



Python Power: Leveraging ArcGIS Notebooks for Efficient Data ETL

Jin Yao, 4/18/2024

References

This slide deck and the demo ArcGIS notebook can be downloaded:

https://github.com/JinMcBurney/MAGIC_2024_Symposium

Census Bureau

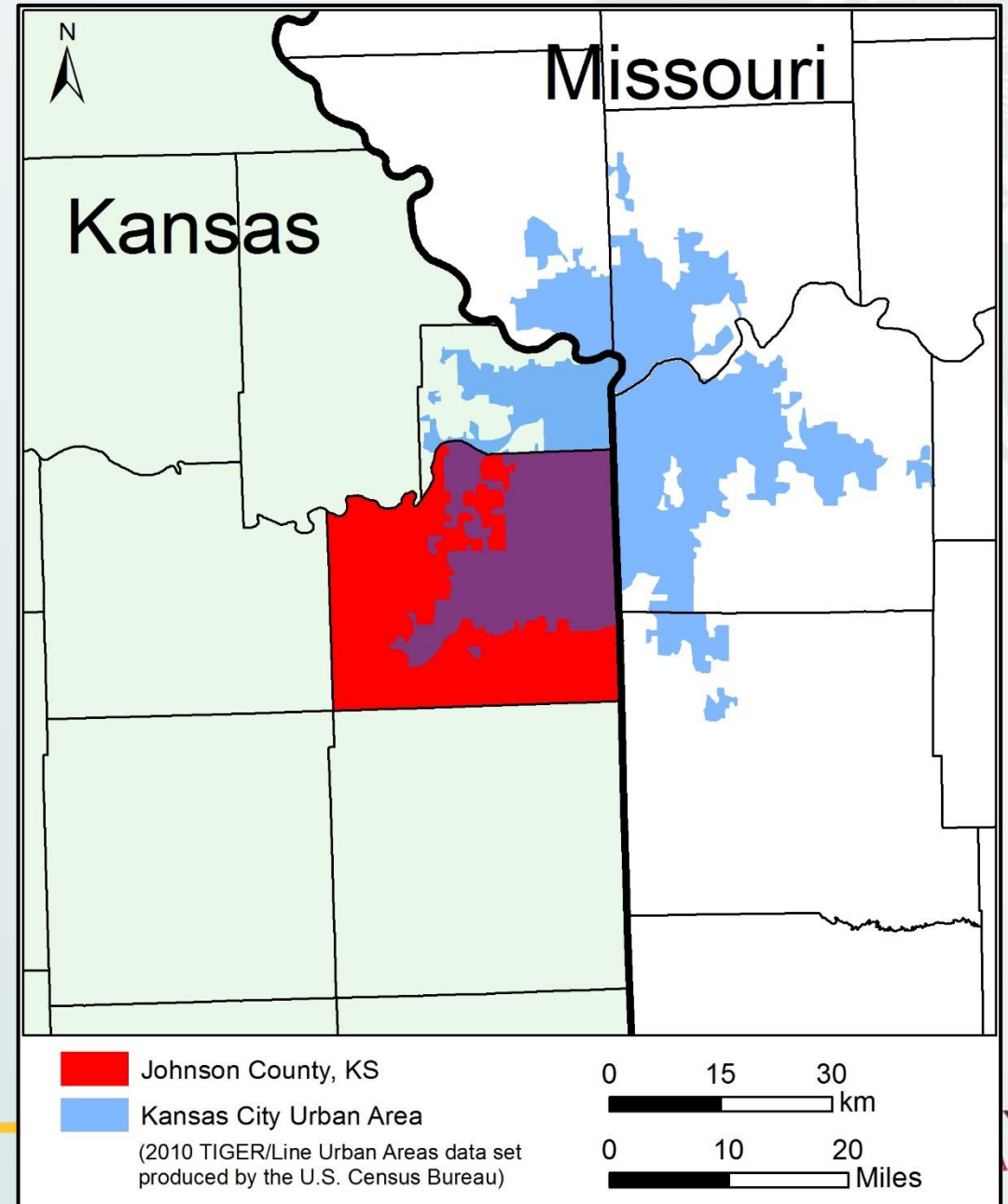
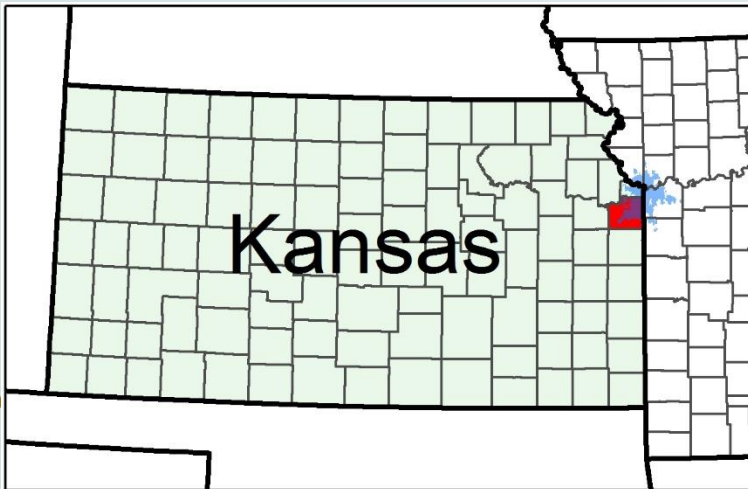
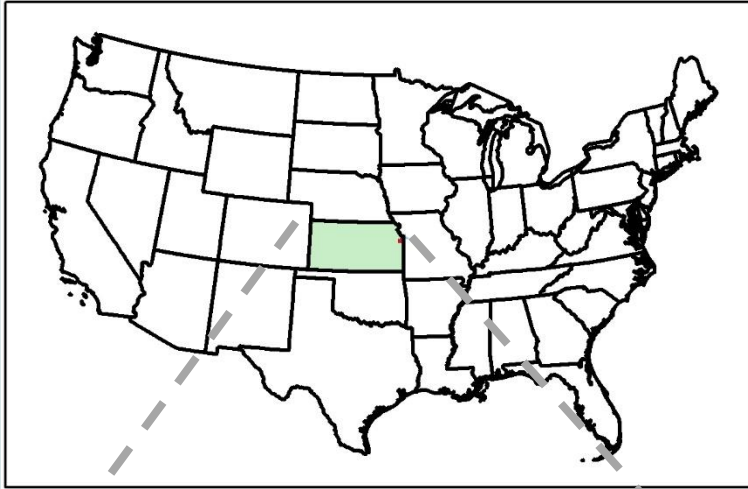
- Data website: <https://data.census.gov/>
- Census Bureau Video Tutorials: <https://www.census.gov/data/what-is-data-census-gov/guidance-for-data-users/video-tutorials.html>
 - How to use Census API
- Data at your door: Census API Decoded: <https://www.census.gov/data/academy/webinars/2022/census-api-decoded.html>
- APIs of available datasets: <https://www.census.gov/data/developers/datasets.html>

