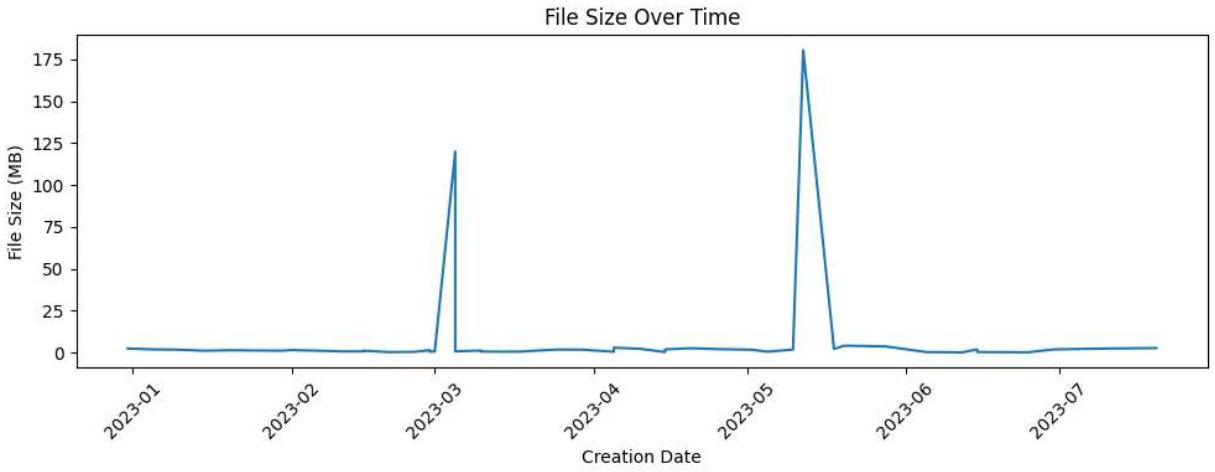
Perform the Exploratory Data Analysis on your domain-based dataset and demonstrate the retrieved insights using "Matplotlib" modules. Visualize hidden insights using appropriate plots (graphs) [Usage of line plot and scatter plot are mandatory]

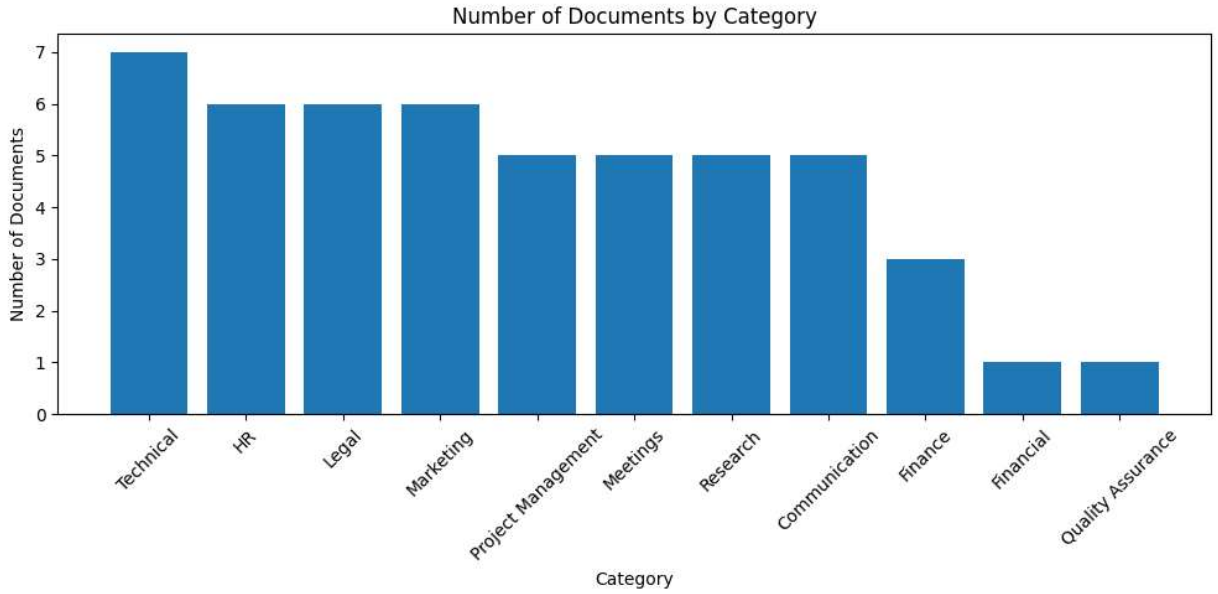
In [ ]:
```
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
import seaborn as sns
```

