In [12]:

```
import requests
from bs4 import BeautifulSoup
import pandas as pd
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.decomposition import TruncatedSVD
from sklearn.cluster import KMeans
import matplotlib.pyplot as plt
import seaborn as sns
```

In [2]:

```
base_urls = [
    ("bukhari", 97),
    ("muslim", 56),
    ("nasai", 51),
]
```

In [3]:

```
hadith = []
for collection, hadith_count in base_urls:
    print(f"Scraping Hadith from {collection} collection...")
    for hadith_number in range(1, hadith_count + 1):
        hadith_url = f"https://sunnah.com/{collection}/{hadith_number}"
        response = requests.get(hadith_url)
        print("Scraping from:", hadith_url)
        if response.status_code == 200:
            soup = BeautifulSoup(response.content, "html.parser")
            hadith_text = soup.find("div", class_="text_details").get_text(strip=True)
            narrated_by = soup.find("div", class_="hadith_narrated").get_text(strip=True
)
            hadith.append((collection, hadith_number, narrated_by, hadith_text))
        else:
            print(f"Error from {hadith_url}")
```

```
Scraping Hadith from bukhari collection...
Scraping from: https://sunnah.com/bukhari/1
Scraping from: https://sunnah.com/bukhari/2
Scraping from: https://sunnah.com/bukhari/3
Scraping from: https://sunnah.com/bukhari/4
Scraping from: https://sunnah.com/bukhari/5
Scraping from: https://sunnah.com/bukhari/6
Scraping from: https://sunnah.com/bukhari/7
Scraping from: https://sunnah.com/bukhari/8
Scraping from: https://sunnah.com/bukhari/9
Scraping from: https://sunnah.com/bukhari/10
Scraping from: https://sunnah.com/bukhari/11
Scraping from: https://sunnah.com/bukhari/12
Scraping from: https://sunnah.com/bukhari/13
Scraping from: https://sunnah.com/bukhari/14
Scraping from: https://sunnah.com/bukhari/15
Scraping from: https://sunnah.com/bukhari/16
Scraping from: https://sunnah.com/bukhari/17
Scraping from: https://sunnah.com/bukhari/18
Scraping from: https://sunnah.com/bukhari/19
Scraping from: https://sunnah.com/bukhari/20
Scraping from: https://sunnah.com/bukhari/21
Scraping from: https://sunnah.com/bukhari/22
Scraping from: https://sunnah.com/bukhari/23
Scraping from: https://sunnah.com/bukhari/24
Scraping from: https://sunnah.com/bukhari/25
Scraping from: https://sunnah.com/bukhari/26
Scraping from: https://sunnah.com/bukhari/27
Scraping from: https://sunnah.com/bukhari/28
Scraping from: https://sunnah.com/bukhari/29
Scraping from: https://sunnah.com/bukhari/30
Scraping from: https://sunnah.com/bukhari/31
Scraping from: https://sunnah.com/bukhari/32
Scraping from: https://sunnah.com/bukhari/33
Scraping from: https://sunnah.com/bukhari/34
Scraping from: https://sunnah.com/bukhari/35
Scraping from: https://sunnah.com/bukhari/36
Scraping from: https://sunnah.com/bukhari/37
Scraping from: https://sunnah.com/bukhari/38
Scraping from: https://sunnah.com/bukhari/39
Scraping from: https://sunnah.com/bukhari/40
Scraping from: https://sunnah.com/bukhari/41
Scraping from: https://sunnah.com/bukhari/42
Scraping from: https://sunnah.com/bukhari/43
Scraping from: https://sunnah.com/bukhari/44
Scraping from: https://sunnah.com/bukhari/45
Scraping from: https://sunnah.com/bukhari/46
Scraping from: https://sunnah.com/bukhari/47
Scraping from: https://sunnah.com/bukhari/48
Scraping from: https://sunnah.com/bukhari/49
Scraping from: https://sunnah.com/bukhari/50
