Visualization Design and Redesign

Maneesh Agrawala

CS 448B: Visualization
Winter 2020

1

Last Time: Data and Image Models

2

The big picture

task
- questions, goals, assumptions

data
- physical type
  - int, float, etc.
- abstract type
  - nominal, ordinal, etc.

domain
- metadata
- semantics
- conceptual model
- conventions

processing algorithms →

image
graphical marks
visual channel

mapping
visual encoding

3

Nominal, ordinal and quantitative

On the theory of scales of measurements
S. S. Stevens, 1946

N - Nominal (labels)
Fruits: Apples, oranges, ...
Operations: =, ≠

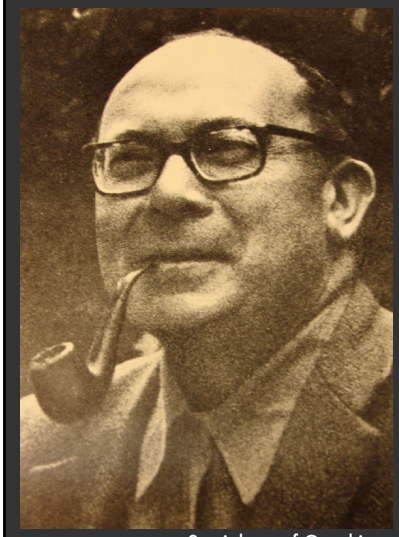
O - Ordered
Quality of meat: Grade A, AA, AAA
Operations: =, ≠, <, >

Q - Interval (location of zero arbitrary)
Dates: Jan, 19, 2016; Loc.: (LAT 33.98, LON -118.45)
Like a geometric point. Cannot compare directly
Only differences (i.e. intervals) may be compared
Operations: =, ≠, <, >, -

Q - Ratio (location of zero fixed)
Physical measurement: Length, Mass, Temp, ...
Counts and amounts
Like a geometric vector, origin is meaningful
Operations: =, ≠, <, >, -, ÷

4

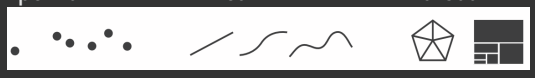
Marks and Visual Variables



Semiology of Graphics
J. Bertin, 1967

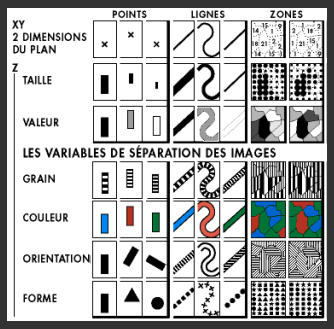
Marks: geometric primitives

points lines areas



Visual Variables: control mark appearance

- Position (2x)
- Size
- Value
- Texture
- Color
- Orientation
- Shape



5

Bertins' "Levels of Organization"

Position	N	O	Q
Size	N	O	Q
Value	N	O	Q
Texture	N	O	
Color	N		
Orientation	N		
Shape	N		

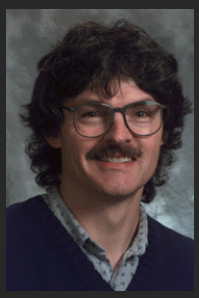
N Nominal
O Ordered
Q Quantitative

Note: Q < O < N

7

Automated design

Jock Mackinlay's APT 86



8

Principles

Challenge:

Assume 8 visual encodings and n data fields

Pick the best encoding from the exponential number of possibilities $(n+1)^8$

Principle of Consistency:

The properties of the image (visual variables) should match the properties of the data

Principle of Importance Ordering:

Encode the most important information in the most effective way

9

Mackinlay's expressiveness criteria

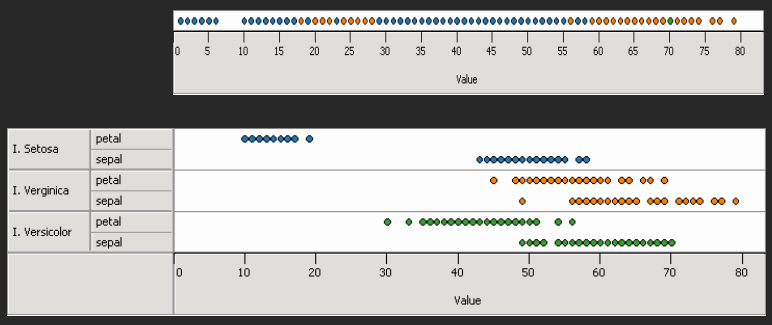
Expressiveness

A set of facts is expressible in a visual language if the sentences (i.e. the visualizations) in the language express *all* the facts in the set of data, and *only* the facts in the data.

10

Cannot express the facts

A one-to-many (1 → N) relation cannot be expressed in a single horizontal dot plot because multiple tuples are mapped to the same position



11

Expresses facts not in the data

A length is interpreted as a quantitative value;

∴ Length of bar says something untrue about N data

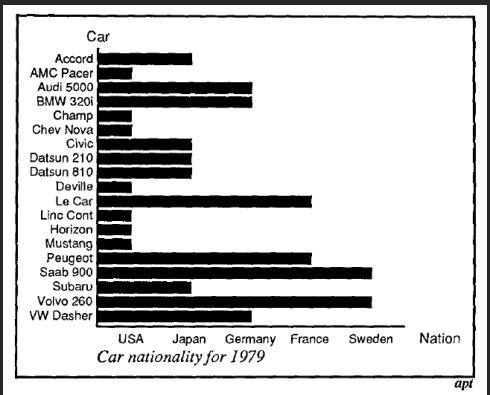


Fig. 11. Incorrect use of a bar chart for the *Nation* relation. The lengths of the bars suggest an ordering on the vertical axis, as if the USA cars were longer or better than the other cars, which is not true for the *Nation* relation.

