Assignment 2 - TM5516

Jonathan Nolan[1,1,*]

[a]*James Cook University, Townsville, Queensland, Australia*

Abstract

Keywords: heart rate, student performance, student age

1. Question 1 - Summary statistics

A total of 250 students across four JCU programs participated in the study, providing data in two surveys and from their smart watch.

The average age of participants was 25. The youngest participant was 18 and the oldest participant 41. Participants were spread across all programs, as shown in Table 1. 56% of students were completely satisfied with University.

Participants had a average resting heart rate outside of semester of 78, took an average 8079 per day and exercised 378 minutes per week.

Samsung was the most popular watch brand, with 73 participants wearing this watch, while watches other than Samsung, Garmin or Apple were least popular, with only51 wearing these brands.

Performance varied widely across the cohort, with the lowest participant scoring 47 and the best scoring 91. Full summary variables across each program are provided in Table 1.

2. Quesiton 2 - Correlations between variables

Pearson correlations for each variable are provided in Graph 1. Where a value is large and green it is suggestive of a positive correlation, whereas when values are large and red it is suggestive of a negative correlation. There appears to be a strong correlation between all measures of exercise and sleep both in and out of semester. Mid-term and final grade area also correlated. Resting heart-rate is negatively correlated with age. These data that made Graph 1 are also provided in table 2 below:

*Corresponding author

Email address: jonathan.a.nolan@gmail.com (Jonathan Nolan)

[1]jc975847

Table 1: Summary statistics

Characteristic	**Overall**, N = 250	**Nursing**, N = 59	**Occupational Therapy**, N = 84	**Dentistry**, N = 42	**Biomedicine**, N = 65
Age					
Median (IQR)	25.0 (23.0, 27.0)	25.0 (22.0, 27.0)	25.0 (23.0, 27.0)	25.0 (23.0, 26.0)	24.0 (23.0, 26.0)
Range	18.0, 41.0	18.0, 32.0	19.0, 41.0	20.0, 32.0	18.0, 30.0
Mean (SD)	25.0 (3.3)	24.7 (3.2)	25.7 (3.7)	25.1 (2.6)	24.2 (3.0)
Resting heart rate outside of semester (BPM)					
Median (IQR)	79 (67, 89)	79 (68, 93)	77 (66, 85)	70 (63, 81)	84 (72, 93)
Range	57, 106	58, 99	57, 106	59, 98	59, 106
Mean (SD)	78 (12)	79 (13)	76 (12)	73 (11)	82 (12)
Resting heart rate during semester (BPM)					
Median (IQR)	83 (71, 94)	84 (72, 96)	81 (70, 90)	74 (66, 87)	89 (76, 98)
Range	60, 100	61, 100	60, 100	62, 100	62, 100
Mean (SD)	82 (13)	83 (13)	80 (13)	77 (12)	86 (12)
Final grade (%)					
Median (IQR)	68 (63, 72)	67 (62, 72)	67 (63, 71)	69 (63, 73)	68 (64, 72)
Range	47, 91	52, 79	50, 88	55, 89	47, 91
Mean (SD)	68 (7)	67 (7)	67 (7)	69 (7)	68 (7)
Watch brand					
Apple	63 (25%)	14 (24%)	20 (24%)	11 (26%)	18 (28%)
Garmin	63 (25%)	12 (20%)	24 (29%)	12 (29%)	15 (23%)
Others	51 (20%)	17 (29%)	17 (20%)	6 (14%)	11 (17%)
Samsung	73 (29%)	16 (27%)	23 (27%)	13 (31%)	21 (32%)
Satisfaction with university					
Not completely satisfied	109 (44%)	26 (44%)	33 (39%)	18 (43%)	32 (49%)
Completely satisfied	141 (56%)	33 (56%)	51 (61%)	24 (57%)	33 (51%)
Steps per day outside of semester (average)					
Median (IQR)	7,995 (7,225, 8,880)	7,847 (6,818, 8,857)	7,764 (7,179, 8,609)	8,589 (7,467, 9,381)	8,048 (7,384, 8,865)
Range	4,398, 11,670	4,398, 11,321	5,784, 11,670	6,580, 11,108	5,864, 10,803
Mean (SD)	8,079 (1,242)	7,831 (1,480)	7,945 (1,125)	8,552 (1,217)	8,172 (1,090)
Hours of sleep per day outside semester (average)					
Median (IQR)	460 (428, 490)	451 (400, 478)	456 (426, 478)	479 (438, 501)	469 (440, 494)
Range	309, 565	309, 534	374, 546	401, 535	412, 565
Mean (SD)	458 (42)	441 (51)	453 (37)	473 (37)	470 (36)
Exercise outside of semester (minutes per week)					
Median (IQR)	379 (340, 415)	369 (311, 400)	370 (336, 400)	407 (357, 435)	385 (354, 424)
Range	202, 497	232, 476	202, 483	314, 477	301, 497
Mean (SD)	378 (50)	361 (55)	370 (47)	400 (46)	388 (45)