Scraping from: https://sunnah.com/bukhari/51
Scraping from: https://sunnah.com/bukhari/52
Scraping from: https://sunnah.com/bukhari/53
Scraping from: https://sunnah.com/bukhari/54
Scraping from: https://sunnah.com/bukhari/55
Scraping from: https://sunnah.com/bukhari/56
Scraping from: https://sunnah.com/bukhari/57
Scraping from: https://sunnah.com/bukhari/58
Scraping from: https://sunnah.com/bukhari/59
Scraping from: https://sunnah.com/bukhari/60
Scraping from: https://sunnah.com/bukhari/61
Scraping from: https://sunnah.com/bukhari/62
Scraping from: https://sunnah.com/bukhari/63
Scraping from: https://sunnah.com/bukhari/64
Scraping from: https://sunnah.com/bukhari/65
Scraping from: https://sunnah.com/bukhari/66
Scraping from: https://sunnah.com/bukhari/67
Scraping from: https://sunnah.com/bukhari/68
Scraping from: https://sunnah.com/bukhari/69
Scraping from: https://sunnah.com/bukhari/70
Scraping from: https://sunnah.com/bukhari/71
Scraping from: https://sunnah.com/bukhari/72
Scraping from: https://sunnah.com/bukhari/73
Scraping from: https://sunnah.com/bukhari/74
Scraping from: https://sunnah.com/bukhari/75
Scraping from: https://sunnah.com/bukhari/76
Scraping from: https://sunnah.com/bukhari/77
Scraping from: https://sunnah.com/bukhari/78
Scraping from: https://sunnah.com/bukhari/79
Scraping from: https://sunnah.com/bukhari/80
Scraping from: https://sunnah.com/bukhari/81
Scraping from: https://sunnah.com/bukhari/82
Scraping from: https://sunnah.com/bukhari/83
Scraping from: https://sunnah.com/bukhari/84
Scraping from: https://sunnah.com/bukhari/85
Scraping from: https://sunnah.com/bukhari/86
Scraping from: https://sunnah.com/bukhari/87
Scraping from: https://sunnah.com/bukhari/88
Scraping from: https://sunnah.com/bukhari/89
Scraping from: https://sunnah.com/bukhari/90
Scraping from: https://sunnah.com/bukhari/91
Scraping from: https://sunnah.com/bukhari/92
Scraping from: https://sunnah.com/bukhari/93
Scraping from: https://sunnah.com/bukhari/94
Scraping from: https://sunnah.com/bukhari/95
Scraping from: https://sunnah.com/bukhari/96
Scraping from: https://sunnah.com/bukhari/97
Scraping Hadith from muslim collection...
Scraping from: https://sunnah.com/muslim/1
Scraping from: https://sunnah.com/muslim/2
Scraping from: https://sunnah.com/muslim/3
Scraping from: https://sunnah.com/muslim/4
Scraping from: https://sunnah.com/muslim/5
Scraping from: https://sunnah.com/muslim/6
Scraping from: https://sunnah.com/muslim/7
Scraping from: https://sunnah.com/muslim/8
Scraping from: https://sunnah.com/muslim/9
Scraping from: https://sunnah.com/muslim/10
Scraping from: https://sunnah.com/muslim/11
Scraping from: https://sunnah.com/muslim/12
Scraping from: https://sunnah.com/muslim/13
Scraping from: https://sunnah.com/muslim/14
Scraping from: https://sunnah.com/muslim/15
Scraping from: https://sunnah.com/muslim/16
Scraping from: https://sunnah.com/muslim/17
Scraping from: https://sunnah.com/muslim/18
Scraping from: https://sunnah.com/muslim/19
Scraping from: https://sunnah.com/muslim/20
Scraping from: https://sunnah.com/muslim/21
Scraping from: https://sunnah.com/muslim/22
Scraping from: https://sunnah.com/muslim/23
Scraping from: https://sunnah.com/muslim/24
Scraping from: https://sunnah.com/muslim/25
Scraping from: https://sunnah.com/muslim/26
Scraping from: https://sunnah.com/muslim/27
Scraping from: https://sunnah.com/muslim/28
Scraping from: https://sunnah.com/muslim/29
