visualization-uhuru-day2.Rmd

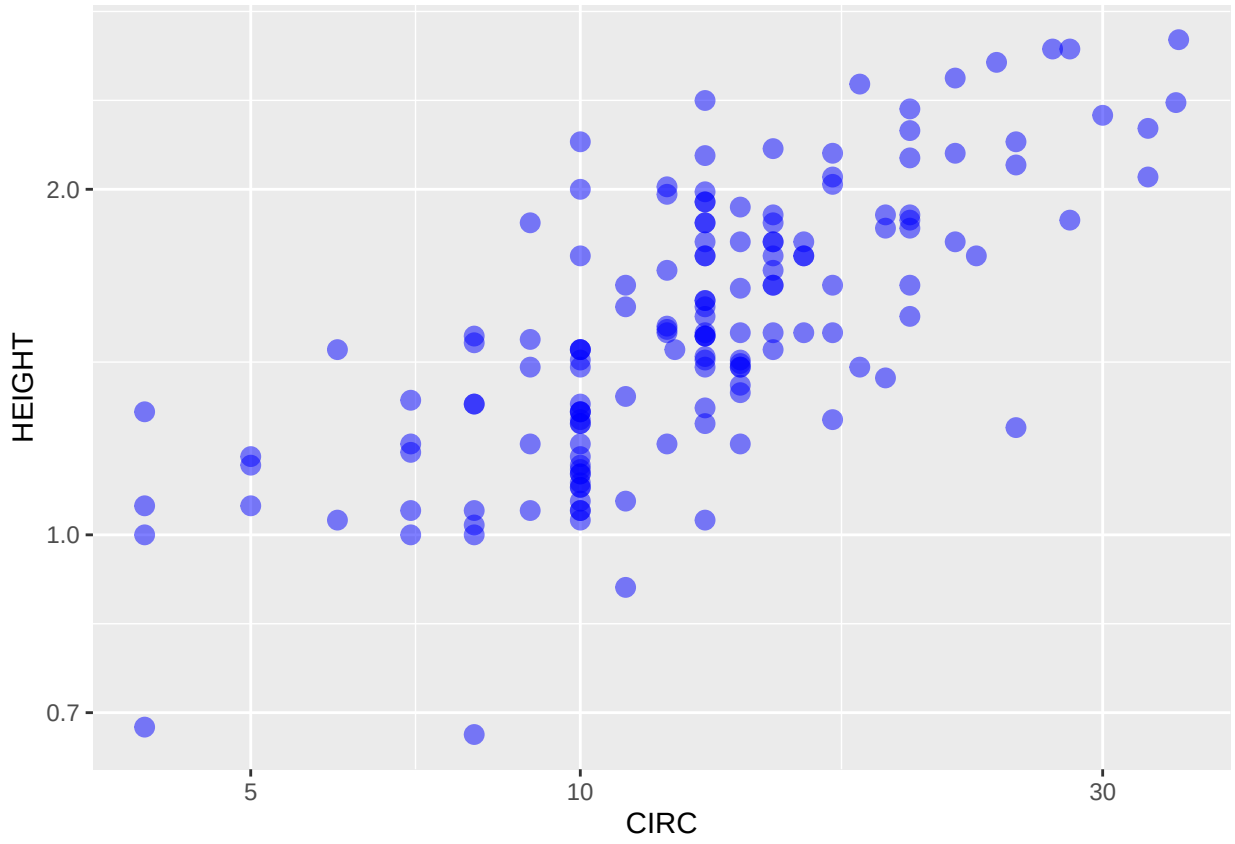
Austin Mercado

2023-02-28

In Class Activity Day:2 ######

```
ggplot(data = ACACIA, mapping = aes(x = CIRC, y = HEIGHT)) +
geom_point(size = 3, color = "blue", alpha = 0.5) +
scale_y_log10() +
scale_x_log10()
```

```
## Warning: Removed 4 rows containing missing values (geom_point).
```

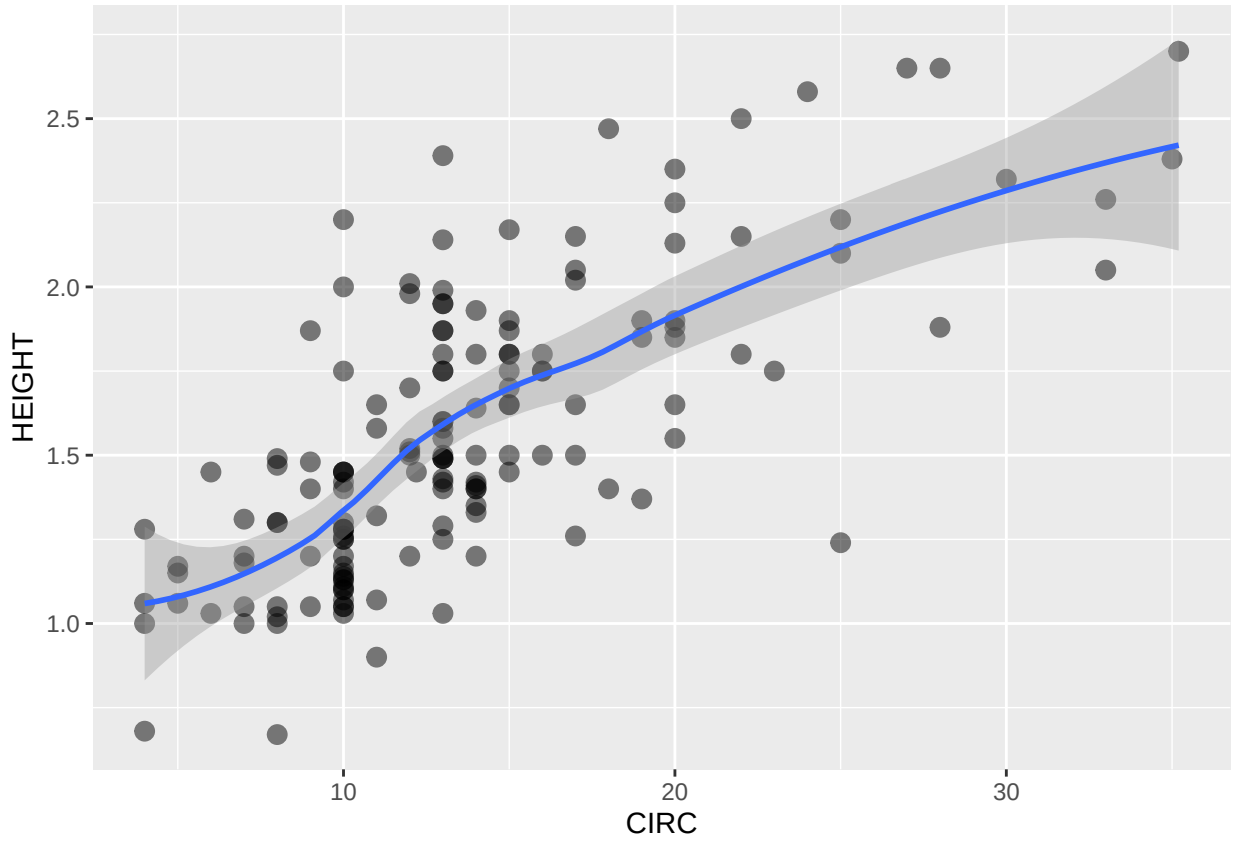



```
ggplot(ACACIA, aes(x = CIRC, y = HEIGHT)) +
geom_point(size = 3, alpha = 0.5) +
geom_smooth()
```

