

NIL

# Scraping NIL Deals

Troy Gilmore – May 18, 2023

ON3

# Variables

1. Athlete name
2. School(s)
3. Sport
4. Sponsorship
5. Social Media

on3.com/nil/deals/


ON3 OS

Subscribe

Undisclosed 5/09/23 SOURCE

NCAA



JR




**Anderson Coby**

WR / 6-2 / 180

Deerfield Beach (Deerfield Beach, FL )


 → 

★★★★★ 

★★★★★ 85.10

NIL VALUE -

COLLECTIVE



Hoosiers For Good

COMPANY/CLIENT

*Stop the Violence Indianapolis*

Undisclosed 5/09/23 SOURCE

LOAD MORE

✓ Verified: Deal contract has been reviewed by On3/NIL

✎ Agent or Agency

# Variables

1. Athlete name
2. School(s)
3. Sport
4. Sponsorship
5. Social Media

on3.com/nil/deals/

ON3 OS

Subscribe

Undisclosed 5/09/23 SOURCE

NCAA

JR

**Anderson Coby**

WR / 6-2 / 180

Deerfield Beach (Deerfield Beach, FL)

T → Ψ

★★★★★ Ψ

★★★★★ 85.10

NIL VALUE -

COLLECTIVE

Hoosiers For Good

COMPANY/CLIENT

Stop the Violence Indianapolis

Undisclosed 5/09/23 SOURCE

LOAD MORE

✓ Verified: Deal contract has been reviewed by On3/NIL

✎ Agent or Agency

on3.com/nil/deals/


ON3 OS

Subscribe

Undisclosed 5/09/23 SOURCE

NCAA

JR



**Anderson Coby**

WR / 6-2 / 180

Deerfield Beach (Deerfield Beach, FL )


