**Linear Models – Part of Sample Exam Question**

For a dataset of size 30 a linear regression model relating the response variable ***M*** to the independent variables ***S, P,A*** and ***E*** is developed using the *R* statistics

software.

The resultant correlation matrix is

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | M | S | P | A |
| S | 0.607 |  |  |  |
| P | 0.795 | 0.788 |  |  |
| A | 0.564 | 0.928 | 0.745 |  |
| E | 0.823 | 0.869 | 0.821 | 0.782 |

**(i)** Comment on the importance of the coefficients shown in the first column **2**

**(ii)** What is meant by multicollinearity? 2

**(iii)**  Is there evidence of multicollinearity in this data? **2**

**(iv)** Why does it make sense to omit independent variable ***A*** from the model? **2**

In fitting a multiple regression model of the form:

***M = bo + b1S+ b2P + b3E***

The following partial print-out emerged:

**TABLE 1: ANOVA TABLE:**

**Source DF SS MS F Value p-value**

Regression ? ? ? ? p<0.01 \*\*\*

Error ? ? 51.3\_\_\_\_\_\_\_\_\_\_\_\_\_

Total ? 6974.1

1. Complete this table, by filling in the unknown values.
2. Describe what this table is used for.
3. Interpret the outcome of the procedure for this model.