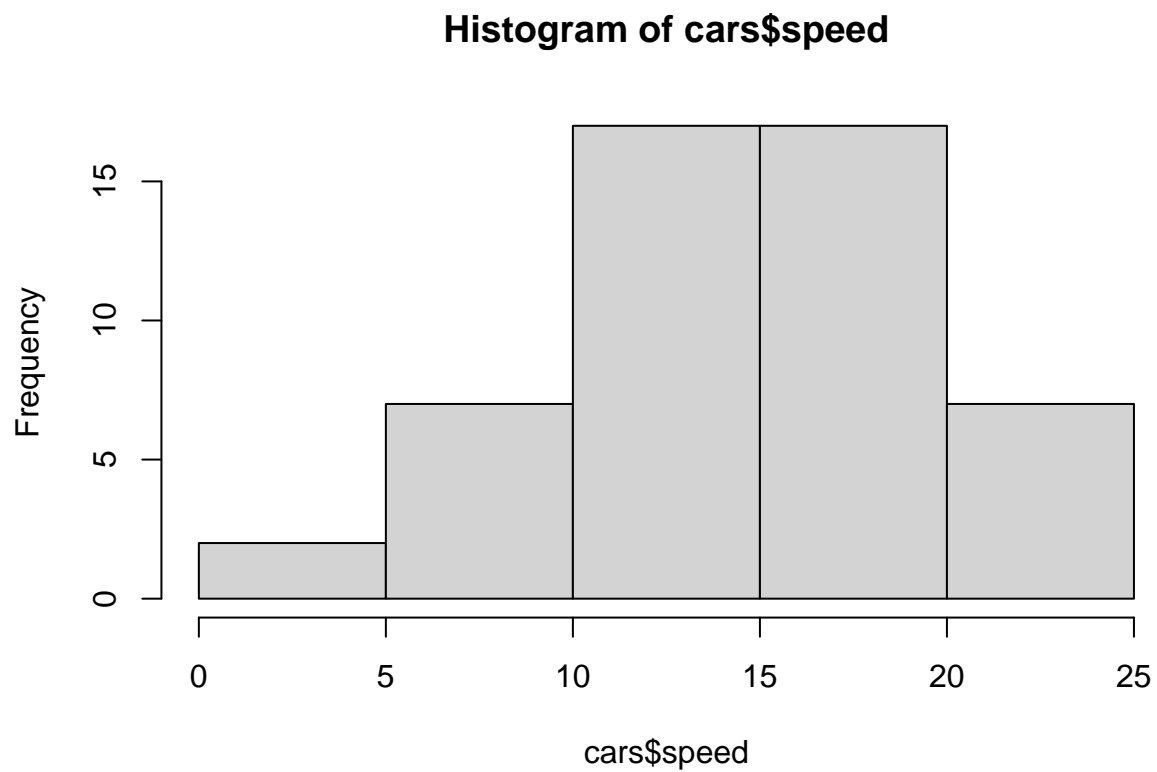
```
## 5      8   16
## 6      9   10
```
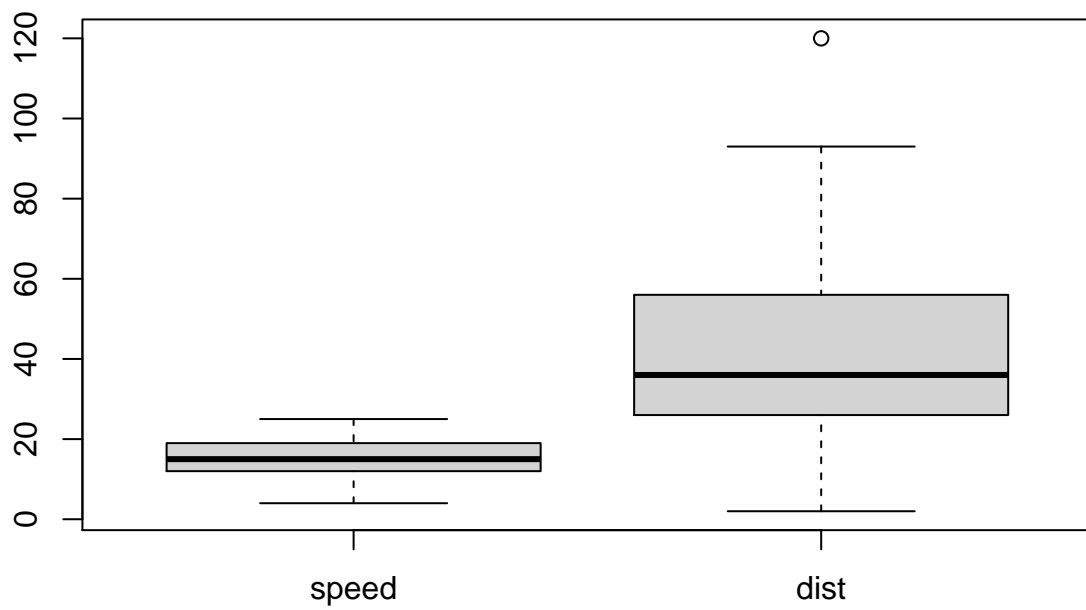
```
summary(cars)
```

```
##      speed          dist
##  Min.   : 4.0   Min.   :  2.00
##  1st Qu.:12.0   1st Qu.: 26.00
##  Median :15.0   Median : 36.00
##  Mean   :15.4   