[Mackinlay, APT, 1986]

12

Mackinlay's effectiveness criteria

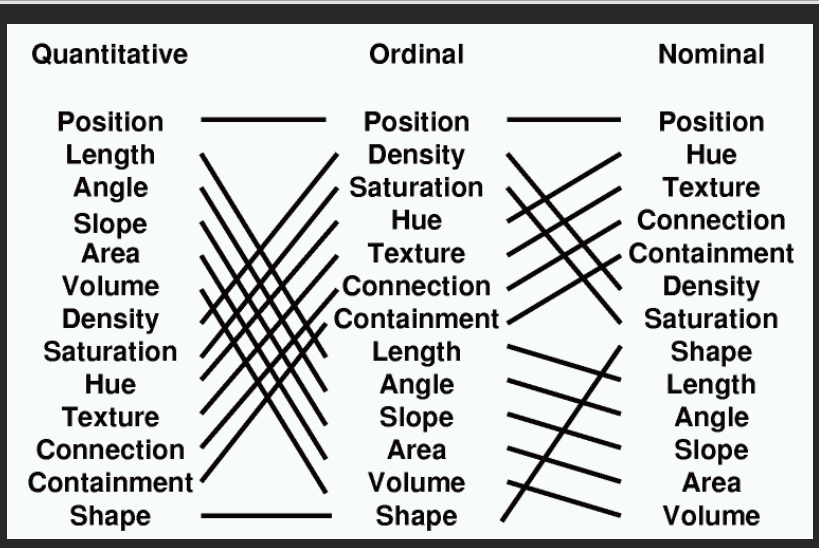
Effectiveness

A visualization is more effective than another visualization if the information conveyed by one visualization is more readily *perceived* than the information in the other visualization.

Subject of perception lecture

13

Mackinlay's ranking



Conjectured *effectiveness* of the encoding

14

Mackinlay's Design Algorithm

User formally specifies data model and type

Input: list of data variables ordered by importance

APT searches over design space

- Tests expressiveness of each visual encoding (rule-based)
- Generates encodings that pass test
- Rank by perceptual effectiveness criteria

Outputs *most effective* visualization

16

Automatic chart construction

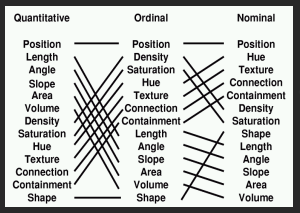
Automating the design of graphical presentation of relational information
J. Mackinlay, 1986

Encode most important data using highest ranking visual variable for the data type

Year	Exports	Imports
1700	170,000	300,000
1701	171,000	302,000
1702	176,000	303,000
...	...	...

→

1. Year (Q)
2. Exports (Q)
3. Imports (Q)



→

mark: lines
Year → x-pos (Q)
Exports → y-pos (Q)
Imports → y-pos (Q)

17

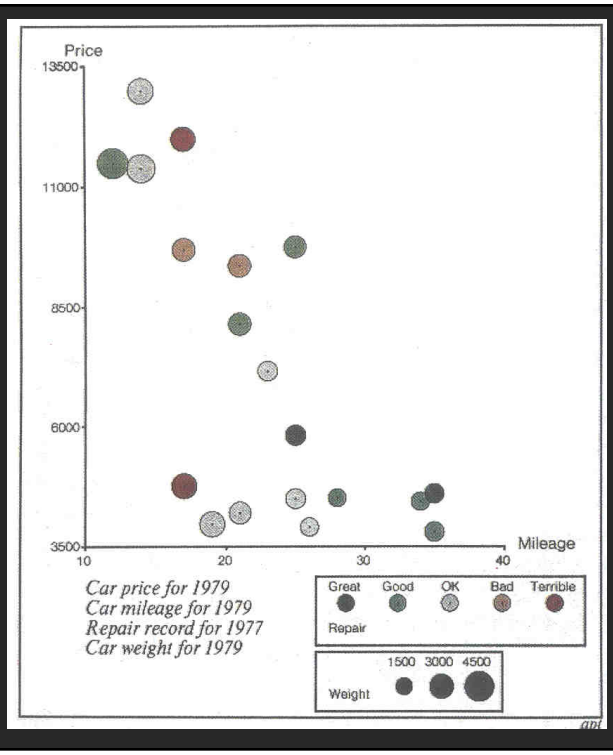
Cars Data

1. Price (Q)
2. Mileage (Q)
3. Weight (Q)
4. Repair (O)

Quantitative

Position
Length
Angle
Slope
Area
Volume
Density
Saturation
Hue
Texture
Connection
Containment
Shape

[Mackinlay, APT, 1986]



18

Limitations

Does not cover many visualization techniques

- Networks, maps, diagrams
- Also, 3D, animation, illustration, …

Does not consider interaction

Does not consider semantics or conventions

Assumes single visualization as output

19

Summary

Formal specification

- **Data model:** relational data, N,O,Q types
- **Image model:** marks, attributes, encodings
- **Encodings mapping data to image**

Choose expressive and effective encodings

- **Rule-based test of expressiveness**
- **Perceptual effectiveness rankings**

20

Announcements

21

Announcements

Class participation requirements

- Complete readings and notebooks before class
- In-class discussion
- Post at least 1 discussion substantive comment/question per week.
  - 1 pass for the quarter

Class website
https://magrawala.github.io/cs448b-wi20

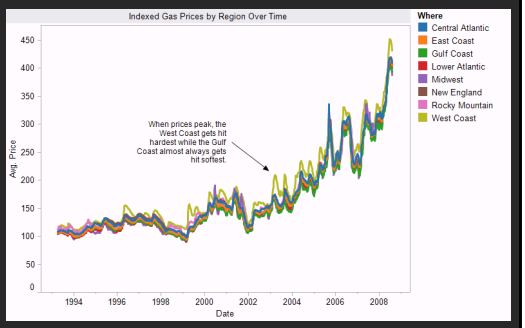
22

A2: Exploratory Data Analysis

Use Tableau to formulate & answer questions

First steps

Step 1: Pick domain & data
Step 2: Pose questions
Step 3: Profile data
Iterate as needed



Create visualizations

Interact with data
Refine questions

Author a report

Screenshots of most insightful views (10+)
Include titles and captions for each view

Due before class on Jan 27, 2020

23

A1 Review

24

Design Considerations

Guides: Title, labels, legend, captions, source!