Scraping from: https://sunnah.com/muslim/30
Scraping from: https://sunnah.com/muslim/31
Scraping from: https://sunnah.com/muslim/32
Scraping from: https://sunnah.com/muslim/33
Scraping from: https://sunnah.com/muslim/34
Scraping from: https://sunnah.com/muslim/35
Scraping from: https://sunnah.com/muslim/36
Scraping from: https://sunnah.com/muslim/37
Scraping from: https://sunnah.com/muslim/38
Scraping from: https://sunnah.com/muslim/39
Scraping from: https://sunnah.com/muslim/40
Scraping from: https://sunnah.com/muslim/41
Scraping from: https://sunnah.com/muslim/42
Scraping from: https://sunnah.com/muslim/43
Scraping from: https://sunnah.com/muslim/44
Scraping from: https://sunnah.com/muslim/45
Scraping from: https://sunnah.com/muslim/46
Scraping from: https://sunnah.com/muslim/47
Scraping from: https://sunnah.com/muslim/48
Scraping from: https://sunnah.com/muslim/49
Scraping from: https://sunnah.com/muslim/50
Scraping from: https://sunnah.com/muslim/51
Scraping from: https://sunnah.com/muslim/52
Scraping from: https://sunnah.com/muslim/53
Scraping from: https://sunnah.com/muslim/54
Scraping from: https://sunnah.com/muslim/55
Scraping from: https://sunnah.com/muslim/56
Scraping Hadith from nasai collection...
Scraping from: https://sunnah.com/nasai/1
Scraping from: https://sunnah.com/nasai/2
Scraping from: https://sunnah.com/nasai/3
Scraping from: https://sunnah.com/nasai/4
Scraping from: https://sunnah.com/nasai/5
Scraping from: https://sunnah.com/nasai/6
Scraping from: https://sunnah.com/nasai/7
Scraping from: https://sunnah.com/nasai/8
Scraping from: https://sunnah.com/nasai/9
Scraping from: https://sunnah.com/nasai/10
Scraping from: https://sunnah.com/nasai/11
Scraping from: https://sunnah.com/nasai/12
Scraping from: https://sunnah.com/nasai/13
Scraping from: https://sunnah.com/nasai/14
Scraping from: https://sunnah.com/nasai/15
Scraping from: https://sunnah.com/nasai/16
Scraping from: https://sunnah.com/nasai/17
Scraping from: https://sunnah.com/nasai/18
Scraping from: https://sunnah.com/nasai/19
Scraping from: https://sunnah.com/nasai/20
Scraping from: https://sunnah.com/nasai/21
Scraping from: https://sunnah.com/nasai/22
Scraping from: https://sunnah.com/nasai/23
Scraping from: https://sunnah.com/nasai/24
Scraping from: https://sunnah.com/nasai/25
Scraping from: https://sunnah.com/nasai/26
Scraping from: https://sunnah.com/nasai/27
Scraping from: https://sunnah.com/nasai/28
Scraping from: https://sunnah.com/nasai/29
Scraping from: https://sunnah.com/nasai/30
Scraping from: https://sunnah.com/nasai/31
Scraping from: https://sunnah.com/nasai/32
Scraping from: https://sunnah.com/nasai/33
Scraping from: https://sunnah.com/nasai/34
Scraping from: https://sunnah.com/nasai/35
Scraping from: https://sunnah.com/nasai/36
Scraping from: https://sunnah.com/nasai/37
Scraping from: https://sunnah.com/nasai/38
Scraping from: https://sunnah.com/nasai/39
Scraping from: https://sunnah.com/nasai/40
Scraping from: https://sunnah.com/nasai/41
Scraping from: https://sunnah.com/nasai/42
Scraping from: https://sunnah.com/nasai/43
Scraping from: https://sunnah.com/nasai/44
Scraping from: https://sunnah.com/nasai/45
Scraping from: https://sunnah.com/nasai/46
Scraping from: https://sunnah.com/nasai/47
Scraping from: https://sunnah.com/nasai/48
Scraping from: https://sunnah.com/nasai/49
Scraping from: https://sunnah.com/nasai/50
Scraping from: https://sunnah.com/nasai/51
```

