

Multiple regression model



Suppose we have to analyze the *Dataset50Startups.xlsx*.

After checking for outliers, missing data, and inaccuracies, import the file.

Define the variables:

```
rd=Dataset50Startups$`R&D expenditure`  
admin=Dataset50Startups$`Administration expenditure`  
mktg=Dataset50Startups$`Marketing expenditure`  
profit= Dataset50Startups$Profit
```

To estimate the multiple regression model with all the variables

```
modC=lm(profit ~ rd+admin+mktg)
```

To select the best model considering the predictors included in *modC*:

```
modAIC=step(modC)
```

To display the output of the regression model

```
summary(modAIC)
```