Programming, Problem Solving, and Algorithms

CPSC203, 2021 W2

Announcements

- Test 2 is this week
- Lab3 is this week
 - It's about DataClasses!
- Problem of the Week 3 is due this week (extended from last week)
 - Assigned Pandas videos from Thursday should help!
- Problem of the Week 4 is also due this week
 - Dataclasses practice

Today's Plan...

- 1. Announcements! (5 mins)
- 2. Review/Questions (10 mins)
- 3. Billboard 100 Demo (65 mins)

Slides from the Assigned Videos

103 to 203

Typical Introductory Data Flow:

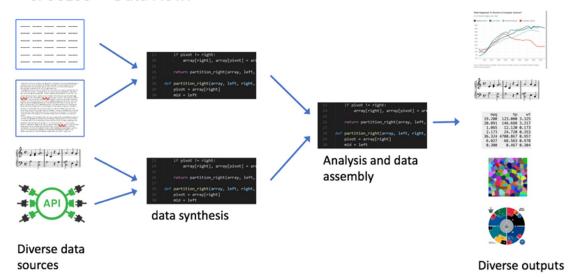


.csv file

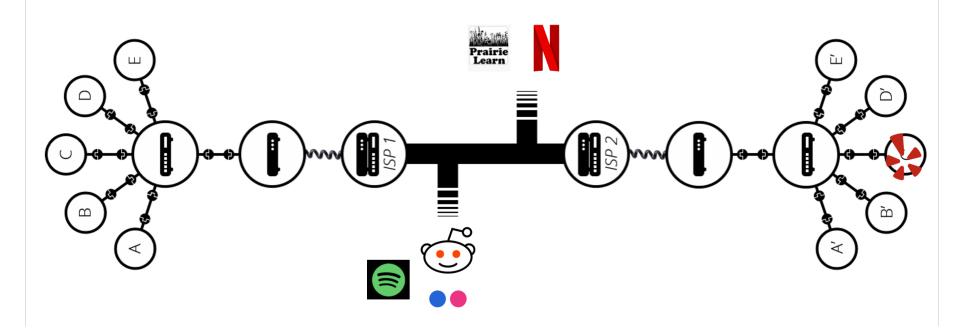
Python problem solution using simple data types and elementary list iteration.

Matplotlib bar or line graph or other summative output illustrating results of computation.

CPSC103++ Data Flow:



The internet...



Billboard Hot 100...

Navigate to https://www.billboard.com/charts/hot-100
What happens to the URL if you load a past week?
What happens to the page if you substitute a different date into the URL?
—— Write one question you would like to ask of this data:

Anatomy of html...

<!DOCTYPE html>

```
<html><head><title>The Dormouse's story</title></head>
   <body><b>The Dormouse's story</b>
      Once upon a time there were two little sisters.
   Their names were <a href="http://example.com/elsie" class="sister"
   id="link1">Elsie</a>, and <a href="http://example.com/lacie"
   class="sister" id="link2">Lacie</a>, and they lived at the bottom of
   a well.
   </body>
</html>
```

Billboard Hot 100... page source

```
<div class="chart-list-item piano-content-overlay_gated-item" data-rank="49" data-artist="Taylor Swift" data-title="Lover" data-has-content="true"> <div
class="chart-list-item first-row chart-list-item cursor-pointer"> < div class="chart-list-item position chart-list-item position--centered">
                                                                                                                                          <div
class="chart-list-item rank"> 49 </div>
                                                                                                                                          <div
class="chart-list-item award"> </div> </div>
                                                                                                                                          <div
class="chart-list-item image-wrapper"> < div class="chart-list-item trend-icon">
                                                                                                                                         <img
src="https://assets.billboard.com/assets/1568911107/images/charts/arrow-down.svg?df89925e3b37f64521bd"
srcset="https://assets.billboard.com/assets/1568911107/images/charts/arrow-down-mobile.svg?df89925e3b37f64521bd 30w,
https://assets.billboard.com/assets/1568911107/images/charts/arrow-down.svg?df89925e3b37f64521bd 38w" sizes="(min-width: 768px) 38px,
30px"></div>
<img src="data:image/png;base64,iVBORw0KGgoAAAANSUhEUgAAAAEAAABCAYAAAAfFcSJAAAADUlEQVR42mNkYPhfDwAChwGA60e6k
gAAAABJRU5ErkJggg==" data-src="https://charts-static.billboard.com/img/2019/08/taylor-swift-p7u-lover-tuk-53x53.jpg" data-srcset="https://charts-
static.billboard.com/img/2019/08/taylor-swift-p7u-lover-tuk-53x53.jpg 53w, https://charts-static.billboard.com/img/2019/08/taylor-swift-p7u-lover-tuk-
106x106.jpg 106w, https://charts-static.billboard.com/img/2019/08/taylor-swift-p7u-lover-tuk-87x87.jpg 87w, https://charts-
```

static.billboard.com/img/2019/08/taylor-swift-p7u-lover-tuk-174x174.jpg 174w" sizes="(max-width: 767px) 72px, (min-width: 768px) 86px" class="chart-list-

<div class="chart-list-item__text-wrapper"> <div class="chart-list-item__text "> <div class="chart-list-item__title">