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

## Warning: Removed 4 rows containing non-finite values (stat_smooth).

## Warning: Removed 4 rows containing missing values (geom_point).
```

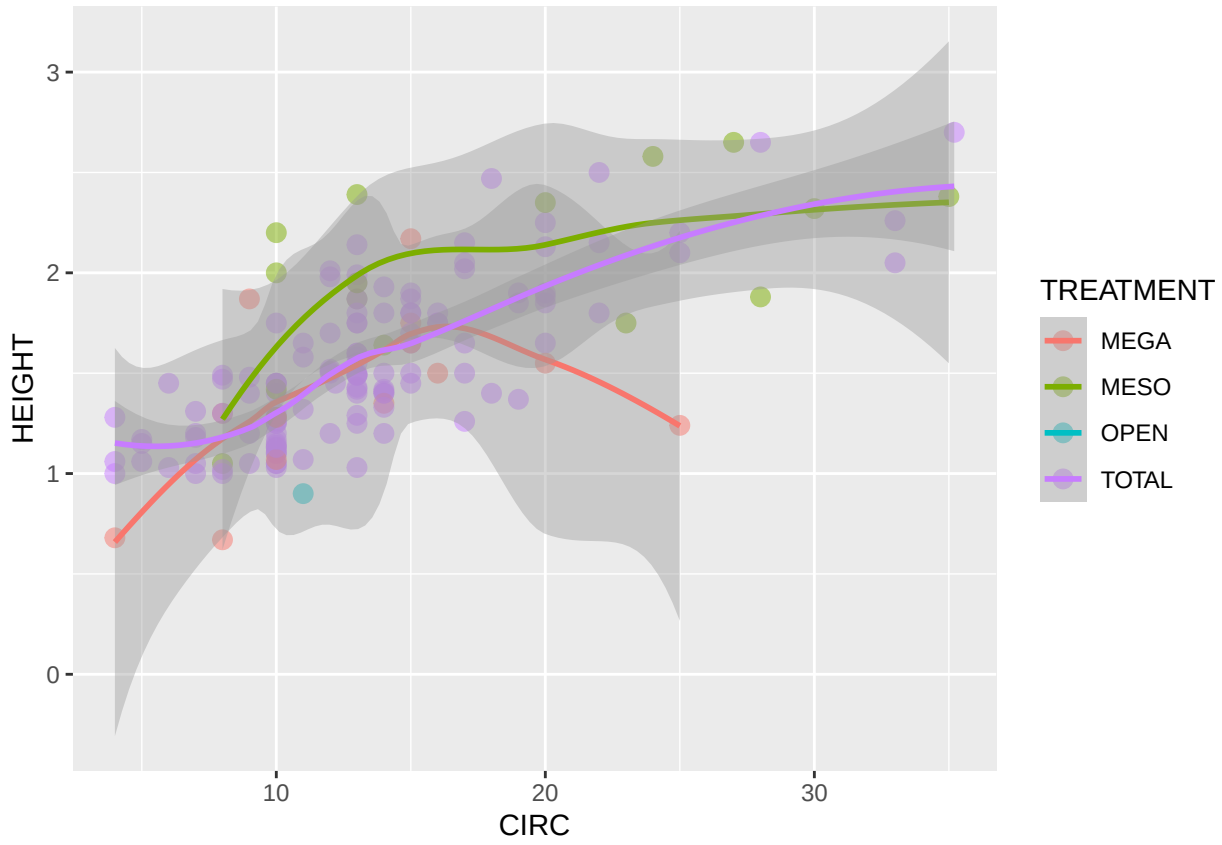



```
ggplot(ACACIA, aes(x = CIRC, y = HEIGHT, color = TREATMENT)) +
geom_point(size = 3, alpha = 0.5) +
geom_smooth()
```

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

## Warning: Removed 4 rows containing non-finite values (stat_smooth).

## Warning: Removed 4 rows containing missing values (geom_point).
```

