

# How do scholars collaborate with each other?

(Comparative study on Co-authorship Networks of  
Scholars in 19 universities worldwide)

Ali (Aliakbar Akbaritabar)

PhD student of Economic Sociology, University of Milan

[Aliakbar.akbaritabar@unimi.it](mailto:Aliakbar.akbaritabar@unimi.it)

@akbaritabar

# research question(s)

## Big questions:

- how scientific community works?
- how science is being produced?  
How scholars interact and collaborate in process of science production?
- How scholars connect to each other to build up their scientific career?
- Is there anything specific to sociology and sociologists in this?

## smaller questions

- Descriptive questions (this report):
  - Does the collaborations change over time? How?
  - Does collaborations differ in different universities and academic contexts & fields? How?
- Explanative and more theory oriented questions (future steps):
  - is scientific collaborations being affected by scholars social capital?
  - can we consider scientific collaborations as social capital of scholars?

# Literature reviewed

Xu, et al.  
Journal of the Association  
for Information Systems  
Vol. 15, Issue 12,  
pp. 835-859,  
December 2014

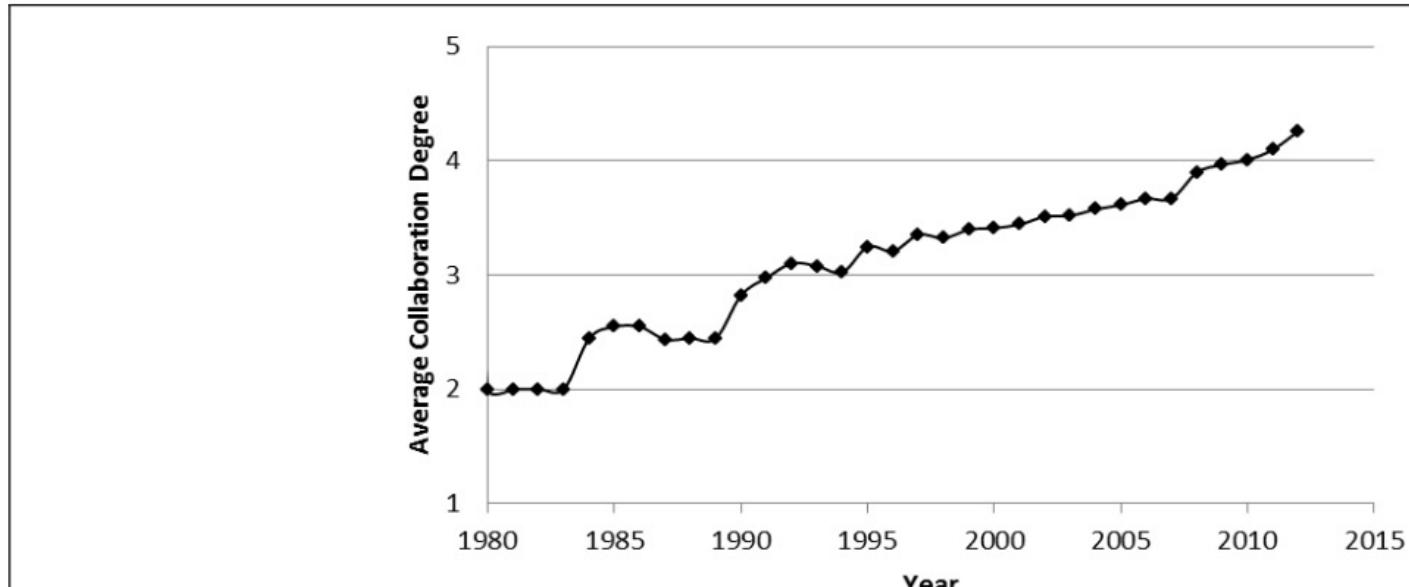
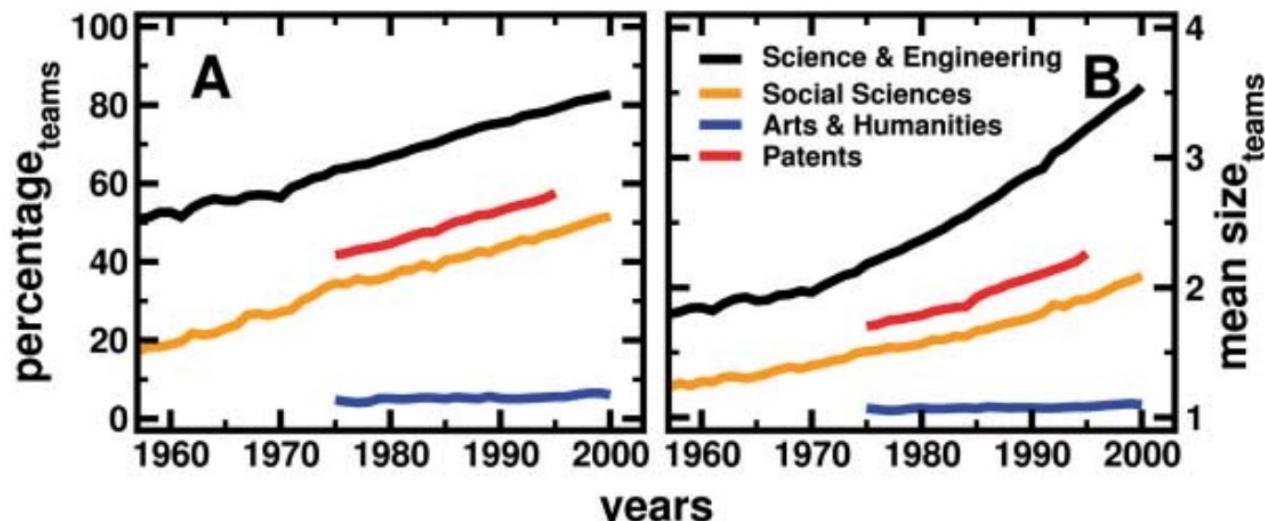


Figure 2. Changes in Average Degree Centrality Over Time (in Yearly Networks)



Wuchty, et al.  
Science 316 ,  
1036 (2007)

Fig. 1. The growth of teams. These plots present changes over time in the fraction of papers and patents written in teams (A) and in mean team size (B). Each line represents the arithmetic average

# Literature reviewed

- Becher and Trowler (2006; 124); **laboratory and facilities needed in different fields** (like **physics**) and their effect on collaborations of necessity of team work (in contrast to situation of **mathematics** and **sociology**; who are known to usually be solo-writers)
- In social sciences, fields closer to natural and hard sciences (like **psychology**) have features similar to them in collaboration patterns
- New policies like **horizon 2020** and **COST projects** to foster collaborations rather than single researches;
  - possible question: how much it has been successful (not just based on names, in action, based on the papers and presentations ... published based on those cost projects ...)

Examples of affecting variables (based on Tie formation versus tie persistence in scientific collaborations Dahlander and McFarland (2013))

### Independent variables

- Organizational foci
- Status and value homophily
- Cumulative advantage
- Triadic closure
- Tie inertia
- Means-ends rationalization

### Dependent variables

- Scientific collaborations
  - Tie formation
  - Tie persistence
  - (they have seen both papers that has been published and grants they have applied for, either successful or not)
- Control variables (tenure status, gender, and ethnicity, Exclusion restriction Appointment year difference)

# Design : case selection, methods, ...

**Cases:** all papers of each individual scholar (as ego); (unit of analysis is individual scholars not journals or papers as usual between researches reviewed)

- Rationale: the agreements between scholars about name order ...

Social network analysis (mostly with **ego-centric approach**) to see the embeddedness of scholars in their context and how they have mobilized their potentially accessible resources to get returns?

**Data:** So far google scholar data; next, web of science, PEERE project data.

co-authorship networks as a proxy of how scholars interact with each other

Checked different structural properties of ego-networks and here some are presented

# Scholar.socialvillage.me

Who are researchers mostly co-authoring with?

see how it works



Copy the link to your person of interest Google Scholar profiles below and see the co-authorships network; go here to see a video of how it works <https://goo.gl/053N2o>

google scholar profile link

<https://scholar.google.com/citations?user=BEzh-pcAAA>

Submit

Reference: data for co-authorships and publications are being extracted from google scholar with 'scholar' package for R

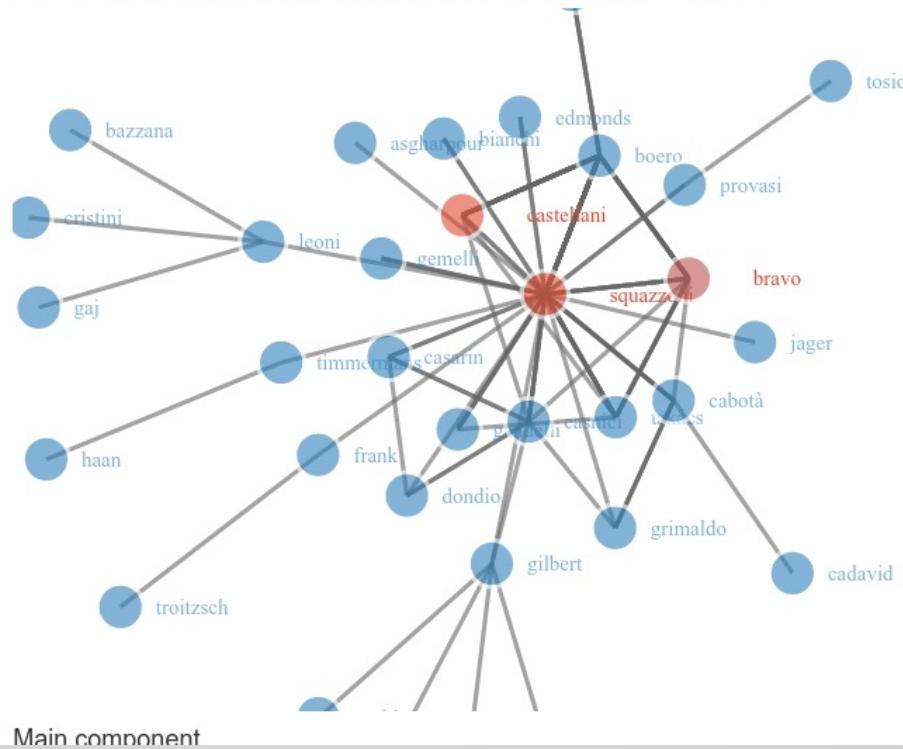
\*\*: It is the first and earliest version, some issues are known and in the process to be solved, some not, I will be happy to hear: akbaritabar[at] gmail.com

author's data

	number of articles	unique journals	oldest article	number of top journals	h-index
1	92	57	2000	0	15

Raw co-authorships Graph

use mouse scroll to zoom in and out; further details (including main component) below



