ANOVA, ANCOVA & Factorial ANOVA Materials

Brier Gallihugh, M.S.

2024-02-23

Table of contents

Introduction to ANOVA

Running ANOVA

```r
library(car)
library(psych)
library(multcomp)
library(effects)
library(tidyverse)
library(sjstats)


data <- dplyr::starwars

data <- data %>%
  dplyr::select(height,mass,hair_color,species,sex) %>%
  na.omit() %>%
  filter(species == "Gungan" | species == "Human" | species == "Wookiee")   ①

height_species_aov <- aov(height ~ species, data = data)                     ②
summary(height_species_aov)                                                  ③

# Post Hoc Test (Tukey)
TukeyHSD(height_species_aov)                                                 ④
# Bonferroni Adjustment vs Tukey
pairwise.t.test(data$height, data$species, p.adjust.method = "bonferroni")   ⑤
```

① Here we are using the `filter()` function to filter for "Gungan", "Human" and "Wookiee"
② This is the `aov()` function. Here we are specifying the ANOVA formula.
③ The `summary()` function will give you the output of the ANOVA
④ Because the `species` variable has more than 2 conditions, we might want to figure out which comparisons are statistically different from each other. The `TukeyHSD()` function allows us to perform a post hoc Tukey test
⑤ Maybe Tukey isn't your favorite correction. Another route might be to use a pairwise t test that corrects for multiple comparisons using a Bonferroni correction. You can do that with the `pairwise.t.test()` function.

```
          Df Sum Sq Mean Sq F value   Pr(>F)
species    2   5841  2920.6   21.11 9.43e-06 ***
Residuals 21   2906   138.4
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = height ~ species, data = data)

$species
                diff        lwr       upr       p adj
Human-Gungan   -29.75 -51.738704 -7.761296 0.0071194
Wookiee-Gungan  21.00  -8.649562 50.649562 0.1987756
Wookiee-Human   50.75  28.761296 72.738704 0.0000258


    Pairwise comparisons using t tests with pooled SD

data:  data$height and data$species

        Gungan Human
Human   0.0079 -
Wookiee 0.2660 2.7e-05

P value adjustment method: bonferroni
```

Assumptions of ANOVA

As ANOVA is just a special case of linear regression, the same assumptions exist for ANOVA
that exist for linear regression. We will test each assumption below. The code will look awfully
similar to the code used for linear regression previously

Model Normality

Graphical

```
density_plot <- data %>% ggplot(aes(x = height_species_aov$residuals)) +
  geom_histogram(aes(y= after_stat(density)),binwidth = 1) +
  stat_function(fun = dnorm,
                args = list(mean = mean(height_species_aov$residuals),
                            sd = sd(height_species_aov$residuals)),
                col = "blue",
                linewidth = 1) +
  labs(title = "Figure 1. Histogram of Model Residual Scores",
```
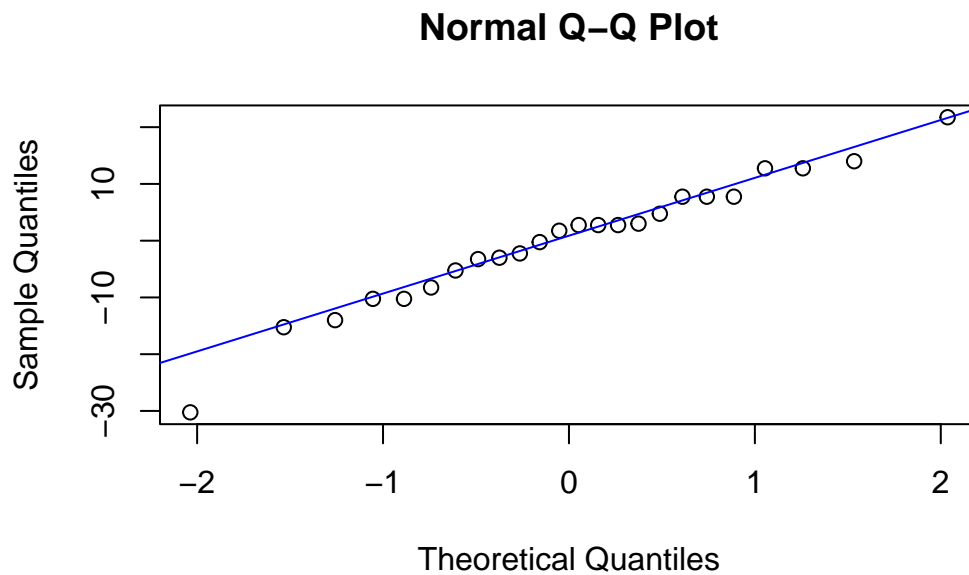
```
        x = "Model Residuals",
        y = "Density")

density_plot
```