In [4]:

```
df = pd.DataFrame(hadith, columns=["Collection", "Hadith Number", "Narrated By", "Hadith Text"])

tfidf_vectorizer = TfidfVectorizer(max_df=0.8, max_features=10000)
tfidf_matrix = tfidf_vectorizer.fit_transform(df["Hadith Text"])
svd = TruncatedSVD(n_components=50)
reduced_matrix = svd.fit_transform(tfidf_matrix)
num_clusters = 5
kmeans = KMeans(n_clusters=num_clusters, random_state=42)
df["Cluster"] = kmeans.fit_predict(reduced_matrix)
pillar_criteria = {
    "Shahada": ["faith", "testimony", "witness", "worshipped"],
    "Salat": ["prayer", "ritual", "worship", "mosque", "clean"],
    "Zakat": ["charity", "almsgiving", "poor", "money", "property"],
    "Sawm": ["fasting", "Ramadan", "abstain", "patience"],
    "Hajj": ["pilgrimage", "Mecca", "Kaaba", "Hajj"]
}
```

In [5]:

```
def categorize_hadith(hadith_text):
    for pillar, keywords in pillar_criteria.items():
        for keyword in keywords:
            if keyword in hadith_text:
                return pillar
    return "Others"

df["Pillar"] = df["Hadith Text"].apply(categorize_hadith)
```

In [10]:

```
df.to_csv("hadith.csv", index = False)
```

In [11]:

```
data = pd.read_csv("hadith.csv")
data
```

Out[11]:

	Collection	Hadith Number	Narrated By	Hadith Text	Cluster	Pillar
0	bukhari	1	Narrated 'Umar bin Al-Khattab:	I heard Allah's Messenger (□) saying, "The rew...	2	Others
1	bukhari	2	Narrated Ibn 'Umar:	Allah's Messenger (□) said: Islam is based on ...	0	Shahada
2	bukhari	3	Narrated Abu Huraira:	While the Prophet (□) was saying something in ...	3	Others
3	bukhari	4	Narrated Abu Huraira:	Allah's Messenger (□) said, "The prayer of a p...	2	Salat
4	bukhari	5	Narrated `Aisha:	Whenever the Prophet (□) took a bath after Jan...	1	Salat
...	...	...	...	...	...	...
199	nasai	47	It was narrated from Abu Hurairah that:	The Messenger of Allah [SAW] was asked: "Which...	1	Others
200	nasai	48	It was narrated from 'Aishah that:	The Messenger of Allah [SAW] said: "Ten things...	1	Others
201	nasai	49	It was narrated from 'Abdullah bin 'Amr bin Al...	The Prophet [SAW] said: "Those who are just an...	1	Others
202	nasai	50	It was narrated from Mu'adh bin 'Abdullah that...	"It was raining and dark, and we were waiting ...	3	Salat
203	nasai	51	It was narrated from 'Umar that:	When the prohibition of Khamr was revealed, 'U...	3	Salat