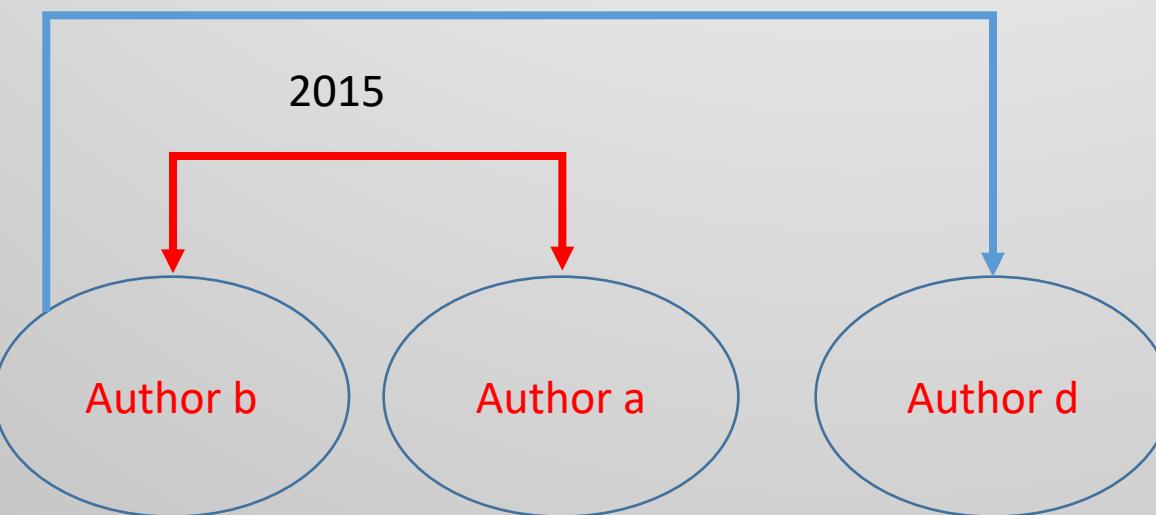
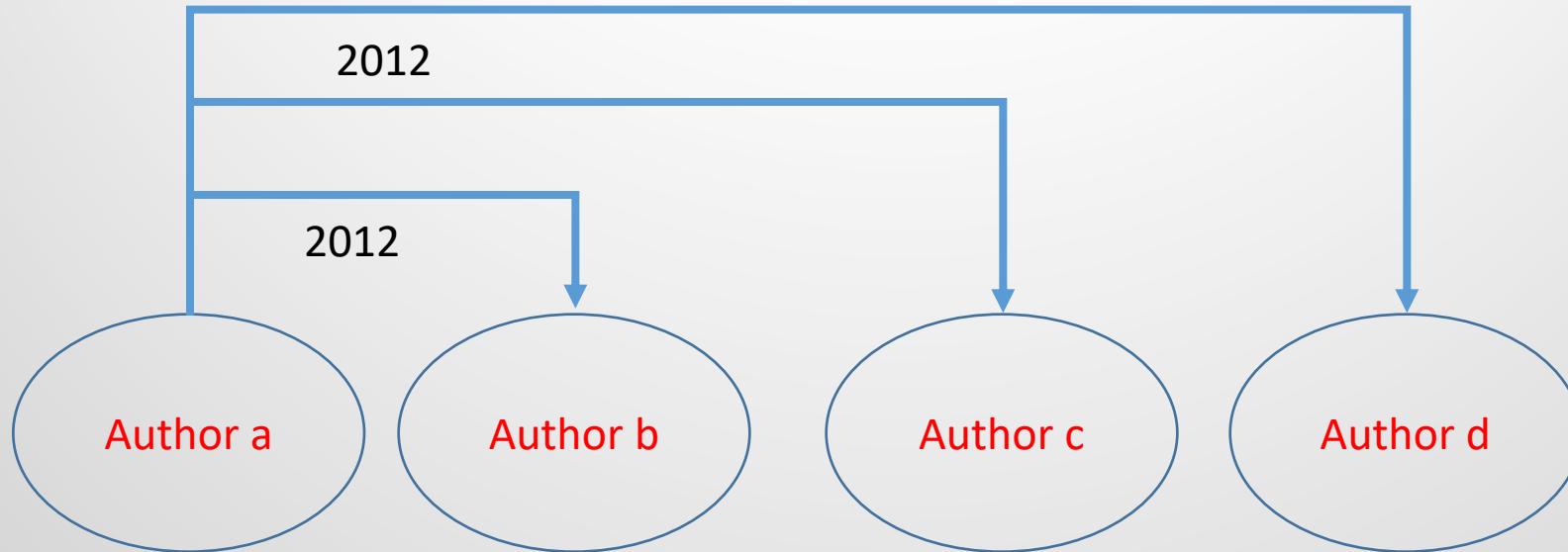
# 2 points of view to co-authorships

- **Choosing some journals and following all papers** published in them to see co-authorships (can be called sociocentric and complete network view)
- **Choosing some scholars, and following all their papers** and publications in different journals to see co-authorships (can be called ego-centric view)
  - Consider scholar like “network manager” type in Burt’s terms, someone who decides and manages his scientific collaborations through time; manages based on what? We shall see ...
- How you define your nodes and ties, changes the way you are capturing the network and model the social phenomenon under study ...
  - Revisiting the Foundations of Network Analysis, Carter T. Butts, Science 325 , 414 (2009)

# Adjacency lists based on papers' author names

Paper 1: author a, author b, author c, author d (2012) ..., ...

Paper 2: author b, author a, author d (2015) ..., ...



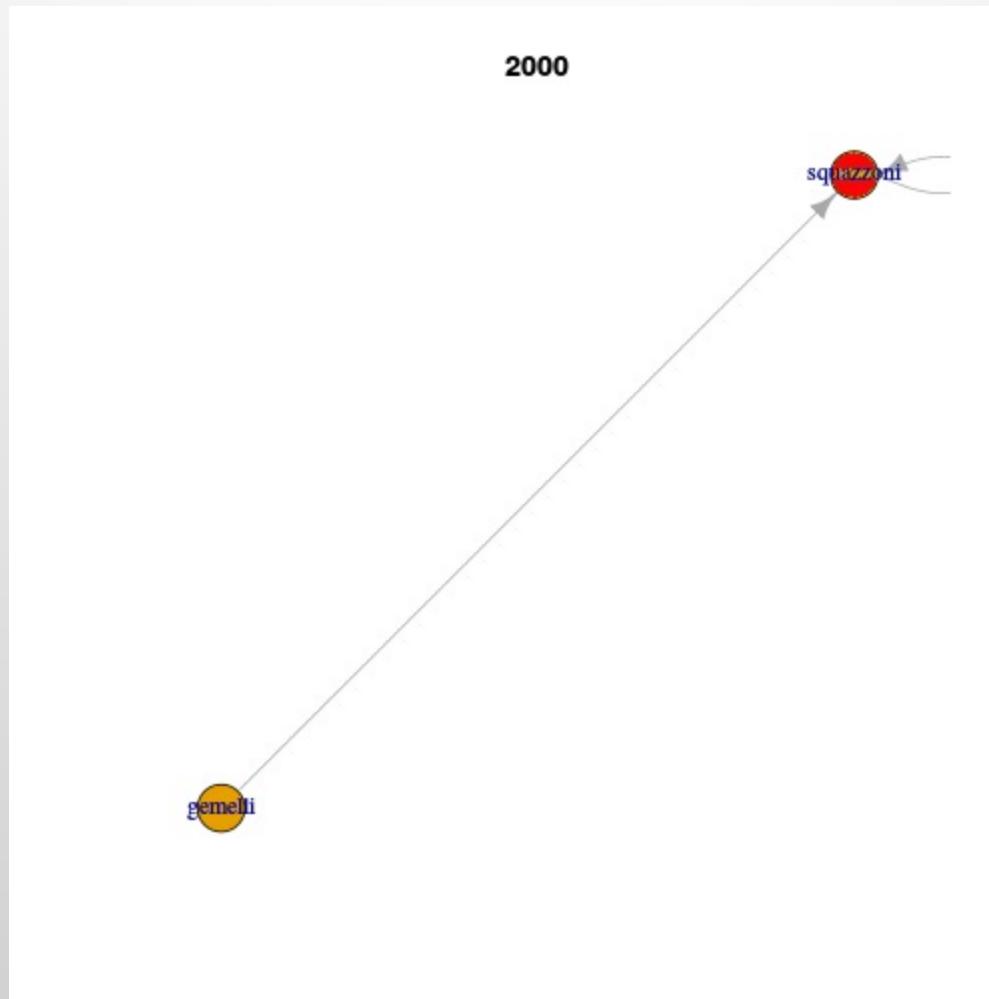
## Innovations:

- ego-centered vs. complete
- One scholar vs. one journal
- Directed vs. undirected

Ego = author b

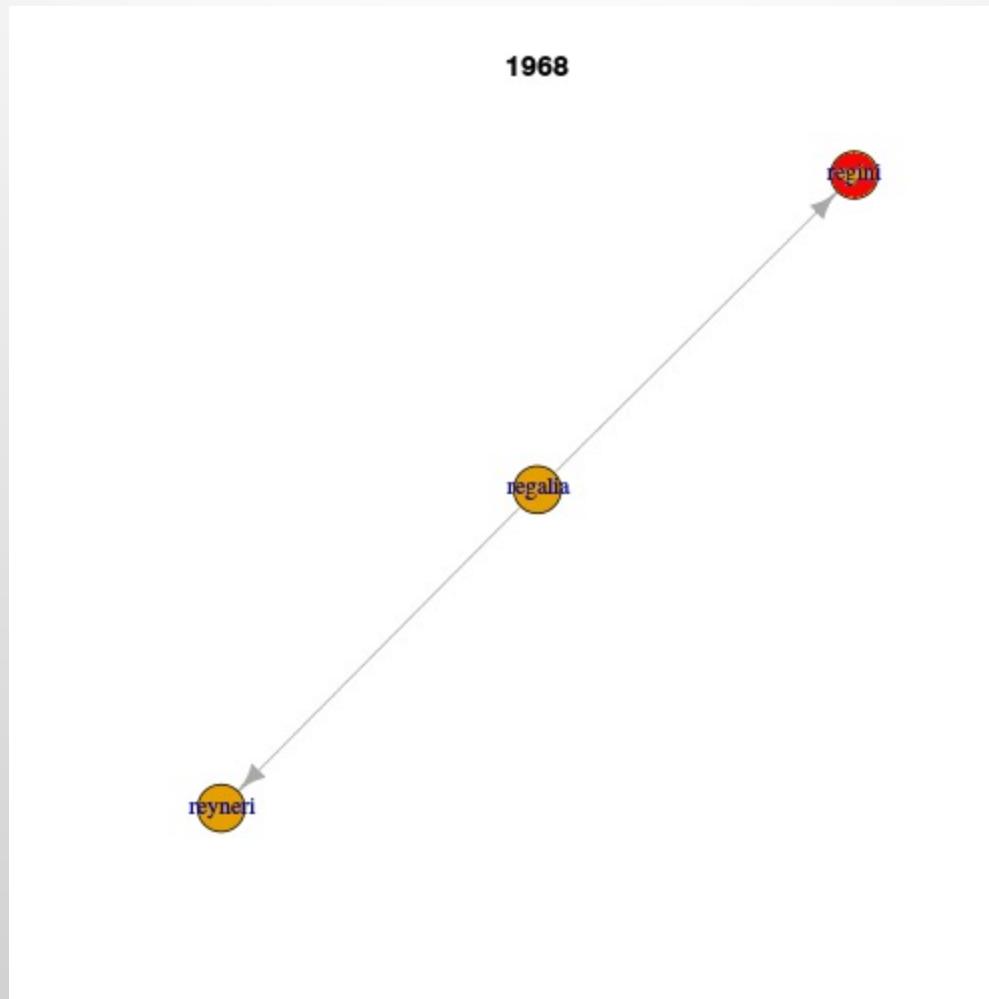
# Trend of changes in co-authorships

(middle aged, associate professor, 2000 – 2015)



# Trend of changes in co-authorships

(aged, full professor, 1968 – 2015)



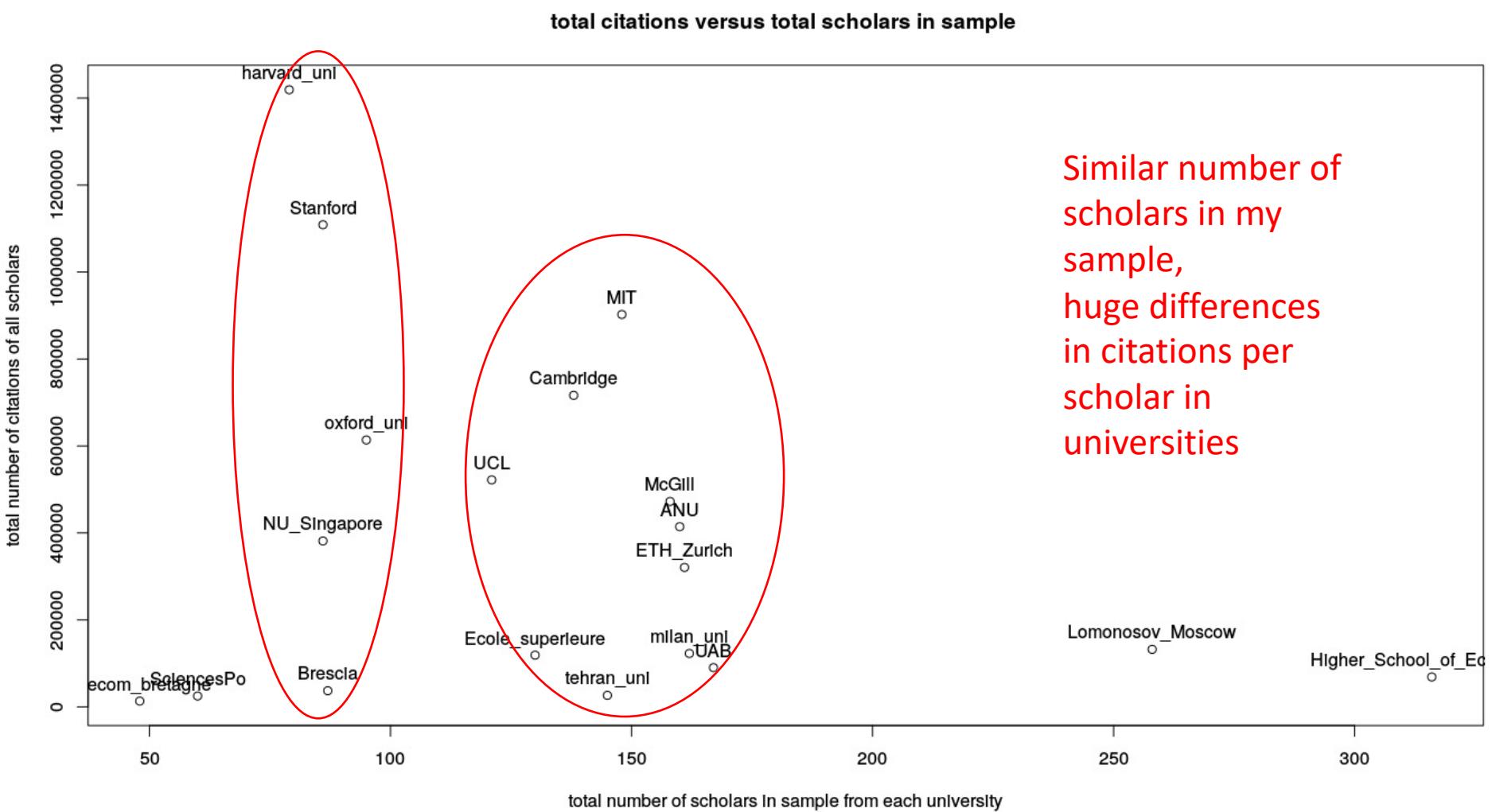
# Study sample

Sample of study	
Total sample of scholars	8373
Number of scholars <b>without name inconsistencies</b> (reported here)	2578