T → Ψ

★★★★★ Ψ

★★★★★ 85.10

NIL VALUE -

COLLECTIVE



Hoosiers For Good

COMPANY/CLIENT

*Stop the Violence Indianapolis*

Undisclosed 5/09/23 SOURCE

LOAD MORE

✓ Verified: Deal contract has been reviewed by On3/NIL

✎ Agent or Agency

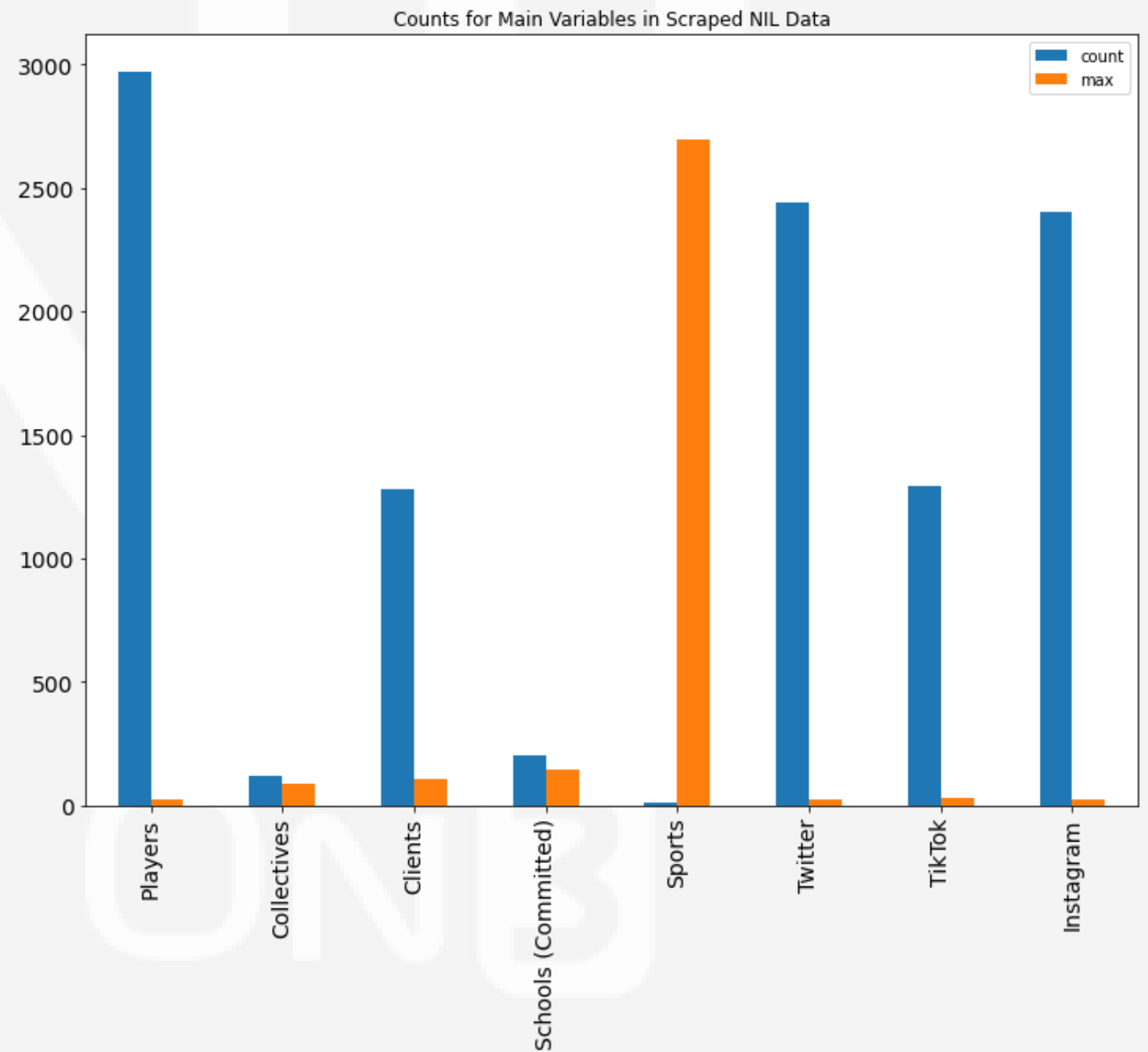
# Methods

- Selenium
- BeautifulSoup
- Pandas
- NumPy
- RegEx (re)