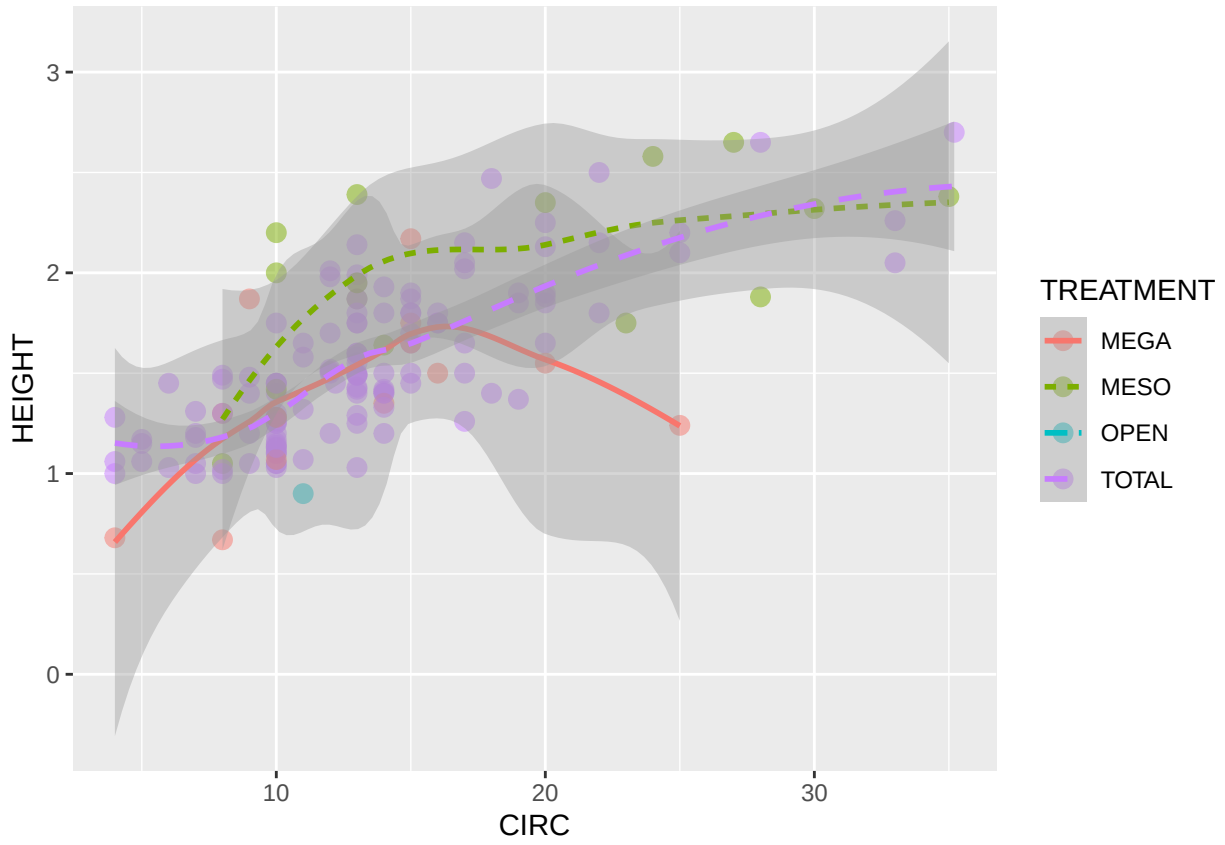



```
ggplot(ACACIA, aes(x = CIRC, y = HEIGHT, color = TREATMENT, linetype = TREATMENT)) +
geom_point(size = 3, alpha = 0.5) +
geom_smooth()
```

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

```
## Warning: Removed 4 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 4 rows containing missing values (geom_point).
```

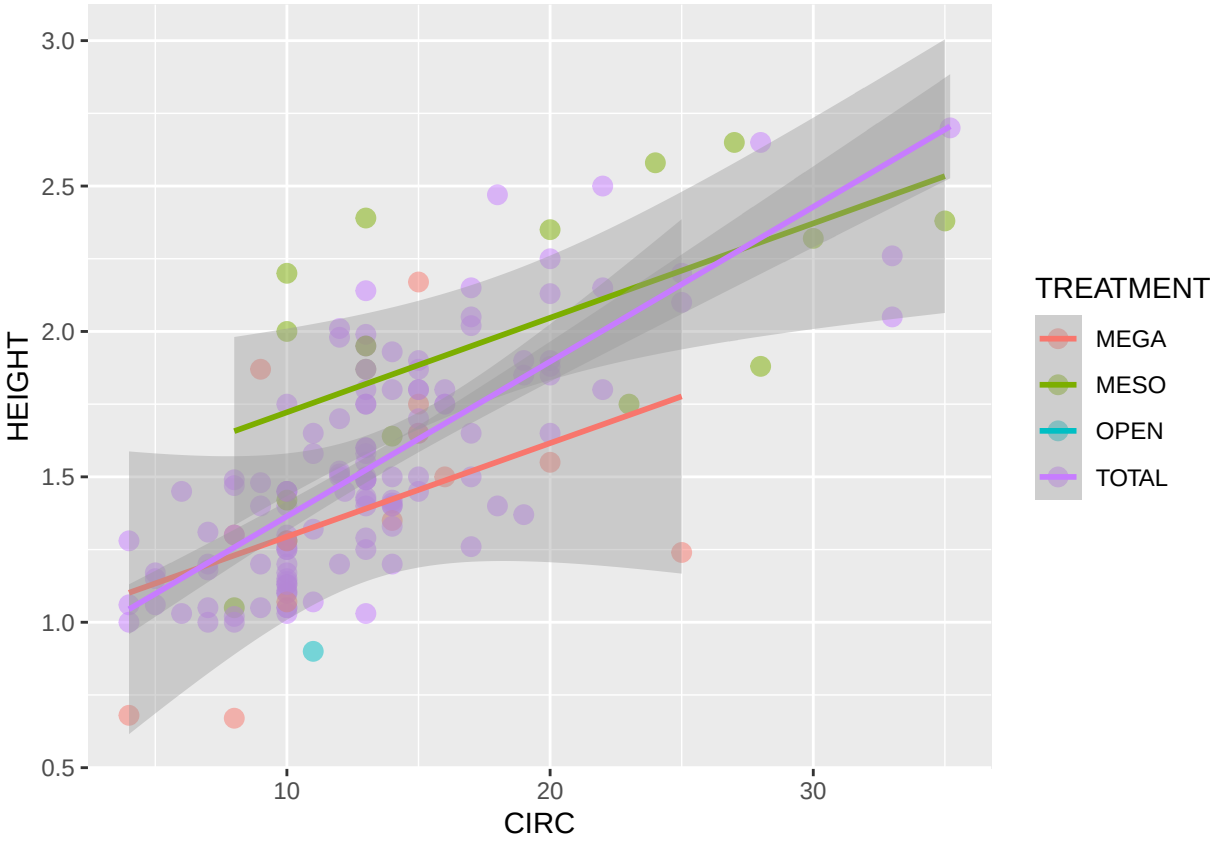



```
ggplot(ACACIA, aes(x = CIRC, y = HEIGHT, color = TREATMENT)) +
geom_point(size = 3, alpha = 0.5) +
geom_smooth(method = "lm") # try with "glm"
```

