

W01D3

APIs

Instructor: Eric Elmoznino

Outline for today

- Introductions
- What are APIs, and why do we need them?
- Navigating an API and examples
- Break (10 min)
- Demo
 - Using the command line (terminal)
 - Using Postman
 - Using Python

Lecture protocol

- Will pause frequently for your questions (and to ask you questions)
 - I may call on people who are participating less
- Interrupt me when you have questions or are lost
- Tell me to slow down
- If you're an experienced coder, give others a chance to answer questions first
- Link to code, slides, and lecture recording are on Compass


Introductions









- Your name
- What did you do pre-bootcamp life?
- Why are you doing this bootcamp?
- Do you have coding experience?

Where do we get our data?

- Publicly available datasets (e.g. check out Kaggle)
 - Good for benchmarking, but limited for real use-case
- Company's database (e.g. transaction history)
 - SQL, MongoDB, etc.
- From the web
 - Collected manually
 - **Collected automatically (code)**

Collecting data from the web




-  Home
-  Explore
-  Notifications
-  Messages
-  Bookmarks
-  Lists
-  Profile
-  More


Tweet

←

TopLatestPeoplePhotosVideos


**la patrona** @robzgalloway · 8h
How are we meant to remember all the number plates & names of these abusive **uber** drivers??? 🚗🚗🚗🚗


33133

**GMF78** @GMF782 · 5h
Revised valuations for @eurasiamining after considering impact of gaining MT flank approval and rising #palladium prices. #EUA sale price : 46-63p - conservative value
94-111p - my new valuation
249-266p - **uber** bull value (v unlikely imo)
Further details : powerof78.com/2020/09/13/eua...

Eurasia Mining			
Current Share Price	Conservative Value	Full Takeover price	Bullish/Max Takeover price
	WACC 10%, PA \$100M avg	WACC 10%, PA \$100M avg	WACC 10%, PA \$100M avg
20.75p	46-63p	94-111p	249-266p

64578

**Jisoo feed** @jskfeed · 3h
your **Uber** has arrived 🚗




Advanced search

What's happening

NFL · LIVE

Dolphins at Patriots

Trending with: #Patriots and Dolphins



Football · Trending


#FlyEaglesFly

Trending with: #Eagles

34.7K Tweets

NBA · LIVE

Clippers at Nuggets




NFL · Trending

Wentz


21.3K Tweets

Show more

Who to follow

**Daniel Dale** @ddale8

Follow

**NASA** @NASA

Follow

Show more

Terms Privacy policy Cookies Ads info More

© 2020 Twitter, Inc.

Collecting data from the web

```
<!DOCTYPE html>
<html dir="ltr" lang="en" style="font-size: 15px;">
  <head>_</head>
  <body style="background-color: rgb(255, 255, 255); overflow-y: scroll;">
    <noscript>_</noscript>
    <div id="react-root" style="height:100%;display:flex;">
      <div class="css-1dbjc4n r-13awgt0 r-12vffkv">
        <div class="css-1dbjc4n r-13awgt0 r-12vffkv">
          <h2 aria-level="2" role="heading" class="css-4rbku5 css-1dbjc4n r-4iw3lz r-1xx2f4g r-109y4c4 r-1udh08x r-wwvuq4 r-u8sld r-92ng3h">
            <div dir="auto" class="css-90loao r-hkyrab r-1qd0xha r-a023e6 r-16dba41 r-ad9z0x r-bcqeeo r-qvutc0">
              <span class="css-90loao css-16my406 r-1qd0xha r-ad9z0x r-bcqeeo r-qvutc0">To view keyboard shortcuts, press question mark</span>
            </div>
            <a href="/i/keyboard_shortcuts" dir="auto" role="link" tabindex="-1" class="css-4rbku5 css-18t94o4 css-90loao r-1n1174f r-1loqt21 r-1qd0xha r-a023e6 r-16dba41 r-ad9z0x r-bcqeeo r-qvutc0">
              <span class="css-90loao css-16my406 r-1qd0xha r-ad9z0x r-bcqeeo r-qvutc0">View keyboard shortcuts</span>
            </a>
          </h2>
          <div id="layers" class="r-1d2f490 r-u8sld r-zchlnj r-ipm5af r-184en5c" style>
            <div class="css-1dbjc4n r-aqfbo4 r-1p0dtai r-1d2f490 r-12vffkv r-1xcajam r-zchlnj">
              <div class="css-1dbjc4n r-12vffkv">_</div>
            </div>
            <div class="css-1dbjc4n r-18u37iz r-13qzluu r-417010" data-at-shortcutkeys="{"n":"New Tweet","Cmd Enter":"Send Tweet","m":"New Direct Message","/":"Search","l":"Like","r":"Reply","t":"Retweet","s":"Share Tweet","b":"Bookmark","u":"Mute account","x":"Block account","Enter":"Open Tweet details","o":"Expand photo","i":"Open/Close Messages dock","?":"Shortcut help","j":"Next Tweet","k":"Previous Tweet","Space":"Page down",".":"Load new Tweets","g h":"Home","g e":"Explore","g n":"Notifications","g r":"Mentions","g p":"Profile","g l":"Likes","g i":"Lists","g m":"Direct Messages","g s":"Settings","g b":"Bookmarks","g u":"Go to user...","g d":"Display settings"}" aria-hidden="false" style="min-height: 897px;">
              <div aria-label="Skip to recommended content" role="button" data-focusable="true" tabindex="0" class="css-18t94o4 css-1dbjc4n r-1niwhzg r-lawa8pu r-sdzlij r-lphboty r-4iw3lz r-1xx2f4g r-109y4c4 r-1vuscf0 r-1dhvaqw r-1ny4l3l r-1udh08x r-wwvuq4 r-1fneopy r-u8sld r-o7ynqc r-6416eg r-lrvibr r-92ng3h">
                <div dir="auto" class="css-90loao r-lawozwy r-1re7ezh r-6koalj r-18u37iz r-16y2uox r-1qd0xha r-a023e6 r-vw2c0b r-1777fci r-eljoum r-dnmrzs r-bcqeeo r-q4m81j r-qvutc0">
                  <span class="css-90loao css-16my406 css-bfa6kz r-1qd0xha r-ad9z0x r-bcqeeo r-qvutc0"></span>
                </div>
              </div>
              <div aria-label="Skip to secondary content" role="button" data-focusable="true" tabindex="0" class="css-18t94o4 css-1dbjc4n r-1niwhzg r-lawa8pu r-sdzlij r-lphboty r-4iw3lz r-1xx2f4g r-109y4c4 r-1vuscf0 r-1dhvaqw r-1ny4l3l r-1udh08x r-wwvuq4 r-1fneopy r-u8sld r-o7ynqc r-6416eg r-lrvibr r-92ng3h">
                <div dir="auto" class="css-90loao r-lawozwy r-1re7ezh r-6koalj r-18u37iz r-16y2uox r-1qd0xha r-a023e6 r-vw2c0b r-1777fci r-eljoum r-dnmrzs r-bcqeeo r-q4m81j r-qvutc0">
                  <span class="css-90loao css-16my406 css-bfa6kz r-1qd0xha r-ad9z0x r-bcqeeo r-qvutc0"></span>
                </div>
              </div>
              <div role="banner" class="css-1dbjc4n r-obd0qt r-16y2uox r-1g40b8q">
                <div class="css-1dbjc4n r-1gymjhz">
                  <div class="css-1dbjc4n r-aqfbo4 r-1pi2tsx r-1xcajam r-ipm5af">
                    <div class="css-1dbjc4n r-1pi2tsx r-lwtj0ep r-lrnoaur r-ou255f r-1gymjhz">
                      <div class="css-1dbjc4n r-lawozwy">
                        ...
                      </div>
                    </div>
                  </div>
                </div>
              </div>
            </div>
          </div>
        </div>
      </div>
    </body>
  </html>
```

