

Project 75

Andrew Seo

Introduction



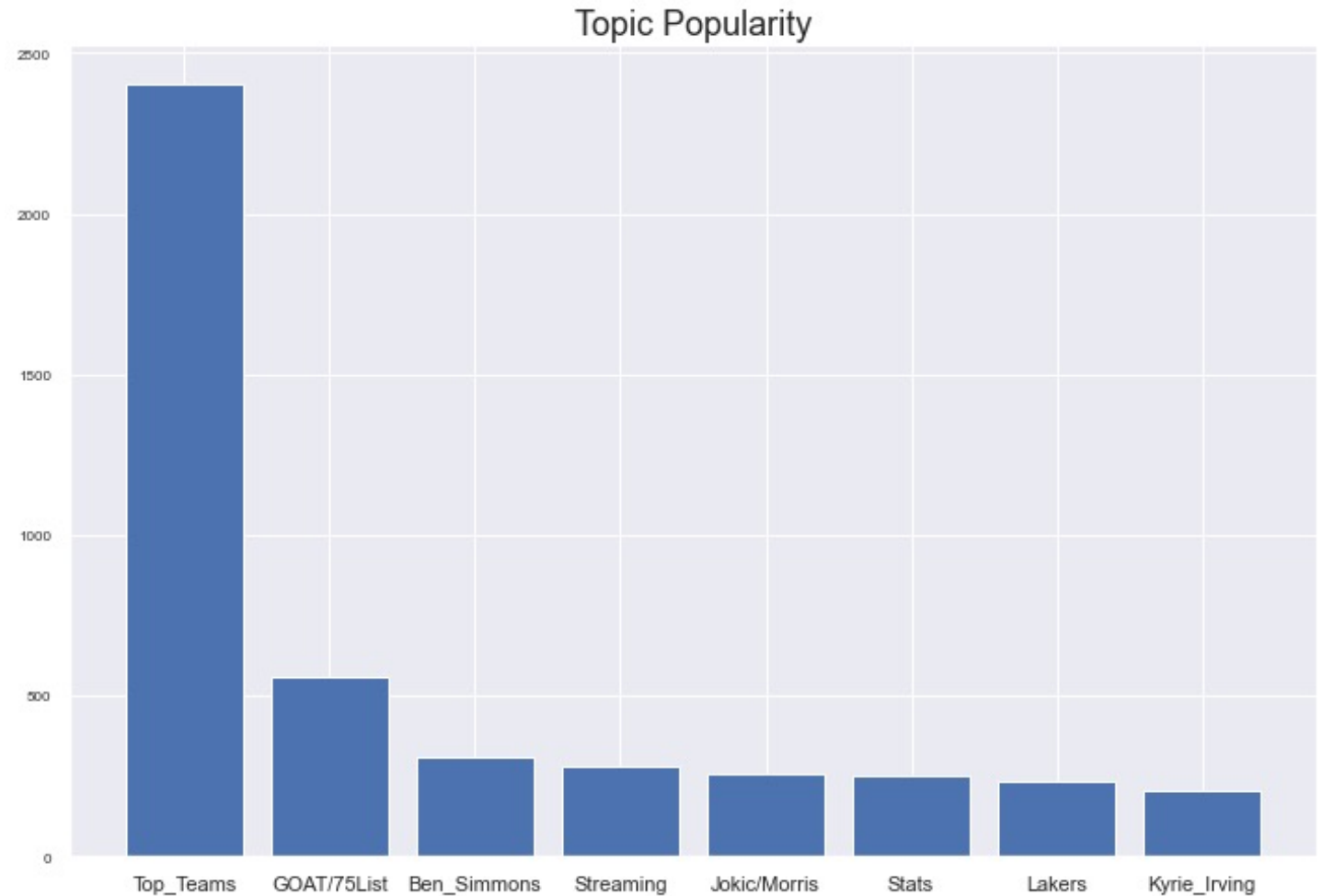
- Question: What do NBA fans really care about?
- How: Reddit - NLP and Unsupervised Learning Methods
- Audience: Advertisers and Front Offices