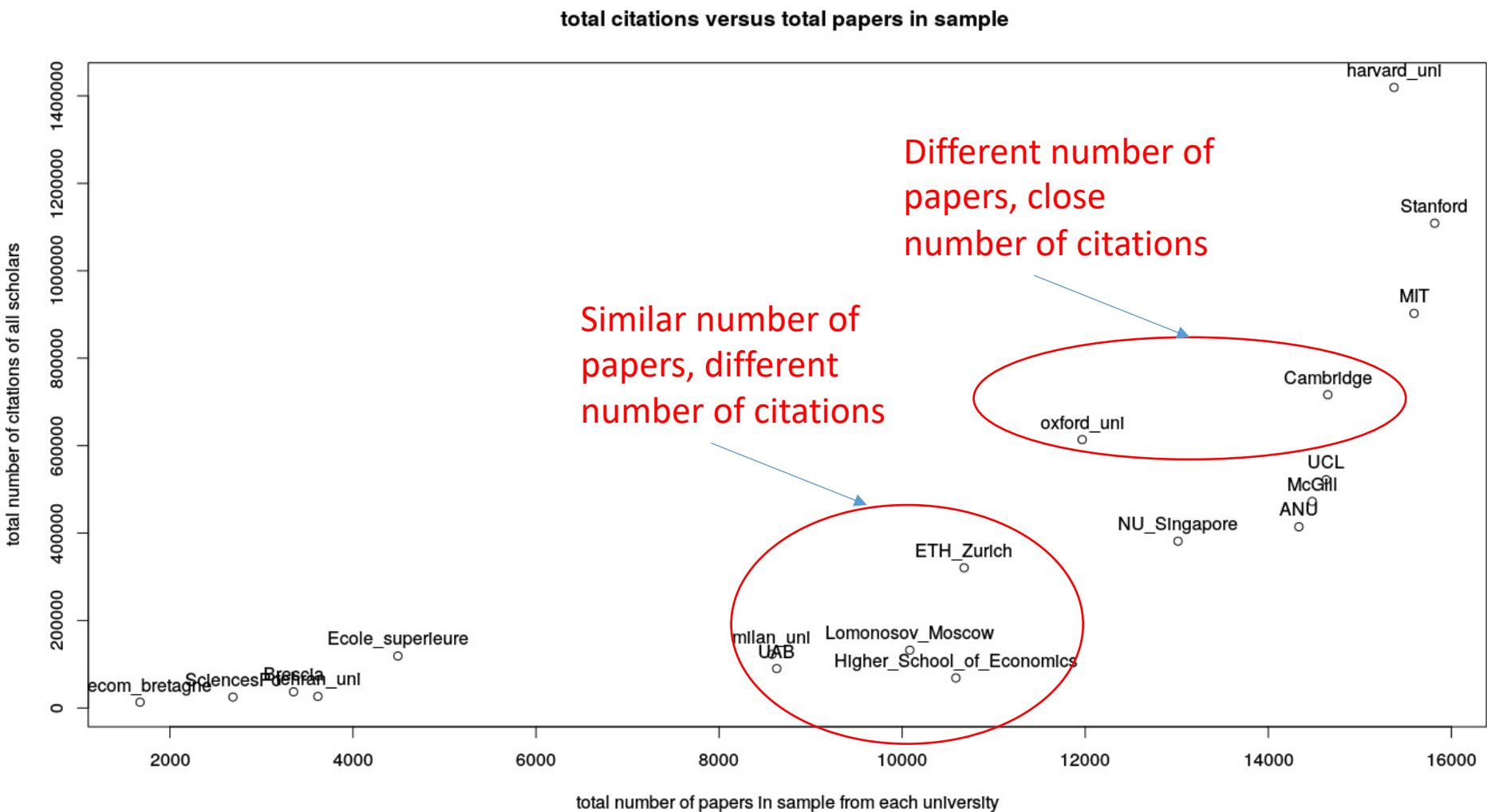
University name	Initial sample before name correction	Number of scholars in report
Higher_School_of_Economics	530	311
Lomonosov_Moscow	433	251
UAB	529	164
ETH_Zurich	530	161
Milan_(unimi)	500	161
McGill	530	158
ANU	529	158
MIT	530	148
Tehran	496	145
Cambridge	530	138
Ecole_superieure	229	126
UCL	530	120
Oxford	498	95
Brescia	223	87
Stanford	530	86
NU_Singapore	530	84
Harvard	504	77
SciencesPo	106	60
Telecom_Bretagne	86	48
<b>Total</b>	<b>8373</b>	<b>2578</b>

- Author names inconsistencies that took long to solve (took same approach to similar research like Dahlander and McFarland (2013) to solve)
- Limiting time span from 1970 to 2015 (45 years)