Mean   : 42.98
##  3rd Qu.:19.0   3rd Qu.: 56.00
##  Max.   :25.0   Max.   :120.00
```
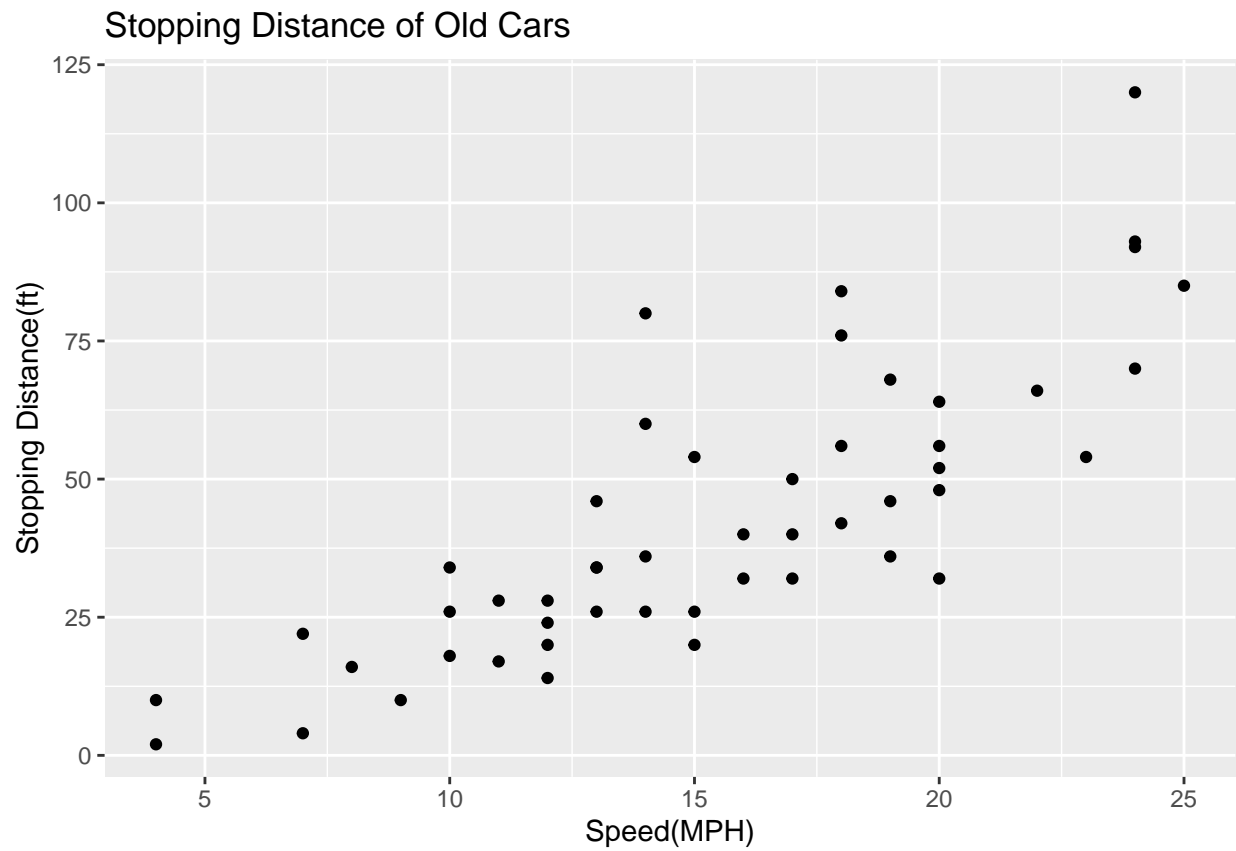
```
hist(cars$speed)
```
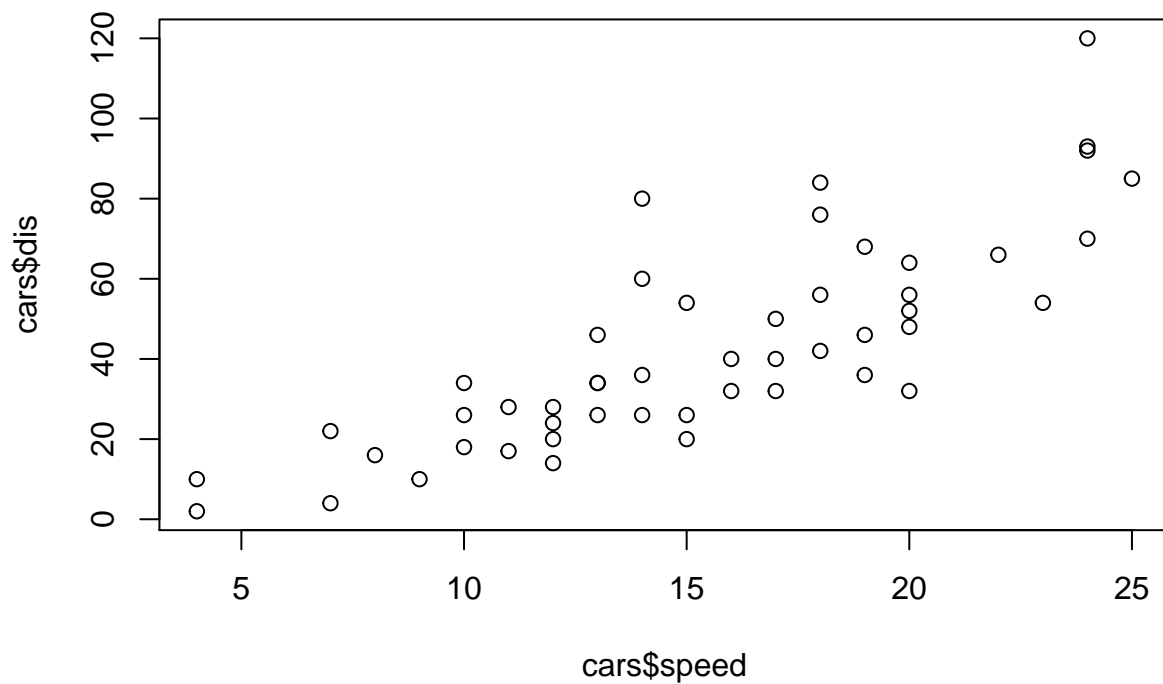
Histogram of cars$speed

```
boxplot(cars)
```

```
ggplot(data=cars) +
  aes(x=speed, y=dist)+
