

# Content Summary Generator

According to our research and background study the main algorithms used for Abstraction Summary are **Frequency Distribution** and **Page Rank**. In Below we will describe the process of Frequency Distribution and Page Rank.

## Content:

Cancer is a DNA (Deoxyribonucleic acid) disease which is due to the process of immortality of one cell which gives birth to cancerous cells. The following colony does not match to the natural rules of cells (1). Cancer is becoming a worldwide problem and a major public health (2). In 2012, 14.1 millions of new cases have been registered around the world with 8.2 millions of death. The incidence of cancer is increasing, consequently the death rate would increase more than 50 % and to be 15 millions by the year 2030 according to the World Health Organization (WHO) (2,3). Epidemiologic studies in Africa estimate about 1.2 millions of new cancer cases by 2030 with more than 970000 deaths if there are no adequate and rapid measures taken. In Ivory Coast, during 1984 to 2009, 12841 cancer cases were registered, with 493 cases in a year (4). In Niger, from 1992 to 2009, 7031 cancer cases have been registered within a significant increase of 186 cases in 1992 to 646 cases in 2009 (5). Concerning the situation in Madagascar, a study has been done between the years 2009 to 2010 at the oncology service of Joseph Ravoahangy Andrianavalona hospital (CHU/JRA) and revealed 1481 new cancer cases, an annual frequency of 740 cases (6). Some cancer cases show a frequency decreased around the world, sex and age mixed up cancer of : lung, breast, colorectum, liver, bladder, esophagus, stomach, prostate and the cervix (7). In Africa, breast cancer and cervix cancer are the most frequent (5). The same case is noticed in Madagascar (8,9).

## Frequency Distribution Table:

Sentences	A	B	C	D	E	F	G	H	I	J	SUM
A	2	1	1	0	1	1	1	1	1	1	10
B	1	1	0	0	0	0	0	0	0	0	2
C	1	0	2	0	2	1	1	1	1	1	10
D	0	0	0	7	3	4	2	2	3	1	22
E	1	0	2	3	7	3	2	2	2	1	23
F	1	0	1	3	3	6	2	3	2	3	27
G	1	0	1	2	2	2	4	2	2	2	18
H	1	0	1	2	2	3	2	6	2	3	23
I	1	0	1	3	2	2	2	3	7	4	25
J	1	0	1	1	1	3	2	3	4	6	22

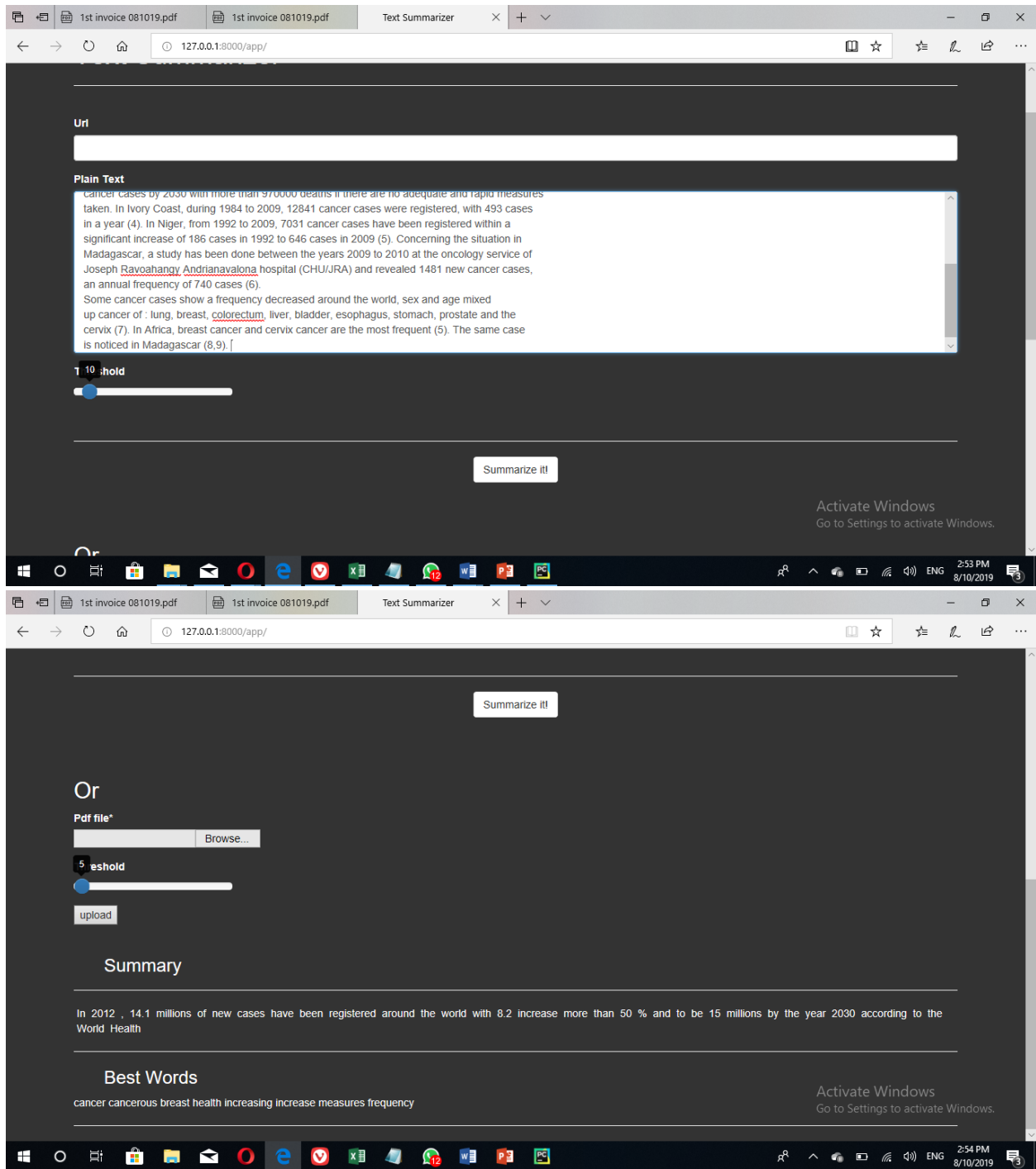
In the above case the frequency distribution algorithm will pick sentence F because it contains highest frequency.

## Page Rank Table:

Sentences	A	B	C	D	E	F	G	H	I	J	SUM	
A	2	1	1	0	1	1	1	1	1	1	10	0.81
B	1	1	0	0	0	0	0	0	0	0	2	0.79
C	1	0	2	0	2	1	1	1	1	1	10	0.83
D	0	0	0	7	3	4	2	2	3	1	22	0.98
E	1	0	2	3	7	3	2	2	2	1	23	0.963
F	1	0	1	3	3	6	2	3	2	3	27	0.824
G	1	0	1	2	2	2	4	2	2	2	18	0.94
H	1	0	1	2	2	3	2	6	2	3	23	0.92
I	1	0	1	3	2	2	2	3	7	4	25	0.93
J	1	0	1	1	1	3	2	3	4	6	22	0.96

In the above case the Page Rank algorithm will pick sentence D because it contains Highest Value.

All these values  
will be divided  
by the total sum  
of their  
frequency



This is the least summarized content by page rank.