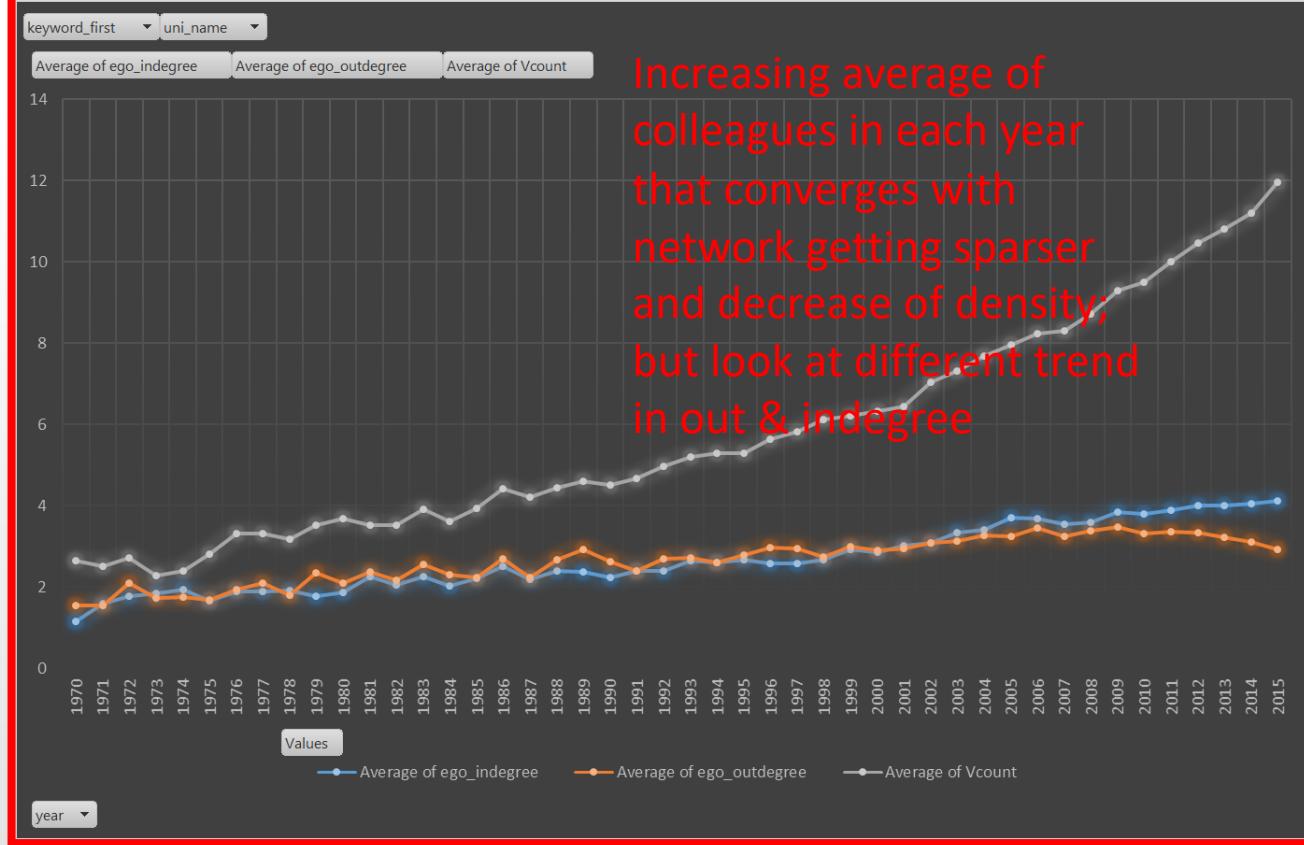
# Preliminary results



As expected, affiliation matters; Number of papers in sample vs. citations

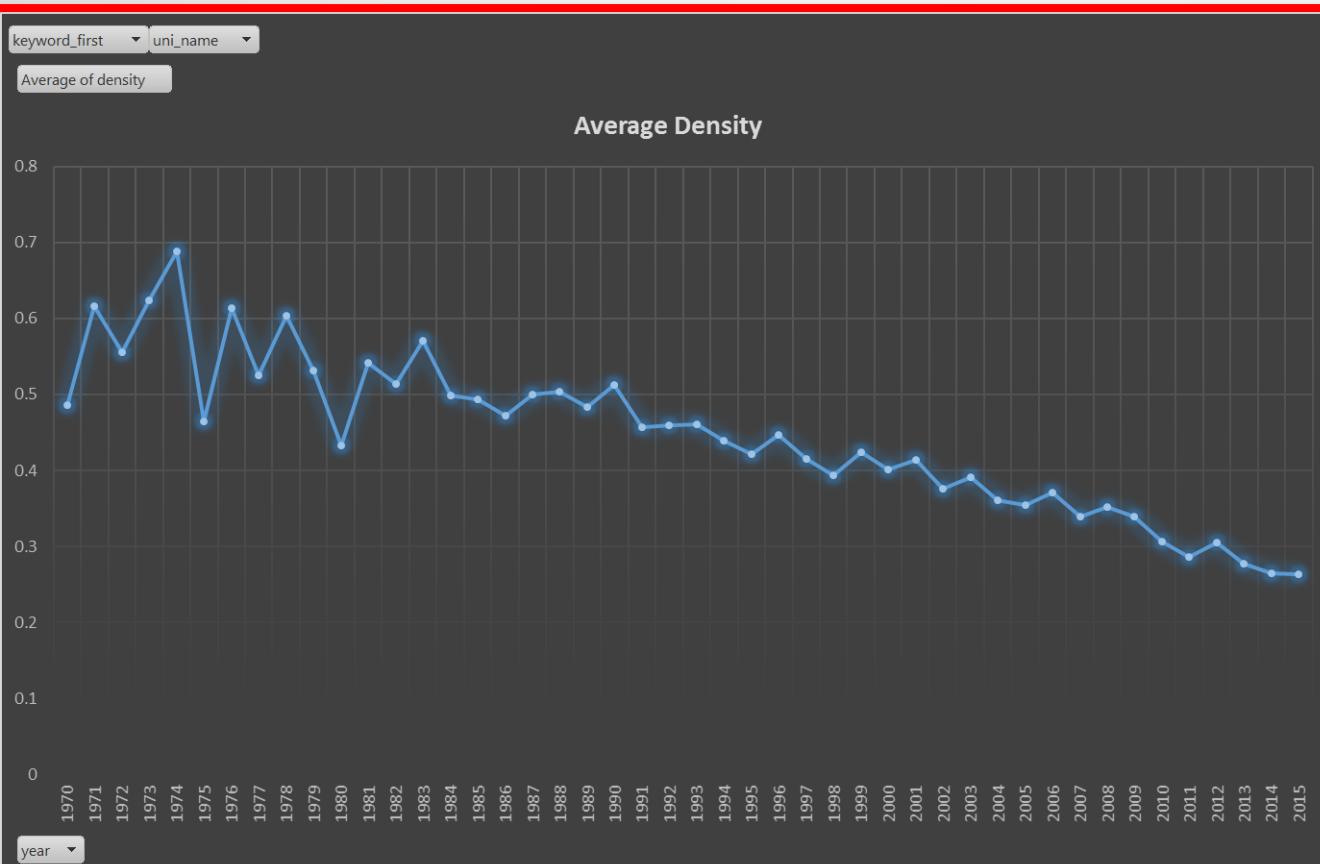


# All fields all universities

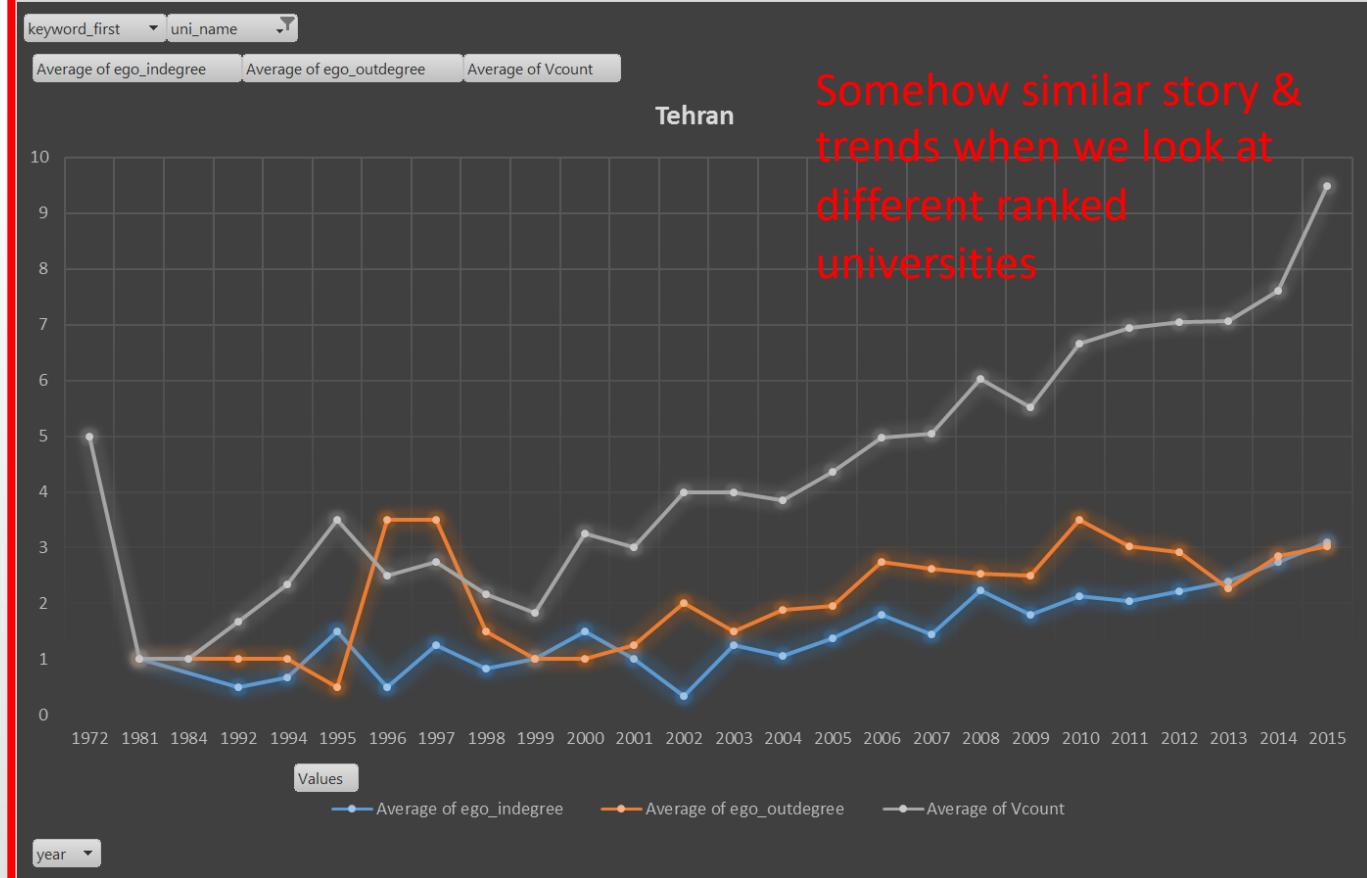


# All fields all universities

Increasing average of colleagues in each year that converges with network getting sparser and decrease of density; but look at different trend in out & indegree

