

ISA 401: Business Intelligence & Data Visualization

08: Connecting to APIs in

Fadel M. Megahed, PhD

Raymond E. Glos Professor in Business
Farmer School of Business
Miami University

 @FadelMegahed


 fmegahed

 fmegahed@miamioh.edu

 Automated Scheduler for Office Hours

Fall 20254

Quick Refresher from Last Week

- ✓ Understand when can we scrape data (i.e., `robots.txt`)
- ✓ Scrape a webpage Using 
- ✓ Utilize loops to scrape data from multiple webpages

Kahoot Competition # 2

To assess your understanding and retention of the topics covered last week, you will **compete in a Kahoot competition (consisting of 9 questions)**:

- Go to <https://kahoot.it/>
- Enter the game pin, which will be shown during class
- Provide your first (preferred) and last name
- Answer each question within the allocated 20-second window (**fast and correct answers provide more points**)

Winning the competition involves having as many correct answers as possible AND taking the shortest duration to answer these questions. The winner 🏆 of the competition from each section will receive: 0.25 on their Web Scraping II Assignment. Good luck!!!

Learning Objectives for Today's Class

- Describe what we mean by an API
- Explain how APIs will be a huge part of your career as a business analyst and/or data scientist
- Use APIs for extracting web data

What is an API? (A Web Server Based Perspective)

What is an API? [1]

- An **API** is an acronym for application programming interface.
- It is a **popular** approach to interact with an application/service or data since it:
 - Defines a set of functionalities independent of implementation (i.e., it only exposes information that a programmer might find useful and keep those parts consistent even if the implementation changes later)
 - Provides some level of privacy/control over one's internal data and the rate at which it can be accessed.

What is an API? [2]

What is an API?



What is an API? [3]

Scenario: Alone, you went into a warehouse and are trying to retrieve 3 screwdrivers, a toolbox, and 15 Phillips screws. But you do not know, where those things are in the warehouse.




The API is the set of instructions provided by the warehouse manager on where/how to retrieve this information without touching/accessing other things in there.

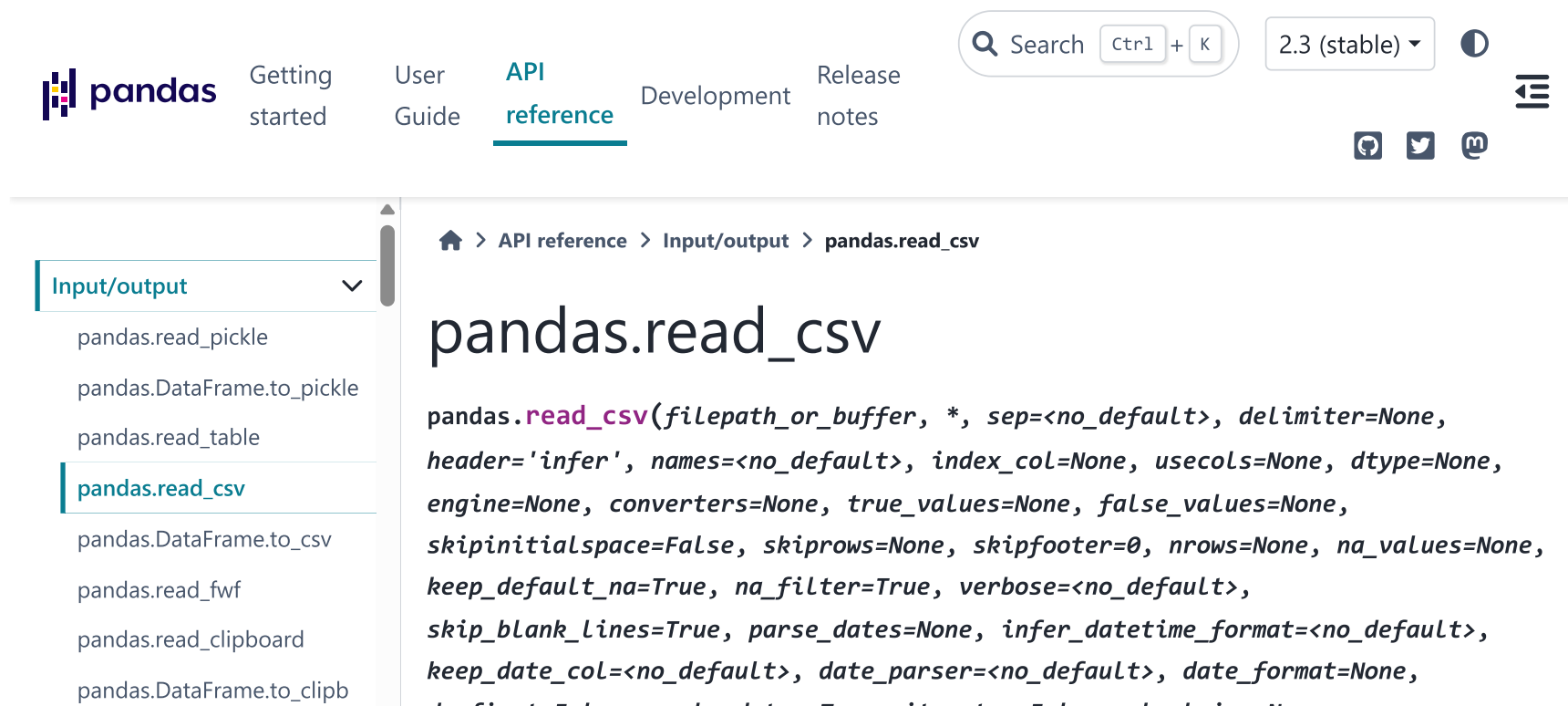
Source: Matt Z. (2018). What is an API? (explanation with cartoon picture so a 5 year old could understand it)

