

College Network: Comparison Groups

Sebastian Daza

University of Wisconsin-Madison

September 15, 2014

Comparison Groups 2012

- ▶ Each year colleges submit *comparison groups* to the U.S. Department of Education
- ▶ Feedback on how these institutions stack up in terms of:
 - ▶ Finances
 - ▶ Enrollment
- ▶ Integrated Postsecondary Education Data System (IPEDS)

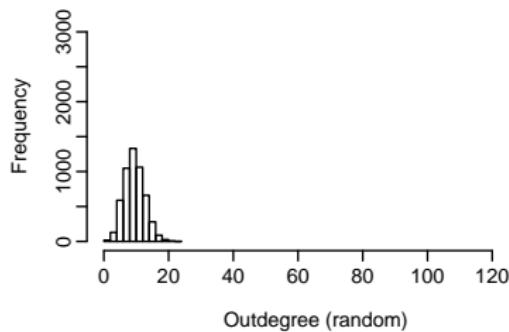
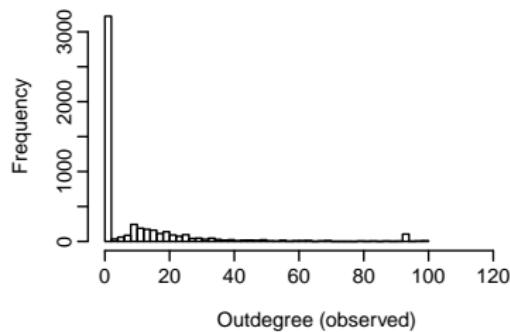
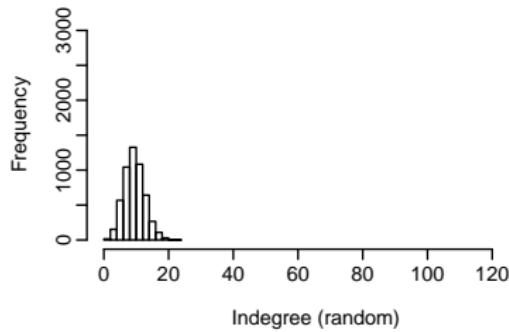
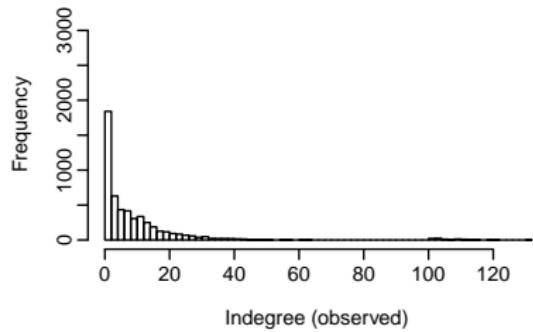
"When colleges look to compare themselves with others, they're not much different from high-school students chasing popularity: Everyone wants to be friends with the Ivy League, but the Ivy League is really picky about whom it hangs out with."

The Chronicle of Higher Education, September 2012

Network Data 2012

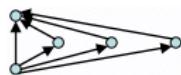
- ▶ Directed
- ▶ 5239 nodes and 52108 edges
 - ▶ Density: 0.0019
 - ▶ Reciprocity: 0.109
 - ▶ Transitivity: 0.117
 - ▶ Degree Centralization: 0.019

Degree Distribution

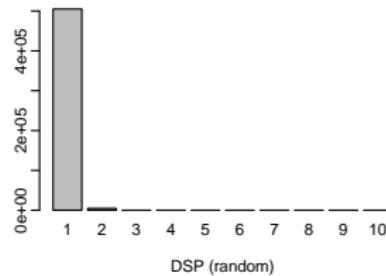
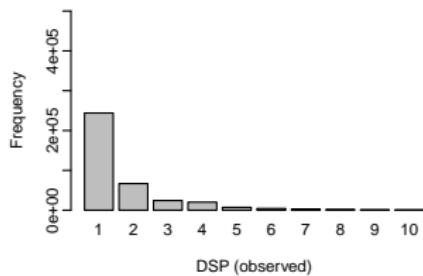
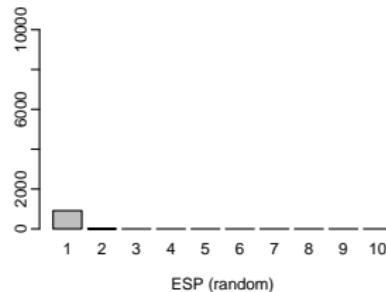
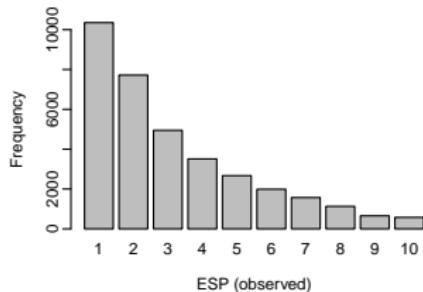
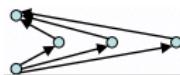


