Introduction to Computers and Programming

Week12 Lab 10

Due date: 2019/12/3 (Thu.)

Extract the information of the <u>57th Golden Horse Awards</u>.

Lab 1

最佳劇情長片 Award	l title
日子	汯
灣 消失的情人節	華文創股份有限公司,影一製作所股份有限公司,美商華納兄弟(遠東)股份有限公司台灣分公司,國賓影城股份有限公司,威像電影有限公司
同學麥娜絲	甜蜜生活製作有限公司,華文創股份有限公司
親愛的房客	一期一會影像製作有限公司
手捲煙	手捲煙電影製作有限公司
最佳紀錄片 Award	l title

Use BeautifulSoup to find out every Award title on this website.

鄧偉

爺爺和父親

Hint: You just need to find table with class name "table special award"

湖南省二友文化傳媒有限公司

Here is an example of the output:

```
['最佳劇情長片',
'最佳紀錄片
'最佳女配角
'最佳新演員
'最佳原著劇本
'最佳改編劇本
'最佳視覺效果
'最佳美術設計
'最佳造型設計
'最佳動作設計
'最佳原創電影音樂 '
'最佳原創電影歌曲',
'年度台灣傑出電影工作者',
'觀眾票選最佳影片 ',
'國際影評人費比西獎',
'終身成就獎 ']
```

最佳原創電影歌曲	
Happy Family	<u>詞</u> : 黃明志, 5forty2, The Real Masta Clan <u>曲</u> : 黃明志 <u>唱</u> : 黃明志, 5forty2, The Real Masta Clan (你是豬)
Lost and Found	<u>詞</u> :李霈瑜, 盧律銘 <u>曲</u> : 盧律銘 <u>唱</u> :李霈瑜 (消失的情人節)
孤味	<u>詞</u> : 林逸心 <u>曲</u> : 柯智豪 <u>唱</u> : 陳淑芳, 于子育 (孤味)
歡迎嚟到呢座城市	<u>詞</u> : 霍嘉豪, 阿弗, 陳心遙 <u>曲</u> : 戴偉, 霍嘉豪 <u>唱</u> : 霍嘉豪, 阿弗, 劉敬雯, Jan Curious, KIDA Choir (狂舞派 ₃)
刻在我心底的名字	<u>詞</u> : 佳旺, 陳文華, 許媛婷 <u>曲</u> : 佳旺, 陳文華, 許媛婷 <u>唱</u> : 盧廣仲, 陳昊森(刻在你心底的名字)

Must use regular expression to extract the information of yellow part in the award of "最佳原創電影歌曲". And print it in a nice way(see below).

(Movie name)
詞:	
曲:	
唱:	

Hint: You just need to find table with class name "table special award"

Hint: You can put? after + or * to make your search non-greedy.

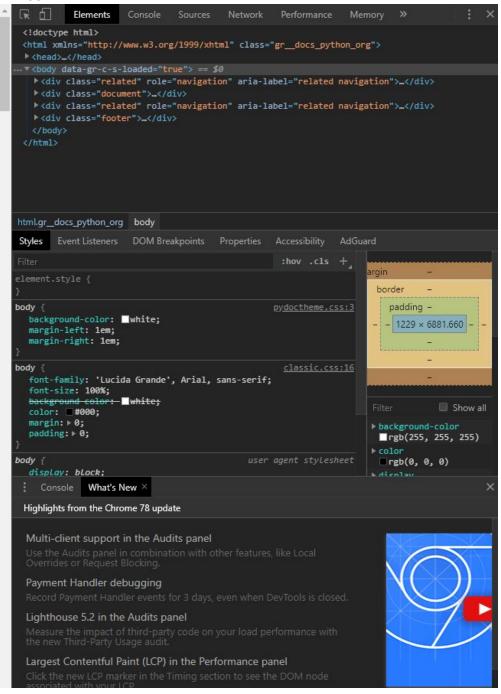
Here is an example of the output:

唱: 盧廣仲, 陳昊森

(你是豬) 詞: 黃明志, 5forty2, The Real Masta Clan 曲: 黃明志 唱: 黃明志, 5forty2, The Real Masta Clan (消失的情人節) 詞:李霈瑜, 盧律銘 曲: 盧律銘 唱: 李霈瑜 (孤味) 詞: 林逸心 曲: 柯智豪 唱: 陳淑芳, 于子育 (狂舞派3) 詞: 霍嘉豪, 阿弗, 陳心遙 曲: 戴偉, 霍嘉豪 唱: 霍嘉豪, 阿弗, 劉敬雯, Jan Curious, KIDA Choir (刻在你心底的名字) 詞: 佳旺, 陳文華, 許媛婷 曲: 佳旺, 陳文華, 許媛婷

Tools

1. In Google Chrome, you can press "F12" > "Elements" or "Ctrl + Shift + I" to open debugger



2. Then by clicking this icon, you can find the element on the website



3. Right-click on the highlighted element in the "Elements" panel, you can select "Copy" > "Copy selector", or "Edit as HTML". It's useful when coding a crawler.

