Flood Response Data Summary Report

About the Dataset:

The dataset 'Flood_response.csv' comprises 3,774 records with 13 columns, capturing various aspects of flood response efforts. Key columns include `id`, `gender`, `location_district`, `location_upazila`, `location_village`, `caller_need`, `note`, `status`, `resolve_organization`, `feedback_status`, `resolve_comment`, `resolve_at`, and `created_at`.

- The `id` column is a unique identifier for each record, with values ranging from 1 to 7,487.
- The `gender` column categorizes the caller's gender, including options like 'Unknown', 'Female', 'Male', and 'O'.
- Location details are captured in `location_district`, `location_upazila`, and `location_village`, providing granular geographic information.
- The `caller_need` column specifies the type of assistance required, such as 'Food', 'Rescue', or 'Medicine'.
- The `note` column provides additional context, while the `status` column indicates whether the request is 'Pending', 'Resolved', or other statuses.

From the first five sample entries, it is evident that most calls were for rescue operations, primarily from the Feni and Noakhali districts. Most of these calls were resolved, with organizations like 'Volunteer' and 'Not applicable' frequently mentioned in the `resolve_organization` column. Feedback statuses varied, with entries like 'Unknown', 'Got no help', and 'Did not answer'. The `resolve_comment` column often included notes like 'Requesting for help' and 'Done', indicating the resolution status.

Overall, the dataset provides a comprehensive overview of the flood response efforts, highlighting the needs and resolutions of affected individuals across various districts.

Relevant Inquiries

Q1. What is the distribution of caller needs (Food, Rescue, Medicine) across different districts?

Distribution of Caller Needs Across Districts
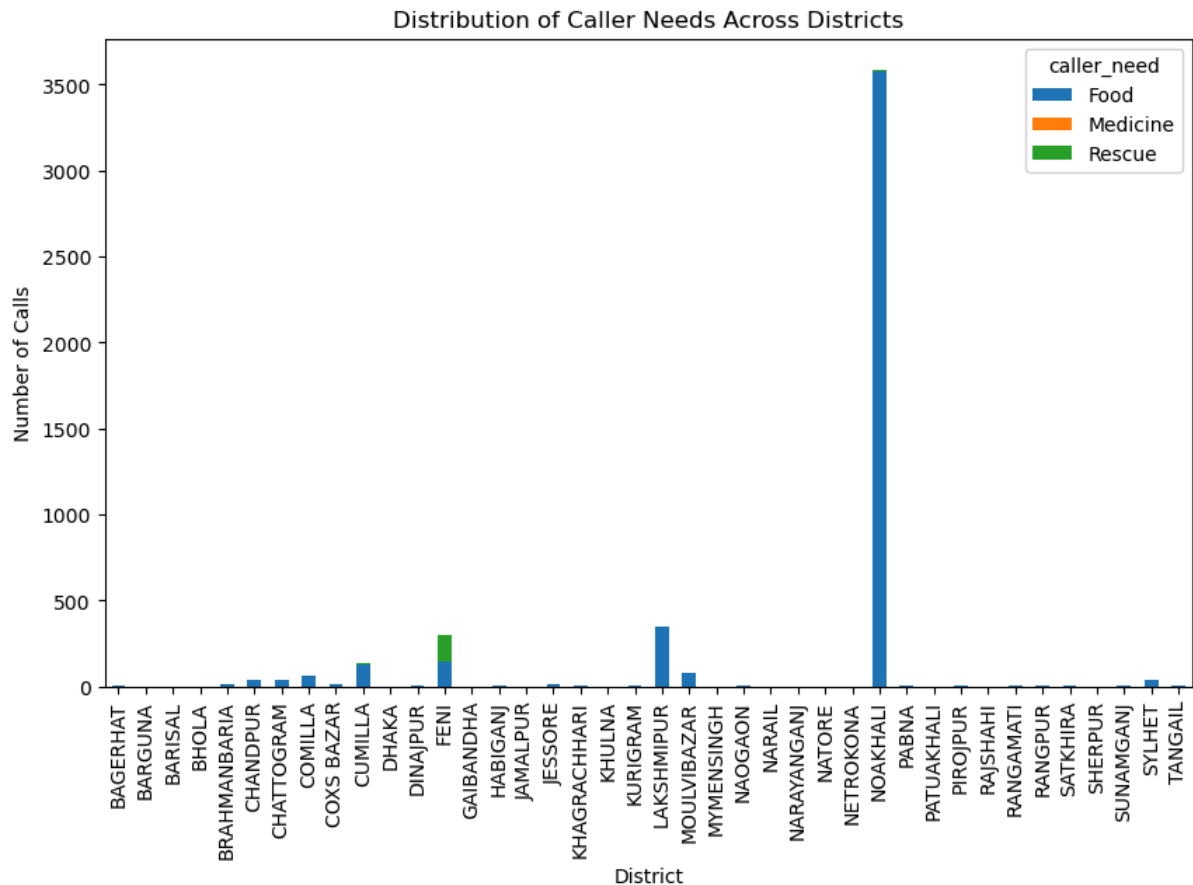
Data Analysis

