

Open Data Dashboards & Python Web Scrapping

PyData DC 2016

Marie Whittaker



mseew



@mariecwhittaker

A bit about me

Studied economics

here



Worked at **cfpb** for 3 years



using data to prevent these

Now with



monitoring the economy of the 51st State



What is data competing with?

Pittsburgh's Youth-Driven Food Boom

By JEFF GORDINIER MARCH 15, 2016

MILLENNIALS PREFER CITIES TO SUBURBS, SUBWAYS TO DRIVEWAYS

CONSUMER | 03-04-2014

[Tweet](#)

[in Share](#)

[f Share](#)

[submit](#)

Millennials Don't Want Work Life Balance

Apr 7, 2015 | 697,768 views | [2,006 Likes](#) | [278 Comments](#)

Millennials And Money: Give Us More Transit Options

by Benét J. Wilson / April 14, 2015

D.C. Sees Significant, "Hip" Population Growth

By Tom Sherwood

The Clearest Explanation Yet for Why Millennials Are Driving Less

Shifting demographics matter, but shifting attitudes may matter more.

ERIC JAFFE | [@e_jaffe](#) | Jul 13, 2015 | [74 Comments](#)

Millennials Investing in Rental Properties

Mortgages

By LISA PREVOST NOV. 20, 2015

THE BLOG

A Trend Story About Millennials, By The New York Times

© 03/24/2016 05:15 pm ET | Updated Mar 24, 2016

Is Your Neighborhood Changing? It Might Be Youthification, Not Gentrification

One urban planning professor has defined this as a process that occurs in discrete stages.

RICHARD FLORIDA | [@Richard_Florida](#) | Feb 5, 2015 | [37 Comments](#)

Small Kitchens, by Choice

By KIM SEVERSON DEC. 7, 2015

The six best cities for single Millennials

Many Millennials are both single and career focused - and there are plenty of cities around the country that cater to this young and diverse demographic.