[1] n (%)

Figure 1 Correlation matrix

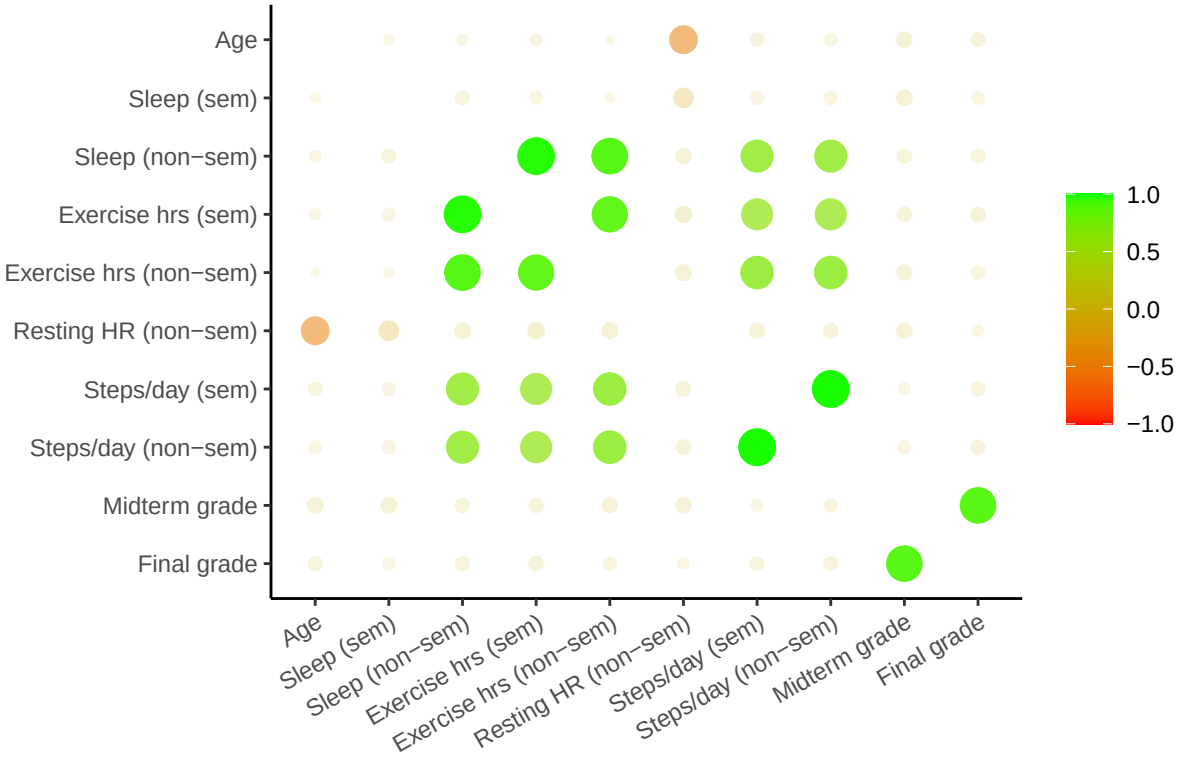


Table 2: Correlation between variables in the study

term	Age	Sleep (sem)	Sleep (non-sem)	Exercise hrs (sem)	Exercise hrs (non-sem)	Resting HR (non-sem)	Steps/day (sem)	Steps/day (non-sem)	Midterm grade
Age									
Sleep (sem)	-0.01								
Sleep (non-sem)	-0.01	0.04							
Exercise hrs (sem)	-0.01	0.03	0.97						
Exercise hrs (non-sem)	0	0.01	0.9	0.87					
Resting HR (non-sem)	-0.46	-0.16	0.07	0.08	0.07				
Steps/day (sem)	0.04	-0.03	0.68	0.62	0.71	0.06			
Steps/day (non-sem)	0.03	-0.03	0.68	0.62	0.71	0.05	0.99		
Midterm grade	0.07	0.07	0.05	0.05	0.06	-0.06	-0.02	-0.03	
Final grade	0.04	0.03	0.04	0.05	0.03	0.01	-0.04	-0.04	0.89