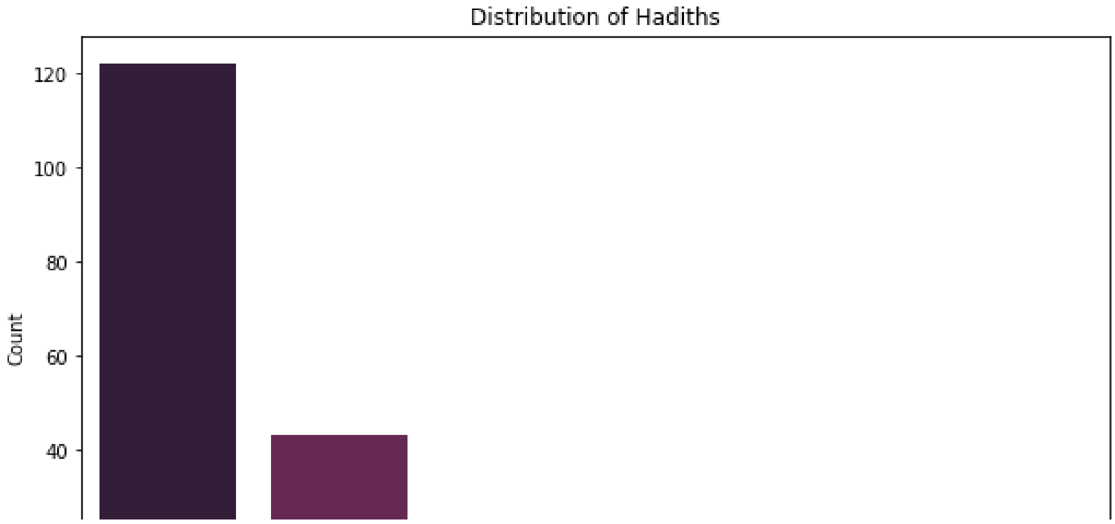
204 rows × 6 columns

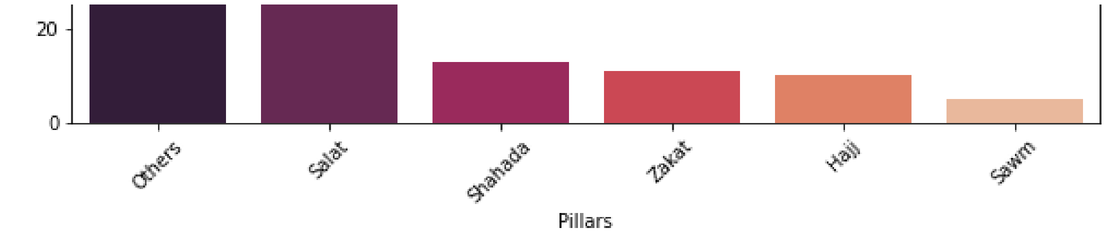
In [13]:

```
counts = data['Pillar'].value_counts()
plt.figure(figsize=(10, 6))
sns.barplot(x = counts.index, y = counts.values, palette = 'rocket')
plt.title('Distribution of Hadiths')
plt.xlabel('Pillars')
plt.ylabel('Count')
plt.xticks(rotation=45)
```

Out[13]:

```
(array([0, 1, 2, 3, 4, 5]),
 [Text(0, 0, 'Others'),
  Text(1, 0, 'Salat'),
  Text(2, 0, 'Shahada'),
  Text(3, 0, 'Zakat'),
  Text(4, 0, 'Hajj'),
  Text(5, 0, 'Sawm')])
```



