

Requirements of the MSO Licensee Register List Scapper

Saturday, February 29, 2020 4:46 PM

I am looking for someone to help me download an online database(directory)-The Licensee Register List of Money Services Operators using python and python libraries only.

The Customs of Hong Kong is responsible for licensing Money Service Operators (A.K.A Money Changers) in Hong Kong.

Objective:

The idea is to download the database into .csv, .json, and .xml data that can be manipulated by Microsoft access or imported into other databases. I need the data of both these in order to promote my services to them.

Programming language used:

1. Python-3

Libraries:

1. Lxml
2. Json
3. Selenium
4. BeautifulSoup

Output:

1. Must Have CSV
2. Optional JSON or XML Output

Where is the data?

Saturday, February 29, 2020 4:51 PM

You can search for the The Licensee Register List of Money Services Operators here <https://eservices.customs.gov.hk/MSOS/wsrh/001s1w>

Once you reach this page you can click "ALL" to find all the licencees, you can also click the pagination to go through all the pages.

Search Licensee Register List

Home > Search Licensee Register List

Search By Name of Licensee or Licence No. or Address

Name of Licensee (in English):

Name of Licensee (in Chinese):

Licence No.:

Address (in English):

Area:

District:

Search By Alphabetical Order

A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z | ALL

Search Result: 2156 Records Update Date: 03/04/2020 15:5

	Name of Licensee	Address	Address Type
1	169 LIMITED	FLAT A, 7/F, GOLDEN HILL MANSION, 209-211 HENNESSY ROAD, WAN CHAI, HONG KONG	Specified Premises
2	247 PAY LIMITED	ROOM 1845, 15/F, RADIO CITY, 505-511 HENNESSY ROAD, CAUSEWAY BAY, HONG KONG	Specified Premises
3	28 NETWORKING LIMITED 28資訊平台有限公司	FLAT 2202, 22/F, TRENDY CENTRE, 682 CASTLE PEAK ROAD, CHEUNG SHA WAN, KOWLOON	Specified Premises
4	3 SING INTERNATIONAL LIMITED 三昇聯匯有限公司	SHOP 28-85, 2/F, JADE PLAZA, 3 ON CHEE ROAD, TAI PO, NEW TERRITORIES (Branch Name : 3 SING CURRENCY EXCHANGE) (分行名稱 : 三昇聯匯)	Specified Premises
5	3 SING INTERNATIONAL LIMITED 三昇聯匯有限公司	G/F, BLOCK C, PO LOK BUILDING, 41-45 PO HEUNG STREET, TAI PO, NEW TERRITORIES (Branch Name : 3 SING CURRENCY EXCHANGE) (分行名稱 : 三昇聯匯)	Specified Premises
6	3 SING INTERNATIONAL LIMITED 三昇聯匯有限公司	SHOP 175, 1/F, KINGS WING PLAZA 1, 3 ON KWAN ST, SHEK MUN, SHA TIN, NEW TERRITORIES (Branch Name : 3 SING CURRENCY EXCHANGE) (分行名稱 : 三昇聯匯)	Specified Premises
7	7 STAR MONEY EXCHANGE COMPANY LIMITED 七星外幣兌換有限公司	SHOP C AND COCKLOFT, G/F, 2B SAN LOK STREET, SHEUNG SHUI, NEW TERRITORIES (Branch Name : 7 STAR MONEY EXCHANGE COMPANY LIMITED) (分行名稱 : 七星外幣兌換有限公司)	Specified Premises
8	7 STAR MONEY EXCHANGE COMPANY LIMITED 七星外幣兌換有限公司	SHOP 71, G/F, CHUNG KING MANSION, 36-44 NATHAN ROAD, TSIM SHA TSUI, KOWLOON	Specified Premises
9	A STAR ASIA COMPANY LIMITED 其星亞洲有限公司	SHOP 10, G/F, YIP CHEONG BLDG, 4-16 HILL RD, SAI YING PUN, HONG KONG	Specified Premises
10	A.C.T.	ROOM B, 24/F, TOWER 8, BAUHINIA GDN, TSEUNG KWAN O, NEW TERRITORIES	Specified Premises

Page 1 of 216

What data should be scrapped?

2020年4月3日 16:00



MSO_2020...

The output data should have a name MSO_YYYYMMDD where YYYY is the year the scrapper scrapped results for, MM the month, DD the day.

In the file there are 8 Columns

MSOID	English Name of Licensee	English Trading As	Chinese Name of Licensee	Chinese Trading As	English Address	English Branch Name	Chinese Branch Name	Address Type
-------	--------------------------	--------------------	--------------------------	--------------------	-----------------	---------------------	---------------------	--------------

I need you to fill in the data of the website into 8 columns,

- MSOID: A unique Identifier that identifies a unique record, a value we will use as a key in the database, I would prefer it to start with 0.
- English Name of Licensee: Is the english name of the licensee, you will find it in the "Name of Licensee" column in your data.
- English Trading as: Is the trading name of the licensee, some licensees will have it, some wont. For those that do, you will find it in the "Name of Licensee" column in your data. I want the data after "TRADING AS" to be included in this column.
- Chinese Name of Licensee: Is the Chinese name of the licensee, you will find it in the "Name of Licensee" column in your data.
- Chinese Trading as: Is the trading name of the licensee, some licensees will have it, some wont. For those that do, you will find it in the "Name of Licensee" column in your data. I want the data after "營業名稱" to be included in this column.
- English Address: Obviously this is the address, which you can find in the "Address" column in the data.
- English Branch Name: For some licensee, they will have a English branch name, look for the branch name in the address column.
- Chinese Branch Name: For some licensee, they will have a Chinese branch name, look for the branch name in the address column. I want the data after "分行名稱".
- Address Type: I want the Address type of the data to be included in this column.



×

Output Formating

2020年4月3日 16:20

Hints:

- Please remember to output the data as .csv, .xml and .json.
- The encode should be UTF-8.



MSO_2020...