3. Question 3 - Relationship between University Program and Resting heart rate

3.1. Step one - Decide on most appropriate test

To decide on the most appropriate statistical test, the data was first assessed for normality. Graphical assessment of data in Figure 2 suggest that there is some right-ward shift in the data. The QQ plot in Figure 3 also shows significant deviation from the QQ line. It is therefore not necessary to conduct other tests, and the data should be considered non-parametric. However since the sample size in all three groups is larger than 30, the Central Limit Theorem applies and we can utilise tests that usually rely on the data being normally distributed.

Figure 2: Distribution of HR across programs

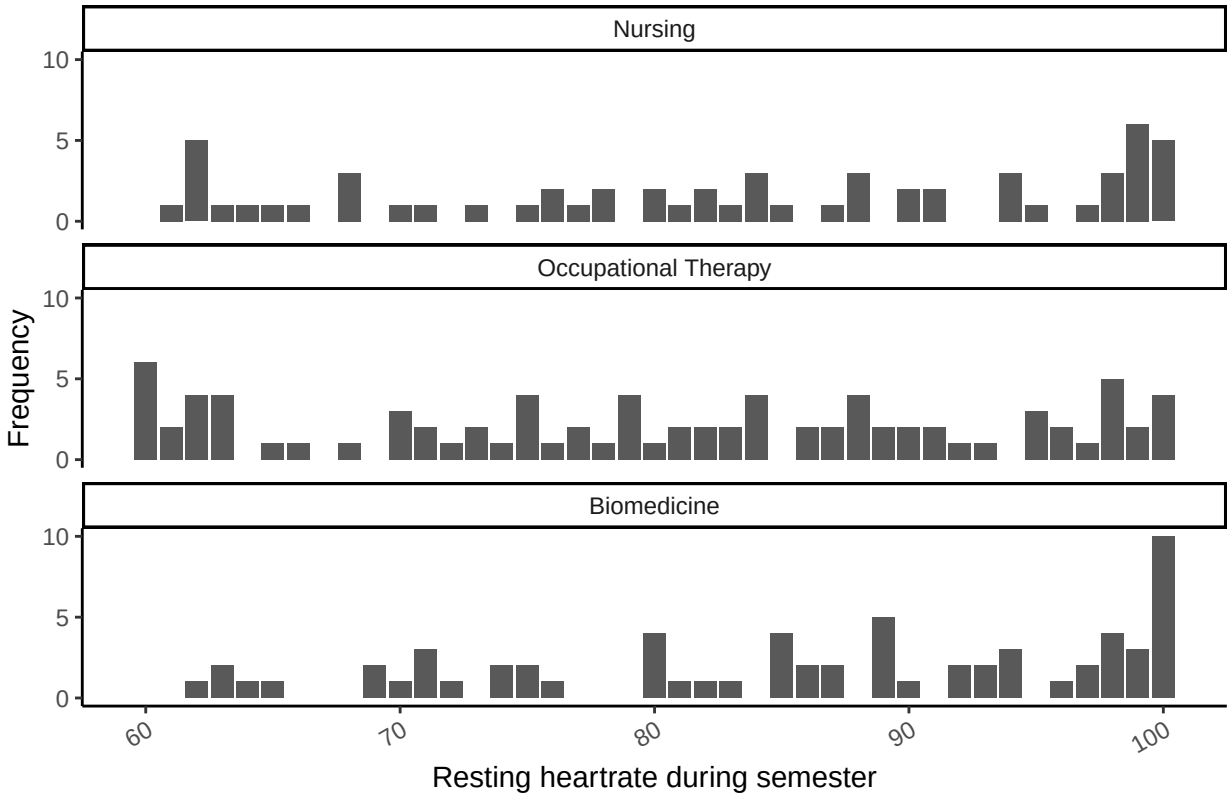
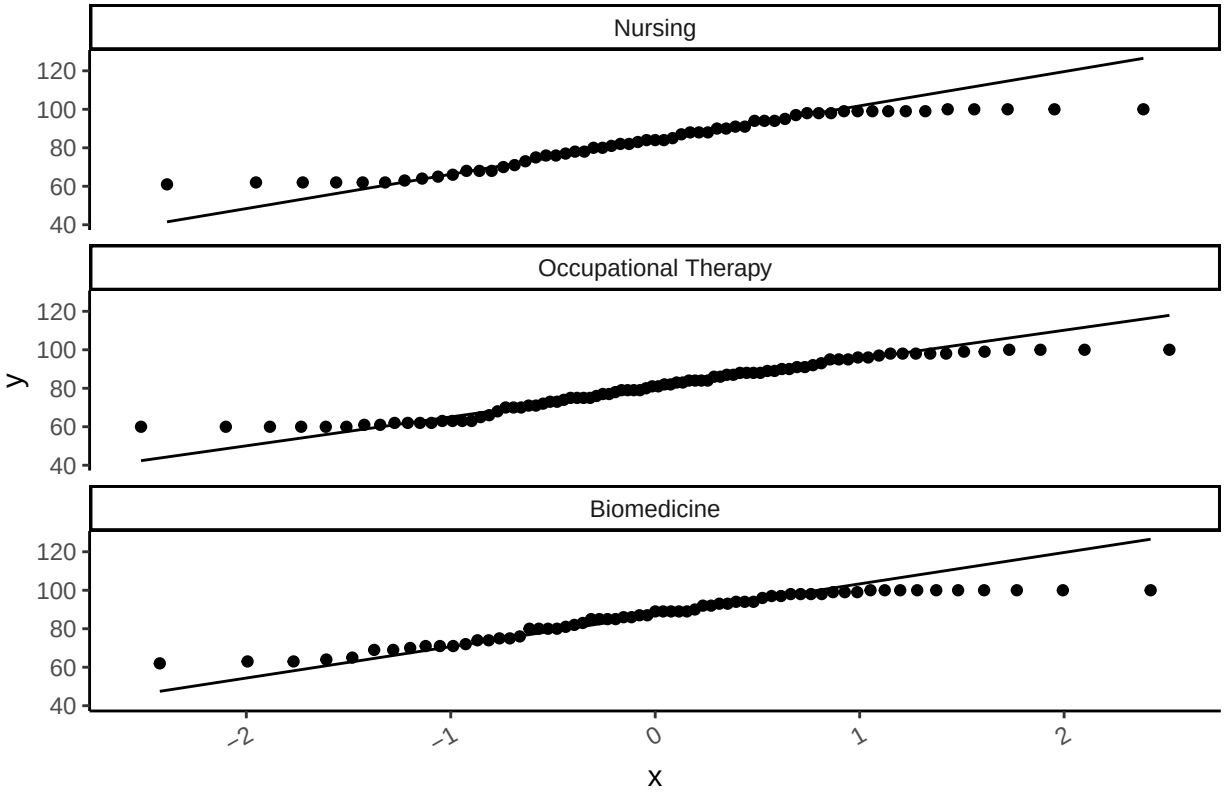


