Getting Started with APIs in R

Reina Chano Murray

R-Ladies Baltimore + Irvine

April 11, 2023

https://bit.ly/3MBht0O

What we'll use today

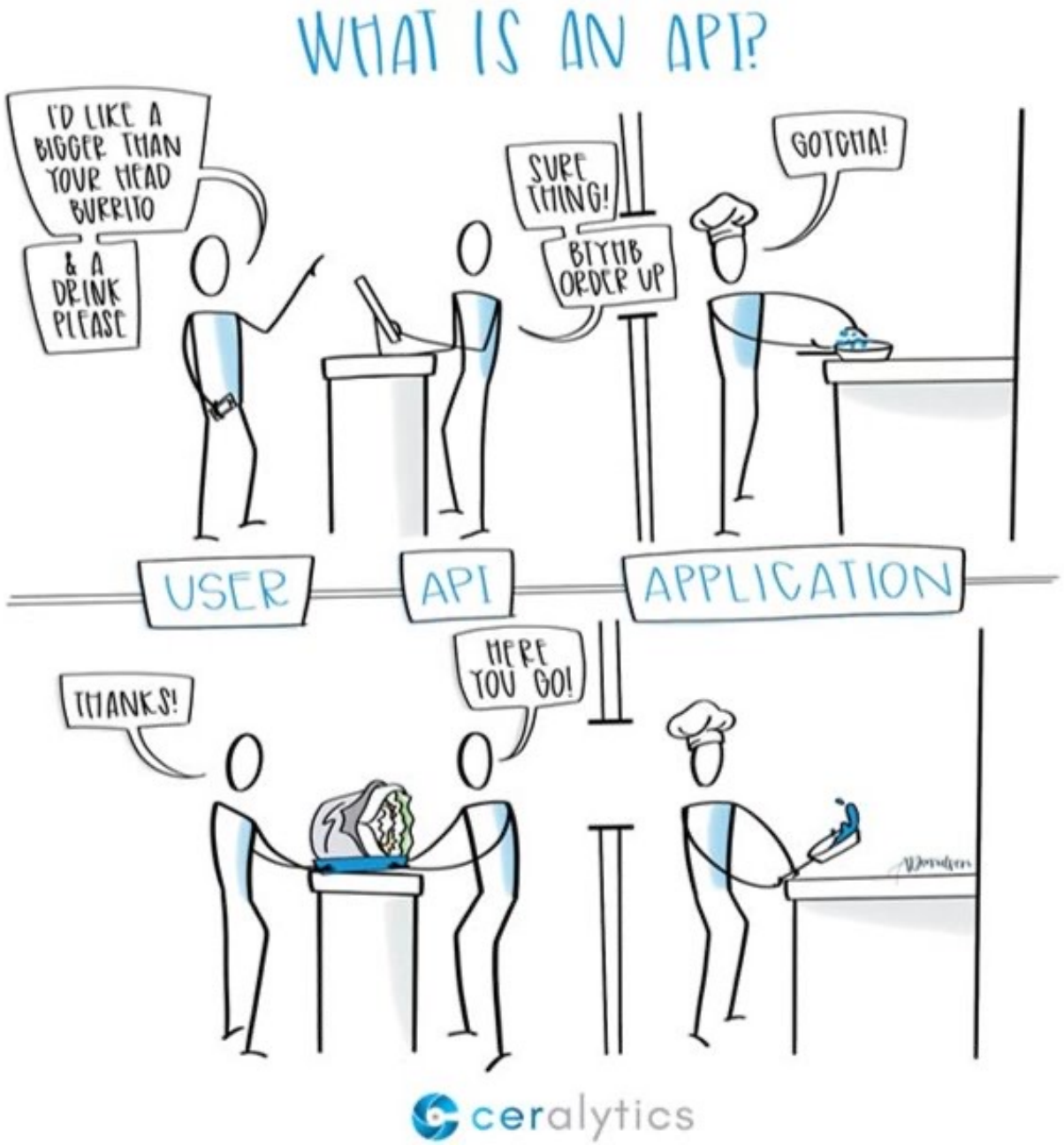
- Rstudio (Posit)
- Packages:
  - `httr`
  - `jsonlite`
  - `censusapi`
- US Census Bureau API Key

