> deg\_child\_06 <- GSS2016 %>%

+ select(degree, childs) %>%

+ filter(!is.na(degree)) %>%

+ filter(!is.na(childs))

Warning message:

package ‘bindrcpp’ was built under R version 3.4.4

> count(deg\_child\_06, degree) %>%

+ mutate(percent = n/sum(n)\*100) %>%

+ arrange(desc(percent)) %>%

+ mutate(cumulative\_prc = cumsum(percent))

# A tibble: 5 x 4

degree n percent cumulative\_prc

<chr> <int> <dbl> <dbl>

1 high school 1459 51.1 51.1

2 bachelor 536 18.8 69.9

3 lt high school 326 11.4 81.3

4 graduate 318 11.1 92.5

5 junior college 215 7.53 100

> #unique responses

> unique(deg\_child\_06$childs)

[1] "3" "0" "2" "4"

[5] "5" "7" "6" "eight or more"

[9] "1"

> deg\_child\_06$childs <- recode(deg\_child\_06$childs,

+ "0" = 0,

+ "1" = 1,

+ "2" = 2,

+ "3" = 3,

+ "4" = 4,

+ "5" = 5,

+ "6" = 6,

+ "7" = 7,

+ "eight or more" = 8

+ )

> count(deg\_child\_06, childs) %>%

+ mutate(percent = n/sum(n)\*100) %>%

+ arrange(desc(percent)) %>%

+ mutate(cumulative\_prc = cumsum(percent))

# A tibble: 9 x 4

childs n percent cumulative\_prc

<dbl> <int> <dbl> <dbl>

1 0 797 27.9 27.9

2 2 731 25.6 53.5

3 3 467 16.4 69.9

4 1 459 16.1 86.0

5 4 211 7.39 93.4

6 5 92 3.22 96.6

7 6 50 1.75 98.4

8 7 25 0.876 99.2

9 8 22 0.771 100

> # Table for 'degree' and 'childs'

> table(deg\_child\_06$degree, deg\_child\_06$childs)

0 1 2 3 4 5 6 7 8

bachelor 190 86 157 63 23 10 4 0 3

graduate 106 46 97 45 16 5 2 1 0

high school 396 247 351 258 113 48 27 9 10

junior college 59 42 55 32 16 4 2 4 1

lt high school 46 38 71 69 43 25 15 11 8

> deg\_child\_06 %>%

+ group\_by(degree) %>%

+ summarize(count= n(), mean = mean(childs), std.dev = sd(childs)) %>%

+ arrange(desc(mean)) %>%

+ print()

# A tibble: 5 x 4

degree count mean std.dev

<chr> <int> <dbl> <dbl>

1 lt high school 326 2.81 1.98

2 high school 1459 1.86 1.65

3 junior college 215 1.77 1.63

4 graduate 318 1.52 1.38

5 bachelor 536 1.45 1.44

> # ANOVA test

> model <- aov(childs ~ degree, data = deg\_child\_06)

> summary(model) # getting the summary of our ANOVA model

Df Sum Sq Mean Sq F value Pr(>F)

degree 4 419 104.84 39.73 <2e-16 \*\*\*

Residuals 2849 7519 2.64

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1