Show me the data!

Week16: Import Data API

Big Data & Social Analysis R

Instructors: Chung-pei Pien

ZU1942001/266868001/Z23937001/ZM1941001

International College of INNOVATION
National Chengchi University
國立政治大學創新國際學院

CONTENTS

1 Import Data: API

2 Final Presentation

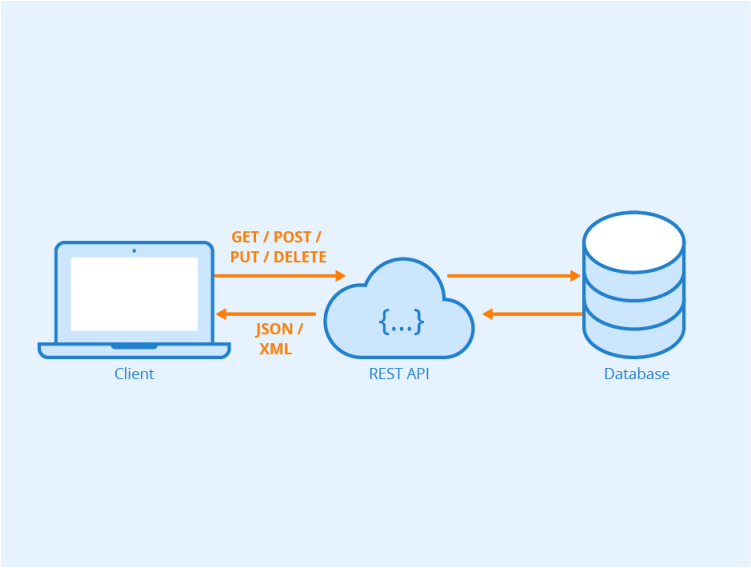
Import Data: API

01 Import Data: API

Introduction

There is rare opportunities to get csv or xlsx data.

Many data providers to avoid web scraping allow users to request Application Programming Interface (API) to get raw data.



01 Import Data: API

World Bank API

Week 16: June 02

Import Data: API

- World Bank Indicators API
- US Census Bureau Economic
- Week16 Practice
- tx income

Let's go to World Bank API

Data

About the Indicators API Documentation

← Developer Information

About the Indicators API

Note: version 2 of the Indicators API was released several years ago, and version 1 was officially retired in November , 2018. Version 1 was discontinued entirely on June 19, 2020 (see link). If you access the older endpoints you will receive a "Resource not found" error. Use the information here to update your code.

The World Bank Indicators API provides access to nearly 16,000 time series indicators. Most of these indicators are available online through tools such as Databank and the Open Data website. The API provides programmatic access to this same data. Many data series date back over 50 years, and can be used to create interesting applications.

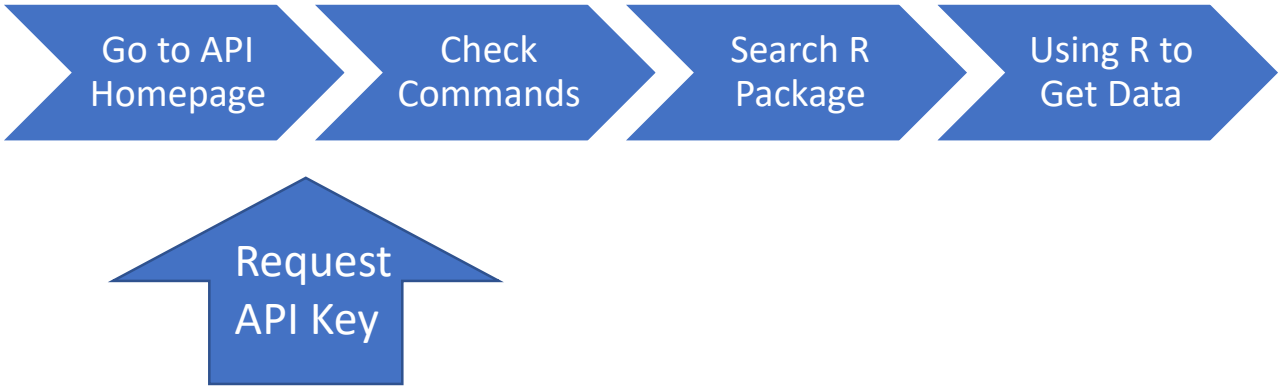
The Indicators API provides access to over 45 databases, including:

- World Development Indicators
- International Debt Statistcs
- Doing Business
- Human Capital Index
- Subnational Poverty
- And many more

01 Import Data: API

World Bank API

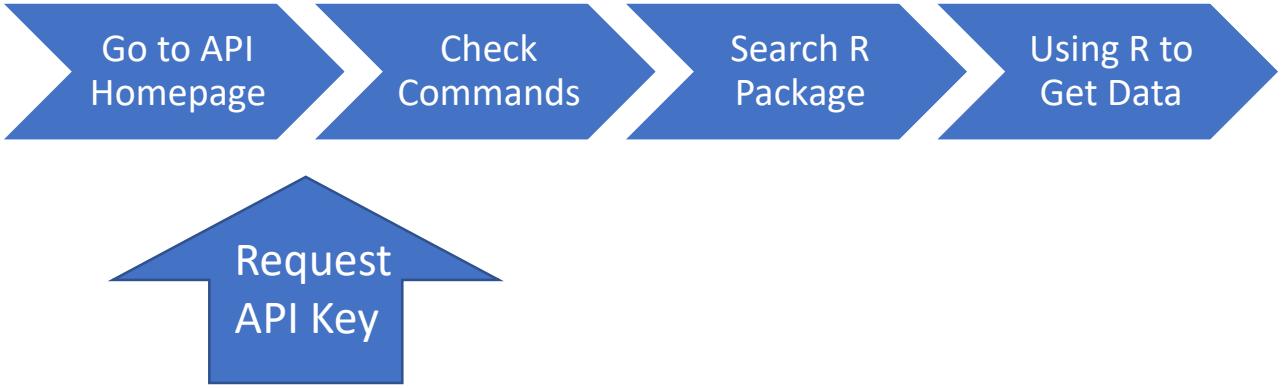
Let's go to World Bank API



01 Import Data: API

World Bank API

Let's go to World Bank API



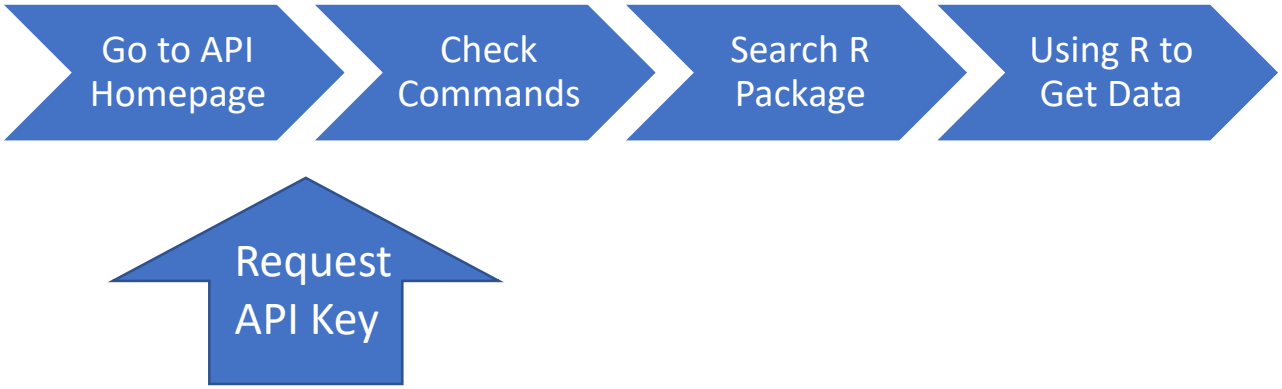
API Access / Authentication

API keys and other authentication methods are no longer necessary to access the API.

01 Import Data: API

World Bank API

Let's go to World Bank API



Date and Date-Range: Date-range by year, month or quarter that scopes the result-set.