Shared Partners

ESP



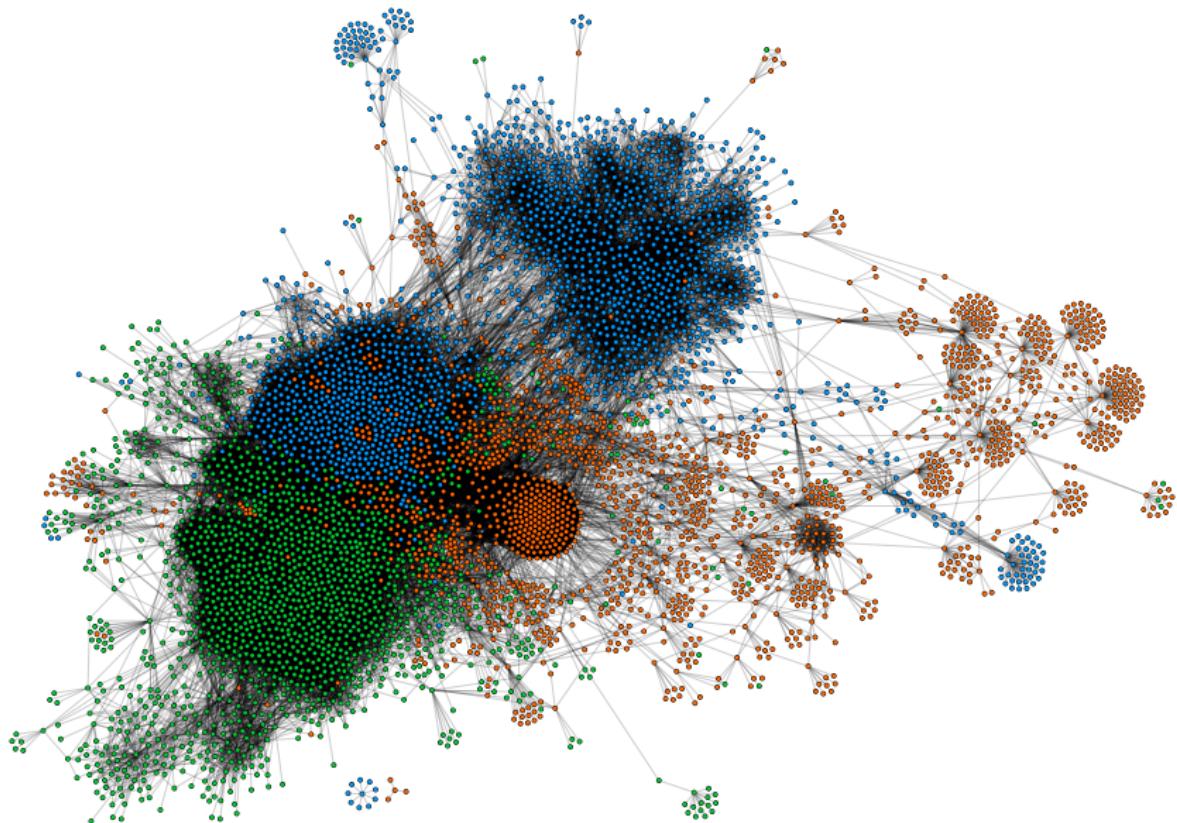
DSP



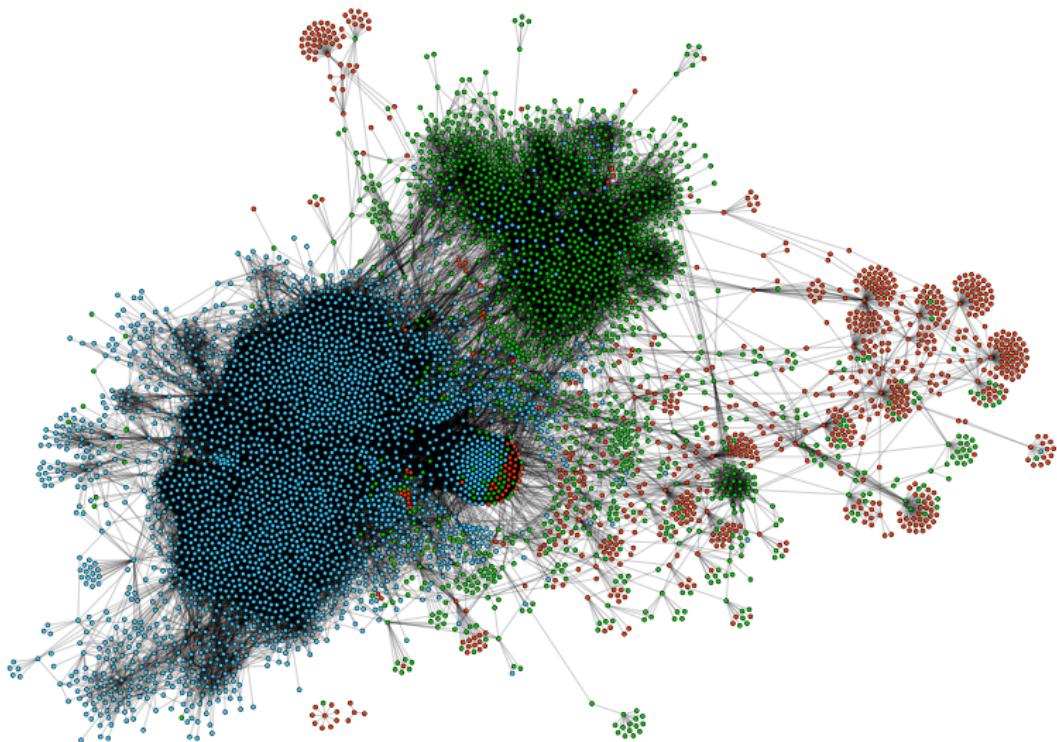
Ranking Eigenvector Centrality

Name	Centrality
Iowa State U	0.15
Strayer U-District of Columbia	0.12
Florida State U	0.11
Spring Hill C	0.10
Wheaton C	0.09
Flagler C-St Augustine	0.08
The U of Tampa	0.08
U of Alabama at Birmingham	0.08
