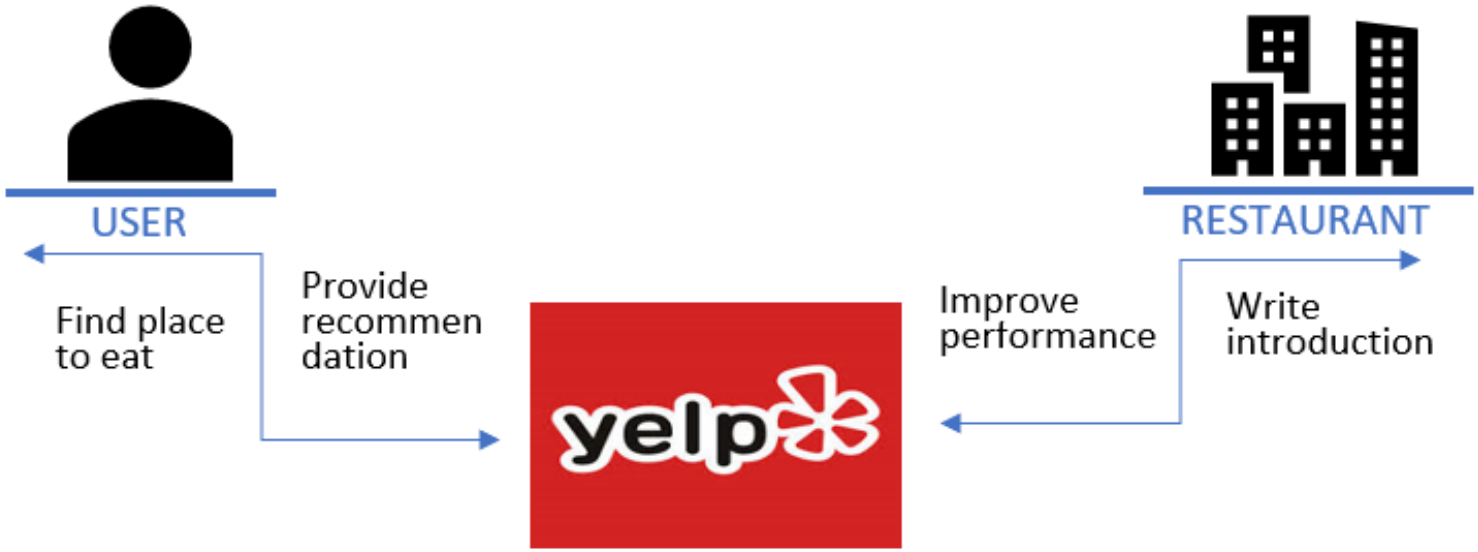
User Recommendation System

Canyang Liu, Jia Liu, Wenjia Xie