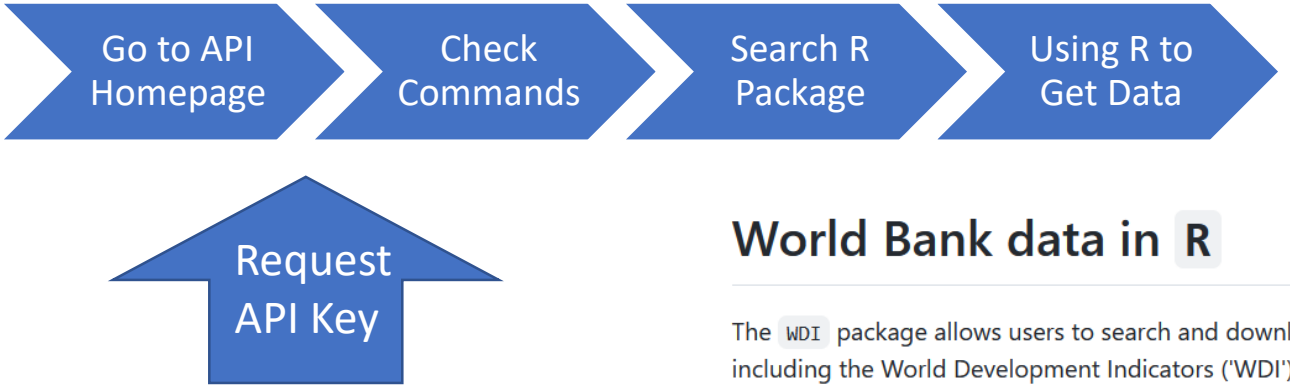
Examples:

- http://api.worldbank.org/v2/country/all/indicator/SP.POP.TOTL?date=2000
- http://api.worldbank.org/v2/country/chn;bra/indicator/DPANUSSPB?date=2012M01

01 Import Data: API

World Bank API

Let's go to World Bank API



01 Import Data: API

World Bank API

https://github.com/vincentarelbundock/WDI

`WDIsearch` uses grep and ignores cases, so you can also use regular expressions. For instance, if you are looking for GDP per capita in constant dollars:

```
WDIsearch('gdp.*capita.*constant')
     indicator           name
[1,] "GDPPCKD"           "GDP per Capita, constant US$, millions"
[2,] "NY.GDP.PCAP.KD"    "GDP per capita (constant 2000 US$)"
[3,] "NY.GDP.PCAP.KN"    "GDP per capita (constant LCU)"
[4,] "NY.GDP.PCAP.PP.KD" "GDP per capita, PPP (constant 2005 international $)"
```

Download and use the data

Download a series you like for the countries you like:

```
dat = WDI(indicator='NY.GDP.PCAP.KD', country=c('MX','CA','US'), start=1960, end=2012)
```

Look at the data:

```
install.packages("WDI")
library(WDI)

#search key words

ws_gini <- WDIsearch("gini")
```

01 Import Data: API

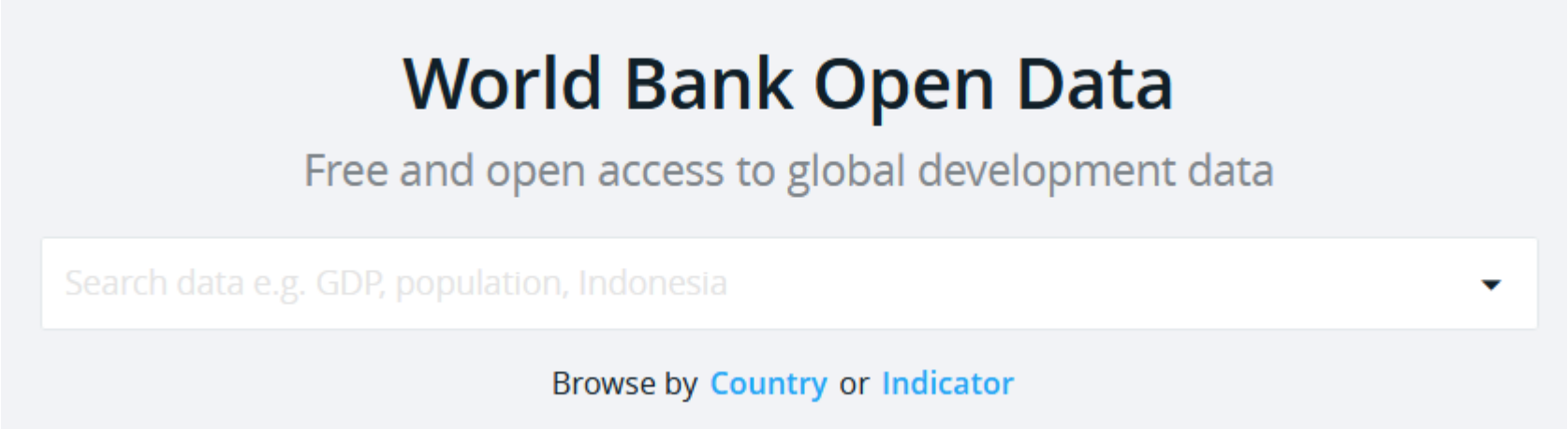
World Bank API

ws_gini <- WDIsearch("gini")

Which indictor we want?