The National Hispanic U	0.08
Northern Arizona U	0.07
Earlham C	0.07
Southeastern U	0.07
Oglethorpe U	0.07
Connecticut C	0.06
U of Arizona	0.06
U of Colorado Boulder	0.06
DeVry U-California	0.05
Knox C	0.05
Colorado C	0.05
Southern Illinois U Edwardsville	0.05

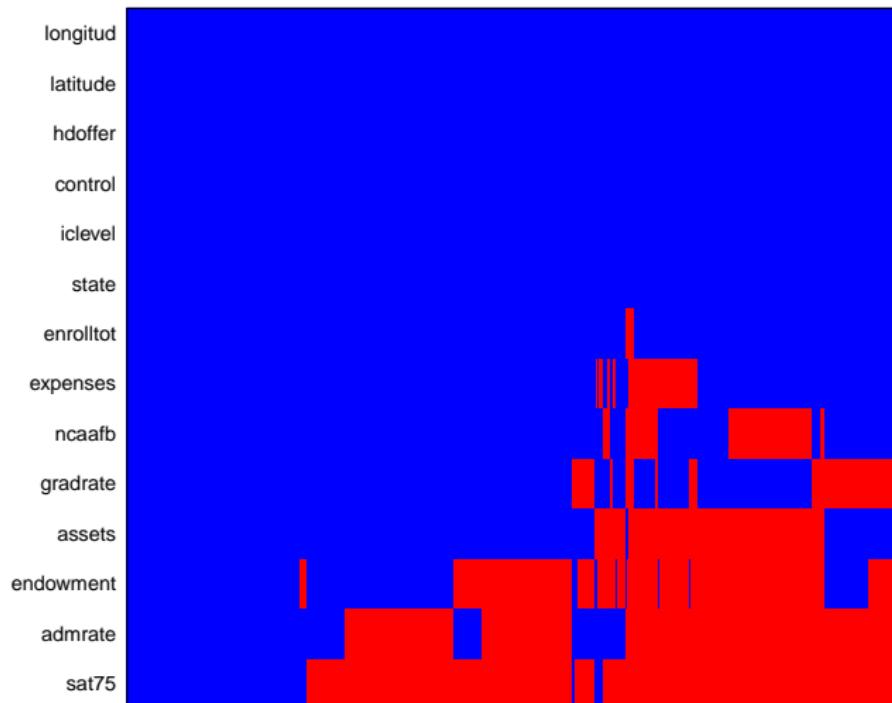
Control: Public, Non-for-Profit, For-Profit



