

```
1 import pandas as pd
2 import matplotlib.pyplot as plt
3 import seaborn as sns
4 import re # Import for regular expressions
5
6 df1 = pd.read_excel('Tech_Info_online.xlsx')
7 df2 = pd.read_excel('Tech_Info_business.xlsx')
8
9 # Function to detect garbled text (add this)
10 def contains_garbled_text(text):
11     if pd.isnull(text):
12         return False
13     return not all(ord(c) < 128 for c in text)
14
15 # Calculate the number and percentage of garbled rows
16 garbled_rows_df1 = df1['Page title'].apply(contains_garbled_text).sum()
17 garbled_percentage_df1 = (garbled_rows_df1 / len(df1)) * 100
18
19 garbled_rows_df2 = df2['Page title'].apply(contains_garbled_text).sum()
20 garbled_percentage_df2 = (garbled_rows_df2 / len(df2)) * 100
21
22 # Print the results
23 print(f'Number of rows with garbled text in Tech_Info_online.xlsx:
24       {garbled_rows_df1}')
25 print(f'Percentage of rows with garbled text in Tech_Info_online.xlsx:
26       {garbled_percentage_df1:.2f}%')
27 print(f'Number of rows with garbled text in Tech_Info_business.xlsx:
28       {garbled_rows_df2}')
29 print(f'Percentage of rows with garbled text in Tech_Info_business.xlsx:
30       {garbled_percentage_df2:.2f}%')
```

Number of rows with garbled text in Tech\_Info\_online.xlsx: 2161  
Percentage of rows with garbled text in Tech\_Info\_online.xlsx: 4.12%  
Number of rows with garbled text in Tech\_Info\_business.xlsx: 8173  
Percentage of rows with garbled text in Tech\_Info\_business.xlsx: 6.01%

```
1 # Drop rows with garbled page titles (add these lines)
2 df1 = df1[~df1['Page title'].apply(contains_garbled_text)]
3 df2 = df2[~df2['Page title'].apply(contains_garbled_text)]
4
5 # Remove rows with NaN in essential columns (like 'Date')
6 df1 = df1.dropna(subset=['Date'])
7 # Convert Date to datetime format
8 df1['Date'] = pd.to_datetime(df1['Date'].astype(int), format='%Y%m%d', errors='coerce')
9 df1 = df1.dropna(subset=['Date']) # Drop rows where date conversion failed
10
11 # Check for any other missing values
12 print(df1.isnull().sum())
13
14 # Remove rows with NaN in essential columns (like 'Date')
15 df2 = df2.dropna(subset=['Date'])
16 # Convert Date to datetime format
17 df2['Date'] = pd.to_datetime(df2['Date'].astype(int), format='%Y%m%d', errors='coerce')
18 df2 = df2.dropna(subset=['Date']) # Drop rows where date conversion failed
19
20 # Check for any other missing values
21 print(df2.isnull().sum())
```

Date 0  
Page title 0  
Device category 0  
Operating system 0  
Browser 0  
Views 0  
Sessions 0  
New users 0  
Returning users 0  
dtype: int64  
Date 0  
Page title 1  
Device category 0  
Operating system 0  
Browser 0  
Views 0  
Sessions 0  
New users 0  
Returning users 0  
dtype: int64

```
1 # Define classification logic
2 def classify_program(title):
3     if not isinstance(title, str):
4         return 'Other'
5     title = title.lower()
6     if 'business analytics' in title or 'msba' in title:
7         return 'MSBA'
8     elif 'finance' in title or 'msf' in title:
9         return 'MSF'
10    elif 'accounting' in title or 'msa' in title:
11        return 'MSA'
12    elif 'management' in title and 'technology' not in title:
13        return 'MSM'
14    elif 'technology management' in title or 'mstm' in title:
15        return 'MSTM'
16    else:
17        return 'Other'
18
19 # Apply classification
20 df2['Program'] = df2['Page title'].apply(classify_program)
21
22 # Group by Date and Program to count engagements
23 daily_engagement = df2.groupby(['Date', 'Program']).size().reset_index(name='Engagements')
```

1 df2

	Date	Page title	Device category	Operating system	Browser	Views	Sessions	New users	Returning users	Program
0	2025-01-15	MSA Admissions   Gies Master's Degree in Accou...	desktop	Windows	Chrome	627	610	531	42	MSA
1	2025-01-15	Masters in Business   Graduate Programs   Mast...	mobile	Android	Android Webview	593	474	429	18	Other
2	2025-01-15	Masters in Finance   Gies College MSF Program ...	mobile	Android	Chrome	421	388	138	38	MSF
3	2025-01-15	MSM Admissions   Gies College of Business   Ma...	desktop	Windows	Chrome	378	354	261	53	MSM
4	2025-01-15	Masters in Business Analytics Admissions   Gie...	desktop	Windows	Chrome	328	317	156	83	MSBA
...	...	...	...	...	...	...	...	...	...	...
135904	2025-03-14	Southern California grapples with economic fal...	desktop	Windows	Chrome	0	1	0	1	Other
135905	2025-03-14	VITA Program	desktop	Macintosh	Edge	0	1	0	0	Other
135906	2025-03-14	What approaches can help improve the US health...	desktop	Macintosh	Safari	0	1	0	0	Other
135907	2025-03-14	Wymer Hall going geothermal to enhance sustain...	desktop	Windows	Firefox	0	1	0	0	Other
135908	2025-03-14	Yunchuan(Frank) Liu	desktop	Macintosh	Chrome	0	2	0	0	Other

