
**Model Code: AQ804-0-M**

**Module Name: Statistics**

**Student's TP: TP084085**

**Student's Name: Seham Saleh Yaslam Bin Tayeb**

# Table of Contents: Page

**Answers for Question1:**

**(a) (i) Show the output for the mean, standard deviation, skewness and kurtosis of the catheter length, height and weight:**

*Table 1 Descriptive Statistics*

| 1 | *Gender* | | *Height (inches)* | | *Weight (pounds)* | | *Catheter Length (cm)* | |
|---|---|---|---|---|---|---|---|---|
| 2 | | | | | | | | |
| 3 | Mean | 1.473684211 | Mean | 42.068421 | Mean | 41.8421053 | Mean | 37.5789474 |
| 4 | Standard Error | 0.117687788 | Standard Error | 2.6984407 | Standard Error | 5.82079513 | Standard Error | 1.75605184 |
| 5 | Median | 1 | Median | 40 | Median | 38.5 | Median | 37 |
| 6 | Mode | 1 | Mode | #N/A | Mode | 40 | Mode | 37 |
| 7 | Standard Deviation | 0.512989176 | Standard Deviation | 11.76223 | Standard Deviation | 25.3722577 | Standard Deviation | 7.65445251 |
| 8 | Sample Variance | 0.263157895 | Sample Variance | 138.35006 | Sample Variance | 643.751462 | Sample Variance | 58.5906433 |
| 9 | Kurtosis | -2.235294118 | Kurtosis | -0.4768585 | Kurtosis | -0.5270216 | Kurtosis | -0.00305376 |
| 10 | Skewness | 0.114668169 | Skewness | 0.0967051 | Skewness | 0.57349086 | Skewness | -0.36057868 |
| 11 | Range | 1 | Range | 41 | Range | 85 | Range | 30 |
| 12 | Minimum | 1 | Minimum | 22.5 | Minimum | 8.5 | Minimum | 20 |
| 13 | Maximum | 2 | Maximum | 63.5 | Maximum | 93.5 | Maximum | 50 |
| 14 | Sum | 28 | Sum | 799.3 | Sum | 795 | Sum | 714 |
| 15 | Count | 19 | Count | 19 | Count | 19 | Count | 19 |

**(ii) Interpret the output that you produced in part (a)(i)**

**Gender**

- **Mean (1.47):** The average indicates that there are more females than males
- **Standard Deviation (0.51):** indicating a little variability from the sample collected
- **Skewness (0.11):** Distribution is almost balanced suggesting a slight positive skew
- **Kurtosis (-2.24):** it is flatter than normal, meaning fewer extreme values.

**Height**

- **Mean (42.07 inches):** The mean height is 42.07, representing the average among participants which is low height that can be related young participants.
- **Standard Deviation (11.76 inches):** suggesting substantial variability in height among participants
- **Skewness (0.10):** Nearly symmetric distribution, with no strong bias.
- **Kurtosis (-0.48):** Flatter distribution can indicate that fewer individuals at extreme heights.

**Weight**

- **Mean (41.84 pounds):** Indicates low average weight, suggesting a younger or smaller population.
- **Standard Deviation (25.37 pounds):** High variability shows a diverse range of weights.
- **Skewness (0.57):** Positively skewed, indicating more lighter individuals with a tail towards heavier weights.
- **Kurtosis (-0.53):** the sample shows flatter distribution that suggests fewer extreme weights.

**Catheter Length:**

- **Mean (37.58 cm):** it shows that normal and typical average length.
- **Standard Deviation (7.65 cm):** little variability means lengths can differ among individuals.
- **Skewness (-0.36):** Slightly left-skewed, indicating a few longer lengths but generally balanced.
- **Kurtosis (-0.003):** shows that the distribution is close to normal, meaning extreme lengths are uncommon

