**SIMPLE LINEAR RIGRESSION**

**Business Problem** = Predict weight gained using calories consumed.

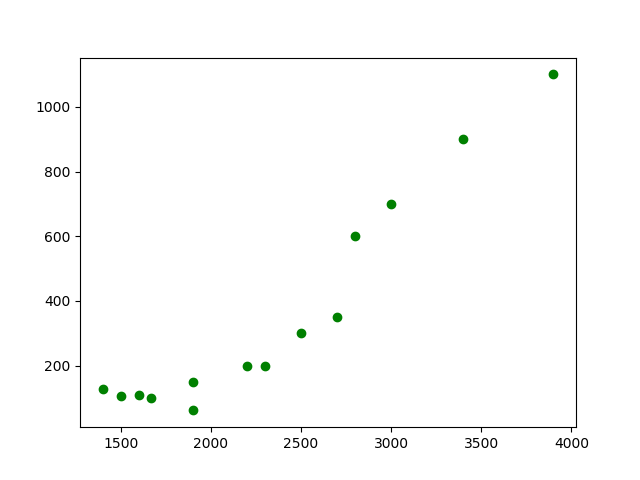
* **Name of the File: -** Calories\_consumed.csv
* **Size of the File: -** 181 bytes
* **Data: -** 14 Observation, 2 Variable
* **Column Name: -** Weight gained (grams), Calories Consumed

**Exploratory data Analysis** =

* **Skewness: -** Both variable have Right Skewed data
* **Outliers: -**  Data don’t have Outliers
* **Missing Value: -** Data don’t have Missing Values
* **Normality: -** Data are not normal
* **Transformation: -**  Required

**Liner Regression Model =**

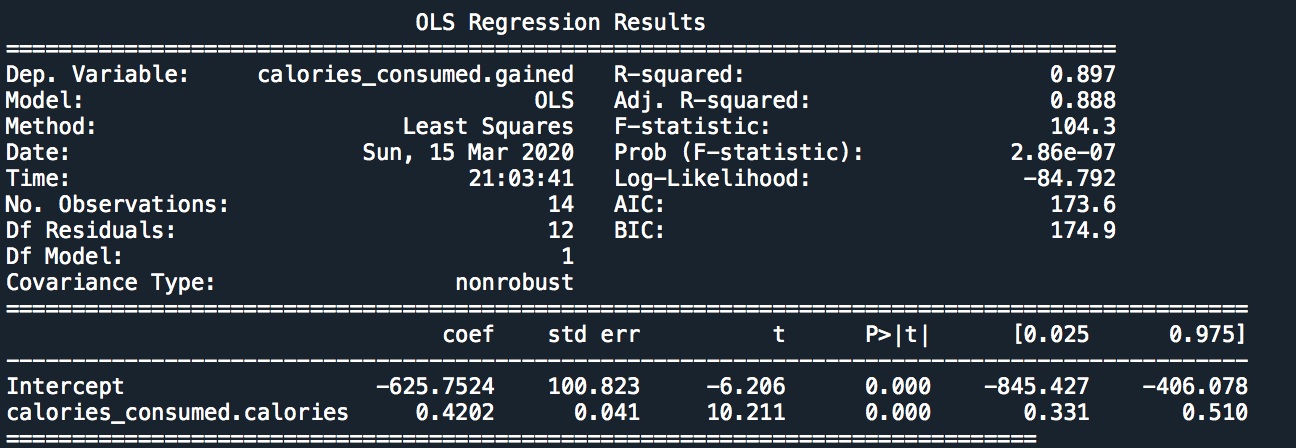
**Scatter plot =** From below scatter plot we can assume that the direction of correlation of data is positive and linear in nature.



**Correlation Coefficient (r) =** Correlation Coefficient In between Weight gain and Calories consumed is equal to 0.94 this means the data have strong positive correlation.

**Model Building =**

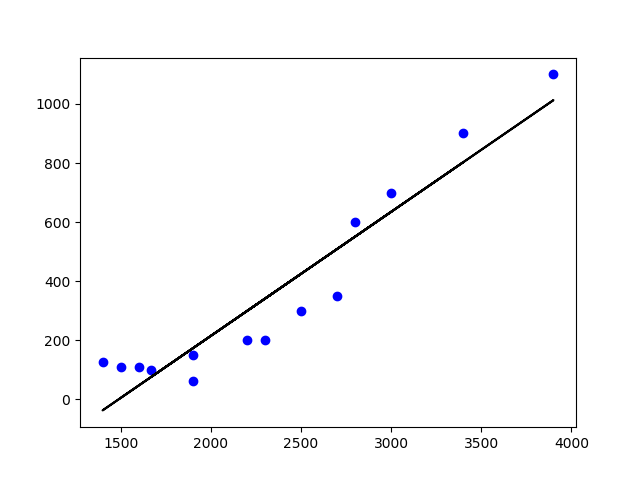
* **Summary: -**

****

* **Intersect(B0): -** -625.75
* **Slope(B1): -** 0.4202
* **Coefficient of**

**Determination (R2): -** 0.90-Strong Correlation

* **Pvalue: -** P value less than 0.05 so we can use B1 and B0 for model building.
* **Model: - weight gained** = 0.42(calories consumed) – 625.75
* **RMSE:** - 103

**Final Scatter Plot with Best Fit line**: -

As P value is significant and the Determination of coefficient is high as well as distribution of data is homoscedasticity we can use this model for Prediction.

**Final Models with 95% confidence interval: -**

* **weight gained** = 0.33(calories consumed) – 845.42
* **weight gained** = 0.50(calories consumed) – 406.42

**Python code file**: - [Calories Consumed Analysis.py](https://github.com/nilaydeshmukh0/Simple-Linear-Regression-With-EDA/blob/master/Calories%20Consumed%20Analysis/Calories%20Consumed%20Analysis.py)