

A decorative graphic on the left side of the slide. It consists of a blue parallelogram and a light green parallelogram, both tilted at an angle. The blue shape is in the foreground, and the green shape is partially behind it. They are set against a dark blue background with faint, lighter blue diagonal stripes.

# Store feedback analysis



# Pre-processing

- Removing of NULL values
- Removal of irrelevant data
- Count of unique product category
- Count of store attributes by value and plotting graphs
- Count of store by location
- Analysis of the top stores



# Individual store analysis

- Counting individual product categories
- Imbalanced data
  - 5.0 - 419
  - 4.0 - 190
  - 3.0 - 70
  - 2.0 - 12
  - 1.0 - 9
- Stop words removal
- Numerics removal
- Word cloud



# Dealing with imbalanced data

1. Undersampling
2. Oversampling
3. Synthetic Data Generation
4. Cost Sensitive Learning



# Individual store analysis

- Created new dataframe with equal no of rows for ratings
- Trained on 70% data
- Tested on remaining 30% data
- Analysed impact of each feature on overall score
- Final results :
  - Overall Experience score is going up by 24 % by incrementing store ambience.
  - 22 % for store collection
  - 17% for staff product knowledge
  - 11.6 % for staff helpfulness.