Agenda

- What is an API?
- Why use APIs?
- Getting started (the general workflow)
- Explore the US Census API (creating queries in a web browser)
- Demos:
  - Using `httr` and `jsonlite`
  - Using `censusapi`

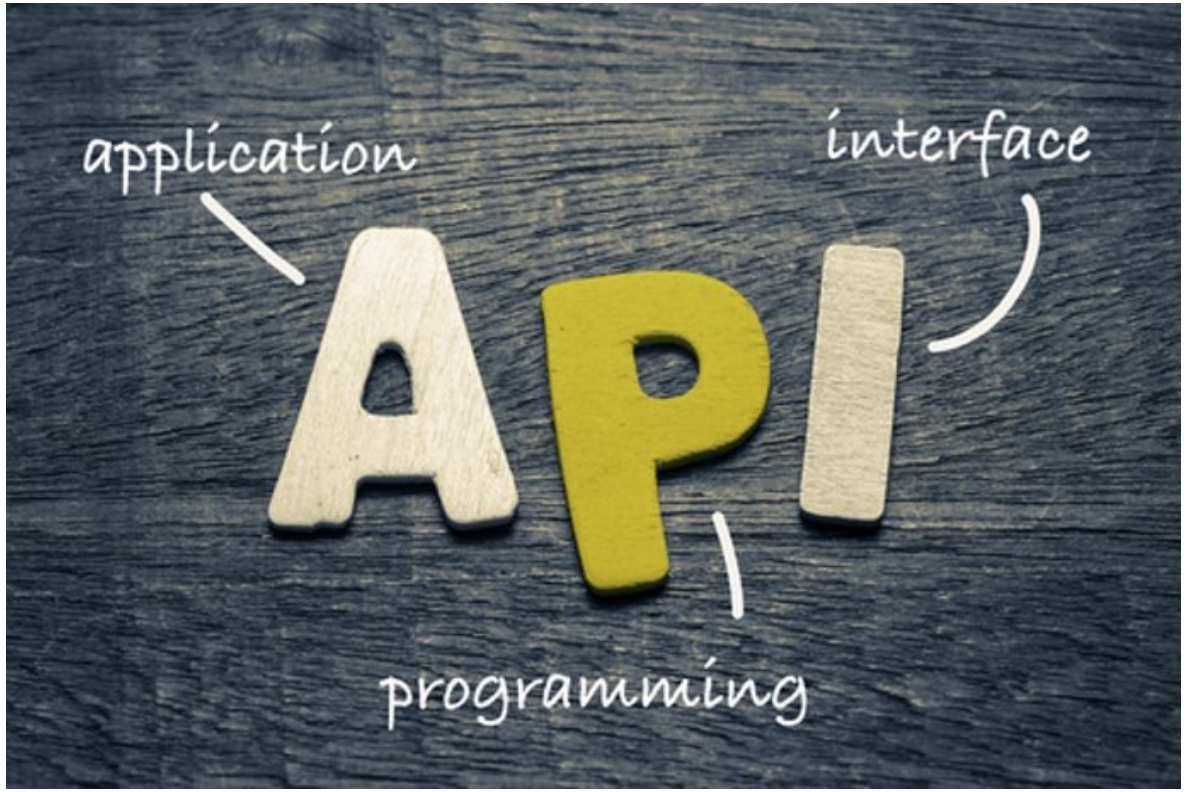
What is an API?