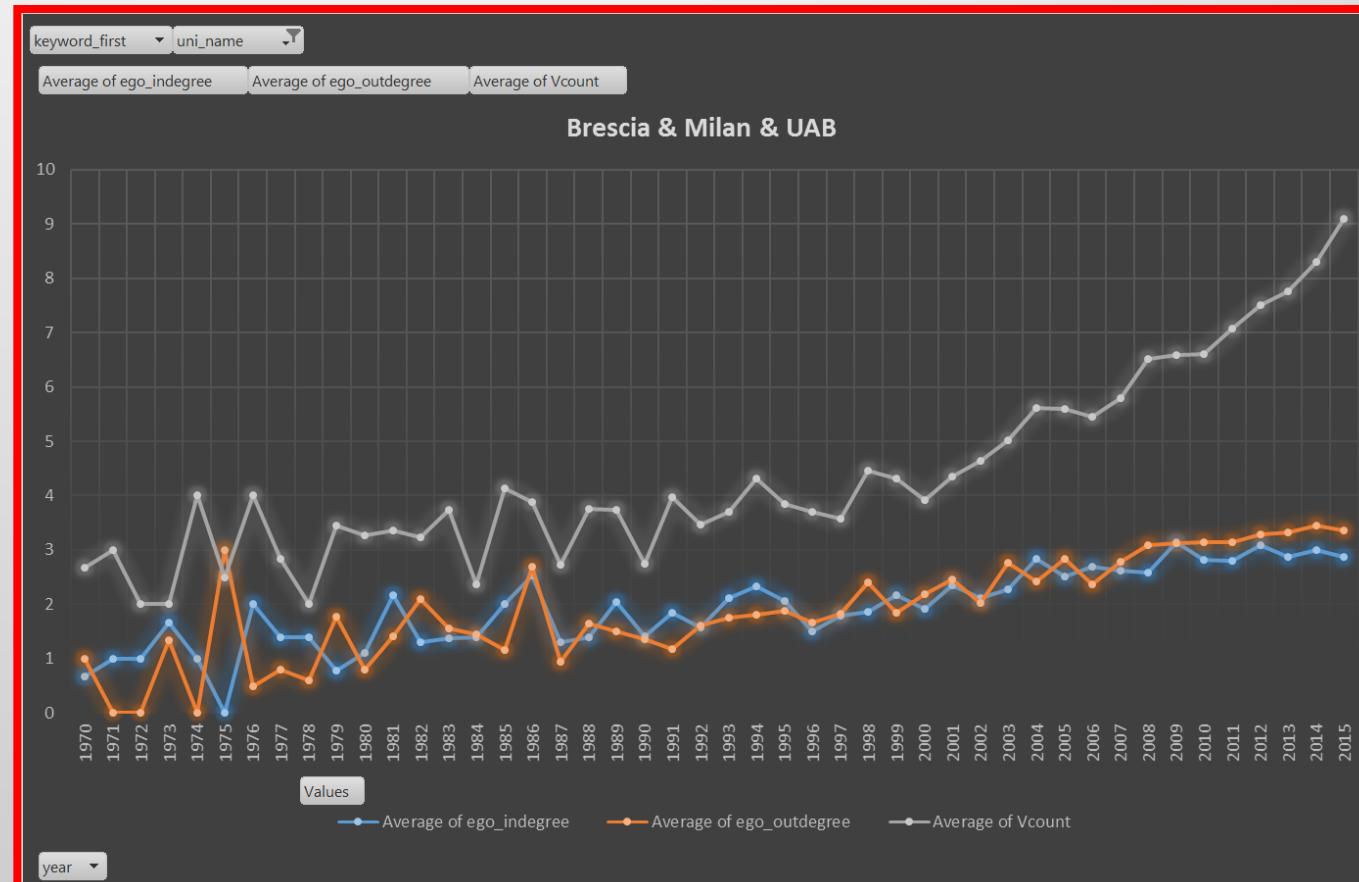


# All fields universities ranked



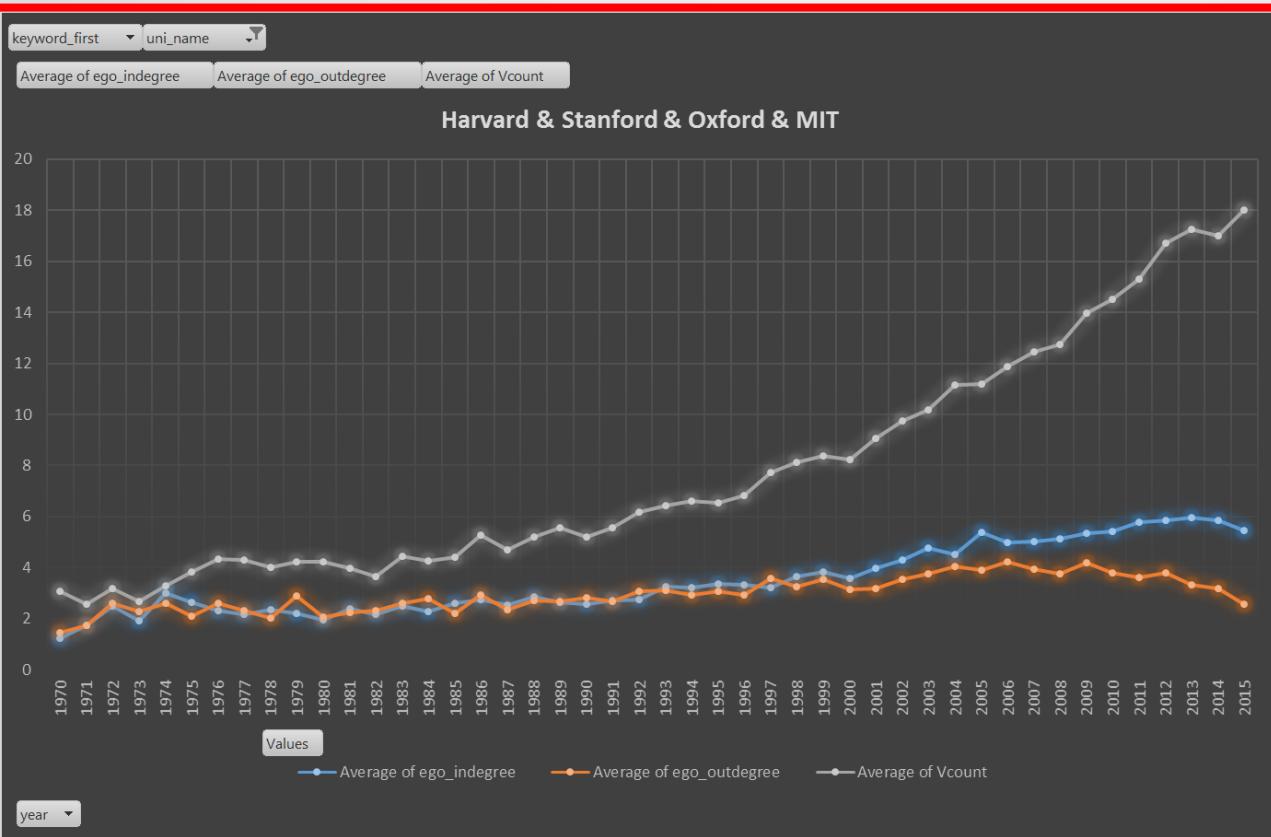
# All fields universities ranked

Somehow similar story & trends when we look at different ranked universities

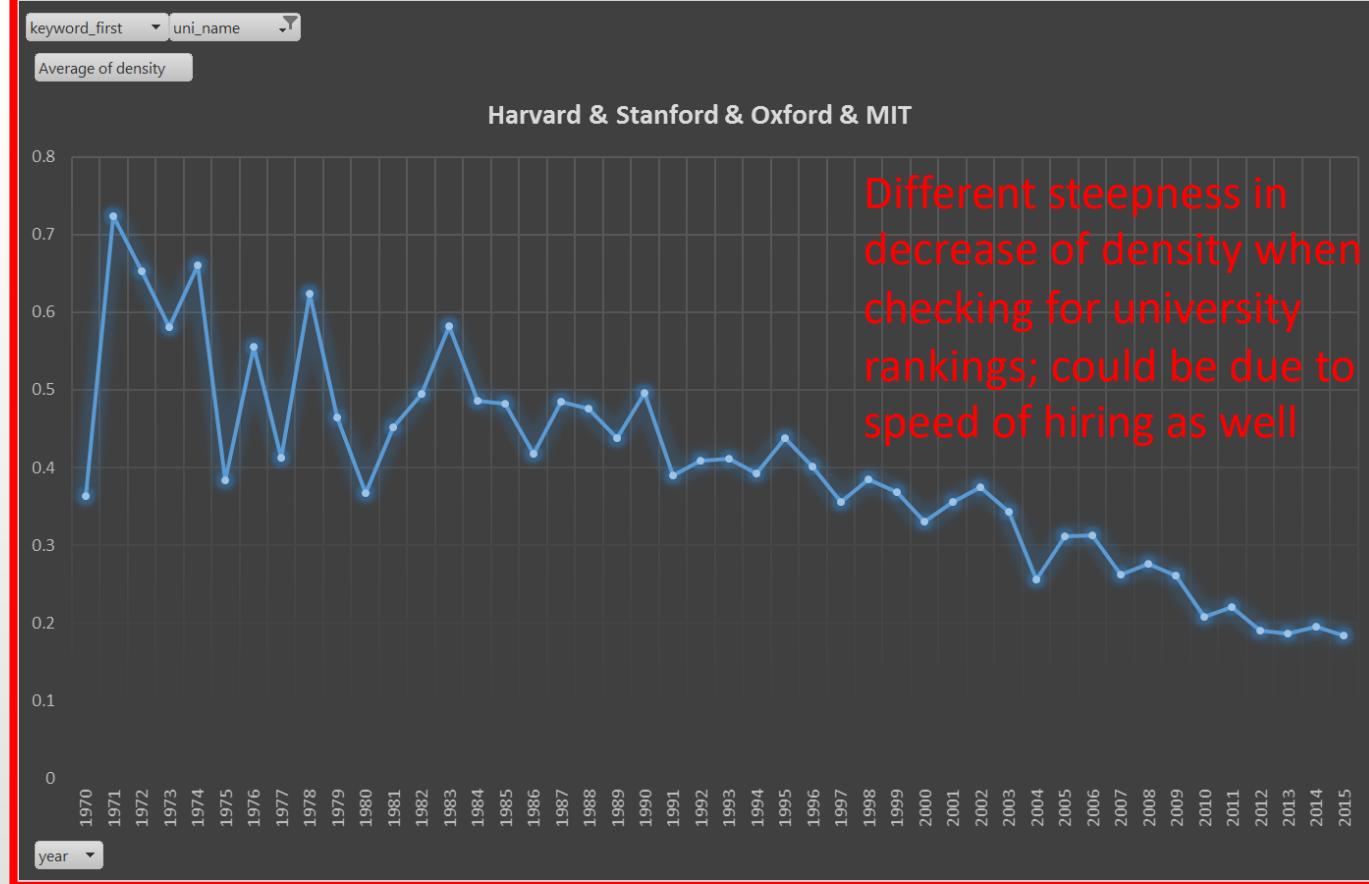


# All fields universities ranked

Somehow similar story & trends when we look at different ranked universities

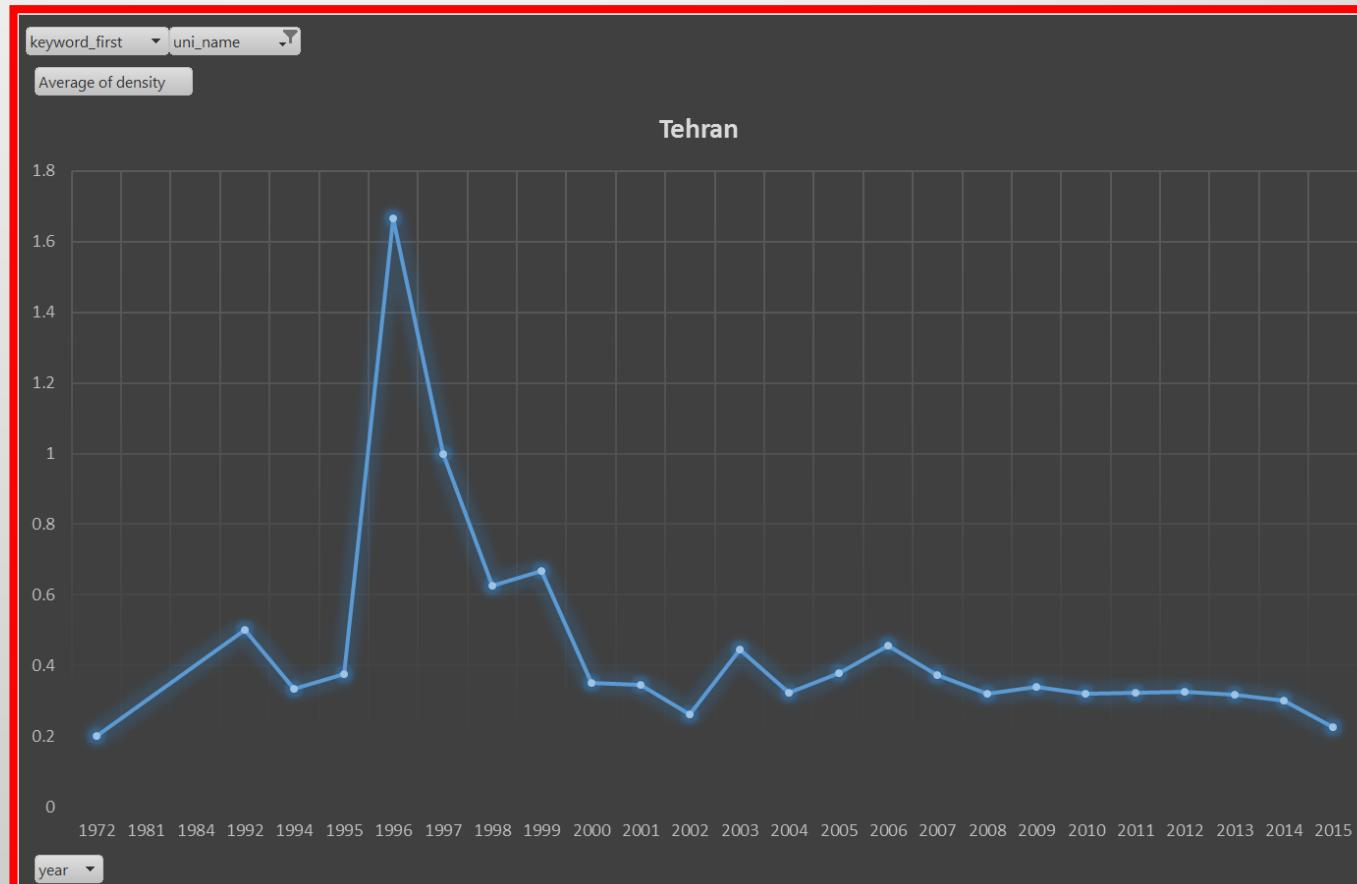


# Average density universities ranked



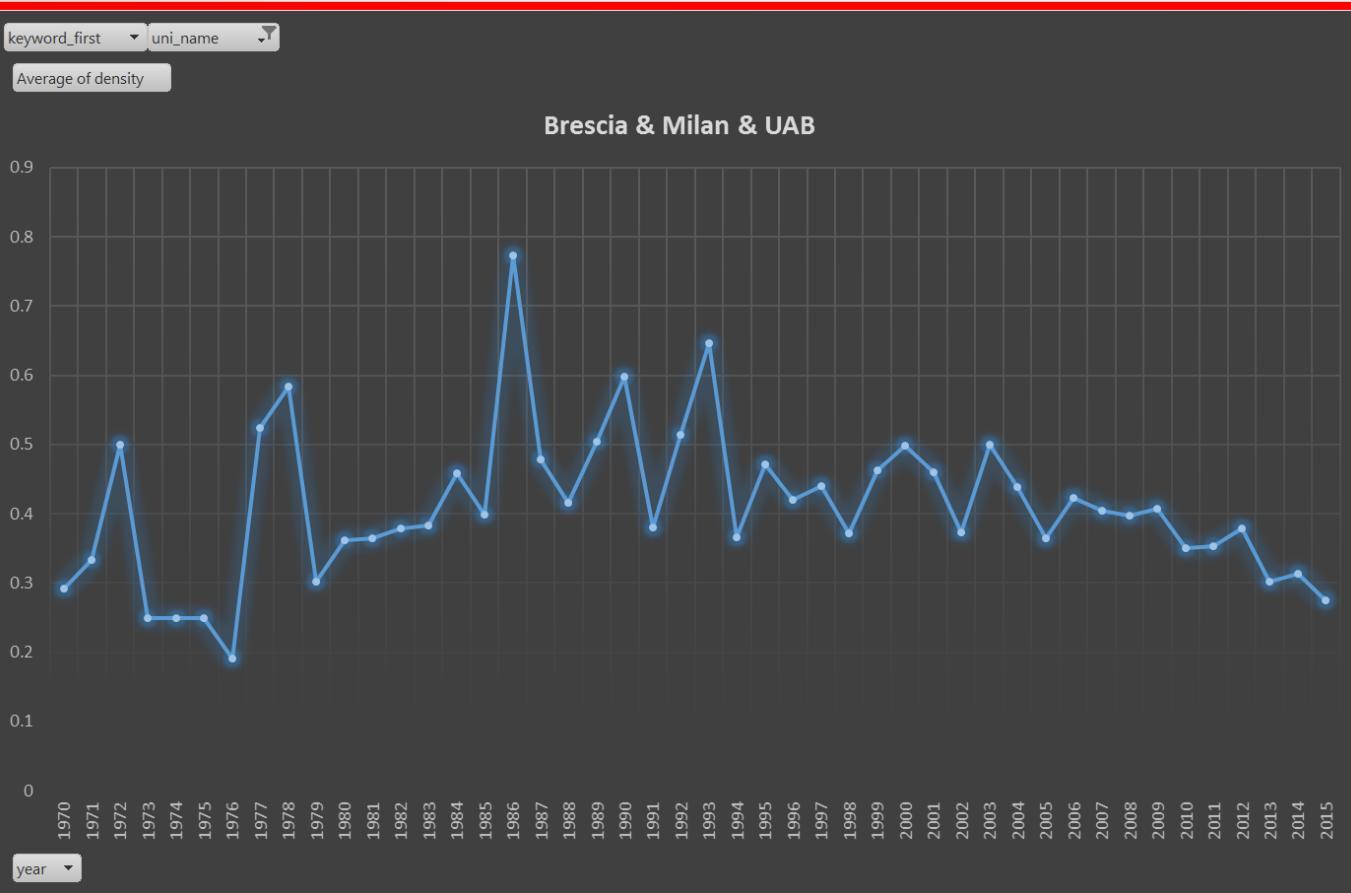
# Average density universities ranked

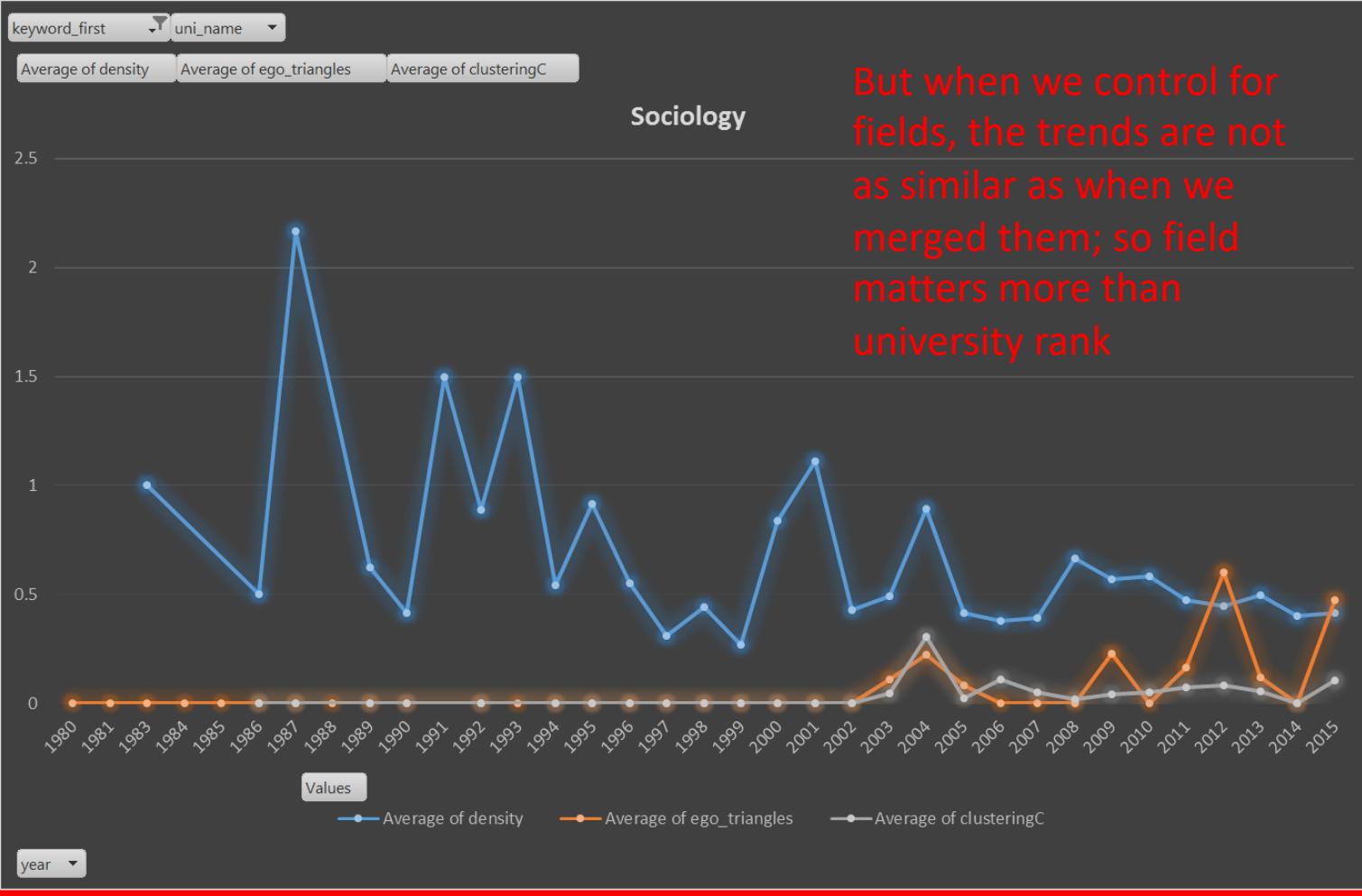