  geom_point() +
  labs(title="Stopping Distance of Old Cars",
       x="Speed(MPH)",y="Stopping Distance(ft)")
```

Stopping Distance of Old Cars


```
#side-note:ggplot is not the only graphics system
plot(cars$speed,cars$dis)
```

```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
```

```r
nrow(genes)
```

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

```r
colnames(genes)
```

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

```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
```

```r
#How many genes
nrow(genes)
```

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

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

```
##
##        down unchanging         up
##          72       4997        127
```
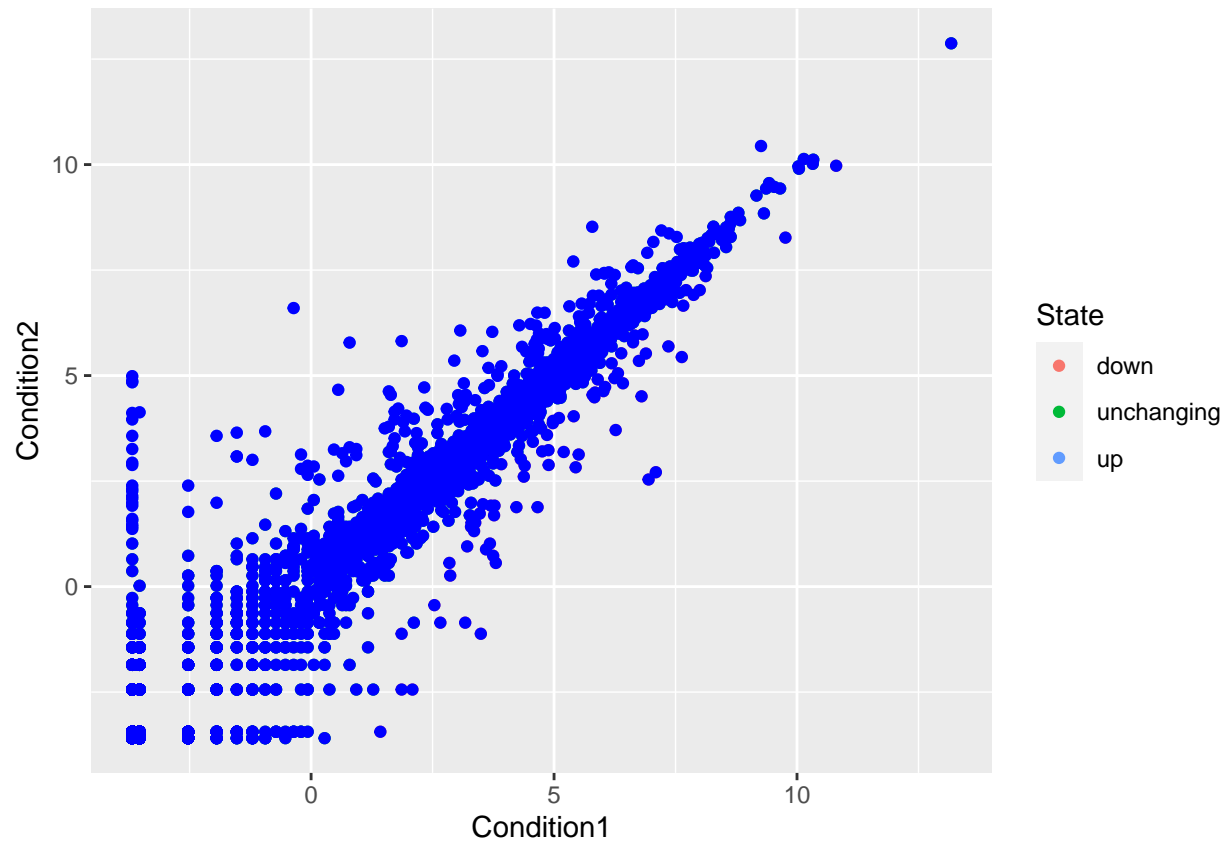
```r
round(table(genes$State)/nrow(genes) * 100,3)
```
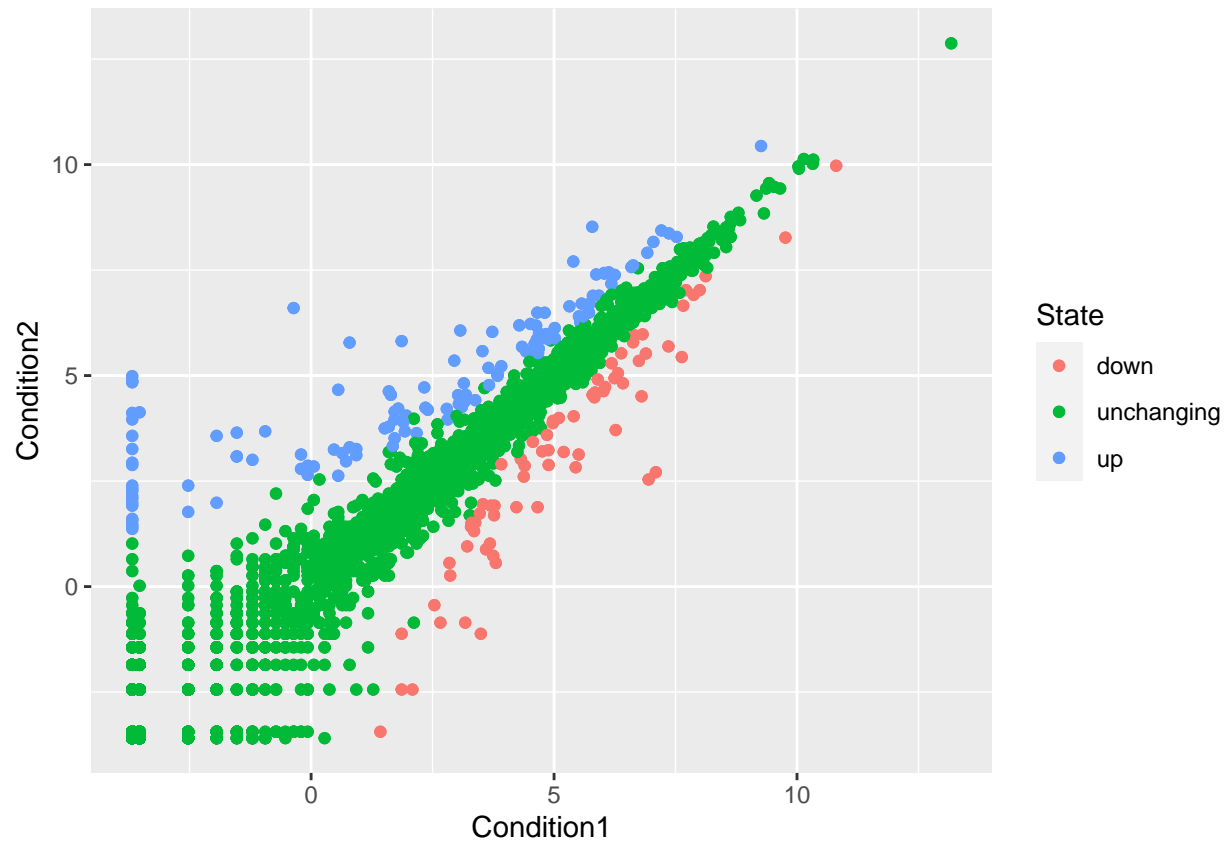
```
