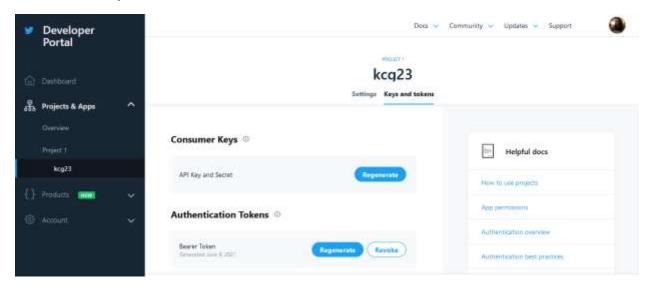
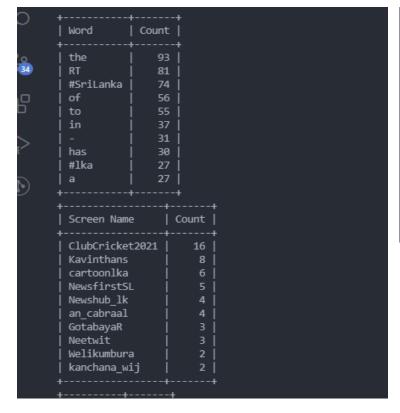
## Lab sheet 6 - Data Extraction

# Kavishka Gamage - 17000475

## Twitter Developer account – Access for Twitter API



Task 1 – Tweets extraction - Prettytable



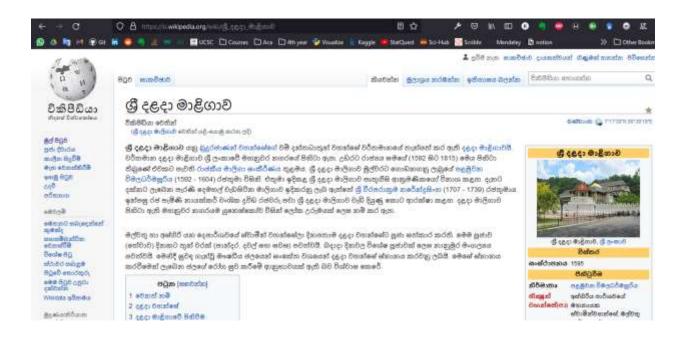
+   Hashtag +	+   Count   +
SriLanka	79
1ka	28
India	16
IndvsSL	16
LKA	12
SLnews	10
COVID19	10
srilanka	4
News	4
Trending	4
+	+ <b>+</b>

Count	Screen Name	Text
55     1	Welikumbura   	RT @Welikumbura: Quite interesting, isn't it?
		We have two different sets of laws, one set for the rich & amp; powerful & amp; one set for the ordina
54   	Kavinthans     	RT @Kavinthans: A nurse who was seriously injured   after attacked by a monk, has been transferred to   Kurunegala hospital. Nurse has conducte
54	Kavinthans	RT @Kavinthans: A nurse who was seriously injured   after attacked by a monk, has been transferred to   Kurunegala hospital. Nurse has conducte
54	Kavinthans	RT @Kavinthans: A nurse who was seriously injured   after attacked by a monk, has been transferred to   Kurunegala hospital. Nurse has conducte
54	Kavinthans	RT @Kavinthans: A nurse who was seriously injured   after attacked by a monk, has been transferred to   Kurunegala hospital. Nurse has conducte

# Trending Hashtags – Pretty table screenshot

+   Hashtag +	+   Count	
lka	55	
SriLanka	32	
JusticeforHishalini	13	
LKA	12	
srilanka	11	
COVID19SL	10	
BlackJuly	6	
TheDivideLK	6	
SLvIND	6	
COVID19LK	5	
<b>+</b>		

### Task 2 – Scrap Sinhala Wikipedia page



#### Link.txt screenshot

```
extractText.py U
pasic_scrape.py U
wikipedia_scrape.py U
links.txt U X
corape_book.py U

1
ග්‍රී දළදා මාළිගාව->විශේෂ:ගවේෂණය/ග්‍රී දළදා මාළිගාව

2
ග්‍රී දළදා මාළිගාව->මළගා

4
ග්‍රී දළදා මාළිගාව->සෙංකඩගල රාජකීය පරිගුය

5
ග්‍රී දළදා මාළිගාව->ශ්‍රී වීර පරාකුම නරේන්දුසිංහ රජ

6
ග්‍රී දළදා මාළිගාව->සෙංකඩගල රාජකීය පරිගුය

8
ග්‍රී දළදා මාළිගාව->සෙංකඩගල රාජකීය පරිගුය

9
ග්‍රී දළදා මාළිගාව->වැලිවිට සරණංකර හිමි
```