

Multi-Language Helpdesk Ticket Analysis using Python and Power BI

Introduction

Customer support teams handle large volumes of helpdesk tickets raised by users across different languages, priorities, and issue categories. Analyzing such data is essential to understand customer pain points, prioritize urgent issues, and allocate support resources efficiently.

This project focuses on analyzing a multi-language helpdesk ticket dataset to uncover trends related to ticket priority, issue categories, language distribution, and support team workload. The analysis aims to provide actionable insights that can help improve operational efficiency and customer satisfaction.

Objective of the Project

The main objectives of this project are:

- To analyze ticket distribution across languages, priorities, and issue types
 - To identify the most common and critical customer issues
 - To understand how ticket workload is distributed among support teams
 - To build an interactive dashboard that supports data-driven decision making
-

Dataset Description

The dataset consists of helpdesk tickets containing both structured and unstructured information.

Key Columns Used

- subject – Ticket subject
- body – Ticket description
- language – Language of the ticket
- priority – Ticket urgency level (High, Medium, Low)
- type – Ticket type (Incident, Request, etc.)
- queue – Assigned support team

- tag_1 to tag_9 – Issue classification tags
-

Tools and Technologies Used

Google Colab

Used as the development environment to run Python code without local setup.

Python

Used for data preprocessing and transformation.

Pandas

Python library used for:

- Loading datasets
- Cleaning missing values
- Creating new derived columns
- Exporting cleaned data

Power BI

Used for:

- Data analysis
 - Aggregation and comparison
 - Interactive visualizations
 - Dashboard creation
-

Methodology

Data Cleaning and Preprocessing (Google Colab)

- Loaded the raw dataset into Pandas
- Handled missing values and inconsistent formatting
- Cleaned text fields to ensure consistency
- Created a new feature main_tag by consolidating multiple tag columns into a single primary issue category
- Exported the cleaned dataset as tickets_clean.csv

Data Analysis (Power BI)

- Imported the cleaned dataset into Power BI
 - Performed aggregation and grouping by language, priority, type, queue, and issue category
 - Applied filters and slicers for interactive segmentation
 - Conducted issue severity and workload analysis
-

Dashboard Design

The Power BI dashboard consists of **two pages**:

Page 1: Helpdesk Overview

- Total ticket volume
- High-priority ticket count
- Ticket distribution by priority
- Ticket distribution by language
- Ticket distribution by type
- Interactive slicers for language, priority, and queue

Page 2: Issue & Priority Analysis

- Top customer issue categories
 - Issue severity by category (priority-wise)
 - Support team workload analysis
 - Insight text summarizing key findings
-

Key Insights

- High-priority tickets are concentrated in service disruption and account-related issues
- English and German tickets account for the majority of support requests
- Technical Support handles the highest volume of critical tickets
- Certain issue categories consistently generate urgent tickets, indicating areas needing proactive attention

Conclusion

This project demonstrates an end-to-end data analysis workflow using Python and Power BI. By transforming raw helpdesk ticket data into meaningful insights, the dashboard helps identify operational bottlenecks and customer pain points. The approach can be extended further by incorporating automation, advanced analytics, or real-time data sources.

Future Enhancements

- Integration of SQL or live databases
- Sentiment analysis on ticket descriptions
- Automated issue classification using NLP
- Real-time dashboard updates