

# Web Scrapping With Reverse Engineering



# Hello!

**I'm David Teather**

I work as a software engineer and  
my speciality is data extraction.

Twitter: @david\_teather

Everything else: @davidteather



## Anonymous Social Media App Yik Yak Exposed Users' Precise Locations

Privacy researchers have found that it's possible to find out the precise location of Yik Yak posts, potentially exposing users to doxing or stalking.

# What I'm Known For

- My research on YikYak (a social media app) has been featured in Vice and The Verge
- Creating data extraction tools
  - Most popular is TikTokApi
    - 600K+ Downloads
    - 2.3K+ Stars

# How You Will Learn

- ⬡ Real website examples
- ⬡ Websites which I've created for this course
- ⬡ Each lesson will have a hands on activity
  - Actual programming and a test suite to check if you've completed it

# How To Learn Effectively

- ⬡ Everybody learns differently
- ⬡ Take notes about the new information presented
- ⬡ Try the hands on activities before watching the solution in the video
  - Don't look at the source code of the mock websites

# Course Topics

- ⬡ Forging API Requests
- ⬡ Proxies
- ⬡ Captchas
- ⬡ Storing Data at Scale
- ⬡ Emulating Human Behavior
- ⬡ And More!

**Note:** Not an exhaustive list, feel free to leave comments on what you'd like to see!



# Prerequisites

- ⬡ A basic understanding of programming
- ⬡ Recommended
  - Some Python experience
    - In the beginning, we won't be doing complex Python



# Learning Objectives

- ⬡ Learners will understand the many different ways websites prevent web scraping
- ⬡ Learners will be able to reverse engineer a real-world website for data extraction



# Tools Required

- ❏ Docker
  - And Docker compose
- ❏ Python 3
  - I'll be using 3.10
- ❏ A web browser
  - I'll be using Brave
- ❏ Download Course Files



# Thanks!

**Any questions?**

Twitter: @david\_teather

Everything Else: @davidteather



# Credits

- ⬡ Photographs by [Unsplash](#)
- ⬡ Presentation template by [SlidesCarnival](#)

**subscribe?**

**First Lesson**

**Playlist**