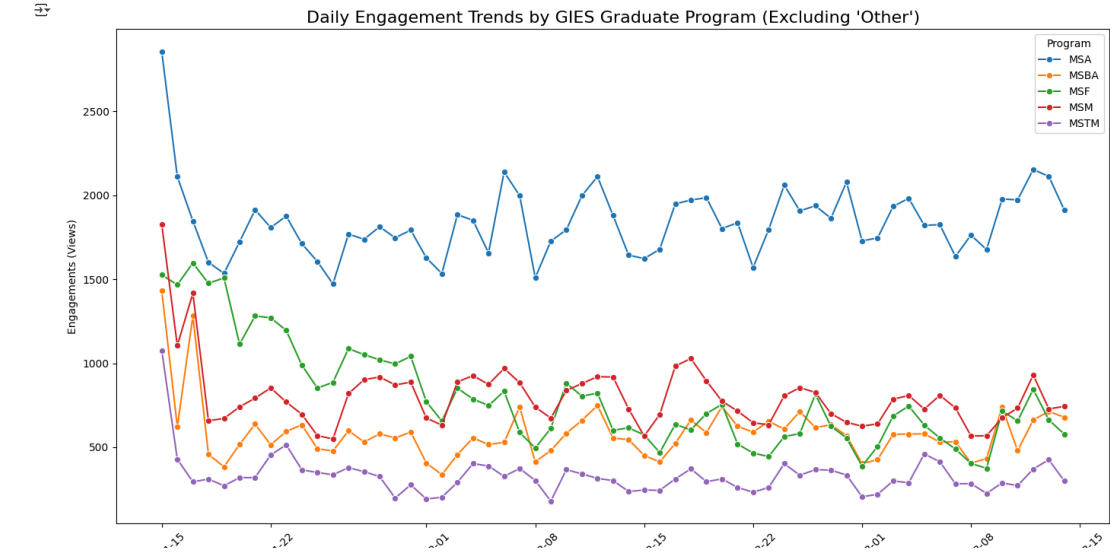
127790 rows × 10 columns

```
1 # Recalculate daily engagement using 'Views'
2 daily_engagement_views = df2.groupby(['Date', 'Program'])['Views'].sum().reset_index(name='Engagements')
3
4 # Compute total views per program
5 total_views = daily_engagement_views.groupby("Program")["Engagements"].sum().reset_index()
6 total_views = total_views.rename(columns={"Engagements": "Total Engagements"})
```

```
7
8 # Merge total into daily data
9 merged_views_df = pd.merge(daily_engagement_views, total_views, on="Program")
10
11 # Sort for clarity
12 merged_views_df = merged_views_df.sort_values(by=["Program", "Date"])
13
14 program_totals_sorted = merged_views_df[['Program', 'Total Engagements']].drop_duplicates().sort_values(by='Total Engagements', ascending=False)
15 program_totals_sorted
```

	Program	Total Engagements
5	Other	344690
0	MSA	108584
3	MSM	47427
2	MSF	46955
1	MSBA	34529
4	MSTM	19429

```
1 # Remove 'Other' category from the data
2 filtered_views_df = merged_views_df[merged_views_df['Program'] != 'Other']
3
4 # Plot daily engagement trends without 'Other'
5 plt.figure(figsize=(14, 8))
6 sns.lineplot(data=filtered_views_df, x="Date", y="Engagements", hue="Program", marker="o")
7
8 # Customize the plot
9 plt.title("Daily Engagement Trends by GIES Graduate Program (Excluding 'Other')", fontsize=16)
10 plt.xlabel("Date")
11 plt.ylabel("Engagements (Views)")
12 plt.legend(title="Program")
13 plt.xticks(rotation=45)
14 plt.tight_layout()
15
16 # Display the plot
17 plt.show()
```



```
1 merged_views_df
```

	Date	Program	Engagements	Total Engagements
0	2025-01-15	MSA	2856	108584
6	2025-01-16	MSA	2112	108584
12	2025-01-17	MSA	1847	108584
18	2025-01-18	MSA	1601	108584
24	2025-01-19	MSA	1535	108584
...	...	...	...	...
329	2025-03-10	Other	5470	344690
335	2025-03-11	Other	6133	344690
341	2025-03-12	Other	7380	344690
347	2025-03-13	Other	5223	344690
353	2025-03-14	Other	4337	344690