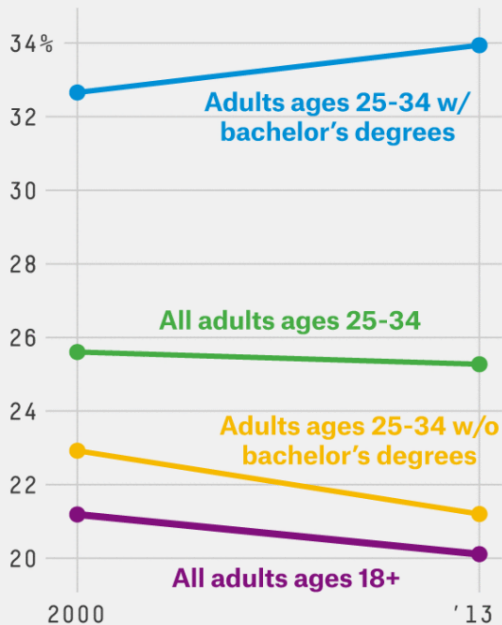
APR 5, 2016 @ 09:25 AM

219

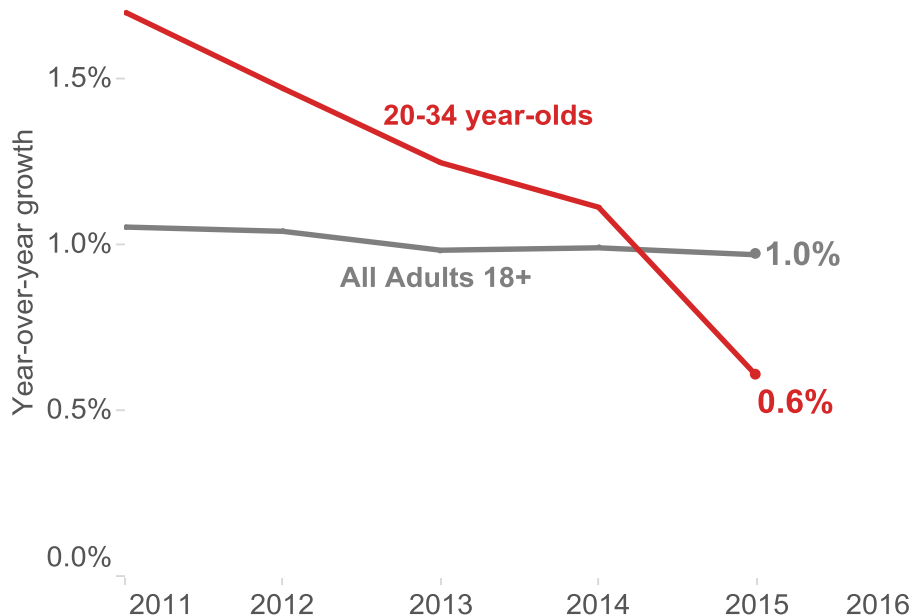
Why Do Millennials Move? The Answers May Surprise You

What is data competing with?

Share of adults who live in **urban** neighborhoods



DC Population Growth Rate (year-over-year)
2010 to 2015



Stevens Elementary School

What makes a good dashboard?

What makes a good dashboard?

1. It's consistent.
2. It includes the right data.

What makes a good dashboard?

3. It answers the questions you get asked most frequently.

4. (Sometimes) it raises more questions for you to dive into elsewhere.

Who are dashboards for?

Who are dashboards for?

- The audience
 - For initial context-setting
 - For a quick reference

Who are dashboards for?

- The audience
 - For initial context-setting
 - For a quick reference
- The **creator**

Problem?

factfinder.census.gov/faces/nav/jsf/pages/searchresults.xhtml?refresh=t

U.S. Department of Commerce
American FactFinder

MAIN COMMUNITY FACTS GUIDED SEARCH **ADVANCED SEARCH** DOWNLOAD CENTER

Search - Use the options on the left (topics, geographies, ...) to narrow your search results

Your Selections
"Your Selections" is empty
load search | save search

Search using the options below:

- Topics (age, income, year, dataset, ...)
- Geographies (states, counties, places, ...)
- Race and Ethnic Groups (race, ancestry, tribe)
- Industry Codes (NAICS industry, ...)
- EEO Occupation Codes (executives, analysts, ...)

To search for tables and other files in American FactFinder:

- 1 Enter search terms and an optional geography and click GO

topic or table name state, county or place (optional) GO ?

☒ topics ☐ race/ancestry ☐ industries ☐ occupations

-- or --

Select from **Topics**, **Race and Ethnic Groups**, **Industry Codes**, **EEO Occupation Codes**.

- these are added to "Your Selections"
- the Search Results are updated

- 2 Next, select **Geographies** (states, counties, cities, towns, etc.)

