



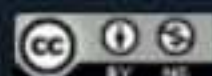
Unit 05.01

## Databases and Visualization: Geocoding

### USING DATABASES WITH PYTHON

Charles Severance Ph.D.

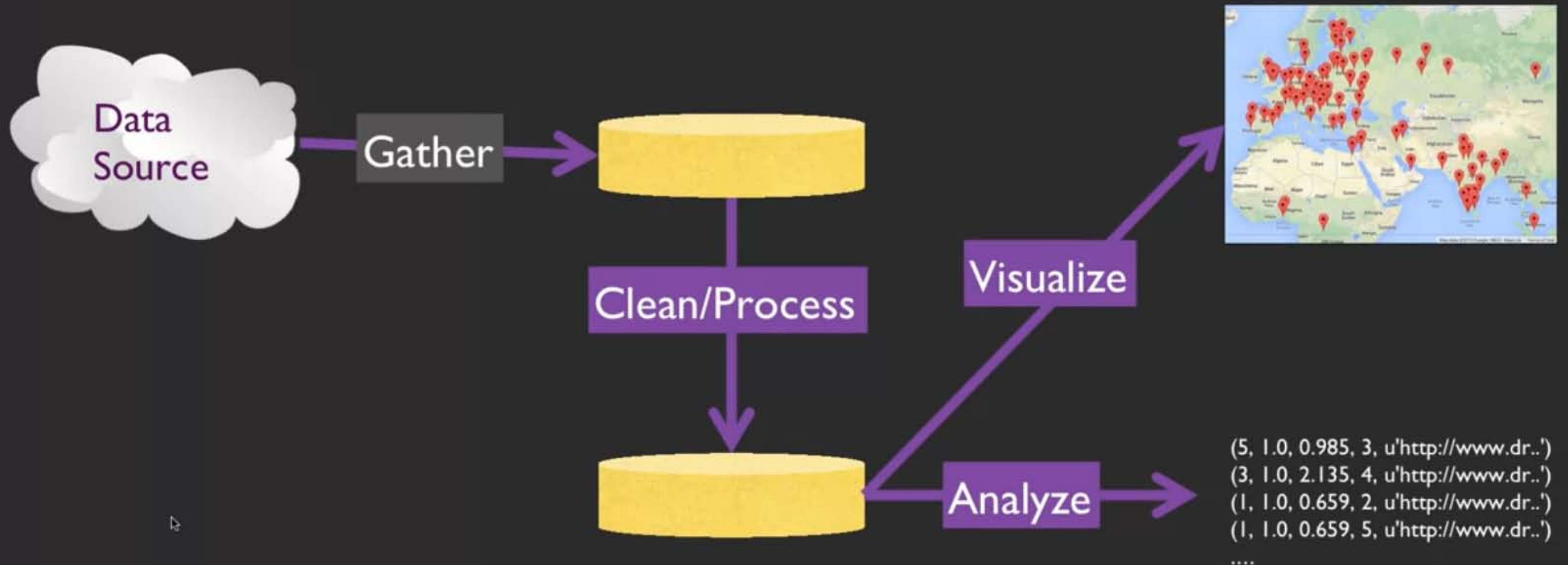
Clinical Associate Professor of Information  
School of Information



© 2015 Charles Severance and The Regents of the University of Michigan  
Except where otherwise noted, this work is licensed under  
<http://creativecommons.org/licenses/by-nc/3.0/>



# Multi-Step Data Analysis



So what we're really going to work on is  
we're going to combine the act of working



# Many Data Mining Technologies

- <https://hadoop.apache.org/>
- <http://spark.apache.org/>
- <https://aws.amazon.com/redshift/>
- <http://community.pentaho.com/>

Amazon has a whole data mining operation called Redshift.

# "Personal Data Mining"

Our goal is to make you better programmers – not to make you data mining experts

And it is not to say that once you're done with this you're a data mining expert,



# GeoData

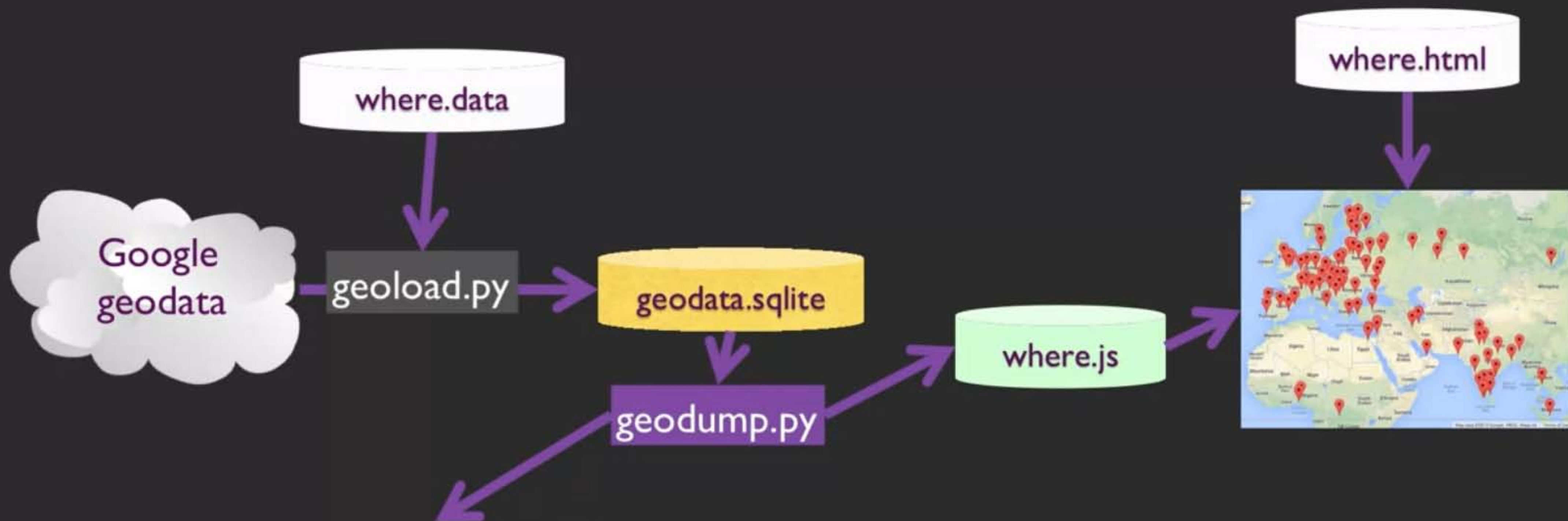
- Makes a Google Map from user entered data
- Uses the Google Geodata API
- Caches data in a database to avoid rate limiting and allow restarting
- Visualized in a browser using the Google Maps API



<http://www.pythonlearn.com/code/geodata.zip>



## 05.01 Databases and Visualization: Geocoding



Northeastern University, ... Boston, MA 02115, USA 42.3396998 -71.08975

Bradley University, 1501 ... Peoria, IL 61625, USA 40.6963857 -89.6160811

...

Technion, Viazman 87, Kesalsaba, 32000, Israel 32.7775 35.0216667

Monash University Clayton ... VIC 3800, Australia -37.9152113 145.134682

Kokshetau, Kazakhstan 53.2833333 69.3833333

...

12 records written to where.js

Open where.html to view the data in a browser

<http://www.pythonlearn.com/code/geodata.zip>