Figure 1. Histogram of Model Residual Scores
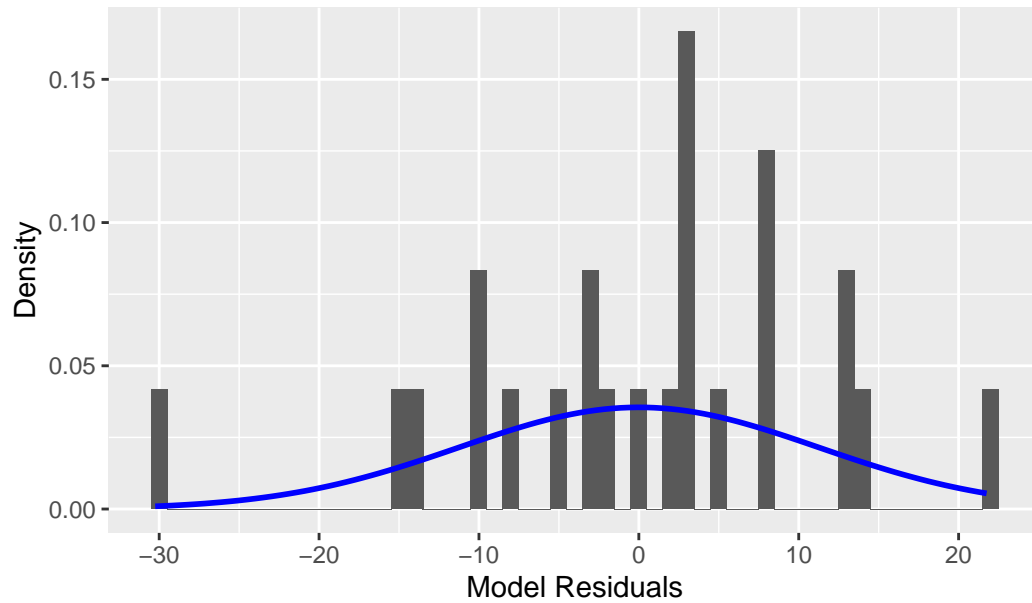

```
# OR

qqnorm(height_species_aov$residuals)                                    ①
qqline(height_species_aov$residuals, col = "blue")
```

4

Normal Q–Q Plot

1.When running the `qqnorm()` and `qqline()` functions at the same time, you will get a qq-plot which will graphically assess the normality of the argument given (in this case the model residuals). The `col` argument simply tells R which color to make the reference line.

Statistical

```
psych::describe(height_species_aov$residuals)                    ①
shapiro.test(height_species_aov$residuals)
```

① Another way to assess normality is to look at the skew and kurtosis of the data. Typically skew and kurtosis between -1 and 1 are acceptable. We can see this using the `describe()` function within the `psych` package.

```
    vars  n mean    sd median trimmed  mad    min   max range  skew kurtosis
X1     1 24    0 11.24   2.25    0.49 8.15 -30.25 21.75    52 -0.53     0.35
      se
X1 2.29

    Shapiro-Wilk normality test

data:  height_species_aov$residuals
W = 0.97026, p-value = 0.6734
```
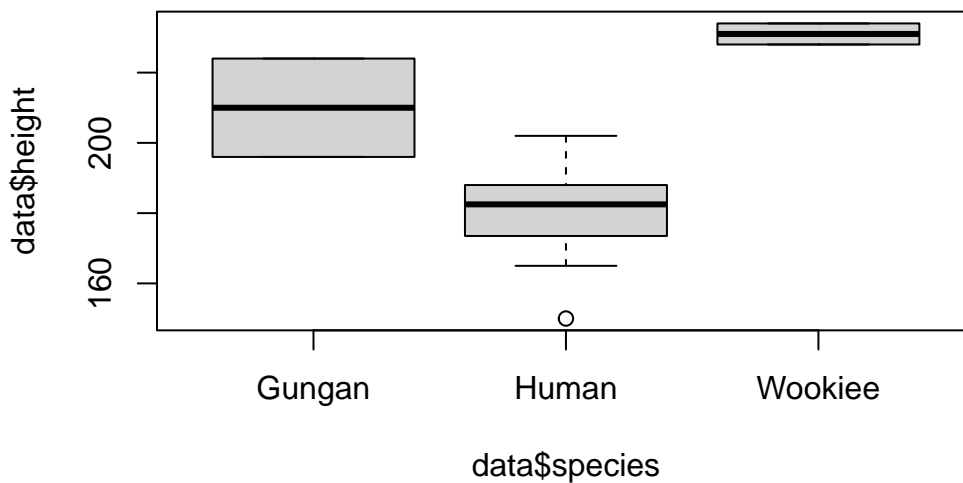
Homogeneity of Variance

As with regression, homogeneity of variance is also important for ANOVA. We will graphically and statistically assess this assumption below.

Graphical

```
boxplot(data$height ~ data$species)                                              ①
```

① The `boxplot()` function will give us a visual representation of the DV at each level of the IV. We are looking for boxplots that are roughly the same height across each group.

