**Documentation of Liquor**

**(Kakuyasu Japan)**

**System Details and Database:**

|  |  |
| --- | --- |
| System IP | 172.27.132.55 |
| System Password | Act@101 |
| Database IP | 172.27.132.55 (localhost) |
| Database Name | kakuyasu |
| Page Save Path | D:\PAGESAVE\KAKUYASU |

**VPN/Proxy Required:**

* Require VPN if not getting response please change VPN.

**Script/Code Details:**

|  |  |
| --- | --- |
| **Project Path** | D:\DATA\HINAL\LIQUOR\_SITES\GERMANY\kakuyasu |
| **Project Back-up path** | D:\BACKUP\Liquor\_backup\kakuyasu.zip |

### Points to be take care before running the script:-

### Change month date:

### Go to the db\_config.py file and change the month\_date variable value according to “DayMonthYear” format.

### Date will be fixed “08”

### For example: month=11, date: 08, year=2024 then month\_date = ‘08112024’

### Page Save file drive available in PC (for example: D) and check the file path as mentioned in db\_config.py file for page save & check EXPORT.py for excel output file.

1. Ensure that no link table exists before running **link\_extractor.py**. There is no need for old or new concepts; simply run the link\_extractor.py script, which will generate the links table each time.

### 

### Step for run the script for data extraction:-

### **Run the **link\_extractor.py** file** for extract links into the table as db\_links\_table name provide in db\_config.py file.

### You will also check during the count of all insert link using mention query as below.

### After running the ****link\_extractor.py****, you have require to run page\_save.py for page save.

### After the table is generated, saving the page is required using Selenium. There are two scripts available for saving the page: one uses threading, and the other does not. If you want to save the page faster, use the threading script. However, if you're experiencing response issues, use the normal page-saving script.

### While running page\_save.py you can check page save status using refer SQL queries mention as below.

### Once page save done then require to run data\_extractor.py.

### The data\_extractor.py script retrieves links from the spirit\_links table where the status is set to "SAVE" and processes them for data extraction.

### In this project new link add and also check old link as well for data extraction.

### For checking how many new links added please refer SQL queries mention as below.

### While running the script please check pages are proper save or not

### During running the script check the status:-

### In SQL execute query:

### for check the count of links added in spirit\_links table.

### “SELECT COUNT(\*) FROM ` spirit\_links`”

### for check page save count of links from spirit\_links table.

### “SELECT COUNT(\*) FROM ` spirit\_links` where status=”SAVE””

### for checking the status of extract data is “Pending”, “Done” or “Not Found”.

### “SELECT "Total", COUNT(\*) FROM `spirit\_links` UNION

### SELECT "Done", COUNT(\*) FROM `spirit\_links` WHERE STATUS="Done" UNION

### SELECT "Pending", COUNT(\*) FROM `spirit\_links` WHERE STATUS="Pending" UNION

### SELECT "SAVE", COUNT(\*) FROM `spirit\_links` WHERE STATUS=" SAVE" UNION

### SELECT "Not Found", COUNT(\*) FROM `spirit\_links` WHERE STATUS="Not Found"”

### How running further process:-

### First, run the file for link extraction. After that, run the data page save file. Finally, run the extraction file.

### Simply run the script; I have already set up the file generation code. Once the process is complete and the data extraction is finished, the file generation script will run automatically, and the Excel file will be available at the specified path.

### Make sure for points for file generation

### I have try to convert the file into English language but don’t depend on this translation make sure and verify from your side. (It’s not proper working for large data)

### If any cases file not generate after completion the process, run EXPORT.py file manual and if facing any issue then contact to developer.

### You have require to change manual translate from native language to English. (Using Google Sheet)

**Store Information:**

* **Primary Developer:** Hinal Patel
* **Secondary Developer:** Siraj Dal

**Note:** Additional point for developers