- **Districts:** The data includes various districts such as Noakhali, Chattogram, Moulvibazar, Feni, and Cumilla, among others.
- **Caller Needs:** The needs are categorized into three types: Food, Rescue, and Medicine.

Visualization Insights

- **NOAKHALI:** This district has the highest count of caller needs, significantly higher than other districts.
- **Other Districts:** Most other districts have relatively low counts of caller needs, with some districts like Bagerhat, Bhola, and Chattogram showing minimal counts.
- **Distribution**: The distribution is highly skewed, with NOAKHALI dominating the count of needs.

Conclusion and Insights

Skewed Distribution: The distribution of caller needs is highly skewed, with NOAKHALI having a disproportionately high count compared to other districts.

Need for Focus: Efforts and resources might need to be concentrated more in Noakhali due to the high volume of needs reported from this district.

Q2. How does the resolution status (Pending, Resolved, Unknown) vary by gender?
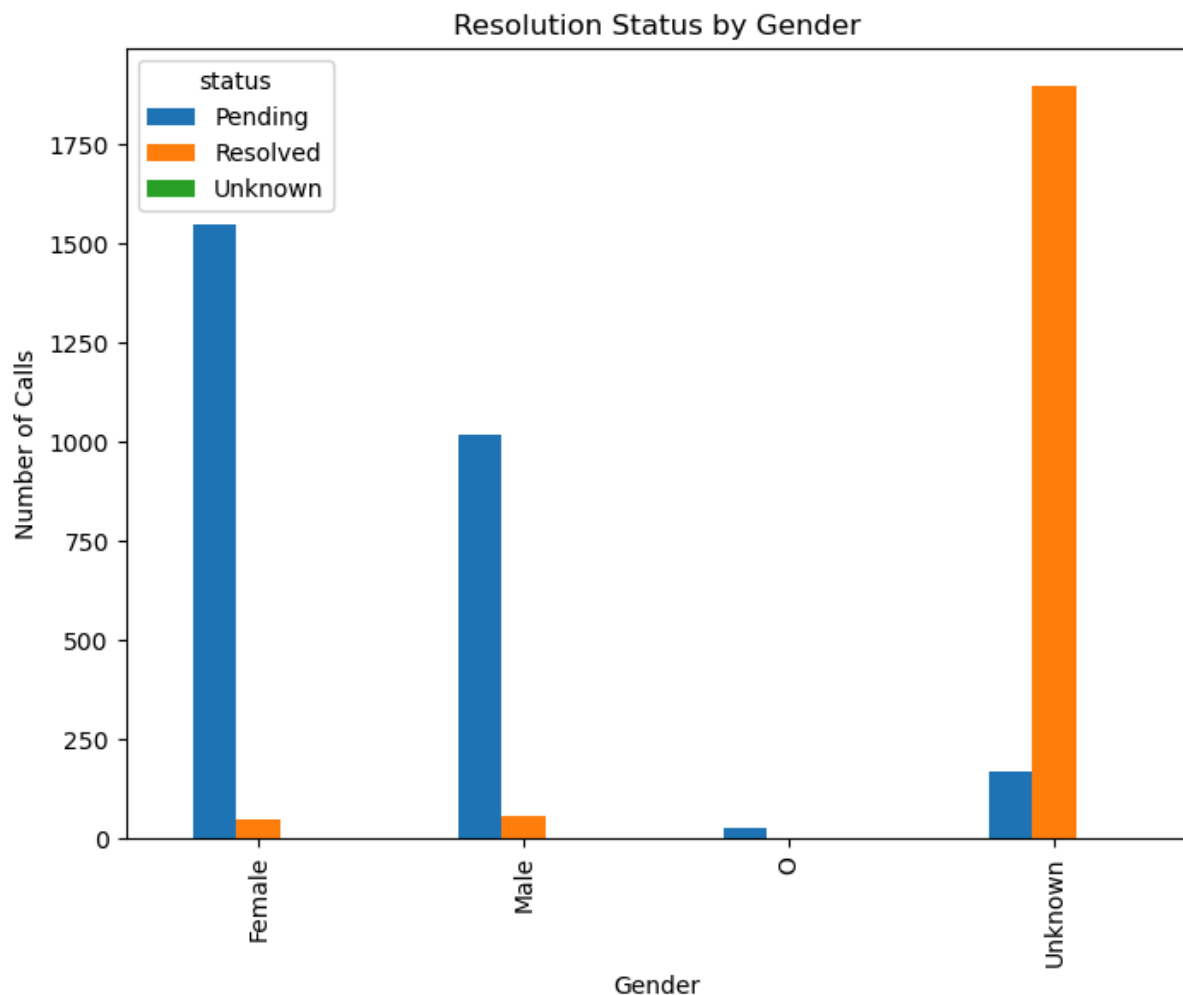
Analysis of Resolution Status by Gender