Statistical

Statistically, we can assess homogeneity of variance using Levene's Test. Here we want the results to not be statistically significant.

```
leveneTest(data$height ~ data$species)                                           ①
```

① The `leveneTest()` function from the `car` package takes a DV and IV as arguments.

```
Levene's Test for Homogeneity of Variance (center = median)
      Df F value Pr(>F)
group  2  1.0572 0.3652
      21
```

Introduction to ANCOVA

Running ANCOVA

ANCOVA refers to any ANOVA model with 2 or more parameters. These models require similar assumptions to ANOVA with a couple of additional ones. Below we will see how to run an ANCOVA and then test its assumptions.

```
data <- data %>% mutate(species = as.factor(species))
# Type I SS
ancova <- aov(height ~ species + mass, data = data)
summary(ancova)                                                          ①
# Reversed Order
ancova2 <- aov(height ~ mass + species, data = data)
summary(ancova2)                                                         ②

car::Anova(ancova2, type = "III")                                        ③

# Std Means
summary_data <- data %>%
  dplyr::group_by(species) %>%
  dplyr::summarise(n = n(),
                   mean = mean(height))

# Adjust means for covariate effect
adjustedMeans<- effect("species", ancova, se=TRUE)                       ④
summary(adjustedMeans)
adjustedMeans$se

# Post Hoc Tests

posthoc <- multcomp::glht(ancova, linfct = multcomp::mcp(species = "Tukey")) ⑤
summary(posthoc)
confint(posthoc)

# Effect size
sjstats::anova_stats(car::Anova(ancova,type = "III"))                    ⑥
```

① This displays our ANCOVA output with species and then mass.
② This displays our ANCOVA output with mass and then species. We can see that 1 and 2 are different. This is because R by default does a sequential ANCOVA so order matters.
③ To get around this, we can use the `Anova()` function in the `car` package to specify that we want Type III sums of squares. This will then generate an ANCOVA where the order of predictors doesn't matter
④ When having covariates, one might wish to report adjusted means. We can do that using the `effect()` function from the `effects` package.
⑤ To run a post hoc analysis, we need to use the `ghlt()` function from the `multcomp` package. There are additional corrections you can do outside of Tukey. To investigate, please see the documentation.
⑥ To get basic ANOVA effect size statistics, we can use the `anova_stats()` function from the `sjstats` package.

```
          Df Sum Sq Mean Sq F value   Pr(>F)
species    2   5841  2920.6  28.662 1.34e-06 ***
mass       1    868   867.8   8.516   0.0085 **
Residuals 20   2038   101.9
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
          Df Sum Sq Mean Sq F value   Pr(>F)
mass       1   3231    3231   31.70 1.64e-05 ***
species    2   3478    1739   17.07 4.74e-05 ***
Residuals 20   2038     102
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Anova Table (Type III tests)

Response: height
            Sum Sq Df  F value    Pr(>F)
(Intercept) 27816.3  1 272.9827 3.998e-13 ***
mass          867.8  1   8.5164    0.0085 **
species      3478.5  2  17.0684 4.736e-05 ***
Residuals    2038.0 20
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

 species effect
species
  Gungan    Human  Wookiee
213.4855 181.2518 217.4968
```

```
 Lower 95 Percent Confidence Limits
species
  Gungan    Human  Wookiee
198.3892 176.4892 199.7528

 Upper 95 Percent Confidence Limits
species
  Gungan    Human  Wookiee
228.5818 186.0143 235.2409
[1] 7.237078 2.283137 8.506395


     Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: Tukey Contrasts


Fit: aov(formula = height ~ species + mass, data = data)

Linear Hypotheses:
                   Estimate Std. Error t value Pr(>|t|)
Human - Gungan == 0    -32.234      7.534  -4.278  < 0.001 ***
Wookiee - Gungan == 0    4.011     11.653   0.344  0.93377
Wookiee - Human == 0    36.245      8.986   4.034  0.00164 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Adjusted p values reported -- single-step method)


     Simultaneous Confidence Intervals

Multiple Comparisons of Means: Tukey Contrasts


Fit: aov(formula = height ~ species + mass, data = data)

Quantile = 2.4993
95% family-wise confidence level


Linear Hypotheses:
                   Estimate lwr       upr
Human - Gungan == 0    -32.2337 -51.0643 -13.4032
Wookiee - Gungan == 0    4.0113 -25.1119  33.1346
```

```
Wookiee - Human == 0    36.2451   13.7868   58.7034
```

term	sumsq	meansq	df	statistic	p.value	etasq	partial.etasq	omegasq
species	3478.457	1739.228	2	17.068	< .001	0.545	0.631	0.505
mass	867.799	867.799	1	8.516	0.008	0.136	0.299	0.118
Residuals	2037.951	101.898	20					

Assumptions of ANCOVA

ANCOVA requires the same assumptions as ANOVA with two additional assumptions: 1) That the predictor and covariate are independent 2) The regression slopes are homogeneous

We will test each of the traditional parametric tests as well as the two additional assumptions below.

