

- I) Configuring collector.config
- II) Starting Collector.py
- III) SQL

I) collector.config must be configured prior to starting Collector.py. The configuration file must contain the following:

1. [db\_info] This section will contain database info. The [db\_info] section will contain the following:
  1. host = MySql host
  2. db\_user = MySql username
  3. db\_pass = MySql pass

This is a template of the [db\_info] section

```
#this is a comment
[db_info]
host          = localhost
db_user       = username
db_pass       = password
```

2. At least one user section must be defined. The user section must be enclosed in two square brackets. The user section will contain the following:

1. con\_key = consumer key
2. con\_secret = consumer secret
3. key = user key
4. secret = user secret
5. db = database to store tweets
6. include\_path = None if empty, path to twitter ids otherwise

This is a template of the user section

```
#This is a comment.
#you can add as many comments as you want
#you can define more than one user. A new thread will be created for each user.
[user_name_goes_here]
con_key      = get con_key from twitter
con_secret   = get con_secret from twitter
key          = get user key from twitter
secret       = get user secret from twitter
db           = database to store tweets(Note: the user defined under [db_info] must
              have access to this db)
include_path = path to a file containing twitter ids. If this path is set, then no ids will be
              retrieved from user account
```

## II) Starting Collector.py

Starting Collector.py will load the config file and start collection tweets. A log file will be created in the same directory. This log file will contain info, warnings and errors. Check this file when debugging the collector. The default config file is 'collector.config'. This file must be stored in the same directory as Collector.py or set at run time using:

“Collector.py -c path\_to\_config\_file”

Run the collector in the background by adding '&' at the end of the command.

- III) A schema image and .sql file can be found in the github directory. Please refer to the image file for more information.