Tools & Methodology

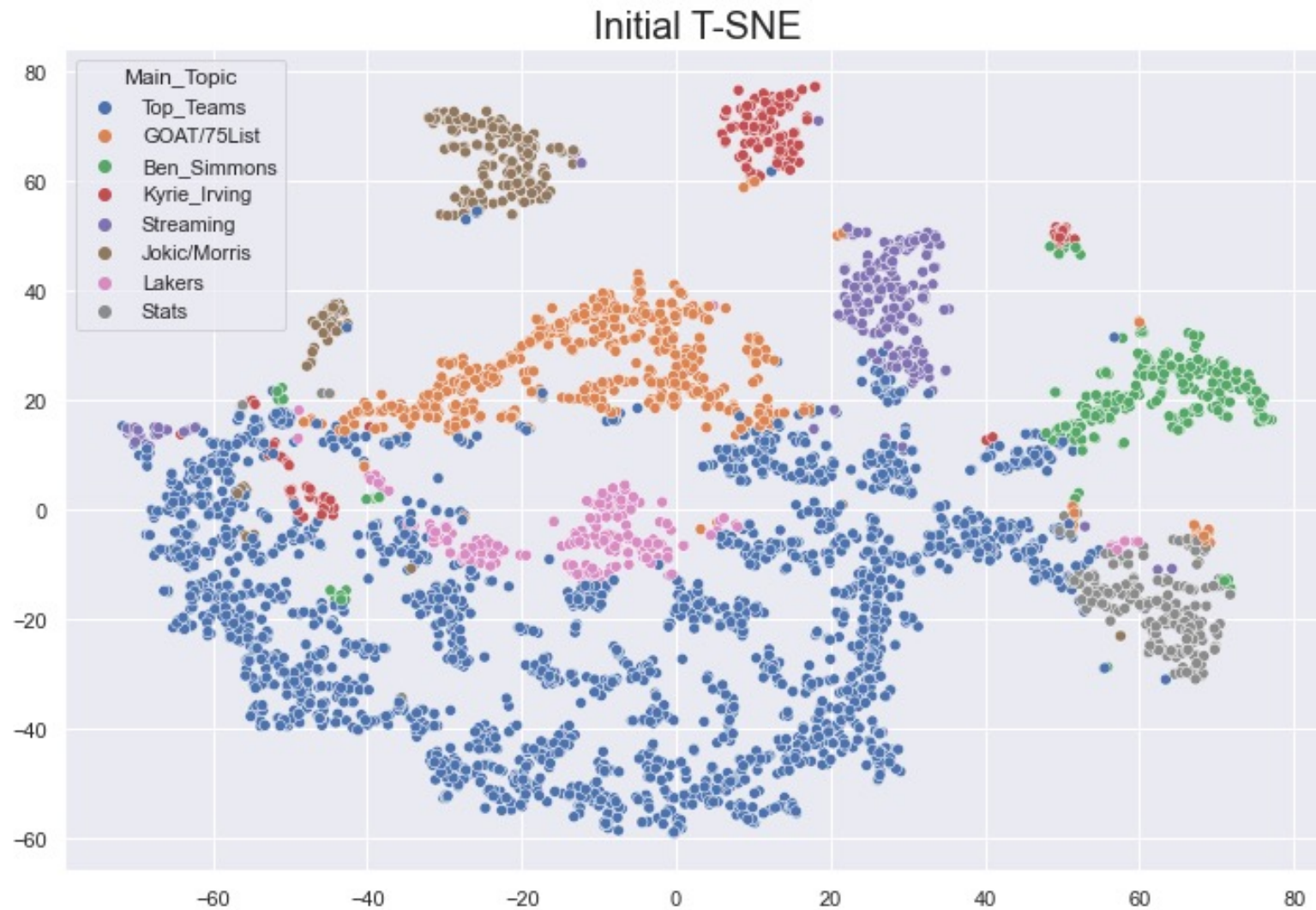
- Data: PSAW to gather NBA Subreddit posts from beginning of season to present (10/17/21-11/11/21), ~12k posts
- Dimensionality Reduction/Topic Modeling: LDA, NMF, LSA and domain knowledge
- Application: Clustering, EDA

