Airline Quality

Sraping information on Airline Quality comments related to airports and fly experience.

|  |  |  |
| --- | --- | --- |
| **Variable name** | **Type** | **Description** |
| Data\_Source | string | Source of scrapping |
| Date\_Review | date | Date of comment |
| Review | string | Comment in anglais |
| Title | string | Vidéo title |
| 'Date\_Flown' | date | Date of fly |
| Description | string | video description |
| 'Cleanliness' | int | Overall on terminal cleanliness |
| 'Food\_And\_Beverages' | int | Overall on Food and Beverages |
| 'Wifi\_And\_Connectivity' | int | Overall on airport connection |
| 'Cabin\_Staff\_Service' | int | Overall on airport staff |
| 'Overall\_Customer\_Rating' | int | Global overall |
| Recommended | string | Recommandation of user |
| Airport | string | Name of airport concerning the fly |
| 'Type\_Of\_Traveller' | string | Situation of user (couple, single) |
| 'Queuing\_Times' | int | Overall on Queuing Times in airport |
| 'Terminal\_Seating' | int | Overall on terminal seating |
| Terminal\_Signs | int | Overall on terminal signs |
| Airport\_Shopping | int | Overall on airport shopping |
| 'Experience\_At\_Airport' | string | Depart or arrival in airport |

Steps to scrap:

- For each airport, select comments with description and overalls (according to date).

- For each scraped comment, we will retrieve and save the different variables in a dataFrame.

- DataFrame will be exported in a json file.

In order to retrieve the comments, you have to go through Beautiful Soup, 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:**

82,5 Mo with 38002 rows X 63 columns (18 columns not NA)

"""Cleanliness

moy 3.139530630905212

med 3.0

min 1

max 5

ecart-type 1.375485603444257

Food\_And\_Beverages

moy 2.4437046004842613

med 2.0

min 1

max 5

ecart-type 1.347665124592081

Wifi\_And\_Connectivity

moy 2.5715202967824786

med 2.0

min 1

max 5

ecart-type 1.472260229069088

Cabin\_Staff\_Service

moy 2.278249406802004

med 2.0

min 1

max 5

ecart-type 1.4610990139330031

Queuing\_Times

moy 2.473302891933029

med 2.0

min 1

max 5

ecart-type 1.53215247994542

Terminal\_Seating

moy 2.4315887645380734

med 2.0

min 1

max 5

ecart-type 1.3953449991820368

Terminal\_Signs

moy 2.78709294269029

med 3.0

min 1

max 5

ecart-type 1.4267573846439106

Airport\_Shopping

moy 2.6824801412180053

med 3.0

min 1

max 5

ecart-type 1.363486640956753

"""