##
##        down unchanging         up
##       1.386     96.170      2.444
```

```r
#let's make a figure
p<-ggplot(genes)+
  aes(x=Condition1,y=Condition2,col=State)+
  geom_point()

p+ geom_point(col="blue")
```
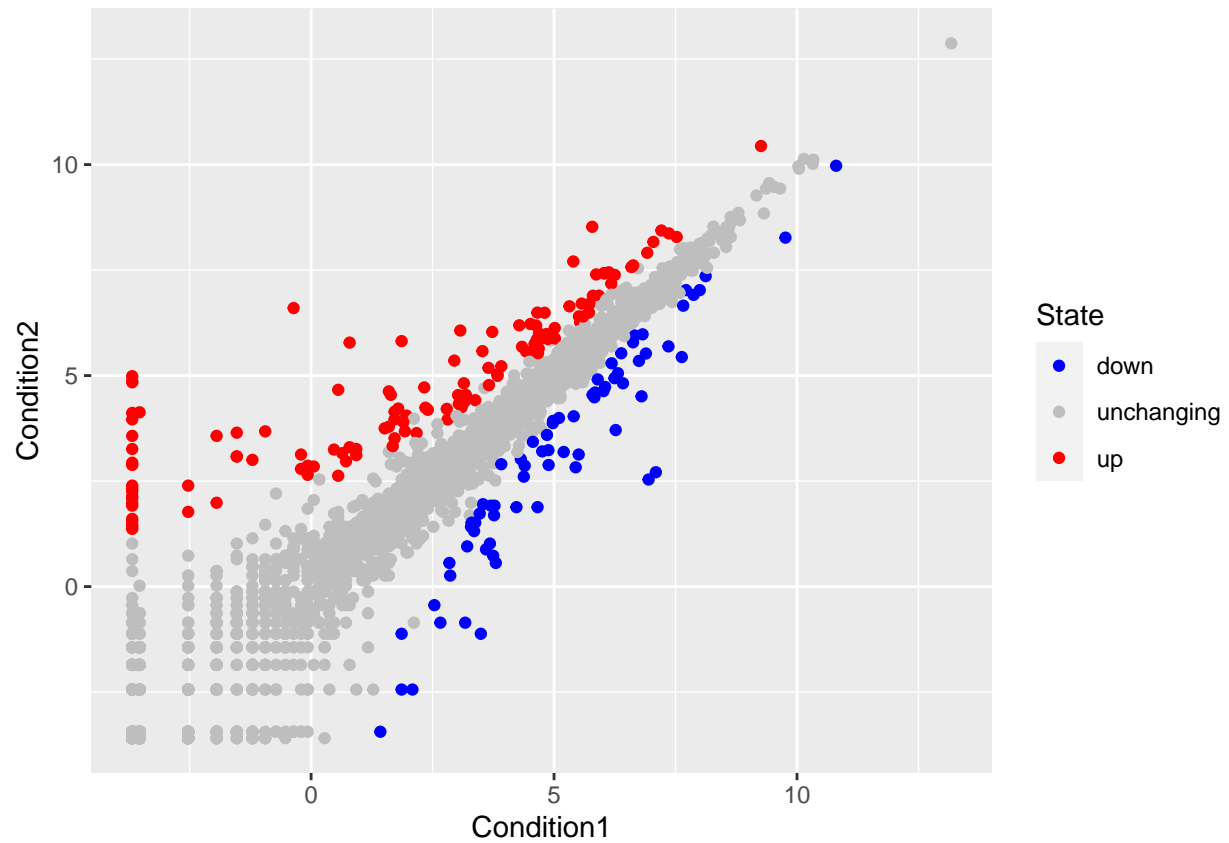
```
p+aes(col=State) + geom_point()
```