STAT 479

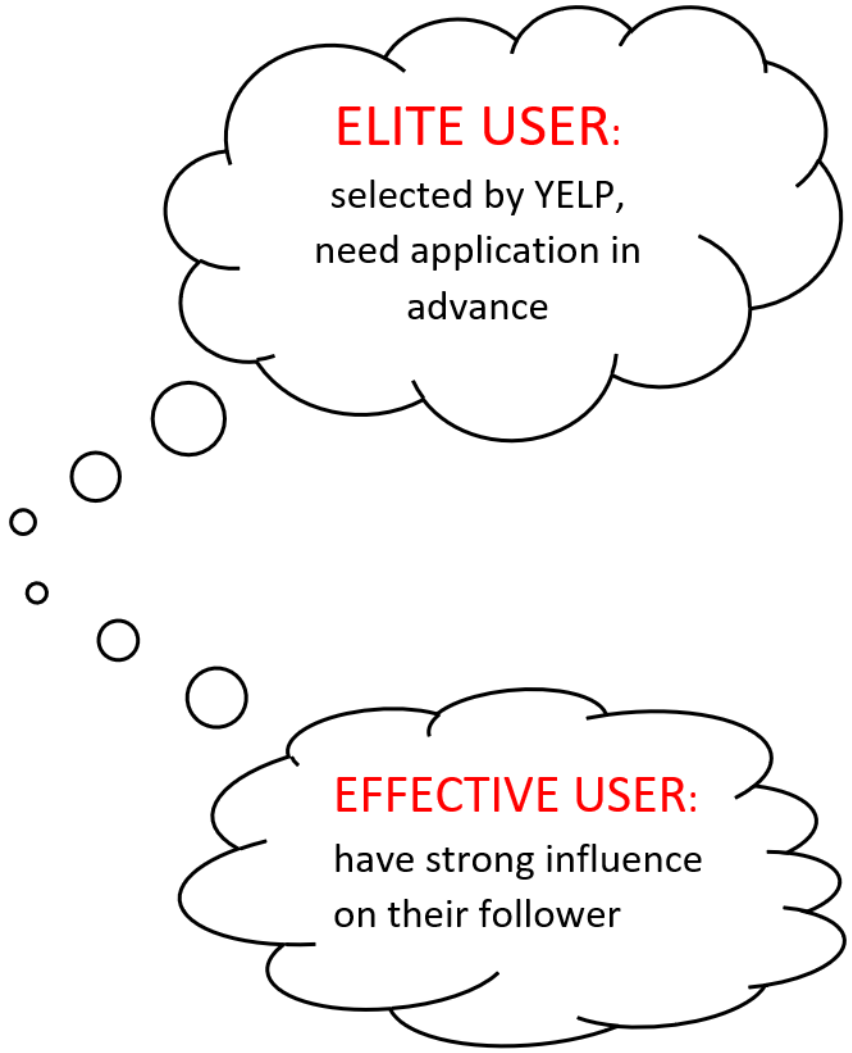
April 26, 2018

Brief Introduction

