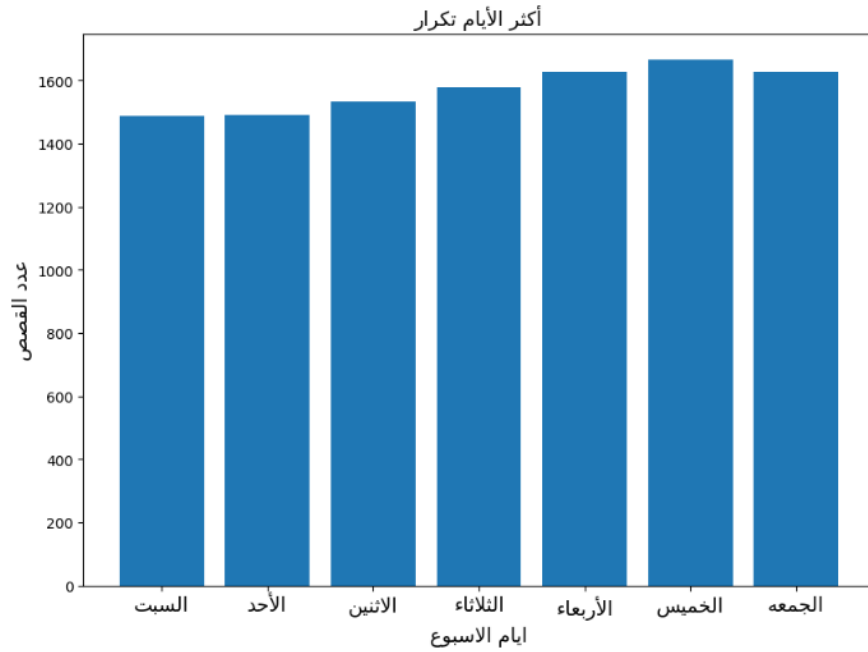


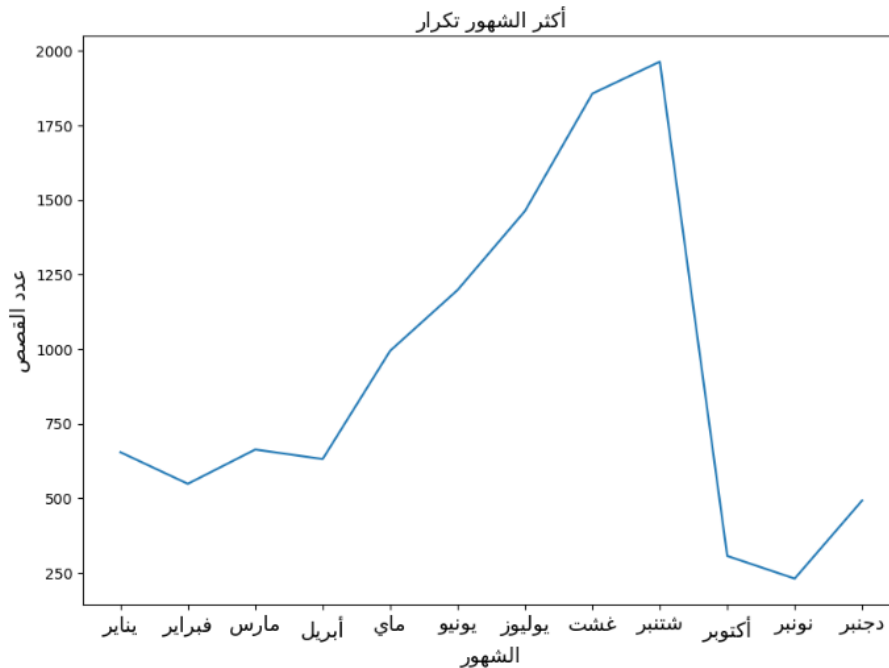
TASK 2

1) Number of stories that was published at the days of week :



