

# Exam 3

Kevin Linares

2024-12-09

---

Please answer whether the following statements (Questions 1-3) are True or False.

1. The Box-Cox transformation can help you identify what transformation of a predictor might benefit the model.

- FALSE

2. A limitation of leave one out (LOO) cross-validation is that it can be computationally intensive.

- TRUE

3. We can use cross-validation methods to estimate the test error of a model.

- TRUE
- 

4. You fit a model and are assessing for the presence of collinearity by estimating VIF for the predictors. The VIF for one of the predictors is 212.678. Roughly how much higher would you expect the standard error of this predictor to be than if there were no collinearity (you must show your work)?

- We can work out how much larger the standard error for the collinear predictor is compared to when not having collinearity by taking the square root of the VIF, and find that this predictor's SE is about 14.5 times larger than not having collinearity.

```
vif <- 212.648  
sqrt(vif)
```

```
[1] 14.58246
```

**5. List 3 ways (discussed in class) that variable selection methods (forward/backward/stepwise) can be problematic.**

- One problem with these variable selection methods is that there is no correction for p-values which may lead to a spurious significant result. A second problem is that we may estimate a biased coefficient in that  $\hat{\beta}$  is not equal to the population  $\beta$ . This would lead to inaccurate findings on the relationship between the predictors and the outcome variable. Finally, these methods are sensitive to small changes in the data, which can lead to different model selections for different runs on slightly modified data.

**6. “PATH\_FINAL.csv” is provided on Canvas.**

- Read in the data (“PATH\_FINAL.csv”).
- The data comes from a project called “Positive Attitudes Towards Health” in 2017 that targeted persons who inject drugs in Southeast Michigan.
- This dataset includes 409 cases and the following variables:
  - SAMPLEID: ID of respondents
  - AGE: Age in years
  - MALE: 1. Male; 0. Female
  - BLACK: 1. Race black; 0. Other race than black
  - EDUC: 1. <High school (HS); 2. HS Education; 3. >HS Education
  - LIFESAT: Life satisfaction score summarized from 5 questions; Higher scores mean higher satisfaction
  - INJECTMULTIPLE: 1. Inject multiple drugs; 0. Injects only one drug
  - AGE\_DIFF: Age one feels (from a question, “How old do you feel?”) minus actual age. Positive values mean feeling older than actual age. Negative values mean feeling younger than actual age.

```
glimpse( path_final <- read_csv(
  "~/repos/UMD_classes_code/stat_mod_ML_a_SURV615/exams/PATH_FINAL.csv") |>
  dplyr::select(-1, -SAMPLEID))
```

Rows: 409

Columns: 7

```
$ AGE          <dbl> 58, 57, 62, 34, 67, 65, 27, 50, 46, 61, 62, 27, 36, 67, ~
$ MALE         <dbl> 1, 0, 1, 1, 1, 1, 0, 0, 1, 1, 1, 1, 0, 1, 1, 0, 1, 0~
$ BLACK        <dbl> 1, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 0, 1, 0~
$ EDUC         <chr> "1.<HS", "2.HS", "2.HS", "3.>HS", "1.<HS", "1.<HS", "3.~
$ LIFESAT      <dbl> 23, 24, 25, 12, 25, 24, 6, 11, 12, 27, 20, 18, 17, 10, ~
$ INJECTMULTIPLE <dbl> 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 1, 0, 0, 1, 1, 1, 1, 1~
$ AGE_DIFF     <dbl> 0, 0, -12, 11, 0, 0, 20, -10, 0, -16, 0, 0, 0, -2, 9, 1~
```

## 6.1 Regress AGE\_DIFF on the remaining variables (except for SAMPLEID). Examine collinearity of the predictors and report your conclusions?

