# Lesson 17 - Patient Satisfaction Surveys

LTC J. K. Starling

Last compiled on 28 February, 2022

## Background

The U.S. healthcare system increasingly uses patient satisfaction scores in payment models and quality assessment. Researchers continue to explore different variables and whether they explain variation in patient satisfaction scores. For example, Tyser, Gaffney, Zhang, and Presson (2018) explored “the association of patient satisfaction with pain, anxiety, and selfreported physical function” (Journal Bone Joint Surgery of America, 1811–1818). In this exploration you will analyze an example data set that includes measures of patient satisfaction, anxiety level, severity of illness, and age (found in Kutner, Nachtsheim, and Neter, 2004).

**Ask a research question.**

In this study, we can ask: Do any/all of age, anxiety, and severity explain variation in patient satisfaction? Does anxiety interact with age?

```
### Read data
survey <- read.table('https://raw.githubusercontent.com/jkstarling/MA376/main/patientSatisfaction.txt', header = TRUE, strin
gsAsFactors = TRUE)

setwd("C:/Users/james.starling/OneDrive - West Point/Teaching/MA376/JimsLessons/Block III/LSN17/")
survey <- read.table('patientSatisfaction.txt', header = TRUE,
                     stringsAsFactors = TRUE)
```

The data set consists of 46 patients at a certain hospital who completed a satisfaction survey upon release from the hospital. The data in PatientSatisfaction assess satisfaction on a (0–100) scale (with larger values indicating higher satisfaction), anxiety (measured as a quantitative variable with higher numbers indicating more anxiety), and age. Age is reported both as a quantitative variable and as a categorical variable (age group: younger = 20–34, middle = 35–49, older = 50–64). We will first focus on the categorical age group variable.

## Single (explanatory) variable model

1. Let’s start with the relationship between satisfaction and anxiety (score).

- What are the hypotheses?

$H_0$ : There is no association between anxiety and patient satisfaction. $H_a$ : There is an association between anxiety and patient satisfaction.

```
anxiety.lm <-lm(satisfaction~anxiety,data=survey)
summary(anxiety.lm)
```

```
##
## Call:
## lm(formula = satisfaction ~ anxiety, data = survey)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -20.369  -9.606  -1.946   9.212  31.631
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  146.449     15.304   9.569 2.55e-12 ***
## anxiety      -37.117      6.637  -5.593 1.33e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 13.33 on 44 degrees of freedom
## Multiple R-squared:  0.4155, Adjusted R-squared:  0.4022
## F-statistic: 31.28 on 1 and 44 DF,  p-value: 1.335e-06
```

```
anova(anxiety.lm)
```

| | Df<br><int> | Sum Sq<br><dbl> | Mean Sq<br><dbl> | F value<br><dbl> | Pr(>F)<br><dbl> |
|---|---|---|---|---|---|
| anxiety | 1 | 5554.913 | 5554.9131 | 31.2777 | 1.334544e-06 |
| Residuals | 44 | 7814.391 | 177.5998 | *NA* | *NA* |

2 rows

- What are the conclusions? Interpret coefficient of determination.

There is a statistically significant association between anxiety and satisfaction. The model using anxiety accounts for 41.55% of the variation in satisfaction scores.

2. We can also look at the relationship between age group and satisfaction.

- What are the hypotheses?

$H_0$ : There is no association between age group and satisfaction $H_a$ : There is an association between age group and satisfaction

```
age.lm <-lm(satisfaction~AgeGroup,data=survey)
summary(age.lm)
```

```
##
## Call:
## lm(formula = satisfaction ~ AgeGroup, data = survey)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -23.947 -10.496   1.955  10.857  27.857
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)        74.947      2.869  26.123  < 2e-16 ***
## AgeGroupage35-49  -19.805      3.960  -5.002 1.01e-05 ***
## AgeGroupage50-64  -33.281      5.856  -5.683 1.06e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 12.51 on 43 degrees of freedom
## Multiple R-squared:  0.497,  Adjusted R-squared:  0.4736
## F-statistic: 21.24 on 2 and 43 DF,  p-value: 3.836e-07
```

```
anova(age.lm)
```

| | Df<br><int> | Sum Sq<br><dbl> | Mean Sq<br><dbl> | F value<br><dbl> | Pr(>F)<br><dbl> |
|---|---|---|---|---|---|
| AgeGroup | 2 | 6644.452 | 3322.2261 | 21.24295 | 3.835654e-07 |
| Residuals | 43 | 6724.852 | 156.3919 | *NA* | *NA* |