- Female:
    Pending: 1,546
    Resolved: 49

- Male:
    Pending: 1,019
    Resolved: 55

- Unknown:
   Pending: 171
    Resolved: 1,896

Unknown: 1

Visualization of Resolution Status by Gender

- **Female:** The majority of cases are pending, with a small number resolved.
- **Male:** Similar to females, most cases are pending, with a few resolved.
- **'O':** All cases are pending.
- **Unknown:** The majority of cases are resolved, with some pending and one unknown.

Conclusion and Insights

- **Pending Status Dominance:** Both females and males have a high number of pending cases compared to resolved ones.
- **Unknown Gender Anomaly:** The 'Unknown' gender category has a significantly high number of resolved cases, which is an anomaly compared to other genders.
- **Minimal Data for 'O' Gender:** The 'O' gender category has very few cases, all of which are pending.

Q3. Are there any trends in the number of calls received over time (created_at)?
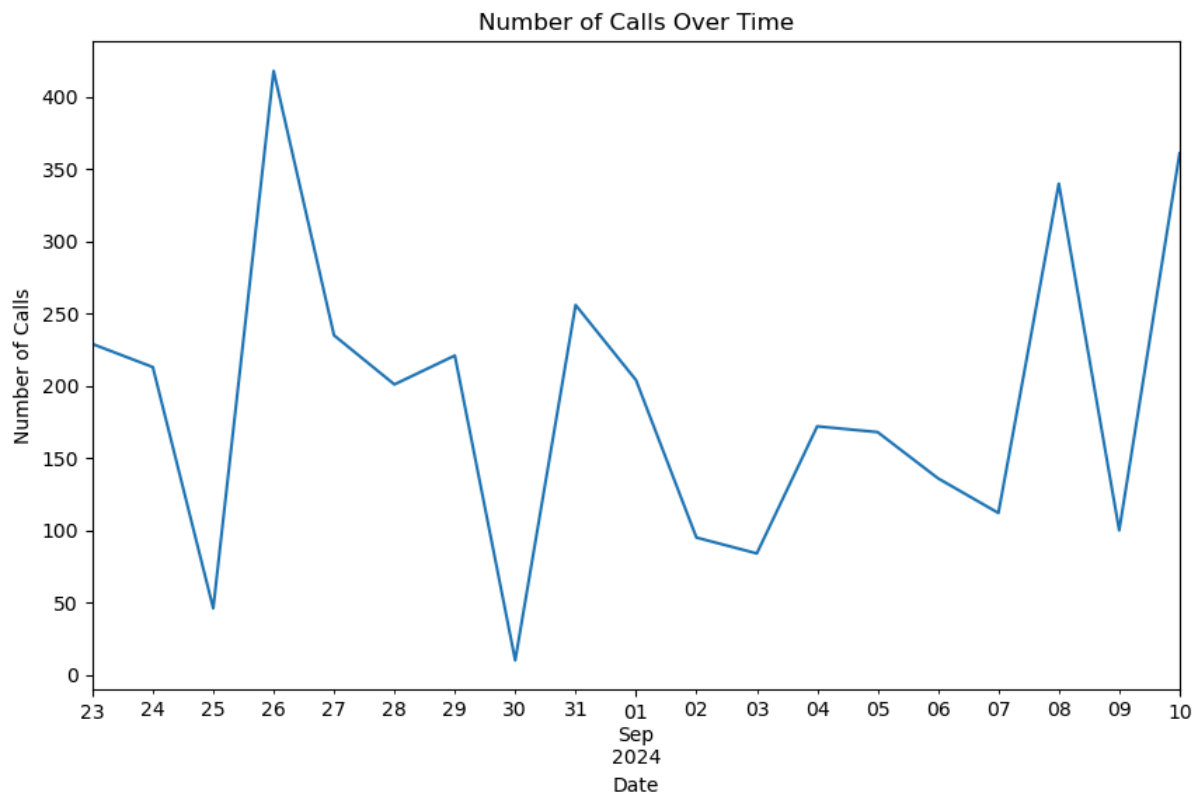
Analysis of Call Data