Duration: Less 2 yrs, 2-4 yrs, 4 or more yrs

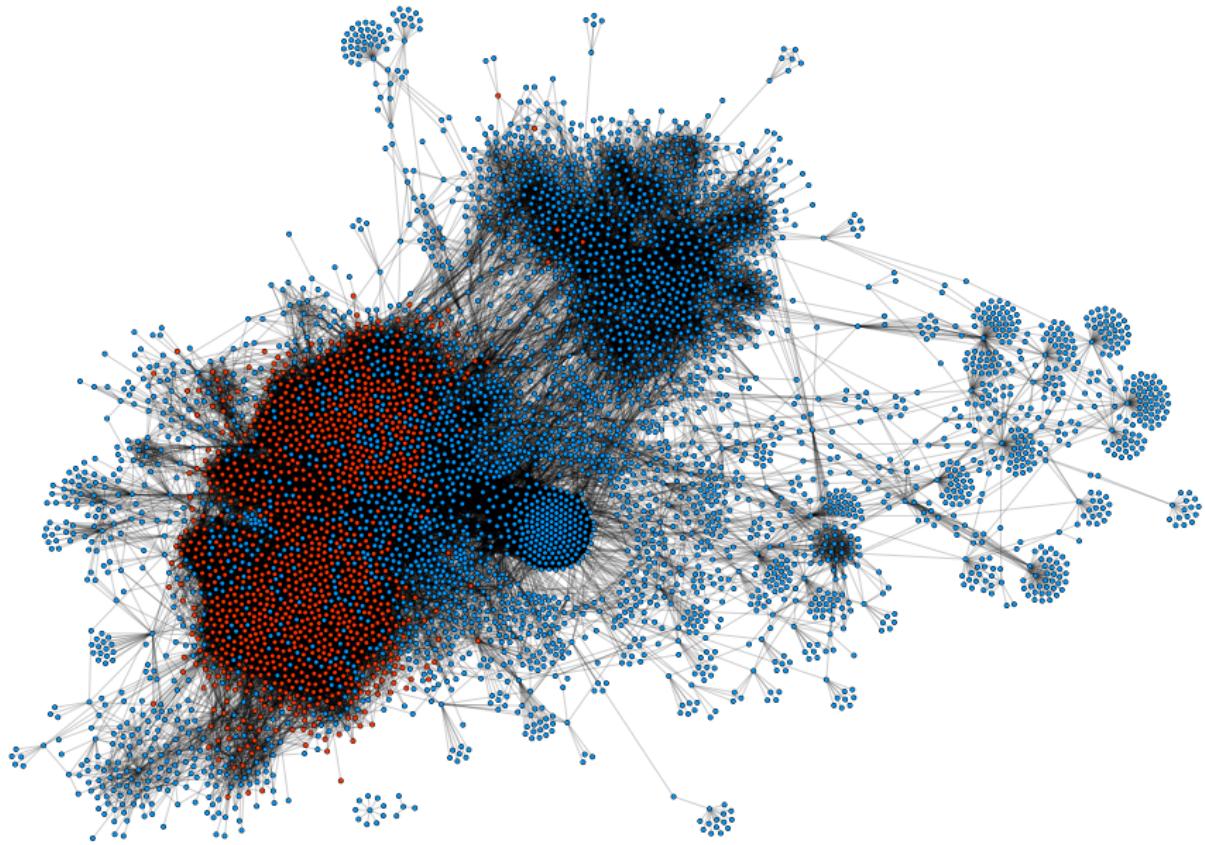


Attributes (IPEDS 2012): Missing data



Ordered by number of missing items

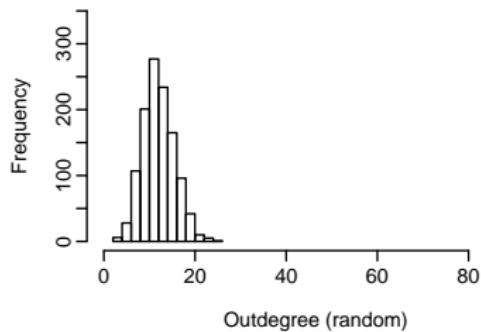
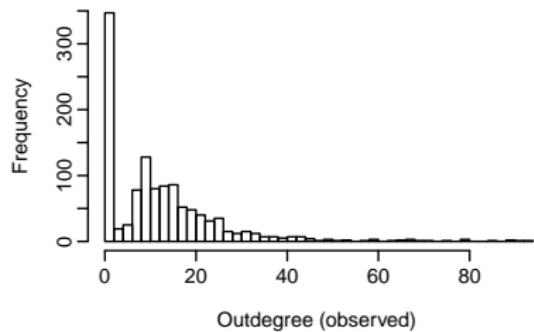
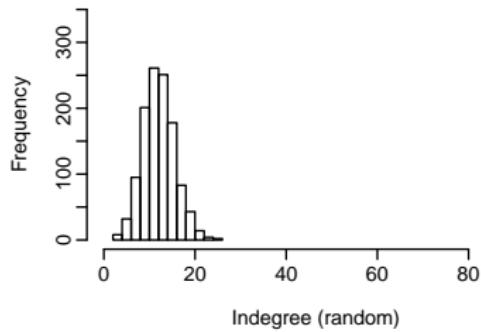
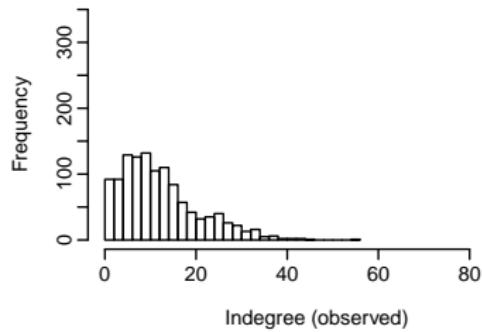
Only Complete Data



Network Data No Missing Attributes

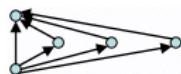