2 rows

```
Anova(age.lm, type = 3)
```

| | Sum Sq<br><dbl> | Df<br><dbl> | F value<br><dbl> | Pr(>F)<br><dbl> |
|---|---|---|---|---|
| (Intercept) | 106725.053 | 1 | 682.42055 | 5.155061e-28 |
| AgeGroup | 6644.452 | 2 | 21.24295 | 3.835654e-07 |
| Residuals | 6724.852 | 43 | *NA* | *NA* |

3 rows

- How do we interpret the coefficients?

The intercept is the predicted average satisfaction for the young age group and the coefficients for the other indicator variables are the predicted change in the average satisfaction from the younger age group.

- Interpret coefficient of determination.

The model using age group accounts for 49.7% of the variation in satisfaction.

## Two-variable model

3. Maybe we should use a model with both explanatory variables (anxiety and AgeGroup) to explain customer satisfaction. Create two models `both.lm` and `both.lm2` for indicator coding and effect coding, respectively.

- What are the hypotheses?

$H_0$ : There is no association between satisfaction and anxiety after adjusting for age group.
$H_a$ : There is an association between satisfaction and anxiety after adjusting for age group.

```
both.lm<-lm(satisfaction~anxiety+AgeGroup,data=survey)
summary(both.lm)
```

```
##
## Call:
## lm(formula = satisfaction ~ anxiety + AgeGroup, data = survey)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -20.525  -7.600   1.469   7.496  19.975
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)       129.041     12.563  10.272 5.01e-13 ***
## anxiety           -25.007      5.700  -4.387 7.57e-05 ***
## AgeGroupage35-49  -15.549      3.457  -4.498 5.33e-05 ***
## AgeGroupage50-64  -24.441      5.305  -4.607 3.77e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 10.48 on 42 degrees of freedom
## Multiple R-squared:  0.6551, Adjusted R-squared:  0.6304
## F-statistic: 26.59 on 3 and 42 DF,  p-value: 8.457e-10
```

```
survey2 <- survey
contrasts(survey2$AgeGroup) = contr.sum
contrasts(survey2$AgeGroup)
```

```
##          [,1] [,2]
## age20-34    1    0
## age35-49    0    1
## age50-64   -1   -1
```

```
both.lm2 = lm(satisfaction~anxiety+AgeGroup, data = survey2)
summary(both.lm2)
```

```
##
## Call:
## lm(formula = satisfaction ~ anxiety + AgeGroup, data = survey2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -20.525  -7.600   1.469   7.496  19.975
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  115.711     13.447   8.605 8.15e-11 ***
## anxiety      -25.007      5.700  -4.387 7.57e-05 ***
## AgeGroup1     13.330      2.485   5.365 3.23e-06 ***
## AgeGroup2     -2.219      2.236  -0.992    0.327
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 10.48 on 42 degrees of freedom
## Multiple R-squared:  0.6551, Adjusted R-squared:  0.6304
## F-statistic: 26.59 on 3 and 42 DF,  p-value: 8.457e-10
```

- What are the conclusions? Interpret the coefficients.

129.041, represents the predicted satisfaction of the younger age group with an anxiety score of 0; -25.007, is the rate of change in the predicted satisfaction as the anxiety score increases by 1 unit (but the age group stays the same). The rate of change is adjusted by -15.55 for ages 35-49 and -24.44 for ages 50-64.

- Interpret coefficient of determination and compare to $R^2$ of the first two models.

The model using both explanatory variables (anxiety and AgeGroup) accounts 65.55% of the variation in satisfaction (compared to 41.55% and 49.7% of the variation in satisfaction (for anxiety and AgeGroup, respectively).