Topic Modeling: Initial

- NMF(8)
- Top_Teams: ~2400
- Rest of Topics: ~2100
- GOAT/75List



T-SNE Attempt



First Dive – GOAT/75List



New Results

NMF(8)

Topic 0

lebron, james, lakers, kuzma, ending, dunk, squid, kyle, takes, nft

Topic 1

kevin, paul, malone, bob, dave, george, anthony, jerry, durant, allen

Topic 2

harden, james, washed, free, foul, throws, changes, rule, planting, twitter

Topic 3

kd, giannis, prime, finals, wins, harden, nets, mj, ad, westbrook

Topic 4

list, greatest, dwight, carmelo, howard, anthony, remove, replace, davis, believe

Topic 5

curry, steph, stephen, prime, seth, shaq, george, paul, durant, shooter

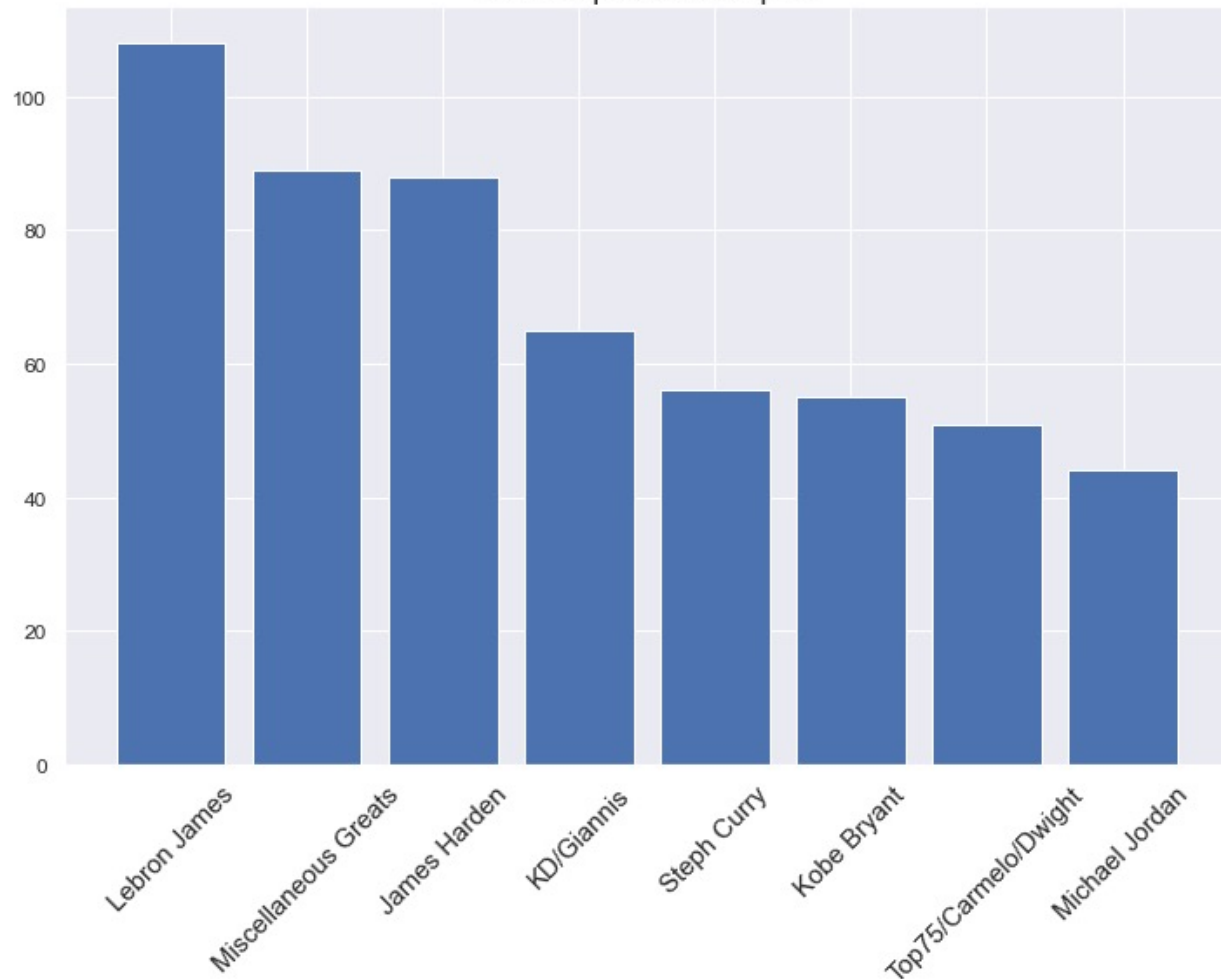
Topic 6

jordan, michael, poole, hakeem, nike, gay, rant, jesus, shorts, goat