- We compute the conditional number kappa, which measures the relative sizes of the eigenvalues where  $k > 30$  is considered large. Our kappa value of 182 is 6 times larger than 30, but can only tell us that at least one of the eigenvalues is small relative to the rest., Therefore, we examine other conditional indices  $\eta_j$  because they indicate whether more than just one independent linear combination is to blame. We see that five of seven conditional indices are above 30, suggesting several independent linear combinations are present in this model. The variance inflation factor (VIF) assesses multicollinearity, which is the ratio of the variance of  $\hat{\beta}_j$  : when fitting the full model divided by the variance of  $\hat{\beta}_j$  if fit on its own. The smallest possible value is 1 indicating absence of collinearity. A rule of thumb of a VIF of over 5 is an indication of a problematic amount of collinearity. VIF ranged between 1.1 and 2.1, yet all VIF values are less than 5 so there does not appear to be any issues of collinearity.

```
summ(mod <- lm(AGE_DIFF ~ ., path_final))
```

MODEL INFO:

Observations: 409

Dependent Variable: AGE\_DIFF

Type: OLS linear regression

MODEL FIT:

$F(7,401) = 13.17$ ,  $p = 0.00$

$R^2 = 0.19$

Adj.  $R^2$  = 0.17

Standard errors:OLS

	Est.	S.E.	t val.	p
(Intercept)	17.99	3.76	4.78	0.00
AGE	-0.29	0.08	-3.62	0.00
MALE	1.03	1.61	0.64	0.52
BLACK	-5.76	2.16	-2.66	0.01
EDUC2.HS	1.19	1.77	0.67	0.50
EDUC3.>HS	-1.15	1.91	-0.60	0.55
LIFESAT	-0.19	0.10	-2.00	0.05
INJECTMULTIPLE	-2.59	1.97	-1.32	0.19

```
#----- handy functions -----
kappa_fun <- function(mod_name){
  # model matrix w/o intercept
  x_mod = model.matrix(mod_name)[,-1]
  # matrix multiplication
  e=eigen(t(x_mod)%*%x_mod)
  print("Inspect wide range in eigenvalues")
  print(e$val)
  # calculate kappa values
  print("Inspect if Kappa conditional number value > 30")
  kappa_mod = sqrt(max(e$val)/min(e$val))
  print(kappa_mod)
  print("Inspect condition index, of at least one linear combination")
  nu = sqrt(max(e$val)/e$val)
  print(nu) }

kappa_fun(mod)
```

```
[1] "Inspect wide range in eigenvalues"
[1] 1327301.72118 25553.47453 131.99333 90.46271 72.69003
[6] 52.77881 39.87940
[1] "Inspect if Kappa conditional number value > 30"
[1] 182.436
[1] "Inspect condition index, of at least one linear combination"
[1] 1.000000 7.207088 100.278734 121.129527 135.128637 158.582414 182.435986
```

```
ols_vif_tol(mod) |> dplyr::select(Variables, VIF) |> arrange(desc(VIF))
```

	Variables	VIF
1	AGE	2.132141
2	BLACK	2.036807
3	EDUC3.>HS	1.338690
4	EDUC2.HS	1.333458
5	INJECTMULTIPLE	1.266157
6	MALE	1.115114
7	LIFESAT	1.084497

**6.2 From the regression model in Q6.1, select predictors using AIC-based step-wise regression. Report the changes in AIC and the selected predictors through all steps.**

- The model with Age, Black, and Lifesat had the lowest AIC (2209.04). Given the lowest AIC value we pick this model as our “best” model, yet these AIC changes across models are small increments. For instance, for the model with AGE + MALE + BLACK + EDUC + LIFESAT + INJECTMULTIPLE the AIC was 2212.19, followed by the model with AGE + MALE + BLACK + LIFESAT + INJECTMULTIPLE AIC= 2210.42, and model with AGE + BLACK + LIFESAT + INJECTMULTIPLE AIC = 2209.19 which may not be significant enough to warrant one model over another.

```
back.mod <- stepAIC(mod, direction = "both")
```

Start: AIC=2212.91

AGE\_DIFF ~ AGE + MALE + BLACK + EDUC + LIFESAT + INJECTMULTIPLE

	Df	Sum of Sq	RSS	AIC
- EDUC	2	324.26	88329	2210.4
- MALE	1	89.24	88094	2211.3
- INJECTMULTIPLE	1	381.60	88386	2212.7
<none>			88005	2212.9
- LIFESAT	1	874.15	88879	2215.0
- BLACK	1	1555.10	89560	2218.1
- AGE	1	2869.27	90874	2224.0

