兒福聯盟議題程式碼說明 應徵者:曹智凱

1. 去除重複及不要的欄位

import pandas as pd

data = pd.read\_csv('合併資料.csv')

# df = data.drop\_duplicates(subset=['Content'], keep='first')

# print(df.info())

df2 = df.drop(['index','Time','Page ID','Post Type','URL'],axis = 1)

print(df2.info())

2. 圖一折線圖製作

import matplotlib.pyplot as plt

import seaborn as sns

plt.rcParams["font.family"] = "Microsoft YaHei"

plt.rcParams["font.size"] = 18

daily\_interaction = df2.groupby('Date').agg({

'Engagement Score': 'sum',

'Comments': 'sum',

'Shares': 'sum'

}).reset\_index()

# 互動量 = 評論數加分享數

daily\_interaction['interaction'] = daily\_interaction['Comments'] + daily\_interaction['Shares']

plt.figure(figsize=(8, 6))

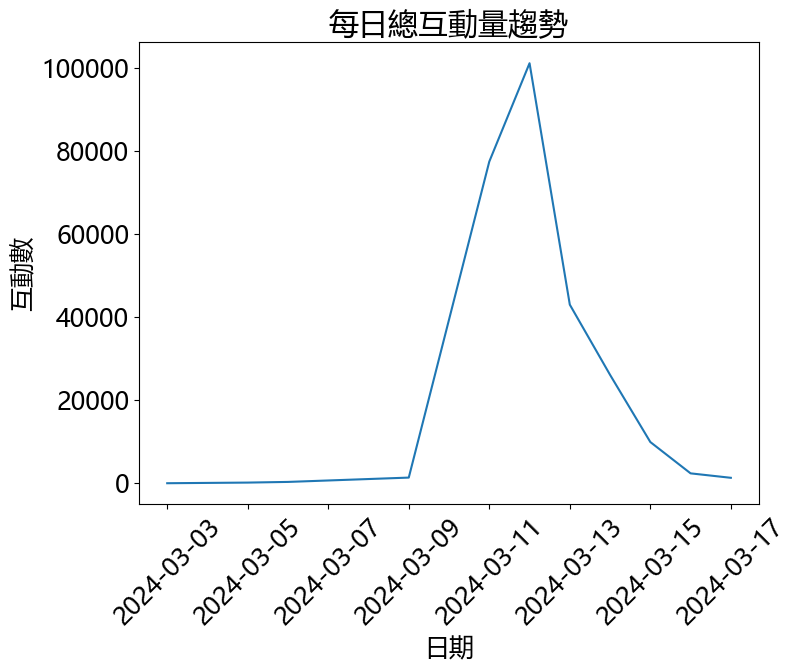
plt.plot(daily\_interaction['Date'], daily\_interaction['interaction'])

plt.title('每日總互動量趨勢')

plt.xlabel('日期')

plt.ylabel('互動數')

plt.xticks(rotation=45)



3.只保留3/11-3-13的資料

# 設定只想保留資料日期範圍

start\_date = '2024-03-11'

end\_date = '2024-03-13'

# 篩選指定日期範圍的資料

mask = (df2['Date'] >= start\_date) & (df2['Date'] <= end\_date)

filtered\_df = df2[mask]

filtered\_df.reset\_index(inplace = True)

print(filtered\_df.info())

4. 三大類議題分類(訪視、管理及監督)

relevant\_keywords = lst\_view

data\_view = filtered\_df[filtered\_df['Content'].apply(is\_relevant)]

data\_view['Issue'] = '訪視'

relevant\_keywords = lst\_man

data\_man = filtered\_df[filtered\_df['Content'].apply(is\_relevant)]

data\_man['Issue'] = '管理'

relevant\_keywords = lst\_sup

data\_sup = filtered\_df[filtered\_df['Content'].apply(is\_relevant)]

data\_sup['Issue'] = '監督'

5.計算表一欄位及數據

def calculate\_stats\_df(df):

# 計算統計數據

stats = {

'影響力分數': [round(df['Engagement Score'].mean(), 0)],

'發文數' : [len(df)],

'互動數': [df['Comments'].sum() + df['Shares'].sum()],

'平均按讚數': [round(df['Like'].mean(),0)]

}

stats\_df = pd.DataFrame(stats)

return stats\_df

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 議題 | 影響力分數 | 發文數 | 互動數 | 平均按讚數 |
| 訪視 | 1093 | 43 | 75912 | 3282 |
| 管理 | 1167 | 60 | 88410 | 4300 |
| 監督 | 1120 | 18 | 26878 | 4189 |

6.製作圖二互動數圓餅圖

result= result.drop('index', axis=1).set\_index('議題')

result2 = result.T

interactions = result2.loc['互動數']

# 計算佔比百分比

percentages = (interactions / interactions.sum()) \* 100

# 創建圓餅圖

plt.figure(figsize=(10, 8))

plt.pie(percentages,

labels=[f'{label}\n{value:.1f}%' for label, value in percentages.items()],

autopct='',

startangle=90)

# 添加標題

plt.title('各議題互動數佔比', pad=20, size=24)

# 添加圖例

plt.legend(interactions.index, title="議題", loc="center left", bbox\_to\_anchor=(1, 0, 0.5, 1))

# 調整布局

plt.tight\_layout()

# 顯示圖表

plt.show()