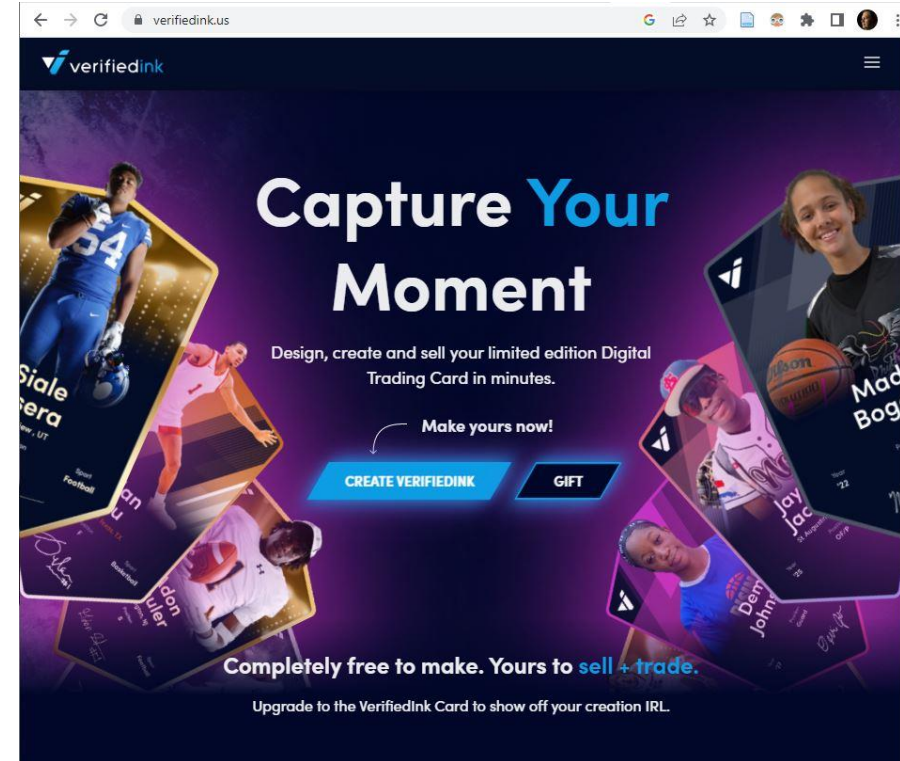
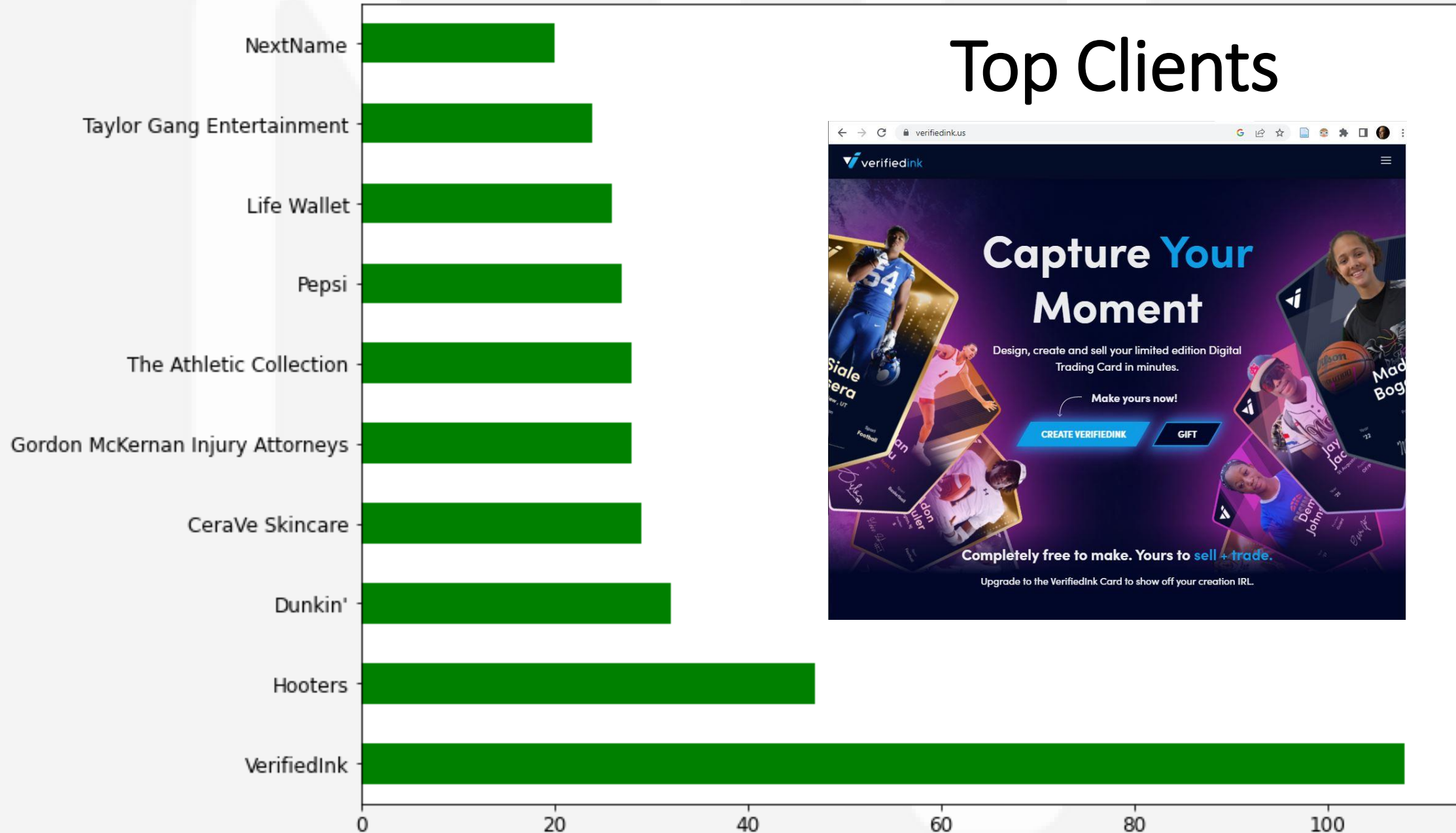
# Approach