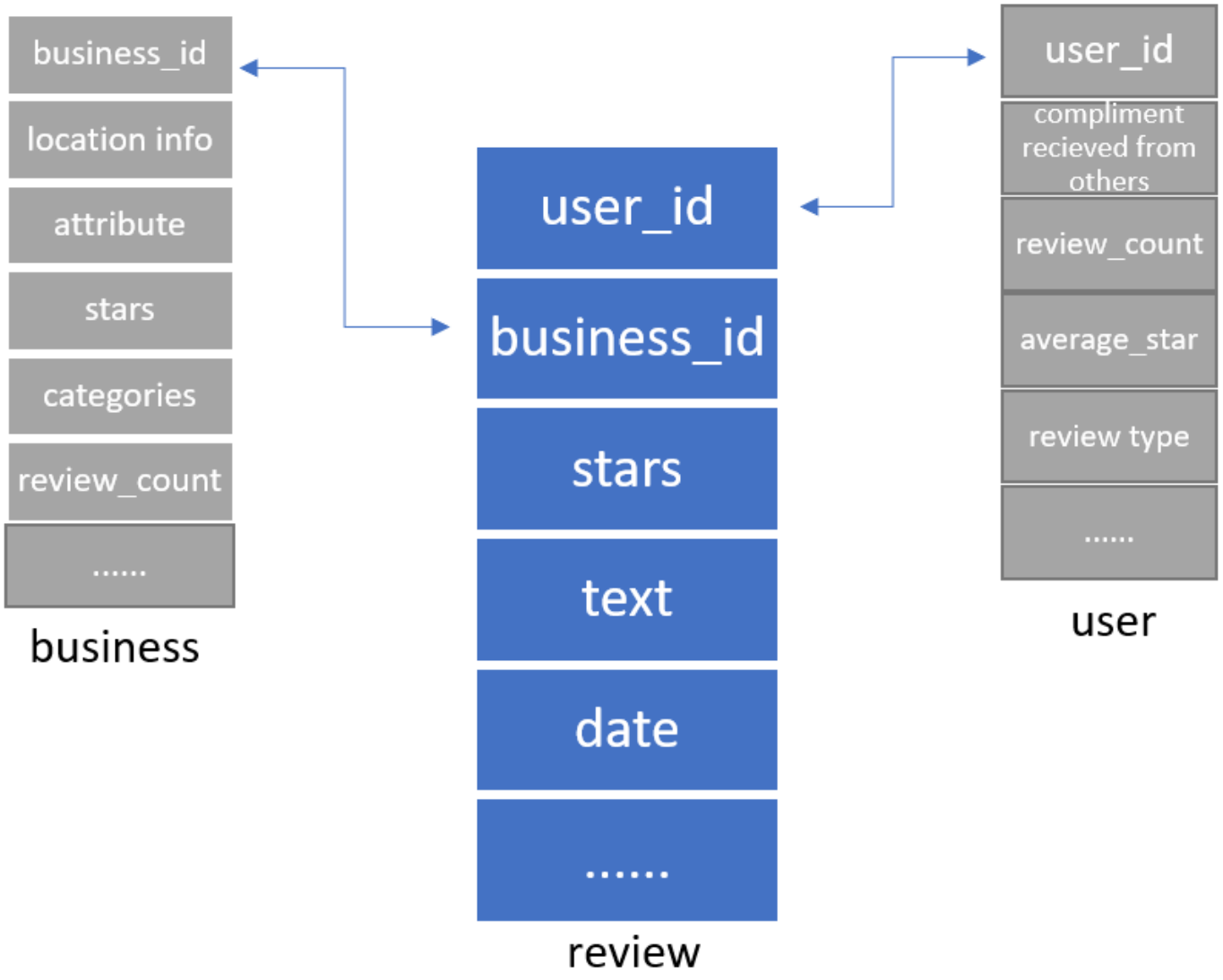


Brief Introduction

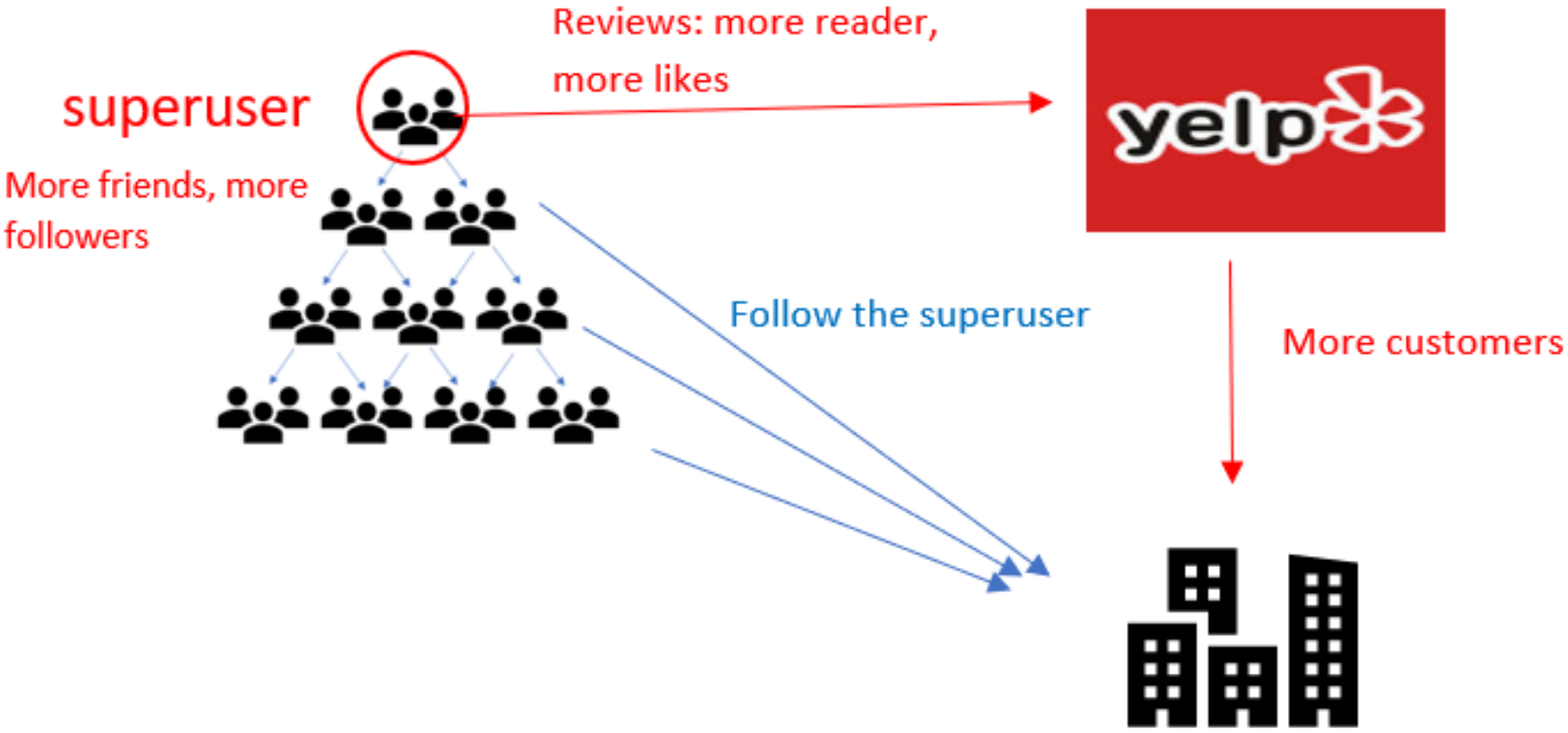
Who is superuser?