Figure 3: QQ Plot



Since we are comparing three different categories, a one way anova test is the most appropriate test.

3.2. State Hypothesis

H_o: There is no difference in the mean resting heart rate during the semester among students in the three university programs

H_α: At least one university program has a different mean resting heart rate during the semester.

3.3. Set alpha

The alpha chosen for this test is .05 as this is the most common in the literature.

3.4. Calulcate test statistic

The results for the one-way anova are presented in Table 3. There is a difference in the results between the resting heart-rate during the semester for the students in the three programs, F = NA and p = NA), meaning that the probability that this result occurred by chance alone was well below the .05 threshold chosen for this test. Since the value of the one-way anova was significant, a pairwise-comparison with Bonferroni adjustment was made in table 4. Here we find that the adjusted p-value for the comparison of means for Biomedical and Occupational Therapy students was 0.02, suggesting that there is a statistically significant difference between these two groups.

3.5. Reject or retain hypothesis

As a result of the analysis of this test above, we have rejected the null hypothesis because there is a statistically significant difference between the Resting heart rate during semester of Occupational Therapy and Biomedical students.

Table 3: One-Way Anova of of heart rate during semester across programs

term	df	sumsq	meansq	statistic	p.value
Program	2	1289.86	644.93	4.03	0.02
Residuals	205	32792.33	159.96	NA	NA

Table 4: Pairwise comparison of heart rate during semester across programs with Bonferroni adjustment

First program	Mean	Second program	Mean	Difference between means	p.value
Occupational Therapy	80.26	Nursing	83.17	-2.91	0.53
Biomedicine	86.18	Nursing	83.17	3.02	0.56
Biomedicine	86.18	Occupational Therapy	80.26	5.92	0.02

4. Question 4 - Sleep during semester

4.1. Decide on most appropriate test

Comparing the heart rate of the same individuals both during and after semester is an example of paired sample, and therefore a paired t-test is the most appropriate analysis.

4.2. State hypothesis

H_o: The average time spent sleeping by participants during the semester is equal to the average time spent sleeping outside of the semester. H_α: The average time spent sleeping by participants during the semester is not equal to the mean time spent sleeping outside of the semester.

4.3. Set alpha

The alpha chosen for this test is .05 as this is the most common in the literature.

4.4. Calulcate test statistic

Table 5 shows that the average time spent sleeping for students in semester is 426 minutes, whereas the average time spent sleeping outside of semester is 458. To test if this value is significant we consider the results of a paired t-test in table 6. The p-value for this test is .00 which suggests that the difference between the two means is statistically significant.

4.5. Reject or retain hypothesis

Because the two means are different we reject null hypothesis that the two means are the same

5. Question 5 University satisfaction and unviersity program

5.1. Decide on most appropriate test

Because we are comparing two categorical variables, the chi squared test is the most appropriate way to check if these variables are related.

5.2. State hypothesis

H_o: University satisfaction status is independent of the university program.
H_α: University satisfaction status is dependent on the university program.

Table 5: Time spent sleeping during and outside semester

variable	Respondents	Mean	sd
Hours of sleep per day during semester (average)	250	426	45
Hours of sleep per day outside semester (average)	250	458	42

Table 6: Results of paired t-test of Sleep during and outside of semester

Difference between means	t	p.value	parameter	conf.low	conf.high	method	alternative
-31.95	-8.38	0	249	-39.45	-24.44	Paired t-test	two.sided

Table 7: Share of participants completely satisfied across programs

Program	Participants	Share completely satisfied (%)	sd
Nursing	59	56	0.5007300
Occupational Therapy	84	61	0.4913188
Dentistry	42	57	0.5008703
Biomedicine	65	51	0.5038315

5.3. Set alpha

The alpha chosen for this test is .05 as this is the most common in the literature.

5.4. Calulcate test statistic

As shown in table 7, a similar number of participants are completely satisfied across all programs.

A chi-squared test was performed, and $\chi 2$ is 1.49 p-value is 0.68, it appears the values are not statistically significantly different.

5.5. Reject or retain hypothesis

Given the p-value is less than .05, we fail to reject the null hypothesis that satisfaction is independent of program.

6. Question 6

6.1. Decide on most appropriate test

Given we are testing the effect of multiple variables on a continuous outcome variable (Resting heart rate outside of semester) a linear regression is the most appropriate test.

The equation for the regression is:

$$\begin{aligned} \text{`Resting heart rate outside of semester (BPM)`} = \alpha + \\ \beta_1(\text{Age}) + \\ \beta_2(\text{`Steps per day during semester (average)`}) + \\ \beta_3(\text{`Exercise during semsester (minutes per week)`}) + \\ \beta_4(\text{`Final grade (\%)`}) + \\ \epsilon \end{aligned}$$

6.2. State hypothesis

H_o: Resting heart rate outside of semester is not correlated with age, steps per day during semester, exercise during semester or final grade

H_α: Resting heart rate outside of semester is correlated with any of age, steps per day during semester, exercise during semester or final grade

6.3. Set alpha

The alpha chosen for this test is .05 as this is the most common in the literature.