Python Tutorial: Using the Census API: <https://youtu.be/l47HptzM7ao>

Do you use open-source software?

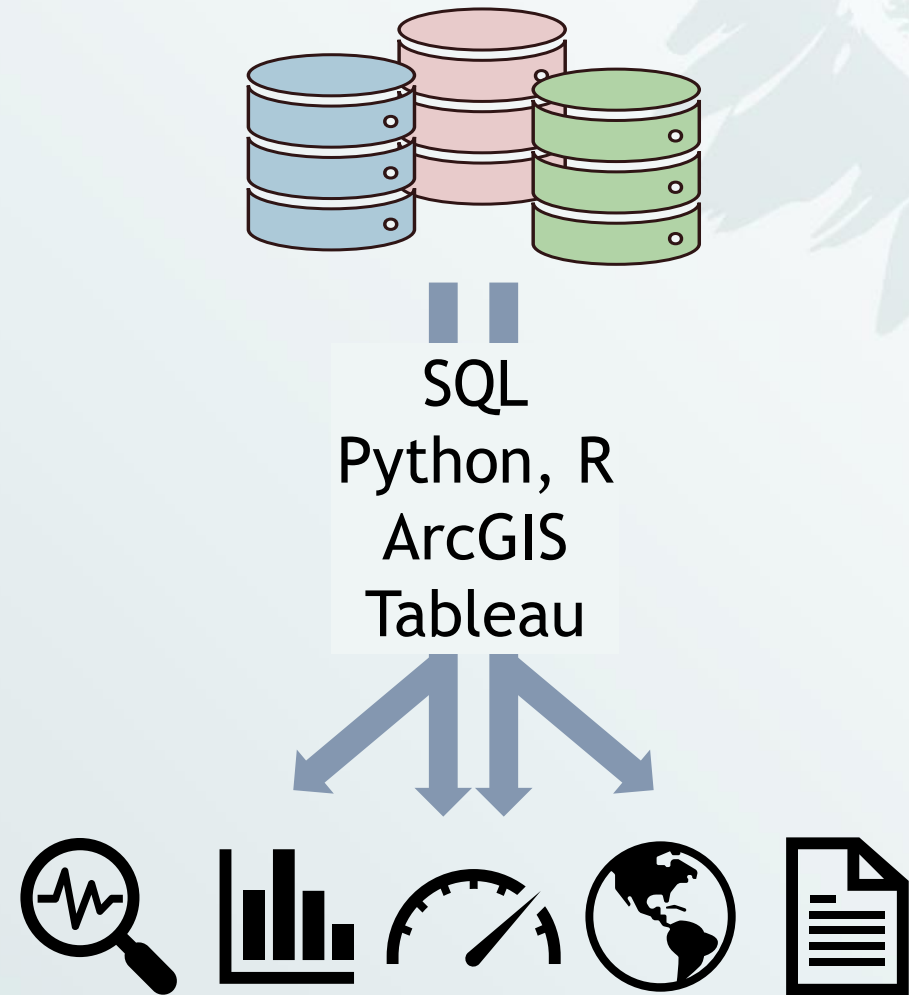
- Python, R, ...?
- IDE?
- Versions of libraries or packages?

