RUG_Information_Retrieval

Our application runs on **Python**.

Packages used

For this Project we will use Scholarly Package for Python https://github.com/scholarly-python-package/scholarly

Installation : By using pip3 to install from pypi

```
$ pip3 install scholarly
```

Also we will use GoogleSearch Package from SerpAPI https://serpapi.com/ To install:

```
$ pip install google-search-results
```

And finally for the GUI of our application we will use tkinter package https://docs.python.org/3/library/tkinter.html#a-very-quick-look-at-tcl-tk Installation:

```
$ pip install tk
```

How does the application work?

From the users perspective :

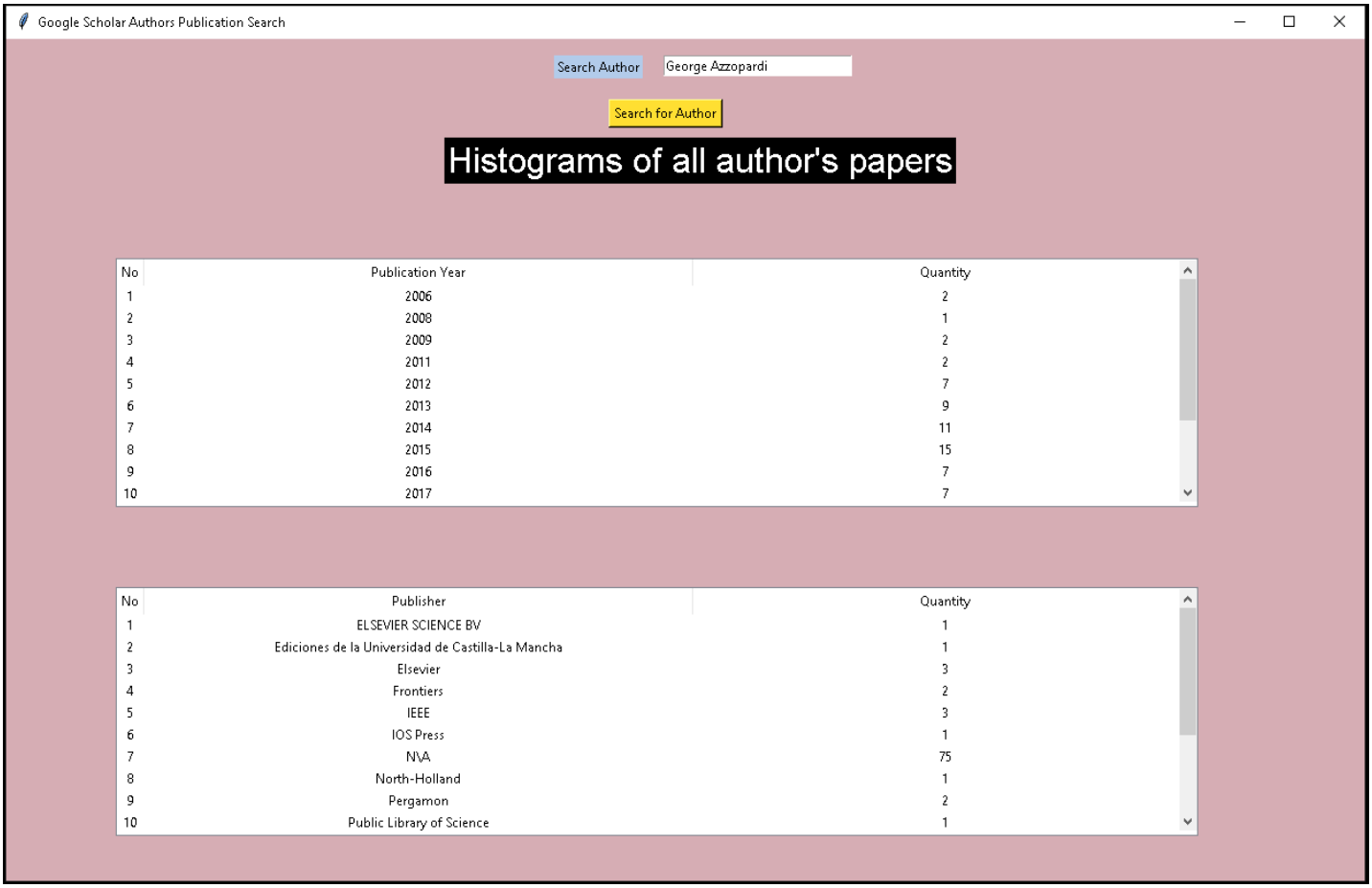
In the Graphical User Interface we have an input field, a menu with one or two choices and a button. The user should input the name of the Author he is interested in searching, then if available input an interval in the papers user is interested in viewing, for example [1,5] meaning user wants to view papers with index 1 and 5, or [1-5] meaning user wants to view papers with index 1,2,3,4 and 5. Afterwards user will press on the button to select his choice. After processing, the results will appear in the same window on user's screen.