application
interface
API



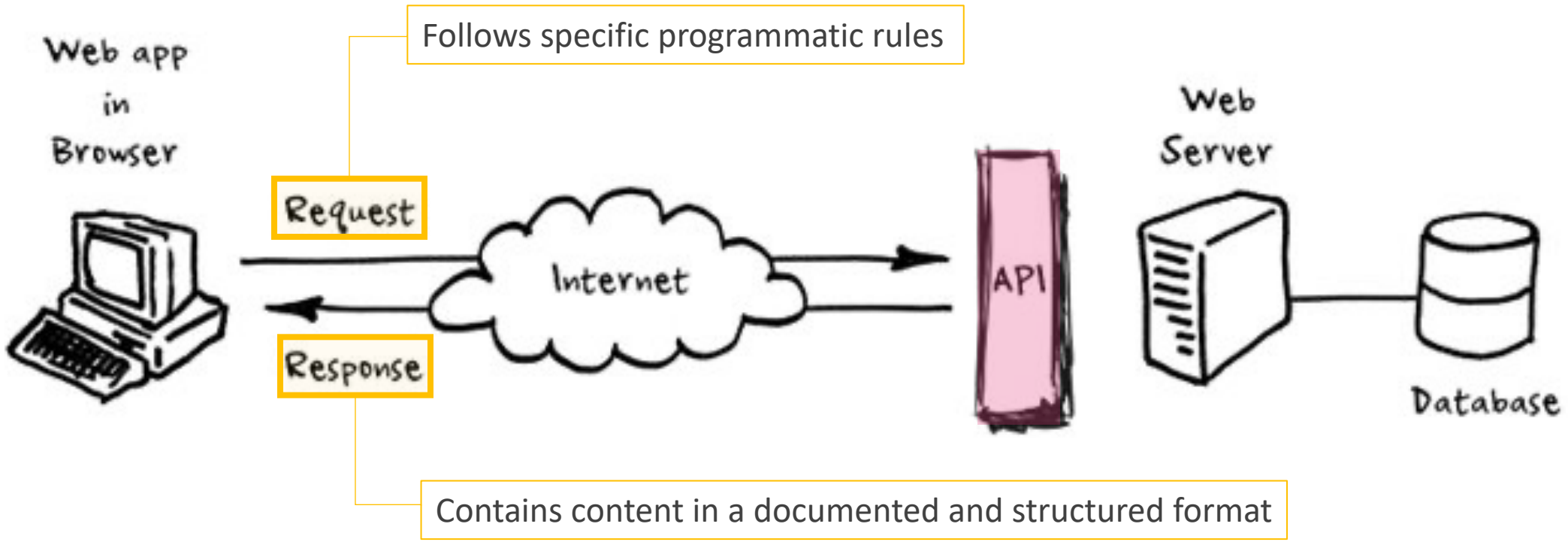
Ordering a Meal

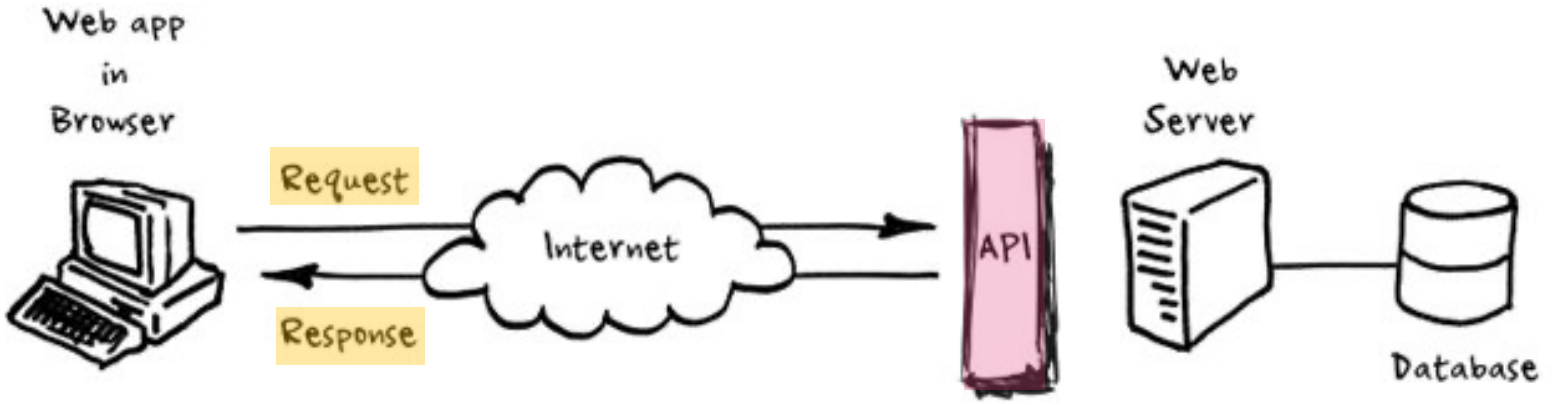
- Customer makes an order
- Server takes down order, processes it, and passes it to cook
- Cook acknowledges order
- Cook makes the order, gives it to the server
- Server gives the customer their order



