Exercise 1

Model Coefficients - Absorbance (2)

Predictor	Estimate	SE	95% Confidence Interval Lower	Upper	t	p
Intercept	0.25796	0.00234	0.25196	0.26397	110.4	<.001
Concentration	0.00530	6.48e–5	0.00513	0.00546	81.7	<.001

Figure 1: Result of descriptives in Jamovi

a. The equation of the standard curve is y= 0.005x+0.258 where y is absorbance and x is concentration. The gradient is 0.005 ng/mL, and the intercept is 0.258.
b. The estimate of the gradient has a 95% CI (0.005, 0.005) and the intercept has a 95% CI (0.252,0.264).
c. The concentration estimated (xe) value is the following:
   X_E = intercept/gradient = 0.258/0.005= 48.67

   The prediction intervals for this concentration that has been determined by standard addition is made with this equation $x_E + t_{df}\, S_{XE}$

   X_E = intercept/gradient = 0.258/0.005= 48.67
   T_{df}= 2.571
   $S_{XE}= \frac{S_{y/x}}{b} \surd\{1\, \frac{1}{n} + \frac{\mathrm{y2}}{\sum\ (\mathrm{xi}-\mathrm{x})2 * b^2}\}$ =0.992

Average y (ybar)	0.416857143
n	7
Residuals Sum of Squares	5.88E-05
b	0.0053
Average x (x bar)	30
intercept	0.25796
Sy/x	3.43E-03
$S_{y/x}$/b	6.47E-01
1/n	0.142857143
(Average y)^2	0.173769878
b^2	0.00002809
$\sum (x_i - \bar{x})^2$	2800
Calculate in MS Excel	
S_{XE}	**0.992351816**

Figure 2: Excel data needed to calculate S_{XE}

Table 1: final calculations to calculate the prediction interval

Final Calculations	
gradient (b or m)	0.0053
intercept (c or a)	0.25796
Xe (Estimate = a/b)	**48.67**
t(df)	2.571
Sxe	0.922
upper	**51.223**
lower	**46.120**

Measurement by standard addition estimated 48.67 (ng/ml). 95% prediction interval (46.120, 51.223) for the sample.

Exercise 2

Model Coefficients - Absorbance

Predictor	Estimate	SE	95% Confidence Interval Lower	95% Confidence Interval Upper	t	p
Intercept	0.00211	0.00479	–0.0102	0.0144	0.440	0.678
Silver solution (ng/ml)	0.02516	2.66e–4	0.0245	0.0258	94.760	<.001

Figure 1: Result of descriptives in Jamovi

a. The equation of the standard curve is y= 0.025x+0.002 where y is absorbance and x is concentration. The slope is 0.025 and the intercept is 0.0021.
b. The estimate of the slope has a 95% CI (0.0245, 0.0258) and the intercept has a 95% CI (-0.0102,0.0144).
c.
   1. An unknown sample has an absorbance of 0.456, using de equation the concentration (x) obtain is the following:
      y= 0.025x+0.002
      0.456-0.002=0.025x
      0.454=0.025x
      X= 0.454/0.025= 18.016 (ng/ml)

   2. The prediction intervals for this concentration that has been determined by standard addition is made with this equation $x_E + t_{df} S_{XE}$

X_E = y-a/b = 18.016 (ng/ml)

T_{df}= 4.303

$$S_{XE}= \frac{Sy/x}{b} \sqrt{\{1\frac{1}{n} + \frac{y2}{\sum \ (xi - x)2 * b^2}\}} = 1.782$$

Average y (y bar)	0.380
n	7.000
Residuals Sum of Squares	0.000247
b	0.025
c	0.00210
Average x	15.000
estimated y for unknown yE	0.456
$(yE-ybar)^2$	0.006
$sy_{/x}$	0.007
$s_{y/x}/b$	0.281
1/n	0.143
$(\text{Average y})^2$	0.144
b^2	0.0006250
$\sum (x_i - \bar{x})^2$	700.000
Calculate in MS Excel	
1/m	40.000
S_{XE}	1.782

Figure 2: Excel data needed to calculate S_{XE}

Table 1: final calculations to calculate the prediction interval

gradient (b or m)	0.025
intercept (c or a)	0.002
Xe (Estimate = (ye-c)/b	18.016
t(df)	2.57
Sxe	1.782
upper	22.596
lower	13.436

Measurement by standard addition estimated 18.016 (ng/ml) and 95% prediction interval (13.436, 22.596) for the sample.

