

Lab exercise: Getting Started with APIs for open data in R

```
## Warning: package 'knitr' was built under R version 4.1.2
```

Bonus opportunity (10 points)

Can you use the API endpoint to get a subset of crimes data from the Chicago data portal, which only includes the data since 2021-01-01 to 2021-07-01 and the primary type is THEFT? Please only select these columns to be returned: date, primary__ type, x-coordinate and y-coordinate. Name the returned subset data as theft_2021. Hint: you can use & inside the `paste0()` function to add additional filtering parameters.

Q1: How many crimes occurred between 2021-01-01 to 2021-07-01 are theft? (5 points)

Answer: 17091

Q2: After retrieving this subset via the API request, use the `head()` function to return the first 6 records of the theft_2021 dataset. A screenshot of the first 6 records is needed as delivery (5 points)

Answer:

```
> head(theft_2021)
```

	date	primary_type	x_coordinate	y_coordinate
1	2021-04-16T20:45:00.000	THEFT	<NA>	<NA>
2	2021-04-02T19:09:00.000	THEFT	1147368	1919108
3	2021-04-16T04:10:00.000	THEFT	1138223	1928483
4	2021-04-16T15:00:00.000	THEFT	1172101	1906785
5	2021-04-16T21:00:00.000	THEFT	1170490	1903271
6	2021-04-16T14:00:00.000	THEFT	1133917	1925929

```
>
```

```
library(tidyverse)
library(rjson)
library(jsonlite)
library(RCurl)
# The base URL below is the API endpoint you copied from the page above.

base_URL = "https://data.cityofchicago.org/resource/crimes.json?"

base_theft_count <- paste0(base_URL, "$limit=900000", "&$where=date between '2021-01-01T00:00:00.000'
and '2021-06-30T23:59:59.000'",
"&primary_type='THEFT'", "&$select=count(date)")

fromJSON(URLEncode(base_theft_count))

##    count_date
## 1      17091
```

```
base_theft_2021 <- paste0(base_URL, "$limit=20000", "&$where=date between '2021-01-01T00:00:00.000' and  
  '2021-06-30T23:59:59.000'",  
  "&primary_type='THEFT'", "&$select=date,  
  primary_type, x_coordinate, y_coordinate")  
theft_2021 <- fromJSON(URLEncode(base_theft_2021))  
head(theft_2021)
```

```
##           date primary_type x_coordinate y_coordinate  
## 1 2021-04-16T20:45:00.000      THEFT      <NA>      <NA>  
## 2 2021-04-02T19:09:00.000      THEFT    1147368    1919108  
## 3 2021-04-16T04:10:00.000      THEFT    1138223    1928483  
## 4 2021-04-16T15:00:00.000      THEFT    1172101    1906785  
## 5 2021-04-16T21:00:00.000      THEFT    1170490    1903271  
## 6 2021-04-16T14:00:00.000      THEFT    1133917    1925929
```