General workflow of the application :

Our application is divided into three different files, each corresponds to the action the user wants to perform. The first action, described in the first part of the assignment is retrieving the histograms of the publication year's and publishers of the articles for the input author. In order to do this, the user must run the file **publications.py**. The second file **citing.py** is for the actions described in the second part of assignment, what it does it is creating three histograms for all the papers citing the selected papers of the author, the histograms contains the journal or the conference name where the article appears, the publication years of the articles, and the publisher of the articles citing the indicated article of the author that was input by the user. Finally, the third file **selfcit.py** calculate the self citations for the indicated article of the input author by the user, meaning that if the author of the article cited his article himself, it will be counted as a self citation.

Below some screenshots will be attached :

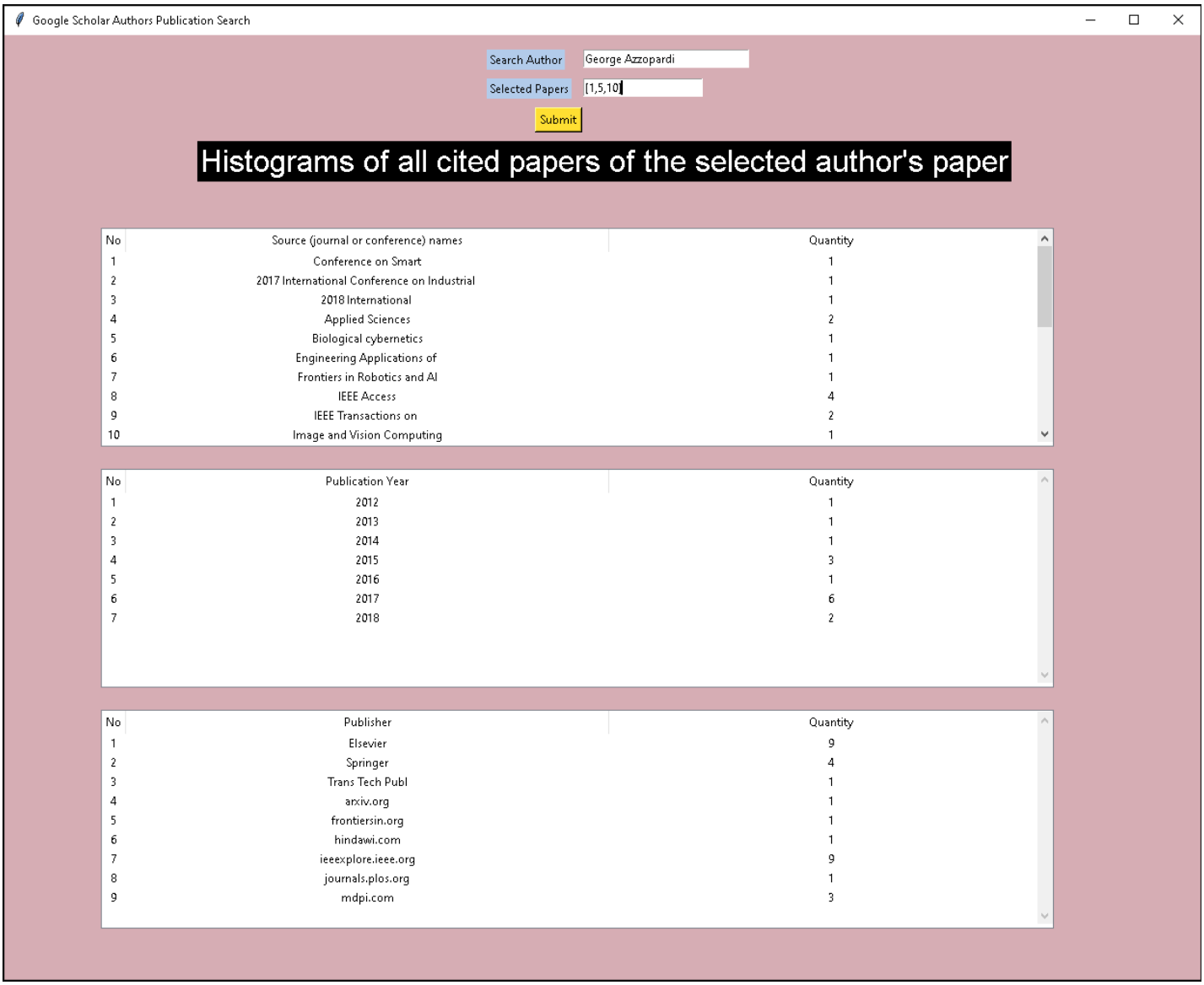
PART 1 (Histograms of publication year's and publishers of the author's articles)



