**Hypothesis 1: Relationship Between Review Score and Sentiment**

**Hypothesis:** There is a positive correlation between the average review score and the sentiment scores of positive reviews.

* **Test:** Calculate the correlation coefficient between Reviewer\_Score and Positive\_Review\_Sentiment.
* **Analysis:** If there is a strong positive correlation, it suggests that higher review scores are associated with more positive sentiments in the reviews.

**Hypothesis 2: Impact of Nationality on Review Scores: Pass**

**Hypothesis:** Review scores vary significantly by reviewer nationality.

* **Test:** Perform an ANOVA test to see if there are significant differences in Reviewer\_Score between different Reviewer\_Nationality.
* **Analysis:** If the test shows significant differences, it indicates that review scores are influenced by the reviewer's nationality.

**Hypothesis 3: Difference in Review Sentiments Between Cities: Pass**

**Hypothesis:** There are significant differences in the sentiment of reviews between different cities.

* **Test:** Use a Kruskal-Wallis test to compare Positive\_Review\_Sentiment and Negative\_Review\_Sentiment across different City groups.
* **Analysis:** Significant results would suggest that the sentiments of reviews differ between cities.

**Hypothesis 4: Correlation Between Review Length and Sentiment: Failed**

**Hypothesis:** Longer reviews (more words) are more likely to have extreme sentiment scores (either very positive or very negative).

* **Test:** Analyze the correlation between Review\_Total\_Positive\_Word\_Counts and Positive\_Review\_Sentiment, and Review\_Total\_Negative\_Word\_Counts and Negative\_Review\_Sentiment.
* **Analysis:** A strong correlation would indicate that the length of the review impacts the sentiment intensity.

**Hypothesis 5: Influence of Additional Scoring on Review Scores: Failed, not impactful**

**Hypothesis:** Hotels with a higher number of additional scores tend to have higher average review scores.

* **Test:** Calculate the correlation between Additional\_Number\_of\_Scoring and Average\_Score.
* **Analysis:** A positive correlation would suggest that additional scores positively influence the average review scores.

**Hypothesis 6: Temporal Changes in Review Sentiment**

**Hypothesis:** Sentiments of reviews have changed over time.

* **Test:** Perform a time series analysis on Positive\_Review\_Sentiment and Negative\_Review\_Sentiment using the Review\_Date field.
* **Analysis:** Identify trends or changes in sentiment over time, which might indicate external factors influencing reviews.

**Hypothesis 7: Differences in Sentiment Based on Trip Type : Pass**

**Hypothesis:** The sentiment of reviews differs based on the type of trip (e.g., leisure, business).

* **Test:** Extract trip types from the Tags field and compare Positive\_Review\_Sentiment and Negative\_Review\_Sentiment for different trip types using t-tests or ANOVA.
* **Analysis:** Significant differences would suggest that the nature of the trip affects the sentiment of the review.

**Hypothesis 8: Location Impact on Review Scores**

**Hypothesis:** Hotels located in more central areas (based on lat/lng) have higher review scores.

* **Test:** Analyze the correlation between Reviewer\_Score and the distance of the hotel from a central point in the city (e.g., city center coordinates).
* **Analysis:** If there is a negative correlation, it suggests that hotels closer to the city center tend to receive higher scores.

Testing these hypotheses will provide deeper insights into your data and help uncover patterns and relationships within the hotel reviews. If you need any specific analysis or further guidance on how to test these hypotheses, feel free to ask!

**Hypothesis 8: Correlation between review length and nationality: Failed**

**Hypothesis 8: Correlation between average review score and individual score: Not very strong**