API Useages for Business Analysts and Data Scientists

Application #1: Request Data from Web Server

This is the **classical** example, which was explained in the previous slides and we will be coding several in-class demos to illustrate how to perform such operations in .

Application #2: Access Python Libraries



The screenshot displays the pandas API reference website. The top navigation bar includes the pandas logo, links for 'Getting started', 'User Guide', 'API reference' (which is underlined and highlighted), 'Development', and 'Release notes'. A search bar with the text 'Search' and a keyboard shortcut 'Ctrl + K' is present, along with a version selector set to '2.3 (stable)'. Social media icons for GitHub, Twitter, and Medium are also visible. On the left side, a sidebar menu shows the 'Input/output' section expanded, with 'pandas.read_csv' selected. The main content area shows the breadcrumb path 'API reference > Input/output > pandas.read_csv' followed by the function name 'pandas.read_csv' in a large font. Below this, the function signature is provided: `pandas.read_csv(filepath_or_buffer, *, sep=<no_default>, delimiter=None, header='infer', names=<no_default>, index_col=None, usecols=None, dtype=None, engine=None, converters=None, true_values=None, false_values=None, skipinitialspace=False, skiprows=None, skipfooter=0, nrows=None, na_values=None, keep_default_na=True, na_filter=True, verbose=<no_default>, skip_blank_lines=True, parse_dates=None, infer_datetime_format=<no_default>, keep_date_col=<no_default>, date_parser=<no_default>, date_format=None, dayfirst=False, cache_dates=True, iterator=False, chunksize=None)`.

Application #3: Use a Pretrained ML Model

02:47

Please click on the ["Watch on Vimeo"](#) to see OpenAI's explanation of their DALL·E 2 model.

Accessing APIs in

The 3 Step Process

Before you dive into the API documentation, you **should first check if there is a R (or Python if you are familiar)** package/library that serves as a wrapper for that API.

- **Find** the **API's documentation** and find information about the following:
 - A. Does the API require an **authentication key**?
 - B. What are the API's **base URL** and **query parameters**?
 - C. How does the request URL look like?
- Craft your **request**. My recommendations are to:
 - A. First, start with a simple request.
 - B. Test that request in your browser and see what results you get.
- Construct that request in **R** by **either**:
 - A. If the generated content seems to be a **JSON** file/webpage, you can capitalize on the reading the content from `jsonlite::fromJSON()`; **OR**
 - B. By passing the **base url** inside the `httr::GET()` and parsing the results with `httr::content()`.

Demo: The CryptoCompare API

Demo Description

Code and Results

- Create a Personal (Free) account at CryptoCompare.com
- Click on create your free key to create your API key and copy the key.
- Go to the [documentation](#), and test their sample call by executing the call after you have pasted your API key in the call.
 - The executed call returns the price of BTC (Bitcoin) in USD, JPY and EUR.
- Now click on the Historical Data Tab on the left
 - Click on Daily Pair OHLCV and Execute the Sample Call for BTC
 - This returns 10 days worth of OHLCV for BTC in USD.
- **Let us obtain the price for \$BTC over the past 100 days.**

Recap

Summary of Main Points

By now, you should be able to do the following:

- Describe what we mean by an API
- Explain how APIs will be a huge part of your career as a business analyst and/or data scientist
- Use APIs for extracting web data

Things to Do to Prepare for Next Class

- Getting Started with http
- Getting Started with http2
- Managing secrets