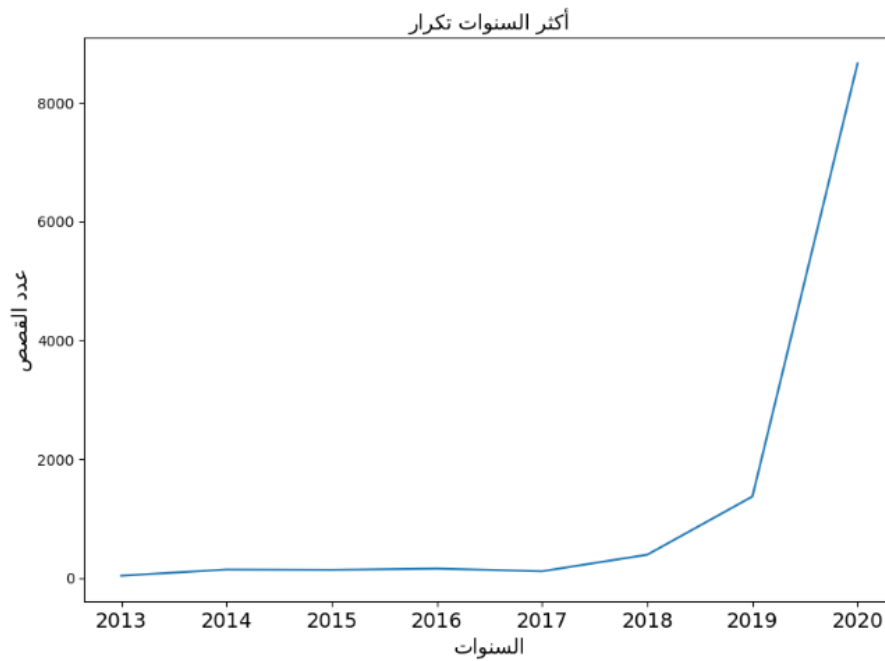
-> All days have the same number of stories approximately. The stop day is Thursday

2) Most common months:



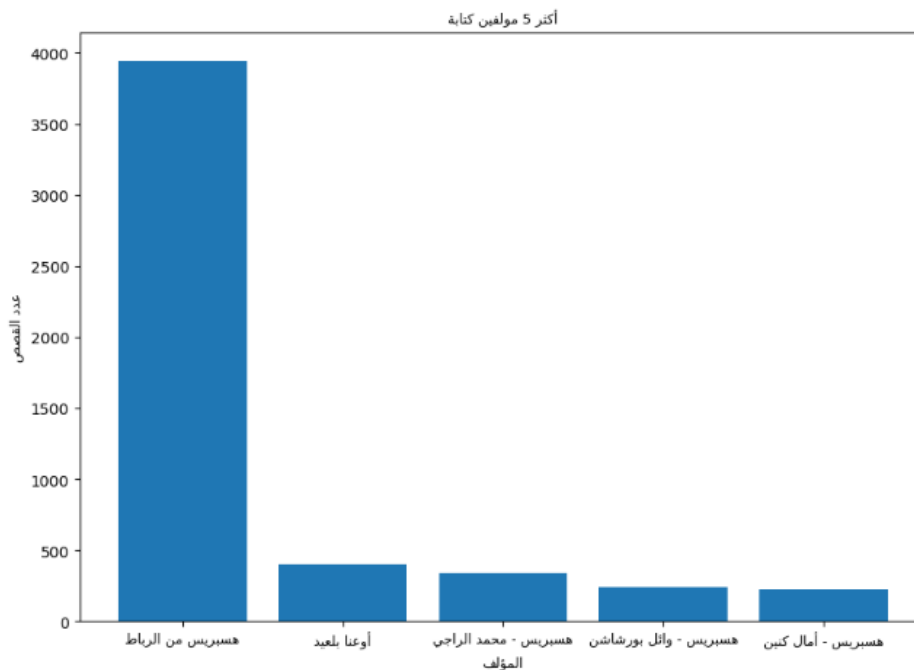
->there are relatively fewer stories at the beginning and end of the year compared to the middle. A noticeable increase in the number of stories occurs between April and September.

3) Most common year:



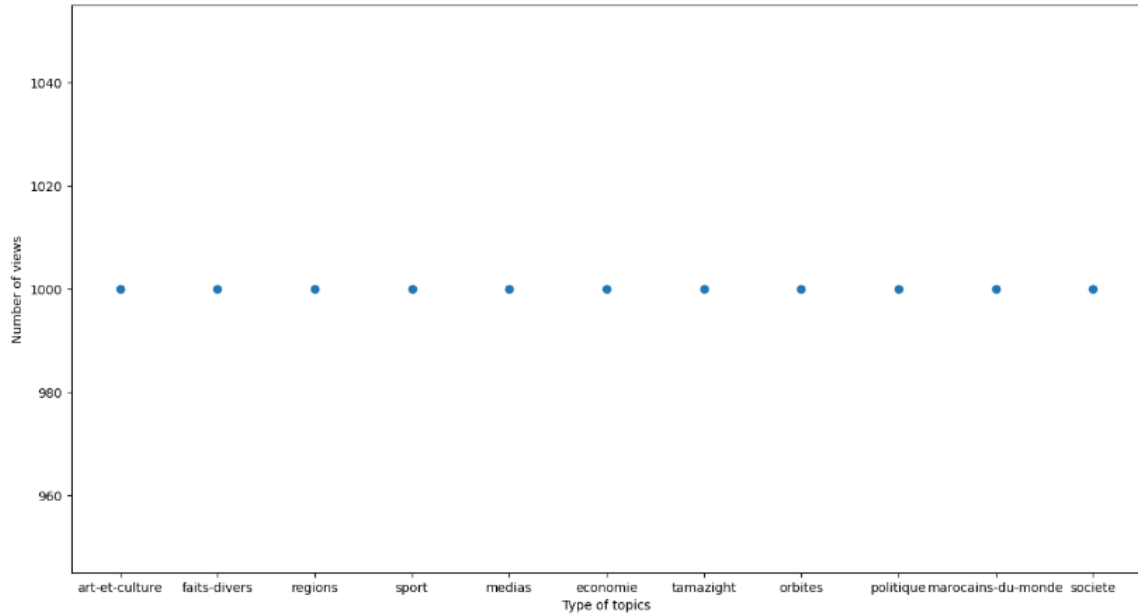
->Through this plot there is increasing in the number of stories starting from 2013. This trend remained constant from 2014 until the end of 2017. However, it return to increase from 2018 a very huge increasing.

4)Top 5 authors:



->There are the top 5 authors who published stories . The most author published is هسيريس من الرباط)

5)Stories with Topics:



->Numbers of stories per topic are the same (1000) with all topics.

6) Top frequent 5-grams generally and per topic

Top frequent 2-grams (General):

[('مجلس نوا', 'حجر الصبح', '1206'), ('ملك محمد', '1393'), ('محمد ساس', '1674'), ('كور مستجد', '1952'), ('دار بيضاء', '2281'), ('كوفيد', '19', '2321'), ('يروس كور', '2361'), ('نياب علم', '1131'), ('لكر قدم', '1148'), ('ب', '1177')]

Top frequent 2-grams for art-et-culture:

[('يروس كور', '116'), ('محمد ساس', '119'), ('زار ثقاف', '136'), ('ثقاف والشباب', '146'), ('دار بيضاء', '171')]

Top frequent 2-grams for faits-divers:

[('مدير علم', '256'), ('شرط قضاء', '375'), ('علم مختصة', '394'), ('درك ملك', '398'), ('نياب علم', '687')]

Top frequent 2-grams for regions:

[('حال اقليم', '357'), ('كوفيد', '19', '448'), ('دار بيضاء', '454'), ('فيروس كور', '476'), ('كور مستجد', '515')]

Top frequent 2-grams for sport:

[('بالغ عمر', '257'), ('كرة قدم', '292'), ('يروس كور', '318'), ('نول مغرب', '336'), ('لكر قدم', '1039')]

Top frequent 2-grams for medias:

[('رصيف صحاف', '272'), ('محمد ساس', '291'), ('دار بيضاء', '449'), ('احداث مغرب', '701'), ('اخبار يوم', '845')]

Top frequent 2-grams for economie:

[('دار بيضاء', '275'), ('مليار درهم', '280'), ('كوفيد', '19', '308'), ('كور مستجد', '314'), ('يروس كور', '423')]

Top frequent 2-grams for tamazight:

[('ملك ثقاف', '504'), ('ثقاف امازيغ', '540'), ('طابع رسم', '540'), ('قان تنظيم', '607'), ('الغ امازيغ', '702')]

Top frequent 2-grams for orbites:

[('ملك محمد', '323'), ('ولا متحد', '327'), ('يروس كور', '333'), ('محمد ساس', '346'), ('كوفيد', '19', '412')]

Top frequent 2-grams for politique:

[('رئيس حكوم', '302'), ('ملك محمد', '334'), ('محمد ساس', '349'), ('الام علم', '353'), ('مجلس نواب', '601')]

Top frequent 2-grams for marocains-du-monde:

[('مغرب مقيم', '462'), ('مغرب عالق', '465'), ('مغرب علم', '498'), ('مقيم خارج', '546'), ('حال مغرب', '604')]

Top frequent 2-grams for societe:

[('يروس كور', '271'), ('دار بيضاء', '287'), ('زار الصبح', '306'), ('ترب وطن', '328'), ('كوفيد', '19', '452')]