**b(i) Plot a histogram with the bin width of 10 by using the height variable.**

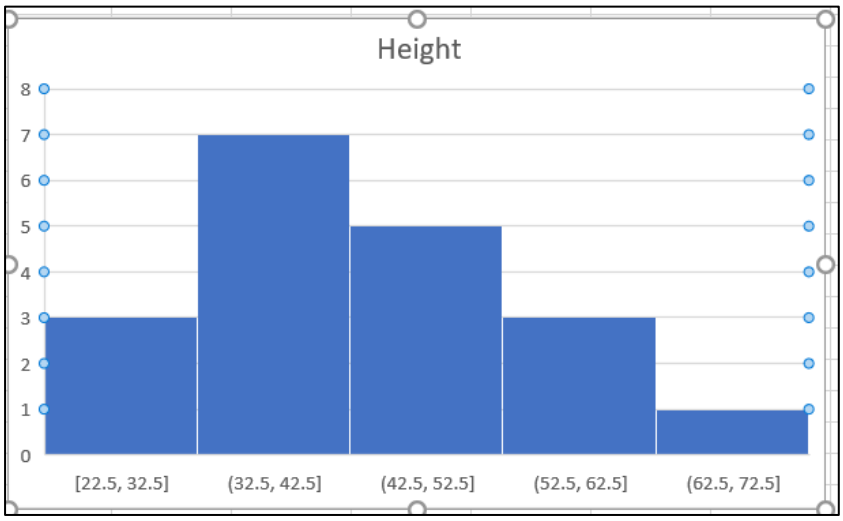


*Figure 1Histogram of Height*

**(ii) Plot a box-plot by using the height variable.**

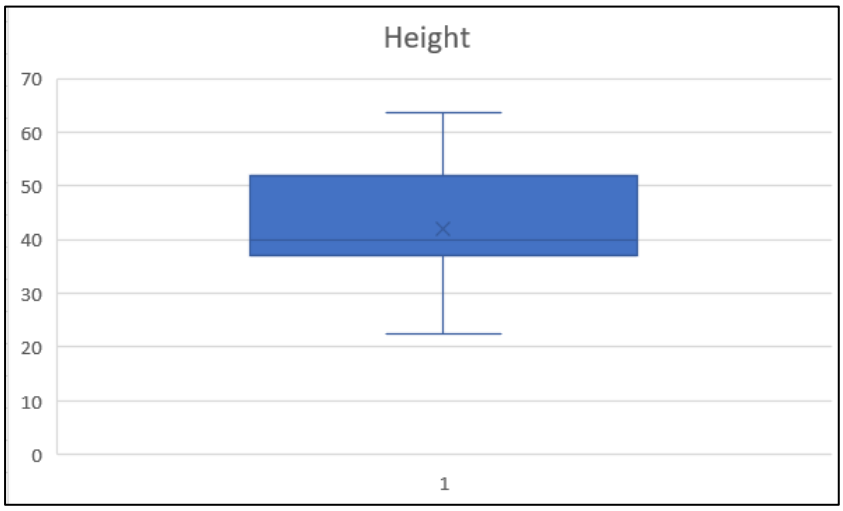


*Figure 2Box Plot of Height in Inches*

**(iii) Describe the skewness of the distribution based on the output in b(ii).**

The median is positioned near the lower section of the rectangle, indicating that the sample distribution is imbalanced. The interquartile range reveals that the lower portion of the data is more compact, suggesting that many data points are clustered around the lower values. Conversely, the upper section of the data is more spread out, signifying that while many measurements are on the lower end, a few individuals

have significantly higher heights. This results in a rightward skew in the distribution. In summary, the data primarily clusters around the lower values, but a small number of taller individuals extend the distribution to the right, creating a longer right tail.

**Answers for Question2:**

**(a) State the null and alternative hypothesis**

**The Null Hypothesis H0:**

There is no difference in the mean catheter lengths between males and females (H0: $\mu 1 = \mu 2$)

**The Alternative Hypothesis H1:**

There is a difference in the mean catheter lengths between males and females (H1: $\mu 1 \neq \mu 2$).