Predictor x Covariate Indepenence

Statistical

```
predictor_assumption <- aov(mass ~ species, data = data)      ①
summary(predictor_assumption)
```

① To assess this statistically, one needs to run an ANOVA looking at each predictor with each additional parameter. We do not want these to be statistically significant.

```
            Df Sum Sq Mean Sq F value Pr(>F)
species      2   3543  1771.6   4.949 0.0173 *
Residuals   21   7517   357.9
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

> 💡 Tip
>
> Here the statistical test for this assumption is statistically significant (which is bad). It actually means we can't do this (unless we have random assignment)

Homogeneity of Regression Slopes

Statistical

```
regression_slope_assumption <- aov(height ~ species*mass, data = data)
car::Anova(regression_slope_assumption, type = "III")                    ①
```

① To test this assumption, one just needs to test the interaction effects within the model. The code for this is shown here.

```
Anova Table (Type III tests)

Response: height
              Sum Sq Df F value  Pr(>F)
(Intercept)   149.72  1  1.5160 0.23407
species       164.06  2  0.8306 0.45183
mass          392.00  1  3.9692 0.06173 .
species:mass  260.25  2  1.3176 0.29241
Residuals    1777.70 18
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

> 💡 Tip
>
> For the above example, the interaction is NOT significant so assumption is met

Residual Normality (DV ~ IV)

Residual normality is an assumption shared with standard regression. We can assess it the same way we did previously when looking at simple ANOVA. The same is true for the homogeneity of variance assumption.

Statistical

```
height_mass_resid <- scale(residuals(aov(height ~ mass, data = data)))   ①
psych::describe(height_mass_resid)                                       ②
shapiro.test(height_mass_resid)                                          ③
```

① The `scale()` function will scale the residuals of the ANOVA. The `residuals()` function pulls out the residuals within the ANOVA

② Statistical summary of the residuals using the `describe()` function in the `psych` package
③ Statistical test (Shapiro Wilks) of the residuals using the `shapiro.test()` function

```
    vars  n mean sd median trimmed  mad   min  max range skew kurtosis  se
X1     1 24    0  1  -0.07   -0.06 0.93 -1.83 2.47   4.3  0.6    -0.15 0.2
```

```
    Shapiro-Wilk normality test

data:  height_mass_resid
W = 0.9563, p-value = 0.3686
```

Graphical

```
hist(height_mass_resid, col = 'beige',                                        ①
     main="", xlab = "ANCOVA Residuals (z-score)",
     probability = TRUE)
curve(dnorm(x, mean = mean(height_mass_resid),
            sd = sd(height_mass_resid)),
            add = TRUE, lwd = 2, col = 'blue')
```

① A graphical histogram (with normal distribution overlay) of the model residuals

ANCOVA Residuals (z–score)
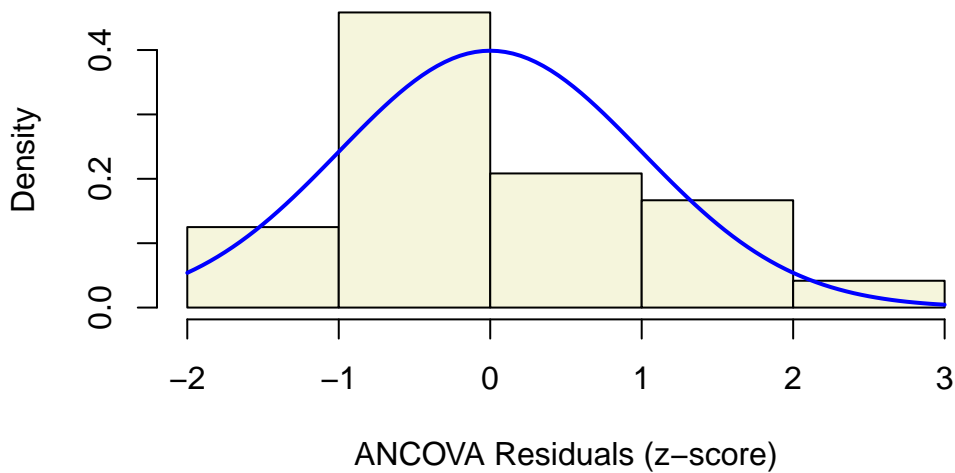
Homogeneity of Variance

Statistical

```
# Levene's Test to Assess Equal Variance for Species
car::leveneTest(data$height ~ data$species)                                    ①
```