# Comparing Models

Certainly our $R^2$ increased, but if we know anxiety do we gain anything by knowing age group? That is, we want to do a single test that evaluates whether the variable age group (the collection of indicator terms) is significant or not. Note that when we have multiple categories, if we test each category separately we are inflating the Type I error. For this we go to the ANOVA:

```
Anova(both.lm2, type=3)
```

| | Sum Sq <dbl> | Df <dbl> | F value <dbl> | Pr(>F) <dbl> |
|---|---|---|---|---|
| (Intercept) | 8129.996 | 1 | 74.04235 | 8.149861e-11 |
| anxiety | 2113.169 | 1 | 19.24527 | 7.574813e-05 |
| AgeGroup | 3202.708 | 2 | 14.58402 | 1.549710e-05 |
| Residuals | 4611.683 | 42 | *NA* | *NA* |

4 rows

4. The final question we want to ask is whether this model is better than the model with only anxiety. (How is this different to the overall F stat information we get from each model?)

- Since we have we can statistically compare the two models. What do I mean by nested models? The *smaller model* has removed some coefficients from the larger model. Assuming our validity conditions are met, we can form the F statistic on page 354 and conduct a partial F test to compare two nested models. What are the hypotheses for the partial F test?

Null- In the population, the extra variables have no effect on the outcome.
Alternative-At least one of the extra variables has an effect on the outcome
OR Null: The full model is not significantly better than the reduced model.
Alternative: The full model significantly better than the reduced model.

```
partF <- ((0.6555-0.4155)/ 2)/((1-0.6555)/42); partF
```

```
## [1] 14.6299
```

```
# partF <- ((summary(both.lm)$r.squared-summary(anxiety.lm)$r.squared) / 2)/ ((1-summary(both.lm)$r.squared)/42)
anova(anxiety.lm,both.lm2)
```

| | Res.Df <dbl> | RSS <dbl> | Df <dbl> | Sum of Sq <dbl> | F <dbl> | Pr(>F) <dbl> |
|---|---|---|---|---|---|---|
| 1 | 44 | 7814.391 | *NA* | *NA* | *NA* | *NA* |
| 2 | 42 | 4611.683 | 2 | 3202.708 | 14.58402 | 1.54971e-05 |

2 rows

```
anova(anxiety.lm,both.lm)
```

| | Res.Df <dbl> | RSS <dbl> | Df <dbl> | Sum of Sq <dbl> | F <dbl> | Pr(>F) <dbl> |
|---|---|---|---|---|---|---|
| 1 | 44 | 7814.391 | *NA* | *NA* | *NA* | *NA* |
| 2 | 42 | 4611.683 | 2 | 3202.708 | 14.58402 | 1.54971e-05 |

2 rows

```
anova(both.lm,anxiety.lm) # does order matter? Should have the "smaller model" listed first.
```

| | Res.Df <dbl> | RSS <dbl> | Df <dbl> | Sum of Sq <dbl> | F <dbl> | Pr(>F) <dbl> |
|---|---|---|---|---|---|---|
| 1 | 42 | 4611.683 | *NA* | *NA* | *NA* | *NA* |
| 2 | 44 | 7814.391 | -2 | -3202.708 | 14.58402 | 1.54971e-05 |