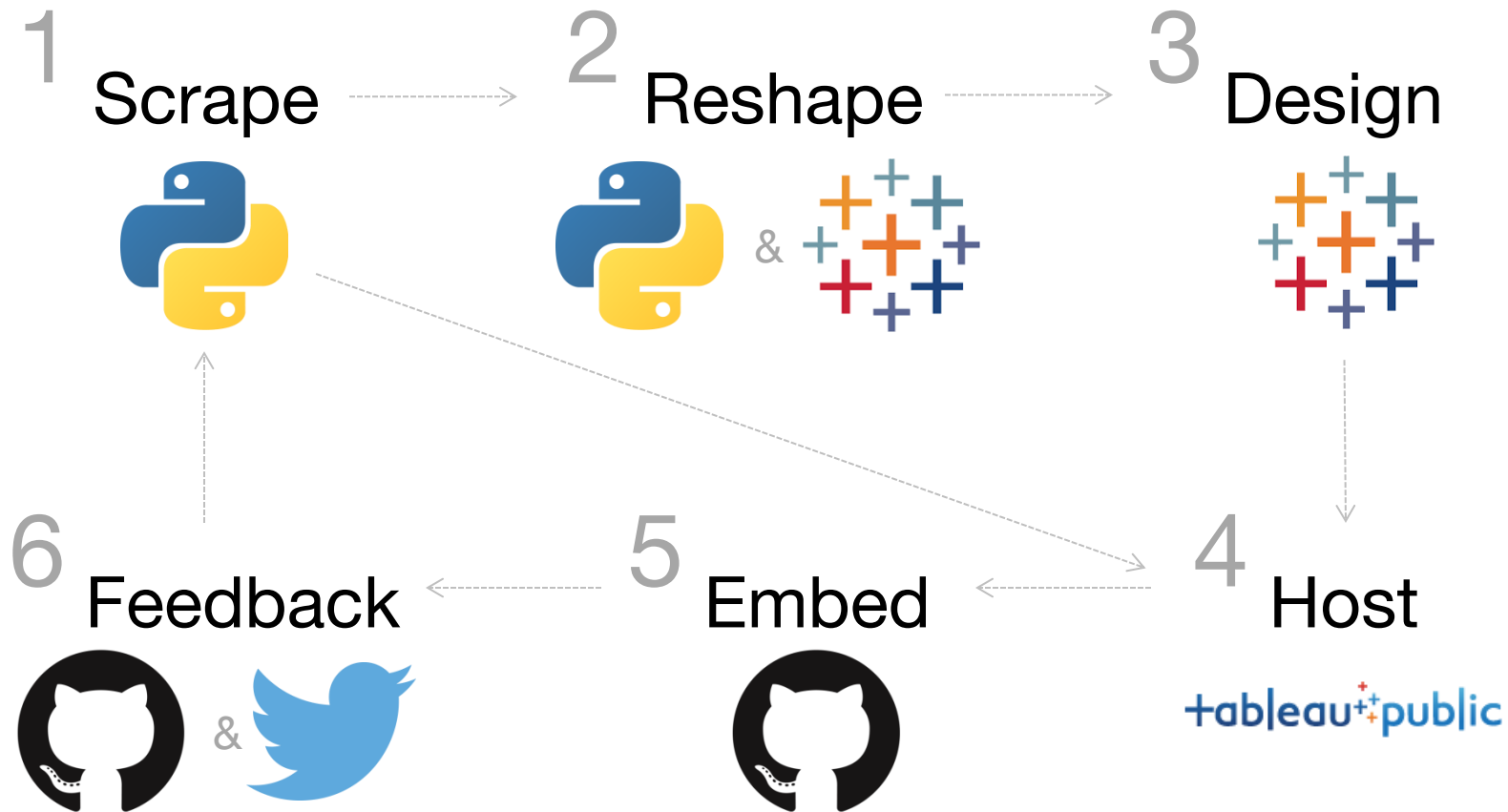
- these are added to "Your Selections"
- the Search Results are updated