Johnson County, KS



About My Team and Me

- AIMS
Johnson County Government, KS
- Me: Senior Data Analyst



Outline

1. ArcGIS Pro's Python Environment
2. Download Data Using Census API
3. Export Data to Microsoft SQL Server



1. ArcGIS Pro's Python Environment

(arctgispro-py3) C:\Program Files\ArcGIS\Pro\bin\Python\envs\arctgispro-py3>pip list

Package	Version
anyio	3.5.0
appdirs	1.4.4
arctgis	2.1.0.2
argon2-cffi	21.3.0
argon2-cffi-bindings	21.2.0
asttokens	2.0.5
attrs	21.4.0
azure-core	1.12.0
azure-storage-blob	12.8.0
Babel	2.11.0
backcall	0.2.0
beautifulsoup4	4.11.1
black	0.0.0
bleach	4.1.0
blinker	1.4
Bottleneck	1.3.5
brotlipy	0.7.0
cachetools	4.2.2
certifi	2022.9.24
cffi	1.15.1
cftime	1.6.2

- ←
- New
- Open
- Info
- Save Project
- Save Project As

- Portals
- Licensing
- Options
- Package Manager**

- Add-In Manager

- Help
- About
- Learning Resources

- Exit

Package Manager

Manage environments and packages for Python, R, and system libraries.

⚠ Cannot modify the default Python environment (arcgispro-py3). Clone then activate a new environment first. [Learn more about cloning](#)

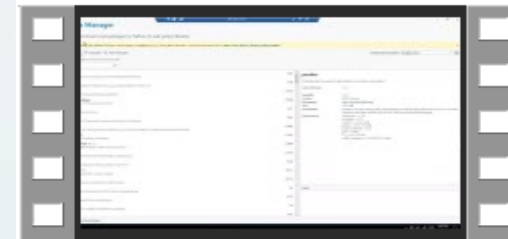
📦 Installed (225)
🔄 Updates
➕ Add Packages

The packages installed in the active environment.

anyio High level asynchronous concurrency and networking framework	3.5.0	anyio High level asyn Installed: License: Homepage: Size: Dependency:
appdirs A small Python module for determining appropriate platform-specific dirs	1.4.4	
argon2-cffi ⓘ The secure Argon2 password hashing algorithm	21.3.0	
argon2-cffi-bindings Low-level Python CFFI Bindings for Argon2	21.2.0	
arrow-cpp ⓘ C++ libraries for Apache Arrow	1.0.1	
asttokens Annotate Python AST trees with source text and token information	2.0.5	
attrs ⓘ ⓘ Bring back the joy of writing classes by relieving you from the drudgery of implementing object protocols	21.4.0	
azure-core ⓘ ⓘ Microsoft Azure Core Library for Python	1.12.0	
azure-storage-blob ⓘ ⓘ Microsoft Azure Blob Storage Client Library for Python	12.8.0	