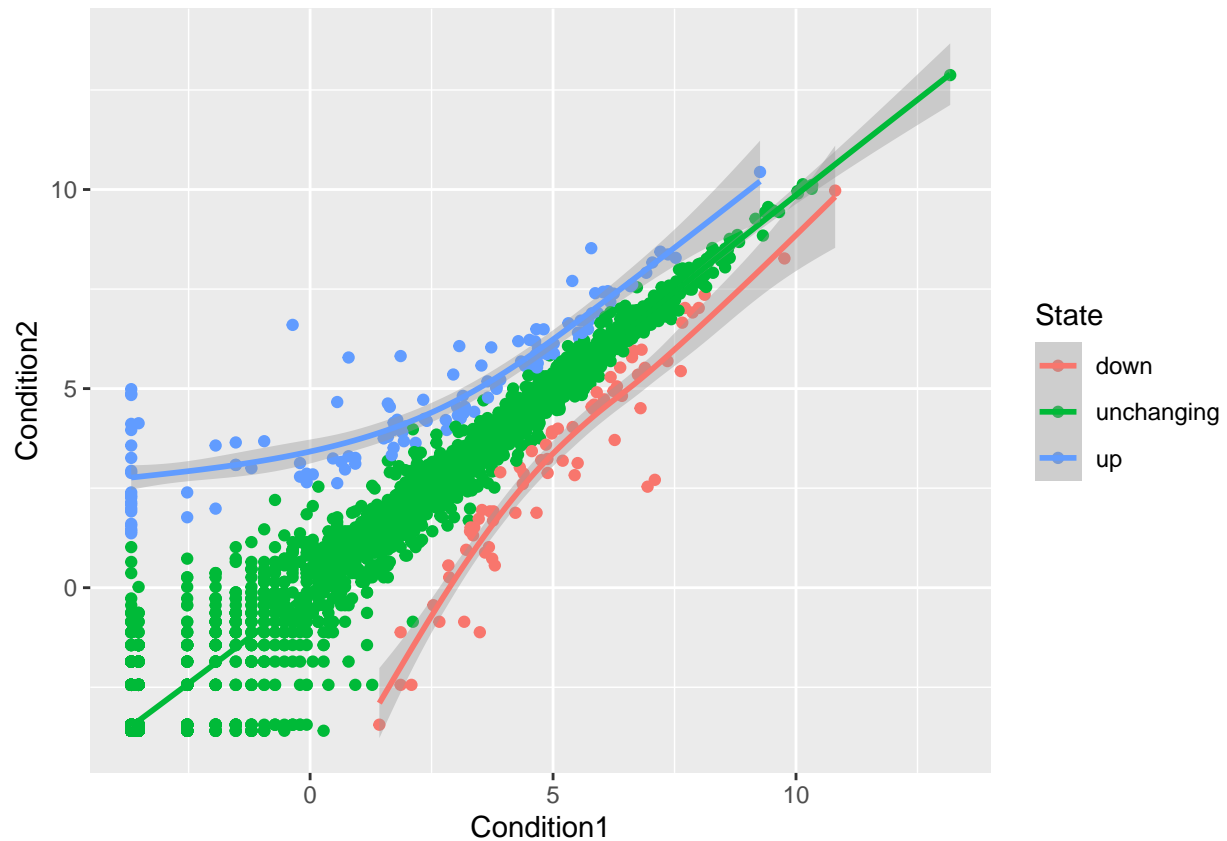
```
#changing color of the dots
p+scale_color_manual(values=c("blue","gray","red"))
```

```
p+geom_smooth()
```

```
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```

```
#install gapminder
#install.packages("gapminder")
library(gapminder)
head(gapminder)
```