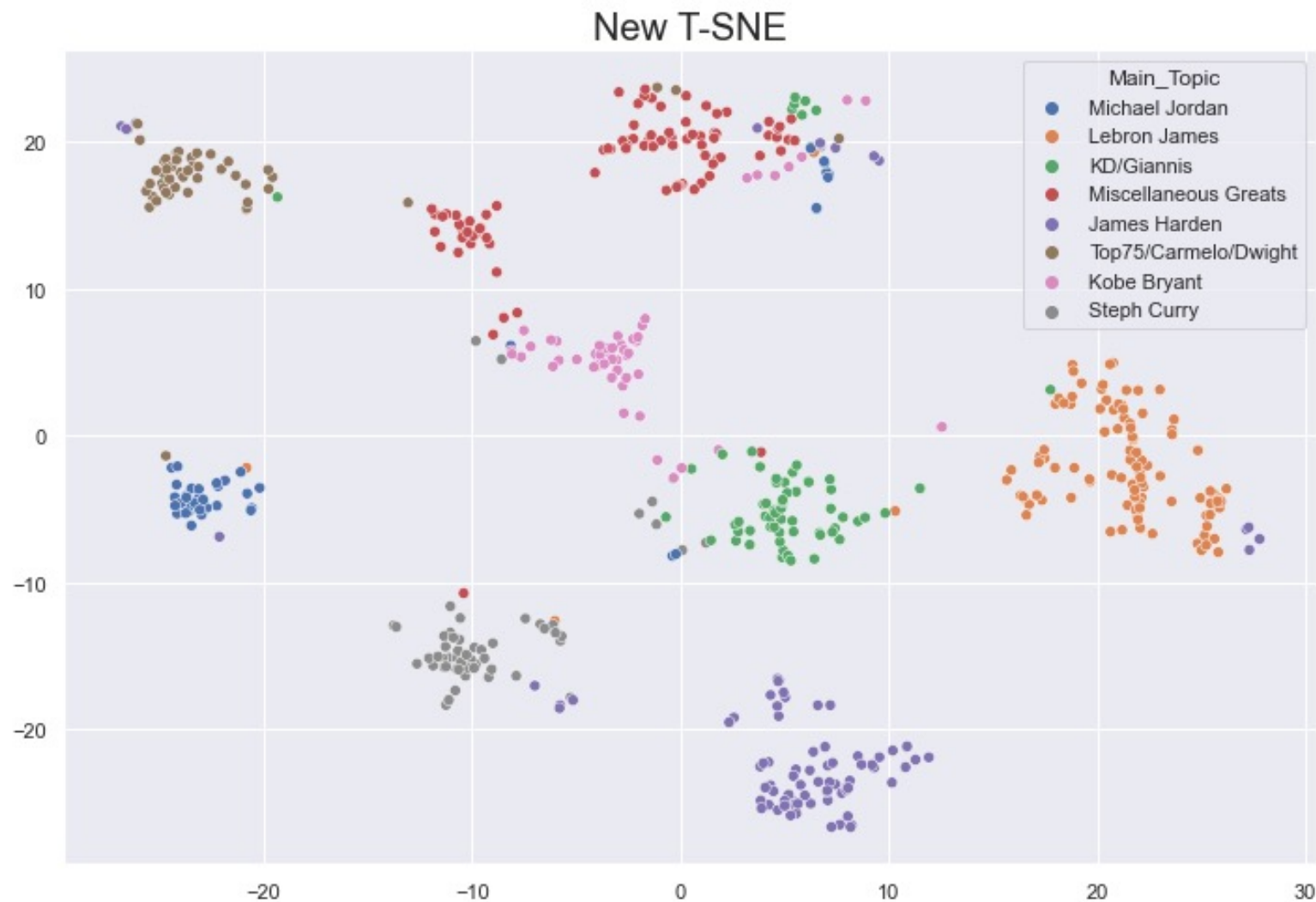
Topic 7

kobe, bryant, kareem, bird, mj, shaq, wilt, bodyarmor, cola, coca

Most Popular SubTopics



New Results: T-SNE Redux



Conclusions

Most Submissions: Present news and reactionary about players and teams

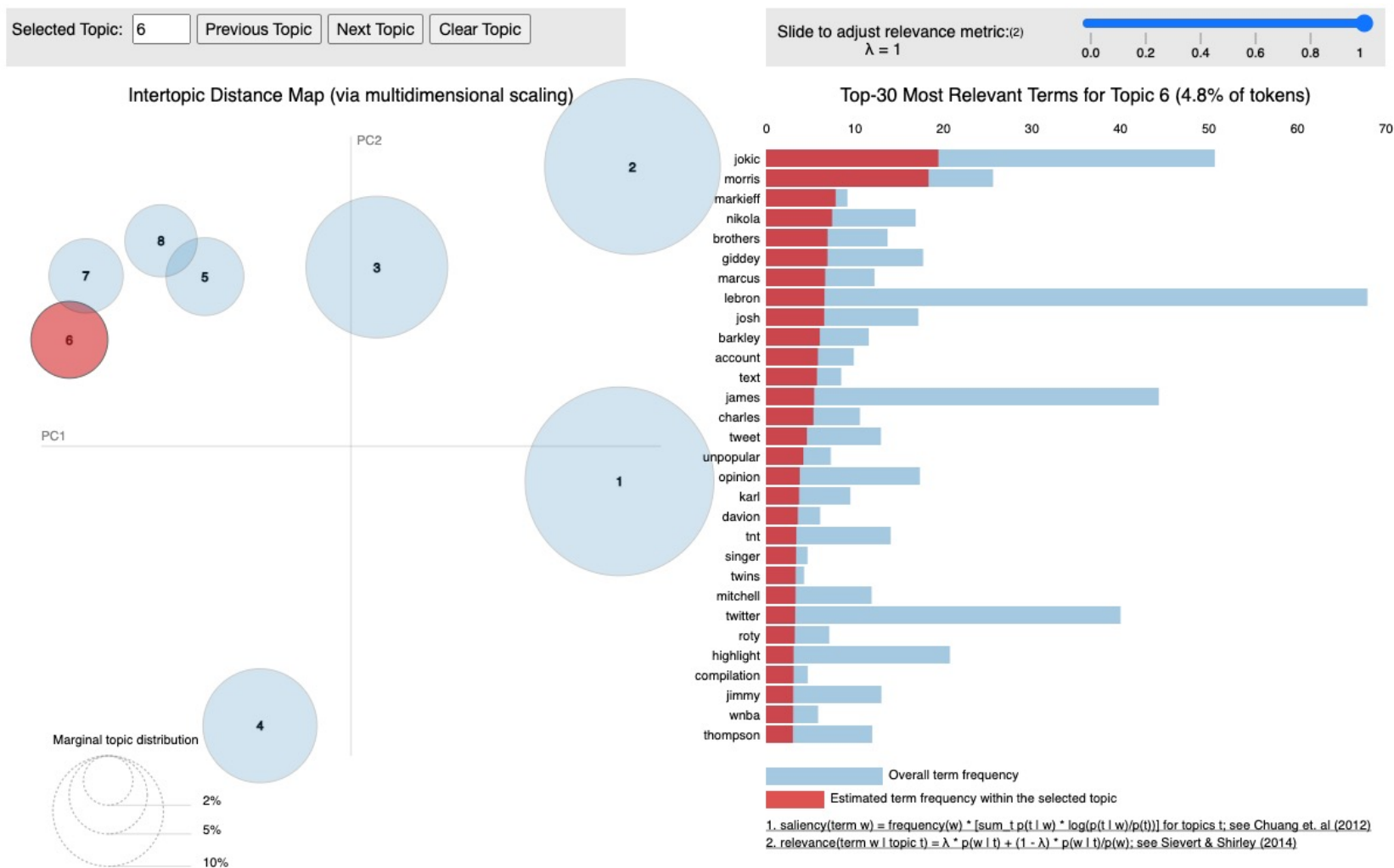
Historical discussion (fueled by 75th Anniversary List)