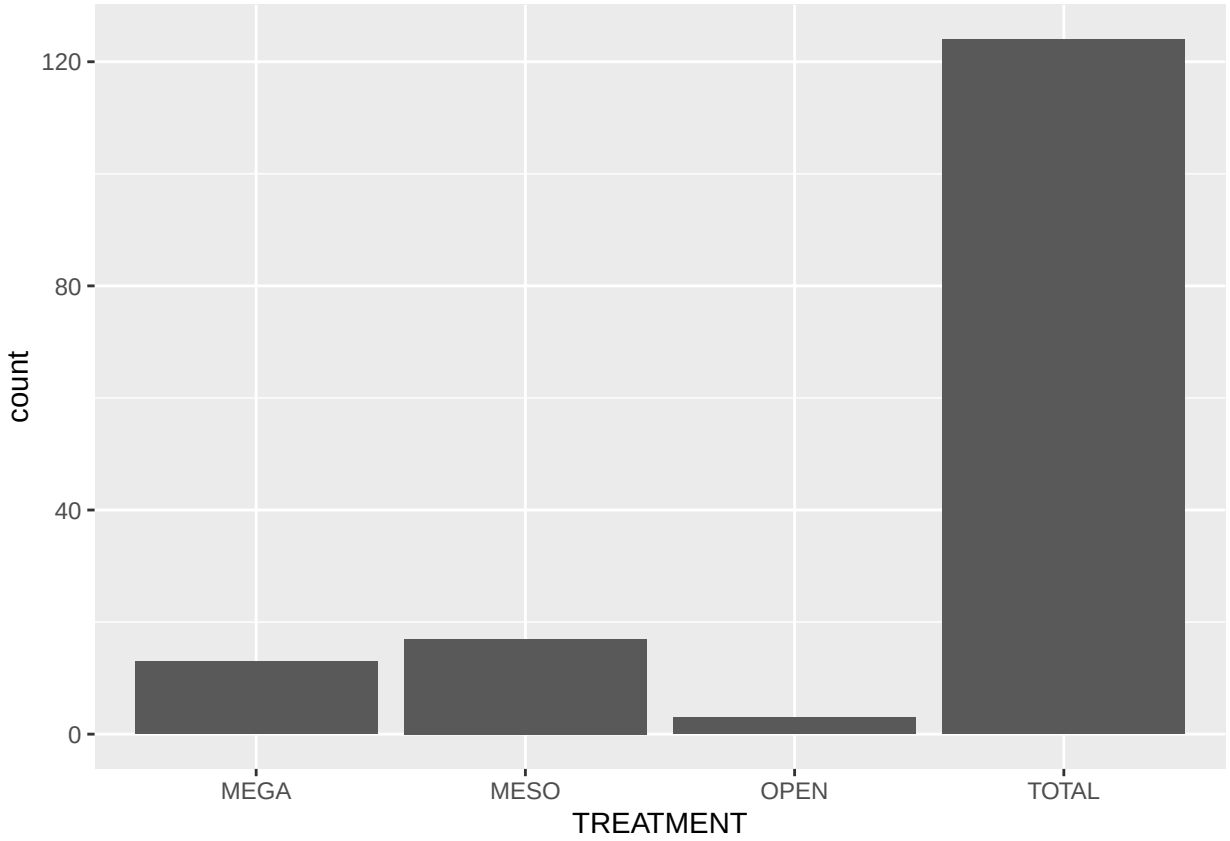
```
## `geom_smooth()` using formula 'y ~ x'
```

```
## Warning: Removed 4 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 4 rows containing missing values (geom_point).
```