① A statistical test (Levene's test) of the homogeneity assumption using the `leveneTest()` function in the `car` package.
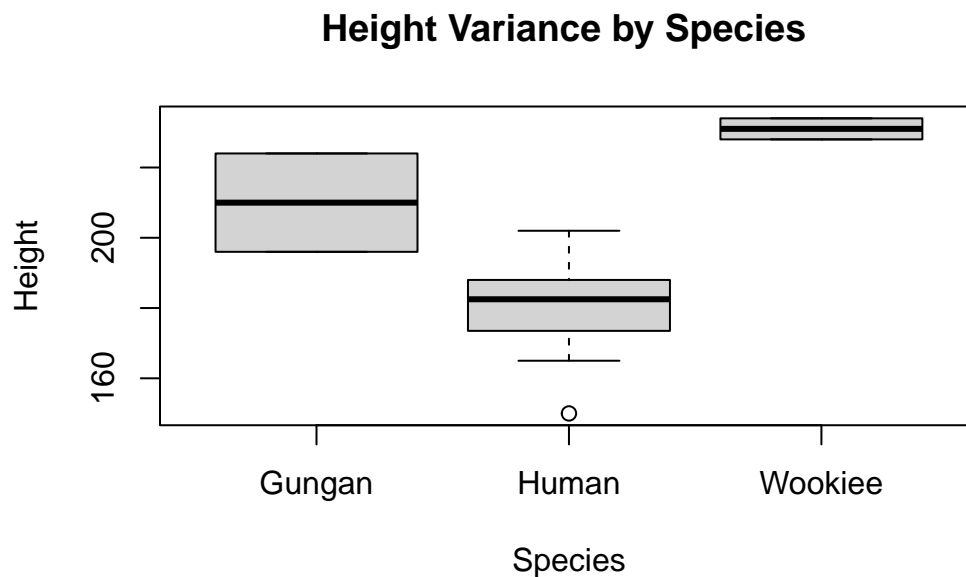
```
Levene's Test for Homogeneity of Variance (center = median)
      Df F value Pr(>F)
group  2  1.0572 0.3652
      21
```

Graphical

```
# Boxplot Height by Species
boxplot(height ~ species,data=data, main="Height Variance by Species ",    ①
   xlab="Species", ylab="Height")
```

① A visual representation of the homogeneity of variance assumption using a boxplot

Height Variance by Species
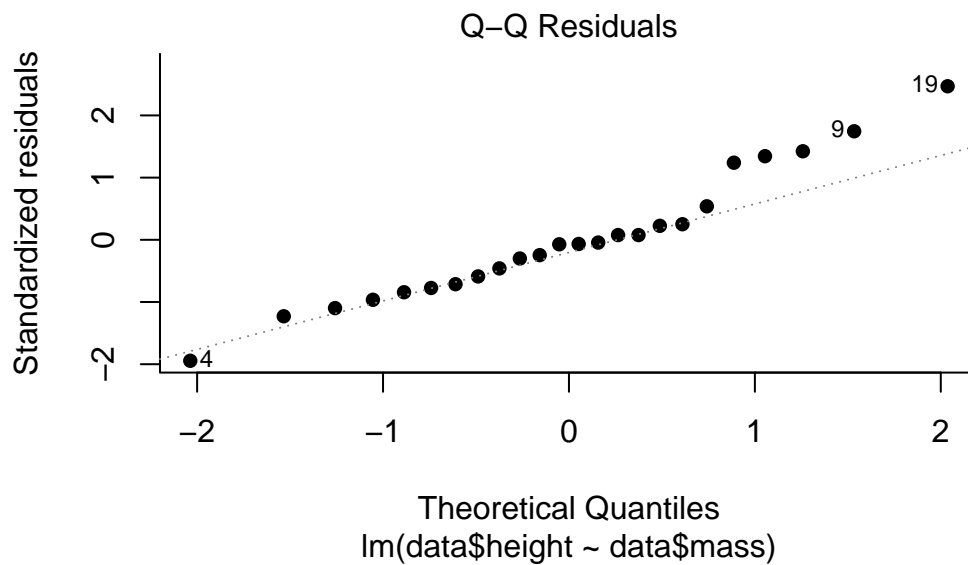
Linearity of CV & DV

The last assumption is that the CV and DV are linearly related. We can test this using the code below graphically.

Graphical

```
plot(lm(data$height ~ data$mass,data = data),          ①
     pch = 16, bty = 'l',2)
```

① Visual representation of the CV to DV linearity assumption

Introduction to Factorial ANOVA

Running Factorial ANOVA

```
library(tidyverse)
factorial_data <- starwars %>%                         ①
  select(mass,homeworld,species) %>%                   ②
  mutate(homeworld = as.factor(homeworld),             ③
```

14

```
        species = as.factor(species)) %>%                              ④
  filter(homeworld == "Tatooine" | homeworld == "Naboo") %>%           ⑤
  filter(species == "Human" | species == "Droid") %>% na.omit()        ⑥
# Modeling As Factorial ANOVA
aov_factorial <- aov(mass ~ species*homeworld, data = factorial_data)  ⑦
Anova(aov_factorial, type = "III")                                     ⑧


# Modeling As A Regression (For SS Analyses)
reg_fanova <- lm(mass ~ homeworld*species,data = factorial_data)       ⑨
summary(reg_fanova)                                                    ⑩
```

