Web Scraping /
Web Application Development 2

Jennifer Helsby, Eric Potash
Computation for Public Policy
Lecture 17: March 1, 2016

computationforpolicy.github.io

Today

- Go over Homework 4
    - If you have not submitted, submit now!
    - Any homeworks handed in after solutions are up *will not receive credit!!!1*
- Cover more web application development if time permits

Announcements

- HW5 optional but online, due by March 8
- Final Project rubric online:

https://computationforpolicy.github.io/assignments/final.html

Final Project Rubric

- Git repository: Due 2/17. We will be grading you based on:
  - (30%) Clear, well-documented code. Unless you used private data that you cannot commit to your repository, we should be able to run your code to replicate your analysis.
  - (10%) README.md: Use Markdown to create a README.md file in your repository that briefly (equivalent of 1 page) summarizes the purpose and findings of your project, including some tables and visualizations.
- Project presentation: During finals week. At least 10 minutes. For groups or 2 or more, at least 20 minutes but no more than 30 minutes. We will be looking for the following (but feel free to use a different outline):
  - Introduction and motivation: Why did you select this problem?
  - Data: Describe your data including how you selected and collected it.
  - (30%) Analysis and Visualization
  - (30%) Result: What does your work mean? What questions does it answer?
  - Further work: If you were to continue this work, what extensions would you pursue?

HW4 solutions

https://github.com/computationforpolicy/lecture-examples/blob/master/Solutions%204.ipynb

Using render_template

```python
from flask import Flask
from flask import render_template
import pandas as pd
app = Flask(__name__)


@app.route('/')
def load_dataframe():
    filename = '../Current_Employee_Names__Salaries__and_Position_Titles.csv'
    df = pd.read_csv(filename)
    return render_template('unfiltered.html', data=df.to_html())

if __name__ == '__main__':
    app.run(debug=True)
```

City of Chicago Employee roster and salaries

	Name	Position Title	Department	Employee Annual Salary
0	AARON, ELVIA J	WATER RATE TAKER	WATER MGMNT	$88968.00
1	AARON, JEFFERY M	POLICE OFFICER	POLICE	$80778.00
2	AARON, KARINA	POLICE OFFICER	POLICE	$80778.00
3	AARON, KIMBERLEI R	CHIEF CONTRACT EXPEDITER	GENERAL SERVICES	$84780.00
4	ABAD JR, VICENTE M	CIVIL ENGINEER IV	WATER MGMNT	$104736.00
5	ABARCA, ANABEL	ASST TO THE ALDERMAN	CITY COUNCIL	$70764.00
6	ABARCA, EMMANUEL	GENERAL LABORER - DSS	STREETS & SAN	$40560.00
7	ABBATACOLA, ROBERT J	ELECTRICAL MECHANIC	AVIATION	$91520.00
8	ABBATEMARCO, JAMES J	FIRE ENGINEER	FIRE	$90456.00
9	ABBATE, TERRY M	POLICE OFFICER	POLICE	$86520.00
10	ABBOTT, BETTY L	FOSTER GRANDPARENT	FAMILY & SUPPORT	$2756.00
11	ABBOTT, LYNISE M	CLERK III	POLICE	$43920.00
12	ABBRUZZESE, WILLIAM J	INVESTIGATOR - IPRA II	IPRA	$72468.00
13	ABDALLAH, ZAID	POLICE OFFICER	POLICE	$69684.00
14	ABDELHADI, ABDALMAHD	POLICE OFFICER	POLICE	$80778.00
15	ABDELLATIF, AREF R	FIREFIGHTER (PER ARBITRATORS AWARD)-PARAMEDIC	FIRE	$98244.00
16	ABDELMAJEID, AZIZ	POLICE OFFICER	POLICE	$80778.00
17	ABDOLLAHZADEH, ALI	FIREFIGHTER/PARAMEDIC	FIRE	$87720.00
18	ABDUL-KARIM, MUHAMMAD A	ENGINEERING TECHNICIAN VI	WATER MGMNT	$106104.00
19	ABDULLAH, DANIEL N	FIREFIGHTER-EMT	FIRE	$91764.00
20	ABDULLAH, KEVIN	LIEUTENANT	FIRE	$110370.00
21	ABDULLAH, LAKENYA N	CROSSING GUARD	POLICE	$16692.00
22	ABDULLAH, RASHAD J	ELECTRICAL MECHANIC-AUTO-POLICE MTR MNT	GENERAL SERVICES	$91520.00
23	ABDUL-SHAKUR, TAHIR	GENERAL LABORER - DSS	STREETS & SAN	$40560.00