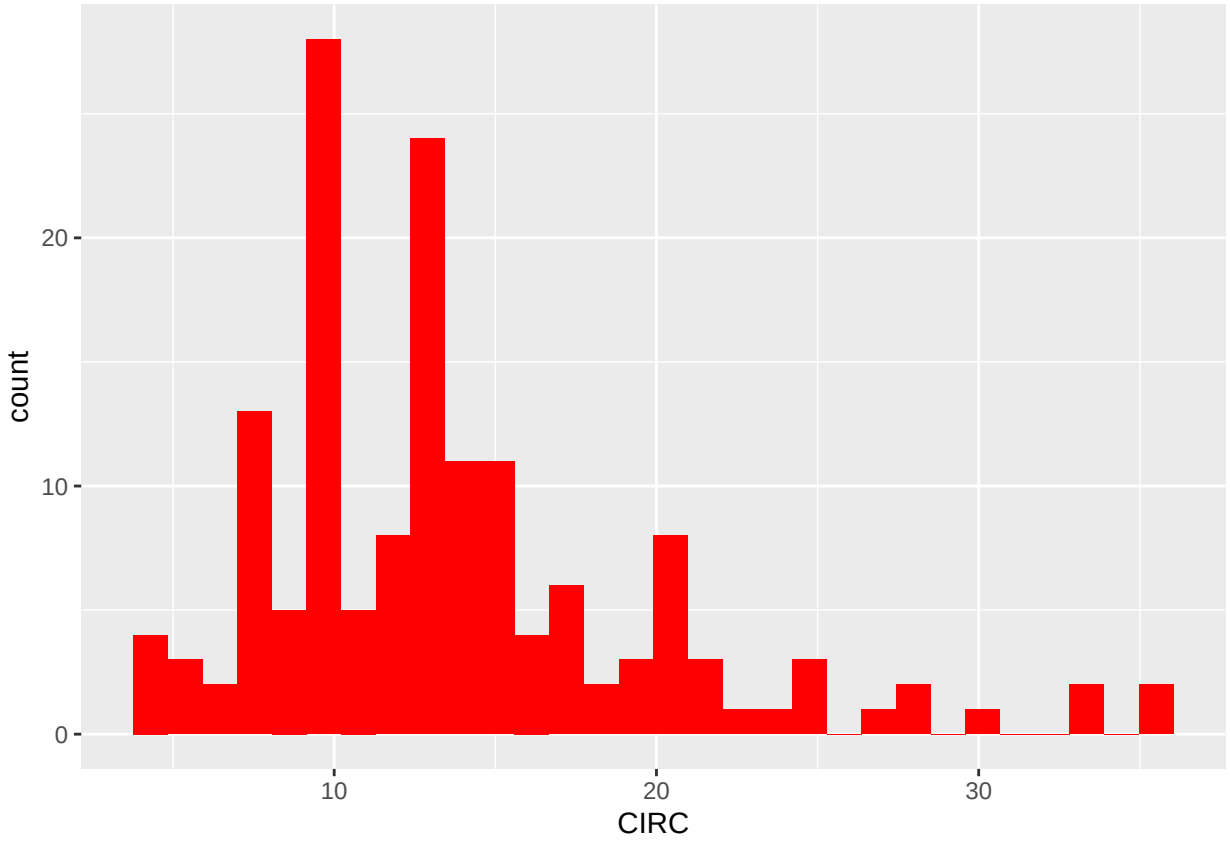

```
ggplot(data = ACACIA, mapping = aes(x = TREATMENT)) +
geom_bar()
```




```
ggplot(ACACIA, aes(x = CIRC)) +
geom_histogram(fill = "red")
```

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

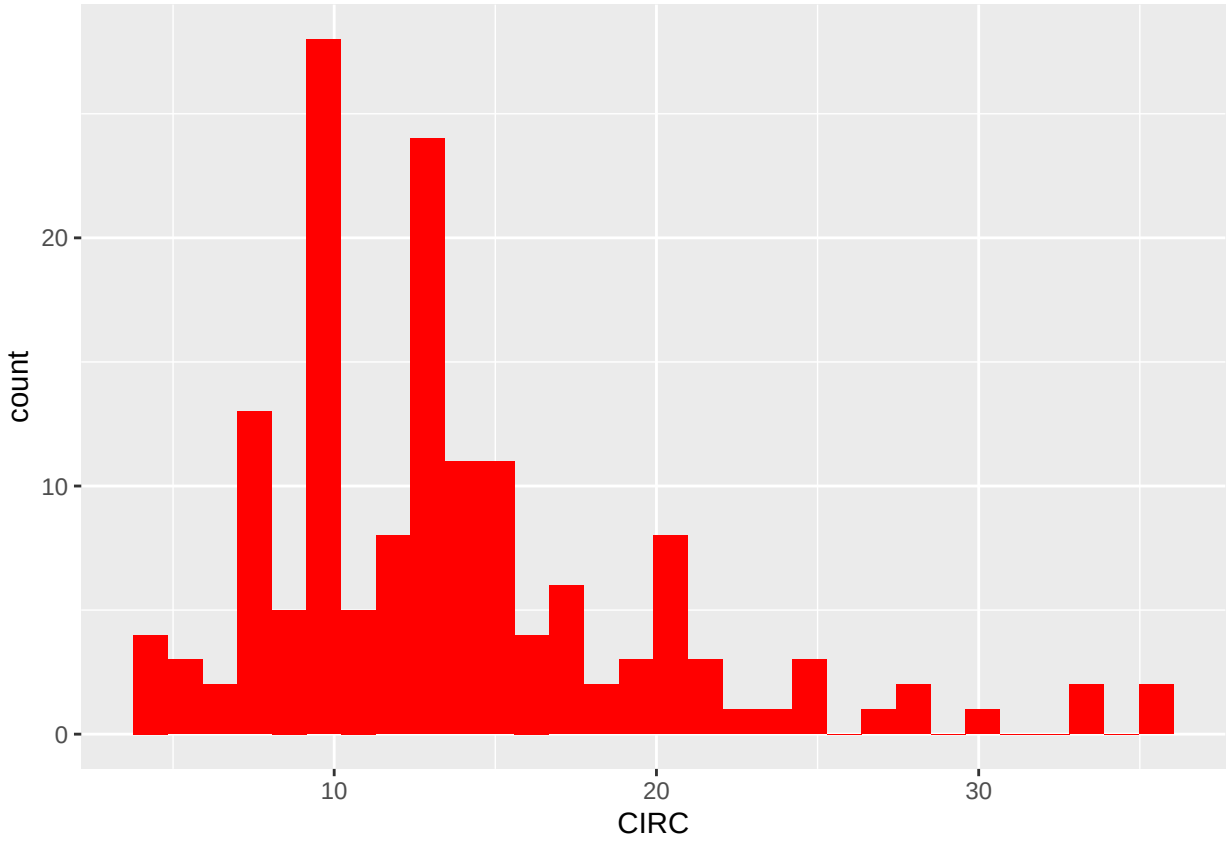
```
## Warning: Removed 4 rows containing non-finite values (stat_bin).
```




```
ggplot(ACACIA, aes(x = CIRC)) +
geom_histogram(fill = "red")
```

