

Wikimedia: Wikipedia Scrapping

Video Link : <https://www.youtube.com/watch?v=b6puvm-QEY0>

Notebook:

<https://colab.research.google.com/drive/1UZky5JdOn2oMYIkls23WefTaT8VinYyg>

Libraries Used:

- wikipedia : Wikipedia is a Python library that makes it easy to access and parse data from Wikipedia.
-

Some important syntax:

To do a Wikipedia search for query

```
wikipedia.search(_query_, _results=10_,  
_suggestion=False_)
```

To generate a plain text summary of the page

```
wikipedia.summary(_query_, _sentences=0_, _chars=0_,  
_auto_suggest=True_, _redirect=True_)
```

sentences - if set, return the first sentences sentences
(can be no greater than 10).

chars - if set, return only the first chars characters

To get a WikipediaPage object for the page

```
wikipedia.page("query")
```

To get the full content

```
print(wikipedia.page("query").content)
```

To get all the URLs of the page

```
print(wikipedia.page("query").url
```

To get all the images

```
print(wikipedia.page("query").images
```