

SCRAPING BIG DATA WITHOUT BREAKING THE BANK

USING TWITTER API FOR ACADEMIC RESEARCH

HANNAH L. FORD

WHAT IS BIG DATA?

Big data is a term describing **large volumes of data** and the way it is strategically analyzed (SAS, 2021).

Social media sites continuously generate information; **Twitter reported 199 million daily active users** and **over a trillion tweets** as of 2019 (Leetaru, 2019).

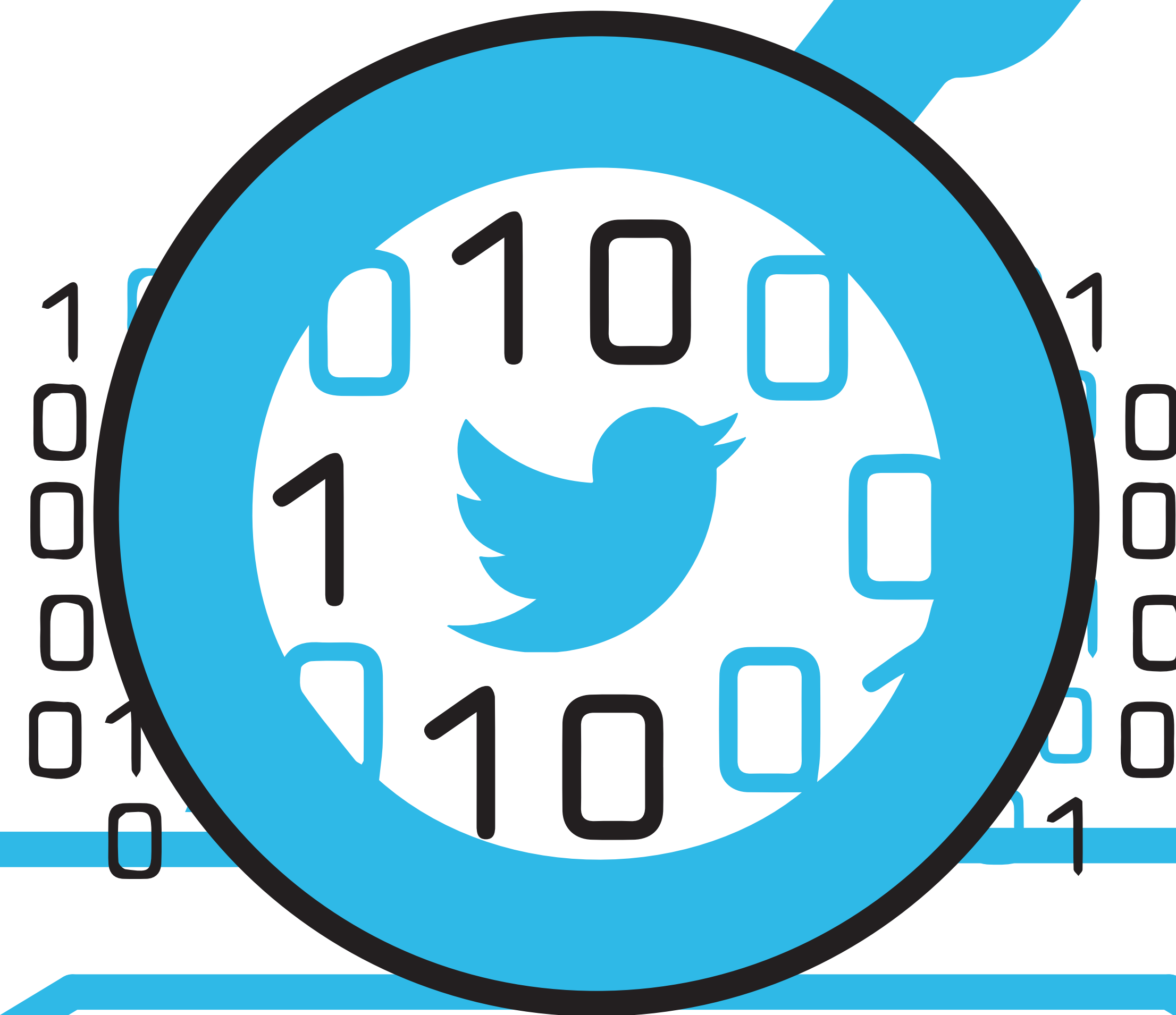
Data are currently characterized globally as the **"most valuable resource"** (Borrero & Zabalo, 2021).

HOW API WORKS FOR YOU

Twitter API v2 now grants individual researchers – **with proper coding experience** – access to the entire Tweet archive: **the first tweet in 2006 until now**.

One COVID study analyzed over **8 million tweets** for sentiment analysis (Aguilar-Gallegos et al., 2020).

Upon building this literature review, no agricultural education and communication research studies implementing API v2 were found.