Finding influencial users

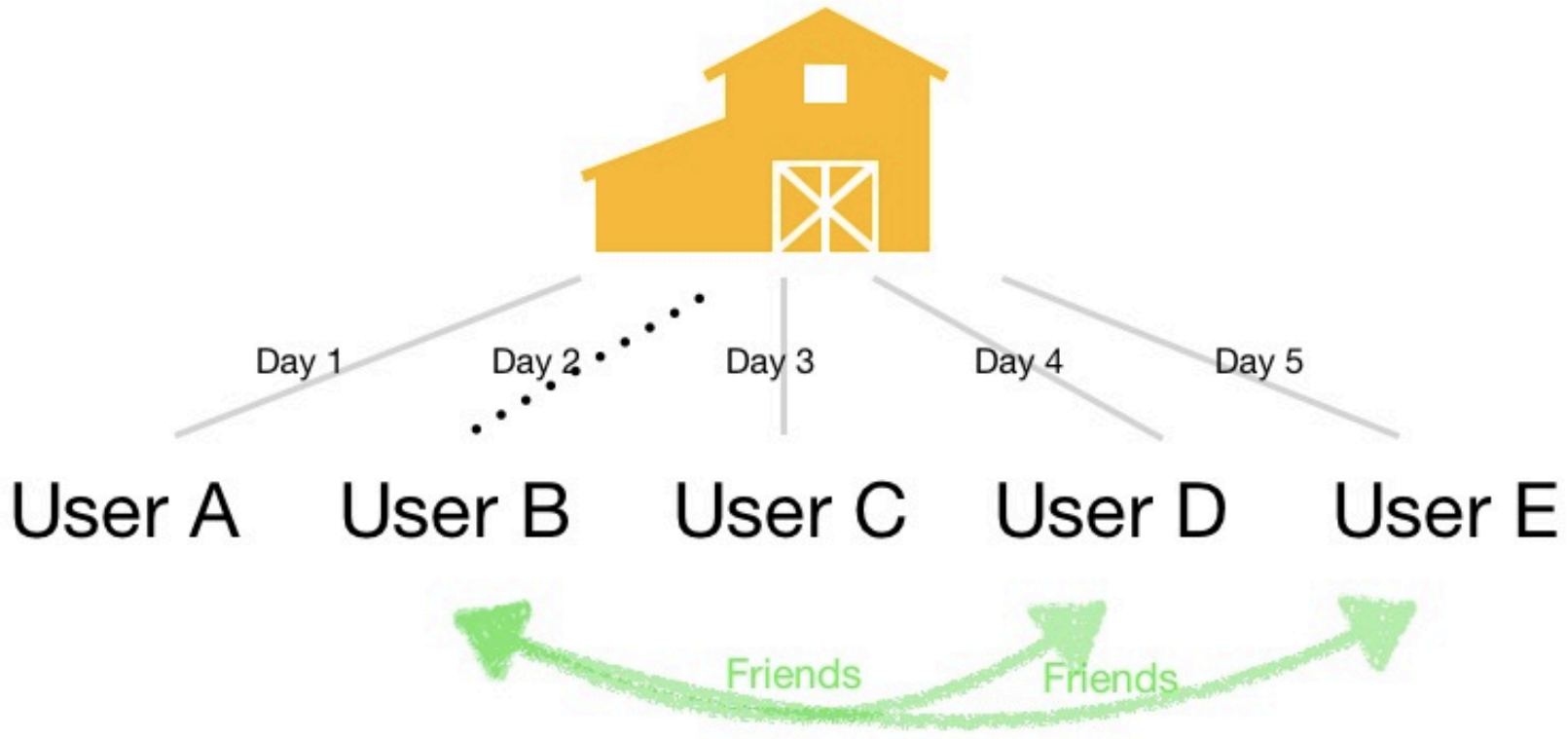


Finding influencial users



Finding influencial users

Influencial users: have influence on their friends.