- ▶ 1172 nodes and 14757 edges
 - ▶ Density: 0.011 (0.0019)
 - ▶ Reciprocity: 0.184 (0.109)
 - ▶ Transitivity: 0.267 (0.117)
 - ▶ Degree Centralization: 0.039 (0.019)

Degree Distribution

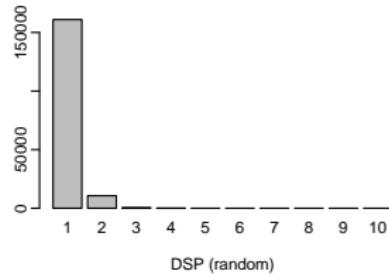
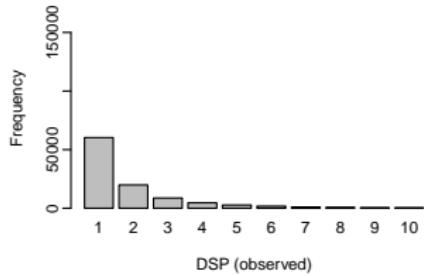
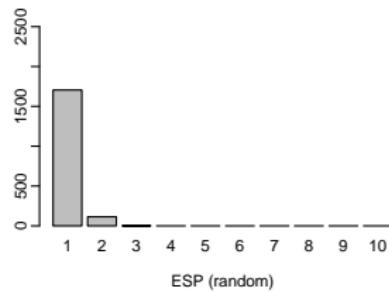
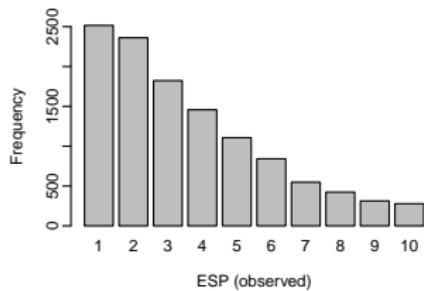
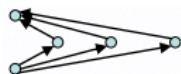