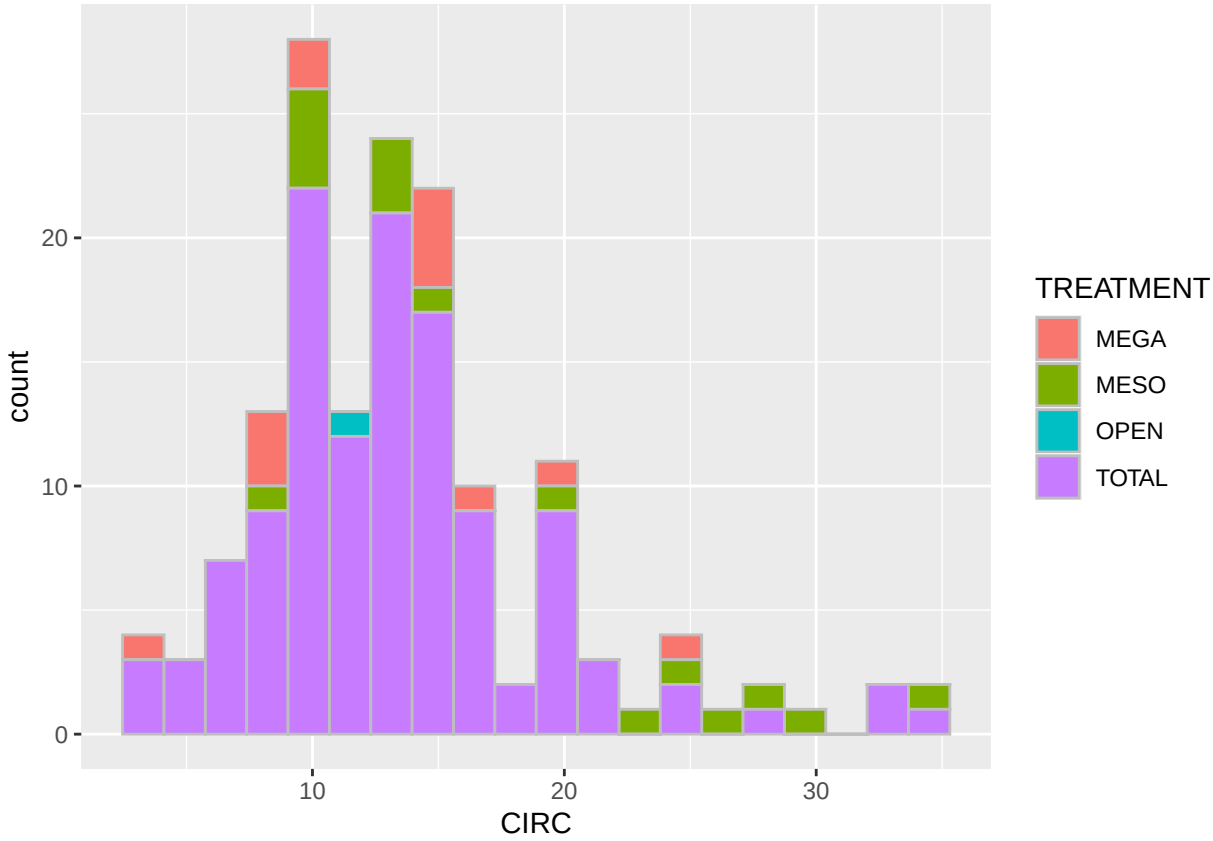
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
## Warning: Removed 4 rows containing non-finite values (stat_bin).
```




```
ggplot(ACACIA, aes(x = CIRC, fill = TREATMENT)) +
geom_histogram(bins = 20, color = "gray")
```

```
## Warning: Removed 4 rows containing non-finite values (stat_bin).
```

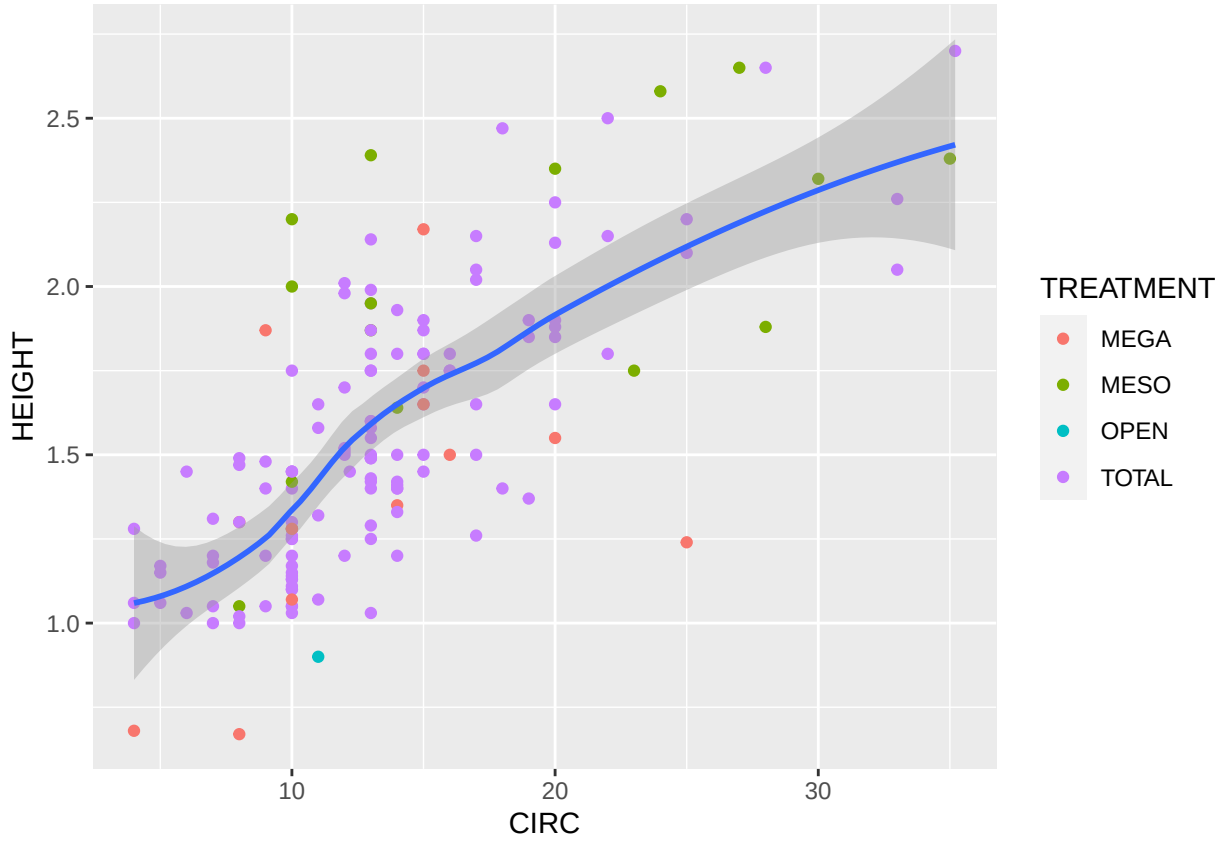



