

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

# Load the dataset
file_path = "project_data.csv"
df = pd.read_csv(file_path)

# Basic info and first look
print("\nDataset Info:\n")
print(df.info())
print("\nFirst 5 Rows:\n")
print(df.head())
print("\nSummary Statistics:\n")
print(df.describe(include='all'))
print("\nMissing Values:\n")
print(df.isnull().sum())

# Check for duplicates
print("\nDuplicate Rows:\n")
print(df.duplicated().sum())

# Visualize response type distribution
plt.figure(figsize=(6, 4))
sns.countplot(x='response_type', data=df, palette='viridis')
plt.title('Distribution of Response Types')
plt.show()

# Visualize topic category distribution
plt.figure(figsize=(8, 5))
sns.countplot(y='topic_category', data=df, palette='mako')
plt.title('Distribution of Topic Categories')
plt.show()

# Visualize user intent distribution
plt.figure(figsize=(8, 6))
sns.countplot(y='user_intent', data=df, palette='rocket')
plt.title('Distribution of User Intents')
plt.show()

# Cross-tabulation of response type and topic category
plt.figure(figsize=(8, 6))
sns.countplot(x='response_type', hue='topic_category', data=df, palette='coolwarm')
plt.title('Response Type by Topic Category')
plt.legend(bbox_to_anchor=(1.05, 1), loc='upper left')
plt.show()

# Cross-tabulation of response type and user intent
plt.figure(figsize=(10, 8))
sns.countplot(y='user_intent', hue='response_type', data=df, palette='Set2')
plt.title('Response Type by User Intent')
plt.legend(bbox_to_anchor=(1.05, 1), loc='upper left')
plt.show()
```



Dataset Info:

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 29 entries, 0 to 28
Data columns (total 5 columns):
#   Column                Non-Null Count  Dtype
---  -
0   response_type          29 non-null    object
1   topic_category         29 non-null    object
2   user_intent            29 non-null    object
3   suggested_action       29 non-null    object
dtypes: object(5)
memory usage: 1.3+ KB
None
```

First 5 Rows:

```

0   How can I file a complaint?      response_type \
1   What are the penalties for drug trafficking?      Text
2   How do I check my case status?      Text
3   I need help with immigration laws.      Form
4   How do I report a crime?      Link
                                Contact

0   Legal Complaints      Filing a Complaint      user_intent \
1   Criminal Law      Seeking Information
2   Case Status/Tracking      Checking Case Status
3   Immigration Law      Seeking Information
4   Criminal Law      Reporting a Crime

0   Visit the complaint form page [https://www.doj...      suggested_action
1   Learn more about drug trafficking penalties [h...
2   Please provide your case number to track your ...
3   Visit the immigration laws page [https://www.d...
4   Click here to speak with a representative for ...
```

Summary Statistics:

```

count      29      29      29
unique      29      4      19
top   How can I file a complaint?      Text      Criminal Law
freq      1      18      5

count      29      29
unique      22      29
top   Seeking Information      Visit the complaint form page [https://www.doj...
freq      4      1
```

Missing Values:

```

0
response_type      0
topic_category      0
user_intent        0
suggested_action   0
dtype: int64
```

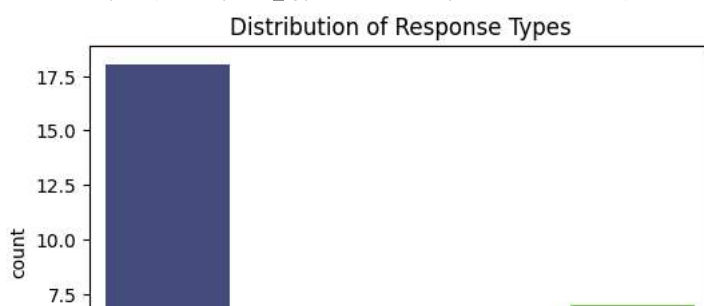
Duplicate Rows:

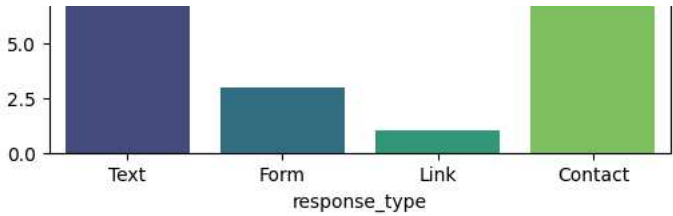
```

0
<ipython-input-1-d59d2298dac5>:25: FutureWarning:
```

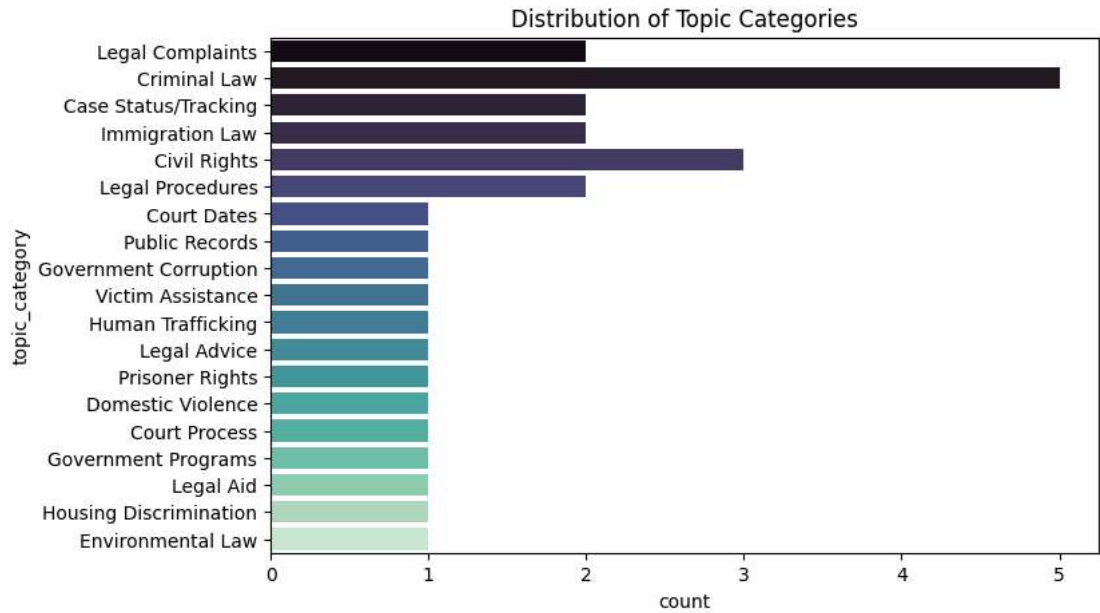
Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `leg`

```
sns.countplot(x='response_type', data=df, palette='viridis')
```





```
<ipython-input-1-d59d2298dac5>:31: FutureWarning:  
Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `y` variable to `hue` and set `leg  
sns.countplot(y='topic_category', data=df, palette='mako')
```



```
<ipython-input-1-d59d2298dac5>:37: FutureWarning:  
Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `y` variable to `hue` and set `leg  
sns.countplot(y='user_intent', data=df, palette='rocket')
```