- **Date Range:** The data spans from August 23, 2024, to September 10, 2024.
- **Extremes:** The maximum number of calls received in a day is 4, and the maximum is 418.

Visualization of Call Trends

- **General Trend:** The number of calls fluctuates over the observed period.
- **Peak:** There is a significant spike in the number of calls on August 26, 2024, with 418 calls.
- **Low Points:** The lowest number of calls is observed on August 30, 2024, with only 10 calls.
- **Variability:** The number of calls shows high variability, with several peaks and troughs throughout the period.

Q4. What is the correlation between the type of caller need and the feedback status?

Contingency Table between Caller Need and Feedback Status
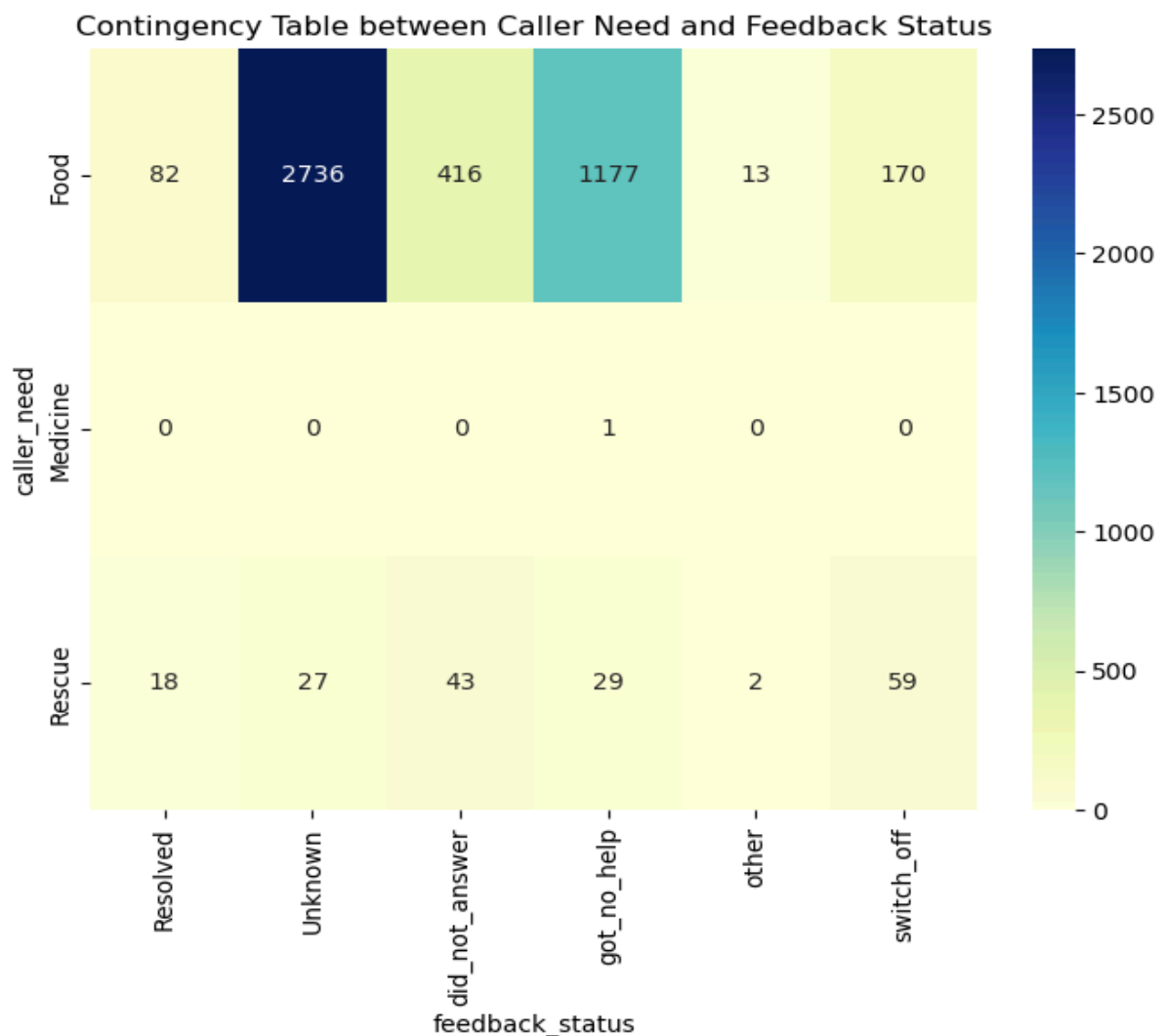
Contingency Table Analysis

