**Visualization:**

Draw the following scatterplots between these variables:

1. Sales and advertising

A graph with blue dots

Description automatically generated

Screenshot 1: Scatter plot for Sales Vs. Advertising

1. Sales and airplay

A graph with blue dots

Description automatically generated

Screenshot 2: Scatter plot for Sales Vs. Airplay

1. Sales and attractiveness

A graph with blue dots

Description automatically generated

Screenshot 3: Scatter plot for Sales Vs. Attractiveness

**Linear Regression**

1. Conduct a linear regression to construct a linear model between Sales and adverts and write down the F-statistic and P-value.

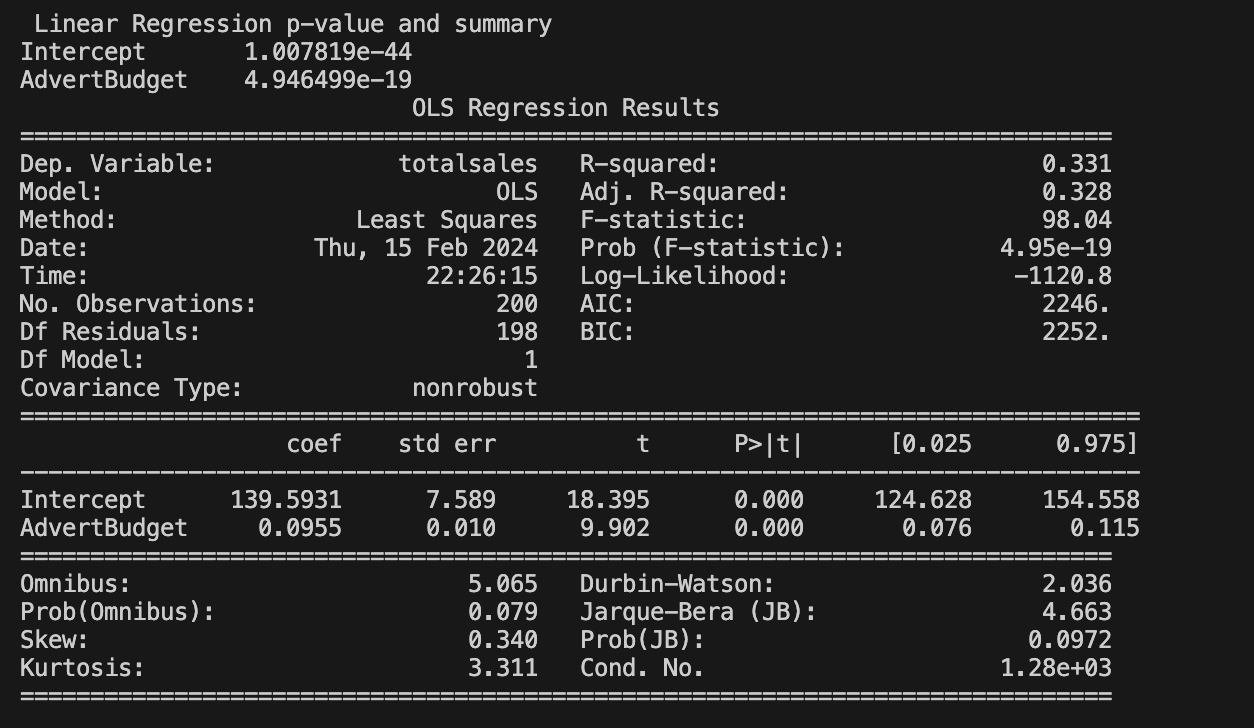
Solution:

F-Statistic: 98.04

P-values: 4.95e-19



Screenshot 4 : F-statistic and P-value



Screenshot 5: Linear Regression Summary

1. Discuss what these values (F-statistic and P-value) describe about our linear regression model? Is it good? Bad? Can’t say?
2. What is the intercept value and coefficient (adverts) value of your linear regression model?

Solution:

**Intercept Value:** 139.593120

**Coefficient (adverts):** 0.095517

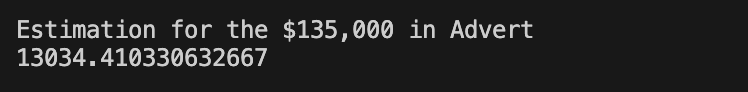
A black background with white text

Description automatically generated

Screenshot 6: Intercept & Coefficient for Adverts (Model Parameters)

1. Using the intercept value and coefficient of your linear model, please calculate how many records will be sold if we spent $135 000 on advertising the latest album “Dear Agony” by Breaking Benjamin.

Solution:



Screenshot 7: Estimation of record sale after $135000 in Advert.

**Multiple Regression**

1. Conduct a multiple regression to construct a model between Sales and the predictors (adverts, airplay, attract) and report the F-statistic and P-value.

F-statistic: 105.2

P-value: 1.27e-40



Screenshot 8: F-statistic and P-value

1. We know that the R-squared value can be used to evaluate the overall fit of a linear model. Also, that higher R-squared values are better if their p-values is < 0.05. Based on this, discuss which one of the two models that you constructed is better?

**Solution:**

Looking at the R-squared value, the multiple regression model between outcome: Sales and the predictors (advertising, airplay, attractiveness) has a good fit in comparison to the linear model between Sales and advertising, as the R-squared value 0.617 for the multiple regression is greater than the R-squared value 0.331 of the linear model between Sales and Advertising.