THEMATIC SENTIMENT ANALYSIS OF HOTEL REVIEWS TO OPTIMIZE OPERATIONAL STRATEGY

1. Background and Problem Statement:

A mid-scale hotel chain operating across four cities was receiving mixed reviews on online travel platforms. While the overall ratings were stable, the management team lacked visibility into **which specific service areas** (e.g., cleanliness, staff behavior, food) were driving satisfaction or dissatisfaction. The hotel needed a **theme-based sentiment analysis** of customer reviews to align operational strategies with actual guest experiences.

2. Objectives:

- To perform sentiment analysis of hotel reviews with categorization by service theme
- To identify which themes (cleanliness, food, staff, location) drive positive and negative sentiments
- To assist management in targeting specific departments for improvement
- To use data to support policy changes in housekeeping, food service, and front-desk training

3. Methodology:

Data Source:

- 3,800 hotel reviews collected from Booking.com and TripAdvisor
- Metadata included review text, review date, star rating, and review location

Tools Used:

- **R**: syuzhet, tidytext, dplyr, tm, ggplot2
- Preprocessing: tokenization, stop-word removal, lemmatization
- Thematic tagging: based on manually curated keyword dictionaries for each theme
- Sentiment scoring: using the **NRC Lexicon** to assign positive and negative sentiments

Themes Defined:

1. **Cleanliness** – clean, dirty, fresh, smell, stains
2. **Staff Behavior** – polite, rude, helpful, welcoming, indifferent

3. **Food Quality** – breakfast, buffet, taste, stale, variety
4. **Location/Accessibility** – metro, airport, nearby, difficult, central
5. **Room Comfort** – bed, noise, space, AC, lighting

4. Results:

Sentiment Breakdown by Theme:

Theme	**Positive**	**Negative**	**Net Sentiment**
Cleanliness	58%	42%	+16%
Staff Behavior	64%	36%	+28%
Food Quality	47%	53%	–6%
Location	74%	26%	+48%
Room Comfort	52%	48%	+4%

Insights from Word Frequency and Sentiment Context:

- Negative mentions in **Food** included "cold breakfast," "limited variety," and "repetitive menu"
- **Room comfort** complaints cited "noise," "small size," and "broken AC"
- Positive mentions in **Location** included proximity to metro and shopping centers
- **Staff behavior** was the strongest positive driver, especially "front desk staff" and "check-in experience"

5. Interpretation and Insights:

- **Food services** were the top contributor to negative experiences, with complaints about variety and temperature
- Guests had **mixed experiences with room comfort**, largely due to noise insulation and climate control issues
- Staff was a **positive differentiator**, often mentioned by name in 5-star reviews
- Cleanliness was acceptable but **showed variation across cities**, with lower scores in older properties

6. Recommendations:

- Redesign the breakfast menu and invest in food quality control procedures
- Audit and upgrade AC systems and soundproofing in older properties
- Use staff mentions to develop internal incentive programs
- Set a cleanliness benchmark chain-wide and conduct quarterly housekeeping audits
- Improve listings by emphasizing location strengths and staff service in marketing content

7. Future Work:

- Conduct repeat sentiment analysis after operational changes to track improvement
- Expand review collection to include Google and Agoda for wider coverage
- Use machine learning classifiers to automate sentiment classification by theme

8. Stakeholder Relevance:

Academic:

- Suitable for courses in hospitality analytics, service operations, or text mining
- Demonstrates thematic sentiment analysis, lexicon-based scoring, and R programming in business context

Corporate:

- Helps hotel brands link guest sentiment directly to department-level operations
- Provides a replicable model for guest feedback analysis and management decision-making