Exercise 3

The aim of this analysis is to compare the mean concentration of metals in soil samples from four different locations. This study is carried out to investigate a suspect at a crime scene based on soil recovered from his footwear.

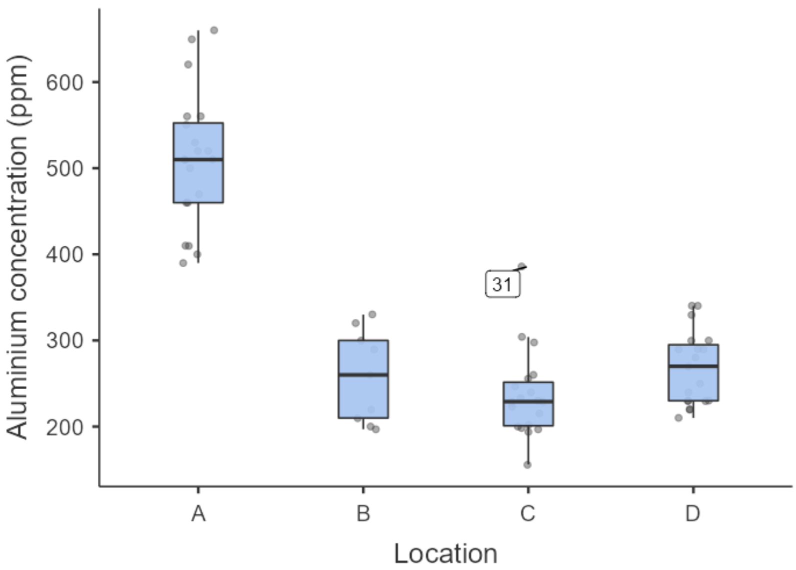

Figure 1: Aluminium concentration (ppm) in four different locations

The mean of location A appears to be higher than the other locations (Location A mean= 510 ppm, location D = 268 ppm, location B mean = 259 ppm, location C = 237ppm). The variance of location A is larger than the rest (IQR (interquartile range) Location A= 92.5, location B=90, location D= 65, location C= 50,5). In addition, highest standard deviation is location A (Standard deviation location A =77.7, location B= 53.2, location C= 50.8, location D= 42.6), this indicates the dispersion of the dataset and confirms the variability shown with the IQR. All boxes appear to not be skewed, so the assumption of normality is likely to be met and there is only one outlier in location C.

Data were analysed with a one–way ANOVA (H0: no difference in the mean of concentration between the locations, H1: there is a difference in the mean of concentration between the locations). This test assumes normally distributed data, roughly equal variance, and independent observations.

There is a significant difference between the mean of the concentration between locations ($F_{(3,63)} = 90.5$; $p<0.001$). H0 is not accepted at the 5% significance level. Tukey's pairwise comparison test confirmed that there is a significant pairwise difference between location A and location B,C,D; but there is not a significant difference between location B,C and D.

The data were normally distributed according to a Shapiro Wilks test ($SW_{(67)} = 0.970$; $p=0.101$) and the QQ plot is approximately linear (Appendix A, figure 2). The variance of the groups are roughly equal ($F_{(3,63)} = 1.66$; $p=0.184$). Therefore, the assumptions are satisfied so the result is valid.

It seems like there is a difference in location A and the rest of locations. So, it can be suggested that the footwear of the suspect was in location A based on the aluminium concentration.

Appendix A

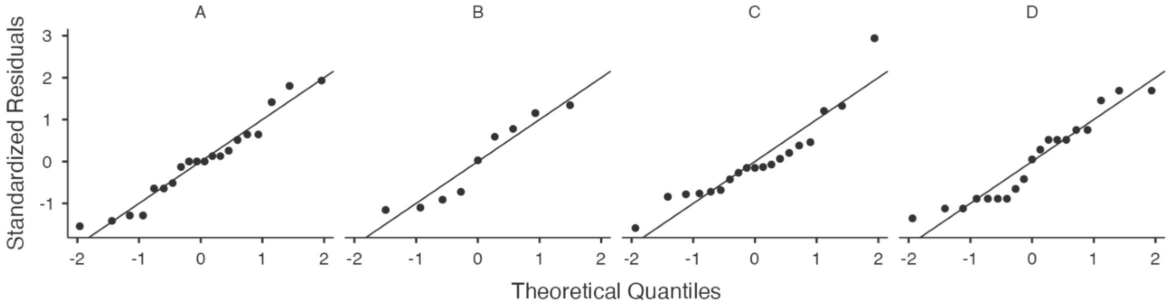

Figure 2: Normal QQ plots to check normality of the data