- 3 Select one or more Search Results and click **View**

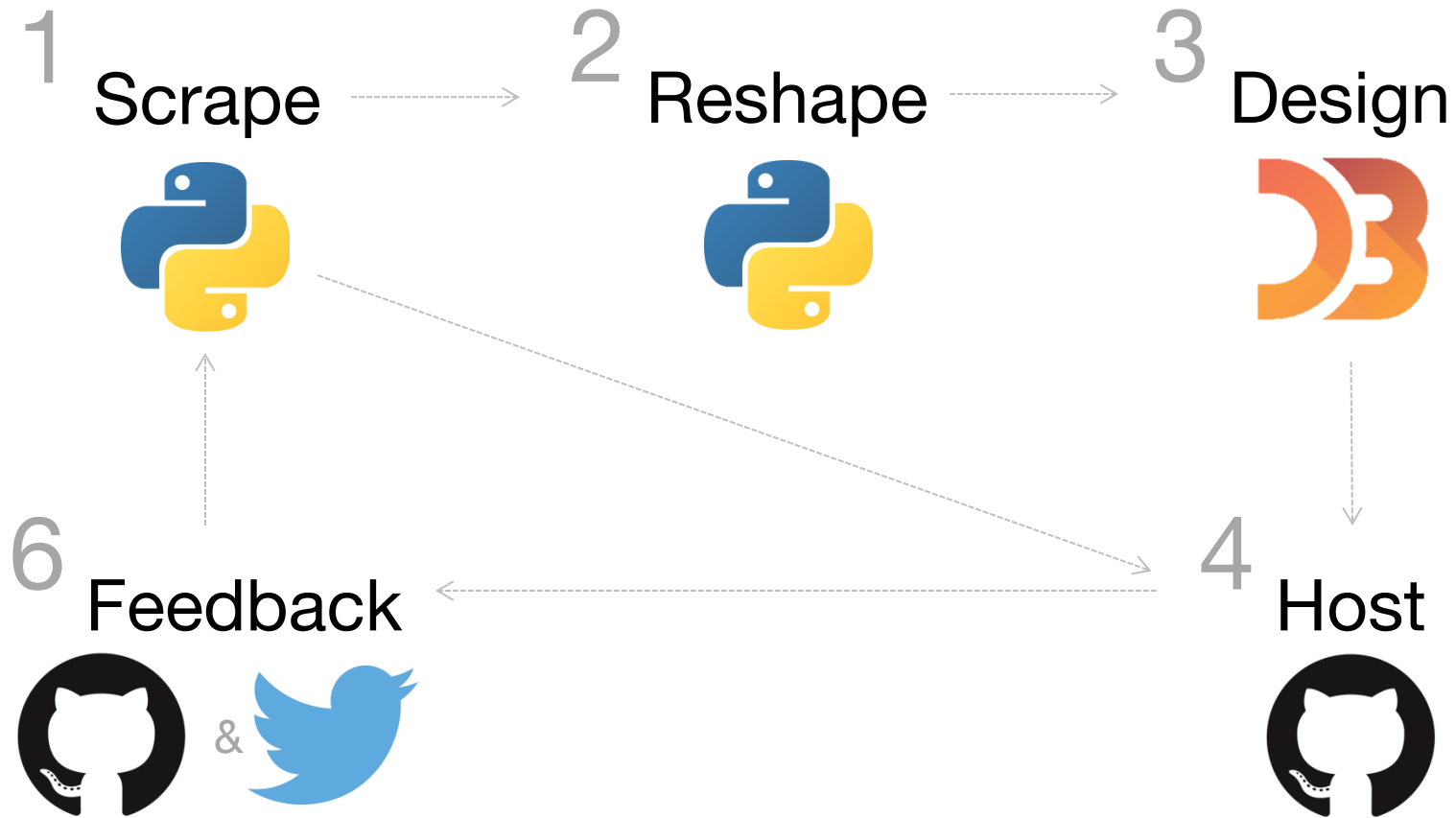
PROPERTY TAX
TOTAL PROPERTY
Real Property Tax
Transfer to Pk
Transfer to TIF
Real Property Tax
Personal Property
Public Spaces
TOTAL PROPERTY
SALES TAXES
TOTAL SALES TA
GENERAL SALES
General Sales And
Convention Ce
Transfer to Ballpark Fun
Transfer to Healthy Schools
Transfer to Prepaid Wireless (U
Transfer to TIF
Transfer to WMATA
General Sales And Use Tax (net)
Plastic Bag Use Fee
TOTAL GENERAL SALES AND USE TAXES
SELECTIVE SALES TAXES
Alcoholic Beverage