```
ggplot() +
geom_point(data = ACACIA,
           mapping = aes(x = CIRC, y = HEIGHT,
                         color = TREATMENT)) +
geom_smooth(data = ACACIA,
            mapping = aes(x = CIRC, y = HEIGHT))
```

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

```
## Warning: Removed 4 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 4 rows containing missing values (geom_point).
```




```
ggplot() +
geom_point(data = ACACIA,
           mapping = aes(x = CIRC, y = HEIGHT,
                         color = TREATMENT)) +
geom_smooth(data = ACACIA,
            mapping = aes(x = CIRC, y = HEIGHT)) +
geom_histogram(data = ACACIA,
               mapping = aes(x = CIRC, color = TREATMENT),
               alpha = 0.1)
```

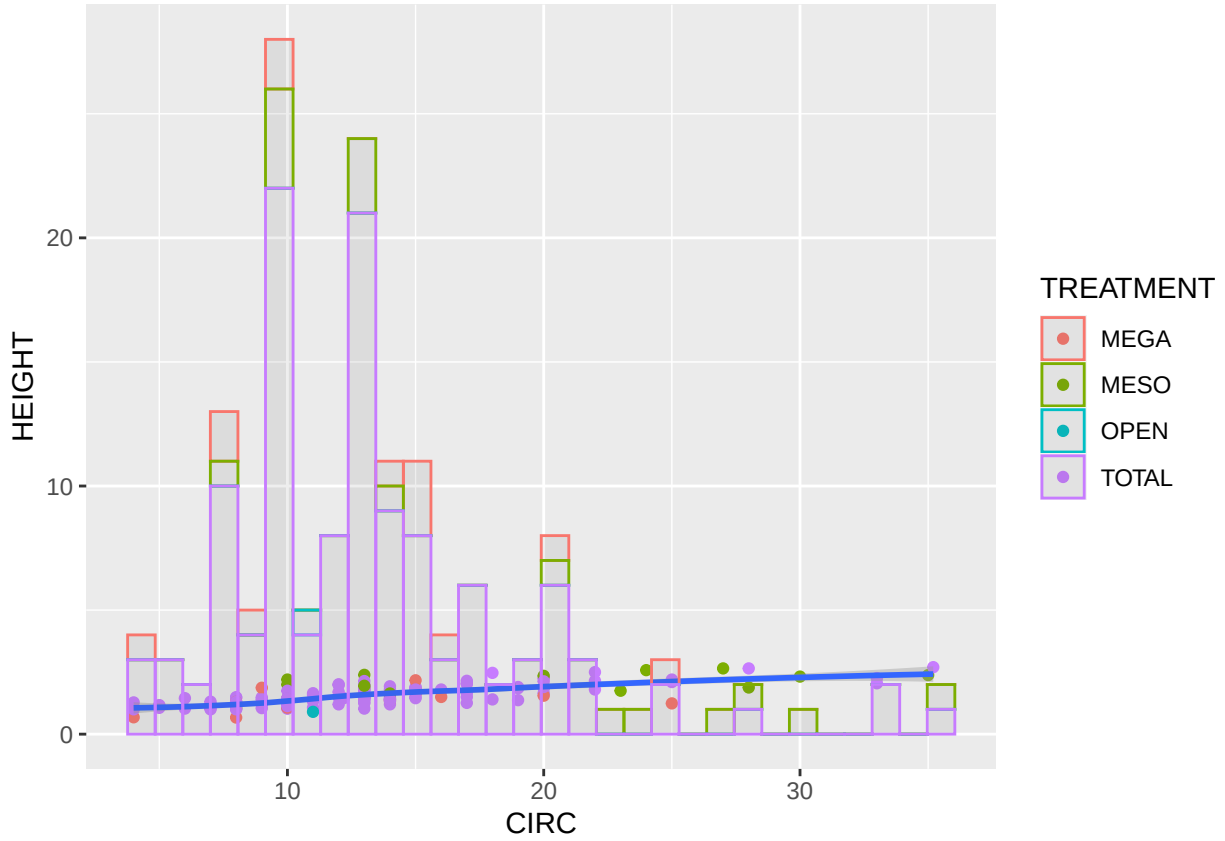
```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 4 rows containing non-finite values (stat_smooth).