Finding influential users

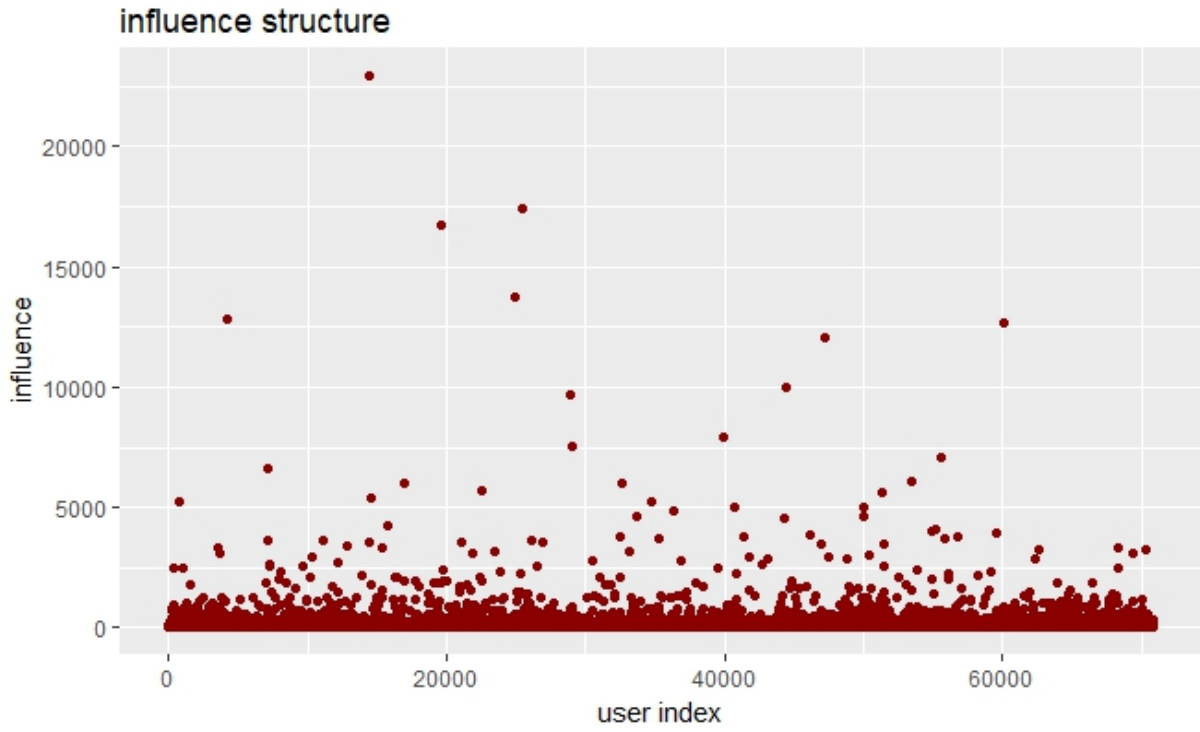
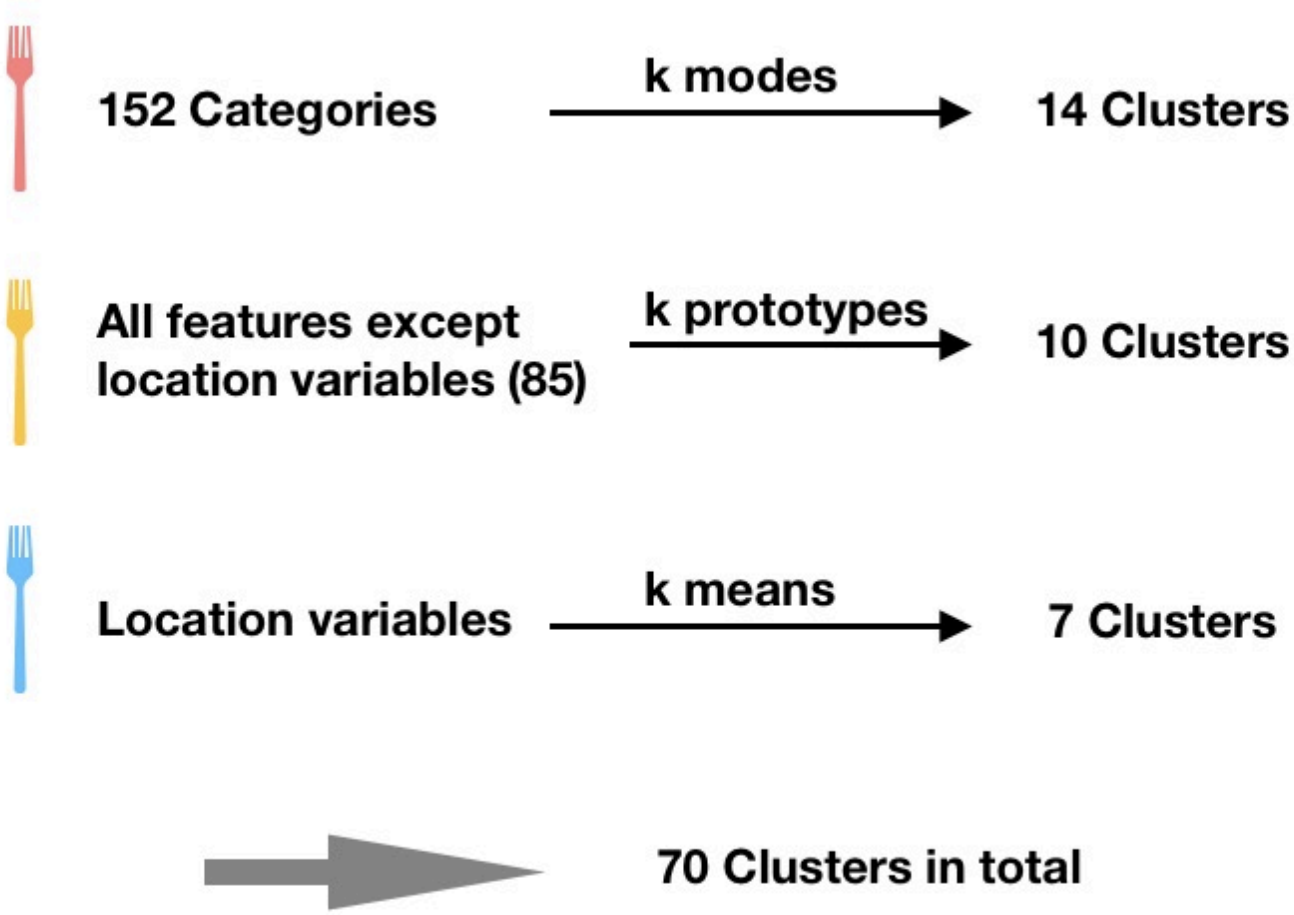