626,884 543,246 83,636 15.4% 626,884 543,246 83,636 15.4%

Workflow



Workflow




```
... ▼ <form method="POST" action="/cgi-bin/bldgprmt/bldgdisp.pl"> == $0
    <input type="hidden" name="o" value="Monthly">
    <input type="hidden" name="S" value="11District of Columbia">
    <input type="hidden" name="M" value="08">
    <input type="hidden" name="Y" value="2016">
    <input type="hidden" name="r" value="Place">
    <input type="hidden" name="A" value="2015 1996 2016">
    <input type="hidden" name="SortOpt" value="Place">
    ► <font size="-1">...</font>
```



```
<input type="hidden" name="S" value="11Distriict of Columbia">
```

Helpful packages

requests

xlrd Reads excel files.

xlsxwriter Writes excel files.

openpyxl Also writes excel files.
Plays well with *pandas*.

pandas Does everything – I use
for reshaping.

json

Various APIs

Various wrappers (Like *fredapi*)

configparser For calling hidden API
keys in a local config file.

BeautifulSoup For parsing HTML.

mechanize For emulating a browser.
Useful when data can
only be reached via web
forms.

Helpful packages

A search input field with a thin grey border. On the right side of the field is a small microphone icon, indicating voice search functionality.

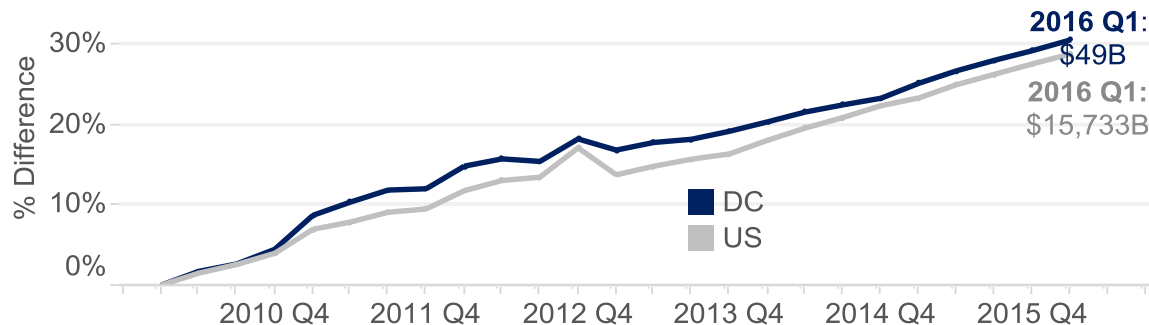
Google Search

I'm Feeling Lucky

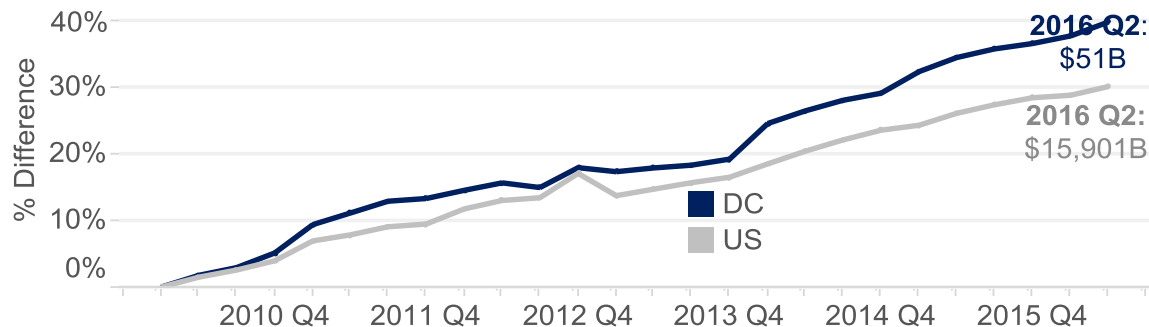
Revisions

Total Personal Income | % Change, Indexed to 2010

2010 Q1 to 2016 Q1



2010 Q1 to 2016 Q2



Source: Bureau of Economic Analysis

Demo

<http://open.dc.gov/economic-intelligence/>

THANKS!

<https://Github.com/dcgov>