

The dataset contains three columns:

1. **statement** : Textual statements or comments.
2. **polarity** : A numerical value indicating the sentiment of the statement, where -1 is extremely negative, 0 is neutral, and +1 is extremely positive.
3. **subjectivity** : A numerical value indicating how subjective or opinionated the statement is, with 0 being very objective and 1 being very subjective.

To summarize the sentiment and visualize this data, I will:

1. Calculate summary statistics for the polarity and subjectivity.
2. Create histograms to show the distribution of polarity and subjectivity.
3. Generate a scatter plot to see the relationship between polarity and subjectivity.

The sentiment analysis of the statements in your dataset can be summarized as follows:

Summary Statistics

- **Polarity:**
 - Mean: 0.118 (slightly positive on average)
 - Standard Deviation: 0.243 (indicating variability in sentiment)
 - Minimum: -1.0 (extremely negative)
 - Maximum: +1.0 (extremely positive)
- **Subjectivity:**
 - Mean: 0.332 (leaning towards objective)
 - Standard Deviation: 0.254 (moderate variability in subjectivity)
 - Minimum: 0.0 (completely objective)
 - Maximum: 1.0 (completely subjective)

Visual Analysis

- **Polarity Distribution:**
 - The histogram shows a concentration of sentiments around the neutral to slightly positive range. There are fewer extremely negative or positive sentiments.
- **Subjectivity Distribution:**
 - Subjectivity scores are more spread out, with many statements having low subjectivity (more objective) and a reasonable spread throughout the range up to 1.0.
- **Polarity vs Subjectivity:**

- The scatter plot does not indicate a strong correlation between polarity and subjectivity. Statements are distributed across different levels of subjectivity regardless of their polarity.

This analysis suggests that while the overall sentiment in the dataset is slightly positive, there's a wide range of opinions and subjectivities expressed in the statements.