Different steepness in decrease of density when checking for university rankings; could be due to speed of hiring as well



# Average density universities ranked

Different steepness in decrease of density when checking for university rankings; could be due to speed of hiring as well

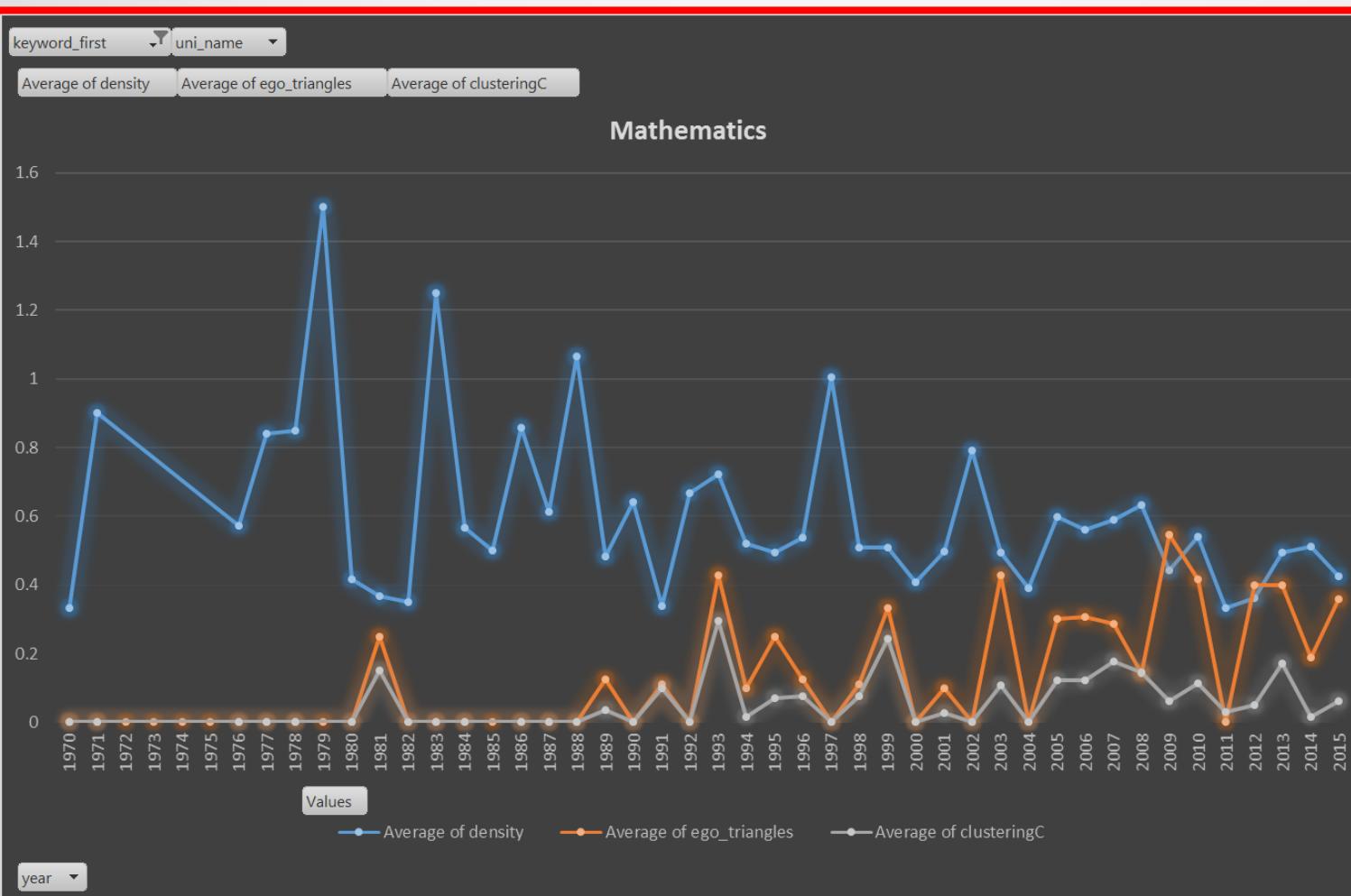




How  
about  
fields?

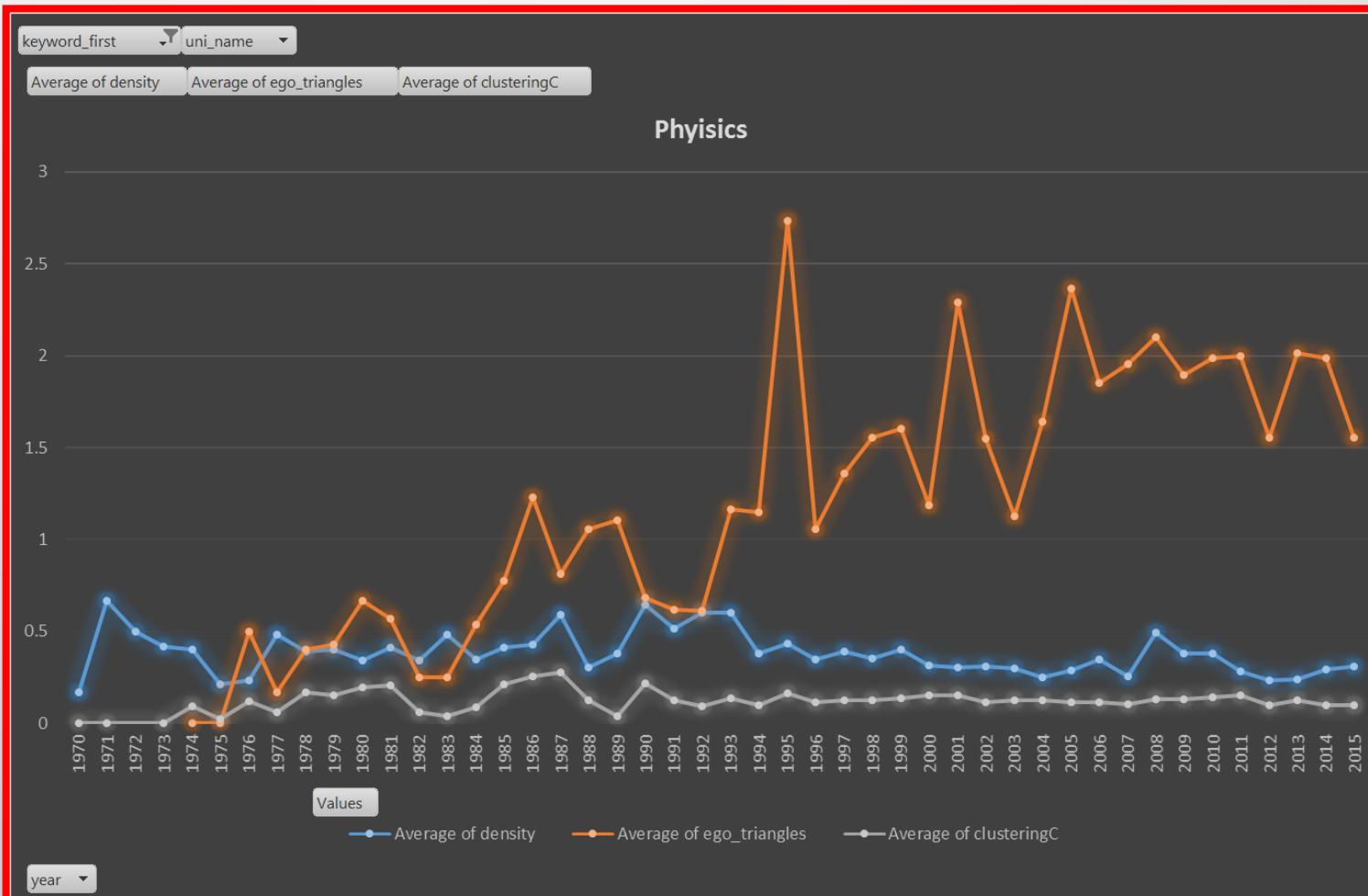
# How about fields?

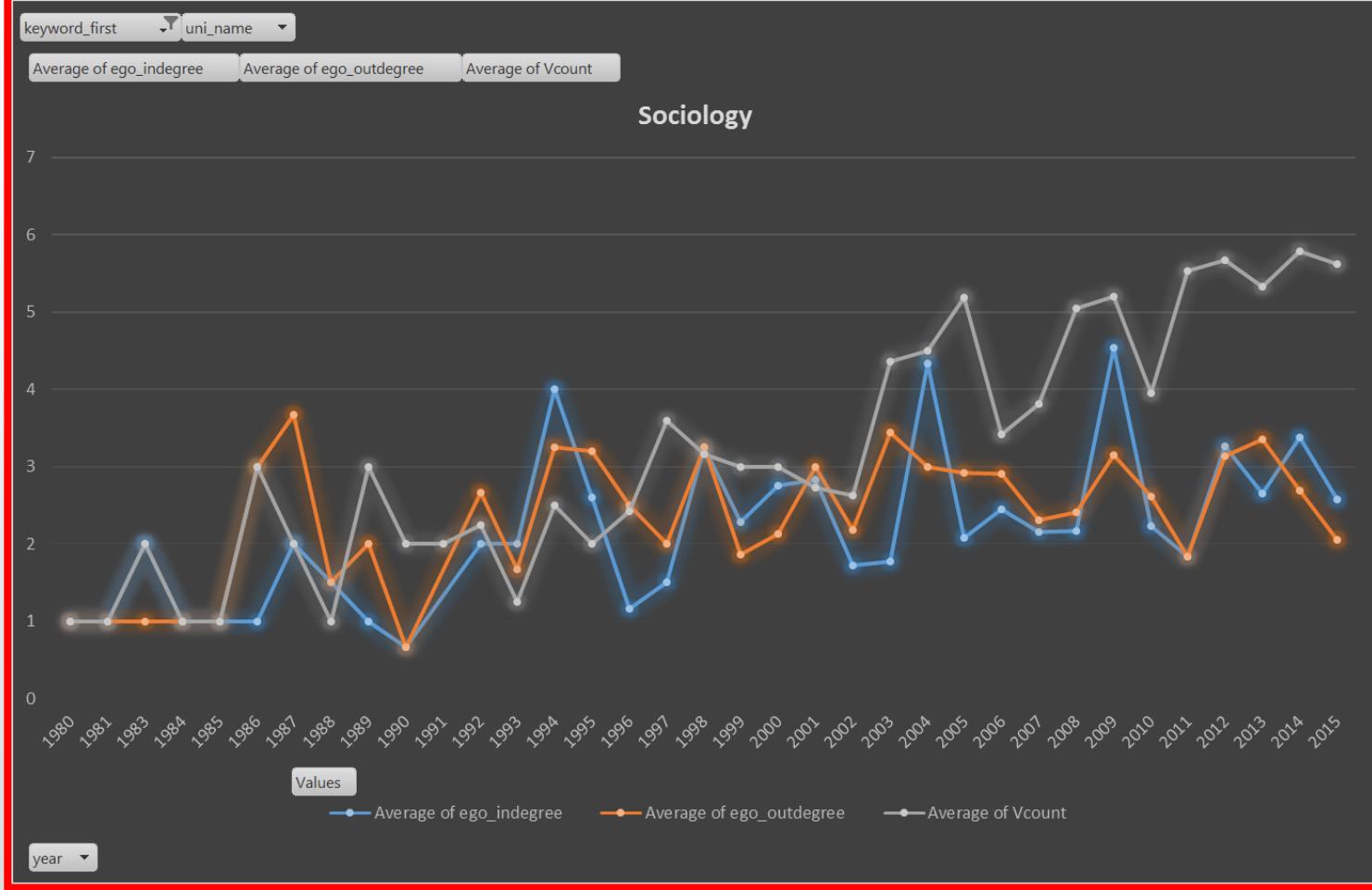
But when we control for fields, the trends are not as similar as when we merged them; so field matters more than university rank



