Introduction:

This utility will extract the information of different comics available in <https://xkcd.com/>. It is a fully configuration driven utility where use can specify all the property details in a config.json file as a result users are not required to modify any code.

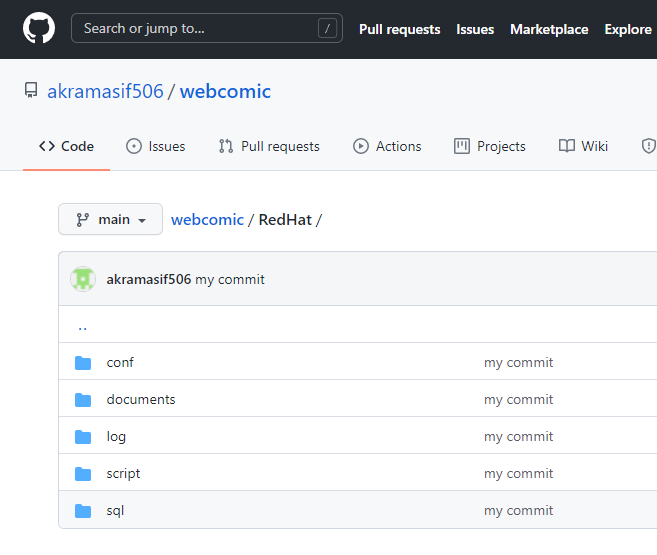
Based on the mentioned properties utility will extract all the comic information along with required metadata and will store in a MySQL database.

Project Setup:

**Git Repository : https://github.com/akramasif506/webcomic.git**

Project structure is look like below:

[GitHub - akramasif506/webcomic: Extract information of Comic from API](https://github.com/akramasif506/webcomic)



|  |  |
| --- | --- |
| **Folder Name** | **Description** |
| Conf | All the configuration files will be stored |
| Documents | Holds documents related to the project |
| Log | All the log files generated by the script |
| Script | Location of the ".py" file |

Prerequisites:

A) Python 3.x setup with below required libraries (mostly available by default):

|  |
| --- |
| Sys |
| Requests |
| Urllib |
| Webbrowser |
| Json |
| Random |
| Pprint |
| Mysql |

B) MySQL database with CREATE TABLE,SELECT and INSERT privileges.

Configuration Setup:

As per the requirement add or modify the properties in config.json.

[webcomic/RedHat/conf at main · akramasif506/webcomic · GitHub](https://github.com/akramasif506/webcomic/tree/main/RedHat/conf)



|  |  |  |  |
| --- | --- | --- | --- |
| Property Details |  |  |  |
| **Property Name** | **Description** | **Sample Value** | **Remark** |
| base\_url | url of the API to extract data from | https://xkcd.com/ |  |
| number\_of\_comic | Number of comic need to fetch in **single execution of job** | 15 |  |
| comic\_number\_range | Comic ID(number) range to be fetch from API | 87 | ID's of the comics fetched will be within this range from 1 to mentioned value |
| table\_name | Name of the target table created in Mysql DB |  |  |
| db\_properties | MySQL databse details |  |  |

Data Base objects:

Execute below DDL statement in database to create the table. <https://github.com/akramasif506/webcomic/tree/main/RedHat/sql>

***create table mysql.comic\_details***

***(***

***comic varchar(400),***

***alt\_text varchar(400),***

***num integer,***

***link varchar(400),***

***image varchar(400),***

***image\_link varchar(400)***

***);***

Script Details:

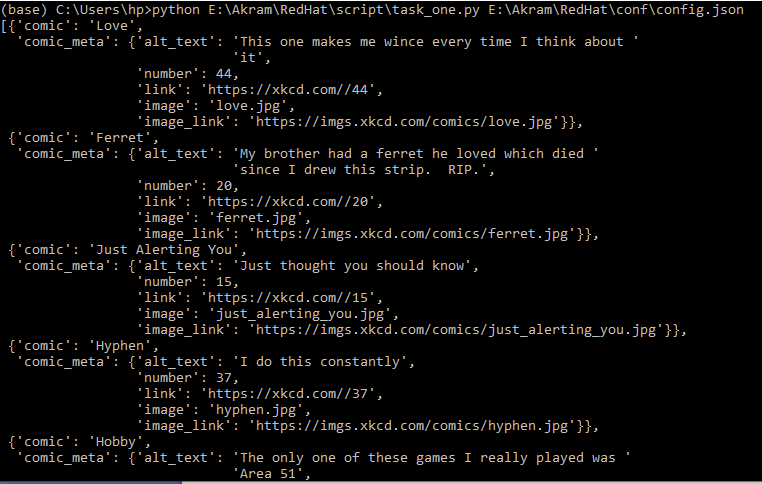
The script **task\_one.py** takes config.json file as input argument.

[webcomic/RedHat/script at main · akramasif506/webcomic · GitHub](https://github.com/akramasif506/webcomic/tree/main/RedHat/script)

**Command to execute:**

python E:\Akram\RedHat\script\task\_one.py E:\Akram\RedHat\conf\config.json

**Console output:**



**Database Table data:**