Expressiveness and Effectiveness

- Express the facts and only the facts
- Avoid unexpressive marks (lines? gradients?)
- Use perceptually effective encodings that match data type
- Don't distract: faint gridlines, pastel highlights/fills
- The "elimination diet" approach – start minimal

Support comparison and pattern perception

- Between elements, to a reference line, or to counts
- Use reader-friendly units and labels

26

Design Considerations

Group / sort data by meaningful dimensions

Transform data (e.g., filter, log, normalize)

- Are model choices (regression lines) appropriate?

Reduce cognitive overhead

- Minimize visual search, minimize ambiguity
- Appropriate size, aspect ratio, legible text
- Avoid legend lookups if direct labeling works
- Avoid color mappings with indiscernible colors

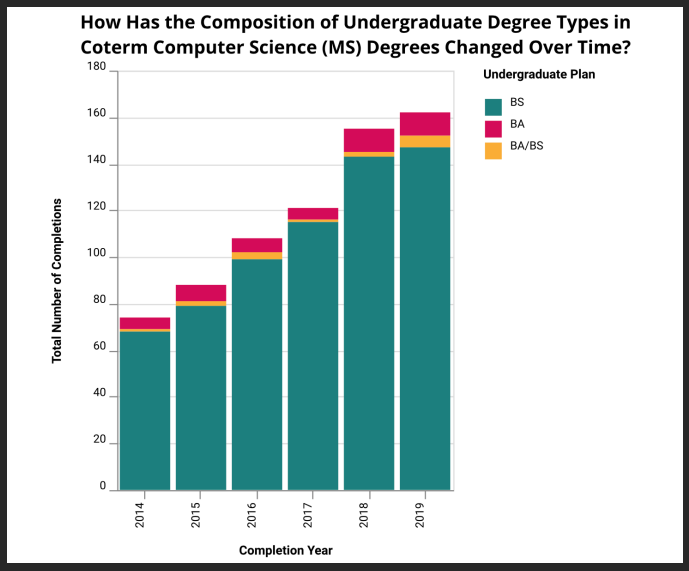
Be consistent! Visual inferences should consistently support data inferences

27

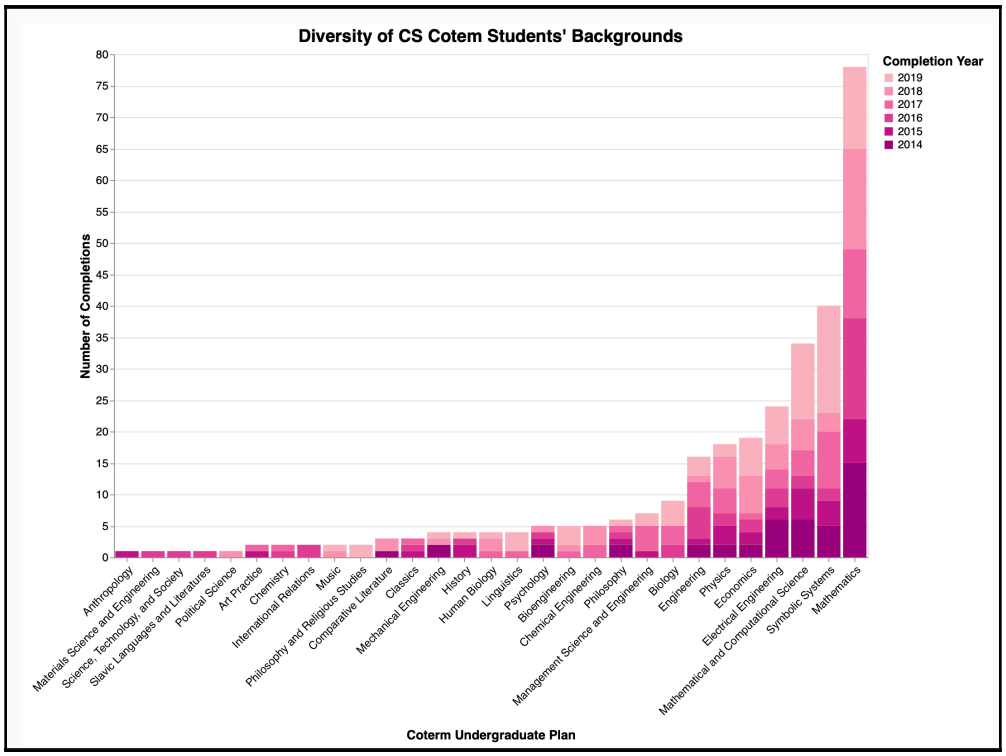
Stacked bar charts

(most common)