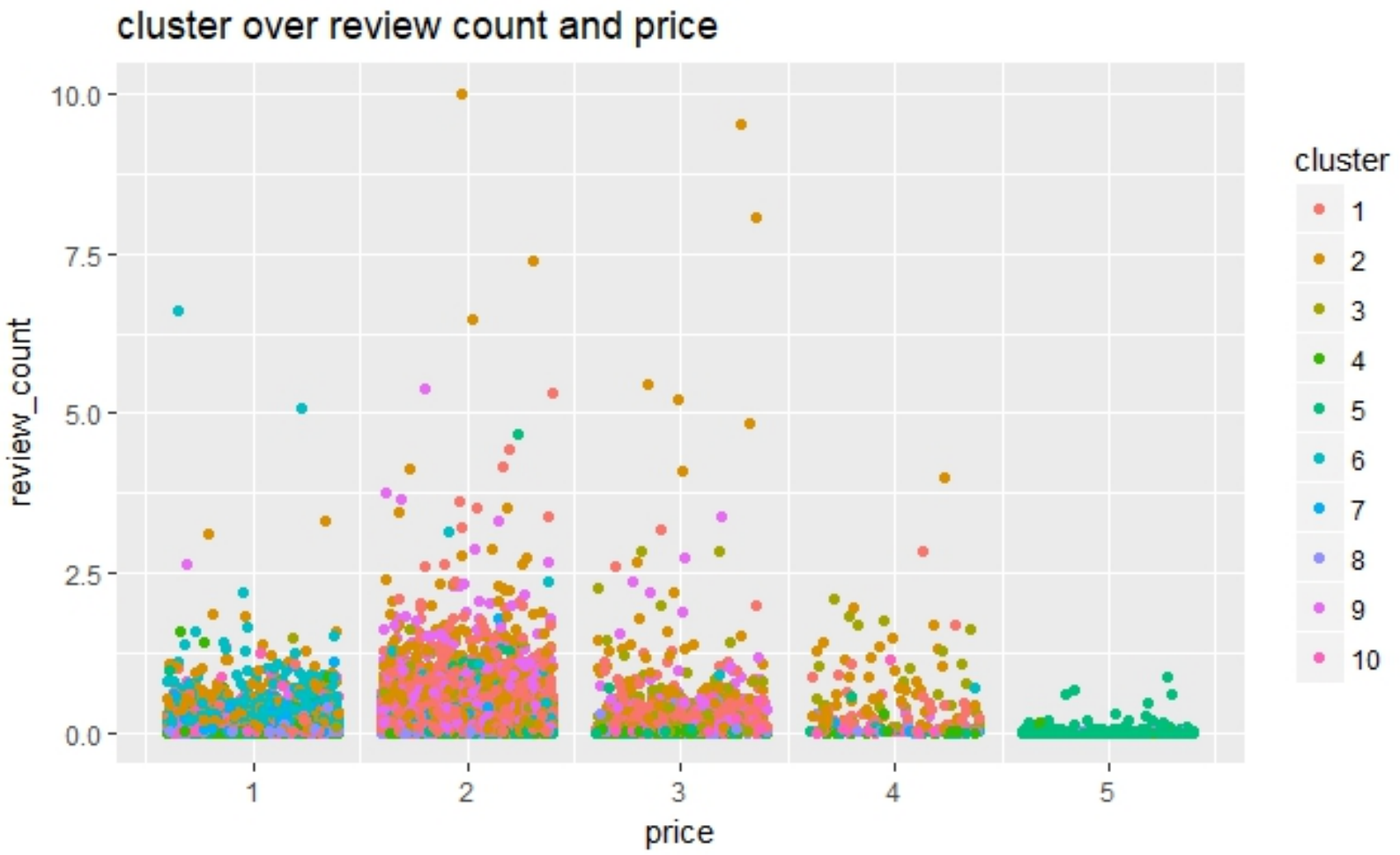
Table: The effection of influential users

Min.	1st Quantile	Median	Mean	3rd Quantile	Max
1.00	1.00	2.00	23.48	5.00	22984.00

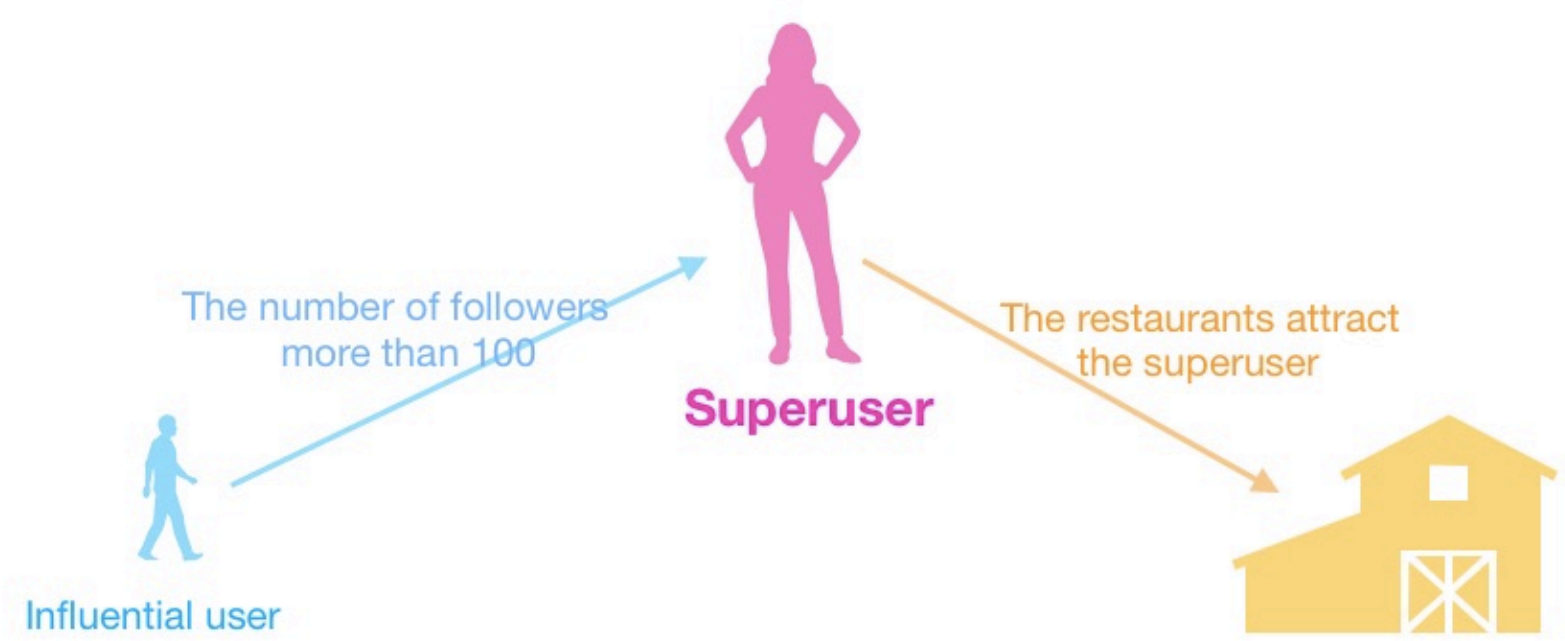
Classify Restaurants



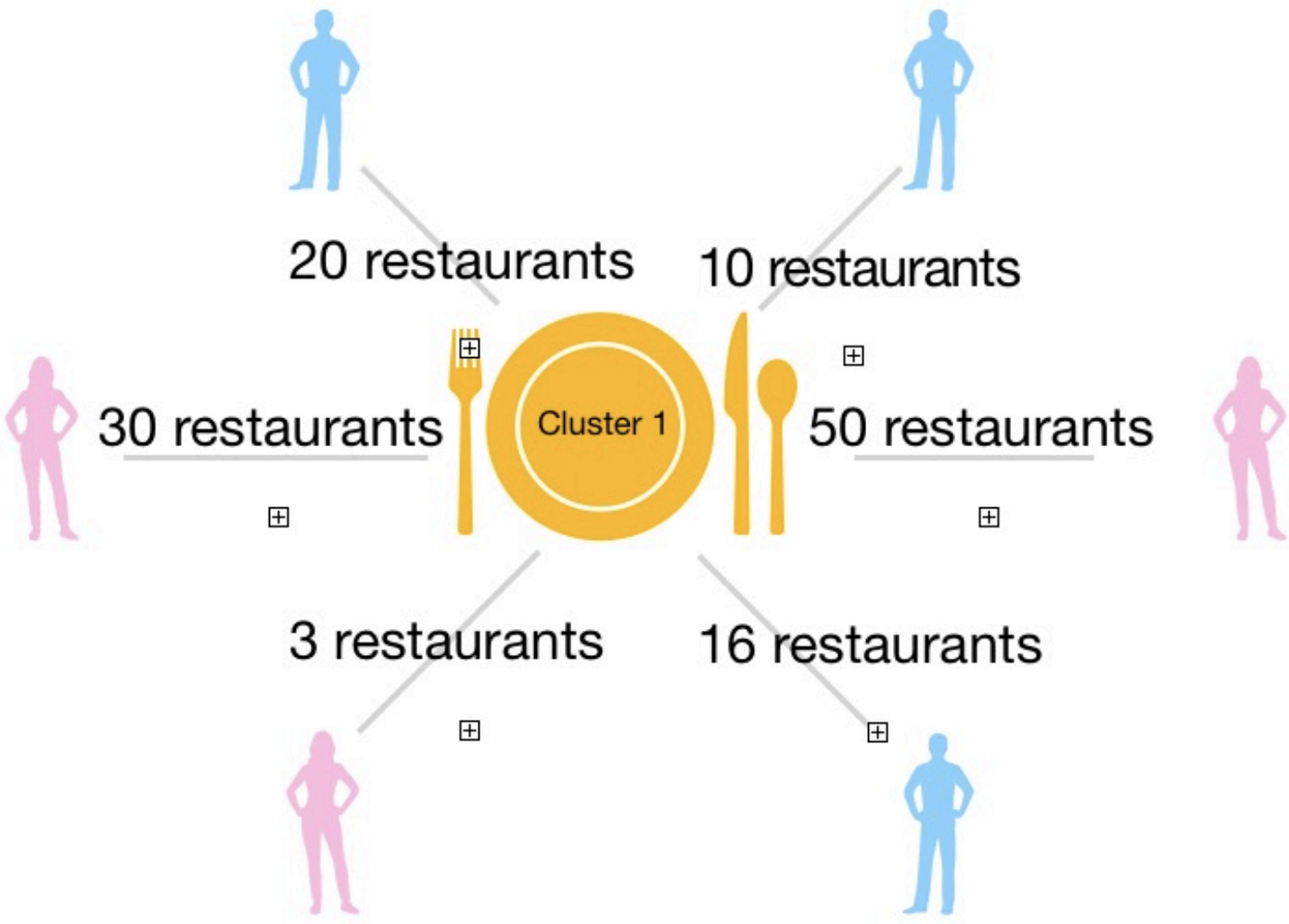
Classify Restaurants



Finding superusers



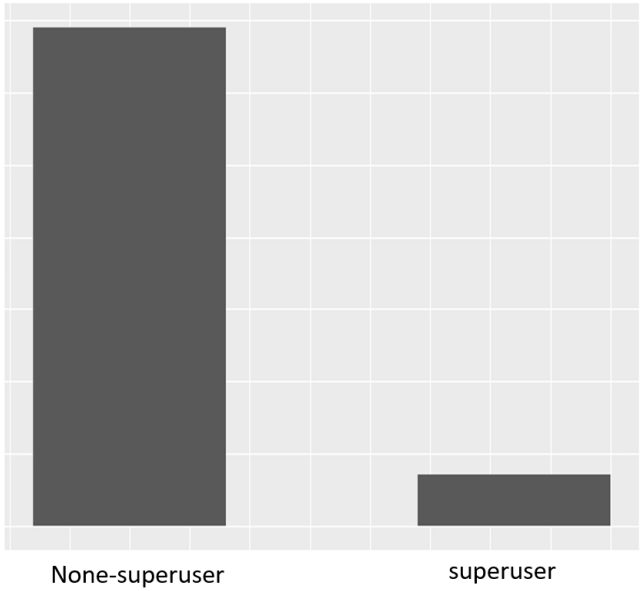
Finding superusers

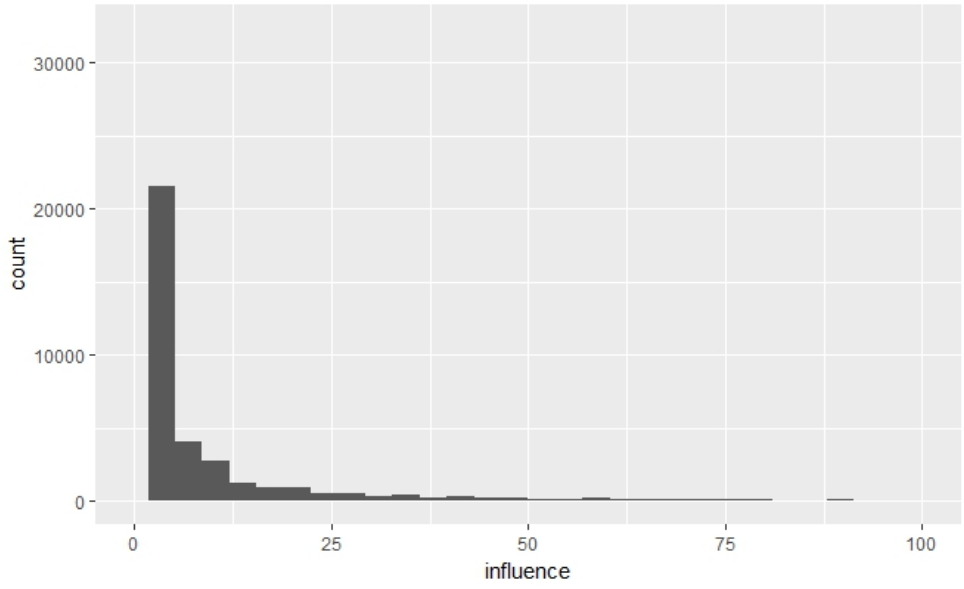


Finding superusers

https://jliu647shiny.shinyapps.io/479proj/

How to become a superuser





How to become a superuser

Table: Logistic Regression Model

(Intercept)	fans	complimet writer
-6.3	-1.2×10^{-3}	2.2×10^{-3}
	compliment profile	compliment funny
	-7.8×10^{-3}	1.4×10^{-3}
	compliment plain	friend number
	-1.9×10^{-3}	1.5×10^{-3}
	review count	
	3.1×10^{-3}	

How to become a superuser

Table: Ordinary Linear Regression Model using Box-cox

(Intercept) -3.2	compliment photos 4.0×10^{-4}	compliment funny 6.3×10^{-4}
compliment hot -4.1×10^{-4}	compliment writer -2.6×10^{-3}	compliment plain -7.7×10^{-4}
compliment list -1.7×10^{-2}	compliment more 1.7×10^{-2}	useful 5.3×10^{-5}
cool -7.0×10^{-5}	fans -3.4×10^{-3}	average stars 1.2×10^{-1}
friend number 2.6×10^{-3}	review count 6.7×10^{-3}	

How to become a superuser

Table: Ordinary Linear Regression Model using Box-cox

(Intercept) 9.4×10^{-4}	compliment list -8.8×10^{-4}	compliment funny 7.6×10^{-4}
compliment plain -6.4×10^{-4}	compliment more 9.9×10^{-4}	compliment hot -1.4×10^{-4}
compliment profile -6.9×10^{-4}	review count 1.7×10^{-3}	friend number 7.7×10^{-4}
fans 6.2×10^{-4}	funny 2.8×10^{-5}	cool -2.5×10^{-5}
average stars 6.6×10^{-3}		

Model Comparison

Coindicence ratio	Logistic Model	v.s.	OLS	: 84%
	OLS	v.s.	NB regression	: 85%
	Logistic Model	v.s.	NB regression	: 82%

Accuracy	Logistic Model	54%
	OLS	50%
	NB regression	45%

QUESTION ?