

Exploratory Data Analysis (EDA) report

Table 1

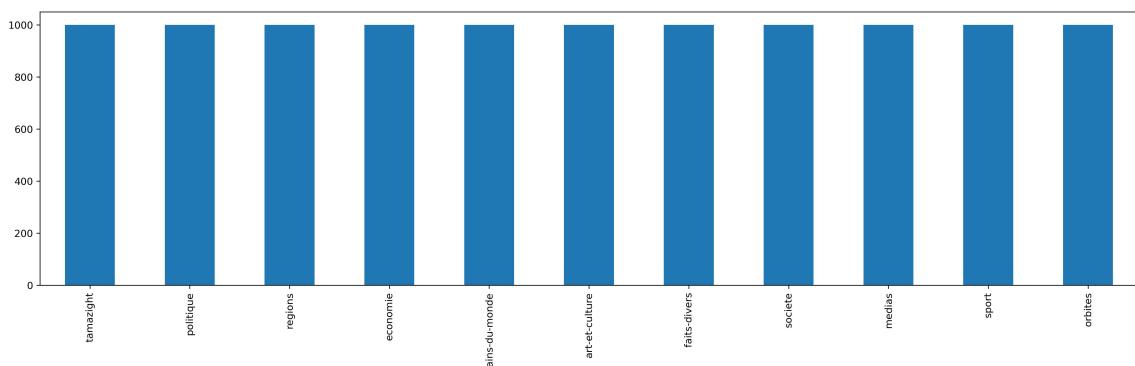
| | id | title | date | author | story | topic |
|---|--|--|---------------------------------------|---------------------|--|--------------|
| 0 | 9de52a46 055311eb 8949646e 69d991ea | وزارة التربية: لا تغير في الغلاف الزمني للأما... | السبت 03 أكتوبر - 2020 00:02 | هسبريس من الرباط | قالت وزارة التربية الوطنية والتكوين المهني وال... | tamazight |

For the Topic column there is 11 unique values:

[‘tamazight’, ‘politique’, ‘regions’, ‘economie’, ‘marocains-du-monde’, ‘art-et-culture’, ‘faits-divers’, ‘societe’, ‘medias’, ‘sport’, ‘orbites’], dtype=object)

Topics Frequencies

| | |
|--------------------|------|
| tamazight | 1000 |
| politique | 1000 |
| regions | 1000 |
| economie | 1000 |
| marocains-du-monde | 1000 |
| art-et-culture | 1000 |
| faits-divers | 1000 |
| societe | 1000 |
| medias | 1000 |
| sport | 1000 |
| orbites | 1000 |