- **High Frequency of 'Unknown' and 'Got no help':** The 'Unknown' and 'Got no help' categories have the highest frequencies, with means of 923 and 402.33, respectively.
- **Zero Values in Some Categories:** Several categories, such as 'Resolved' and 'Did not answer', have zero values in some instances, indicating no feedback was recorded for those caller needs in certain cases.

Conclusion and Insights

- **Consistent Feedback Patterns:** The strong correlations across most categories suggest that feedback status is consistently related to the type of caller need.
- **Variability in 'Switch Off' Category:** The slightly lower correlation for the 'Switch Off' category indicates that this caller needs might have more varied feedback statuses, potentially due to different handling or outcomes.

Q5. Most Common Needs in Districts with the Highest Number of Unresolved Cases

Most Common Needs in Districts with Highest Unresolved Cases
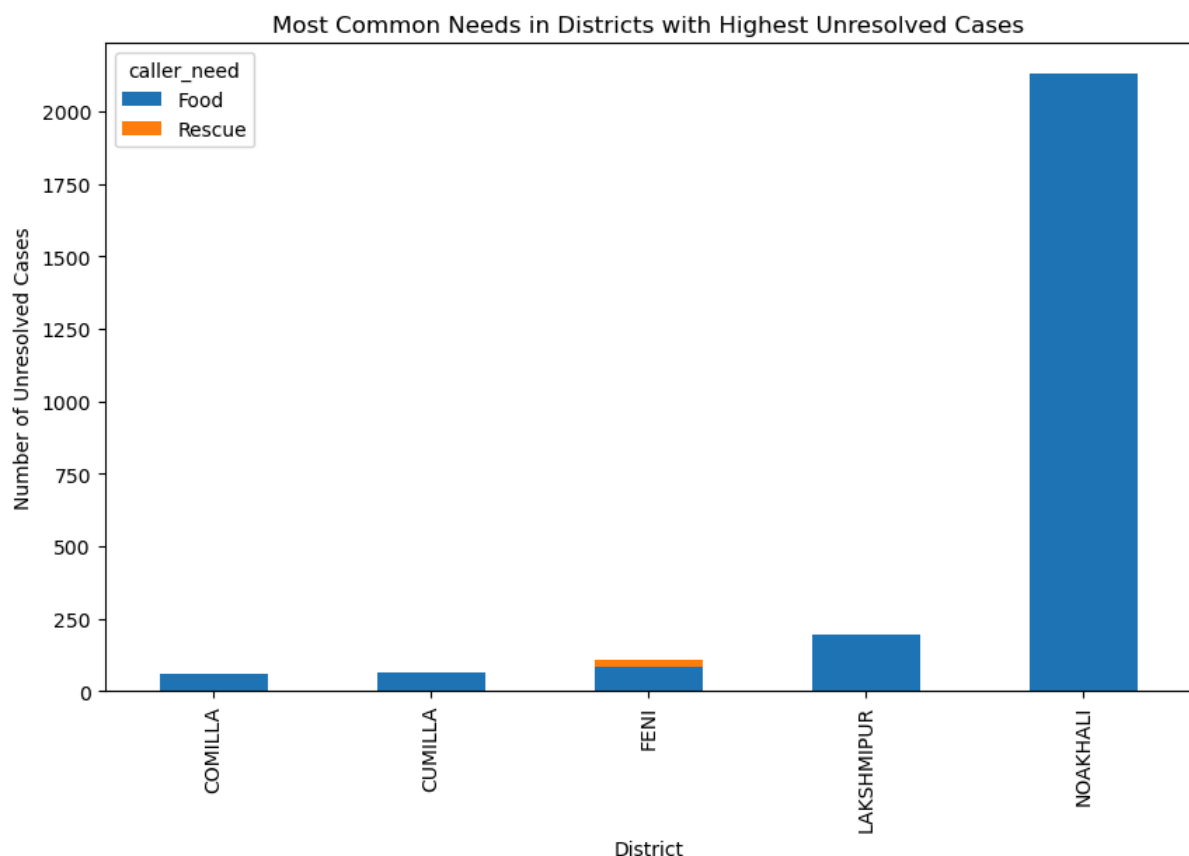
The analysis focused on districts with the highest number of unresolved cases and their most common needs. The top districts with unresolved cases were:

1. Noakhali
2. Lakshmipur
3. Feni
4. Cumilla
5. Comilla

The most common needs across these districts were:
- **Food:** This was the primary need in all districts, with Noakhali having the highest demand (2,129 cases).
- **Rescue:** Only Feni and Cumilla reported rescue needs, with Feni having a higher number (25 cases).

A stacked bar chart was used to visualize this data, clearly showing the dominance of food needs across all districts, with Noakhali standing out significantly.

Q6. Difference in Resolution Time Across Different Types of Caller Needs

Resolution Time Distribution by Caller Need
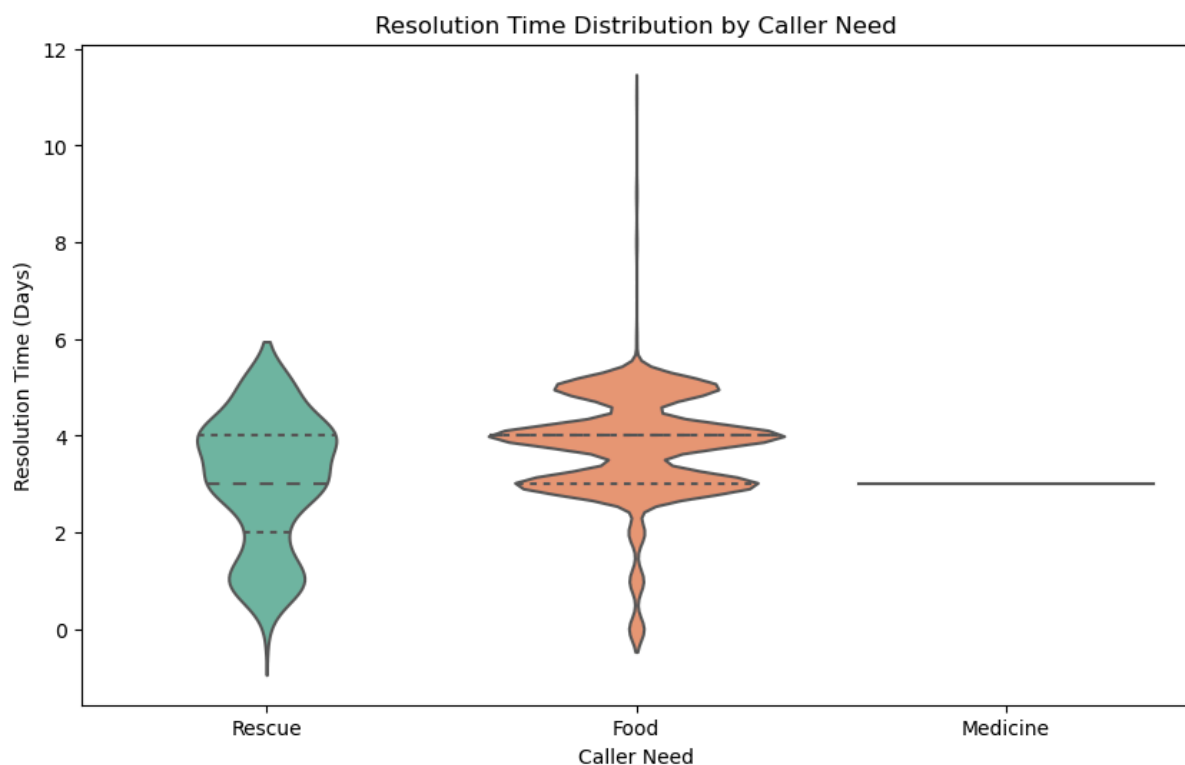
This analysis examined how resolution times varied based on the type of need reported. The data was visualized using a violin plot, which showed the distribution of resolution times for different caller needs. While the specific results aren't provided in the code snippet, this type of analysis can reveal if certain needs are addressed more quickly than others.