Step: AIC=2210.42

AGE\_DIFF ~ AGE + MALE + BLACK + LIFESAT + INJECTMULTIPLE

	Df	Sum of Sq	RSS	AIC
- MALE	1	166.04	88495	2209.2
- INJECTMULTIPLE	1	417.78	88747	2210.3
<none>			88329	2210.4
- LIFESAT	1	918.13	89247	2212.7
+ EDUC	2	324.26	88005	2212.9
- BLACK	1	1459.70	89788	2215.1
- AGE	1	2835.71	91164	2221.3

Step: AIC=2209.19

AGE\_DIFF ~ AGE + BLACK + LIFESAT + INJECTMULTIPLE

	Df	Sum of Sq	RSS	AIC
- INJECTMULTIPLE	1	402.66	88897	2209.0
<none>			88495	2209.2
+ MALE	1	166.04	88329	2210.4
+ EDUC	2	401.05	88094	2211.3
- LIFESAT	1	973.52	89468	2211.7
- BLACK	1	1472.34	89967	2213.9
- AGE	1	2680.17	91175	2219.4

Step: AIC=2209.04

AGE\_DIFF ~ AGE + BLACK + LIFESAT

	Df	Sum of Sq	RSS	AIC
<none>			88897	2209.0
+ INJECTMULTIPLE	1	402.7	88495	2209.2
+ MALE	1	150.9	88747	2210.3
+ EDUC	2	437.7	88460	2211.0
- LIFESAT	1	1070.0	89967	2211.9
- BLACK	1	1809.0	90706	2215.3
- AGE	1	3226.3	92124	2221.6

```
back.mod$anova
```

Stepwise Model Path

Analysis of Deviance Table

Initial Model:

AGE\_DIFF ~ AGE + MALE + BLACK + EDUC + LIFESAT + INJECTMULTIPLE

Final Model:

AGE\_DIFF ~ AGE + BLACK + LIFESAT

	Step	Df	Deviance	Resid. Df	Resid. Dev	AIC
1				401	88004.52	2212.914
2	- EDUC	2	324.2558	403	88328.78	2210.418
3	- MALE	1	166.0385	404	88494.82	2209.186
4	- INJECTMULTIPLE	1	402.6610	405	88897.48	2209.043

### 6.3 Based on the model selected in Q6.2, check whether the OLS assumptions are met. What do you conclude and what are the implications?

- We plot  $\hat{\epsilon}$  against  $\hat{y}$  and observe that there is a discernible pattern funnel-shape observed on the plot - more tightly clustered on the right half and more dispersed and spread apart on the left half of the plot. It is suggestive of non-constant variance or heteroscedasticity thus violating the constant variance assumption. We can speculate that some values do not always center around 0 despite our residual mean = 0, and the variation in values varies depending on the fitted values.
  - Zero mean error ( $E(\epsilon_i) = 0 \text{ for } i = 1, \dots, n$ ) assumption means the average error value should be 0 and determined by random chance. Non-zero mean error can be a result of error dependence, non-normality, or non-constant error variance.
  - Constant error variance ( $V(\epsilon_i) = \sigma^2$ , aka homoscedasticity) assumption refers to the residual variance being constant meaning that the error term does not vary much as the predictor variance changes. Generally, we expect the variance of the data points to be about the same for all data points. If we do not have constant variance, meaning the residuals are heteroscedastic, the error terms may increase with the value of the response. Graphically this shows up as a funnel shape or some pattern in the residual plot. Ignoring this assumption leads to biased standard errors, confidence intervals, predictions, and hypothesis testing.
- We further examine the constant variance assumption by plotting  $\sqrt{|\epsilon|}$  against  $\hat{y}$  (Faraway, 2014, p. 75) and the density and homogeneity of variance plots also shows evidence of non-constant variance. We can also see that higher fitted values have residuals higher than 0, while some mid point values are below 0. Therefore, we are seeing evidence for heteroscedasticity.
- Our residual histogram shows a bimodal distribution around the midpoint. We finally see that our qq-plot at the bottom shows several observations outside of what is expected.
  - Additionally, we used the Shapiro-Wilk normality test and failed to reject the null hypothesis that the residuals are normally distributed in favor of the alternative hypothesis that they are probably not normal. We conclude that both graphically

and statistically the assumption of error normality is violated and place more weight on the statistical test over our graphics.