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

## Warning: Removed 4 rows containing non-finite values (stat_bin).

## Warning: Removed 4 rows containing missing values (geom_point).
```




```
ggsave("acacia_by_treatment.jpg")
```

```
## Saving 6.5 x 4.5 in image
```

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

```
## Warning: Removed 4 rows containing non-finite values (stat_smooth).
```

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

```
## Warning: Removed 4 rows containing non-finite values (stat_bin).
```

```
## Warning: Removed 4 rows containing missing values (geom_point).
```

```
ggsave("acacia_by_treatment.pdf", height = 5, width = 5)
```

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

```
## Warning: Removed 4 rows containing non-finite values (stat_smooth).
```

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

```
## Warning: Removed 4 rows containing non-finite values (stat_bin).
```

```
## Warning: Removed 4 rows containing missing values (geom_point).
```

We are working with the file 'ACACIA_DREPANOLOBIUM_SURVEY.txt' file that currently lives in the 'data-raw' folder.

```
#make sure to provude file name as relative path
read.csv(file = "../data-raw/ACACIA_DREPANOLOBIUM_SURVEY.txt", sep = "\t", na.strings = "dead") -> ACACI
```

```
head(ACACIA2)
```

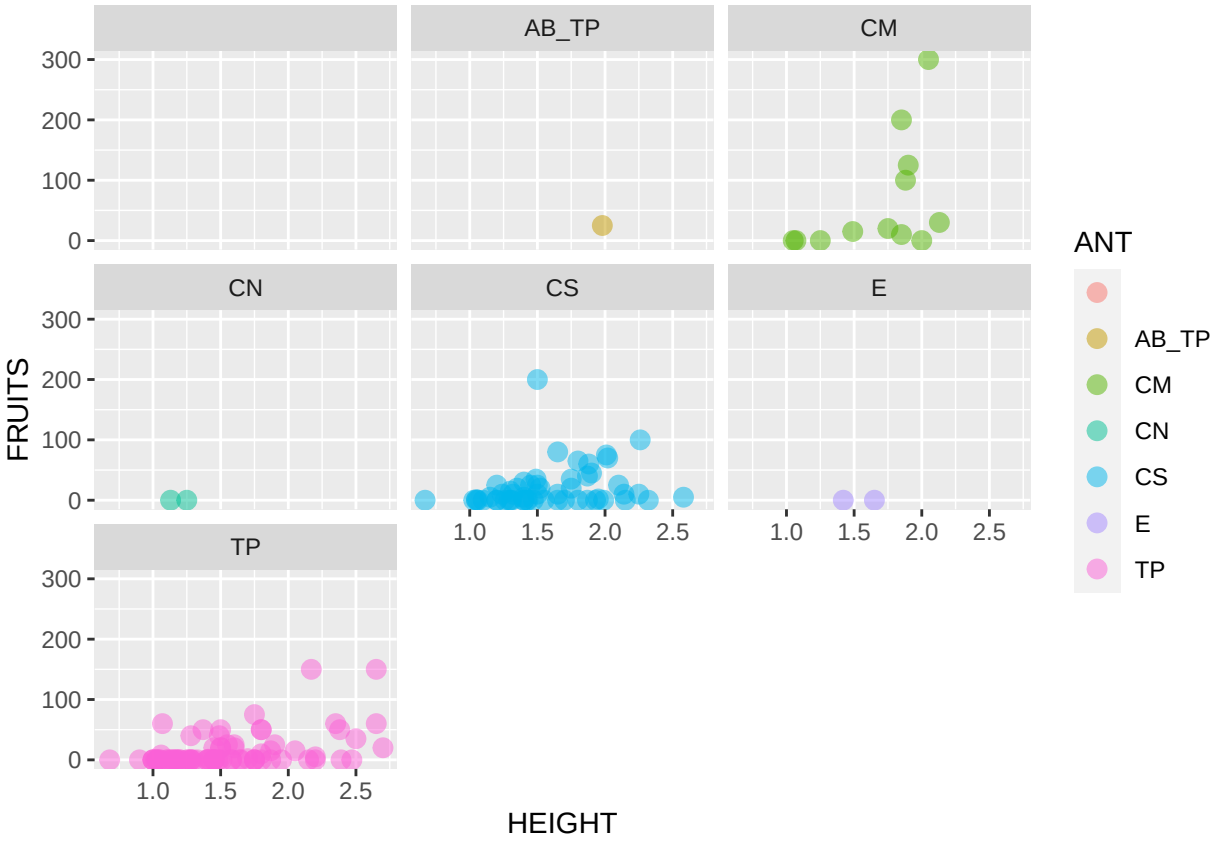
```
##   SURVEY YEAR  SITE BLOCK TREATMENT    PLOT   ID HEIGHT AXIS1 AXIS2 CIRC
## 1      1 2012 SOUTH     1     TOTAL S1TOTAL  581   2.25  2.75  2.15   20
## 2      1 2012 SOUTH     1     TOTAL S1TOTAL  582   2.65  4.10  3.90   28
## 3      1 2012 SOUTH     1     TOTAL S1TOTAL 3111   1.50  1.70  0.85   17
## 4      1 2012 SOUTH     1     TOTAL S1TOTAL 3112   2.01  1.80  1.60   12
## 5      1 2012 SOUTH     1     TOTAL S1TOTAL 3113   1.75  1.84  1.42   13
## 6      1 2012 SOUTH     1     TOTAL S1TOTAL 3114   1.65  1.62  0.85   15
##   FLOWERS BUDS FRUITS ANT
## 1       0    0     10  CS
## 2       0    0    150  TP
## 3       2    1     50  TP
## 4       0    0     75  CS
## 5       0    0     20  CS
## 6       0    0      0   E
```

##Plot the data as a scatterplot

For this we use the function 'geom_point()'

```
library(ggplot2)
ggplot(data = ACACIA2, mapping = aes(x = HEIGHT, y = FRUITS, color = ANT)) + geom_point(size = 3, alpha
```

```
## Warning: Removed 4 rows containing missing values (geom_point).
```


AB_TP
CM
CN
CS
E
TP
FRUITS
HEIGHT
0
100
200
300
1.0
1.5
2.0
2.5
ANT
AB_TP
CM
CN
CS
E
TP