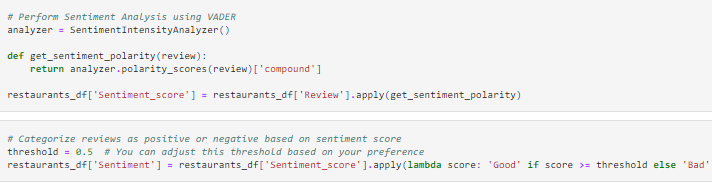
**ANALYSIS OF REVIEWS**

We used Sentiment Intensity Analyzer to categorize the reviews as ‘Good’ or ‘bad’ based on their sentiment scores.

Good score = > 0.5

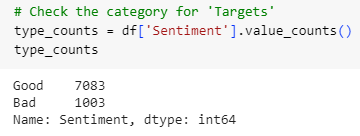
Bad score < 0.5

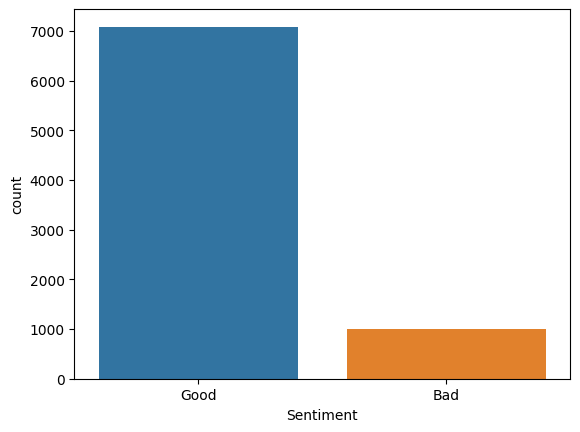


**Machine Learning Models**

1. Random Forest Model
2. Logistic Regression Model

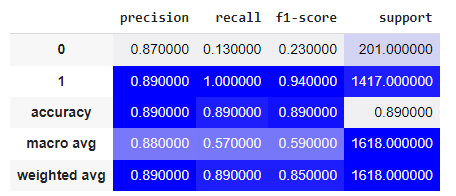
* For our y variable(target), we have a very imbalanced sample. The value count is shown below;



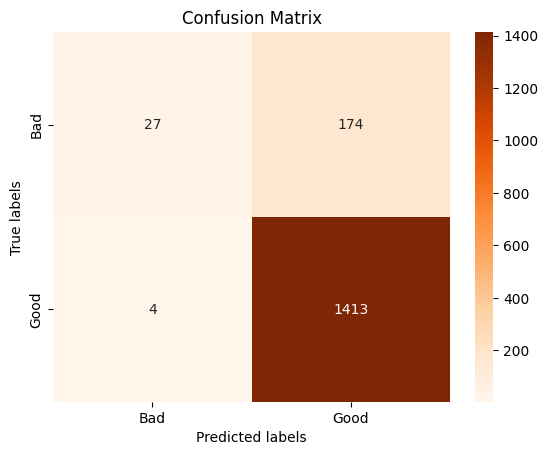


**Random Forest Model**

* Predictions using training and testing data to display Accuracy, F1-Score and precision of the sentiment classifiers

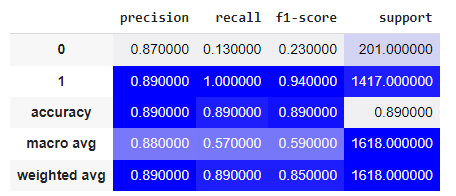


* Confusion Matrix

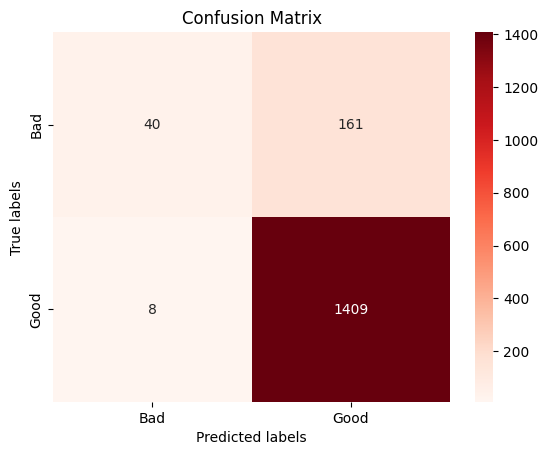


**Logistic Regression Model (LR)**

* Predictions using training and testing data to display Accuracy, F1-Score and precision of the sentiment classifiers



* Confusion Matrix



Results for LR

* **Precision** Measuring our model’s accuracy told us overall how many times we were correct when predicting the reviews. For the ‘Good’ reviews, the precision is 0.90, indicating that the model predicts this label with an excellent precision. For the "bad" reviews, the precision is 0.83, meaning that 83% of the predicted bad reviews are correct.
* **Recall**, also known as sensitivity or true positive rate tells you how many times the model was able to detect a specific label. The recall for the ‘Good Reviews’ is 0.99, indicating that the model captures almost all the actual good reviews. However, the recall for the “bad reviews” is terribly low with just 0.20, meaning that the model identifies 20% of the bad reviews correctly.
* The **F1-score** is the mean of precision and recall, and provides a balanced measure of the model's performance. For the ‘good reviews’, the F1-score is 0.94, indicating a great performance. The F1-score for the "bad reviews" is 0.32which is sadly very low.
* The **accuracy score** is 0.90 correct predictions.

**Conclusion**

* Of the two models, Logistic regression model gives us better results as its accuracy rate is 90% and has 169 false predictions in total.
* The two above models were chosen because they produce good predictions, take less training time, good methods for binary classification problems and can be understood easily.

**WORD CLOUD**

Furthermore, we went ahead and created a world cloud image to show the Frequency of the most popular words from the reviews; the bigger the font size of a word, the more repetitively the word has been used.