- The Shapiro-Wilk test is a formal test for normality as the Null hypothesis being tested is that residuals come from a normal distribution. However, as the sample size increases mild deviations from non-normality may appear, suggesting that the likelihood of rejecting the Null is more likely as the sample size increases. On the other hand, with small samples there may not be enough power to reject the Null hypothesis, meaning that the test could fail to detect non-normal distributions. Finally, we have to take caution using these tests of normality since the p-value is not very helpful as an indicator of how to approach the violation of non-normal residuals. Therefore, also examining QQ-plots is recommended for a visual inspection of normality of the residuals. Of course, this violation, according to Gelman 2021, really only matters if using the model coefficients to make predictions.
- We perform a t-test and fail to reject the null hypothesis that the mean of our  $\hat{\epsilon}$  is equal to 0, our confidence interval contains 0, suggesting that our residual mean is statistically close to 0.
- Finally, we examine statistically equal variances for fitted values greater than 0 and values less than or equal to 0. The F-test for equal variances we see that the variances in the residuals are not equal based on that the confidence intervals (95% [1.5, 2.9]) does not contain 0, meaning that we reject the null that true ratio of variances = 1 in favor of the alternative hypothesis = > unequal variances, according to our F-test. We find that the error term varies as the predictor variance changes, in other words we do not observe variance of the data points to be about the same across all data points.
- In OLS, when these assumptions are violated, the implications are that estimate accuracy from model results become bias leading to incorrect inferences such as p-values or confidence intervals.

```
summ(mod_2 <- lm(AGE_DIFF ~ AGE + BLACK + LIFESAT, path_final))
```

MODEL INFO:

Observations: 409

Dependent Variable: AGE\_DIFF

Type: OLS linear regression

MODEL FIT:

F(3,405) = 29.38, p = 0.00

R<sup>2</sup> = 0.18

Adj. R<sup>2</sup> = 0.17

Standard errors:OLS

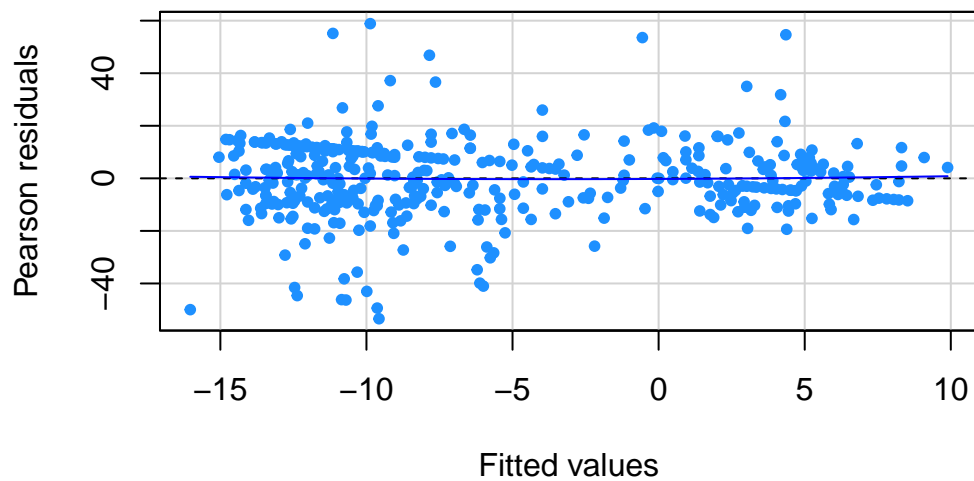


	Est.	S.E.	t val.	p
(Intercept)	17.60	3.49	5.04	0.00
AGE	-0.30	0.08	-3.83	0.00
BLACK	-6.04	2.10	-2.87	0.00
LIFESAT	-0.21	0.09	-2.21	0.03

```
summary(resid(mod_2))
```

```
      Min.   1st Qu.   Median     Mean   3rd Qu.     Max.
-53.4293  -8.2250   0.1924   0.0000   9.8091  58.8688
```