① Start with the **starwars** data set
② Use the **select()** function to pull out mass, homeworld and specices variables
③ Call the **mutate()** function to format the homeworld variable using the **as.factor()** function.
④ Format the species variables as a factor using the **as.factor()** function.
⑤ Use the **filter()** function to filter for observations with homeworld = "Tatooine" OR "Naboo"
⑥ Use the **filter()** function to filter for observations with species = "Human" OR "Droid"
⑦ Create an ANOVA object using the **aov()** function
⑧ Use the **Anova()** function from the **car()** package to specifiy Type III Sums of Squares
⑨ Create an Regression object using the **lm()** function
⑩ Show output of the regression object using the **summary()** function.


```
Anova Table (Type III tests)

Response: mass
                 Sum Sq Df F value Pr(>F)
(Intercept)      1024.0  1  1.5447 0.2539
species           522.7  1  0.7884 0.4041
homeworld         308.2  1  0.4649 0.5173
species:homeworld  97.0  1  0.1464 0.7134
Residuals        4640.5  7

Call:
lm(formula = mass ~ homeworld * species, data = factorial_data)

Residuals:
   Min     1Q Median     3Q    Max
-21.50 -17.00 -12.00  18.25  40.00

Coefficients:
```

```
                     Estimate Std. Error t value Pr(>|t|)
(Intercept)                     32.00       25.75   1.243    0.254
homeworldTatooine               21.50       31.53   0.682    0.517
speciesHuman                    28.00       31.53   0.888    0.404
homeworldTatooine:speciesHuman  14.50       37.90   0.383    0.713


Residual standard error: 25.75 on 7 degrees of freedom
Multiple R-squared:  0.5581,    Adjusted R-squared:  0.3687
F-statistic: 2.947 on 3 and 7 DF,  p-value: 0.108
```
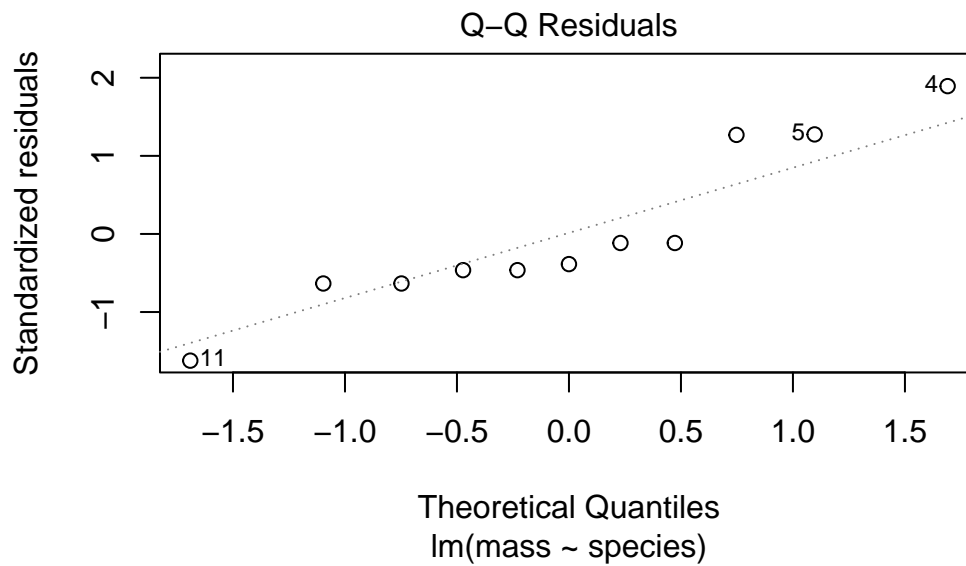
Assumptions of Factorial ANOVA

Linearity of IV to DV (Regression)

Graphical

```
plot(lm(mass ~ species, data = factorial_data),2)                              ①
```
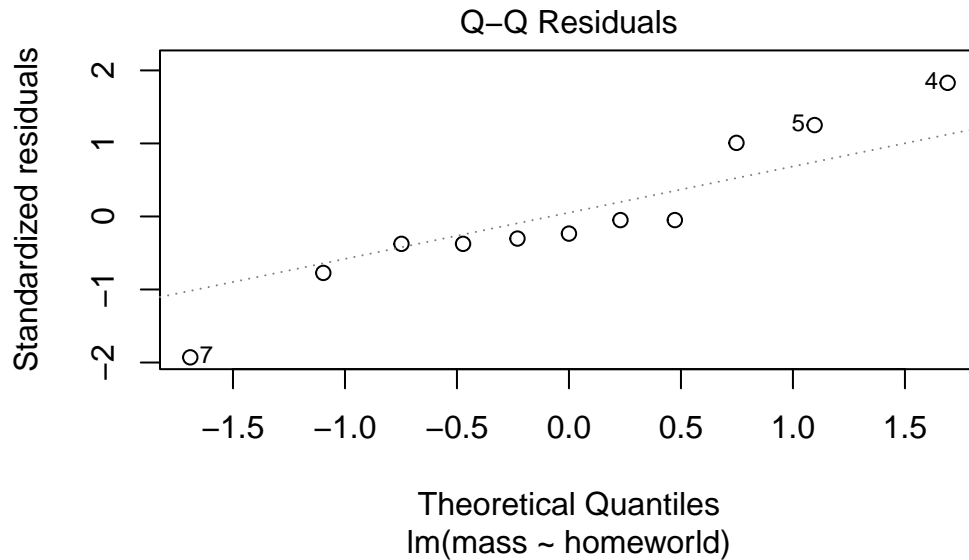
② Plot the residuals of the model using the `plot()` and `lm()` functions for mass on homeworld


```
plot(lm(mass ~ homeworld, data = factorial_data),2)                            ②
```