visualize the frequent words in all of the stories using ArabicWordCloud

Word cloud for GENERAL



**visualize the frequent words of the stories for each topic
using ArabicWordCloud**

Title





faits-divers



society



medias

sport

orbites

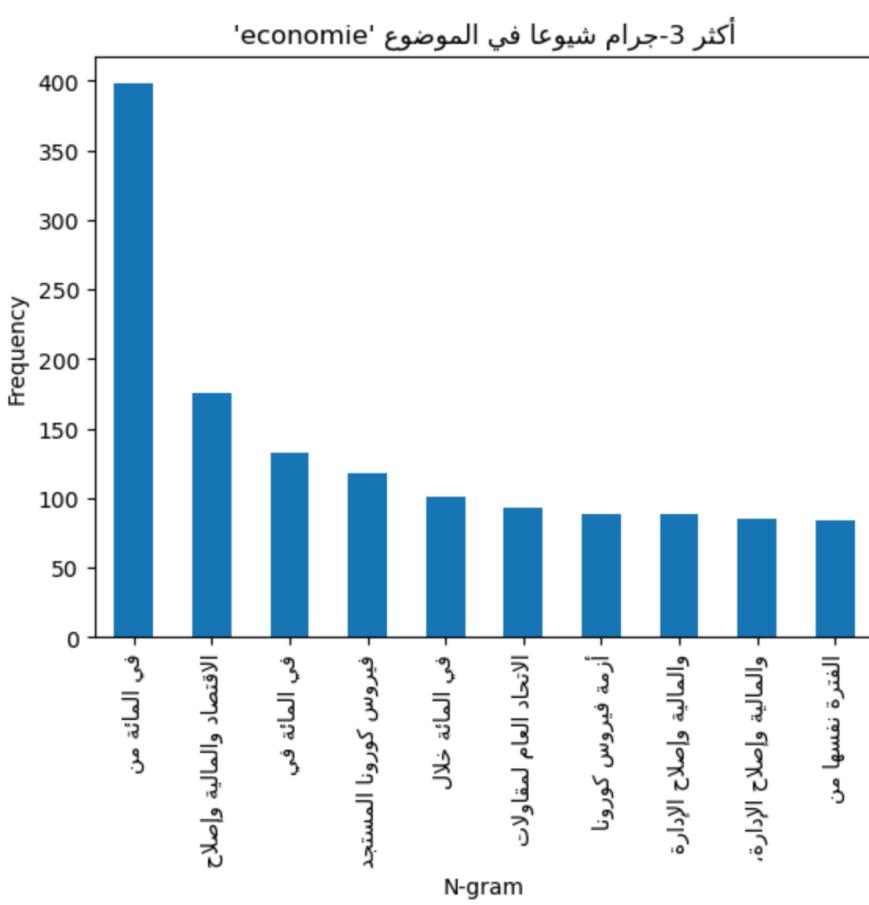
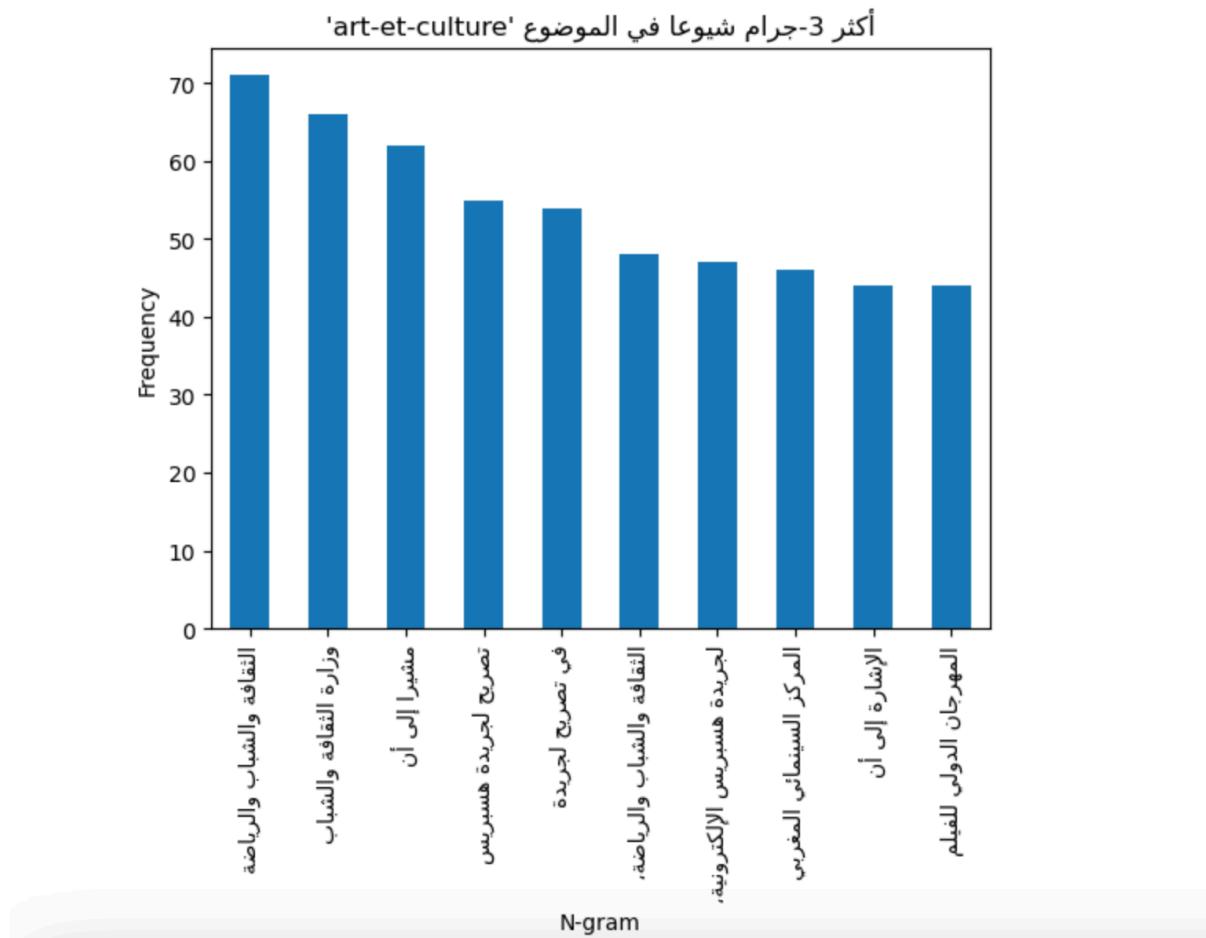
A dense collage of Arabic text from various news articles, forming a central image for the page.