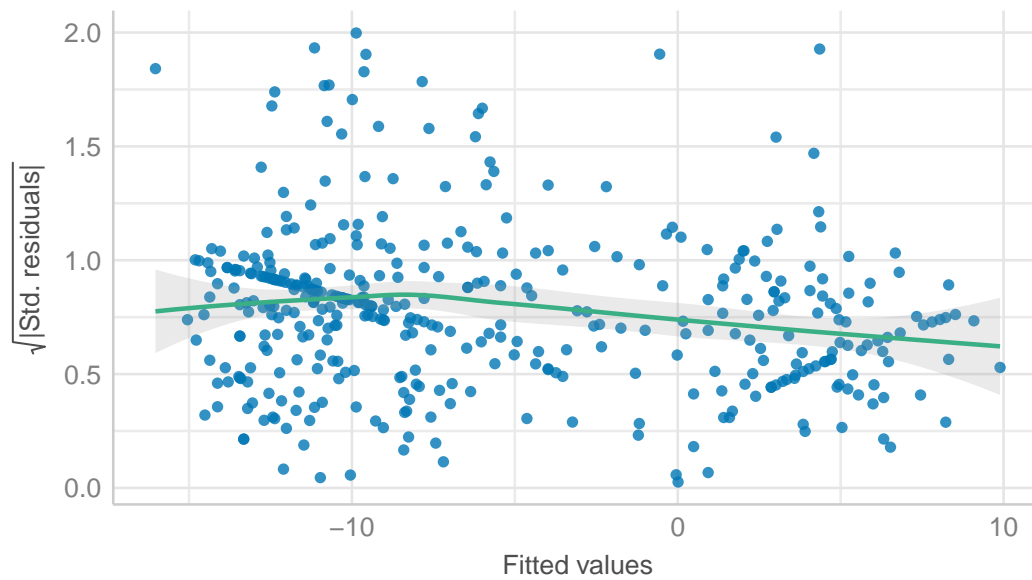
```
# residual plot
residualPlot(mod_2, pch=20, col= "dodgerblue")
```



```
# check for homogeneity: ref line is not flat
plot(check_heteroskedasticity(mod_2))
```

## Homogeneity of Variance

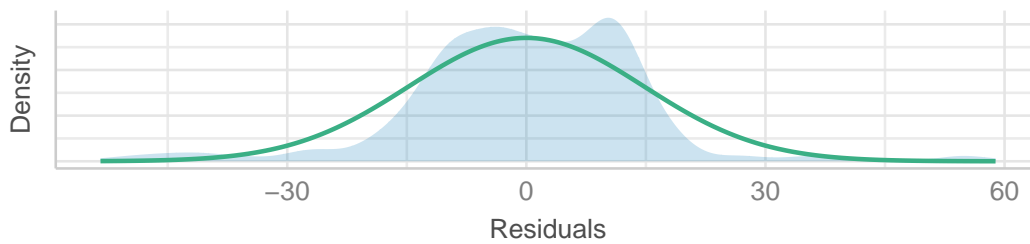
Reference line should be flat and horizontal



```
grid.arrange(  
  # check for normality of residuals,  
  plot(check_normality(mod_2), type="density"),  
  plot(check_normality(mod_2), type="qq")  
)
```

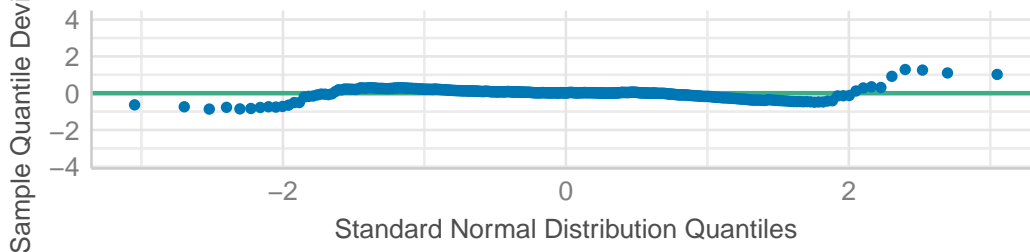
## Normality of Residuals

Distribution should be close to the normal curve



## Normality of Residuals

Dots should fall along the line



```
# shapiro test  
shapiro.test(resid(mod_2))
```