## Q7. Frequency of Different Feedback Statuses Over Time
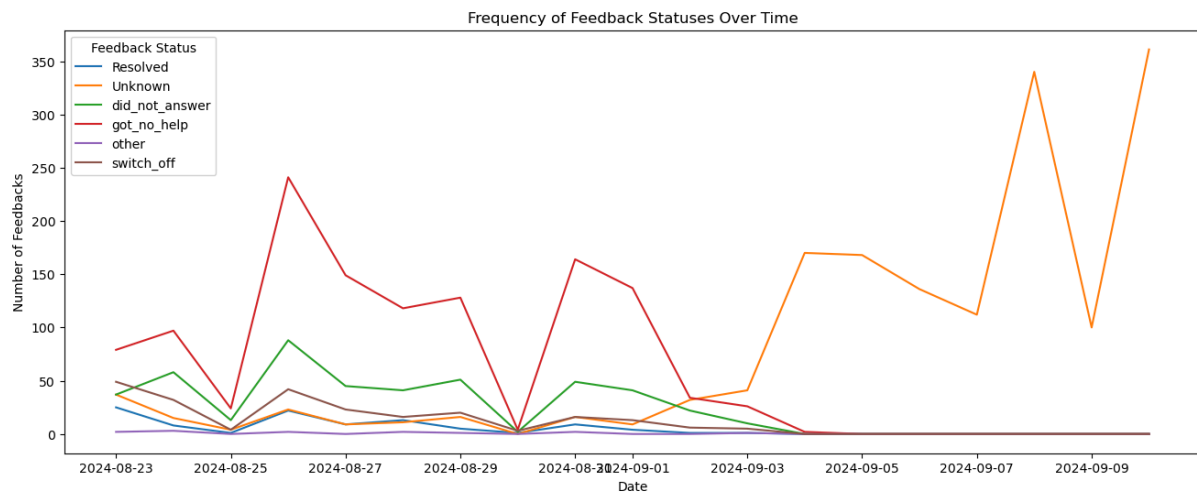
The analysis tracked how different feedback statuses changed over time. The feedback statuses included:
- Resolved
- Unknown
- Did not answer
- Got no help
- Other
- Switch off

A line chart was created to visualize these trends, allowing for the identification of patterns in feedback over the course of the flood response.