# How about fields?

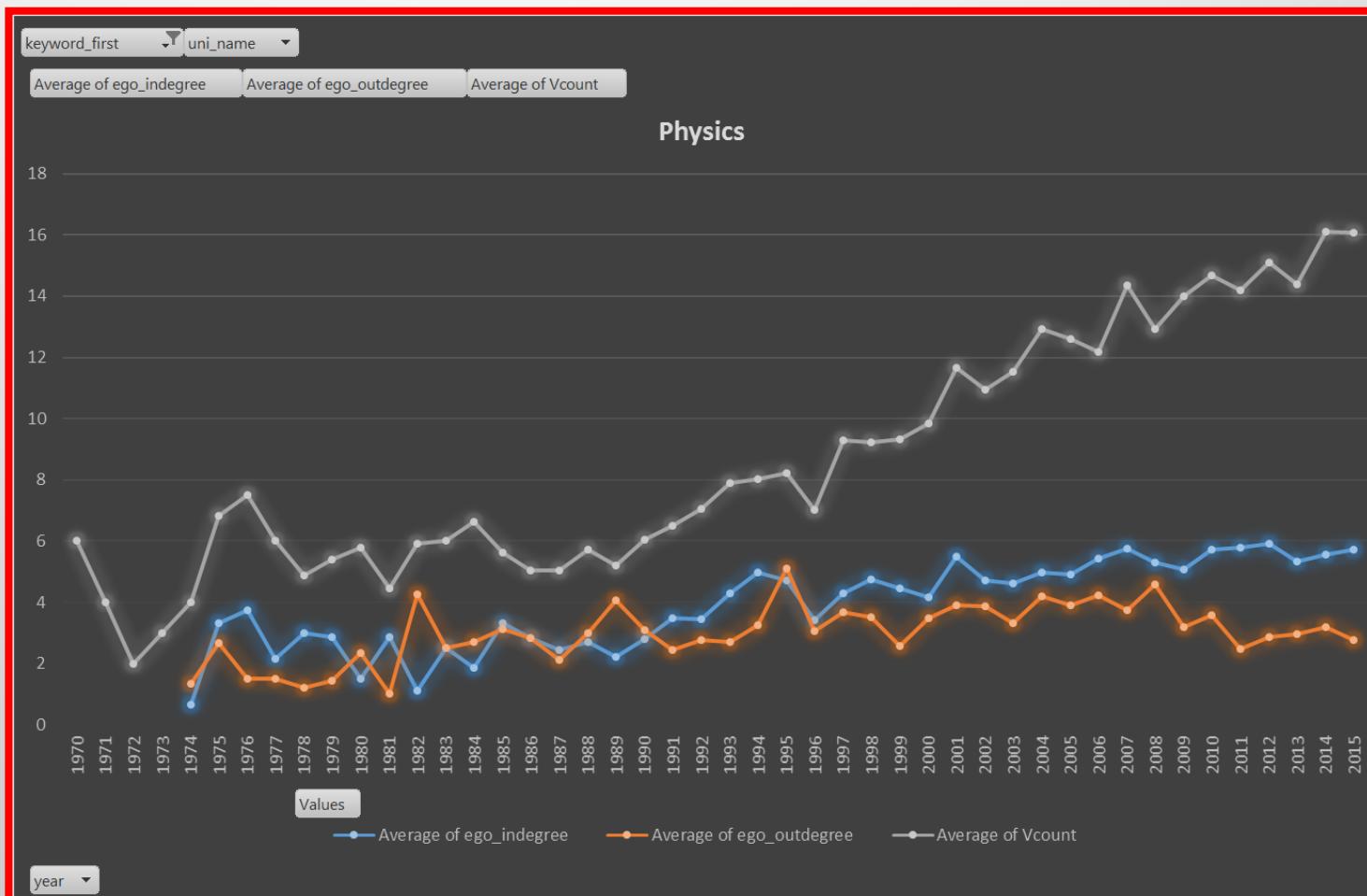
But when we control for fields, the trends are not as similar as when we merged them; so field matters more than university rank

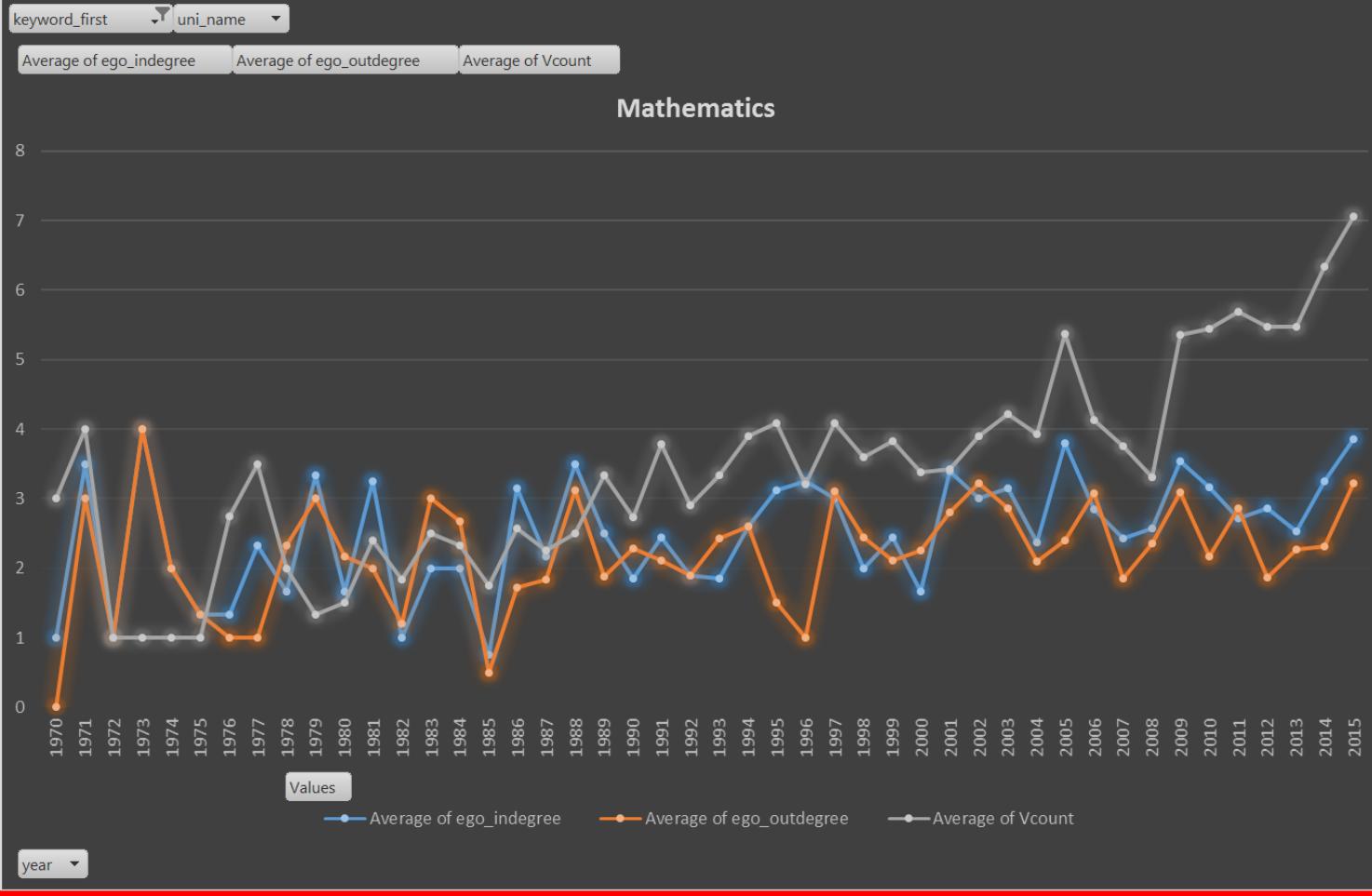




How  
about  
fields?

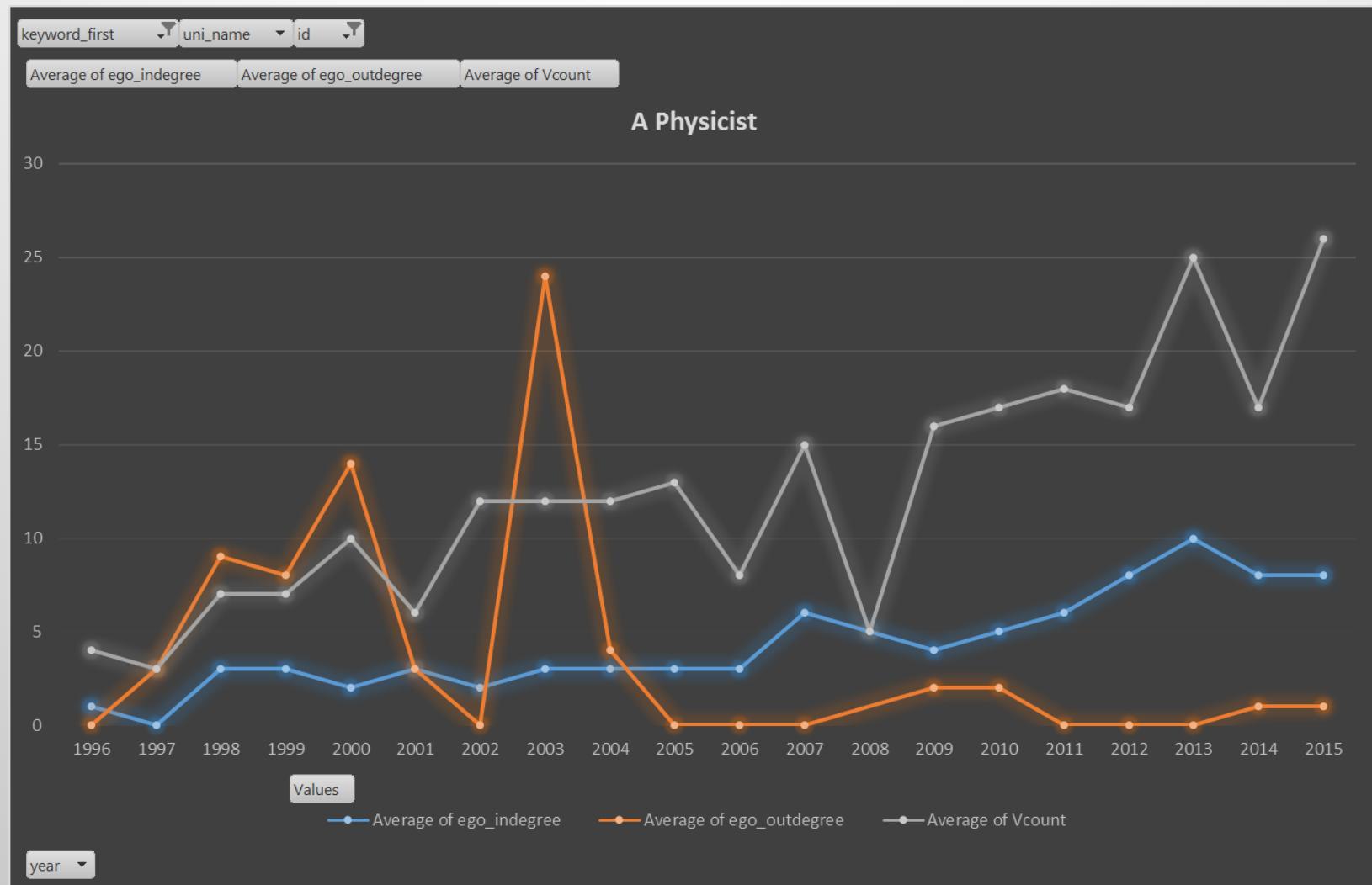
How  
about  
fields?





