

Facebook post analysis

The goal is to analyze how the post type changed through time.

Step 0: Using Facebook Graph API

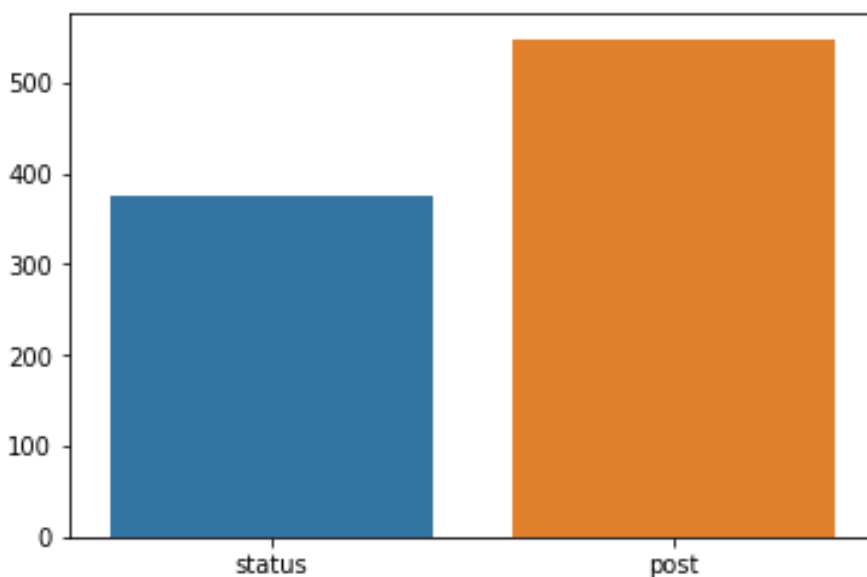
Use the Facebook API to get information about my posts. In this case, I use a file already downloaded through cURL from the Graph API explorer.

Number of status update: 374

Number of total posts: 548

Out[5]:

<matplotlib.axes._subplots.AxesSubplot at 0x7f8c2a95f908>



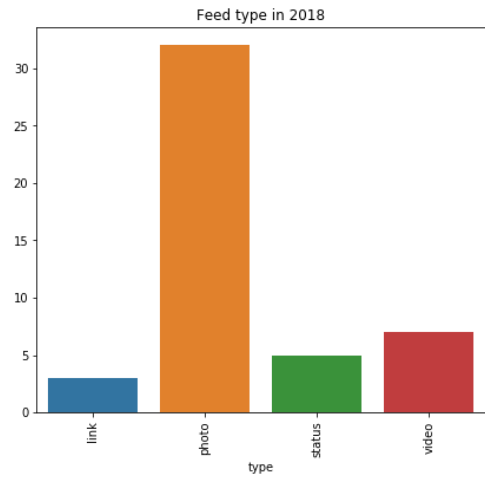
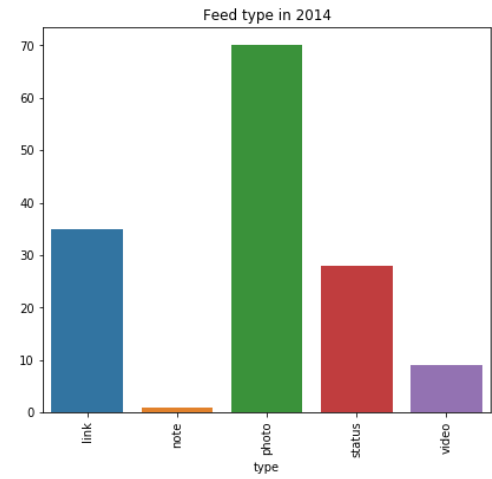
How things changed from 2014?

Let's plot the difference in post *type* and *status_type* for year 2014 and 2018. What we're interested in is to look, among the 500 posts loaded, how the distribution of *type* and *status_type* has changed:

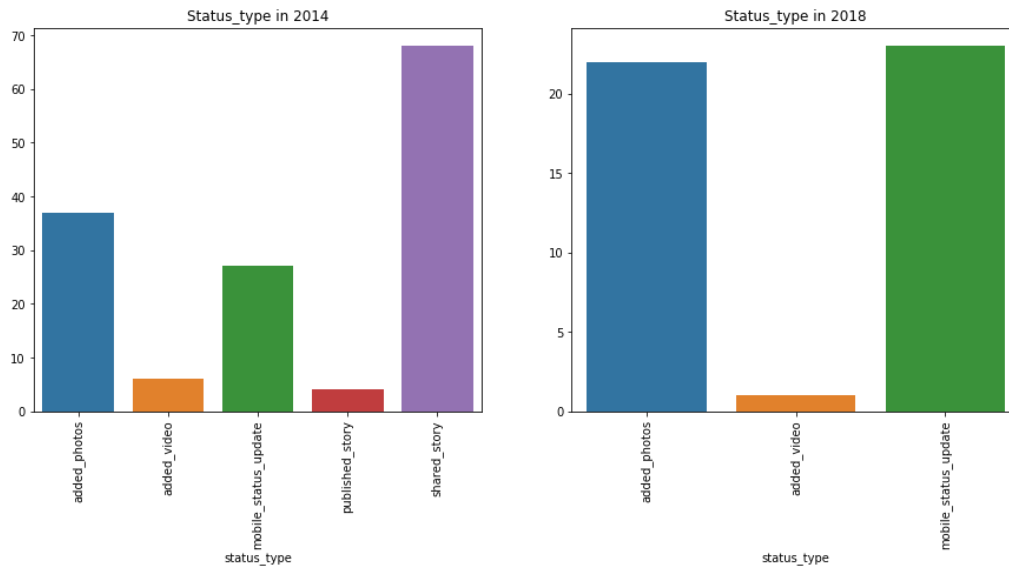
- did I post more images or status updates in 2014? What about now?
- did I do as many updates from mobile in 2014 as I do now?

For this analysis, I'll use all posts published during year 2014 and all those posted in 2018.

```
### DATA FROM 2014 ###
  type  id
0  link  35
1  note   1
2  photo 70
3  status 28
4  video  9
#####
### DATA FROM 2018 ###
  type  id
0  link   3
1  photo 32
2  status  5
3  video  7
#####
```



```
### DATA FROM 2014 ###
      status_type  id
0      added_photos 37
1      added_video   6
2  mobile_status_update 27
3      published_story   4
4      shared_story  68
#####
### DATA FROM 2018 ###
      status_type  id
0      added_photos 22
1      added_video   1
2  mobile_status_update 23
#####
```



Metadata Analysis

Not all information contained in Facebook feed data is actually shown on the platform itself.

Goal: show the metadata associated with the different posts in 2014 and 2018.

J'ai essayé de voir si il y a des differences entre les metadata des posts 2014 et 2018. Pour faire ça, c'est possible de afficher toutes les fields associé à un post, avec la requete:

```
"https://graph.facebook.com/v3.0/" + (post_id) + "?
metadata=1&access_token="+token
```

où *post_id* est l'ID du post. En faisant comme ça, c'est possible de voir que entre un post du 2014 et un du 2018 il n'y a pas differents metadata fields.

```
Types_2014: ['photo' 'link' 'video' 'status']
Types_2018: ['photo' 'link' 'video' 'status']
Status_types_2014: ['shared_story' 'published_story' nan 'added_photos'
'mobile_status_update' 'added_video']
Status_types_2018: ['added_photos' 'mobile_status_update' 'added_video' nan]
```

Different metadata: None

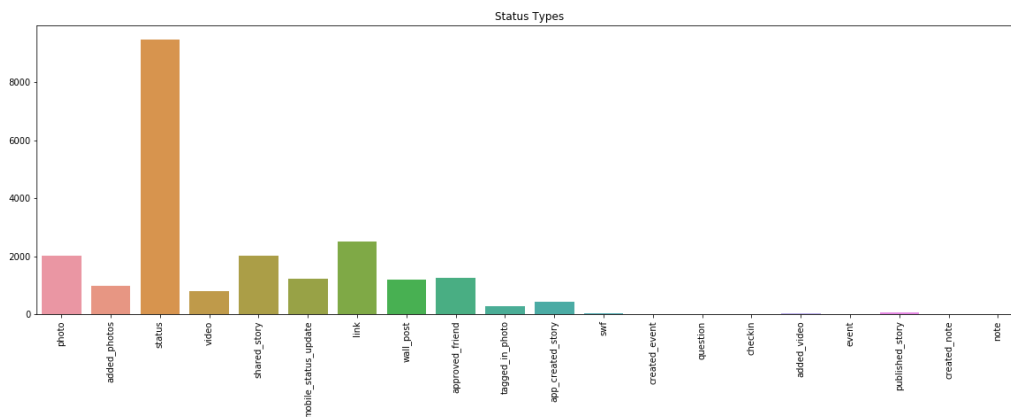
TP2

15k statuts Facebook, extraites de 15k profils Facebook différents, antérieurs à avril 2015 (date de basculement à la v2 de l'API Graph). Chaque ligne contient un post, en format json.