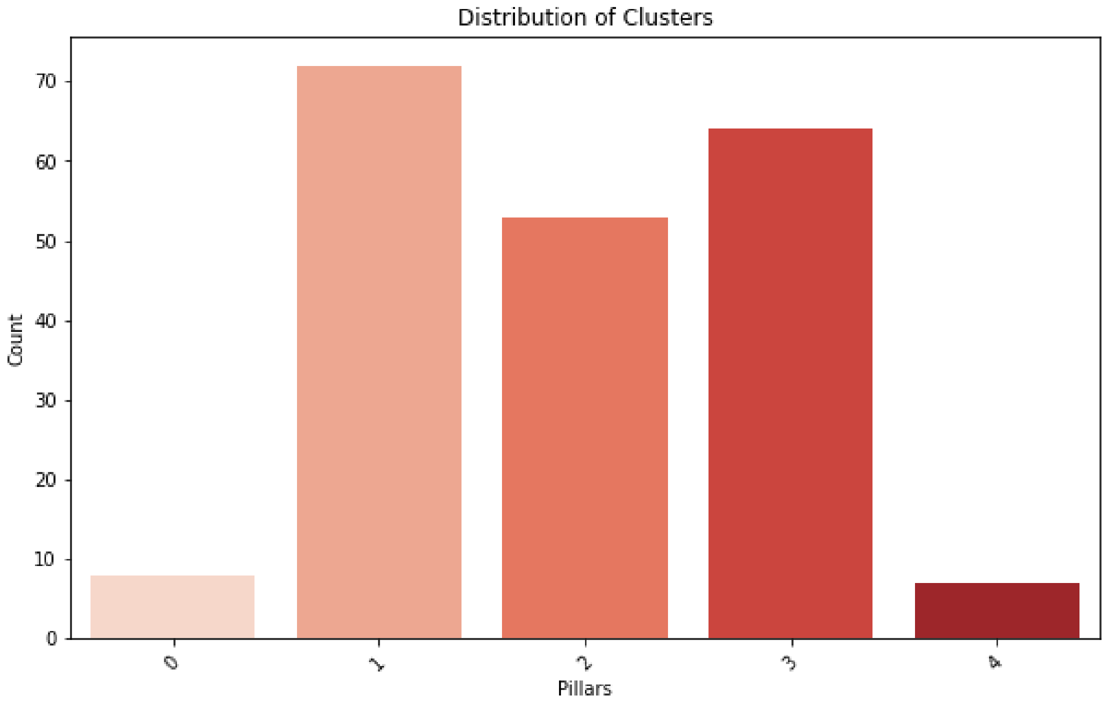


In [16]:

```
counts = data['Cluster'].value_counts()
plt.figure(figsize=(10, 6))
sns.barplot(x = counts.index, y = counts.values, palette = 'Reds')
plt.title('Distribution of Clusters')
plt.xlabel('Pillars')
plt.ylabel('Count')
plt.xticks(rotation=45)
```

Out[16]:

```
(array([0, 1, 2, 3, 4]),
 [Text(0, 0, '0'),
  Text(1, 0, '1'),
  Text(2, 0, '2'),
  Text(3, 0, '3'),
  Text(4, 0, '4')])
```

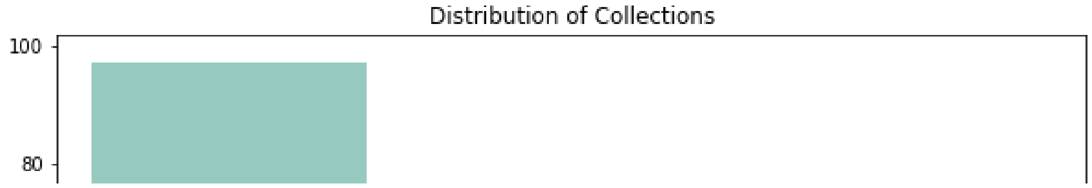


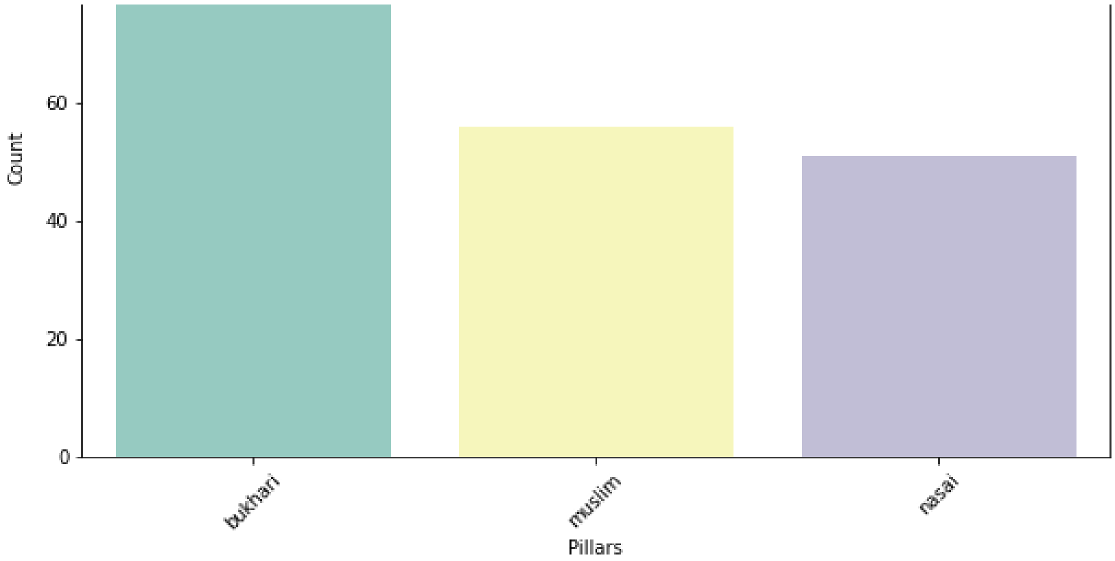
In [18]:

```
counts = data['Collection'].value_counts()
plt.figure(figsize=(10, 6))
sns.barplot(x = counts.index, y = counts.values, palette = 'Set3')
plt.title('Distribution of Collections')
plt.xlabel('Pillars')
plt.ylabel('Count')
plt.xticks(rotation=45)
```