Shapiro-Wilk normality test

```
data: resid(mod_2)  
W = 0.94287, p-value = 0.00000000001915
```

```
t.test(resid(mod_2))
```

One Sample t-test

```
data: resid(mod_2)  
t = 0.00000000000000046218, df = 408, p-value = 1  
alternative hypothesis: true mean is not equal to 0  
95 percent confidence interval:  
-1.4348 1.4348  
sample estimates:  
mean of x  
0.000000000000000337338
```

```
var.test(resid(mod_2)[fitted(mod)<=0], resid(mod)[fitted(mod_2)>0])
```

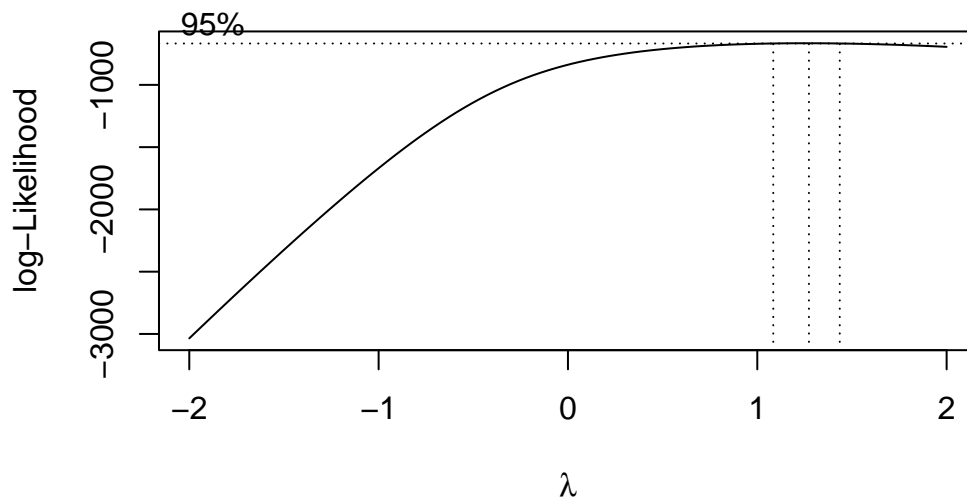
F test to compare two variances

```
data: resid(mod_2)[fitted(mod) <= 0] and resid(mod)[fitted(mod_2) > 0]
F = 2.1168, num df = 297, denom df = 112, p-value = 0.000008137
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 1.537923 2.850857
sample estimates:
ratio of variances
      2.116831
```

#### 6.4 Would a transformation of the outcome variable help the model selected in Q6.3? If so, which transformation?

- The Lambda values is close to 1; therefore, we may wish to leave the outcome variable untransformed.

```
boxcox(mod_2 <- lm((
  AGE_DIFF - min(AGE_DIFF)) +1 ~
  AGE + BLACK + LIFESAT, path_final))
```



**6.5 Focus on the slope coefficient for AGE in the model from Q6.3. Report its standard error and 95% confidence interval using the following three approaches. What do you observe?**

**a. OLS**

- We observe the age slope to be  $-.30$ , for every one year change in age, the expected average age difference one feels from their actual age goes down  $.3$  years with a standard error of  $.078$ , and a 95% confidence interval between  $-.450$  and  $-.145$  which does not contain 0 and the estimated coefficient is statistically significantly different from 0.

```
# SE
summary(mod_2)$coefficients[2, 2]
```

```
[1] 0.07763039
```

```
# 95% CI
confint(mod_2)[2,]
```

```
      2.5 %      97.5 %
-0.4502312 -0.1450136
```