29



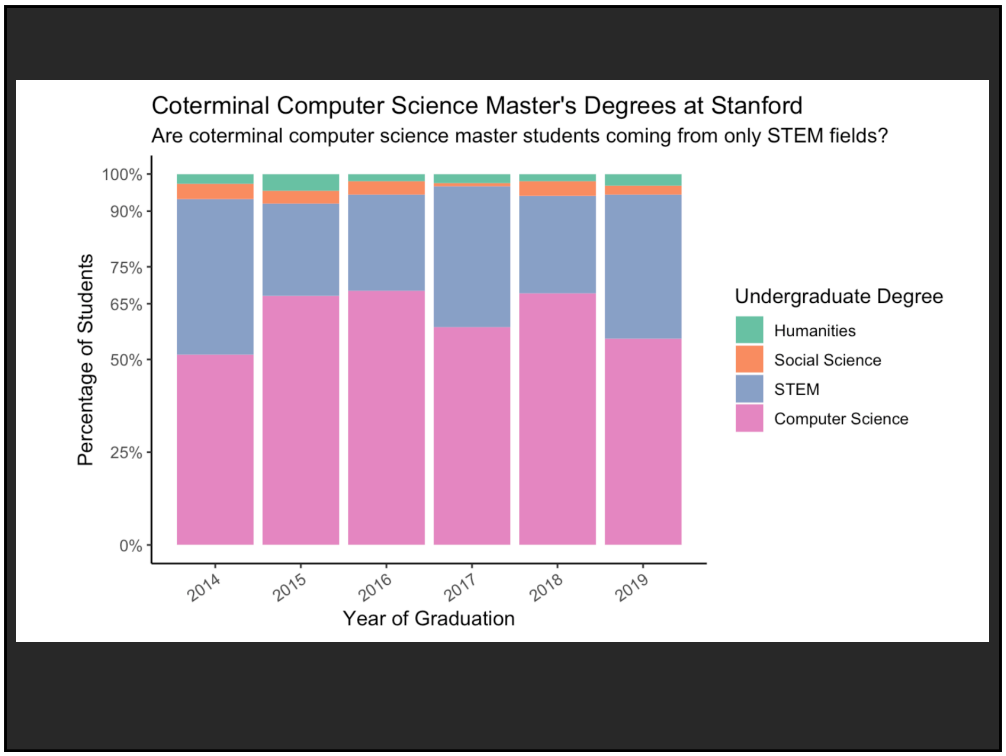
30



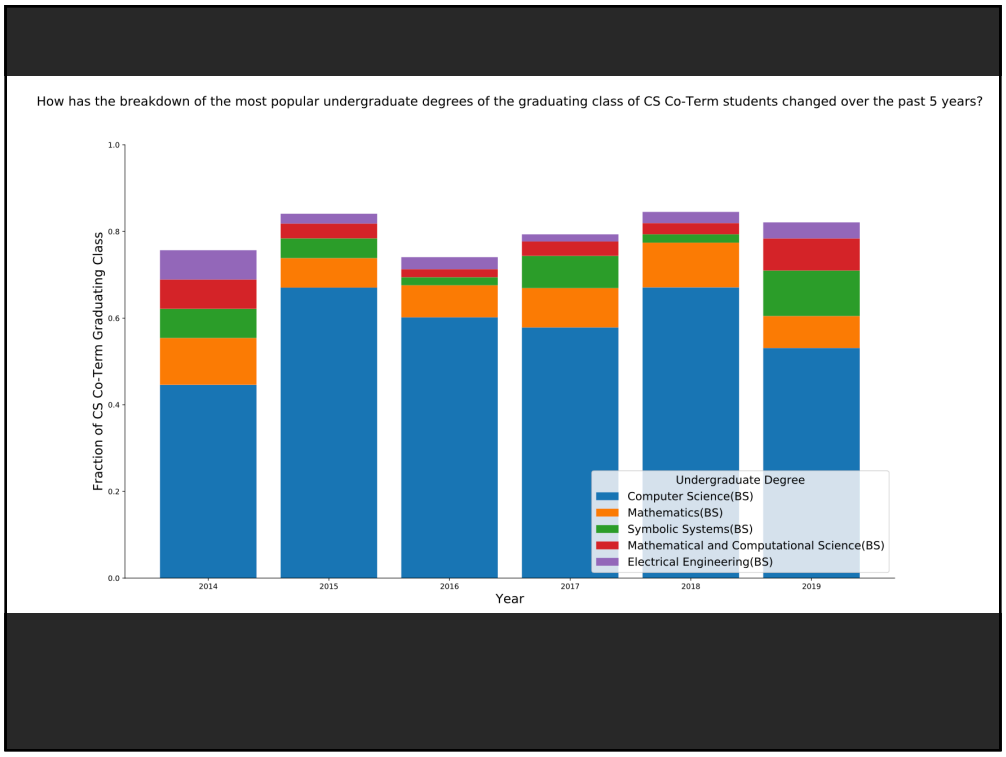
31

Stacked bar charts – for percentages

32


Coterminal Computer Science Master's Degrees at Stanford
Are coterminal computer science master students coming from only STEM fields?
Percentage of Students
100%
90%
75%
65%
50%
25%
0%
2014
2015
2016
2017
2018
2019
Year of Graduation
Undergraduate Degree
Humanities
Social Science
STEM
Computer Science