APIs

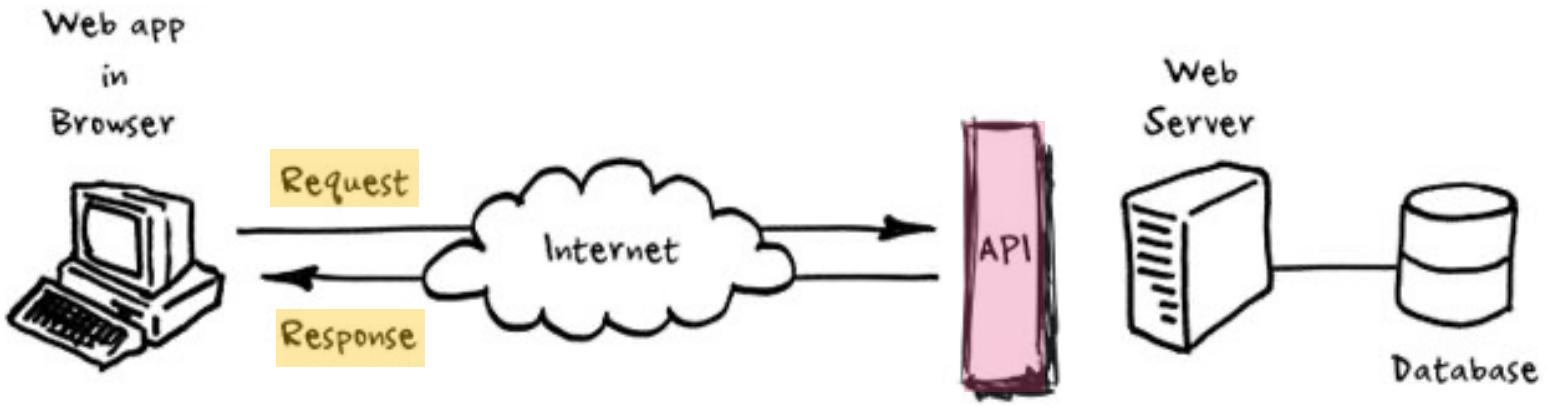
- User (client) makes a **request**
- API takes down request, processes it, and passes it to the application
- Application acknowledges request
- Application processes request and signals API
- API returns a **response** to user


Follows specific programmatic rules
Web app in Browser
Request
Response
Internet
API
Web Server
Database
Contains content in a documented and structured format



APIs

- Is **NOT** the database or the server – it is the code that governs the access point(s) for the server
- APIs cover a broad category that includes all interfaces that facilitate communication between computer applications
- We're talking about **Web APIs** – APIs that expose an application's data and functionality over the internet, allowing two computers (the client and the server) to interact with each other to request and provide data.

