

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

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

# Basic info and structure
print("Dataset Info:")
print(df.info())
print("\nFirst 5 Rows:")
print(df.head())

# Check for missing values
print("\nMissing Values:")
print(df.isnull().sum())

# Length of questions and responses
df['pattern_length'] = df['pattern'].apply(len)
df['response_length'] = df['response/0'].apply(len)

# Distribution of question lengths
plt.figure(figsize=(8, 6))
sns.histplot(df['pattern_length'], bins=20, kde=True)
plt.title('Distribution of Question Lengths')
plt.xlabel('Length of Questions')
plt.ylabel('Frequency')
plt.show()

# Distribution of response lengths
plt.figure(figsize=(8, 6))
sns.histplot(df['response_length'], bins=20, kde=True)
plt.title('Distribution of Response Lengths')
plt.xlabel('Length of Responses')
plt.ylabel('Frequency')
plt.show()

# Most common words in patterns
from collections import Counter
import itertools

all_words = list(itertools.chain.from_iterable([pattern.split() for pattern in df['pattern']]))
word_freq = Counter(all_words)
common_words = pd.DataFrame(word_freq.most_common(10), columns=['Word', 'Frequency'])

plt.figure(figsize=(8, 6))
sns.barplot(x='Frequency', y='Word', data=common_words)
plt.title('Top 10 Most Common Words in Questions')
plt.show()

print("\nTop 10 Most Common Words in Questions:")
print(common_words)
```

```

Dataset Info:
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 54 entries, 0 to 53
Data columns (total 2 columns):
#   Column      Non-Null Count  Dtype
---  ---
0   pattern      54 non-null     object
1   response/0    54 non-null     object
dtypes: object(2)
memory usage: 996.0+ bytes
None

```

First 5 Rows:

```

      pattern \
0   what is the Department of Justice?
1  what does the Department of Justice do?
2           who heads the DOJ?
3   what is the mission of the DOJ?
4  how does the DOJ protect civil rights?

```

```

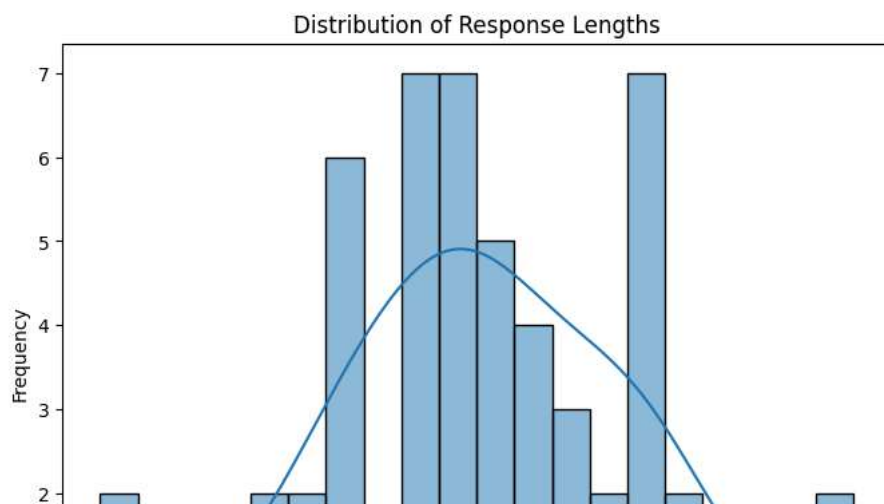
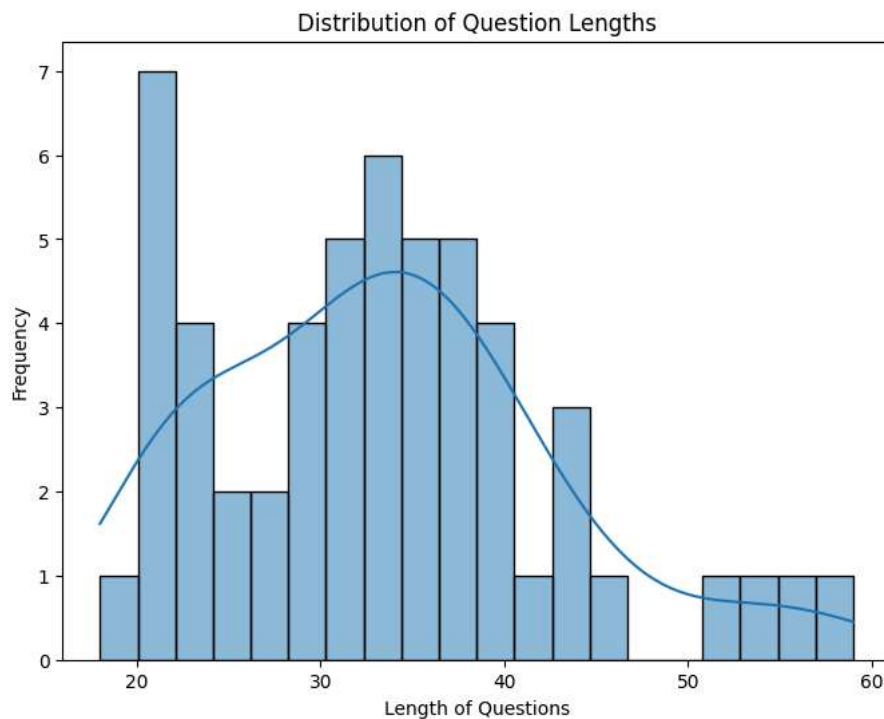
      response/0
0  The Department of Justice (DOJ) is a U.S. gove...
1  The DOJ oversees federal law enforcement, pros...
2  The DOJ is led by the Attorney General of the ...
3  The mission of the DOJ is to enforce federal l...
4  The DOJ protects civil rights through the Civi...