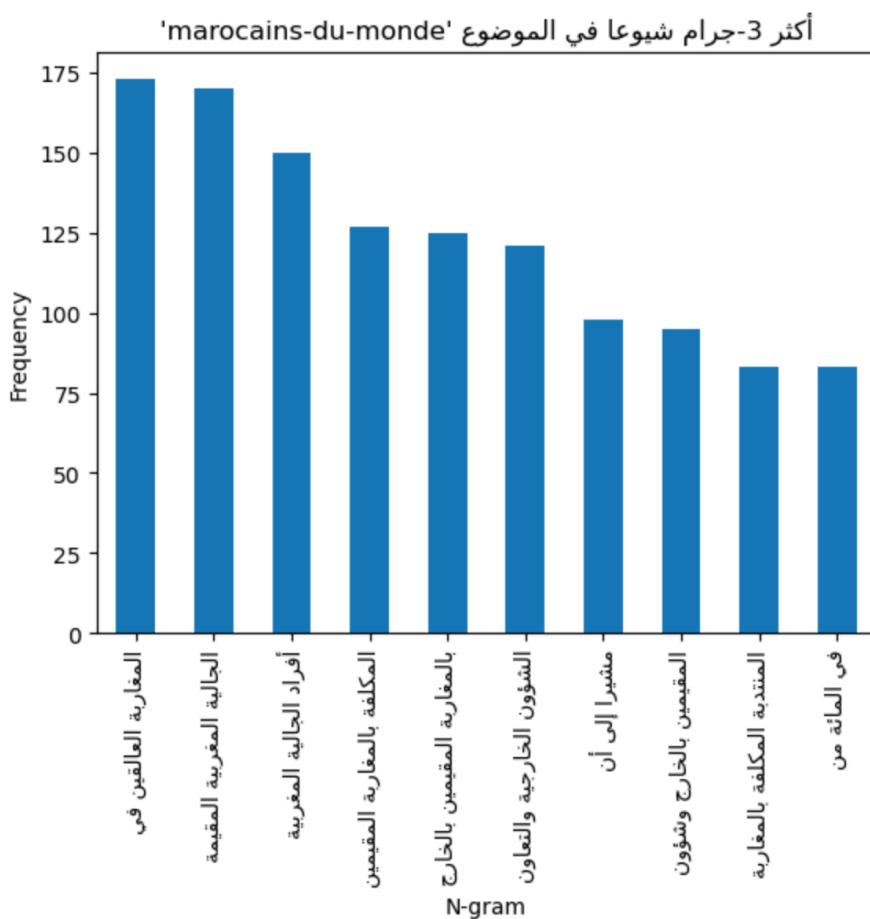
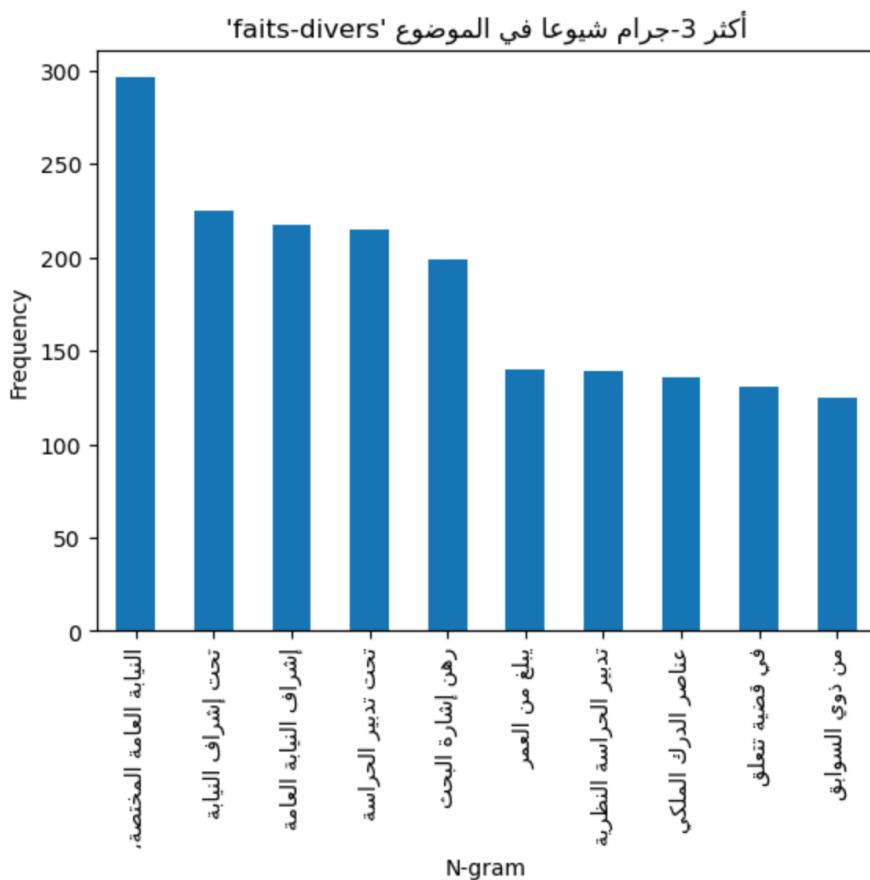