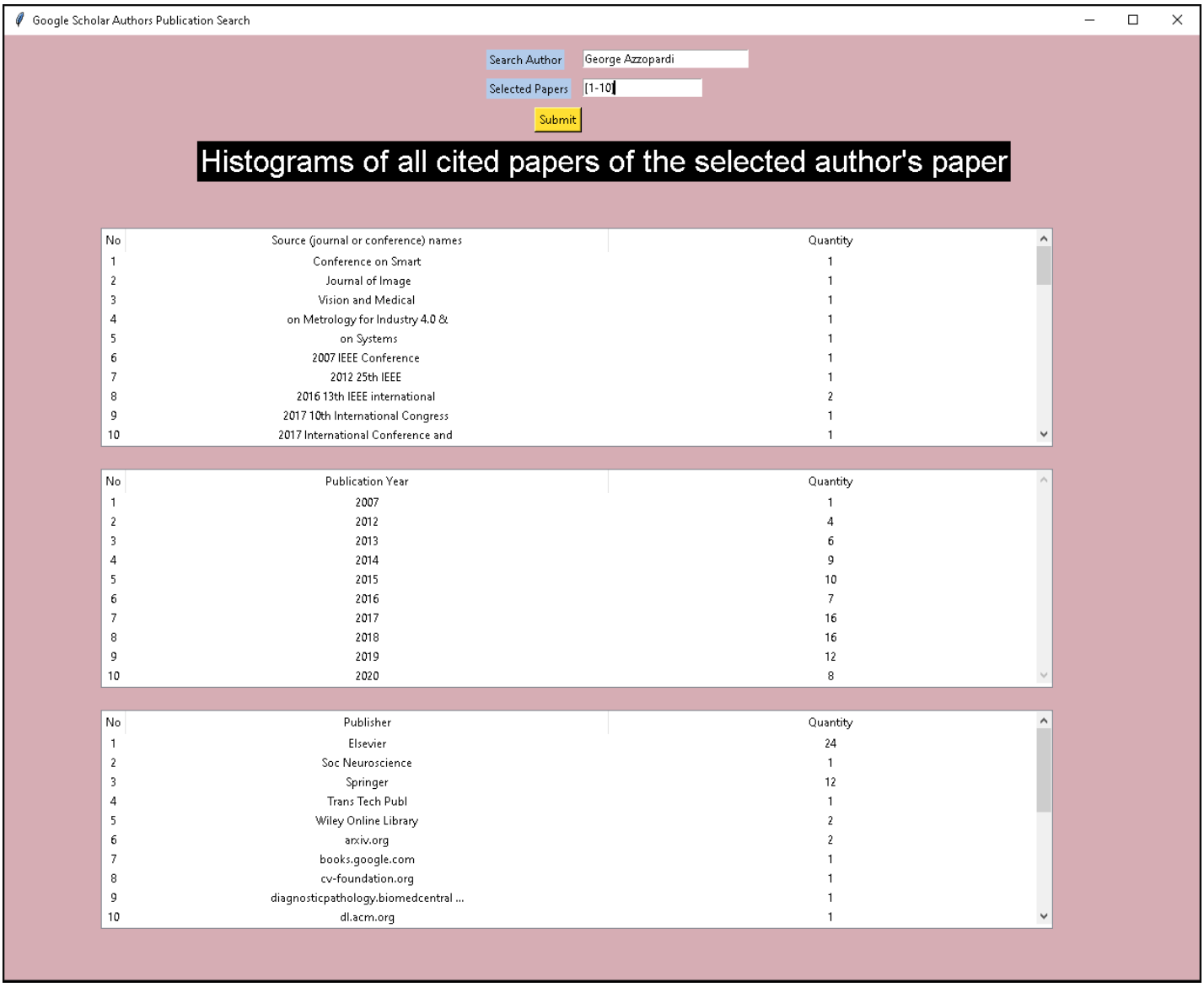
No	Publication Year	Quantity
1	2006	2
2	2008	1
3	2009	2
4	2011	2
5	2012	7
6	2013	9
7	2014	11
8	2015	15
9	2016	7
10	2017	7