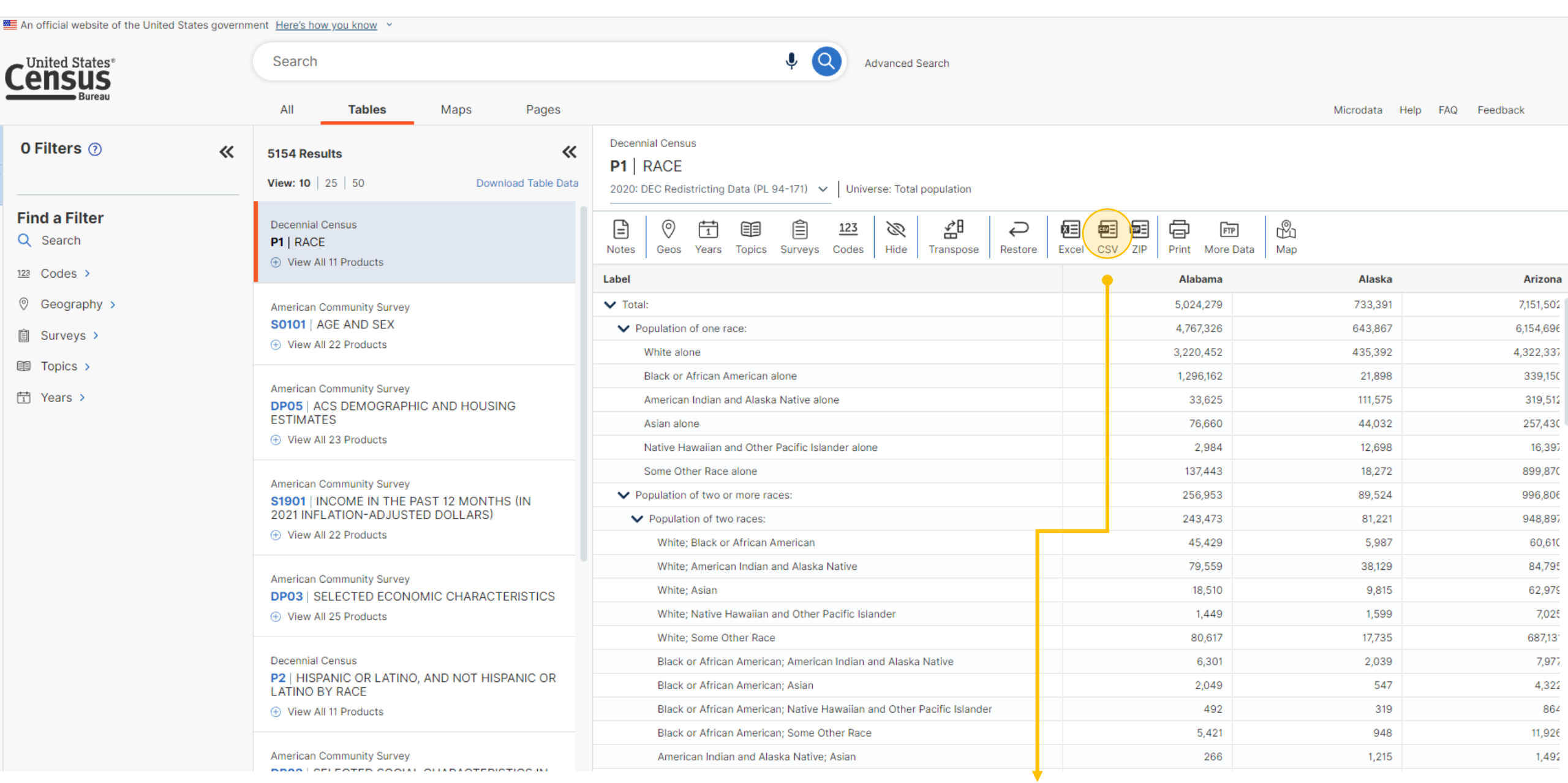


4 main types of web APIs

1. Open / public APIs – no restrictions to access*
2. Partner APIs – requires specific rights and/or licenses
3. Internal APIs – designed for internal use within an organization
4. Composite APIs – combines different data and service APIs

