Facebook

Sraping information on Facebook posts on specifics pages related to airline companies and plane.

|  |  |  |
| --- | --- | --- |
| **Variable name** | **Type** | **Description** |
| Data\_Source | string | Source of scraping |
| Date\_publication | date | Date of publication |
| Likes | int | Number of likes on publication |
| Description | string | Description of publication |
| Nb\_sharing | int | Number of share on publication |
| Nb\_comments | int | Number of comments on publication |
| Review | string | Comment in english |

Steps to scrap:

- For each pages, select informations publication thanks to soup and PhantomJS method.

- You select the description of the publication as well as its date and the number of shares, likes and comments. A comment is also taken.

- DataFrame will be exported in a json file.

In order to retrieve the comments, you have to go through Beautiful Soup and PhantomJS (scroll pages), so that you can to retrieve the page soup. In the soup, we have all tags and information pages. In order to retrieve the data we go through the htlm code of the page .

**Some statistics:**

395 rows × 72 columns (7 not NA), 751 Ko