Jinja2 Templates

- Jinja2 syntax very similar to regular Python
- Text between {{ and }} or {% and %} rendered

template

```
<table class="table table-striped table-hover
condensed">
    <tr>
    {% for col in cols %}
      <th>{{ col }}</th>
    {% endfor %}
    </tr>

    {% for row in range(nrows) %}
      <tr>
      {% for col in cols %}
        <td>
        {{ data[col][row] }}
        </td>
      {% endfor %}
      </tr>
    {% endfor %}

</table>
```

python

```
table-
```

template

```html
<table class="table table-striped table-hover
condensed">
    <tr>
    {% for col in cols %}
      <th>{{ col }}</th>
    {% endfor %}
    </tr>

    {% for row in range(nrows) %}
      <tr>
      {% for col in cols %}
        <td>
        {{ data[col][row] }}
        </td>
      {% endfor %}
      </tr>
    {% endfor %}

</table>
```

python

```python
@app.route('/data')
    def load_entire_df():
        cols = list(df.columns)
        return render_template('dataframe.html',    data=d
to_dict(), cols=cols, nrows=len(df))
```

City of Chicago Employee roster and salaries

Name	Position Title	Department	Employee Annual Salary
AARON, ELVIA J	WATER RATE TAKER	WATER MGMNT	$88968.00
AARON, JEFFERY M	POLICE OFFICER	POLICE	$80778.00
AARON, KARINA	POLICE OFFICER	POLICE	$80778.00
AARON, KIMBERLEI R	CHIEF CONTRACT EXPEDITER	GENERAL SERVICES	$84780.00
ABAD JR, VICENTE M	CIVIL ENGINEER IV	WATER MGMNT	$104736.00
ABARCA, ANABEL	ASST TO THE ALDERMAN	CITY COUNCIL	$70764.00
ABARCA, EMMANUEL	GENERAL LABORER - DSS	STREETS & SAN	$40560.00
ABBATACOLA, ROBERT J	ELECTRICAL MECHANIC	AVIATION	$91520.00
ABBATEMARCO, JAMES J	FIRE ENGINEER	FIRE	$90456.00
ABBATE, TERRY M	POLICE OFFICER	POLICE	$86520.00
ABBOTT, BETTY L	FOSTER GRANDPARENT	FAMILY & SUPPORT	$2756.00
ABBOTT, LYNISE M	CLERK III	POLICE	$43920.00
ABBRUZZESE, WILLIAM J	INVESTIGATOR - IPRA II	IPRA	$72468.00
ABDALLAH, ZAID	POLICE OFFICER	POLICE	$69684.00
ABDELHADI, ABDALMAHD	POLICE OFFICER	POLICE	$80778.00
ABDELLATIF, AREF R	FIREFIGHTER (PER ARBITRATORS AWARD)-PARAMEDIC	FIRE	$98244.00
ABDELMAJEID, AZIZ	POLICE OFFICER	POLICE	$80778.00
ABDOLLAHZADEH, ALI	FIREFIGHTER/PARAMEDIC	FIRE	$87720.00
ABDUL-KARIM, MUHAMMAD A	ENGINEERING TECHNICIAN VI	WATER MGMNT	$106104.00
ABDULLAH, DANIEL N	FIREFIGHTER-EMT	FIRE	$91764.00
ABDULLAH, KEVIN	LIEUTENANT	FIRE	$110370.00