No	Publisher	Quantity
1	ELSEVIER SCIENCE BV	1
2	Ediciones de la Universidad de Castilla-La Mancha	1
3	Elsevier	3
4	Frontiers	2
5	IEEE	3
6	IOS Press	1
7	N\A	75
8	North-Holland	1
9	Pergamon	2
10	Public Library of Science	1

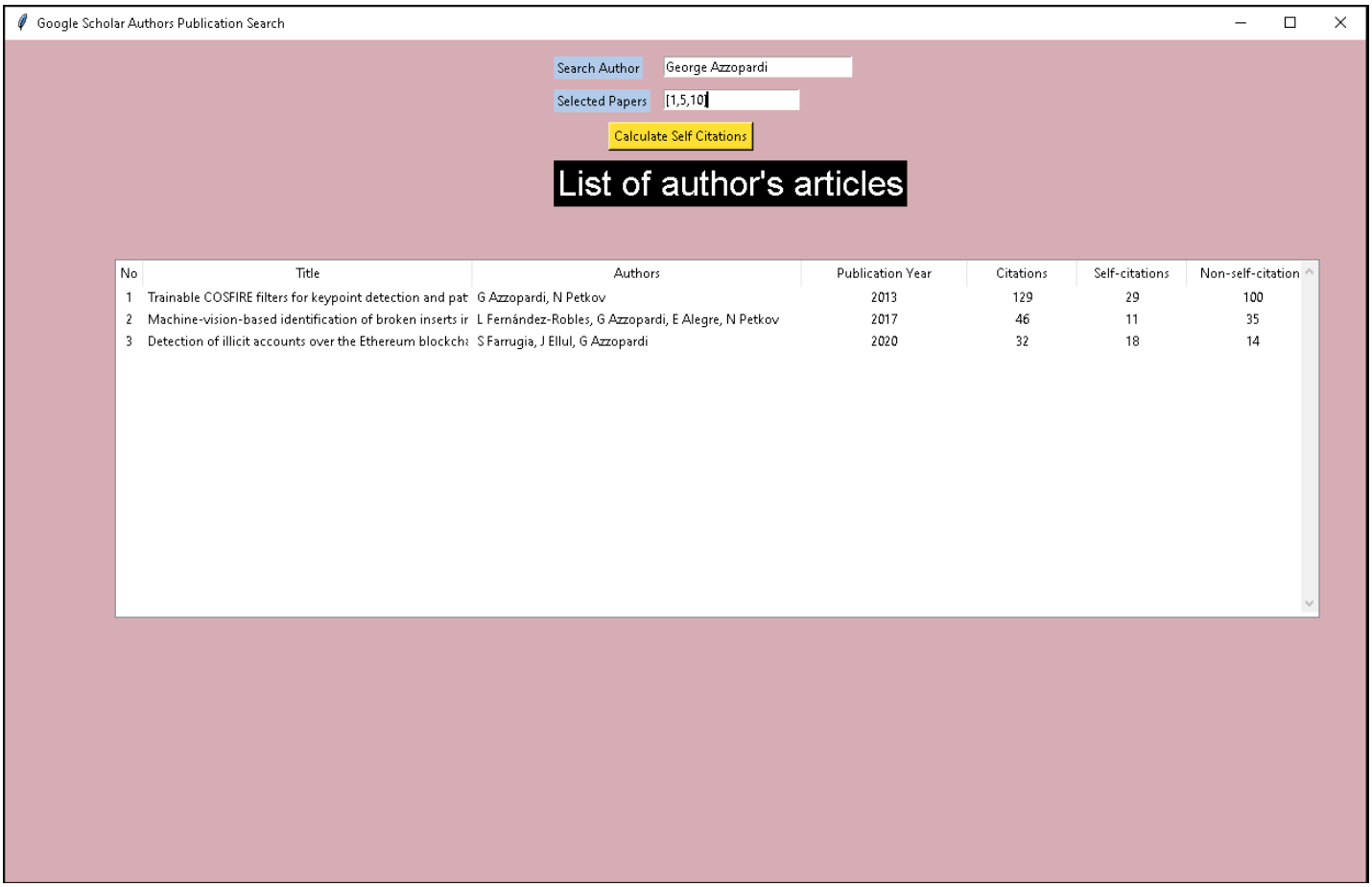
PART 2 (Histograms for the citing papers of the indicated paper of the author) Indicated only the index numbers of the author's articles :



Indicating an interval of indexes of the author's articles :



PART 3 (Self-citations of the author's articles) :



And now by interval :