Most common 3-grams across all the stories:

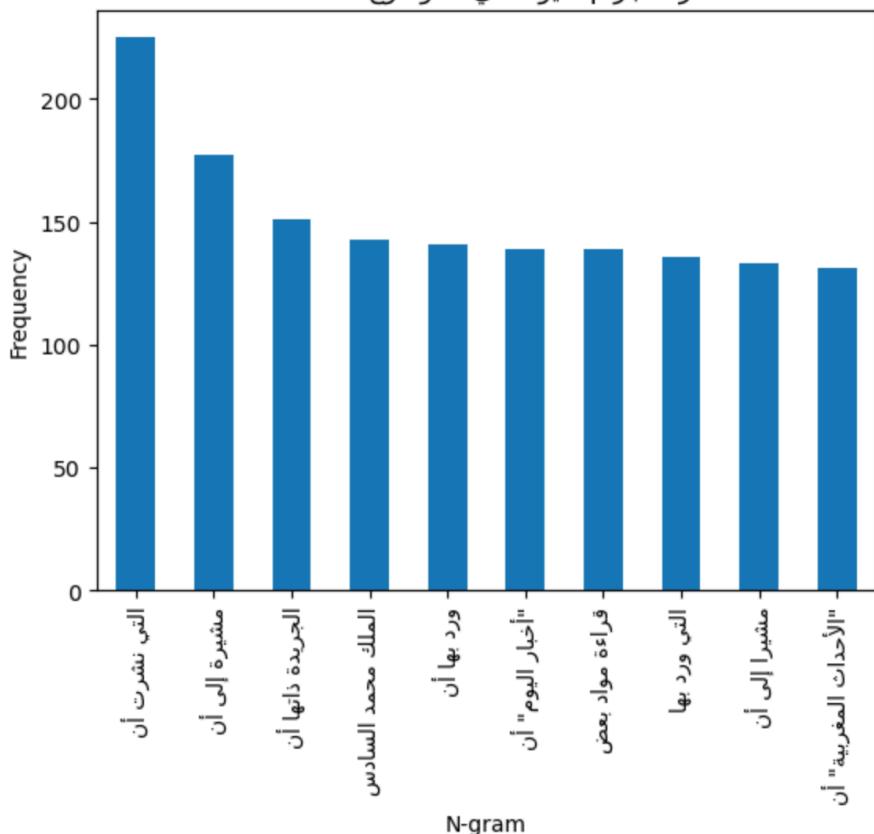
- الملك محمد السادس 741
تصريح لجريدة هسبيريس 638
المعهد الملكي للثقافة 401
فيروس كورونا المستجد 340
حالة الطوارئ الصحية 336
وزارة التربية الوطنية 332
الشؤون الخارجية والتعاون 324
حزب العدالة والتنمية 299
بفيروس كورونا المستجد 262
الاقتصاد والمالية وإصلاح 256

Most common 3-grams across each topic:

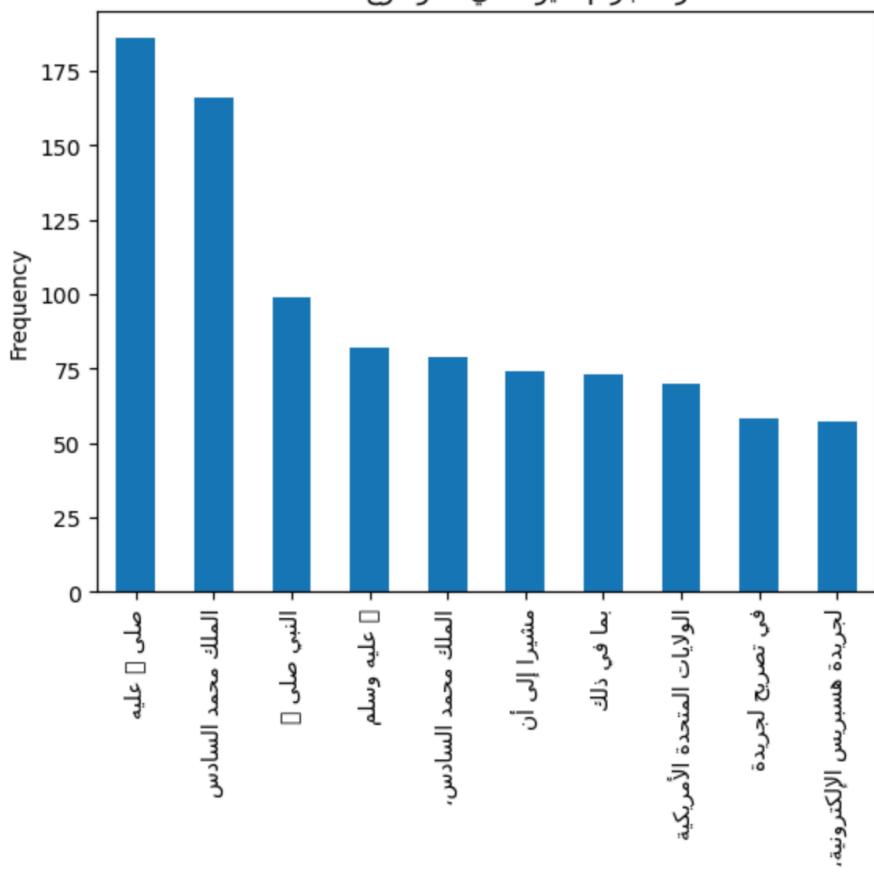




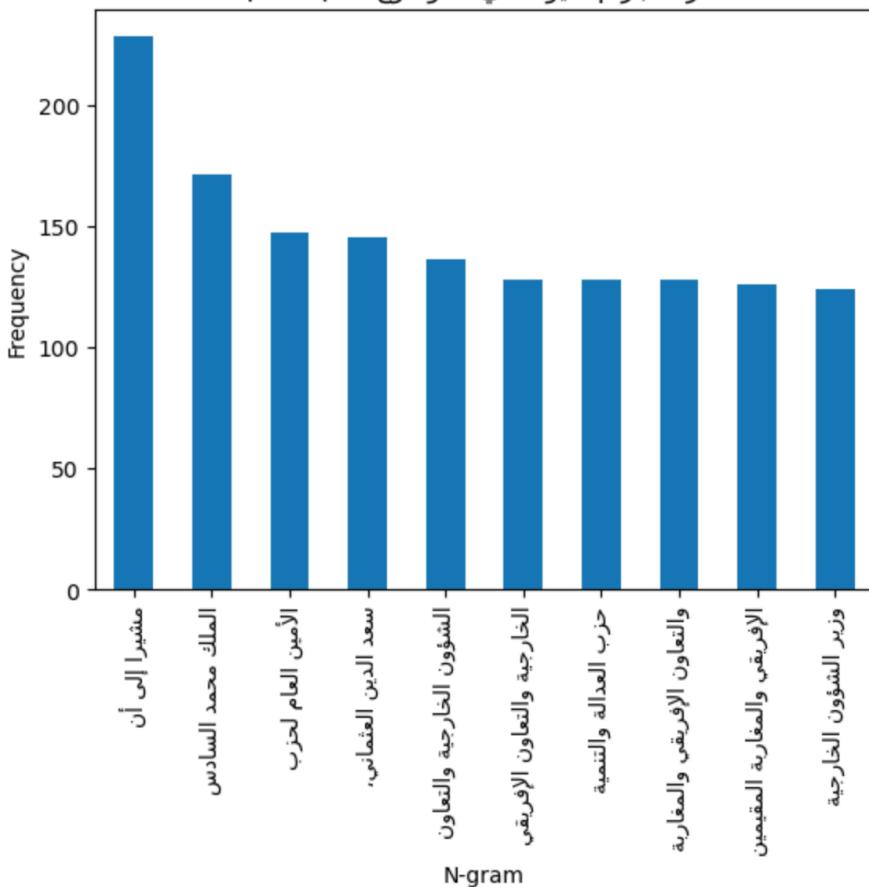
أكثر 3-جرام شيوعا في الموضوع 'medias'