Table 8: Results of linear regression against resting hear-rate outside of semester

Characteristic	Beta	95% CI	p-value
Age	-1.8	-2.2, -1.3	<0.001
Steps per day during semester (average)	0.00	0.00, 0.00	0.5
Exercise during semsester (minutes per week)	0.01	-0.04, 0.07	0.6
Final grade (%)	0.06	-0.14, 0.26	0.6

[1] CI = Confidence Interval

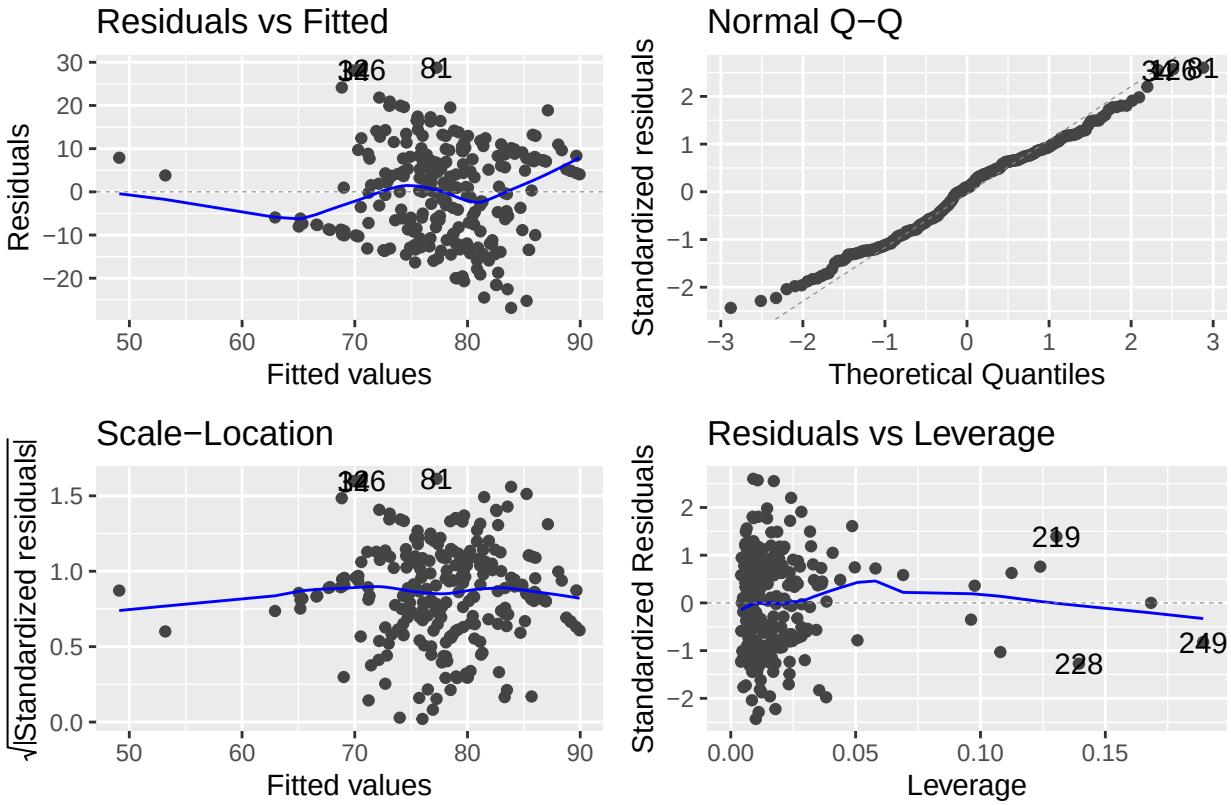
R^2 = 0.219; Adjusted R^2 = 0.207; Sigma = 11.1; Statistic = 17.2; p-value = <0.001; df = 4; Log-likelihood = -954; AIC = 1,920; BIC = 1,941; Deviance = 30,183; Residual df = 245; No. Obs. = 250

6.4. Calulcate test statistic

The results in the table 8 show that after controlling for exercise, steps per day, and final grade, resting heart rate in this population is 1.8 beats per minute lower for every year of extra age. The p-value for this estimate is less than .001, suggesting a statistically significant relationship. For all other variables p is greater than .05 and there is therefore no evidence of a relationship.

This model underwent several robustness tests and their results in figure 4. In the first plot that the residuals are evenly spread above and below the fitted line, suggesting normality. In the second (Normal Q-Q) plot the residuals form a straight line without much deviation, and the third plot shows even spread across the fitted line - suggesting minimal heteroskedasticity. In the final graph there are few outliers that have significant residuals on the model. While not perfect outputs, these four tests suggest high robustness of the model.

Figure 4 − Robustness checks for linear model



6.5. Reject or retain hypothesis

Since there is a statistically significant relationship between age and heart-rate, we have rejected hte null-hypothesis.