L'objectif du TP est de tenter de faire le lien entre métadonnées et catégories de posts, en lisant des posts.

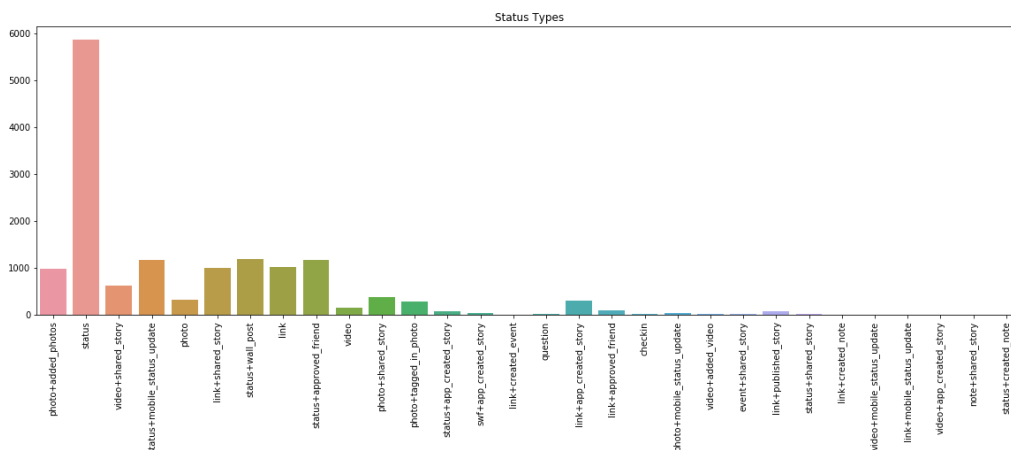
Question 1

Observer les valeurs possibles du champ type, produire un histogramme de toute les valeurs



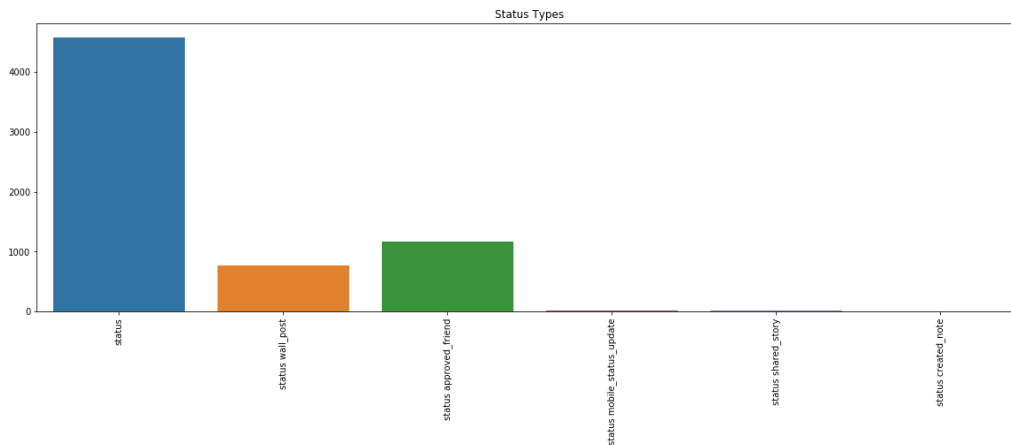
Question 2

Énumérer toutes les *combinaisons* de types de posts (puisque'un post peut avoir plusieurs types)



Question 3

Tenter d'énumérer les valeurs possibles du champ *type* associées à un post contenant uniquement du texte (un "status update"), donc pas un lien. Déterminer si ces valeurs sont suffisantes pour caractériser cette catégorie de posts.



Question 4

Vérifier que le type "approved_friend" ne caractérise pas tous les posts annonçant l'apparition de nouveaux amis.

Number of posts of new friends without 'approved_friend': 194

```
Example: [{ 'story_tags': [[{ 'type': 'user', 'id': '0017a9594e0e1a68c22c632550aaea69' }], [{ 'type': 'user', 'id': 'fe99a0c1421df2136808d6332e19bf05' }]], 'story': ' 0017a959 et fe99a0c1 sont désormais amis.', 'tags': [{ 'id': 'fe99a0c1421df2136808d6332e19bf05' }], 'id': 'f02ee68920dcd4add2af6a91174488d3a18f267a3856962a3a96daa2840707f3', 'from': { 'id': '0017a9594e0e1a68c22c632550aaea69' }, 'created': 1322501341000, 'updated': 1322501341000, 'type': ['status']}]
```