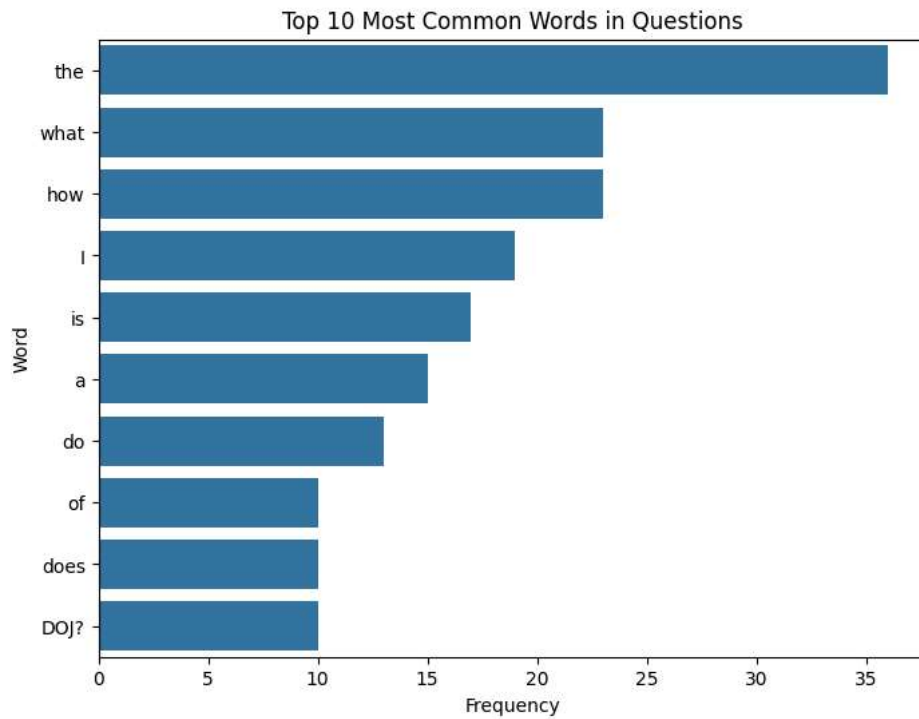
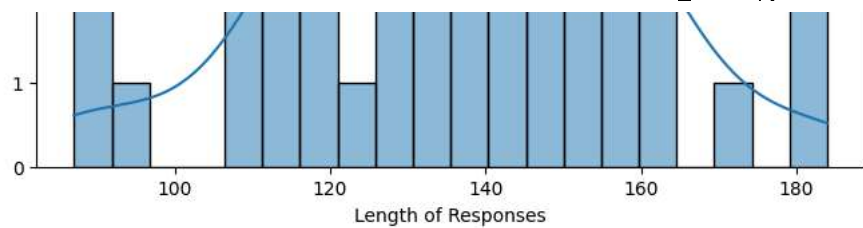
```

```

Missing Values:
pattern      0
response/0    0
dtype: int64

```





Top 10 Most Common Words in Questions:

	Word	Frequency
0	the	36
1	what	23
2	how	23
3	I	19
4	is	17
5	a	15
6	do	13
7	of	10
8	does	10
9	DOJ?	10

Start coding or [generate](#) with AI.