Q–Q Residuals

Theoretical Quantiles
lm(mass ~ homeworld)

Statistical

```
psych::describe(resid(aov_factorial))                                        ①
shapiro.test(resid(aov_factorial))                                           ②
```

① Statistical summary of the model residuals using the `describe()` function in the `psych` package
② Statistical test (Shapiro Wilk) of the residuals using the `resid()` and `shapiro.test()` functions

```
   vars  n mean    sd median trimmed   mad   min max range skew kurtosis  se
X1    1 11    0 21.54    -12   -2.06 14.08 -21.5  40  61.5 0.53     -1.4 6.5
```

```
	Shapiro-Wilk normality test

data:  resid(aov_factorial)
W = 0.87299, p-value = 0.08468
```
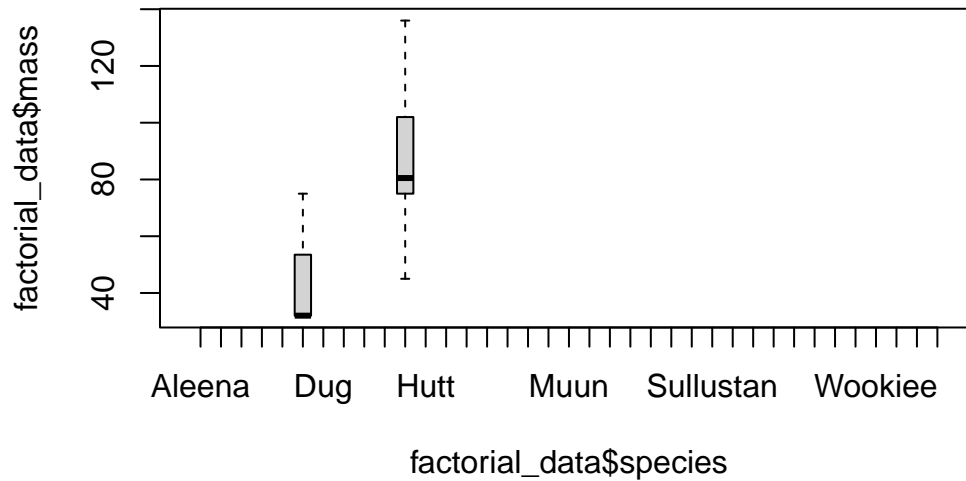
Homogeneity of Variance (Regression)

Graphical

```
boxplot(factorial_data$mass ~ factorial_data$species)                              ①
```
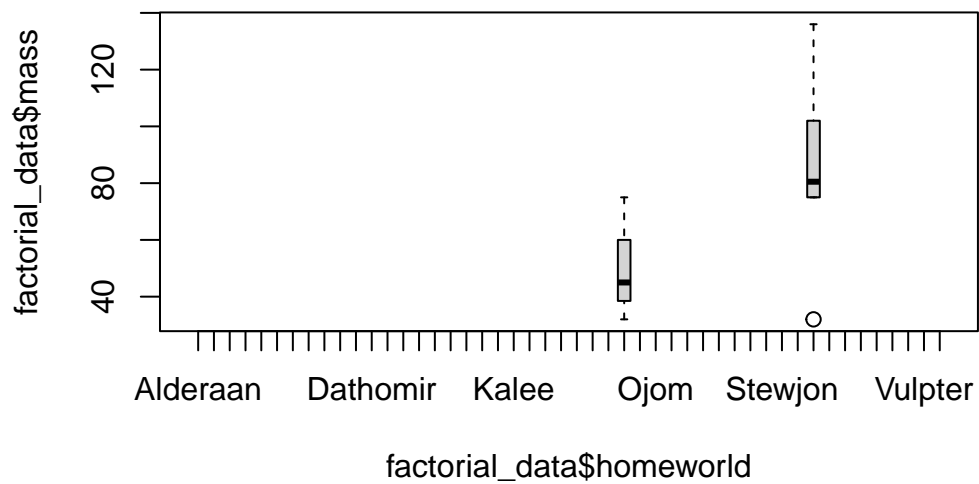
③ Fitted object of the ANCOVA
④ Plot of the residual x fitted data using the `plot()` function
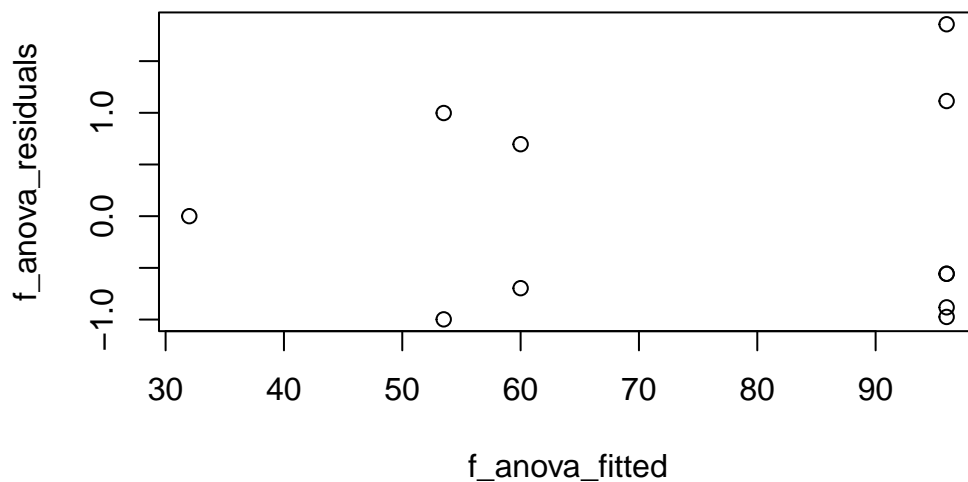


```
boxplot(factorial_data$mass ~ factorial_data$homeworld)                             ①
```

```
# Residuals
f_anova_residuals <- scale(residuals(lm(mass ~ species*homeworld, data = factorial_data))) ③
f_anova_fitted <- fitted(lm(mass ~ species*homeworld, data = factorial_data))

plot(f_anova_residuals ~ f_anova_fitted)                                                    ④
```