* May need to create an account or request an API key

Why use an API?




```
# read in csv
acsDat1 <- read.csv("acsst5y2017_20210119.csv", header=T, na.strings=c("", "NA"))
```

Why Use an API

Get data directly from the source

- Data pulled from an API is usually formatted to be usable with some cleaning
- Get access to most recent, up-to-date data

Integrate and automate your data collection workflow

- APIs can let you forgo the process of manually downloading datasets from a website, managing the data in a file management system, and loading it into the software you are using for analysis.

Getting Started

The general workflow

1. Explore the documentation for the API
2. Make your request
3. Review response

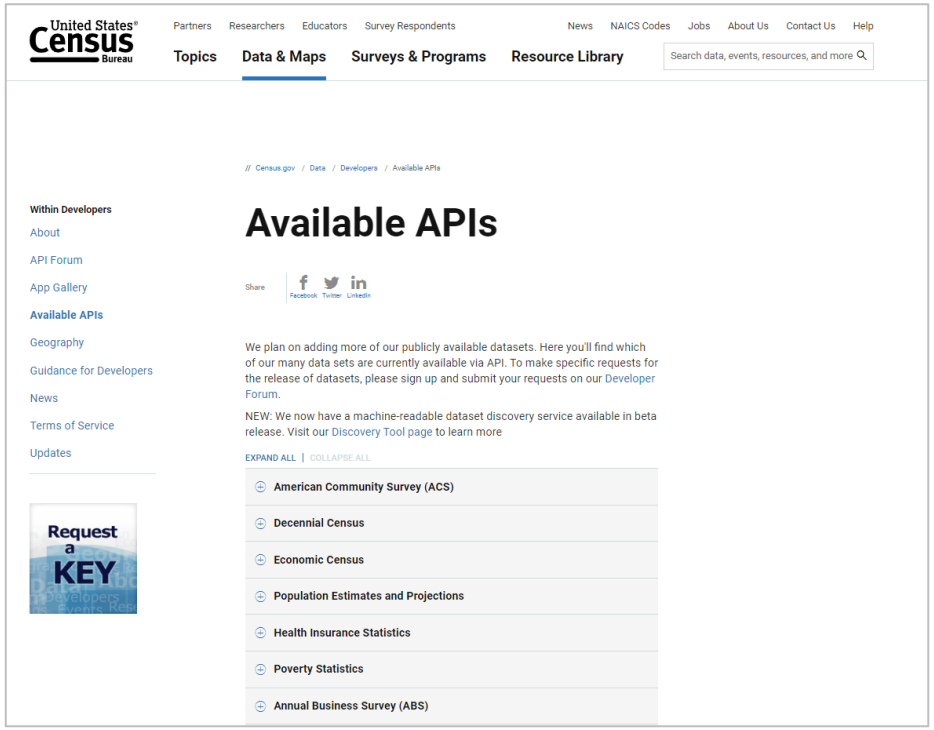
1

Explore the documentation for the API

- Do they supply an API that does what you're looking for?
- Review the process for getting access to the API:
  - Do you need to apply / create an account?
  - Applying for an API key – unique identifier to authenticate your particular requests
  - Query limits (if any)

2 Make your request

3 Review response



https://www.census.gov/data/developers/data-sets.html

1

Explore the documentation for the API

- Do they supply an API that does what you're looking for?
- Review the process for getting access to the API:
  - Do you need to apply / create an account?
  - Applying for an API key – unique identifier to authenticate your particular requests
  - Query limits (if any)

2

Make your request

- Identify the URI/URL (uniform resource identifier/locator)
- Prepare your query
- Send a `GET request` to the API

3

Review response

- Parse through and clean response
- Convert response to a dataframe

Explore the US Census API

US Census Data API

- https://www.census.gov/data/developers/guidance.html

- Available data
  - American Community Survey (ACS), Economic Indicators Time Series, Decennial Census, Economic Census, County Business Patterns and Nonemployer Statistics, Population Estimates and Projections, International Trade, etc…