- Selenium → click “Load More” 200x = ~5,000 athlete containers
- BeautifulSoup →
  - Extract variables directly
  - Extract athlete URL for every athlete
- BeautifulSoup + RegEx (re) →
  - Load athlete URL
  - Extract social media URLs
- “Sport” variable →
  - Build library of 41 unique positions
  - Manually create dictionary {position : sport}
  - Use conditionals to fill in sport for each athlete

# Overall Counts

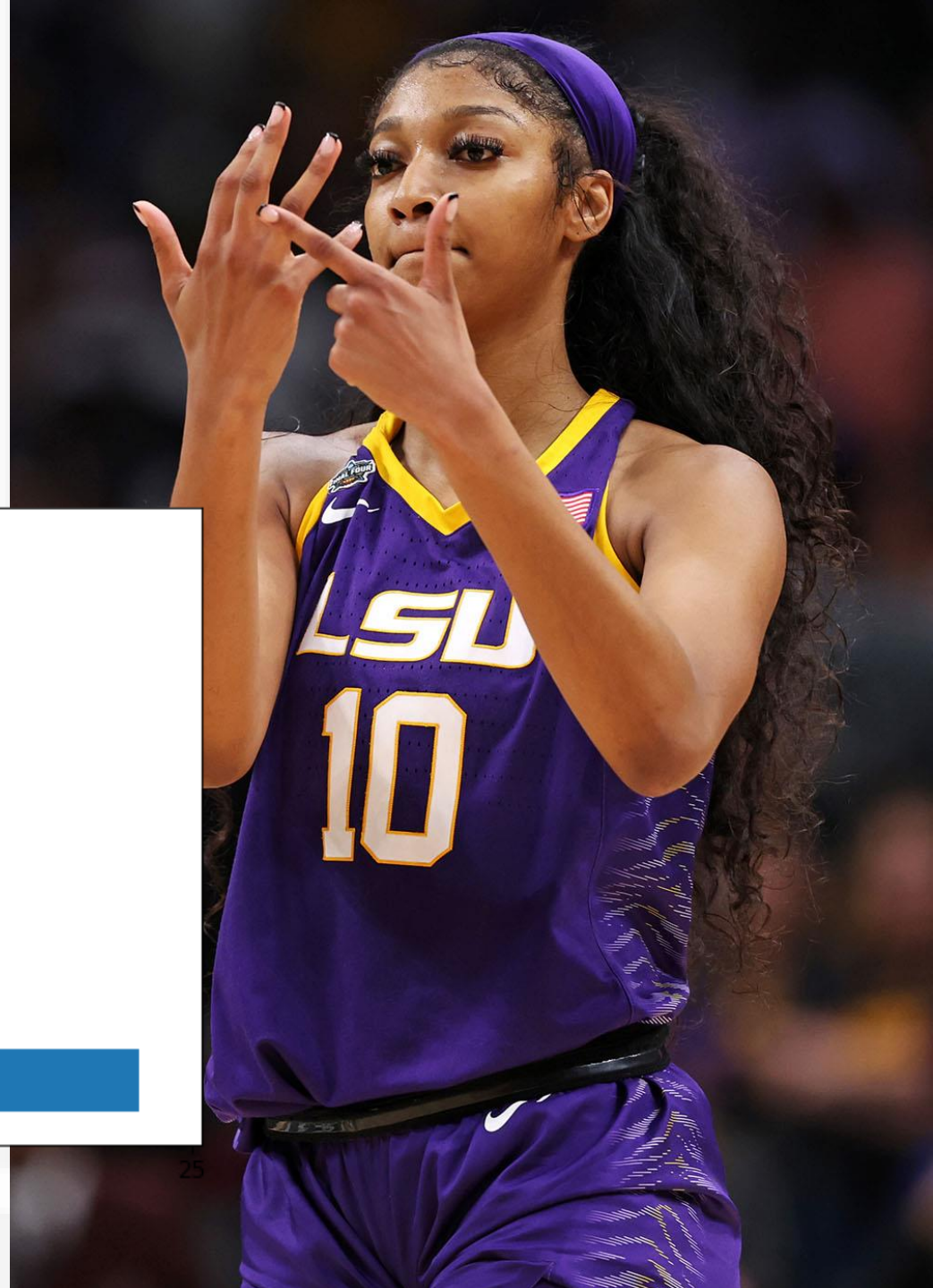
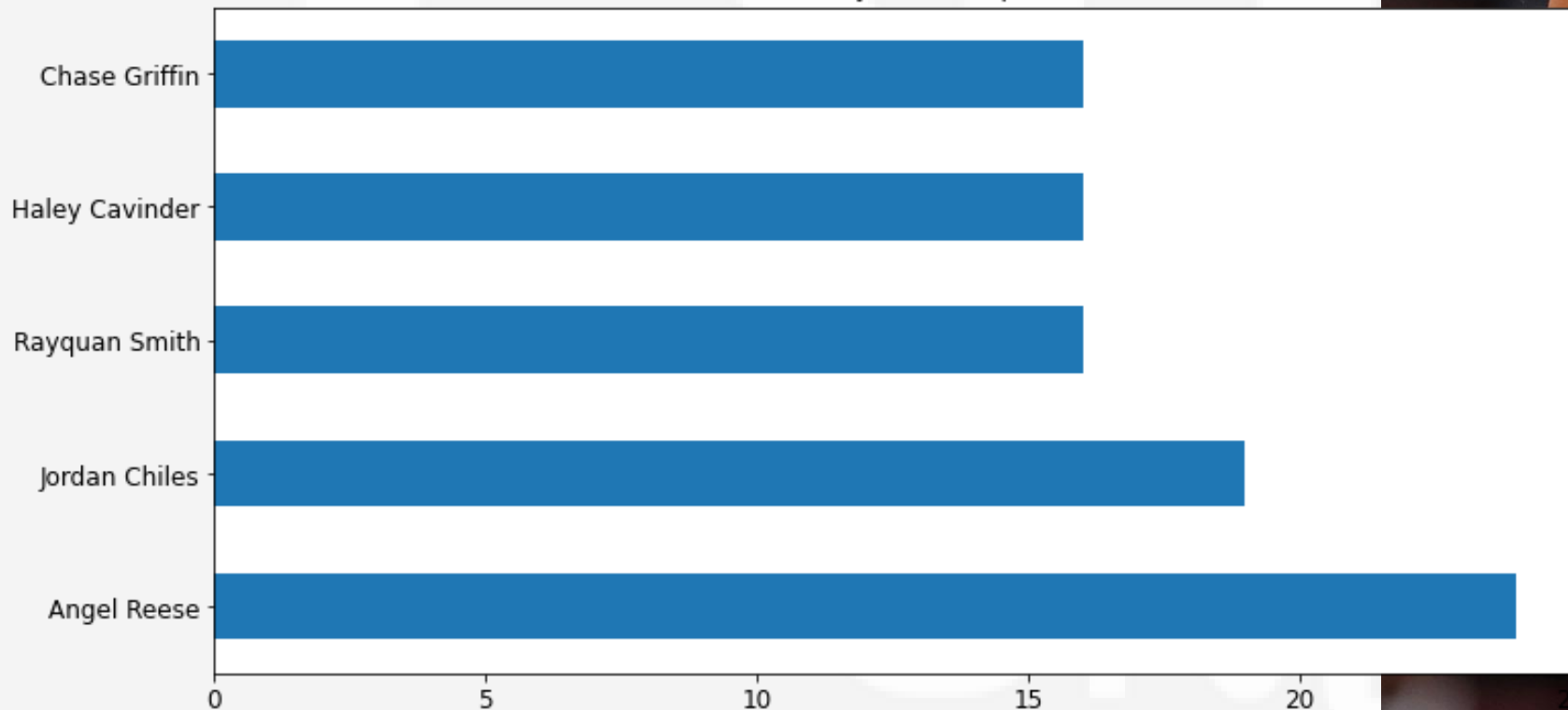


# Top Clients



# Top Players

Most Common Players in Scraped NIL Data





# Summary

- 5,018 athlete records
- 2,972 unique player names
- “sport” was most difficult variable
- Time variable would be interesting