Start a New Notebook

Start in ArcGIS Pro



ArcGISPro_open_new_notebook.mp4

Start at Python
Command Prompt



Python_prompt_open_new_notebook.mp4

2. Download Data Using Census API

Example: download Average Household Size

- from 2022 ACS 5 year data
- at all census tracts in JoCo

<https://data.census.gov/>

B25010 | Average Household Size of Occupied Housing Units by Tenure

American Community Survey | Universe: Occupied housing units | 2022: ACS 5-Year Estimates De... | Notes | Geos ³

Census Tract 500; Johnson County; Kansas		
Label	Estimate	Margin of Error
▼ Average household size --		
▼ Total:	2.43	±0.16
Owner occupied	2.50	±0.18
Renter occupied	2.16	±0.23

Census Data API, 5 components

https://api.census.gov

Host

/data/2022/acs/acs5

Dataset

?get=GEO_ID,B25010_001E,B25010_002E,B25010_003E

Variables

&for=tract:*

Geo unit

&in=state:20;county:091

Geo extent

Census Data API, result of API call

```
api.census.gov/data/2022/acs/acs5?get=GEO_ID,B25010_001E,B25010_002E,B25010_003E&for=tract:*&in=state:20;county:091

MyRC - secure RC - public AIMS JoCo Maps Gmail

Pretty-print 

[[ "GEO_ID", "B25010_001E", "B25010_002E", "B25010_003E", "state", "county", "tract" ],
[ "1400000US20091050000", "2.43", "2.50", "2.16", "20", "091", "050000" ],
[ "1400000US20091050100", "2.11", "2.10", "2.13", "20", "091", "050100" ],
[ "1400000US20091050200", "2.06", "2.09", "1.97", "20", "091", "050200" ],
[ "1400000US20091050301", "1.62", "2.00", "1.52", "20", "091", "050301" ],
[ "1400000US20091050302", "1.66", "1.90", "1.42", "20", "091", "050302" ],
[ "1400000US20091050400", "2.04", "2.18", "1.88", "20", "091", "050400" ],
[ "1400000US20091050500", "2.43", "2.39", "2.63", "20", "091", "050500" ],
[ "1400000US20091050600", "2.46", "2.42", "2.77", "20", "091", "050600" ],
[ "1400000US20091050700", "2.41", "2.43", "2.29", "20", "091", "050700" ],
[ "1400000US20091050800", "3.03", "3.05", "-666666666.00", "20", "091", "050800" ],
[ "1400000US20091050900", "2.25", "2.35", "1.87", "20", "091", "050900" ],
```

To find Census API: Datasets in /data/2022/acs/acs5 and its descendants

- <https://api.census.gov/data/2022/acs/acs5.html>

Title	Description	Vintage	Dataset Name	Dataset Type	Geography List	Variable List	Group List	SortList	Examples	Developer Documentation	API Base URL
American Community Survey: 5-Year Estimates: Detailed Tables 5-Year	The American Community Survey (ACS) provides information that covers a wide range of demographic, economic, and housing characteristics for the United States (in 5-year increments) and for congressional districts, state, and county. The data is available for 64,000 geographic areas.	2022	acs acs5	Aggregate	geographies	variables	groups	sorts	examples	documentation	http://api.census.gov/data/2022/acs/acs5

API Key or not

- Any user may query small quantities of data with minimal restrictions
 - up to 50 variables in a single query, and
 - up to 500 queries per IP address per day
- However, more than 500 queries per IP address per day requires that you register for an API key.
- Source:
https://www.census.gov/content/dam/Census/library/publications/2020/acs/acs_api_han_dbook_2020_ch02.pdf

Python code to request data

See Demo Notebook



3. Export Data to Microsoft SQL Server

See Demo Notebook

Outline

1. ArcGIS Pro's Python Environment
2. Download Data Using Census API
3. Export Data to Microsoft SQL Server



References

This slide deck and the demo ArcGIS notebook can be downloaded:

https://github.com/JinMcBurney/MAGIC_2024_Symposium

Census Bureau

- Data website: <https://data.census.gov/>
- Census Bureau Video Tutorials: <https://www.census.gov/data/what-is-data-census-gov/guidance-for-data-users/video-tutorials.html>
 - How to use Census API
- Data at your door: Census API Decoded: <https://www.census.gov/data/academy/webinars/2022/census-api-decoded.html>
- APIs of available datasets: <https://www.census.gov/data/developers/datasets.html>

Python Tutorial: Using the Census API: <https://youtu.be/l47HptzM7ao>

Contact Us

AIMS website: <https://aims.jocogov.org/>

About AIMS: <https://arcg.is/1mKivb0>

Jin Yao: jin.yao@jocogov.org

