亚马逊

输入文件:csv(utf-8)

输出文件:csv(utf-8)

Step1.先打开网页: https://www.amazon.com/, 如图 1。同时根据图示 1 与图示 2 将语言、配送地址分别调整成英语和英国。然后根据图示 3 输入待搜索信息。

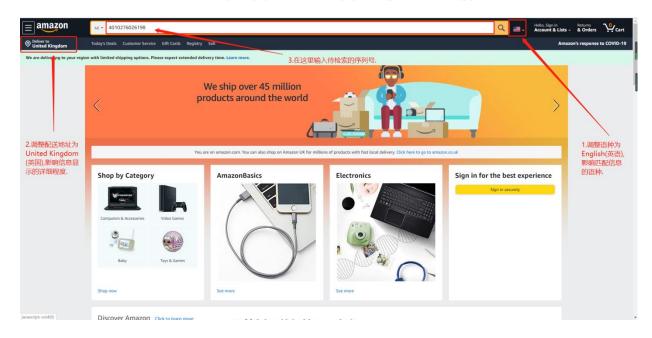


图 1

Step2.如果出现图 2 的情况,则说明该待搜索信息没有任何结果。那么在该编号前添加一个 0,返回 Step1 重新输入带搜索信息再次搜索,直到重复三次失败则放弃该编号。



图 2

Step3.如果出现图 3 的情况,点击第一个蓝色按钮以进入详情页。



图 3

Step4.如图 4.如若出现 Audio CD 的按钮,先点击,如图示 1。否则默认为当前页。随后按照图示 2 与图示 3 的位置抓取'245a'与'245c'字段。注意,'245c'字段只取分号前的第一个人名,且根据逗号前后互换;再互换回来的名字作为'100a'字段.

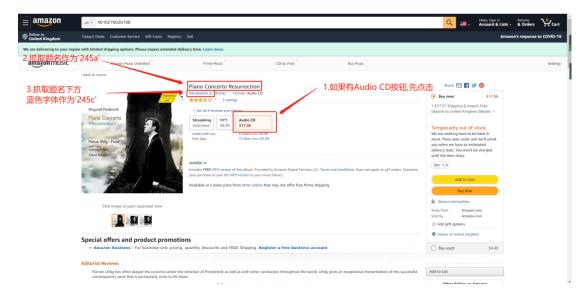


图 4

Step5.页面下拉,看 Product details 信息.

- 如果有 Label 信息,写入'028b'与'260b'字段,如图 5 图示 1
- 如果有 Number of Discs 信息,写入'024c'字段,并以"{} audio disc(s)"形式写入'300a'字段,如图 5图示 2
- 如果有 Date First Available 信息,将后面信息连续四位数字写入'260c'字段,
 如图 5 图示 3
- 设置'300b'、'300c'字段分别为"digital","4 3/4 in."
- 如果有 Total Length 信息,用六位数字写入'306a'字段,缺实例?
- ◆ 如果有 Genres 字段,将后面信息填入'650a'字段,缺实例?



图 5(例:4010276026198)

Step 6.页面拉回来,尝试点击 Streaming 按钮,如图 6.

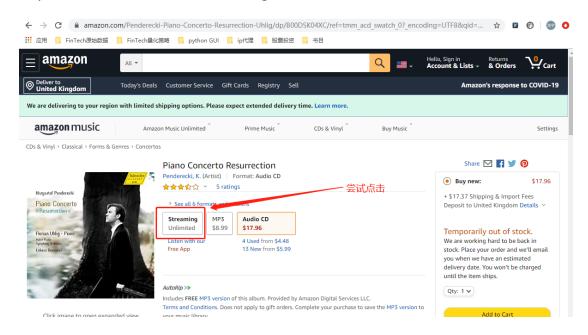


图 6

Step7.尝试抓取节目单,并录入字段'505a'.如图 7.合并字段规则如下:

- 如若有多个标题,则不同题目之间用' -- '隔开
 - 同一标题下的节目以";"隔开
- 如若没有多个标题,则节目之间用' -- '隔开,用" {}.{} ({})" (序号、内容、时间)形式填写

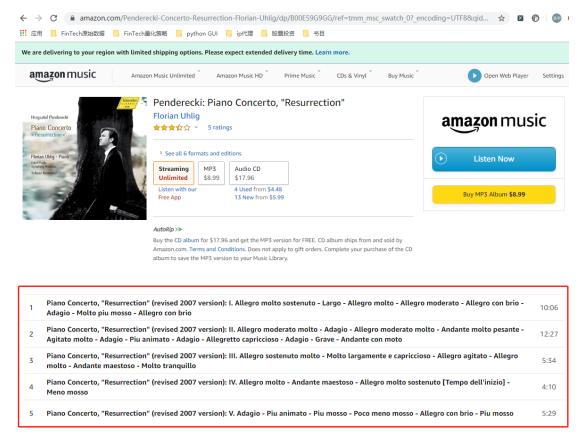


图 7

最终结果示意:

024a:4010276026198

245a:Piano Concerto Resurrection

100a:Penderecki, K.

245c: K. Penderecki

024c:[CD 1]

300a:1 audio disc

028b:Haenssler Classics

260b: Haenssler Classics

300b:digital

300c:4 3/4 in.

505a:1.Piano Concerto, "Resurrection" (revised 2007 version): I. Allegro molto sostenuto - Largo - Allegro molto - Allegro moderato - Allegro con brio - Adagio - Molto piu mosso - Allegro con brio (10:06) -- 2.Piano Concerto, "Resurrection" (revised 2007 version): II. Allegro moderato molto - Adagio - Allegro moderato molto - Andante molto pesante - Agitato molto - Adagio - Piu animato - Adagio - Allegretto capriccioso - Adagio - Grave - Andante con moto (12:27) -- 3.Piano Concerto,

"Resurrection" (revised 2007 version): III. Allegro sostenuto molto - Molto largamente e capriccioso - Allegro agitato - Allegro molto - Andante maestoso - Molto tranquillo (5:34) -- 4.Piano Concerto, "Resurrection" (revised 2007 version): IV. Allegro molto - Andante maestoso - Allegro molto sostenuto [Tempo dell'inizio] - Meno mosso (4:10) -- 5.Piano Concerto, "Resurrection" (revised 2007 version): V. Adagio - Piu animato - Piu mosso - Poco meno mosso - Allegro con brio - Piu mosso (5:29)