**Homework -3 Multiple Linear Regression**

**Name : Shadeeb Hossain**

**ID: sh7492**

1. (a) The target variable could be ***“Quantity of Products sold”*** .

(b) The suggested Linear Model could be equivalent to :

Where, is the Quantity of Products sold in 1 month.

is the frequency of the occurrence of the word “good” .

is the frequency of the occurrence of the word “bad” .

is the frequency of the occurrence of the word “does not work” .

is the number of total reviews .

(c) **Normalization** can be used as shown below:

**A screenshot of a computer program

Description automatically generated**

**(d)** For missing data. dropna() command can be used as shown below

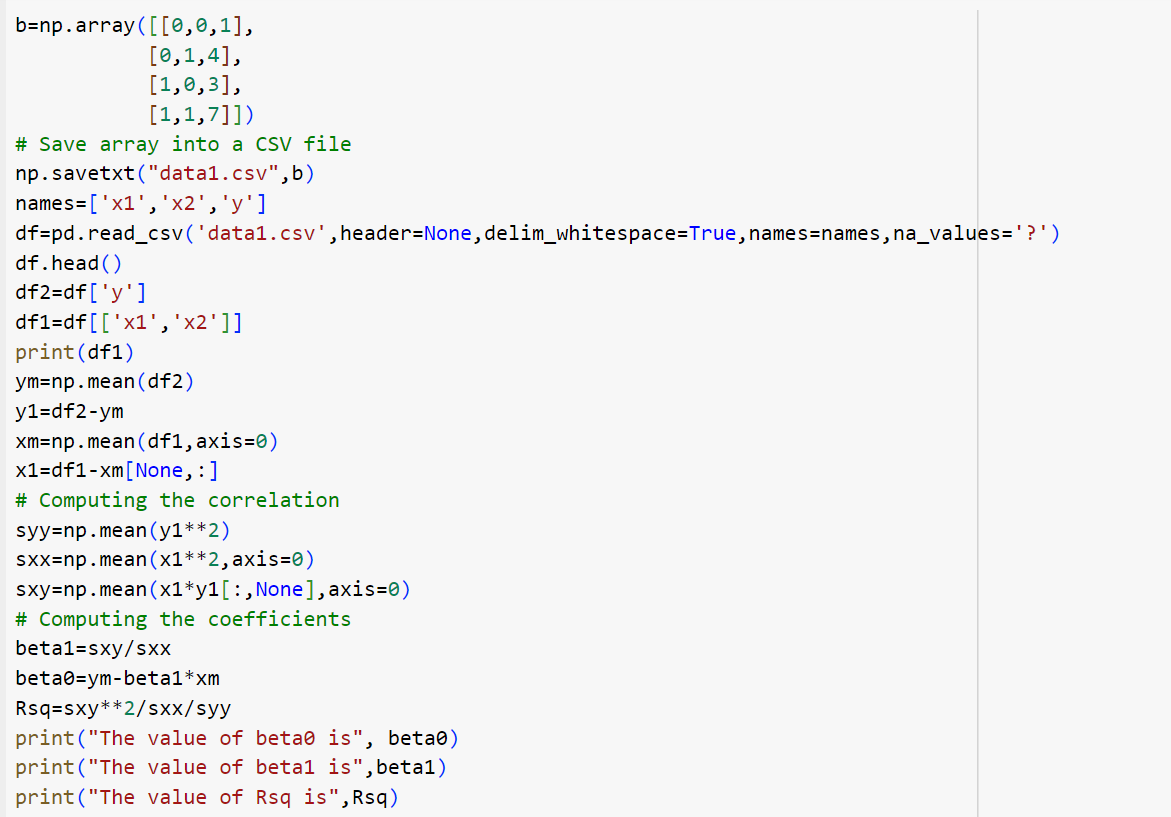
A close-up of a computer code

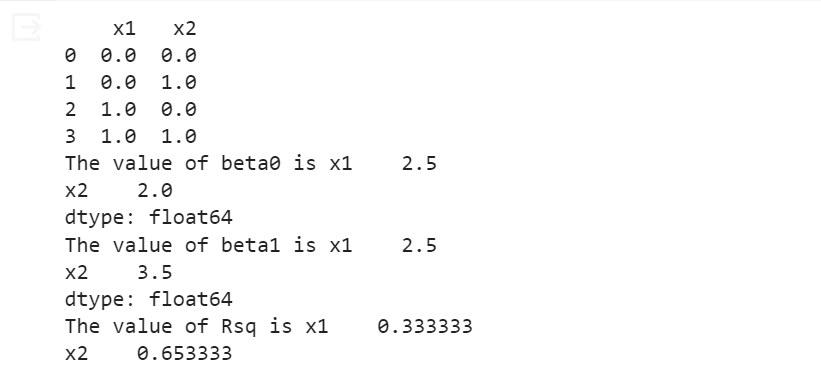
Description automatically generated

**(e)**  Fraction of reviews with the word *“good”* is a better choice between the two. This is because it accounts for normalization. Some products might have a higher number of “*good”* reviews, but their *“bad”* reviews might also be significantly higher. This error can be reduced by opting for fraction of reviews with the word *“good”.*

1. (a) Linear model for y in terms of x1 and x2 is:

(b)





A piece of paper with writing on it

Description automatically generated

A piece of paper with blue lines

Description automatically generated

**Please turn to next page**

**Page left intentionally blank**

A close-up of a graph

Description automatically generated

**Please continue to the next page for Number 5**

**Blank Page**

A piece of paper with writing on it

Description automatically generated

A piece of lined paper with writing on it

Description automatically generated

6

A close-up of a computer screen

Description automatically generated

A computer code with black text

Description automatically generated

A computer code with text

Description automatically generated