2 rows

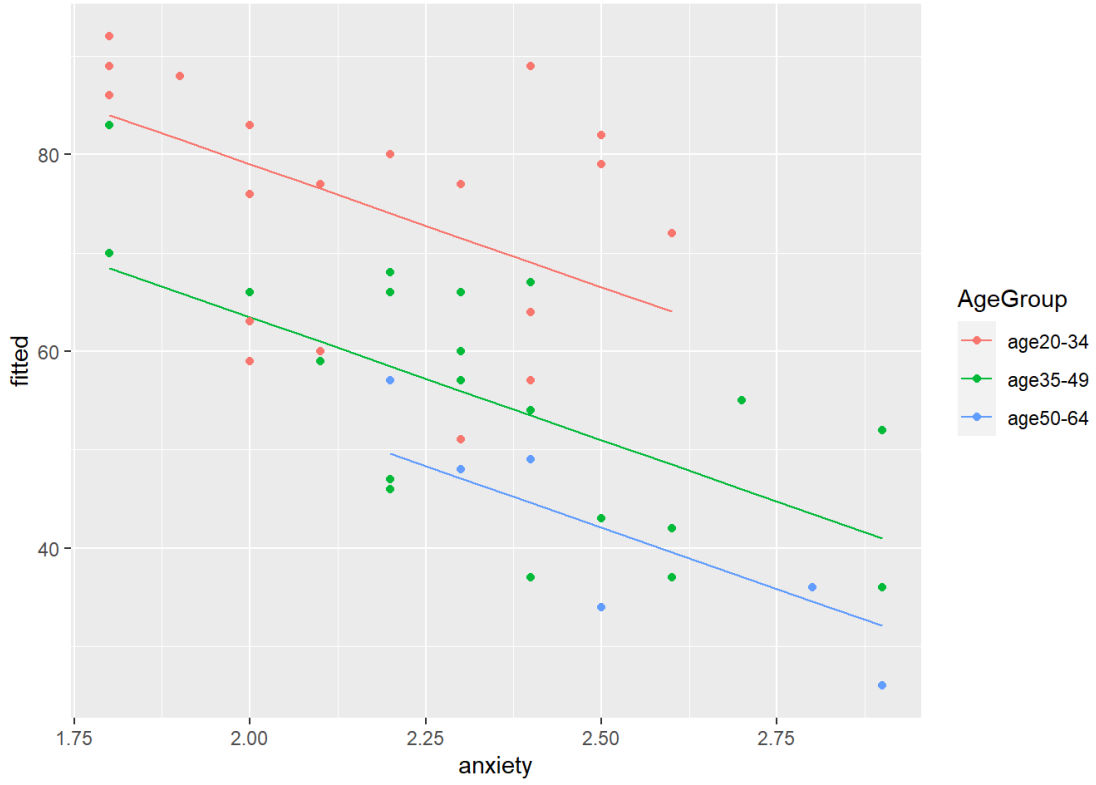
- Note: for the partial F test we aren't concerned with types of Sums of Squares as we are, by default, conducting a conditional test.
- What is our conclusion about which model is better?

*Because the F statistic is statistically significant, our conclusion is that the model with age group is preferred.*

- Lets look what the indicator and effect coding models look like in a plot.

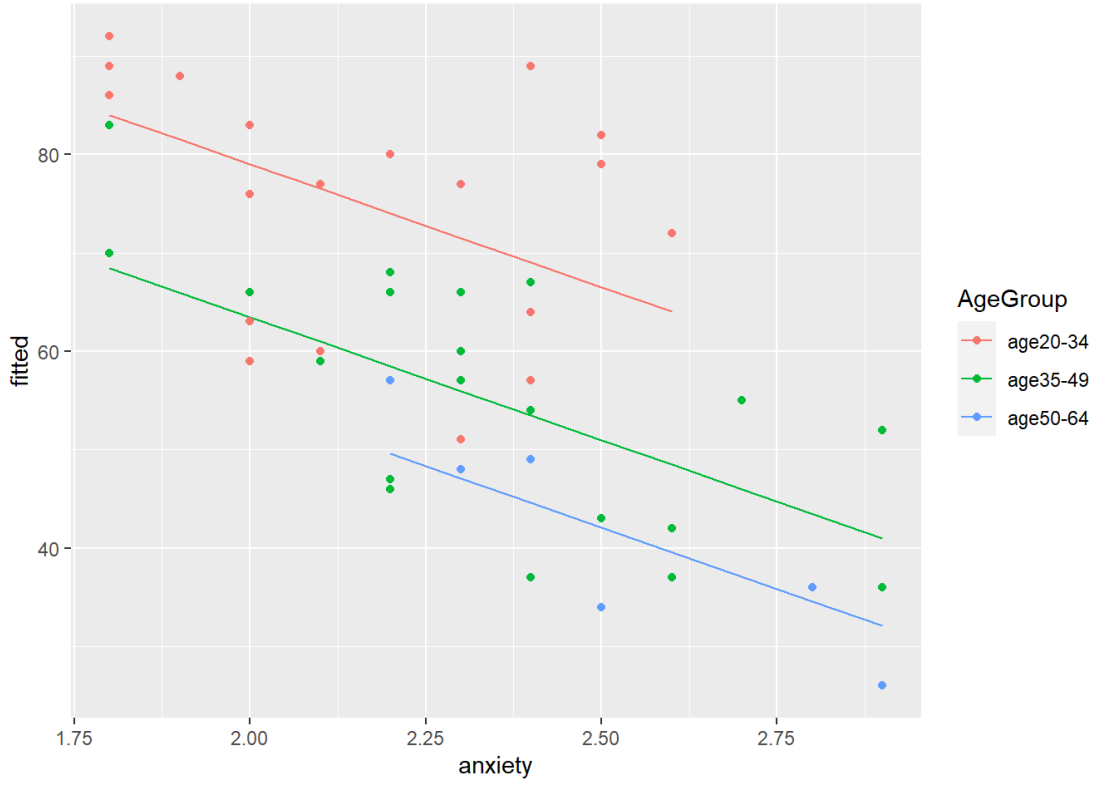
```
survey3 = survey %>%
  mutate(fitted = predict(both.lm, newdata = .),
         residuals = satisfaction - fitted)

survey3 %>%
  ggplot(aes(x = anxiety, y = fitted, color = AgeGroup)) +
  geom_line() +
  geom_point(aes(y = satisfaction))
```




```
survey4 = survey %>%
  mutate(fitted = predict(both.lm2, newdata = .),
         residuals = satisfaction - fitted)

survey4 %>%
  ggplot(aes(x = anxiety, y = fitted, color = AgeGroup)) +
  geom_line() +
  geom_point(aes(y = satisfaction))
```