Shared Partners

ESP



DSP



Ranking Eigenvector Centrality

Name	Centrality
Iowa State U	0.31
Rutgers U-New Brunswick	0.30
Florida State U	0.25
Texas Tech U	0.23
Brigham Young U-Provo	0.19
U of Wyoming	0.18
U of Alabama at Birmingham	0.18
Northern Arizona U	0.15
U of Michigan-Ann Arbor	0.13
U of Virginia-Main Campus	0.13
U of Wisconsin-Madison	0.13
U of Missouri-Columbia	0.13
Stony Brook U	0.13
Purdue U-Main Campus	0.13
U of Pittsburgh-Pittsburgh Campus	0.13
U of Arizona	0.13
The U of Alabama	0.12
Virginia Polytechnic Ins and State U	0.12
U of St. Thomas	0.10
Montana State U	0.10

Attribute Descriptives

	Mean	Median	SD	Min	Max
Admission Rate	62.46	64.30	18.44	5.81	100.00
SAT 75th pct.	586.05	580.00	69.88	395.00	795.00
Graduation Rate	54.84	54.00	18.55	0.00	100.00
NCAA Football	0.51	-	0.50	0.00	1.00
Enrollment*	8760	4126	11084	72	81789
Assets*	536568745	111236699	2536864389	0	57714643000
Endowment	305043091	36660102	1490720278	65712	32012729000
Exp/Enroll Ratio*	25424	17874	36973	3899	944950

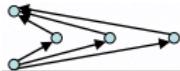
Attribute Descriptives (Correlations)

	admitrate	sat75	graduate	enrolltot	assets	endowment	eeratio
admitrate		-0.33	-0.28	-0.03	-0.33	-0.33	-0.35
sat75			0.77	0.23	0.38	0.39	0.47
graduate				0.14	0.32	0.32	0.40
enrolltot					0.18	0.22	0.09
assets						0.96	0.46
endowment							0.44
eeratio							

ERG Models (Specification)

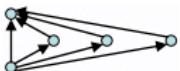
- ▶ Purely Structural Effects (Endogenous)

- ▶ Reciprocity?
 - ▶ Two paths? Multiple two paths (DSP)?



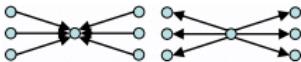
- ▶ Transitive closure (ESP)?

Partners of my partner are my partners



- ▶ Popularity spread? Activity spread?

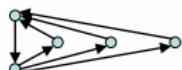
Tendency for centralization in the in-degree and out-degree distribution



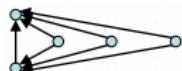
ERG Models (Specification: Future!)

► Other Structural Effects (Endogenous)

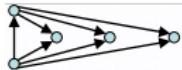
- Cyclic closure (non-hierarchical or generalized exchange)



- Popularity closure (status homophily)



- Activity closure (joint dependence or structural homophily, cooperation among potential competitors?)



ERG Models (Specification)

- ▶ Actor-relation effects (Exogenous)
 - ▶ Sender effect (attribute-based activity)
 - ▶ Receiver effect (attribute-based popularity)
 - ▶ Homophily (uniform, differential)
- ▶ Covariate edge
 - ▶ Geographic distance

Hypotheses More Than Results

- ▶ Structural effects
 - ▶ Ties not more likely to be reciprocated
 - ▶ Ties tends to occur in regions of transitive closure
gwesp+ gwdsp-
 - ▶ Levels of popularity and activity tend to be homogeneous
gwidegree- gwodegree-
 - ▶ High-degree nodes seem to be arising from triangulation not from star effects

Hypotheses More Than Results

- ▶ Actor-relation effects
 - ▶ Positive attribute => less active, more popular
 - ▶ Homophily by state, public/private, highest degree offered, NCAA
 - ▶ Heterophily by SAT, graduation rate, expenses/students ratio, enrollment, endowments
 - ▶ Negative effect of distance on tie formation

Next steps

- ▶ Get models right!
- ▶ Deal with missing data (experiences?)
- ▶ Other important attributes?
- ▶ Community detection
 - ▶ `linkcomm` ≈ 1000
 - ▶ `mcode` ≈ 85
- ▶ Ideas? Hypotheses?

Thank you!