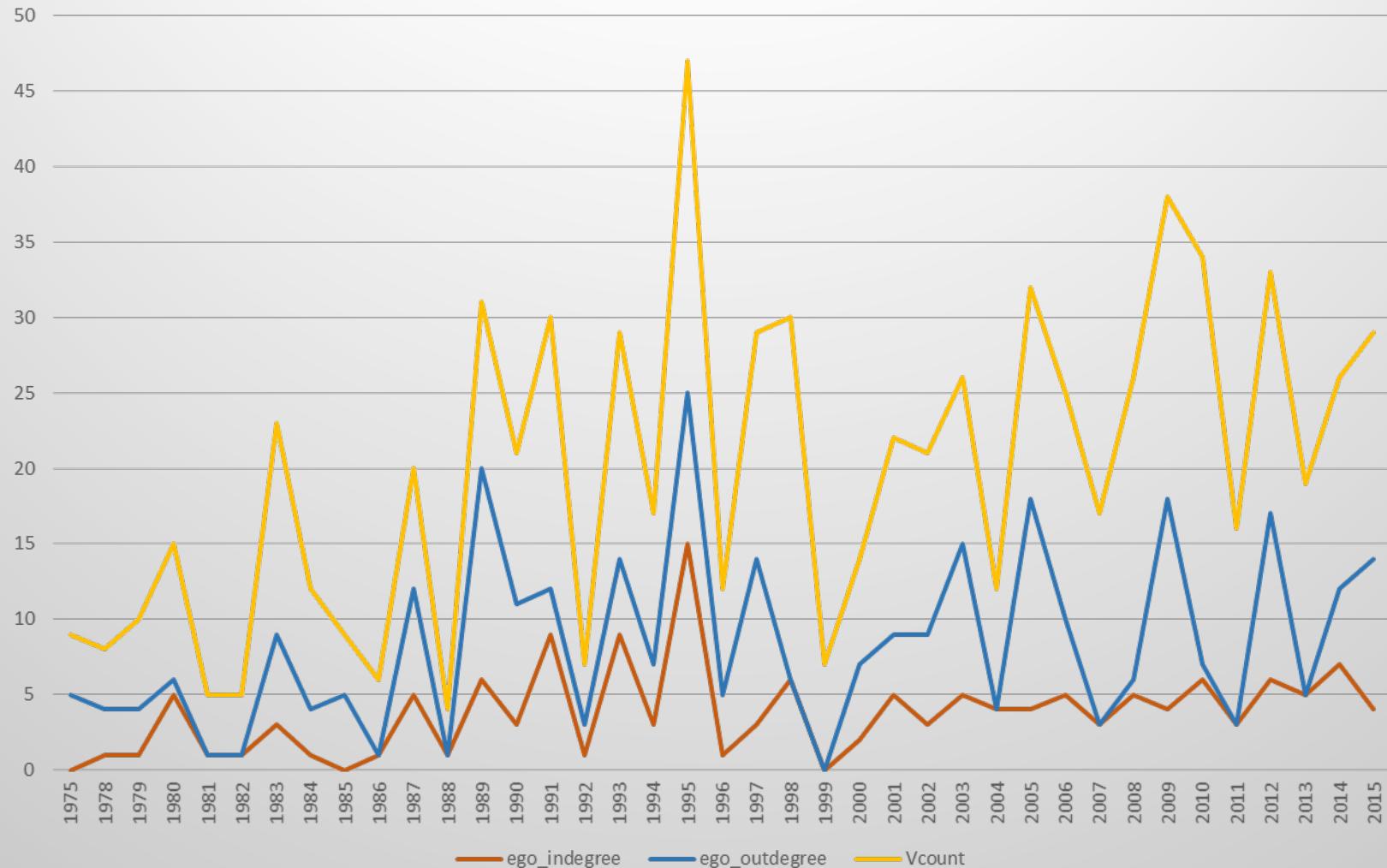
How  
about  
fields?

# A physicist

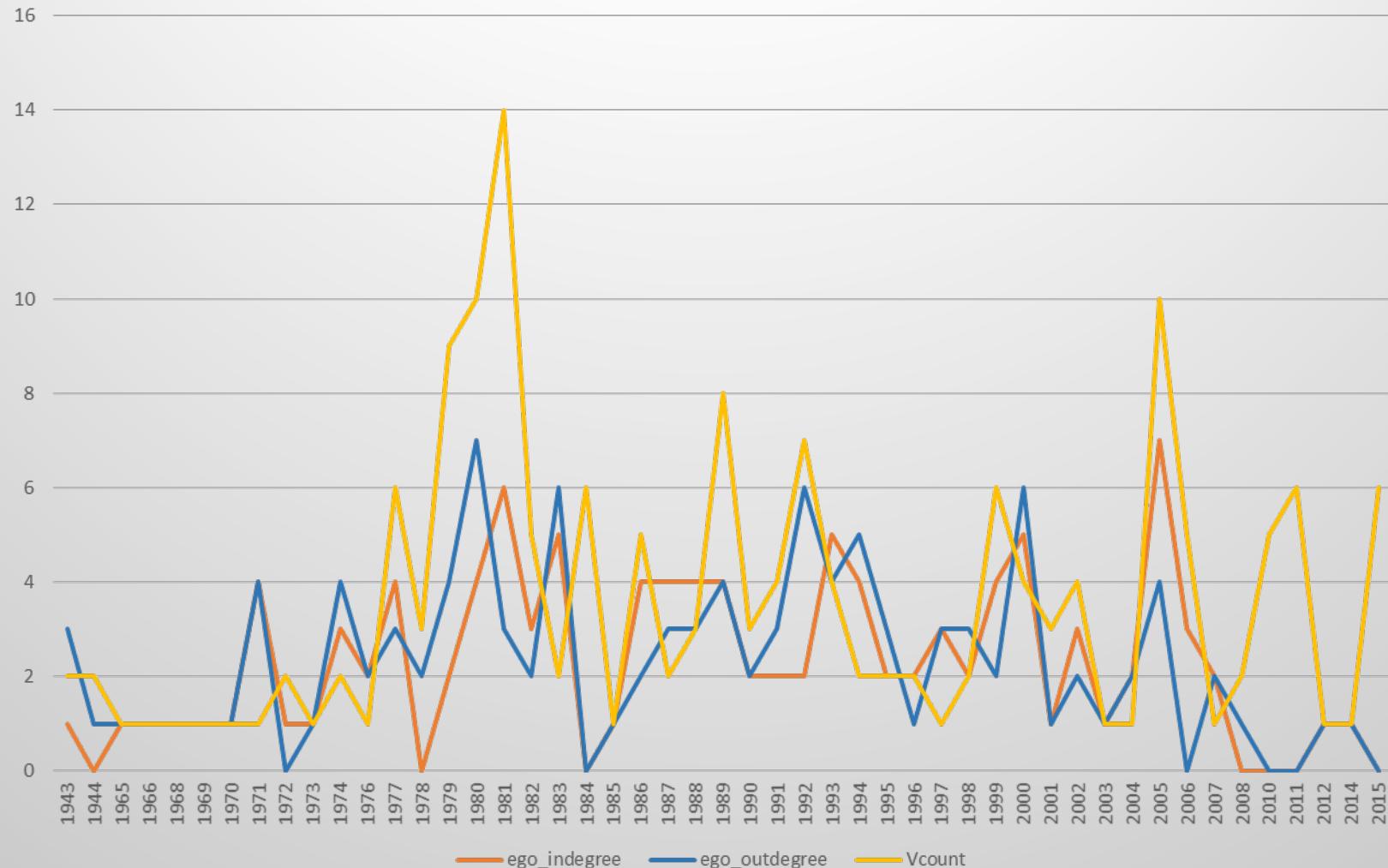


Obvious difference in in and out degree (that shows order of names in papers); usual increase in vcount

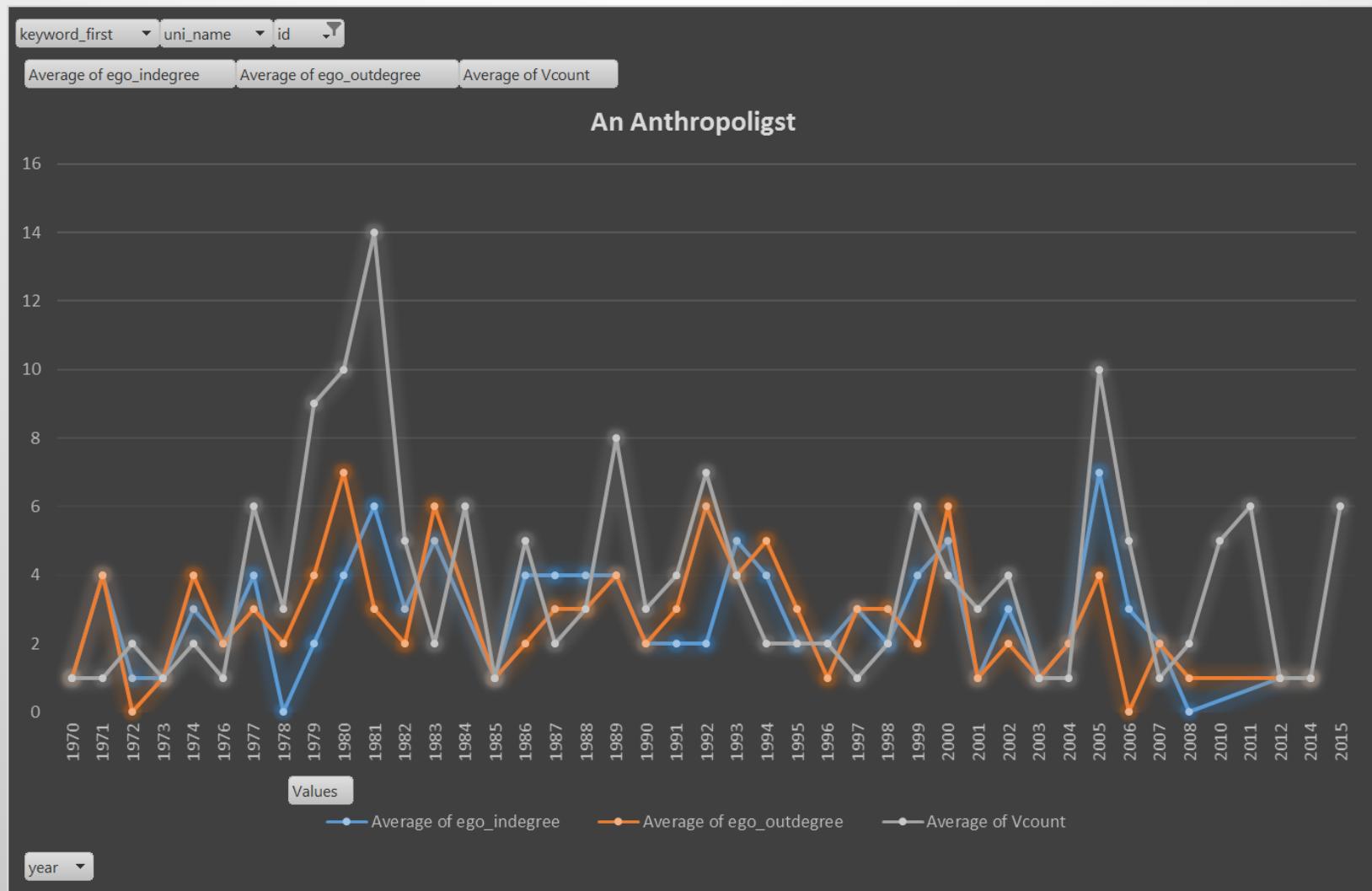
# A plasma physicist



# An anthropologist



# An anthropologist 2



# Future steps

- 2 questions,
- first, my main and big question is how scientific community works? How scholars connect to each other to build up their scientific career, the first way I tried to answer this question was **through co-authorship networks**, the next ways I am thinking to add to this (to build up the parts of the puzzle) is **citation networks**, and **conceptual map of papers keywords** (that can bee seen as a kind of indirect relationship between authors/scholars); if possible **scholars personal networks**; what are other ways you think would help me explore and answer this question? PEERE data (**peer review as indirect interaction between scholars**) probably ...
- 2nd question is now I am at the end of 1st year of my PhD and I am looking for places, institutes and scholars working on similar subjects of how science (and more focused on social sciences) grow through scholars interactions, so please introduce me if you are aware of any options of a fruitful research visit ...
- I have been reading about stochastic actor oriented models (SAOM), TERGM, and I have heard about longitudinal data analysis with Rsiena; I am thinking to apply them

# Nan Lin's (1999) model of Social Capital