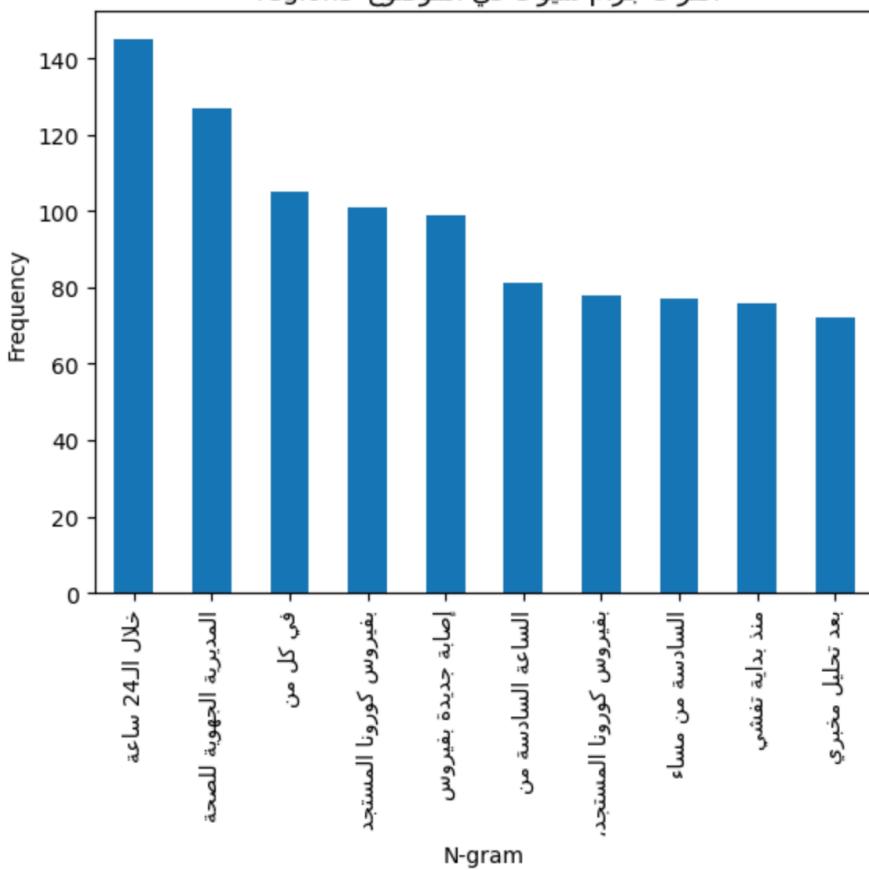
أكثر 3-جرام شيوعا في الموضوع 'orbites'