GET Requests
New Feature: Filter by

- Add a new feature that lets us filter by a chosen field
- Use an HTTP GET request to submit parameters.
- Recall GET parameters appear in the URL:

`ip/filter?parameter1=value1&parameter2=value2`

- Use Flask's requests subpackage to do this

http://127.0.0.1:5000/filter?column=Department&value=POLICE

```python
from flask import request

@app.route('/filter', methods=['GET', 'POST'])
def load_subset():
    selected_column = request.args.get('column')
    value = request.args.get('value')
    df_subset = df[df[selected_column] ==   value].reset_index
(drop=True)
    return render_template('dataframe.html', data=df_subset.to_dict(),
cols=cols, nrows=len(df_subset))
```

http://127.0.0.1:5000/filter?column=Department&value=POLICE

```
from flask import request

@app.route('/filter', methods=['GET', 'POST'])
def load_subset():
    selected_column = request.args.get('column')   store column to filter by
    value = request.args.get('value')   store value in that column to filter by
    df_subset = df[df[selected_column] ==   value].reset_index
(drop=True)
    return render_template('dataframe.html', data=df_subset.to_dict(),
cols=cols, nrows=len(df_subset))
```

`http://127.0.0.1:5000/filter?column=Department&value=POLICE`

City of Chicago — Employee roster and salaries

Name	Position Title	Department	Employee Annual Salary
AARON, JEFFERY M	POLICE OFFICER	POLICE	$80778.00
AARON, KARINA	POLICE OFFICER	POLICE	$80778.00
ABBATE, TERRY M	POLICE OFFICER	POLICE	$86520.00
ABBOTT, LYNISE M	CLERK III	POLICE	$43920.00
ABDALLAH, ZAID	POLICE OFFICER	POLICE	$69684.00
ABDELHADI, ABDALMAHD	POLICE OFFICER	POLICE	$80778.00
ABDELMAJEID, AZIZ	POLICE OFFICER	POLICE	$80778.00
ABDULLAH, LAKENYA N	CROSSING GUARD	POLICE	$16692.00
ABEJERO, JASON V	POLICE OFFICER	POLICE	$86520.00
ABRAHAM, NANCY A	POLICE OFFICER	POLICE	$46206.00
ABRAMS, HENRY L	POLICE OFFICER	POLICE	$92316.00
ABRON, FLOYD	POLICE OFFICER	POLICE	$86520.00
ABSTON, KATHY A	POLICE OFFICER	POLICE	$89718.00
ABUDAYEH, ELIAS	POLICE OFFICER	POLICE	$46206.00
ABUZANAT, ABDALLA H	POLICE OFFICER (ASSIGNED AS EVIDENCE TECHNICIAN)	POLICE	$90846.00
ACCARDO, JENNIFER A	POLICE OFFICER	POLICE	$83616.00

Variable Routes
New Feature: Detail Page

```python
@app.route('/detail/<int:item>')
def load_detail_page(item):
    details = df.iloc[item].to_dict()
    return render_template('detail.html', data=details, cols=df.columns, title=config.title, subtitle=config.subtitle)
```

Variable Routes
New Feature: Detail Page

```
@app.route('/detail/<int:item>') < type : variable_name >
def load_detail_page(item):
    details = df.iloc[item].to_dict()
    return render_template('detail.html', data=details, cols=df.
columns, title=config.title, subtitle=config.subtitle)
```

City of Chicago Employee roster and salaries

Details: AARON, KIMBERLEI R

Name	AARON, KIMBERLEI R
Position Title	CHIEF CONTRACT EXPEDITER
Department	GENERAL SERVICES
Employee Annual Salary	$84780.00