Out[18]:

```
(array([0, 1, 2]),
 [Text(0, 0, 'bukhari'), Text(1, 0, 'muslim'), Text(2, 0, 'nasai')])
```



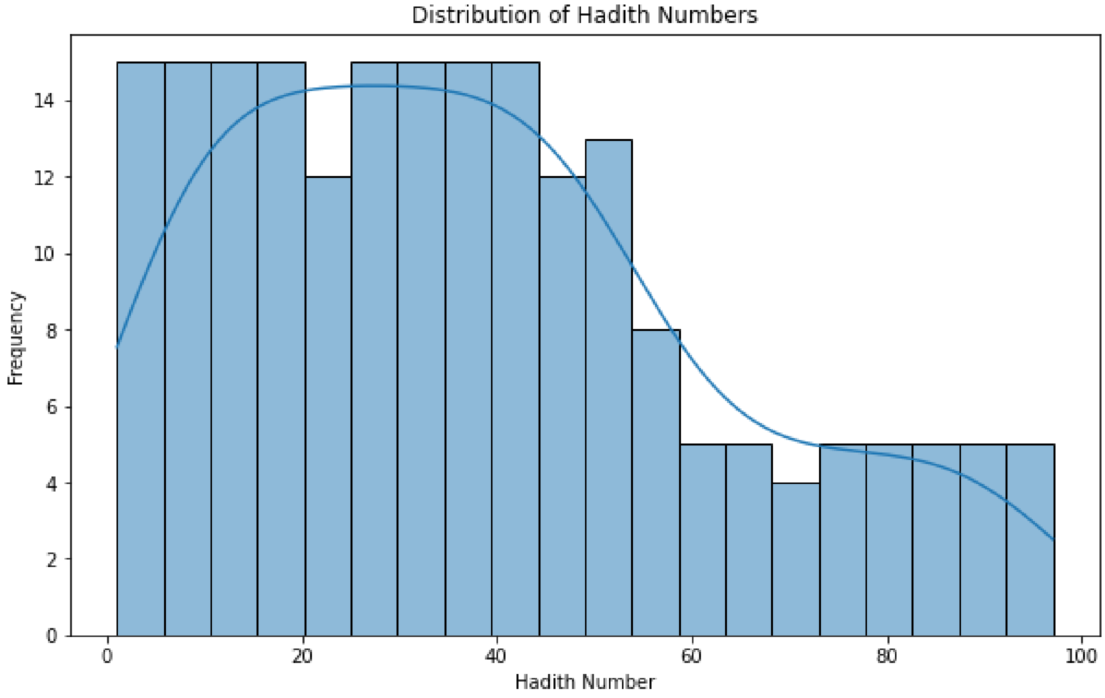


In [19]:

```
plt.figure(figsize=(10, 6))
sns.histplot(data['Hadith Number'], bins=20, kde=True)
plt.xlabel('Hadith Number')
plt.ylabel('Frequency')
plt.title('Distribution of Hadith Numbers')
```

Out[19]:

```
Text(0.5, 1.0, 'Distribution of Hadith Numbers')
```



In [ ]: