# FOURSQUARE AWARE

Prashant Gupta  Ricky Hopper  Divya Jain
Nikhil Nayak  Sandhya Nayak

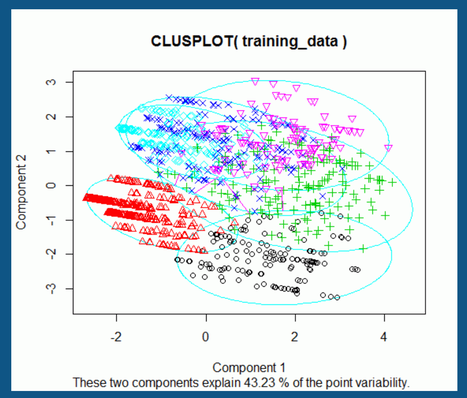
Recommendation systems are a prominent application of big data, used everywhere from targeted advertising to online dating. We clustered users to make these systems more efficient and give users easy access to new venues based on their individual interests.

We first receive a user's data and category preference, and aggregate their check-in counts by venue category.

We then group the user into a cluster of similar users, reducing the user pool we must compare to.

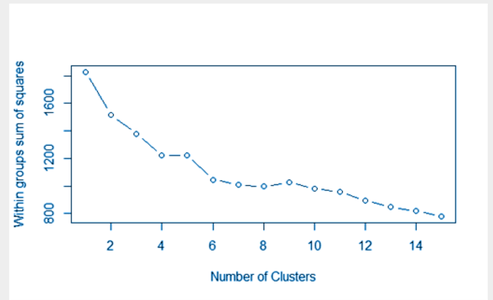
Finally, we compare the most similar user's top venue choices against the input user, and recommend the top 3 unvisited venues with the highest score.

Preprocessing: Using Foursquare's API, we grouped check-ins to individual venues by category for easier comparisons, and scaled these values down to indicate relative frequency.

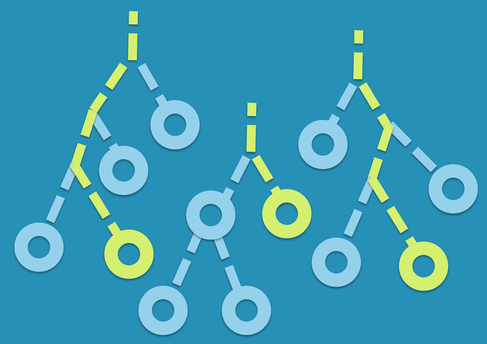



PCA after running k-means w/ 6 centers

Finding the optimal number of clusters (vs. sum of squares error)




We fit querying users into clusters using a decision tree generated during our clustering.



Our dataset was a collection of 227,428 Foursquare check-ins from 1,083 users in New York City over 10 months.

We used cosine similarity to find similar users after considering a number of different metrics and optimizing system performance.

Our system successfully returns venue recommendations based on similar users' preferences. In the future, this system can be extended by factoring in timestamp data to recommend trending locations, or finding new friends at suggested venues.