Data mining

- Program that goes to URLs and **parses the HTML to extract data**
- Effortful
 - HTML is difficult to parse
 - Almost all information is irrelevant (e.g. UI-related)
 - Websites often require interaction (e.g. “Load More” button)
 - When websites update, your code will break
 - Every website is different
 - Companies try to stop data miners

APIs to the rescue

- Application Programming Interface
- Programmer-friendly version of websites
- Go to a URL composed of

- API root endpoint

<https://api.translink.ca/rttiapi/v1/stops?apikey=YOURKEY&lat=49.18&long=-122.85>

- API function

- API key (like a login)

- Parameter keys

- Parameter values

- Returns data (JSON, XML, csv, etc.)

```
JSON
{
  "StopNo": 54997,
  "Name": "SB KING GEORGE BLVD FS 98 AVE",
  "BayNo": "N",
  "City": "SURREY",
  "OnStreet": "KING GEORGE BLVD",
  "AtStreet": "98 AVE",
  "Latitude": 49.179601,
  "Longitude": -122.845814,
  "WheelchairAccess": 1,
  "Distance": 307,
  "Routes": "314, 321, 329"
}
```

Examples of APIs

- Translink API: <https://developer.translink.ca/>
- GitHub API: <https://developer.github.com/v3/>
- Weather API: <https://openweathermap.org/api>
- Stock price API (Alpha Vantage): <https://www.alphavantage.co/>
- Many, many others!

Activity: find an example

- Look through an API you're interested in
- Find an API call (URL), explain what data you would get back
- Post this in Slack
- 10 minutes
- Don't be afraid of documentation!
- Documentation usually provides examples

HTTP Requests

- When you access a website (through an URL), you are:
 - Sending a HTTP GET request to the server to retrieve data
 - "data" can be a webpage that is displayed, it can be JSON, etc.
- GET is not the only request type (e.g. POST for sending information)
- Status codes are helpful when you're working with code
 - 200 OK
 - 400 Bad Request
 - 401 Unauthorized
 - 404 Not Found
 - ...

Regarding final projects...

Some data science problems (e.g. stock prediction, twitter sentiment, etc.) will require data from APIs, but the APIs **might not let you query dates in the past!**

If you know you want to work with data that has a temporal component for your final project:

- Look for an API early
- Check if it lets you query past dates
- If not, write a python program that queries the API on the current date and run it every day (or automate it!)

Demo: using the command line

- `curl https://api.translink.ca/rttiapi/v1/stops/55612?apikey=Dnmlzo7VEguzBUIE5W37`
- It is best practice to set your API key as an **environment variable**
 - `export TRANSLINK_API=YOURAPIKEY`
 - `echo $TRANSLINK_API` (should print out your API key)
 - **Note:** not permanent, have to redo this every time you close the command line! (Google how to “set environment variable permanently” for your operating system.)
- `curl https://api.translink.ca/rttiapi/v1/stops?apikey=$TRANSLINK_API`
- Want to know more about curl? [Understand curl more | Documenting APIs](#)

Demo: using Postman

- Get it here: <https://www.postman.com/>
- Easy to use GUI which:
 - Stores your API key
 - Let's you neatly specify parameter keys + values
 - Saves your past API calls

Demo: using Python

- https://github.com/EricElmoznino/lighthouse_api_tutorial