## b. Heteroscedasticity consistent variance estimator

- The standard error assuming heteroscedasticity is .068, and a 95% confidence interval between -.432 and -.163 which does not contain 0 and the estimated coefficient is statistically significantly different from 0.

```
hetvar <- mod_2 |>
# calculate Heteroscedasticity-consistent estimation
## of the covariance matrix for coefficients
vcovHC() |>
# gives variances as they are the diagonal of the covariance matrix
diag() |>
# variances are the diagonal of the covariance matrix
sqrt()

# SE
hetvar[2]
```

AGE  
0.06829585

```
# variances are the diagonal of the covariance matrix
hetero_ci <- mod_2$coefficients[2] + c(-1,1) *
qt(0.975,mod_2$df.residual)*hetvar[2]

hetero_ci
```

[1] -0.4318810 -0.1633638

## c. Bootstrap method (use set.seed(97) and bootstrap 1000 times; do not use boot() or boot.ci()); do not use quantile-based confidence intervals for this question)

- The standard error for the bootstrap method is .046, and a 95% confidence interval between -.567 and -.387 which does not contain 0 and the estimated coefficient is statistically significantly different from 0.

```

diab <- path_final %>%
  dplyr::select(AGE_DIFF, AGE) %>%
  na.omit() # drops missing observations

n_diab <- dim(diab)[1]

set.seed(97)

n_boot<-1000

beta_chol<-rep(0,n_boot) # create vector the length of n_boot to give something for for loop

# this code might take a second to run
for(i in 1:n_boot){
  samp_diab <- diab%>%
    sample_n(n_diab,replace=T)
  bmod<-lm(AGE_DIFF~AGE,samp_diab)
  beta_chol[i]<-coef(bmod)[2]
}

sd(beta_chol)

```

```
[1] 0.04555169
```

```

t.score<-qt(p=.05/2, df=n_diab-2, lower.tail=F)

Lci<-mean(beta_chol)-t.score*sd(beta_chol)
Uci<-mean(beta_chol)+t.score*sd(beta_chol)

print(c(Lci,Uci))

```

```
[1] -0.5660279 -0.3869360
```

- Overall, the OLS had the largest standard error, and the bootstrap method had the smallest which impacts how the confidence intervals are calculated. Therefore, the OLS had the widest confidence interval while the bootstrap method had the narrowest.

## 6.6 Among the three CI's in Q6.5, which one would you recommend to a non-statistician and how would you explain your rationale?

- Our recommended CI comes from the bootstrapped method. Imagined trying to collect the acidity in rain water in DC, but you cannot collect every single rain drop. Instead

you randomly select 1000 locations within DC and place a bucket at each location. After the rain, we collect all 1000 buckets filled with rain water and test in each one the acidic levels. Would we be more confidence in this approach over simply placing just one bucket in DC and testing its content, the answer is yes. This is why we would recommend to report the bootstrap CI over the OLS or Heteroscedasticity consistent variance estimator.

**7. Questions 7.1-7.3 do not require you to do any coding, but instead ask you to interpret different types of output. PATH researchers conducted an experiment about the location of self-rated health question.**

**7.1. PATH researchers analyzed SRH with the experimental factor, SRH\_LOC}. The outputs are as follows. Report all hypotheses tested in these outputs mathematically or in text and describe the substantive conclusions in layman's terms. Use outputs from both summary() and anova().**

