CS 133 - Introduction to Computational and Data Science

Instructor: Renzhi Cao
Computer Science Department
Pacific Lutheran University
Spring 2017

Processing data

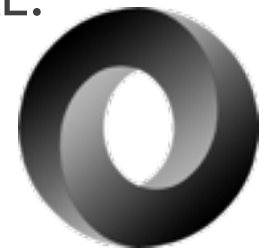
- In the previous class, we learned how to get data.
- Today we are going to continue talk about processing data through APIs

APIs

API = Application programming interface

Allow you to request data in a structured format.

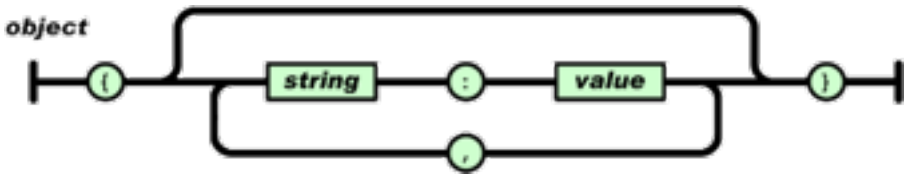
Most popular format is JSON, but if the API provider hates you they will use XML.



JSON

JavaScript Object Notation.

Serialized strings. In other words: Data that follows specific rules in the way of how is portrayed.



Example:

Objects

{ “title” : “Data Science Book”,

“author” : “Joel, Grus”,

“publicationYear” : “2014”,

“topics” : [“data”, “science”, “data science”]

}

Does this look similar to something we have used before?

JSON and Python

JSON Objects look like Python dictionaries!

To parse them, we can use **json** module in python.

```
import json
se = """ {"title": "Data Science Book",
        "author" : "Joel Grus",
        "publicationYear": 2014,
        "topics" : ["data", "science","data science"]
        }"""
des = json.loads(se)
print des
```

(You can download in the course website)

Authenticated vs unauthenticated APIs

There is a huge number of unauthenticated APIs out there.

They are not bad or good.

Just be careful

Finding APIs

http://www.pythonforbeginners.com/api/list-of-python-apis

You will need to get API keys to use most of these. You can find instructions on how to do it by googling “API key ____”

Twitter API

Using Anaconda's command prompt, type the words:

pip install twython

Get credentials.

Page 117 has instructions.

I am assuming at this point that you have had created your account.

Twitter API

Using Search API

- Example 1

Make sure to use .encode('utf-8').

Some tweets have Unicode values

Using the twitter.search API only returns a few handful of most recent results. To be able to extract more we will need to use the Streaming API

- Example 2

Mid-term exam review

- Mid-term exam: go through the study guide
- Codingbat practice

Exercise

Book: 117-120

Coding bat website: http://codingbat.com/python

Example practices and questions for mid-term:

Implement a function to calculate the minimum value in a list

```
def mymin(l):

    return “”;
```