Building a simple linear model

```
> # Let's do regression with just the one predictor

> model0 <- lm (points ~ 1, data = dataset1)
> model1 <- lm (points ~ investment, data = dataset1)
```

We have built two models - *model0* is a model with just the intercept (so the mean of our outcome) predicting the outcome (*points*) while *model1* is a model with *investment* predicting the outcome (*points*).

```
> # You can compare the two models to each other

> anova(model0, model1)
```

```
> anova (model0, model1)
Analysis of Variance Table

Model 1: points ~ 1
Model 2: points ~ investment
  Res.Df    RSS Df Sum of Sq      F    Pr(>F)
1     19 120827
2     18  22046  1     98781 80.654 4.547e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

The F-ratio comparing our two models is 80.654 indicating our model with our predictor (*investment*) is a better fit than our model with just the intercept (the mean).

```
> summary(model1)

Call:
lm(formula = points ~ investment, data = dataset1)

Residuals:
    Min      1Q  Median      3Q     Max
-55.936 -20.840  -2.978  28.212  60.615

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -50.92329   23.44967  -2.172   0.0435 *
investment    0.24166    0.02691   8.981 4.55e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 35 on 18 degrees of freedom
Multiple R-squared:  0.8175,   Adjusted R-squared:  0.8074
F-statistic: 80.65 on 1 and 18 DF,  p-value: 4.547e-08
```

Here we have our parameter estimates.
Here we have the t-test associated with our predictor (investment).
Here are the R-squared and Adjusted R-squared values (which reflects the number of predictors in our model).