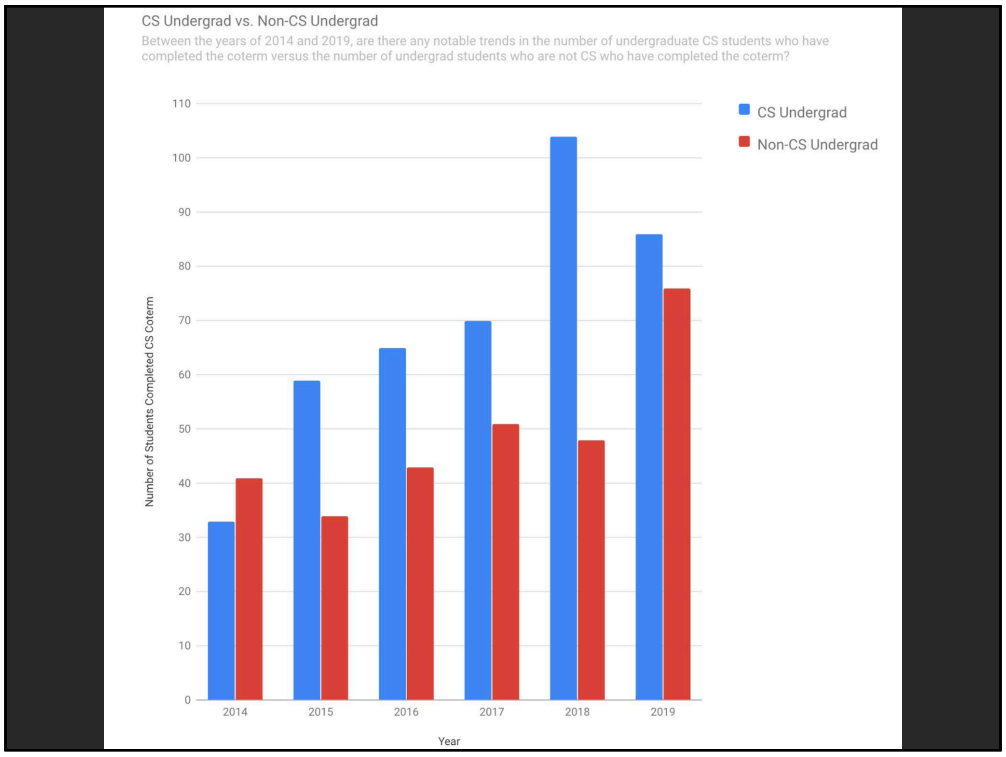
33


How has the breakdown of the most popular undergraduate degrees of the graduating class of CS Co-Term students changed over the past 5 years?
Fraction of CS Co-Term Graduating Class
1.0
0.8
0.6
0.4
0.2
0.0
2014
2015
2016
2017
2018
2019
Year
Undergraduate Degree
Computer Science(BS)
Mathematics(BS)
Symbolic Systems(BS)
Mathematical and Computational Science(BS)
Electrical Engineering(BS)

34

Grouped bar charts

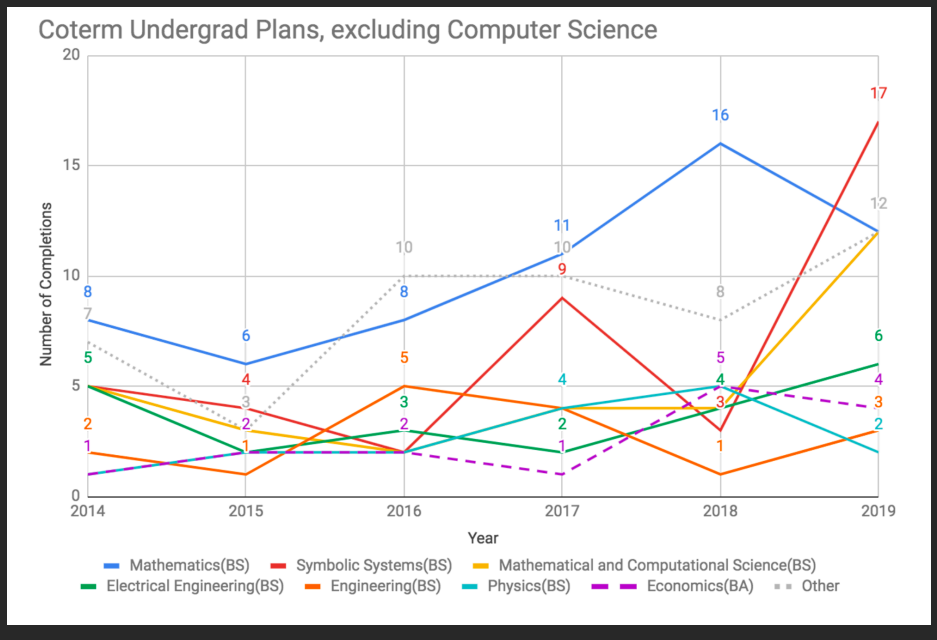
35



36

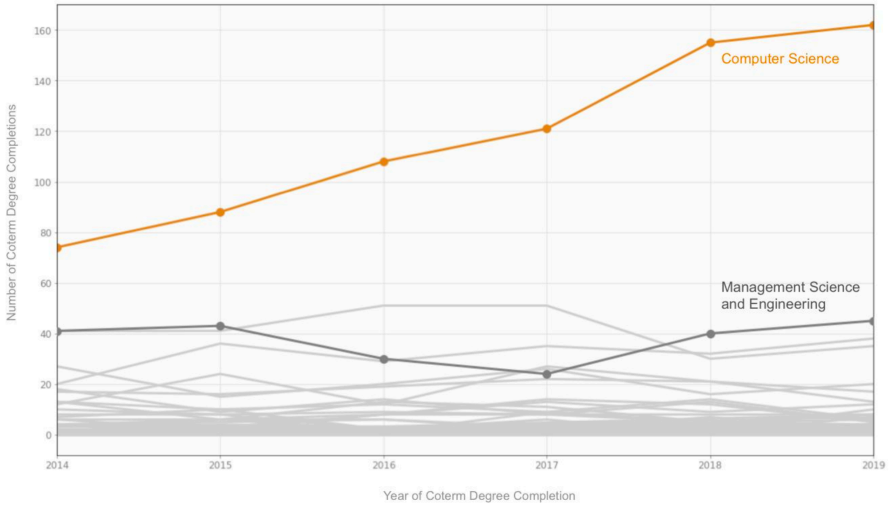
Line charts

37



38

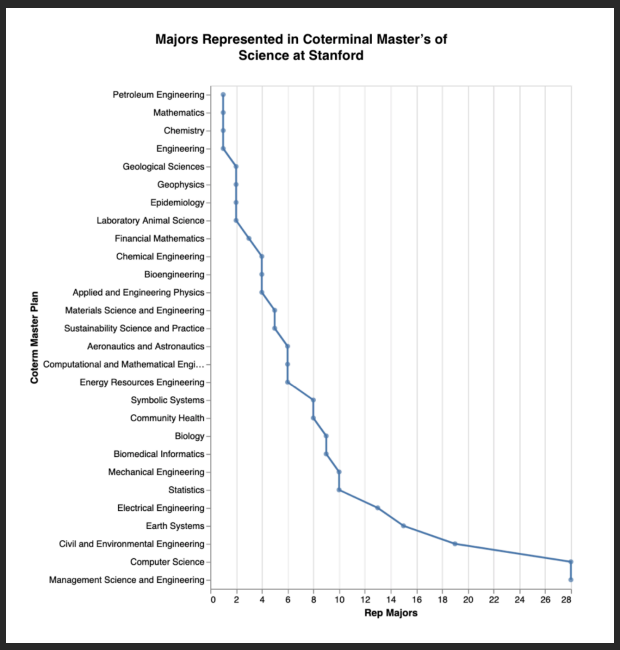
Q: How has the popularity of the computer science coterm grown over time compared to other coterm programs at Stanford?



