

DojinHakusho User Manual

DojinHakushoをダウンロード頂きありがとうございます。

本ツールは、DLsiteで活躍するクリエイターの情報を収集・解析するフリーソフトであり、特に同人音声で活動するサークル様・声優様の作品を解析対象とした利用を想定しています。

本ファイルにはツールのインストール方法と基本的な使用方法が記載されます。内容を熟読されたうえでご利用いただきますようお願い申し上げます。

注意事項：

ツールも解説もいたって真面目ですが、取り扱う分野の性質上一部アダルトな内容が含まれます。18歳未満の方はご使用をお控えください。

利用規約

本ツールはDLsiteを利用するクリエイター様を支援することを目的に開発されたツールです。従って、本ツールによって得られるデータを、その目的から大きく逸脱して使用することはお控えください。（例えば、データを用いた対立煽りや中傷行為などがこれに該当します。）

また、本ツールはDLsite様にアクセスしHTMLソースを取得することで動作するクローラーです。従って、本ツールの利用は常識の範囲内で行い、サーバ管理者から攻撃と見做されるような行為はお控えください。

上記規約に従う限り、本ツールは完全にフリーです。リンク、商用利用、転載、改変、二次配布などは自由に行っていただいて構いませんし、連絡の必要もありません。なお、製作者はその結果生じたいかなる事態にも責任を負いかねます。

インストール

環境構築

pythonおよび必要なライブラリをインストールしてください。

必要なライブラリはnetworkx, matplotlib, pandas, numpyなど基本的なものが大半です。

インストール

[DojinHakushoの頒布場所](#)へアクセスし、ツールを任意のディレクトリにダウンロードしてください。

使用方法

まず『main.py』を適当なエディタで開き、この行に調べたいサークル名と声優名を入力してください。名前は必ずDLsiteのクリエイタータグ名と一致させる必要があります。検索対象となるデータは、変数で設定されたサークルと声優の積集合です。どちらも不問の場合は空欄で構いませんが、両方を空欄にすると実行時間が長くなるのでオススメしません。

なお、今回はサークル名を空欄、声優名を『砂糖しお』様に設定したと仮定して解説を行います。

```
# ここに調べたいサークルの名前を入力する。
# 名前はDLsiteのサークル名と一致させる必要がある
# サークル不問の場合は空欄にすること
circle_name = ""

# ここに調べたい声優の名前を入力する。
# 名前はDLsiteのクリエイタータグと一致させる必要がある。
# 声優不問の場合は空欄にすること
creator_name = "砂糖しお"
```

データを取得するにはコマンド上でインストール先のディレクトリに移動し、以下のコマンドを入力します。

```
> cd DojinHakusho
> python3 main.py
```

上記の命令でツールを実行すると、このような実行結果が得られるかもしれません。

```
> 検索条件：
> 検索条件：
> サークル名      : ""
> 声優名          : "砂糖しお"
> 販売状況        : "販売中"
> 対象性別        : "男性向け"
> ファイル形式    : "オーディオファイル"

> 探索終了までお待ちください...

> 探索終了！
> 作品数          : 127 [個]
> 実行時間        : 23.5 [sec]
```

実行が終了したため、得られた生データ(.csvファイル)を確認してみましょう。ツールが問題なく終了した場合、生データはdataフォルダに生成されるはずです。

(ファイルが文字化けする際は、一度テキストファイルで開き、文字コードをANSIに変更してください。)

(当方の環境だとなんかpipが衝突しているようで、面倒なのでグラフの描画はpython2で動かしています。)

データに対する評価

まずは主要統計量についての解説を行います。

	price	sales
count	127.000000	127.000000
mean	1063.622047	986.755906
std	474.051550	1564.399041
min	110.000000	22.000000
25%	880.000000	229.000000
50%	1100.000000	460.000000
75%	1100.000000	921.000000
max	3520.000000	8972.000000

countはデータの個数, meanは平均値, stdは標準偏差です. min,maxはそれぞれ最小値と最大値を示します. 25%, 50%, 75%はそれぞれ第一から第三四分位数を指します. 今回の解説では, 第二四分位数, つまり中央値をベンチマークとして用います. これは, 同人音声作品は一部の大手と言われるサークルが大量の売上本数を出す傾向にあり, 平均値が実情と剥離すると判断したためです.

出演作の50%が880円から1100円の範囲でリリースされていること, 出演作の50%が460本以上の売上を出していること等を読み取ることができます.

製作者の一言:

勘違いを防ぐために言っておきますが, 中央値は『高い=人気声優』『低い=不人気声優』という意味の指標ではありません. 例えば, 新規サークルだろうが大量に出まくる様な方は, 必然的に中央値が下がる傾向にあります. つまり, 中央値の低い方にも素晴らしい演技をする方はたくさんいるということです...声優さんの能力というのは定量的な評価が難しいのです.

さて, 最後に統計を扱う上での素晴らしい格言を引用しておきます. 『数字は嘘をつかないが, 嘘つきは数字を使う.』

次に売上の箱ひげ図を見ていきます. 結構外れ値が多い. pythonはデフォルトで箱の1.5倍以上の距離にある値を外れ値として扱います.

この箱ひげ図から言えば, 売上2000本を超えればかなりのヒット作と言えます.

最多ヒットは072LABO様の早漏遺伝子滅亡計画1で約9000本, 更に同サークルのマゾヒストプランナーシリーズが7500本, 6500本と続きます.

以下定性的評価(つまり製作者の感想):

何故マニュアルの事例として砂糖しお様を用いたかという点, 製作者が個人的にファンだからです. 無邪気で天真爛漫な小悪魔と言った風情の声や, 「キャハハ」みたいな笑い声などが魅力的な方だと認識しています. ちなみに, 製作者は砂糖しお様を『あなたを壊す強制寸止めオナサポ/アルファートルル(RJ195177)』で知りました.

まあそんな話は端に置いて, まずはベンチマークとなる作品を視聴します. これは解析対象への最低限の礼儀です. タイトルは[磁性体性愛情論理](#)

R-15作品だがこれはいいねクッソかわいい. 抜きはないがしあわせな気持ちになった.

それでは, 砂糖しお様の今後の益々のご活躍を祈りつつ, 筆を置くことにします. あと下に何作かオススメの出演作を紹介しておくのでご参考ください.

Issue / Request