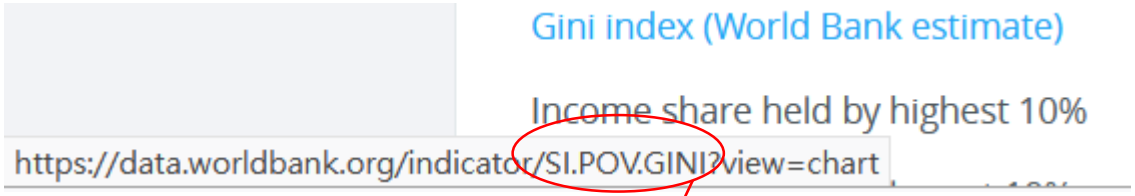
	indicator	name
1	3.0.Gini	Gini Coefficient
2	3.0.Gini_nozero	Gini Coefficient (No Zero Income)
3	3.1.Gini	Gini, Rural
4	3.2.Gini	Gini, Urban
5	per_allsp_gini_rur	Gini inequality index reduction (%) - All Social Protection an...
6	per_allsp_gini_tot	Gini inequality index reduction (%) - All Social Protection an...
7	per_allsp_gini_urb	Gini inequality index reduction (%) - All Social Protection an...
8	per_lm_ac_gini_rur	Gini inequality index reduction (%) - Active Labor Market -r...

https://data.worldbank.org/



01 Import Data: API

World Bank API



Get the name from the link

```
gini <- WDI(indicator='SI.POV.GINI', start=1960, end=2012)
```

	iso2c	country	SI.POV.GINI Gini index (World Bank estimate)	year
1	ZH	Africa Eastern and Southern	*NA*	2012
2	ZH	Africa Eastern and Southern	*NA*	2011
3	ZH	Africa Eastern and Southern	*NA*	2010
4	ZH	Africa Eastern and Southern	*NA*	2009
5	ZH	Africa Eastern and Southern	*NA*	2008
6	ZH	Africa Eastern and Southern	*NA*	2007
7	ZH	Africa Eastern and Southern	*NA*	2006
8	ZH	Africa Eastern and Southern	*NA*	2005
9	ZH	Africa Eastern and Southern	*NA*	2004
10	ZH	Africa Eastern and Southern	*NA*	2003
11	ZH	Africa Eastern and Southern	*NA*	2002

01 Import Data: API

World Bank API

Specific countries' data

```
gini <- WDI(indicator = 'SI.POV.GINI', start = 1960, end =2012)

gini_country <- WDI(indicator = 'SI.POV.GINI', country = c('TR', 'TH'),
      start=1960, end=2012)
```

https://www.iban.com/country-codes

Thailand	TH	THA
Timor-Leste	TL	TLS
Togo	TG	TGO
Tokelau	TK	TKL
Tonga	TO	TON
Trinidad and Tobago	TT	TTO
Tunisia	TN	TUN
Turkey	TR	TUR
Turkmenistan	TM	TKM

01 Import Data: API

US Census Bureau API

Week 16: June 02

Import Data: API

World Bank Indicators API

US Census Bureau Economic Indictors API

Let's go to US Census Bureau's Economic Indictors API.

01 Import Data: API

US Census Bureau API

Go to API Homepage → Check Commands → Search R Package → Using R to Get Data

Request API Key

Click "Request a key"

Request A Key

Organization Name: NCCU ICI

Email Address: torrent@nccu.edu.tw

☑ I agree to the terms of service

Submit Key Request

01 Import Data: API

US Census Bureau API

Census Data API Service <no-reply@census.gov> 5:24 PM (0 minutes ago)

to torrent

Hello!

Thank you for your interest in the Census Data API. Your API key is *fb56ed955* [redacted] Please click here to activate your key.

Save this email for future reference.