**(b)Show the t-statistics output and identify the value of test statistics:**

| t-Test: Two-Sample Assuming Equal Variances | | | |
|---|---|---|---|
| | | | |
| | *g1= Femal* | *g2=Male* | |
| Mean | 36.4 | 38.88889 | |
| Variance | 63.82222 | 56.36111 | |
| Observations | 10 | 9 | |
| Pooled Variance | 60.31111 | | |
| Hypothesized Mean Difference | 0 | | |
| df | 17 | | |
| t Stat | -0.69751 | | |
| P(T<=t) one-tail | 0.247456 | | |
| t Critical one-tail | 1.739607 | | |
| P(T<=t) two-tail | 0.494913 | | |
| t Critical two-tail | 2.109816 | | |

*Table 2 T-test*

I conduct the t-statistic using Excel (t statistic= -0.69751) and (df= 17)

**(c) Find the pooled variance and critical value at $\alpha = 0.05$.**

Using Excel, I found the pooled Variance= 60.31111.

The critical values= (+2.109816, -2.109816)

**(d) Interpret the result using results in (b) and (c).**

The calculated t statistic was -0.69751. To evaluate this statistic, we compared it to the critical value of 2.109816 at a significance level of $\alpha = 0.05$. Since the absolute value of the t statistic (0.69751) is less than the critical value (2.109816), we fail to reject the null hypothesis at significance level of 5%.

This outcome indicates that there is not enough statistical evidence to conclude that there is a significant difference in mean catheter lengths between males and females. Thus, we can say that any observed difference in means is likely due to random variation rather than a true difference between the groups. To summarize, we fail to reject the null hypothesis.

(**e) Identify the p value for the t test for 2 tailed tests from the output in (b).**

P(T<=t) two-tail**:** 0.4949

**(f) Draw conclusion based on (d).**

**Conclusion**: as we can see, the p-value (0.4949) is greater than the significance level $\alpha = 0.05$, we fail to reject the null hypothesis. Therefore, we conclude that there is no difference in catheter length between males and females in this dataset.

**Answers for Quoestion3:**

**(a) Briefly discuss one-way Analysis of Variance (ANOVA)**

One-way Analysis of Variance (ANOVA) is a statistical technique employed to compare the means of three or more independent groups to identify any statistically significant differences among them. This method evaluates the effect of a single categorical independent variable (or factor) on a continuous dependent variable. The null hypothesis posits that all group means are equal, whereas the alternative hypothesis suggests that at least one group mean differs.

**(b) Obtain the mean for each group and ANOVA table output using EXCEL**

| SUMMARY | | | | |
|---|---|---|---|---|
| *Groups* | *Count* | *Sum* | *Average* | *Variance* |
| Group 1 | 11 | 81 | 7.363636 | 0.454545 |
| Group 2 | 12 | 80 | 6.666667 | 1.151515 |
| Group 3 | 11 | 68 | 6.181818 | 2.963636 |

*Table 3 Mean Table*

| ANOVA | | | | | | |
|---|---|---|---|---|---|---|
| *Source of Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
| Between Groups | 7.769162 | 2 | 3.884581 | 2.570457 | 0.092709 | 3.304817 |
| Within Groups | 46.84848 | 31 | 1.511241 | | | |
| | | | | | | |
| Total | 54.61765 | 33 | | | | |

*Table 4 ANOVA Table*

**(c) Conduct ANOVA to test for any significant differences between the means. Use level of significance of 5%**

| ANOVA | | | | | | |
|---|---|---|---|---|---|---|
| *Source of Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
| Between Groups | 7.769162 | 2 | 3.884581 | 2.570457 | 0.092709 | 3.304817 |
| Within Groups | 46.84848 | 31 | 1.511241 | | | |
| | | | | | | |
| Total | 54.61765 | 33 | | | | |

*Table 5 ANOVA Table*

**(d) Discuss the implication of the result in part (c).**

Null Hypothesis ($H_0$): The mean intelligence ratings are the same for all groups (Group 1, Group 2, Group 3).

Alternative Hypothesis ($H_1$): At least one group has a different mean intelligence rating.