```
## # A tibble: 6 x 6
##   country     continent  year lifeExp        pop gdpPercap
##   <fct>       <fct>     <int>   <dbl>      <int>     <dbl>
## 1 Afghanistan Asia       1952    28.8    8425333      779.
## 2 Afghanistan Asia       1957    30.3    9240934      821.
## 3 Afghanistan Asia       1962    32.0   10267083      853.
## 4 Afghanistan Asia       1967    34.0   11537966      836.
## 5 Afghanistan Asia       1972    36.1   13079460      740.
## 6 Afghanistan Asia       1977    38.4   14880372      786.
```

```
#visualizing gapminder
 ggplot(gapminder) +
  aes(x=year,y=lifeExp,col=continent)+
  geom_point(alpha=0.3)+
  geom_jitter(width=0.2,alpha=0.4)+
 # geom_boxplot(aes(group=year),alpha=0.2)+
   geom_violin(aes(group=year),alpha=0.2,draw_quantiles = 0.5)
```
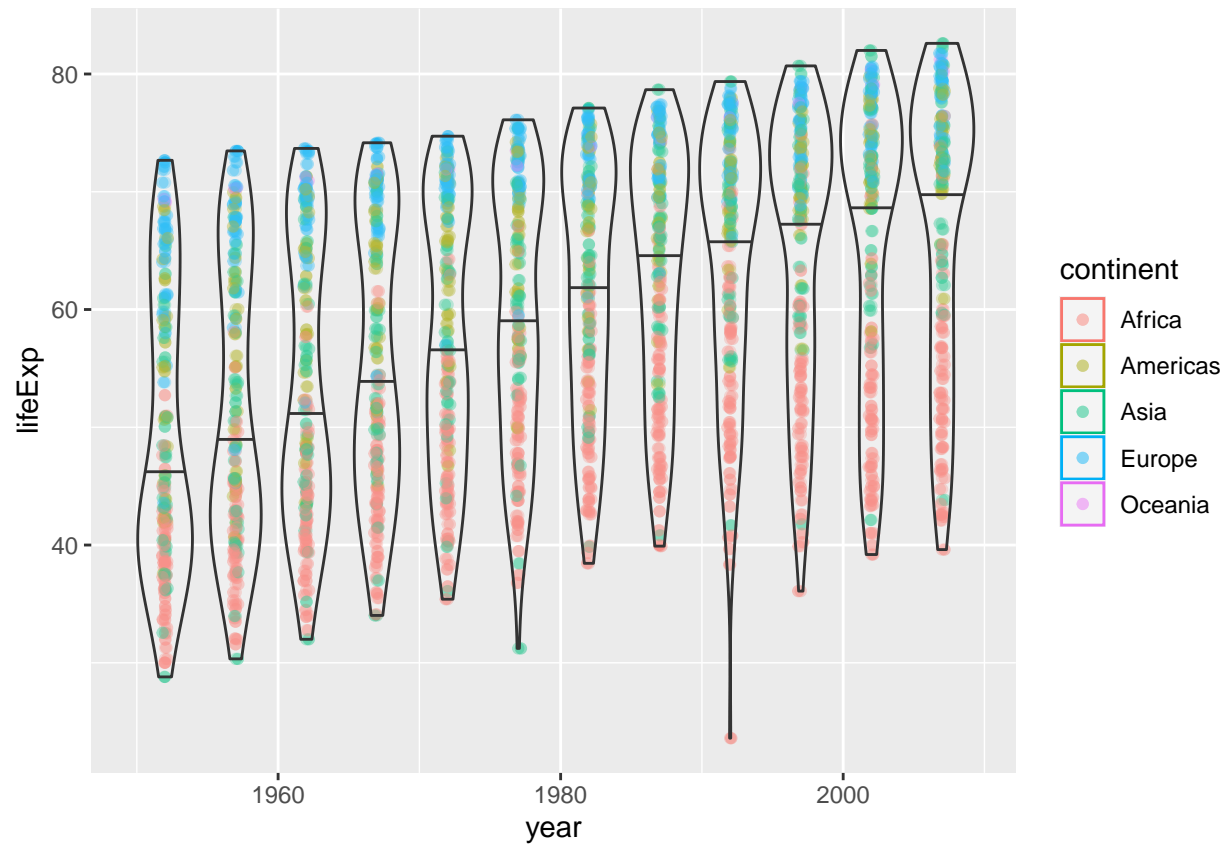
```
#install the plotly
#install.packages("plotly")
# library(plotly)
# ggplotly()
#install.packages("tinytex")
#tinytex::install_tinytex()
```