Have Fun,

The Census Bureau API Team

Follow @uscensusbureau on twitter for API updates.

Congratulations! Your key has been activated. You may now use it to query the Data API.

Happy querying!

01 Import Data: API

US Census Bureau API

```
install.packages("censusapi")
library(censusapi)
```

US Census Bureau provides thousands databases.
The website of US Census Bureau just likes a maze.

Censusapi provides tools to get data.

01 Import Data: API

US Census Bureau API

#Get Database list

apis <- listCensusApis()

	title	name	vintage	url
821	Economic Surveys: Annual Business Survey: Annual Business...	abscb	2017	http://api.census.gov/data/2017/abscb
986	Economic Surveys: Annual Business Survey: Annual Business...	abscb	2018	http://api.census.gov/data/2018/abscb
1131	Economic Surveys: Annual Business Survey: Annual Business...	abscb	2019	http://api.census.gov/data/2019/abscb
827	Economic Surveys: Annual Business Survey: Annual Business...	abscbo	2017	http://api.census.gov/data/2017/abscbo
987	Economic Surveys: Annual Business Survey: Annual Business...	abscbo	2018	http://api.census.gov/data/2018/abscbo
1132	Economic Surveys: Annual Business Survey: Annual Business...	abscbo	2019	http://api.census.gov/data/2019/abscbo
829	Annual Business Survey: Company Summary: 2017	abscs	2017	http://api.census.gov/data/2017/abscs
988	Economic Surveys: Annual Business Survey: Annual Business...	abscs	2018	http://api.census.gov/data/2018/abscs
1133	Economic Surveys: Annual Business Survey: Annual Business...	abscs	2019	http://api.census.gov/data/2019/abscs
1139	Economic Surveys: Annual Business Survey: Annual Business...	absnesd	2018	http://api.census.gov/data/2018/absnesd
1140	Economic Surveys: Annual Business Survey: Annual Business...	absnesdo	2018	http://api.census.gov/data/2018/absnesdo
1028	Annual Economic Surveys: Annual Business Survey	abstcb	2018	http://api.census.gov/data/2018/abstcb
751	American Community Survey: 1-Year Estimates: Detailed Tab...	acs/acs1	2005	http://api.census.gov/data/2005/acs/acs1
756	American Community Survey: 1-Year Estimates: Detailed Tab...	acs/acs1	2006	http://api.census.gov/data/2006/acs/acs1

01 Import Data: API

US Census Bureau API

#Get a database's variable list

```
vars_list <- listCensusMetadata(
  name = "acs/acs5",
  vintage = 2017,
  type = "variables")
```

	name	label	concept	p
1	B24022_060E	Estimate!!Total!!Female!!Service occupations!!Food preparati...	SEX BY OCCUPATION AND MEDIAN EARNINGS IN THE PAST...	in
2	B19001B_014E	Estimate!!Total!!$100,000 to $124,999	HOUSEHOLD INCOME IN THE PAST 12 MONTHS (IN 2017 I...	in
3	B07007PR_019E	Estimate!!Total!!Moved from different municipio!!Foreign bo...	GEOGRAPHICAL MOBILITY IN THE PAST YEAR BY CITIZENSH...	in
4	B19101A_004E	Estimate!!Total!!$15,000 to $19,999	FAMILY INCOME IN THE PAST 12 MONTHS (IN 2017 INFLATI...	in
5	B24022_061E	Estimate!!Total!!Female!!Service occupations!!Building and g...	SEX BY OCCUPATION AND MEDIAN EARNINGS IN THE PAST...	in
6	B19001B_013E	Estimate!!Total!!$75,000 to $99,999	HOUSEHOLD INCOME IN THE PAST 12 MONTHS (IN 2017 I...	in
7	B07007PR_018E	Estimate!!Total!!Moved from different municipio!!Foreign bo...	GEOGRAPHICAL MOBILITY IN THE PAST YEAR BY CITIZENSH...	in
8	B19101A_005E	Estimate!!Total!!$20,000 to $24,999	FAMILY INCOME IN THE PAST 12 MONTHS (IN 2017 INFLATI...	in
9	B19001B_012E	Estimate!!Total!!$60,000 to $74,999	HOUSEHOLD INCOME IN THE PAST 12 MONTHS (IN 2017 I...	in
10	B24022_062E	Estimate!!Total!!Female!!Service occupations!!Personal care ...	SEX BY OCCUPATION AND MEDIAN EARNINGS IN THE PAST...	in
11	B01001B_029E	Estimate!!Total!!Female!!65 to 74 years	SEX BY AGE (BLACK OR AFRICAN AMERICAN ALONE)	in