Using the ANOVA output:

- **F-statistic:** 2.57
- **P-value:** 0.0927

P-value is greater than the significance level 5%, there is no significant difference between the perceived intelligence ratings for people wearing different colors (blue/grey, brown/yellow, pink/orange). Therefore, based on the data, we conclude that the color of clothing does not significantly influence how intelligent a person appears. To summarize, we fail to reject the null hypothesis.

**Answers for Question4:**

**(a)(i) Show your EXCEL outputs**.

| SUMMARY OUTPUT | |
|---|---|
| | |
| *Regression Statistics* | |
| Multiple R | 0.914079582 |
| R Square | 0.835541482 |
| Adjusted R Square | 0.802649778 |
| Standard Error | 3.400422983 |
| Observations | 19 |

*Table 6 R Table*

**(ii) Interpret the meaning of R-square**

R-square measures the percentage of variance in the dependent variable (catheter length) that is accounted for by the independent variables. In this instance, an $R^2$ of 0.8355 indicates that 83.55% of the variability in catheter length can be attributed to the predictors. This high R-square value suggests a linear relationship between the

independent and dependent variables. In other words, changes in gender, height, and weight are closely related to changes in catheter length.

**b)For ANOVA table:**

**(i) Show your EXCEL outputs.**

| ANOVA | | | | | |
|---|---|---|---|---|---|
| | *df* | *SS* | *MS* | *F* | *Significance F* |
| Regression | 3 | 881.188432 | 293.7294773 | 25.40280338 | 3.95492E-06 |
| Residual | 15 | 173.443147 | 11.56287646 | | |
| Total | 18 | 1054.631579 | | | |

**(ii) Write down the relevant hypotheses:**

**Null Hypothesis ($H_0$)**: The independent variables do not predict the catheter length

**Alternative Hypothesis ($H_1$)**: At least one independent variable significantly predicts the catheter length.

**(iii) State the conclusion.**

A p-value of 3.95492E−06 indicates that the likelihood of obtaining the results is extremely low, assuming the null hypothesis is true. Because this p-value is significantly lower than 0.05, it suggests that at least one of the independent variables has a substantial impact on catheter length. Therefore, we reject the null hypothesis.

**(c)For Coefficient of Regression table:**

**(i) Show your EXCEL output.**

| | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* | *Lower 95.0%* | *Upper 95.0%* |
|---|---|---|---|---|---|---|---|---|
| Intercept | 24.27772974 | 6.798207251 | 3.571195881 | 0.002785448 | 9.787693981 | 38.76776549 | 9.787693981 | 38.76776549 |
| Gender | 0.351678635 | 1.607993036 | 0.218706566 | 0.829826489 | -3.07567739 | 3.77903466 | -3.07567739 | 3.77903466 |
| Height (inches) | 0.05583421 | 0.286771341 | 0.194699406 | 0.848240784 | -0.555404434 | 0.667072853 | -0.555404434 | 0.667072853 |
| Weight (pounds) | 0.249368364 | 0.132023433 | 1.888818965 | 0.07840833 | -0.032032922 | 0.53076965 | -0.032032922 | 0.53076965 |

**(ii) Write down the respective hypotheses.**

**Gender**

- **Null Hypothesis ($H_0$):** There is no relation between gender and catheter length.
- **Alternative Hypothesis (H1):** There is relation between gender and catheter length.

**Height (inches)**

- **Null Hypothesis ($H_0$):** There is no relation between height and catheter length.
- **Alternative Hypothesis (H1):** There is a relation between height and catheter length.

**Weight (pounds)**

- **Null Hypothesis ($H_0$):** There is no relation between weight and catheter length.
- **Alternative Hypothesis (H1):** There is a relation between weight and catheter length.

**(iii)Conclusions from the Coefficient of Regression Table**

In this analysis, the p-value for height is 0.8482, indicating no significant impact on catheter length, leading us to fail to reject the null hypothesis. Similarly, the p-value for gender is 0.8298, suggesting that gender does not significantly affect catheter length, which also results in failing to reject the null hypothesis. Regarding weight, the p-value is 0.0784, exceeding 0.05, so we fail to reject the null hypothesis here as well. In conclusion, none of the independent variables have a significant influence on catheter length.