354 rows x 4 columns

```

1 # For each program, calculate total views by page title and select top 5 and bottom 5
2 summary_by_program = []
3
4 for program in df2['Program'].unique():
5     if program == 'Other':
6         continue
7     prog_df = df2[df2['Program'] == program]
8     views_summary = prog_df.groupby('Page title')['Views'].sum().reset_index()
9     views_summary = views_summary.sort_values(by='Views', ascending=False)
10
11     top_5 = views_summary.head(5).assign(Rank='Top 5', Program=program)
12     bottom_5 = views_summary.tail(5).assign(Rank='Bottom 5', Program=program)
13
14     summary_by_program.append(pd.concat([top_5, bottom_5]))
15
16 page_view_extremes = pd.concat(summary_by_program)[['Program', 'Rank', 'Page title', 'Views']]
17 page_view_extremes = page_view_extremes.sort_values(by=['Program', 'Rank', 'Views'], ascending=[True, True, False])
18
19 # Arrange so that Top 5 rows appear before Bottom 5 for each program
20 page_view_extremes['Rank'] = pd.Categorical(page_view_extremes['Rank'], categories=['Top 5', 'Bottom 5'], ordered=True)
21 page_view_extremes_sorted = page_view_extremes.sort_values(by=['Program', 'Rank', 'Views'], ascending=[True, True, False])
22
23 # Print again with Top 5 appearing above Bottom 5 for each program
24 for program in page_view_extremes_sorted['Program'].unique():
25     print(f"\n=== {program} ===")
26     print(page_view_extremes_sorted[page_view_extremes_sorted['Program'] == program][['Rank', 'Page title', 'Views']].to_string(index=False))

```



```

=== MSA ===
Rank      Page title  Views
Top 5    Accountancy | Gies College of Business Degree in Accounting 33958
Top 5    MSA Degree | Master's Degree in Accounting 20415
Top 5    Master of Accounting Science | Gies College of Business | MAS 16809
Top 5    Graduate Accounting Programs | Gies College of Business 16658
Top 5    MSA Admissions | Gies Master's Degree in Accountancy 5154
Bottom 5 Afgestudeerden Accountancy Programma's | Gies College of Business 0
Bottom 5 Gelar MSA | Gelar Magister Akuntansi 0
Bottom 5 Grau MSA | Mestrado em Contabilidade 0
Bottom 5 Grau de MSA | Mestrado em Contabilidade 0
Bottom 5 Penerimaan MSA | Gelar Magister Akuntansi Gies 0

=== MSBA ===
Rank      Page title  Views
Top 5    Business Analytics | Masters in Business Analytics | MSBA 23912
Top 5    Masters in Business Analytics Admissions | Gies College of Business 5284
Top 5    Master's of Business Analytics Curriculum | Gies College of Business 2628
Top 5    Master's of Business Analytics Tuition & Fees | MSBA 1437
Top 5    MSBA Career Outcomes | Gies Master's in Management 830
Bottom 5 Analyse commerciale | Master en analyse commerciale | MSBA 1
Bottom 5 Falañaynta Ganacsiga | Masters in Business Analytics | MSBA 0
Bottom 5 Business Analytics | Mestrado em Business Analytics | MSBA 0
Bottom 5 Ammissioni al Master in Business Analytics | Gies College of Business 0
Bottom 5 Mestrado em Business Analytics - Mensalidade e taxas | MSBA 0

=== MSF ===
Rank      Page title  Views
Top 5    Masters in Finance | Gies College MSF Program | Masters of Finance 21781
Top 5    Master's in Finance Admissions | Gies College of Business | MSF 6006
Top 5    MSF Curriculum | Gies College of Business | Master's in Finance 3290
Top 5    BS in Finance | Gies College of Business Bachelor's in Finance 2189
Top 5    MSF Tuition & Scholarships | Gies Master's in Finance 1741
Bottom 5 Pennacchi Invested as Bailey Memorial Chair of Money, Banking, and Finance 2
Bottom 5 DRAFT PhD in Finance - Placement 1
Bottom 5 Gies MBA fueling Nieves in shift from finance to tech 1
Bottom 5 From radio to finance: How one JMSA learner became a forensic accountant 1
Bottom 5 Master en finance | Programme MSF du Gies College | Master en finance 0

=== MSM ===
Rank      Page title  Views
Top 5    Masters in Management | Gies College of Business | MSM Degree 29208
Top 5    MSM Admissions | Gies College of Business | Master's in Management 4064
Top 5    Master's in Management Programs | Gies College of Business 3571
Top 5    MSM Curriculum | Gies College of Business | Master's in Management 2237
Top 5    BS in Management | Gies College of Business | UIUC 1221
Bottom 5 First Midwest Healthcare Management Conference a lifetime in the making 3
Bottom 5 Gies College of Business Establishes the AXIS Risk Management Academy at the University of Illinois 3
Bottom 5 New Master of Science in Management Degree Program at Illinois 3
Bottom 5 Supply Chain Management Conference 2024 - https://giesbusiness.illinois.edu/ 1
Bottom 5 Strategy Management, Entrepreneurship, and International Business 1

```