- Query small quantities of data (up to 50 variables in a single query, and up to 500 queries per IP address per day)
  - To make more than 500 queries per IP address per day, request an API key

Creating Queries in a Web Browser

https://api.census.gov/data/2018/pep/charagegroups?get=POP,GEONAME,DATE_DESC&DATE_CODE=11&RACE=10&for=county:*&in=state:24

Creating Queries in a Web Browser

https://api.census.gov/data/2018/pep/charagegroups?get=POP,GEONAME,DATE_DESC&DATE_CODE=11&RACE=10&for=county:*&in=state:24

1. Start with the host name

Creating Queries in a Web Browser

https://api.census.gov/data/2018/pep/charagegroups?get=POP,GEONAME,DATE_DESC&DATE_CODE=11&RACE=10&for=county:*&in=state:24

1. host name

2. Add dataset name
(list available at https://api.census.gov/data.html)

Creating Queries in a Web Browser

https://api.census.gov/data/2018/pep/charagegroups?get=POP,GEONAME,DATE_DESC&DATE_CODE=11&RACE=10&for=county:*&in=state:24

1. host name

2. Add dataset name
(list available at https://api.census.gov/data.html)

3. Add variable requests using `?get=`

Creating Queries in a Web Browser

https://api.census.gov/data/2018/pep/charagegroups?get=POP,GEONAME,DATE_DESC&DATE_CODE=11&RACE=10&for=county:*&in=state:24

1. host name

2. Add dataset name
(list available at https://api.census.gov/data.html)

3. Add variable requests using `?get=`

4. Filter variables using `&for=`

Creating Queries in a Web Browser

https://api.census.gov/data/2018/pep/charagegroups?get=POP,GEONAME,DATE_DESC&DATE_CODE=11&RACE=10&for=county:*&in=state:24

1. host name

2. Add dataset name
(list available at https://api.census.gov/data.html)

3. Add variable requests using `?get=`

4. Filter variables using `&for=`

5. If needed, add your key using `&key=`

Click here to see RESULT

Creating Queries in a Web Browser

https://api.census.gov/data/2018/pep/charagegroups?get=POP,GEONAME,DATE_DESC&DATE_CODE=11&RACE=10&for=county:*&in=state:24

1. host name

2. Add dataset name
(list available at https://api.census.gov/data.html)

3. Add variable requests using `?get=`

4. Filter variables using `&for=`

5. If needed, add your key using `&key=`

- Making queries in your browser is a good way to practice and get comfortable with the variables and filters available in the Census Data API
- Works best on Chrome or Firefox
  - Use an extension like JSON Formatter, JSON Viewer, or JSONView to make JSON easier to read in your browser
- In web browser, right-click and select “Save As” to save results as a csv
- Sample queries:
https://www.census.gov/data/developers/guidance/api-user-guide.Example_API_Queries.html

Demo 1

Using `httr` and `jsonlite`

Demo 2

Using `censusapi`

censusapi

- Link
- R wrapper for US Census Data API.
- Covers more datasets than tidycensus
  (another R package for US Census data)
- Returns data frames of census data AND metadata

Other Demos

- rjsonio demo:
  - RJSONIO – a package that converts data to and from JSON format
  - Similar to using `httr` and `jsonlite`

- tidycensus demo:
  - Tidycensus - an integrated package that allows users to access decennial census and ACS data as tidyverse-ready data frames.

  - Offers access to less datasets than `censusapi`:

    The main intent of the tidycensus package is to return population characteristics of the United States in tidy format allowing for integration with simple feature geometries. Its intent is not, and has never been, to wrap the universe of APIs and datasets available from the US Census Bureau.

  - Walkthrough

Thank you!

Reina Chano Murray

R-Ladies Baltimore + Irvine

April 11, 2023