Final Report — Employee Sentiment Analysis

# 1. Approach and Methodology

The project analyzes employee email communications to evaluate sentiment and engagement. The workflow follows six main tasks:  
1) Sentiment Labeling: Using NLTK's VADER sentiment analyzer to classify messages into Positive, Negative, or Neutral.  
2) Exploratory Data Analysis (EDA): Understanding data structure, distribution, and sentiment trends.  
3) Employee Scoring: Assigning scores (+1, -1, 0) to messages and aggregating monthly sentiment scores per employee.  
4) Employee Ranking: Identifying top three positive and negative employees each month.  
5) Flight Risk Detection: Flagging employees with ≥4 negative emails within any rolling 30-day window.  
6) Predictive Modeling: Training a linear regression model using message count and length features to predict sentiment scores.

# 2. Key Findings from EDA

- The dataset is dominated by neutral messages, which is expected for professional communication.  
- Positive and negative messages occur less frequently but provide useful insight into employee tone.  
- Monthly sentiment trends reveal fluctuations in employee mood and communication activity.  
- Certain employees are much more active, likely managers or coordinators.

Visualizations such as sentiment distribution, monthly sentiment trends, and top senders charts were generated.

# 3. Employee Scoring and Ranking

Each message was assigned a score based on sentiment: Positive = +1, Negative = -1, Neutral = 0. These scores were aggregated per employee on a monthly basis. The top three employees with the highest scores were designated as 'Top Positive', while the bottom three with the lowest scores were 'Top Negative'. An extended version also reports 2 Positive, 2 Neutral, and 2 Negative employees monthly.

Example (2011-12):

|  |  |  |  |
| --- | --- | --- | --- |
| Employee | Month | Monthly Score | Rank Type |
| eric.bass@enron.com | 2011-12 | 5 | Top Positive |

# 4. Flight Risk Identification

Criteria: An employee is flagged as a flight risk if they send 4 or more negative messages within any rolling 30-day window. This method is stricter than monthly aggregation and ensures short-term spikes in negativity are detected. In the provided dataset, no employees met this criterion, meaning no flight risks were identified.

# 5. Predictive Modeling

A linear regression model was built using two features: (1) number of messages sent per month, and (2) average message length. The target variable was the employee’s monthly sentiment score. The model was trained on 80% of the data and tested on 20%. Performance was evaluated using Mean Squared Error (MSE) and R² score.

Findings:  
- The model achieved low R² (close to 0), meaning sentiment scores are poorly explained by message count and length alone.  
- MSE values indicate predictions deviate by about 1 point from true scores.  
- Interpretation: Sentiment is influenced by richer linguistic and contextual features beyond message length and frequency.

# 6. Conclusion

This project successfully implemented the six required tasks:  
- Automatic sentiment labeling with VADER.  
- Comprehensive EDA with supporting visualizations.  
- Monthly sentiment scoring and employee ranking.  
- Flight risk detection using rolling 30-day analysis.  
- Linear regression predictive model with interpretation.  
The analysis indicates that while most communication is neutral, ranking and scoring highlight outliers in employee tone. No flight risks were flagged in this dataset. Predictive modeling performance suggests additional features (e.g., NLP embeddings, topic modeling) would improve accuracy.