Statistical

```
leveneTest(factorial_data$mass ~ factorial_data$species)                              ①

leveneTest(factorial_data$mass ~ factorial_data$homeworld)                            ②

leveneTest(factorial_data$mass ~ interaction(factorial_data$homeworld,factorial_data$species)
```

① Levene test of mass x species using `leveneTest()` function
② Levene test of mass x homeworld using `leveneTest()` function
③ Levene test of mass x species/homeworld interaction using `leveneTest()` function

```
Levene's Test for Homogeneity of Variance (center = median)
      Df F value Pr(>F)
group  1  0.0985 0.7608
       9
Levene's Test for Homogeneity of Variance (center = median)
      Df F value Pr(>F)
group  1  0.1878  0.675
       9
Levene's Test for Homogeneity of Variance (center = median)
      Df F value Pr(>F)
```
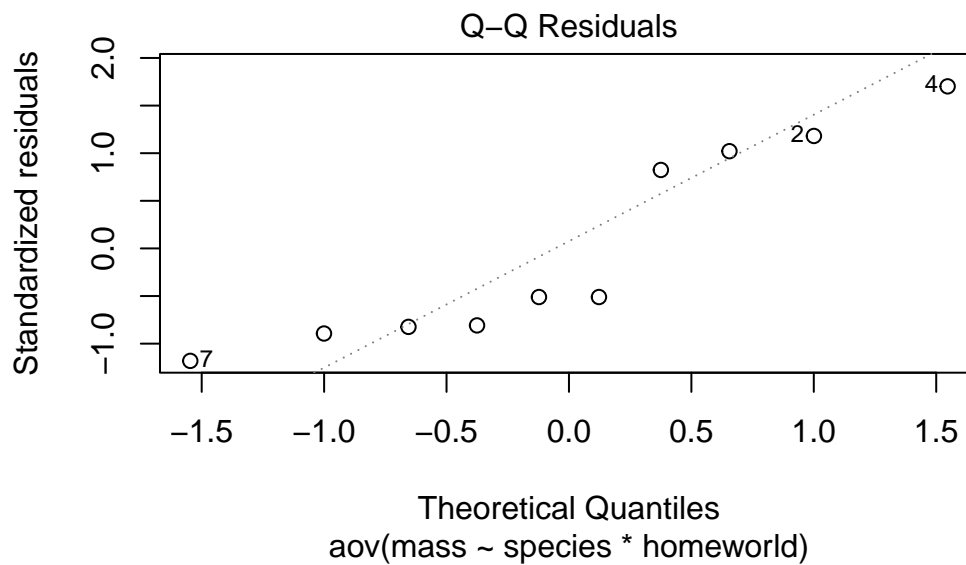
```
group  3  0.3299 0.8043
       7
```

Normallity of Residuals

Graphical

```
plot(aov_factorial,2)                                                        ①
```

① Plot ANOVA residuals using `plot()` function

Q–Q Residuals

Theoretical Quantiles
aov(mass ~ species * homeworld)

Statistical

```
fanova_residuals <- aov_factorial$residuals                                  ①

shapiro.test(fanova_residuals)                                               ②

psych::describe(fanova_residuals)                                            ③
```

① Create residual data set
② Statistical test of residual normality using `shapiro.test()` function
③ Descriptive statistics of residuals using `describe()` function in the `psych` package

```
        Shapiro-Wilk normality test

data:  fanova_residuals
W = 0.87299, p-value = 0.08468


   vars  n mean    sd median trimmed    mad   min max range skew kurtosis  se
X1    1 11    0 21.54    -12   -2.06 14.08 -21.5  40  61.5 0.53     -1.4 6.5
```