Note: our model for all three groups have the same slope.

## Interactions

5. How do we model the effect of anxiety changing across the age group categories?

*Constructing a model that considers the interactions.*

- What are the hypotheses?

*Null: There is no interaction between anxiety and age group, after adjusting for anxiety and age group in the population. Alternative: There is an interaction between anxiety and age group, after adjusting for anxiety and age group in the population.*

We can then test if this model is preferred to a model without interactions by:

```
inter.lm<-lm(satisfaction~anxiety*AgeGroup,data=survey)
summary(inter.lm)
```

```
##
## Call:
## lm(formula = satisfaction ~ anxiety * AgeGroup, data = survey)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -21.818  -7.341   1.187   8.061  17.738
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)               108.605     20.860   5.206 6.12e-06 ***
## anxiety                   -15.559      9.579  -1.624    0.112
## AgeGroupage35-49           13.404     27.877   0.481    0.633
## AgeGroupage50-64           27.464     47.494   0.578    0.566
## anxiety:AgeGroupage35-49  -13.097     12.393  -1.057    0.297
## anxiety:AgeGroupage50-64  -21.951     19.398  -1.132    0.265
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 10.51 on 40 degrees of freedom
## Multiple R-squared:  0.6694, Adjusted R-squared:  0.6281
## F-statistic:  16.2 on 5 and 40 DF,  p-value: 1.016e-08
```

```
inter.lm2<-lm(satisfaction~anxiety*AgeGroup,data=survey2)
Anova(inter.lm, type = 3)
```

| | Sum Sq <dbl> | Df <dbl> | F value <dbl> | Pr(>F) <dbl> |
|---|---|---|---|---|
| (Intercept) | 2994.89659 | 1 | 27.1050549 | 6.123321e-06 |
| anxiety | 291.53478 | 1 | 2.6385106 | 1.121565e-01 |
| AgeGroup | 47.41895 | 2 | 0.2145806 | 8.078026e-01 |
| anxiety:AgeGroup | 191.99594 | 2 | 0.8688214 | 4.272101e-01 |
| Residuals | 4419.68717 | 40 | *NA* | *NA* |

5 rows

```
anova(both.lm,inter.lm)
```

| | Res.Df <dbl> | RSS <dbl> | Df <dbl> | Sum of Sq <dbl> | F <dbl> | Pr(>F) <dbl> |
|---|---|---|---|---|---|---|
| 1 | 42 | 4611.683 | *NA* | *NA* | *NA* | *NA* |

| | Res.Df | RSS | Df | Sum of Sq | F | Pr(>F) |
|---|---|---|---|---|---|---|
| | <dbl> | <dbl> | <dbl> | <dbl> | <dbl> | <dbl> |
| 2 | 40 | 4419.687 | 2 | 191.9959 | 0.8688214 | 0.4272101 |

2 rows

Since the interaction is not statistically significant we should use the model without the interaction term.