- For  $\hat{\beta}_0$ , the intercept, for SRH ratings between SRH asked before any health and well-being questions the expected average rating is 2.62, and is statistically significantly different from 0.
- For  $\hat{\beta}_1$ , the Null hypothesis ( $H_0 : \beta_1 = 0$ ) is the expected mean difference of SRH ratings between SRH asked before any health and well-being questions (reference group) and after chronic conditions is equal to zero. The alternative hypothesis ( $H_A : \beta_2 \neq 0$ ) is the expected mean difference in SRH ratings between these two groups is not equal to zero. We failed to reject the Null hypothesis ( $p > .05$ ), the expected mean for SRH before health and well-being ( $\hat{\beta}_0$ ) of 2.65 is not statistically different from SRH after chronic conditions ( $\hat{\beta}_0 + \hat{\beta}_1$ ) of 2.77, while holding other predictors variables constant.
- For  $\hat{\beta}_2$ , the Null hypothesis ( $H_0 : \beta_2 = 0$ ) is the expected mean difference of SRH ratings between SRH asked before any health and well-being questions (reference group) and after life satisfaction is equal to zero. The alternative hypothesis ( $H_A : \beta_2 \neq 0$ ) is the expected mean difference in SRH ratings between these two groups is not equal to zero. We reject the Null hypothesis ( $p < .05$ ), the difference in SRH ratings for SRH before health and well-being ( $\hat{\beta}_0$ ) of 2.65 is statistically different from SRH after chronic conditions ( $\hat{\beta}_0 + \hat{\beta}_2$ ) of 2.44, while holding other predictors variables constant.
- Since we have multiple parameters we are testing the global Null hypothesis,  $H_0 : \beta_1 = \beta_2 = \dots \beta_p = 0$ , and each  $\beta_p = 0$ , meaning that our Null hypothesis is that each  $\beta$  coefficient is equal to each other and to 0. Our alternative hypothesis,  $H_A : \beta_i \neq 0 (i = 1, \dots, p)$  is that at least one of the  $\beta$  coefficients are not equal to 0. Here we reject the Null hypothesis, and there is a useful linear relationship between SRH ratings and one of the predictors in the model.



- In the One way ANOVA, the Null hypothesis  $H_0 : \alpha_1 = \alpha_2 = \dots \alpha_{alpha}$  that each treatment effect is equal to each other versus the alternative hypothesis that at least one of the  $\alpha$  is different. We take note that the degrees of freedom, F-statistic, and p-value in the ANOVA table are the same values reported in the linear model summary table which we interpret as at least one of the treatment effects is statistically significantly different and reject the Null, ( $p < .05$ ).
- SRH ratings varied by conditions but only one condition was significantly different from the reference group, after life satisfaction. However, based on the adjusted R-squared, this predictor only explains about 2% of variation in SRH ratings.

**7.2 PATH researchers also analyzed SRH as a function of SRH\_LOC as follows. Interpret SRH\_LOC's impact on SRH and the hypothesis tested mathematically or in text. Describe the substantive conclusions in layman's terms. How does the hypothesis tested (for SRH\_LOC) differ in this model when compared to a fixed effects model?**

- This model is also known as the intercept only model as there is only one outcome variable (SRH ratings), and one grouping variable (SRH\_LOC). However, we cannot make any conclusions about the hypothesis being tested from just looking at the summary table, but instead we would want to look at the confidence intervals for the fixed intercept, sigma squared which is the within variation, and the between variation eta squared. What we can make out from this model is that the intraclass correlation is  $.02817 / (.02817 + 0.79072) = 3.4\%$  which is the proportion of variance in SRH values explained by the grouping structure of SRH\_LOC, which is minimal.

**7.3 PATH researchers were curious about whether responses to the global life satisfaction (GLOBALSAT) interact with SRH\_LOC and conducted the following analysis. Report all hypotheses tested in the output of anova() below (mathematically or in text). Describe the substantive conclusions of each hypothesis in layman's terms. Be sure to incorporate your observations about the plot.**

- The two way Anova for this model is testing three Null hypotheses: There is no difference in the means of factor SRH\_LOC, there is no difference in the means of factor GLOBALSAT, and there is no interaction between factors SRH\_LOC and GLOBALSAT. This model given the asterisk interact the two factors to create a synergistic effect rather than an additive model. We can see that the two factors are statistically significantly different ( $p < .05$ ), at least one treatment effect is different. In the plot we can also see that the GLOBALSAT lines move in parallel without crossing each other, visually implying that there might not be an interaction of interest. The p-value of the interaction 0.997695

confirms statistically that it is not statistically significant ( $p > .05$ ), suggesting that way may want to drop the interaction and go back to the additive model. The p-value of SRH\_LOC is 0.002963 indicating that the levels in this factor are associated with significantly different SRH ratings. Additionally, the p-value of GLOBALSAT is 0.028957, and again indicating that the levels in this factor are associated with significantly different SRH ratings.