Looking Ahead

- Data: Take posts from a longer date range, filter team names
- CorEx to take advantage of domain knowledge and use anchoring
- Models: Advanced Clustering, Recommendation Systems, Supervised (Classification, Regression)

Questions?


Appendix



Appendix: Cont.





Appendix: Cont.



NBA
r/nba


Joined


PostsPredictionsWikiRules & GuidelinesTools▼Old Reddit

Create Post

HotNewTop...


↑ 40 ↓



 **Daily Discussion Thread + Game Thread Index** Index Thread

Posted by **[NBA] Scott Foster** u/NBA_MOD 20 hours ago 

45 ...


↑ 53 ↓



 **[SERIOUS NEXT DAY THREAD] Post-Game Discussion (November 10, 2021)** Discussion

Posted by **[NBA] Scott Foster** u/NBA_MOD 21 hours ago  

92 ...


↑ 6.5k ↓

 **[Charania] Ben Simmons believes the 76ers' tactics are worsening his mental health instead of supporting him, and want to deem him mentally fit in order to force him to play. Quotes from his agent, Klutch Sports CEO Rich Paul, and response from team officials** twitter.com/ShamsC... News

Posted by **Washington Bullets** u/GuyCarbonneauGOAT 8 hours ago  3  2

2.2k ...

↑ 1 ↓

 **Stream top-25 action from Kansas, Baylor, Auburn, Virginia, Houston, and Alabama—all on ESPN+**


u/ESPNplus · Promoted


0 ...

<https://plus.espn.com>

Sign Up


↑ 1.2k ↓




 **[Highlight] A Skirmish breaks out in the backcourt between Gobert and Turner** streamable.com/3n25b0 Highlight

Posted by **[UTA] Joe Ingles** u/NitroXYZ 5 hours ago 

578 ...


↑ 1.6k ↓


 **[Post Game Thread] The Toronto Raptors (7-6) defeat the Philadelphia 76ers (8-5), 115 - 109** Post Game Thread

Posted by **[MIA] Udonis Haslem** u/loonspace 6 hours ago  3  2 

368 ...

↑ 1.3k ↓

 **[Highlight] OG Anunoby drives and kicks it to Gary Trent Jr., who drills a tough triple in the clutch. Anunoby then gets a rebound on the other end, brings it up the court and runs almost exactly the same drive and kick play to find VanVleet a triple, who celebrates the shot with a... unique taunt.** streamable.com/bp110f Highlight

Posted by **Tampa Bay Raptors** u/twistedlogicx 6 hours ago 

138 ...

Appendix: Cont. - Data Collection/Cleaning

- PSAW (Reddit API)
- Submissions
- Group by Users and Combine into Documents
- Clean Documents
- Vectorize (CountVectorizer, TF-IDF)