In [ ]:
```
df= pd.read_csv(r'./DataSet.csv')
```

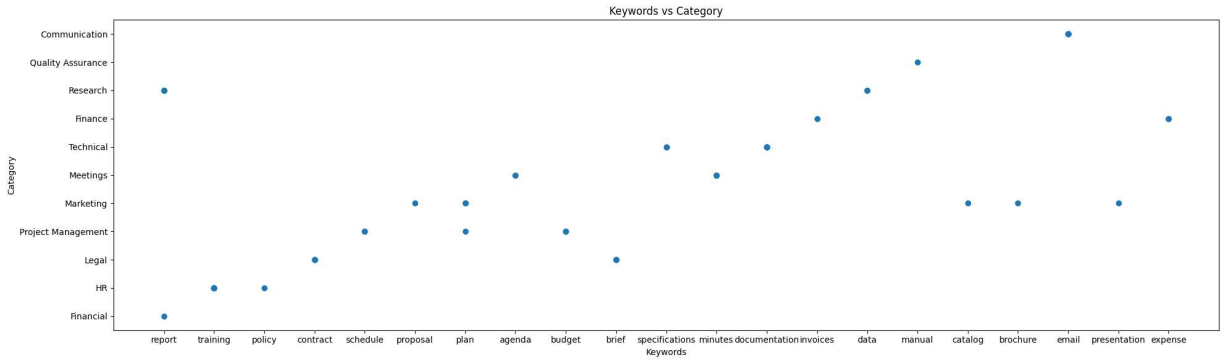
In [ ]:
```
plt.figure(figsize=(10, 4))
df['Creation Date'] = pd.to_datetime(df['Creation Date'], format='%d-%m-%Y')
df.sort_values(by='Creation Date', inplace=True)
plt.plot(df['Creation Date'], df['File Size (MB)'])
plt.title('File Size Over Time')
plt.xlabel('Creation Date')
plt.ylabel('File Size (MB)')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```



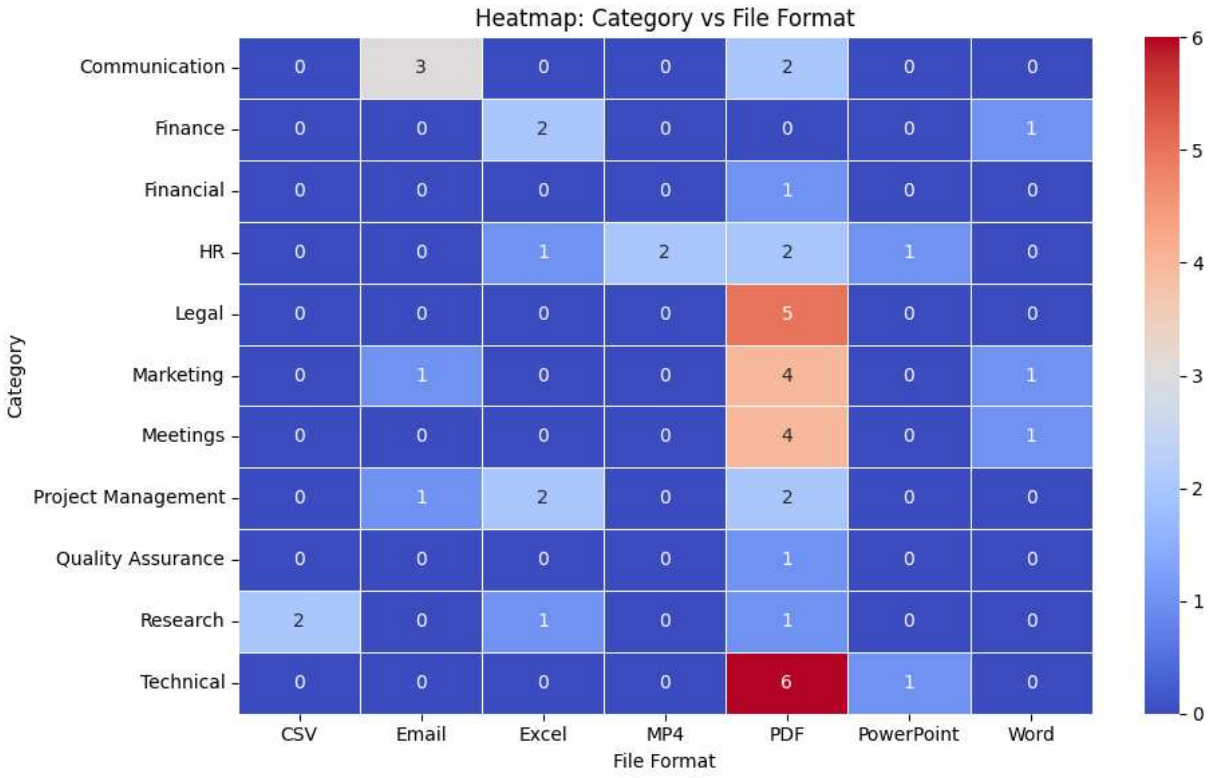
In [ ]:
```
plt.figure(figsize=(10, 5))
category_counts = df['Category'].value_counts()
plt.bar(category_counts.index, category_counts.values)
plt.title('Number of Documents by Category')
plt.xlabel('Category')
plt.ylabel('Number of Documents')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```



In [ ]:
```
plt.figure(figsize=(20, 6))
plt.scatter(df['Keywords'], df['Category'], alpha=1)
plt.title('Keywords vs Category')
plt.xlabel('Keywords')
plt.ylabel('Category')
plt.tight_layout()
plt.show()
```



In [ ]:
```
import seaborn as sns
contingency_table = pd.crosstab(df['Category'], df['File Format'])
plt.figure(figsize=(10, 6))
sns.heatmap(contingency_table, annot=True, cmap='coolwarm', fmt='d', linewidths=0.5
plt.title('Heatmap: Category vs File Format')
plt.xlabel('File Format')
plt.ylabel('Category')
plt.tight_layout()
plt.show()
```


Heatmap: Category vs File Format
Communication 0 3 0 0 2 0 0
Finance 0 0 2 0 0 0 1
Financial 0 0 0 0 1 0 0
HR 0 0 1 2 2 1 0
Legal 0 0 0 0 5 0 0
Marketing 0 1 0 0 4 0 1
Meetings 0 0 0 0 4 0 1
Project Management 0 1 2 0 2 0 0
Quality Assurance 0 0 0 0 1 0 0
Research 2 0 1 0 1 0 0
Technical 0 0 0 0 6 1 0
Category
CSV Email Excel MP4 PDF PowerPoint Word
File Format
6 5 4 3 2 1 0