# Trust Pilot Dataset Appendix

## 1.1: Dataset Overview

This dataset contains Trust Pilot reviews referencing four major U.S. shipping companies:

- USPS
- UPS
- DHL
- FedEx

The first ten entries were scraped using ChatGPT, then the data was collected by manually putting in the other 20 most recent reviews from each page. Raw reviews and comments were cleaned to make the formatting uniform. Sentiment scores were computed using the VADER (Valence Aware Dictionary and sEntiment Reasoner) model.

Total observations analyzed: **120 reviews**

Company Breakdown:

| Company | Number of Comments |
|---|---|
| USPS | 30 |
| UPS | 30 |
| DHL | 30 |
| FedEx | 30 |

## 1.2: Unit of Observation

Each row in the Trust Pilot dataset represents one individual review referencing a shipping company.

- Observational unit = single review
- Each observation contains metadata and computed sentiment measures

# 1.3: Variable Documentation

## 1. Variable: Shipping Company

- Type: Categorical (string)
- Description: Identifies which shipping company the Trust Pilot review references.
- Possible Values: USPS, UPS, FedEx, DHL

## 2. Variable: Posted date

- Type: Integer (Unix timestamp)
- Description: Time the comment was created (UTC format).

## 3. Variable: Star rating

- Type: Integer (Unix timestamp)
- Description: The numerical rating the reviewer gave that was attached to their text review.

## 4. Variable: Review text

- Type: String
- Description: Text of the review left on Trust Pilot.

## 5. Variable: Source

- Type: String
- Description: Website name, shipping company, and website.

## 6. Variable: Year Month

- Type: Datetime
- Description: The data and month the comment was posted in the following format: Year-Month.

## 7. Variable: Review Length

- Type: Integer
- Description: Number of characters in the review text.

## 8. Variable: cleaned_review_text

- Type: String
- Description: Review text without special characters, indents and special spacing replaced with normal spaces, and special characters removed.

## 9. Variable: neg

- Type: Continuous numeric
- Description: Proportion of the cleaned_review_text that falls into the negative category.

## 10. Variable: neu

- Type: Continuous numeric
- Description: Proportion of the cleaned_review_text that falls into the neutral category.

## 11. Variable: pos

- Type: Continuous numeric
- Description: Proportion of the cleaned_review_text that falls into the positive category.

## 12. Variable: compound

- Type: Continuous numeric
- Description: Compound sentiment score from VADER.
- Range: -1 to +1
- Description: VADER compound sentiment score computed from the cleaned_review_text.
- Interpretation:
    - Values close to +1 → Highly positive sentiment
    - Values close to -1 → Highly negative sentiment
    - Values near 0 → Neutral sentiment
- Classification rule:
- compound ≥ 0.05 → Positive
- compound ≤ -0.05 → Negative

- Otherwise → Neutral

## 13. Variable: sentiment_label

- Type: String
- Description: Based on the scores, this label identifies the sentiment as net negative, neutral, or positive for each review.

## 14. Variable: Company_Type

- Type: String
- Description: Describes if the company is public or private.

Classification rule:

- compound ≥ 0.05 → Positive
- compound ≤ -0.05 → Negative
- Otherwise → Neutral

```
                                        Review text  \
0  Low-level service. Not recommended to order or...
1  I contacted FedEx Denmark to request an export...
2  Absolutely garbage service. Package has had st...
3  This is the complete worst company ever first ...
4   The worst shipping and customer lack of service.

                               cleaned_review_text    neg    neu    pos  \
0  lowlevel service not recommended to order or l...  0.076  0.810  0.114
1  i contacted fedex denmark to request an export...  0.119  0.881  0.000
2  absolutely garbage service package has had sta...  0.058  0.901  0.041
3  this is the complete worst company ever first ...  0.122  0.760  0.118
4    the worst shipping and customer lack of service  0.516  0.484  0.000

   compound sentiment_label
0    0.7176        positive
1   -0.8126        negative
2   -0.4019        negative
3   -0.4137        negative
4   -0.7506        negative
```

# 1.4: Figures and Visualization

1. Sentiment Distribution of Reviews for All the Companies Combined (Pie Chart)

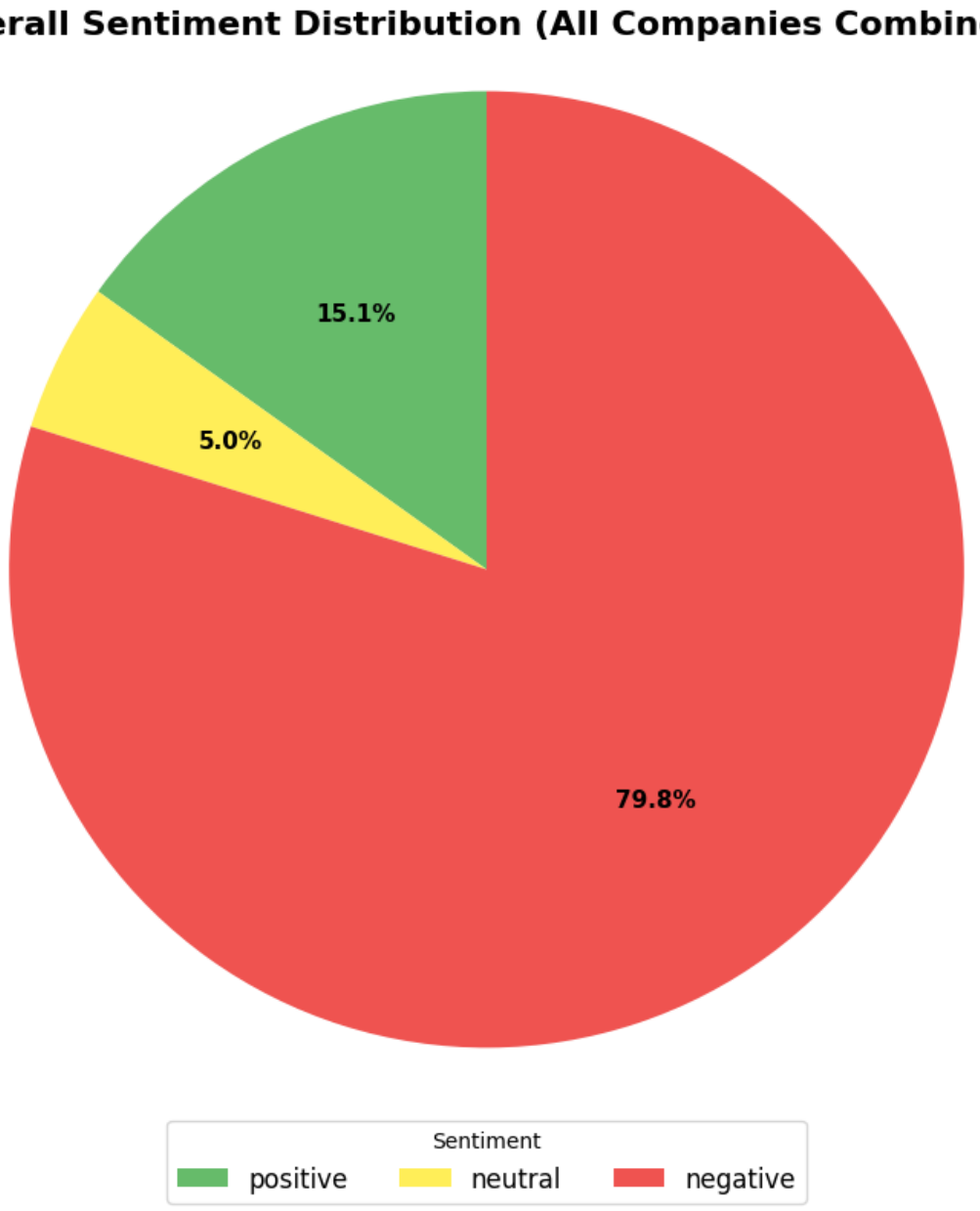



**Figure Description:**
This pie chart displays the proportion of positive, neutral, negative sentiment in the cleaned review text from the VADER analysis of all the shipping companie combined (USPS, FedEx, UPS, DHL).

**Variables Used:**

- Companies grouped
- Sentiment_label

**What the Figure Shows:**

- The majority of the reviews were labelled as negative, 79.8%.
- Second highest proportion of sentiment labels are positive, representing 15.1%.
- Few comments are processed as neutral, 5.0%

**Interpretation:**

Most of the reviews left on Trust Pilot recently have negative sentiment. The proportions we see could be explained by people's motivation to leave a review based on an extreme experience, either very good or very bad.

# 2. Distribution of VADER Sentiment Labels by Shipping Company (Pie Charts)

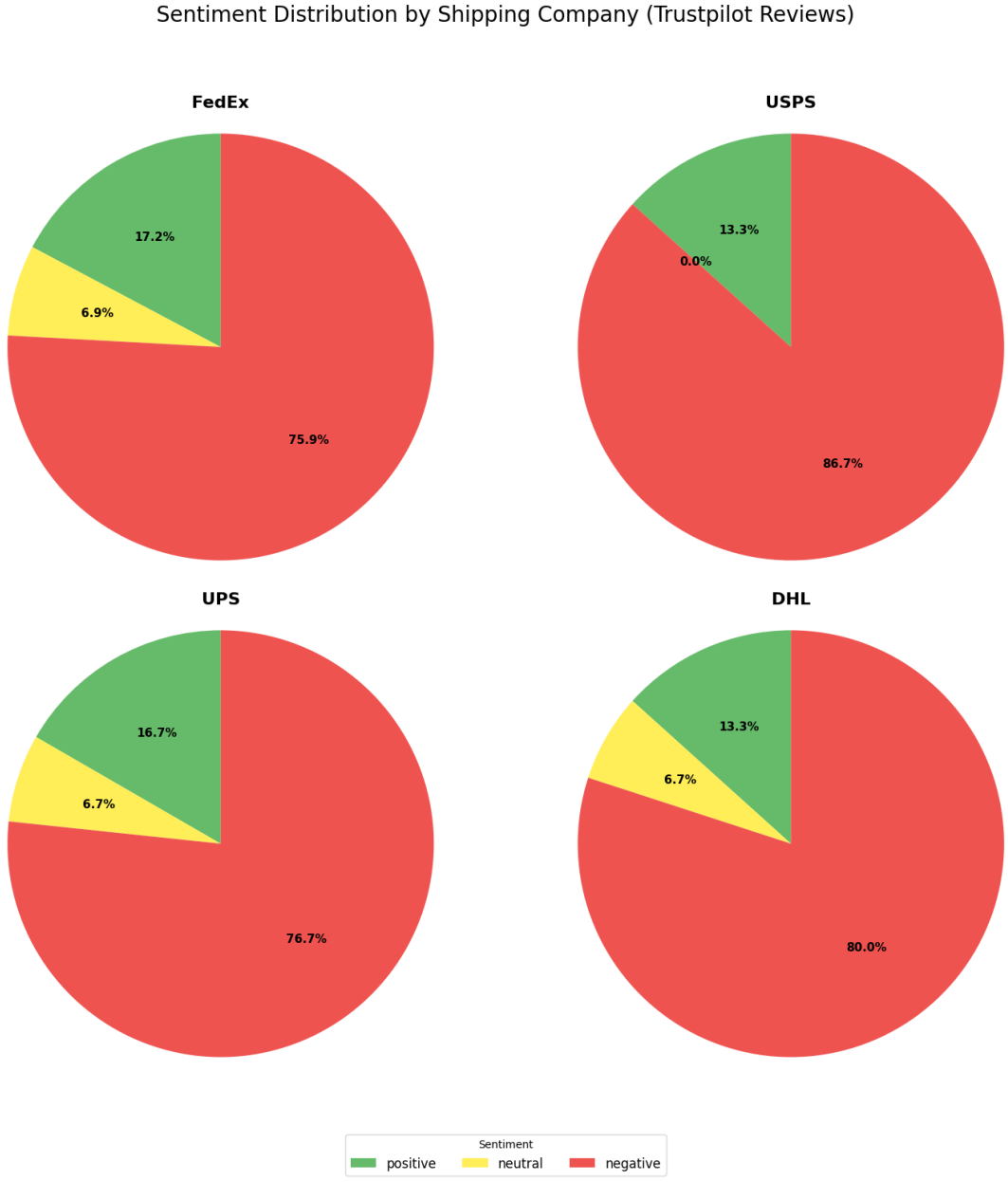



**Figure Description:**
These pie charts show the distribution of the sentiment_label for each of the shipping companies respective to their reviews.

**Variable Used:**

- Shipping Company

- Sentiment_label

**What the Figure Shows:**

- A large proportion of sentiment labels for all the companies were negative with the second most being positive, followed by neutral.

**Interpretation:**
Most Trust Pilot reviews are negative. The second largest review sentiment is positive and few reviews are neutral.

## 3. Frequency of VADER Compound Scores Across Various Reviews (Histogram)

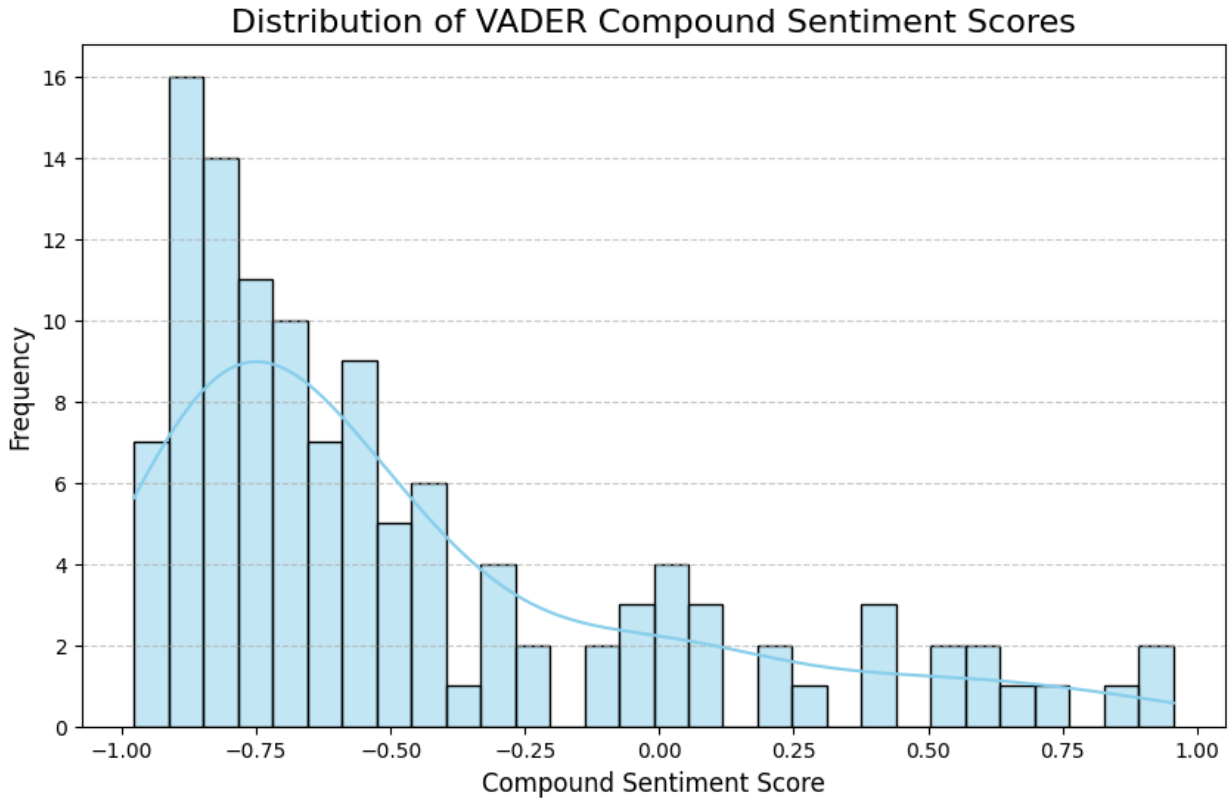



**Figure Description:**
This histogram shows the distribution of the vader_compound sentiment score across all 120 reviews.

This set of four histograms shows the sentiment distribution separately for DHL, FedEx, UPS, and USPS.

**Variable Used:**

- Vader_compound
- Shipping Company

**What the Figure Shows:**

- Higher frequency of negative sentiment scored reviews overall.

**Interpretation:**
The shape shows that there appears to be relatively more negative comments because the graph is left skewed.

# 4. Average VADER Compound by Company (Bar Graph)

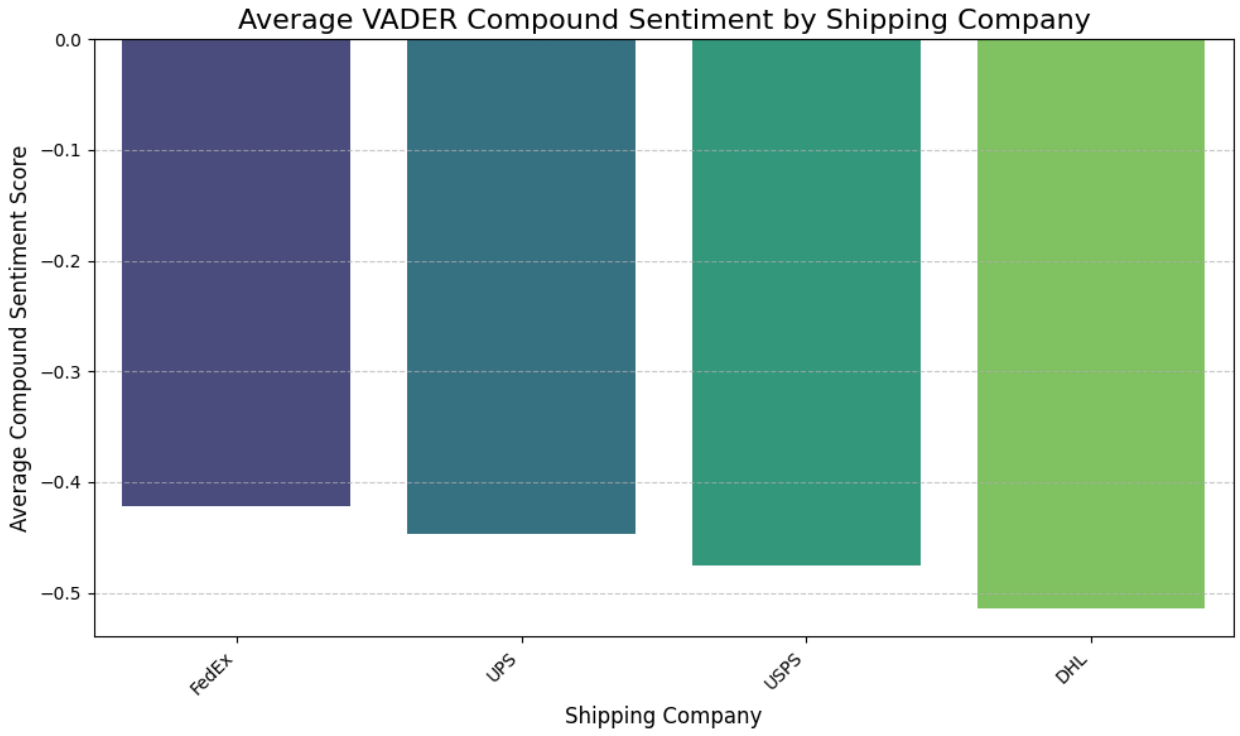


**Figure Description:**
This set of four histograms shows the sentiment average of the compound scores separately for DHL, FedEx, UPS, and USPS.

**Variables Used:**

- Vader_compound
- Shipping Company

**What the Figure Shows:**

- FedEx mean ≈ -0.415
- USPS mean ≈ -0.450
- UPS mean ≈ -0.485
- DHL mean ≈ -0.510

**Interpretation:**

DHL has the highest negative sentiment score followed by USPS, UPS, and then FedEx.

# 5. VADER Compound Distribution by Company (Boxplot)


**Figure Description:**
This boxplot compares the distribution of compound sentiment scores across companies.

**Variables Used:**

- Vader_compound
- Shipping Company

**What the Figure Shows:**

- Median sentiment is fairly similar for all.
- UPS has the largest range of scores.
- USPS has the smallest range of scores with more positive outliers.

**Interpretation:**

Some places have more concentrated negative scores (USPS) while others have more of a range (UPS).

# 6. Overall Sentiment Share (Pie / Table Visualization)

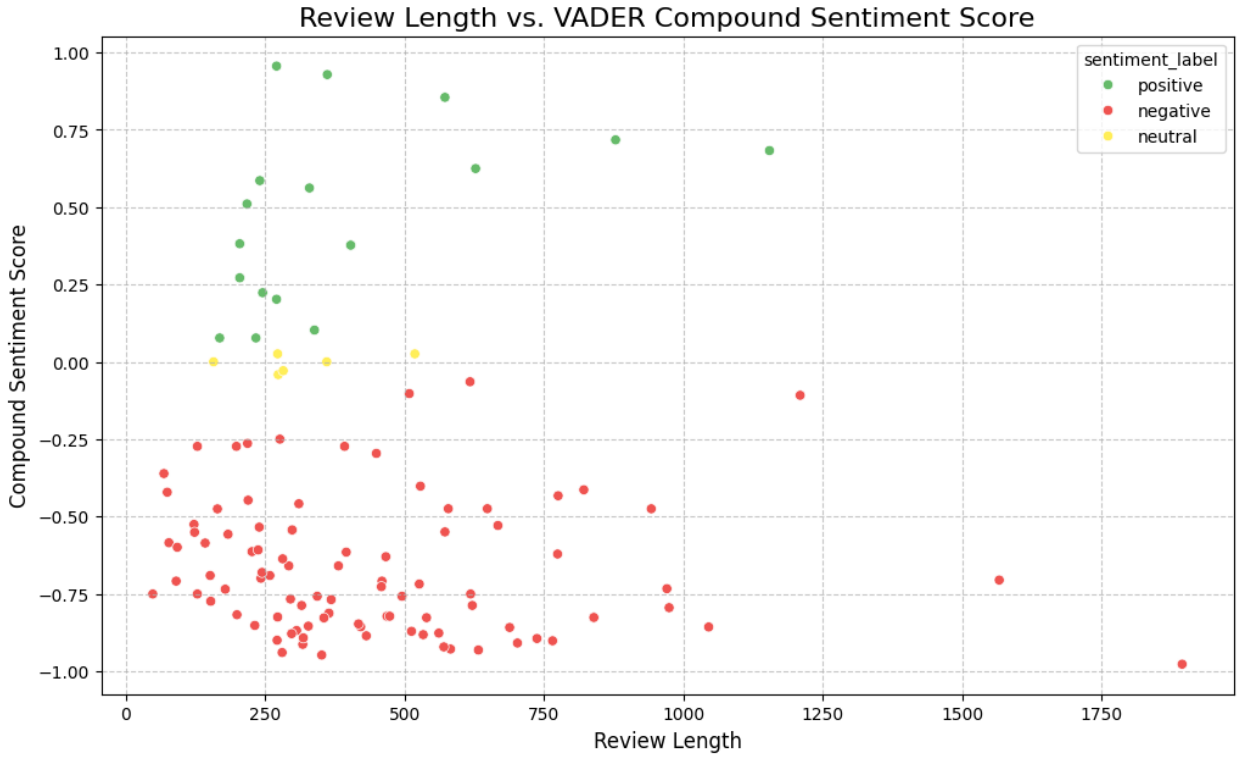



**Figure Description:**
This visualization shows the relationship between review length of reviews and compound sentiment scores.

**Variables Used:**

- Review_length
- sentiment_label

**Interpretation:**

It seems there is no correlation between review length and the sentiment length which shows that the length of the review does not have a great bearing on the sentiment score.

# 1.5: Statistical Procedures

## 1. One-Way ANOVA

```
ANOVA F-statistic: 0.198
ANOVA P-value: 0.897
```

Purpose: Test whether mean VADER sentiment differs across companies.

Results: F-statistic = 0.198, p-value < 0.05

Conclusion:
Mean sentiment does not differ significantly across at least two companies.

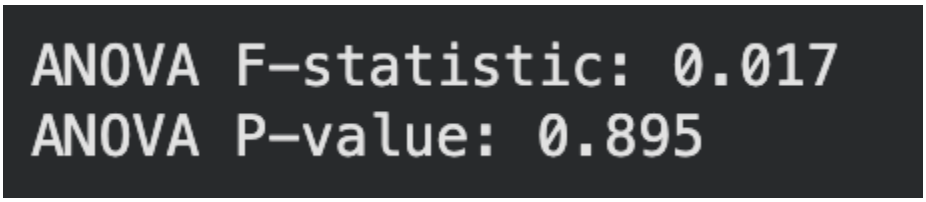
## 2. Binomial Test (>50% Negative)

```
Binomial Test Results for Negative Sentiment (>50%):

Company: DHL
  Total Reviews: 30
  Negative Reviews: 24
  Observed Negative Proportion: 0.800
  Binomial Test P-value (alternative='greater'): 0.001
  Conclusion: With p=0.001, the proportion of negative comments for DHL is significantly greater than 50%.

Company: FedEx
  Total Reviews: 29
  Negative Reviews: 22
  Observed Negative Proportion: 0.759
  Binomial Test P-value (alternative='greater'): 0.004
  Conclusion: With p=0.004, the proportion of negative comments for FedEx is significantly greater than 50%.

Company: UPS
  Total Reviews: 30
  Negative Reviews: 23
  Observed Negative Proportion: 0.767
  Binomial Test P-value (alternative='greater'): 0.003
  Conclusion: With p=0.003, the proportion of negative comments for UPS is significantly greater than 50%.

Company: USPS
  Total Reviews: 30
  Negative Reviews: 26
  Observed Negative Proportion: 0.867
  Binomial Test P-value (alternative='greater'): 0.000
  Conclusion: With p=0.000, the proportion of negative comments for USPS is significantly greater than 50%.
```

Purpose:
Test whether any company has a majority (>50%) of negative comments.

Result:
All companies have more than 50% negative sentiment (all p-values are less than 0.05).

## 3. Chi-Square Test of Independence




Purpose:
Test whether there is a statistically significant difference in mean compound sentiment scores between Private and Public shipping companies.

Results:

- F-statistic = 0.017
- p-value = 0.895

Conclusion: There is no statistically significant difference in mean compound sentiment scores between Private and Public shipping companies.