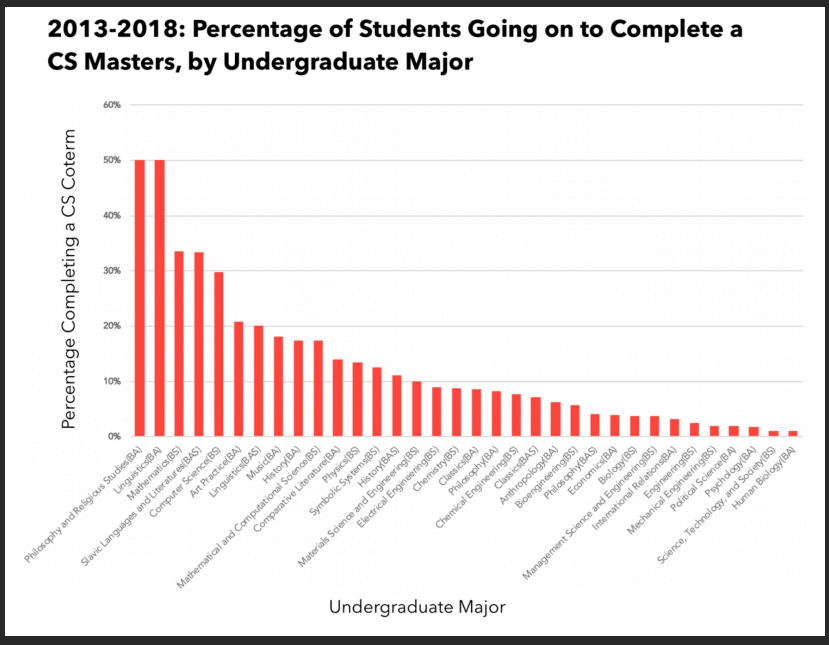
39

Additional Data and Transformations

40


Majors Represented in Coterminal Master's of Science at Stanford
Coterm Master Plan
Petroleum Engineering
Mathematics
Chemistry
Engineering
Geological Sciences
Geophysics
Epidemiology
Laboratory Animal Science
Financial Mathematics
Chemical Engineering
Bioengineering
Applied and Engineering Physics
Materials Science and Engineering
Sustainability Science and Practice
Aeronautics and Astronautics
Computational and Mathematical Engi...
Energy Resources Engineering
Symbolic Systems
Community Health
Biology
Biomedical Informatics
Mechanical Engineering
Statistics
Electrical Engineering
Earth Systems
Civil and Environmental Engineering
Computer Science
Management Science and Engineering
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28
Rep Majors

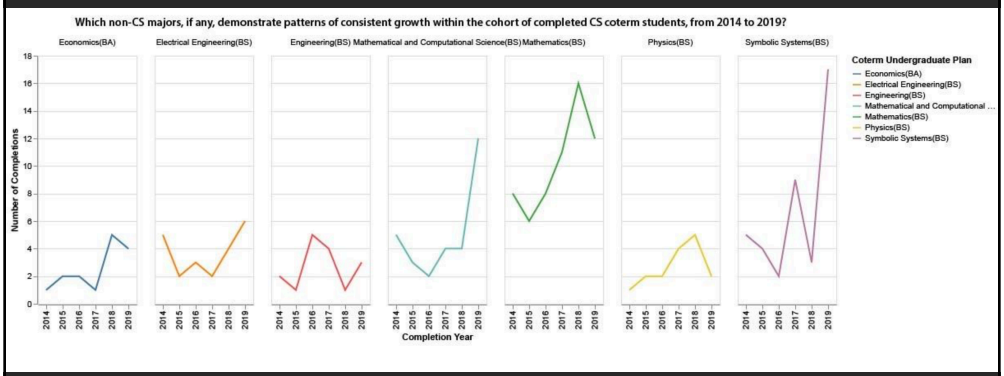
41


2013-2018: Percentage of Students Going on to Complete a CS Masters, by Undergraduate Major
Percentage Completing a CS Coterm
60%
50%
40%
30%
20%
10%
0%
Philosophy and Religious Studies(BA)
Linguistics(BA)
Mathematics(BS)
Slavic Languages and Literature(BAS)
Computer Science(BS)
Art Practice(BA)
Linguistics(BAS)
Music(BA)
History(BA)
Mathematical and Computational Science(BS)
Comparative Literature(BA)
Physics(BS)
Symbolic Systems(BS)
History(BAS)
Materials Science and Engineering(BS)
Electrical Engineering(BS)
Chemistry(BS)
Classics(BA)
Philosophy(BA)
Chemical Engineering(BS)
Classics(BAS)
Anthropology(BA)
Bioengineering(BS)
Philosophy(BAS)
Economics(BA)
Biology(BS)
Management Science and Engineering(BS)
International Relations(BA)
Engineering(BS)
Mechanical Engineering(BS)
Political Science(BA)
Psychology(BA)
Science, Technology, and Society(BS)
Human Biology(BA)
Undergraduate Major