Lover

item image" alt="Taylor Swift Lover Billboard Hot 100"></div>

```
Lover
</span> </div>
<div class="chart-list-item" artist">
<a href="/music/taylor-swift">
Taylor Swift
</a>
</div>
<div class="chart-list-item lyrics">
<a href="https://www.billboard.com/articles/news/lyrics/7950218/ready-for-it-taylor-swift-lyrics">
<span class="hidden-mobile show-expanded-mobile-inline">Song </span>Lyrics
</a></div></div>
<div class="chart-list-item chevron-wrapper"><i class="fa fa-chevron-down"></i></div></div>
<div class="chart-list-item extra-info"><div class="chart-list-item extra-info-shadow"></div>
<div class="chart-list-item stats">
<div class="chart-list-item stats-cell basic-user chart-list-item stats-cell--first-cell"> <div class="chart-list-item stats-icon fa fa-arrow-up fa-rotate-
45"></div>
<div class="chart-list-item last-week">23</div>
LAST WEEK </div>
<div class="chart-list-item" stats-cell basic-user "> <div class="chart-list-item" stats-icon fa fa-arrow-up fa-rotate-45"></div>
<div class="chart-list-item last-week">10</div>
TWO WEEKS AGO</div>
<div class="chart-list-item stats-cell basic-user"> <div class="chart-list-item stats-icon fa fa-line-chart"> </div>
<div class="chart-list-item"> weeks-at-one">10</div>
PEAK POSITION </div>
<div class="chart-list-item stats-cell basic-user chart-list-item stats-cell--no-border-right"><div class="chart-list-item stats-icon fa fa-clock-o"></div>
<div class="chart-list-item" weeks-on-chart">4</div>
NAMERICO ON CHART Adding Adding Adding
```

Beautiful Soup

Reads the html source into a data structure that's easy to query!

https://www.crummy.com/software/BeautifulSoup/bs4/doc/

```
html = simple_get("https://www.billboard.com/charts/hot-100" + '/' + date)
mydivs = html.findAll("div", {"class": "chart-list-item"}) // all the data is here!!

for div in mydivs:
    s = Song(div.attrs['data-title'], div.attrs['data-artist'], int(div.attrs['data-rank']))
```

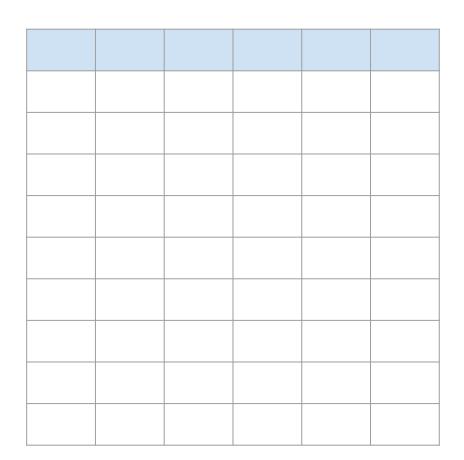
Still too messy for us! Remedy? https://github.com/guoguo12/billboard-charts demo...

Given last week's chart,

- 1) How many new songs were there?
- 2) What's the average peak?
- 3) Among those who were on the list for more than 10wk, what's the average peak? (is it very different than the previous answer?)
- 4) Which song changed the most? Was it rising or falling?
- 5) Write and answer your own question:

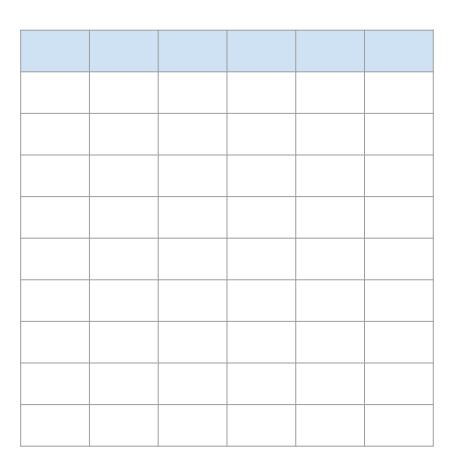
Given last week's chart,

How many new songs were there?



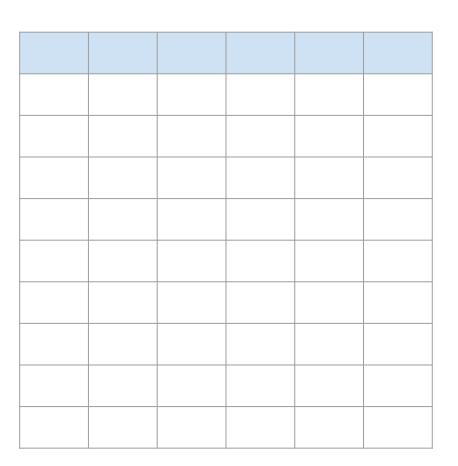
Given last week's chart,

What's the average peak?



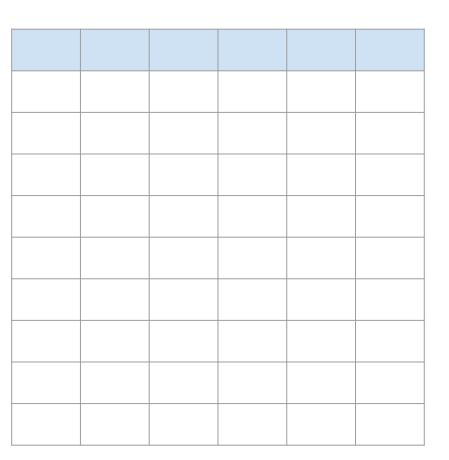
Given last week's chart,

Among those who were on the list for more than 10wk, what's the average peak? (is it very different than the previous answer?)



Given last week's chart,

Which song moved the most? Did it rise or fall?



Given last week's chart,

Write and answer your own question:

ToDo for next class...

POTD: Continue every weekday! Submit to PL.

References:

https://pandas.pydata.org/Pandas_Cheat_Sheet.pdf

https://www.crummy.com/software/BeautifulSoup/bs4/doc/