

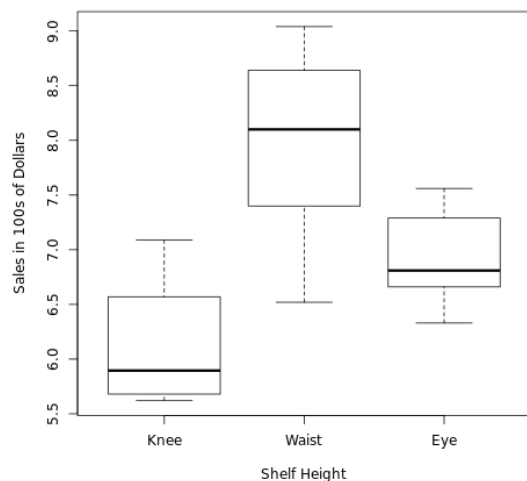
**Problem 1.** (20 points)

A study investigated the effect of shelf placement on supermarket sales of Mellow Moments, a brand of instant coffee. The study ran over a period of fourteen weekdays, and on a given day the coffee brand was shelved in one of three different heights: knee, waist, or eye level. The assignment of shelf position to day was at random. On each day, the sales of Mellow Moments (in \$ 100) were recorded and were as follows:

Knee Level	5.74	7.09	5.62	6.05	
Waist Level	7.40	8.64	8.10	9.04	6.52
Eye Level	7.56	6.33	6.66	7.29	6.81

Box plots displaying the data are shown below:

Box Plots of Mellow Moments coffee by Supermarket Shelf Height



**Part a) TRUE or FALSE:**

If applying the analysis of variance (ANOVA) to these data, we must assume...

i)

The sample mean sales for the three shelf heights are equal.

- A. True
- B. False

ii)

The daily sales from each shelf height are from a Normal distribution.

- A. True
- B. False

iii)

The daily sales for each shelf height are independent.

- A. True
- B. False

iv)

The sample standard deviations of the sales for the different shelf heights are independent.

- A. True
- B. False

v)

The daily sales for the different shelf heights are independent.

- A. True
- B. False

**Part b)** A table summarizing the data is given below. Provide the missing standard deviation (to three decimal places):

Height	Sample size	Mean	Standard deviation
Knee Level	4	6.125	—
Waist Level	5	7.940	1.004
Eye Level	5	6.930	0.493

**Part c)**

Taking that the underlying assumptions of ANOVA hold and that the approach will be applied, what is the estimate of the common variance of the daily sales for the three shelf heights? (Give your answer to three decimal places);

—

**Part d)**

Complete the ANOVA table below, giving answers to **one decimal place for F-ratio** and **three decimal places for everything else (i.e. for SS and MS)**. In particular, when calculating the F-ratio, make sure you keep all intermediate steps to at least 6 decimal places in order to arrive at a final answer that is accurate to one decimal place.

Source of Variation	df	SS	MS	F
Factor	2	—	—	—
Error	11	—	—	—
Total	13	—	—	—

#### Part e)

Which of the following inferences can be made when testing the null hypothesis that the three shelf heights give the same mean daily sales.

- A. Since the observed F statistic is greater than  $F_{2,11,\alpha=0.05}$  we do not reject the null hypothesis that the three shelf heights give the same mean daily sales.
- B. Since the observed F statistic is less than  $F_{2,11,\alpha=0.05}$  we can reject the null hypothesis that the three shelf heights give the same mean daily sales.
- C. Since the observed F statistic is greater than  $F_{2,11,\alpha=0.05}$  we can reject the null hypothesis that the three shelf heights give the same mean daily sales.
- D. Since the observed F statistic is less than  $F_{2,11,\alpha=0.05}$  we do not reject the null hypothesis that the three shelf heights give the same mean daily sales.

**Solution:** Throughout, we will denote the  $i$ th data point of knee level (with mean  $\bar{y}_1$ ), waist level (with mean  $\bar{y}_2$ ), and eye level (with mean  $\bar{y}_3$ ) by  $y_{1i}$ ,  $y_{2i}$ , and  $y_{3i}$  respectively.

#### Part a)

The true statements are (ii), (iii), and (v).

#### Part b)

Denoting  $y_i$  to be the  $i$ th data point for knee level, the standard deviation is  $\sqrt{\frac{1}{4-1} \sum_{i=1}^4 (y_{1i} - \bar{y}_1)^2} = 0.668$ .

#### Part c)

The common variance is the error mean square (MSE):

$$MSE = \frac{\sum_{i=1}^4 (y_{1i} - \bar{y}_1)^2 + \sum_{i=1}^5 (y_{2i} - \bar{y}_2)^2 + \sum_{i=1}^5 (y_{3i} - \bar{y}_3)^2}{(4-1) + (5-1) + (5-1)} = 0.577.$$

#### Part d)

The MSE was found in (c). The SS factor is:

$$4(\bar{y}_1 - \bar{y})^2 + 5(\bar{y}_2 - \bar{y})^2 + 5(\bar{y}_3 - \bar{y})^2 = 7.453.$$

The total SS is the sum of the error SS and the factor SS, and has 13 degrees of freedom. Each mean square SS is the SS divided by its degrees of freedom and the test statistic is the ratio MSA/MSE.

#### Part e)

The 95th percentile of the  $F_{2,11}$  distribution is 3.982. If the observed F statistic is greater than this number, we can reject the null hypothesis; if the observed F statistics is smaller than this number, we do not reject the null hypothesis. In this case, the F statistic is 6.458, so the correct answer is:

Since the observed F statistic is greater than  $F_{2,11,\alpha=0.05}$  we can reject the null hypothesis that the three shelf heights give the same mean daily sales.

Answer(s) submitted:

- B
- A
- A
- B
- A
- 0.668
- 0.577
- 7.453
- 3.727
- 6.458
- 6.347
- 0.577
- 13.8
- C

(correct)

Correct Answers:

- B
- A
- A
- B
- A
- 0.668
- 0.577
- 7.453
- 3.727
- 6.458
- 6.347
- 0.577
- 13.8
- C