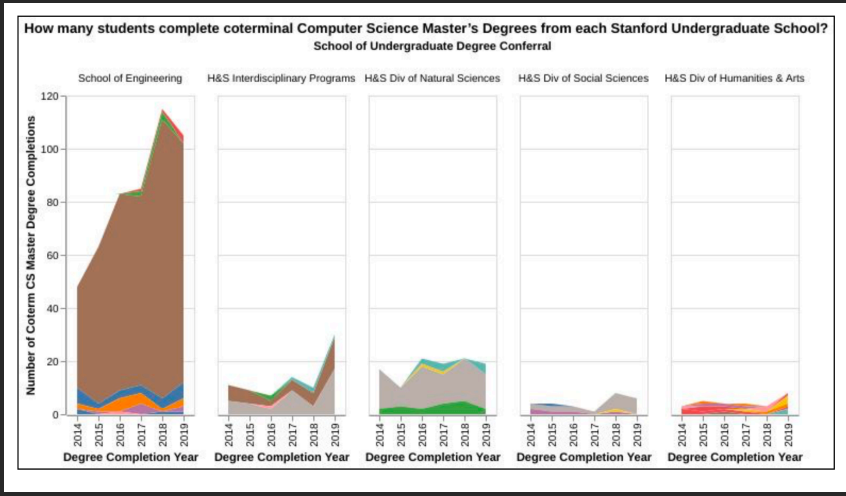
42

Small Multiples

43



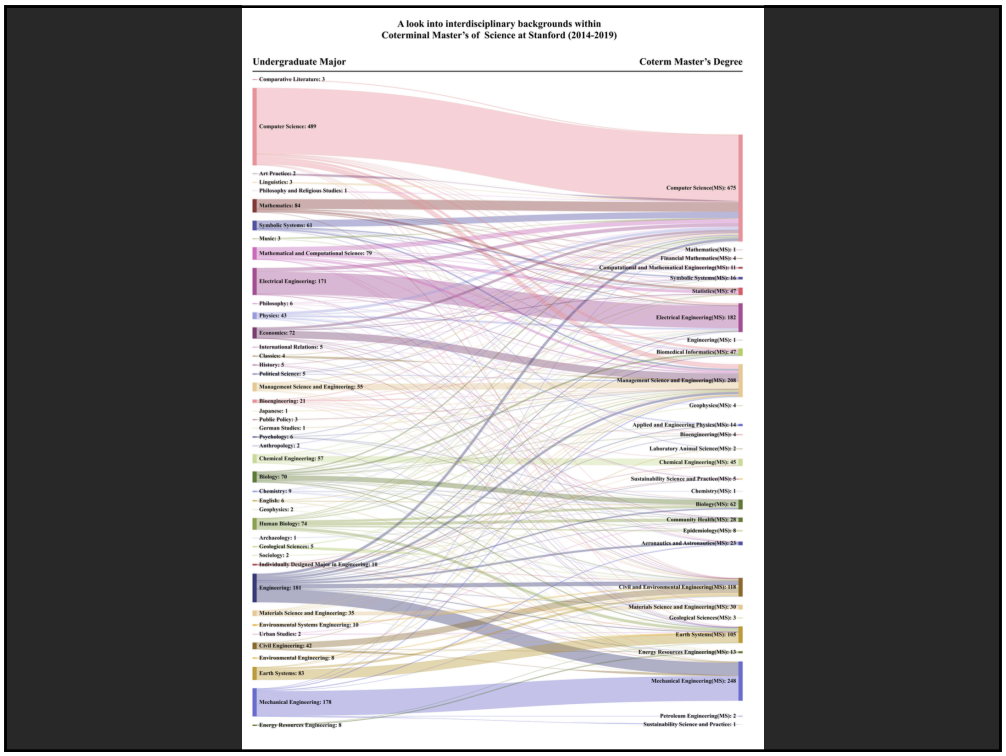
44



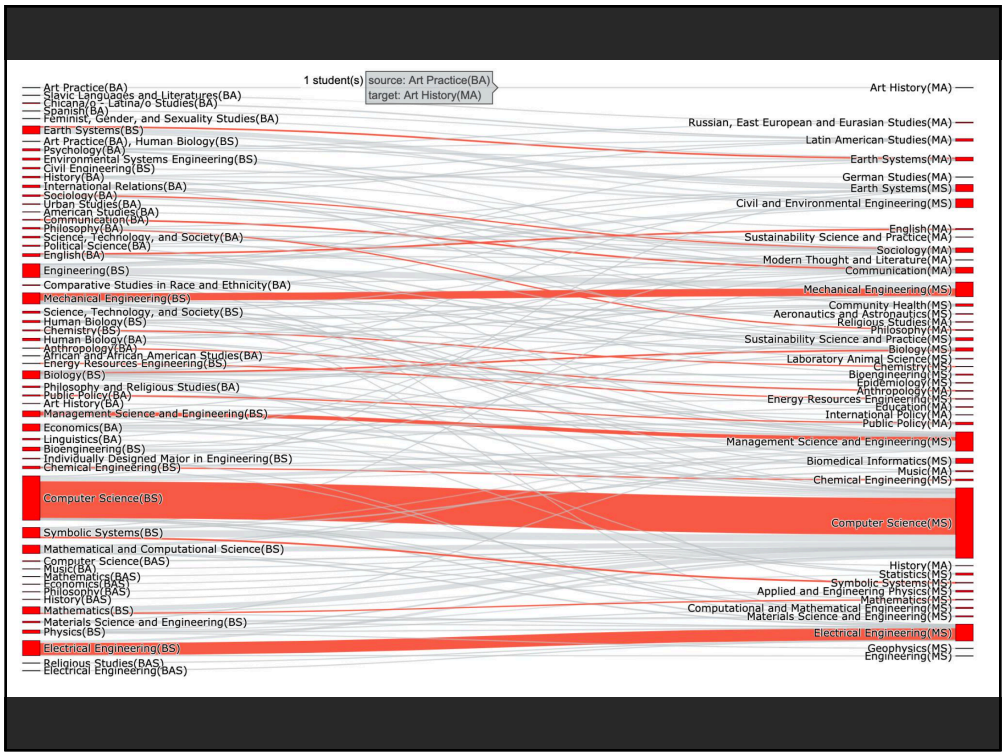
45

Sankey diagrams

46


A look into interdisciplinary backgrounds within
Coterminal Master's of Science at Stanford (2014-2019)
Undergraduate Major
Coterm Master's Degree
Comparative Literature: 3
Computer Science: 489
Art Practice: 2
Linguistics: 3
Philosophy and Religious Studies: 1
Mathematics: 84
Symbolic Systems: 61
Music: 3
Mathematical and Computational Science: 79
Electrical Engineering: 171
Philosophy: 6
Physics: 43
Economics: 72
International Relations: 5
Classics: 4
History: 5
Political Science: 5
Management Science and Engineering: 55
Bioengineering: 21
Japanese: 1
Public Policy: 3
German Studies: 1
Psychology: 6
Anthropology: 2
Chemical Engineering: 57
Biology: 70
Chemistry: 9
English: 6
Geophysics: 2
Human Biology: 74
Archaeology: 1
Geological Sciences: 5
Sociology: 2
Individually Designed Major in Engineering: 10
Engineering: 101
Materials Science and Engineering: 35
Environmental Systems Engineering: 10
Urban Studies: 2
Civil Engineering: 42
Environmental Engineering: 8
Earth Systems: 83
Mechanical Engineering: 178
Energy Resources Engineering: 8
Computer Science(MS): 675
Mathematics(MS): 1
Financial Mathematics(MS): 4
Computational and Mathematical Engineering(MS): 11
Symbolic Systems(MS): 16
Statistics(MS): 47
Electrical Engineering(MS): 182
Engineering(MS): 1
Biomedical Informatics(MS): 47
Management Science and Engineering(MS): 208
Geophysics(MS): 4
Applied and Engineering Physics(MS): 14
Bioengineering(MS): 4
Laboratory Animal Science(MS): 2
Chemical Engineering(MS): 45
Sustainability Science and Practice(MS): 5
Chemistry(MS): 1
Biology(MS): 62
Community Health(MS): 28
Epidemiology(MS): 8
Aeronautics and Astronautics(MS): 23
Civil and Environmental Engineering(MS): 118
Materials Science and Engineering(MS): 30
Geological Sciences(MS): 3
Earth Systems(MS): 105
Energy Resources Engineering(MS): 13
Mechanical Engineering(MS): 248
Petroleum Engineering(MS): 2
Sustainability Science and Practice: 1

47