01 Import Data: API

US Census Bureau API

If we want all California counties':

1. SEX BY AGE
2. MEDIAN AGE BY SEX
3. Median household income in the past 12 months

The names of variables are:

1. B01001_001E
2. B01002_001E
3. B19013_001E

	name	label
1	B24022_060E	Estimate!!Total!!Female!!Service occupa
2	B19001B_014E	Estimate!!Total!!$100,000 to $124,999
3	B07007PR_019E	Estimate!!Total!!Moved from different r
4	B19101A_004E	Estimate!!Total!!$15,000 to $19,999
5	B24022_061E	Estimate!!Total!!Female!!Service occupa
6	B19001B_013E	Estimate!!Total!!$75,000 to $99,999
7	B07007PR_018E	Estimate!!Total!!Moved from different r
8	B19101A_005E	Estimate!!Total!!$20,000 to $24,999
9	B19001B_012E	Estimate!!Total!!$60,000 to $74,999
10	B24022_062E	Estimate!!Total!!Female!!Service occupa
11	B01001B_029E	Estimate!!Total!!Female!!65 to 74 years
12	B20005A_021E	Estimate!!Total!!Male!!Worked full-time

```
asc_df <- getCensus(name = "acs/acs5", vintage = 2020,
                    key = "fb56ec[redacted]",
                    vars = c("B01001_001E", "NAME", "B01002_001E", "B19013_001E", "GEO_ID"),
                    region = "county:*", regionin = "state:06")
```

01 Import Data: API

Other API

Two US Election Data's API:

1. **ProPublica: https://www.propublica.org/datastore/**

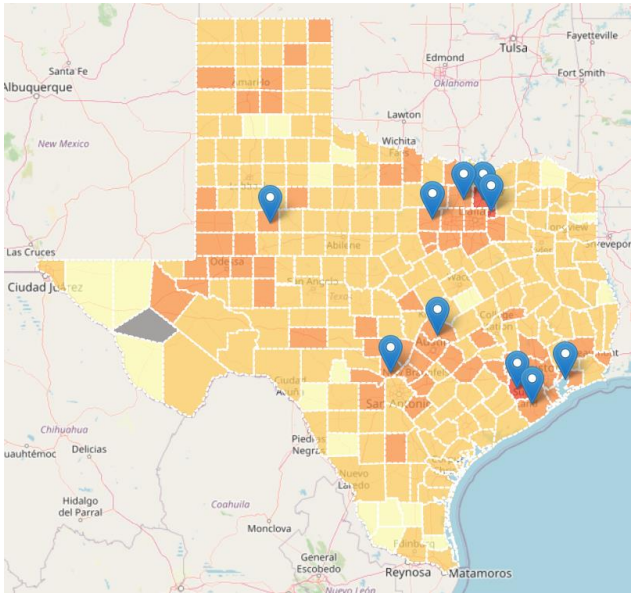
2. **Opensecrets: https://www.opensecrets.org/open-data/api**

01 Import Data: API

Practice

Please create a Texas income leaflet map

1. Use censusapi to collect Texas counties' 2022 income median data
2. Merge the income data into TX shapefile
3. Create a leaflet map to show all counties' income median in Texas and put blue balloons on the top 10 income countries.



Final Presentation

02 Final Presentation

Items	Points	%
Final Presentation	200	15.38%

- **Week 18, we will have a online joint presentation in Gather Town on June 15 (Wed.) 11:00am – 1:20pm.**
- **Every team has a booth and uses a poster to present your final project.**
- **I will announce technological issues as soon as possible.**

02 Final Presentation

Your final presentation should provide the following information:

1. Introduction: Introducing your data project to us, including your research questions.

2. Data sources: Providing information about your data sources, where they come from and how you clean them.

3. Models: Describing the models you use.

4. Results: Using texts, tables, or plots to show and explain your results and findings.

5. Conclusion: Telling us the contributions of your project's finding, such as suggestions for regulations, solutions for problems, or information for the public.

Data Project