## Bonferonni corrections

We can get 95% CI for each of our $\beta$ terms in this model, though might we want to adjust these CIs using what is called Bonferonni Corrections so as not to inflate Type I error. This technique uses $\alpha/k$ in lieu of $\alpha$ where $k$ is the number of comparisons or tests being performed. Here we have 3 Confidence intervals, so Bonferonni corrections would say to use $.05/3 = .016$, or in order to guarantee an overall $\alpha = 0.05$, we should use $1 - 0.016 = 99.84\%$ CI instead of $95\%$ CI.

```
confint(inter.lm)
```

```
##                               2.5 %    97.5 %
## (Intercept)                66.44430 150.76549
## anxiety                   -34.91907   3.80019
## AgeGroupage35-49          -42.93762  69.74489
## AgeGroupage50-64          -68.52502 123.45256
## anxiety:AgeGroupage35-49  -38.14536  11.95081
## anxiety:AgeGroupage50-64  -61.15620  17.25362
```

```
confint(inter.lm,level=0.9984)
```

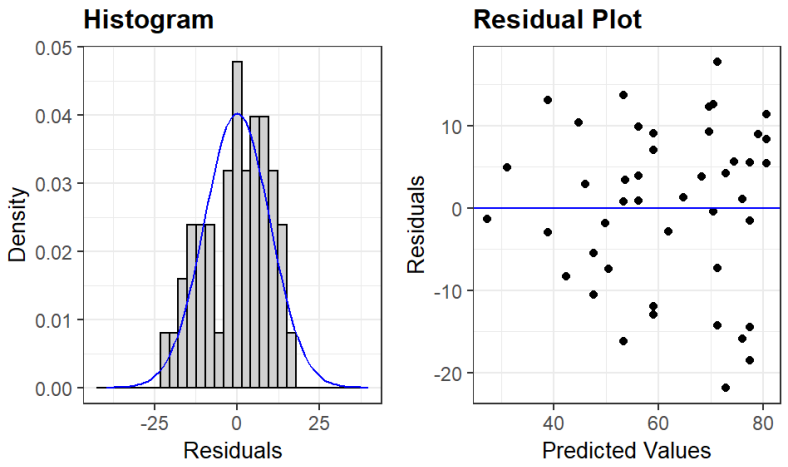
```
##                               0.08 %   99.92 %
## (Intercept)                37.96562 179.24417
## anxiety                   -47.99613  16.87725
## AgeGroupage35-49          -80.99506 107.80233
## AgeGroupage50-64         -133.36361 188.29116
## anxiety:AgeGroupage35-49  -55.06487  28.87032
## anxiety:AgeGroupage50-64  -87.63836  43.73578
```

What can we conclude if we use this technique? Is it different to the unadjusted CI?

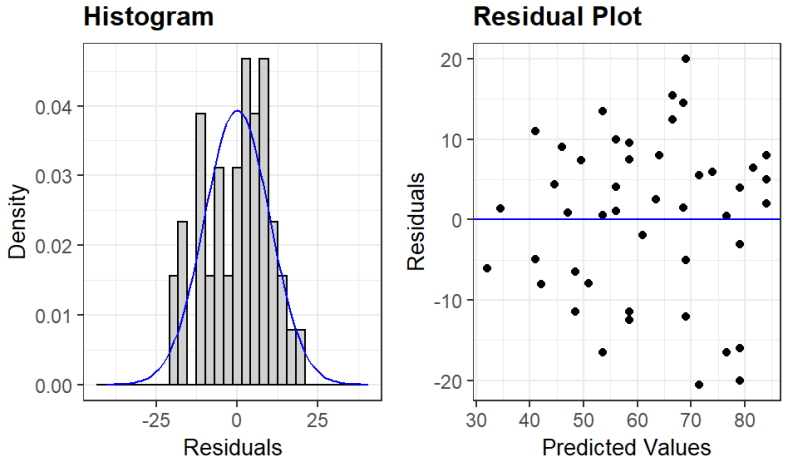
*Bonferonni Corrections is a very very conservative approach.*

6. Now that we have a final model what should we check before we can trust our conclusions?

```
# Validity conditions
resid_panel(inter.lm, plots = c('hist', 'resid'))
```




```
resid_panel(both.lm2, plots = c('hist', 'resid'))
```




## References

Nathan Tintle et al.(2019). Intermediate Statistical Investigations for U.S. Military Academy at West Point.