

# Tutorial for Install and Run the Client

## Installation and setting up the application

1. Install python 3.6
2. Install an IDE such as (Pycharm) or a code editor (Atom, Sublime)
3. Install pip and dependencies
  - a. Install pip using these instructions: <https://pip.pypa.io/en/stable/installing/>
  - b. Install urllib3 using this command: `pip install urllib3`  
(<https://pypi.org/project/urllib3/>)

## How to run the application

There are two ways to run the client:

1. The first using a console application
  - a. Open a console or cmd
  - b. Run the command `python`:
  - c. `python application_ebi_technical_test.py`

```
(alkanzaprojections) [~/Desktop]
[ ] python3 application_ebi_technical_test.py
Choose a menu
1 Get total number of samples
2 Get Name for an given Accession
3 Get Filtered Accessions for an given Attribute and Value
4 Examples of such queries and results
5 Exit
Option: 1
Total number of samples: 5795298

Choose a menu
1 Get total number of samples
2 Get Name for an given Accession
3 Get Filtered Accessions for an given Attribute and Value
4 Examples of such queries and results
5 Exit
```

2. The second option is using each functionality calling each function.
  - a. Open a console or cmd in the root folder (The client folder) `api_ebi_client/src`

```
(alkanzaprojections) [~/Desktop/ebi_client]
[ ] pwd
/Users/okgarces/Desktop/ebi_client
(alkanzaprojections) [~/Desktop/ebi_client]
[ ] _
```

- b. Open a python console using the command:
  - i. python
- c. Import the file and library:
  - i. from client\_api\_ebi import \*
- d. Call a function:
  - i. To get total number of samples
    - 1. Api\_client.total\_number\_of\_samples()
  - ii. To get the name by an accession.
    - 1. Api\_client.get\_name\_by\_accession() and pass as a param the accession
  - iii. To get the accessions using a attribute:value filter
    - 1. Apli\_client.get\_accessions\_by\_filter() and pass two params.  
The first is the attribute and the second is the value

```
(alkanzaprojections) [~/Desktop/ebi_client]
[ ]> pwd
/Users/okgarces/Desktop/ebi_client
(alkanzaprojections) [~/Desktop/ebi_client]
[ ]> python
Python 3.6.4 (default, Mar 1 2018, 18:36:50)
[GCC 4.2.1 Compatible Apple LLVM 9.0.0 (clang-900.0.39.2)] on darwin
Type "help", "copyright", "credits" or "license" for more information.
[>>> from client_api_ebi import *
[>>> Api_client.total_number_of_samples()
5795298
[>>> Api_client.get_name_by_accession('SAMEA1507265')
'source GSM1000387 1'
[>>> Api_client.get_accessions_by_filter('organism','homo sapiens')
'SAMN08025863, SAMN08025864, '
>>>
```

## Bonus questions

### 1. Do you have any suggestions for improvements to the BioSamples API or documentation?

R/ The documentation is complete enough and it is well explained for engineers or people that has experience using APIs, however there are not examples for other kind of users, the errors codes.

Related to API I would create more Error code for applications, I would add documentation using swagger and collections to use the APIs like postman collection.

Actually, the filters are case sensitive. This kind of filters can difficult the interaction with the API and data can be filtered using multiple filters.

### 2. Searching for “lung cancer” using the top right search box at

<https://www.ebi.ac.uk/biosamples/> returns

<https://www.ebi.ac.uk/biosamples/samples/SAMEA1481290>, which does not contain the string “lung cancer”. Can you explain why?

R/ When I tried to use the filter by text using a REST client and the URL

<https://www.ebi.ac.uk/biosamples/samples?text=lung+cancer>. The results do not include the SAMEA1481290.

The first sample was SAMEA1507264, and the results are ordered Ascending.