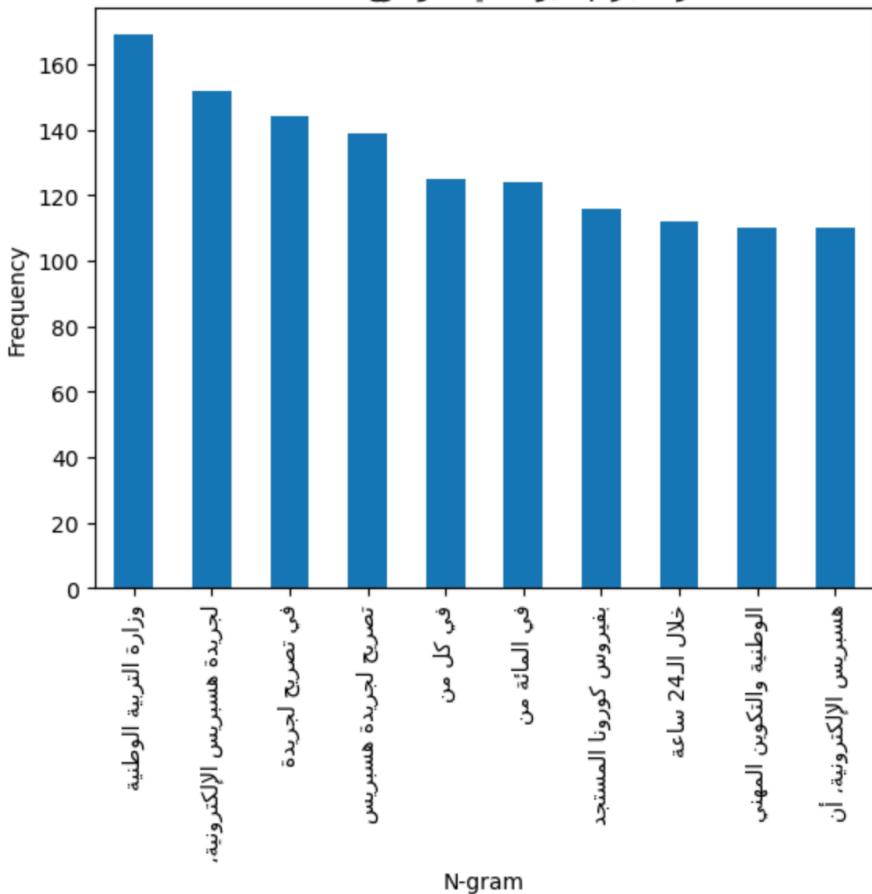
أكثر 3-جرام شيوعا في الموضوع 'polistique'



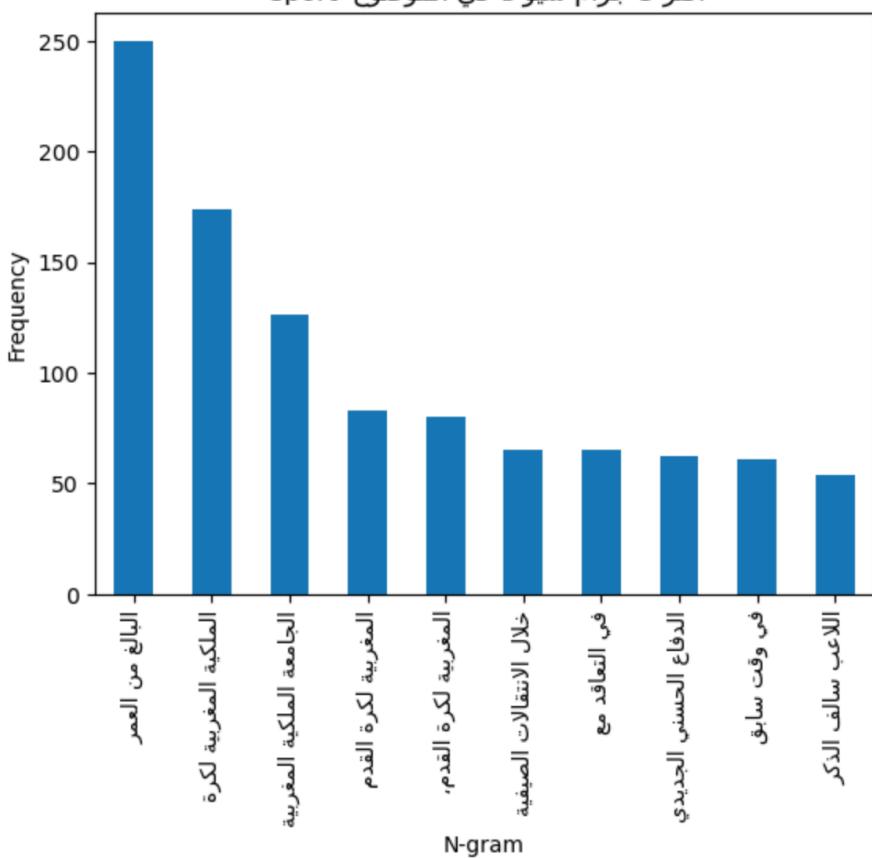
أكثر 3-جرام شيوعا في الموضوع 'regions'



أكثر 3-جرام شيوعا في الموضوع 'societe'



أكثر 3-جرام شيوعا في الموضوع 'sport'



أكتر 3-جرام شيوعا في الموضوع 'tamazight'