COST SAVED WITH API

Cost of API: Free

\$2,499 - \$7,200: paid monthly databases

Paid examples: Meltwater & Cision

Cons of paid: Data aged 7 - 30 days

Cons of paid: Max 380k tweets

RESOURCES NEEDED

High-end processing PC - \$2,000

Time - 30 seconds 498 tweets

10 hours 1,000,000 tweets

50 hours 10,000,000 tweets

Coding experience - Javascript, Java, R, Python, cURL, HTTP

RESULTS TO DATE

TWITTER DISCUSSIONS ON FOOD DESERTS & INSECURITY

USDA FOOD DESERT MAP (ATLAS, USDA, 2021)

TWEET INFO

- 📍 Coordinates*
 - Full text
 - Hashtags
 - Engagement
- 📍 Location
 - Biography
 - Followers
 - Following
 - Friends (*mutual*)
 - Tweet count
- ✅ Verified status
 - Time of tweet
 - Date of tweet

*With security permissions

PROCESS

1. Research topic
2. Create abstract
3. Apply for API v2
4. Grant access
5. Code timeline
6. Run programs
7. Dehydrate
8. Process data
9. Rehydrate
10. Run Analysis
11. Save money

Parameters
When to when?

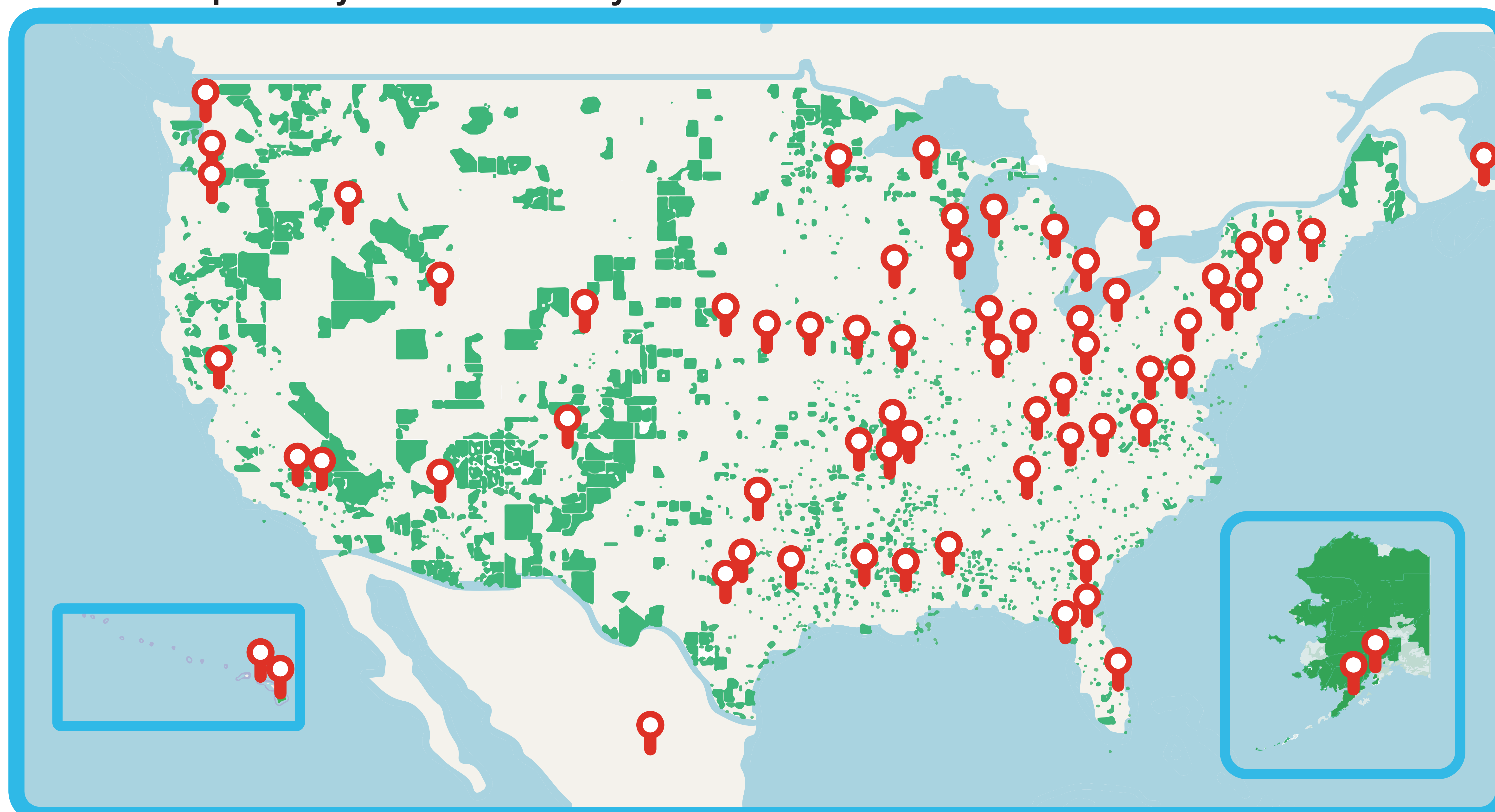
Exceptions

What doesn't matter?

Coding tasks

Step by step for PC

Map Overlay of Food Insecurity Twitter Conversations and USDA Atlas Food Deserts



*498 Tweets from June 2021, "Food insecurity" & "Food desert" (Free, 30 seconds)

