**Employee Feedback Analysis** (Word Clouds / Trends)

Analyze employee survey responses or public queries to extract key concerns and visualize them (e.g., word cloud).

Steps:

1. Load the data. ([employee\_feedback\_data.csv](https://indianinstituteoftechnol299-my.sharepoint.com/:x:/g/personal/24mb0051_iitism_ac_in/EV6SWFH6xjtGvBusMQdQfFoBymagFwFd9dlYTlVpZZFHXg?e=X59uTy))
2. Clean and Prepare the Text
3. Generate Word Cloud Visualization
4. Sentiment Analysis (Positive / Negative / Neutral)

This uses **VADER (Valence Aware Dictionary for sentiment Reasoning)** — a built-in sentiment analyzer in nltk.

It will add sentiment scores and sentiment label for the feedbacks

1. Find Top Trending Words (Words apprearing the most)
2. Proportion of positive vs negative feedback (in a pie chart)

Why NLTK

**NLTK (Natural Language Toolkit)** is one of the most popular open-source libraries for **Natural Language Processing (NLP)** in Python.  
It helps analyze and process human language data — such as feedback, survey responses, or customer reviews.

**NLTK is used here because it provides:**

* Pre-built **linguistic tools** (stopwords, tokenizers)
* A reliable **sentiment analyzer (VADER)**
* Easy, interpretable **text preprocessing** for beginners

This makes it **ideal for internal feedback or survey analysis**, where understanding tone and main themes is more valuable than complex deep learning.

VADER (Valence Aware Dictionary for Sentiment Reasoning):

* Uses a **lexicon (dictionary)** of English words rated for positivity/negativity.
* Also accounts for:
  + **Negations** → “not good” becomes negative
  + **Intensifiers** → “very good” becomes more positive
  + **Punctuation and capitalization** → “GREAT!!!” scores higher
  + **Emojis & slang** → understands 🙂 😢 💯 etc.

For eg.  
Statement 1. I love working here! Great team and culture.

→ {'neg': 0.0, 'neu': 0.32, 'pos': 0.68, 'compound': 0.8519}

Statement 2. The workload is overwhelming and management doesn't listen.

→ {'neg': 0.49, 'neu': 0.51, 'pos': 0.0, 'compound': -0.6696}

| **Key** | **Meaning** | **Typical Range** |
| --- | --- | --- |
| **pos** | Proportion of positive words | 0.0 – 1.0 |
| **neu** | Proportion of neutral words | 0.0 – 1.0 |
| **neg** | Proportion of negative words | 0.0 – 1.0 |
| **compound** | Overall sentiment score (weighted sum, normalized between -1 and +1) | -1 = very negative → +1 = very positive |

* compound ≥ 0.05 → Positive
* compound ≤ -0.05 → Negative
* otherwise → Neutral

So, **Statement 1: Positive & Statement 2 : Negative**

**Why It’s Useful**

This numeric scoring lets you:

* Quantify employee mood (average sentiment per team/department)
* Track sentiment trends over time
* Compare positive vs. negative feedback ratios

**Interpretation:**

**Word Cloud**

* Larger words = more frequent terms
* Example: if “communication” is large → major employee concern

A close-up of words

AI-generated content may be incorrect.

A graph of a number of blue bars

AI-generated content may be incorrect.A green and red bar graph

AI-generated content may be incorrect.

A green circle with a point of view of the same point

AI-generated content may be incorrect.

[Employee\_Feedback\_Analysis.ipynb](https://indianinstituteoftechnol299-my.sharepoint.com/:u:/r/personal/24mb0051_iitism_ac_in/Documents/AI-ML%20Projects/Employee_Feedback_Analysis.ipynb?csf=1&web=1&e=1D5WaD)

Colab Link: [https://colab.research.google.com/drive/1G-qz6mkIUOhvbohaE6STBXYMnc-6ON8a#scrollTo=dK7N83qbspk9](https://colab.research.google.com/drive/1G-qz6mkIUOhvbohaE6STBXYMnc-6ON8a%23scrollTo=dK7N83qbspk9)