Q8. Correlation Matrix Including Feedback Status and Other Variables

Correlation Matrix Including Feedback Status and Other Variables
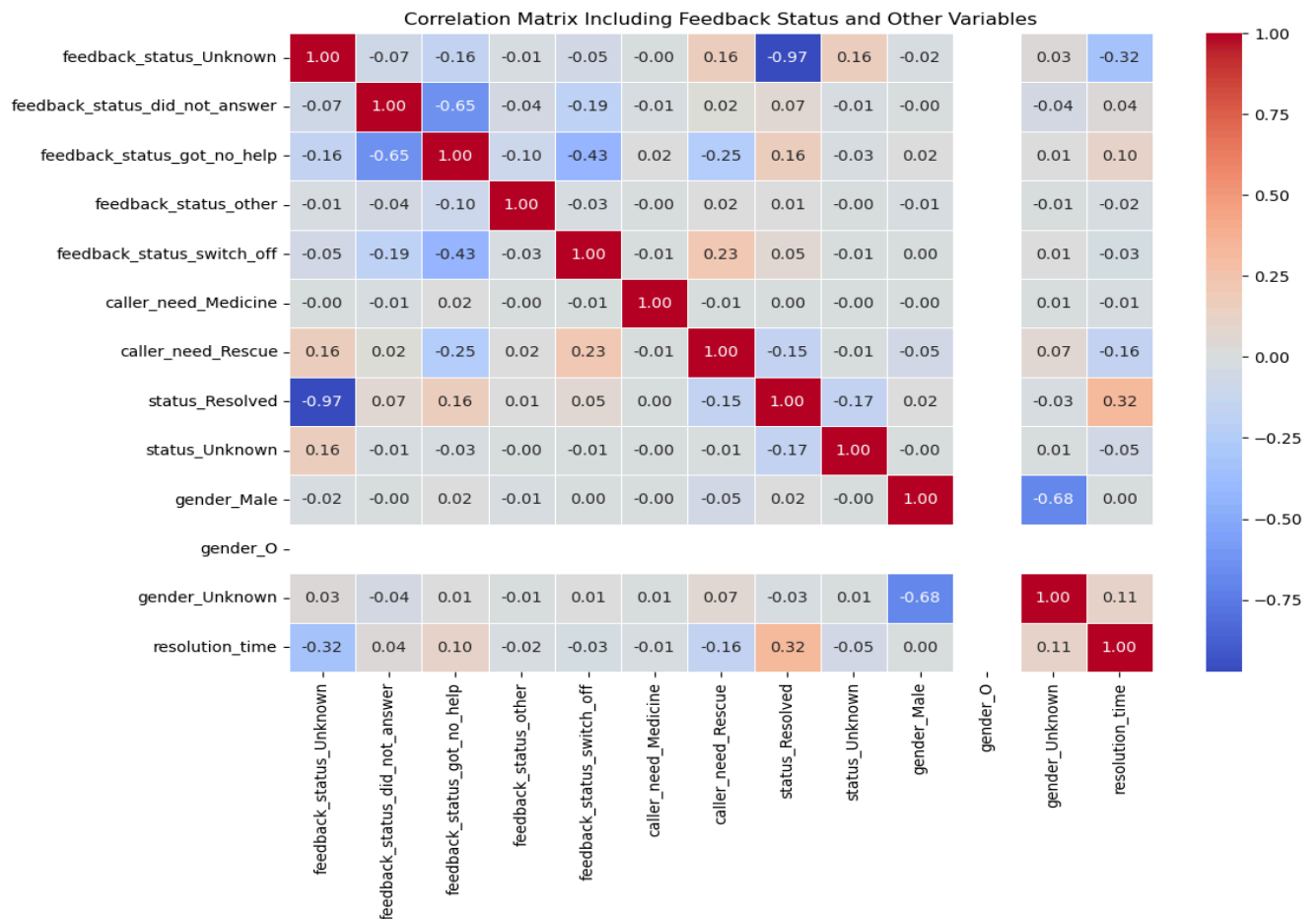
A correlation matrix was computed to understand the relationships between various factors, including feedback status, caller needs, and resolution time. This was visualized using a heatmap, which can reveal strong positive or negative correlations between different variables in the dataset.

Predictive Modeling for Caller Needs

The final analysis used linear regression to predict caller needs for the next two days:
1. The data was prepared by converting dates to a numerical format.
2. A linear regression model was trained on the historical data.
3. The model was then used to predict caller needs for the next two days.

The results showed:
Predicted caller needs for the next two days: [19.61109253, 19.60455028]

This suggests that the model expects around 19-20 calls per day for the next two days, assuming the trend continues as per the historical data. This predictive model could be valuable for resource allocation and preparedness in ongoing flood response efforts.