APPENDIX B. DETAILED STEPS OF LINEAR REGRESSION

With the regression model: $Y_i = \beta_0 + \beta_1 X_{F,i} + \beta_2 X_{C,i} + \beta_3 X_{R,i} + \varepsilon_i$

Our aim is to estimate the coefficients β_0, β_1, β_2 and β_3.

$$\text{Let } \boldsymbol{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_{10} \end{pmatrix}, \ \boldsymbol{X} = \begin{pmatrix} 1 & X_{F,1} & X_{C,1} & X_{R,1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & X_{F,10} & X_{C,10} & 10 \end{pmatrix}, \ \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}, \ \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_{10} \end{pmatrix}$$

$$\therefore \boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$$

Note: $\varepsilon_i \sim \mathrm{N}(0, \sigma^2)$ and $Y_i \sim \mathrm{N}(\boldsymbol{X}_i^T \boldsymbol{\beta}, \sigma^2)$

The Residual Sum of Squares (RSS) is given by:

$$\mathrm{RSS} = \sum_{i=1}^{10} \left(Y_i - \boldsymbol{X_i}^T \boldsymbol{\beta}\right)^2 \ and \ \ \boldsymbol{X_i} = \begin{pmatrix} 1 \\ X_{i1} \\ X_{i2} \\ X_{i3} \end{pmatrix}, i = 1, \ldots, 10$$

RSS is equivalent to the sum of the square of the residuals:

$$\sum_{i=1}^{10} \varepsilon_{\mathrm{i}}^2 = (\varepsilon_1 \quad \cdots \quad \varepsilon_{10}) \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_{10} \end{pmatrix} = \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} = (\boldsymbol{Y} - \boldsymbol{X\beta})^T (\boldsymbol{Y} - \boldsymbol{X\beta}) \ \ where \ \boldsymbol{\varepsilon} = (\boldsymbol{Y} - \boldsymbol{X\beta})$$

$$\therefore \mathbf{RSS} = (\boldsymbol{Y} - \boldsymbol{X\beta})^T (\boldsymbol{Y} - \boldsymbol{X\beta})$$

The RSS is minimised to obtain the regression coefficients, given by:

$$\widehat{\boldsymbol{\beta}} = \left(\boldsymbol{X^T X}\right)^{-\mathbf{1}} (\boldsymbol{X^T Y})$$

The covariance matrix is

$$\mathrm{Var}(\widehat{\boldsymbol{\beta}}) = \sigma^2 (\boldsymbol{X}^T \boldsymbol{X})^{-1} \quad where \quad \widehat{\sigma^2} = \frac{\widehat{RSS}}{df} \quad and \quad t - \text{statistic for } \beta_j = \frac{\widehat{\beta_J}}{\sqrt{Var(\widehat{\beta_J})}}$$

where the degree of freedom $df = n - p - 1$ where *n* is the total number of data points and *p* is the number of predictor variables $\boldsymbol{X_i}$ not including the intercept term.

The model selection criterion is based on the values of AIC and BIC that are defined as follows:

$$\mathrm{AIC} = 2(p+1) - 2\ \mathrm{Log}(L)$$
$$\mathrm{BIC} = (p+1)\ Log(n) - 2\ \mathrm{Log}(L)$$

where $Log(L)$ is the log-likelihood of the regression model.