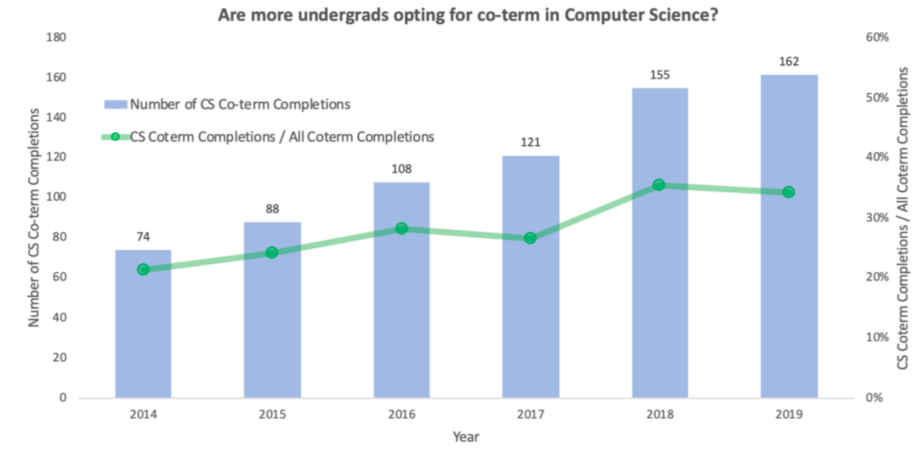

1 student(s)
source: Art Practice(BA)
target: Art History(MA)
Art Practice(BA)
Slavic Languages and Literatures(BA)
Chicana/o - Latina/o Studies(BA)
Spanish(BA)
Feminist, Gender, and Sexuality Studies(BA)
Earth Systems(BS)
Art Practice(BA), Human Biology(BS)
Psychology(BA)
Environmental Systems Engineering(BS)
Civil Engineering(BS)
History(BA)
International Relations(BA)
Sociology(BA)
Urban Studies(BA)
American Studies(BA)
Communication(BA)
Philosophy(BA)
Science, Technology, and Society(BA)
Political Science(BA)
English(BA)
Engineering(BS)
Comparative Studies in Race and Ethnicity(BA)
Mechanical Engineering(BS)
Science, Technology, and Society(BS)
Human Biology(BS)
Chemistry(BS)
Human Biology(BA)
Anthropology(BA)
African and African American Studies(BA)
Energy Resources Engineering(BS)
Biology(BS)
Philosophy and Religious Studies(BA)
Public Policy(BA)
Art History(BA)
Management Science and Engineering(BS)
Economics(BA)
Linguistics(BA)
Bioengineering(BS)
Individually Designed Major in Engineering(BS)
Chemical Engineering(BS)
Computer Science(BS)
Symbolic Systems(BS)
Mathematical and Computational Science(BS)
Computer Science(BAS)
Music(BA)
Mathematics(BAS)
Economics(BAS)
Philosophy(BAS)
History(BAS)
Mathematics(BS)
Materials Science and Engineering(BS)
Physics(BS)
Electrical Engineering(BS)
Religious Studies(BAS)
Electrical Engineering(BAS)
Art History(MA)
Russian, East European and Eurasian Studies(MA)
Latin American Studies(MA)
Earth Systems(MA)
German Studies(MA)
Earth Systems(MS)
Civil and Environmental Engineering(MS)
English(MA)
Sustainability Science and Practice(MA)
Sociology(MA)
Modern Thought and Literature(MA)
Communication(MA)
Mechanical Engineering(MS)
Community Health(MS)
Aeronautics and Astronautics(MS)
Religious Studies(MA)
Philosophy(MA)
Sustainability Science and Practice(MS)
Biology(MS)
Laboratory Animal Science(MS)
Chemistry(MS)
Bioengineering(MS)
Epidemiology(MS)
Anthropology(MA)
Energy Resources Engineering(MS)
Education(MA)
International Policy(MA)
Public Policy(MA)
Management Science and Engineering(MS)
Biomedical Informatics(MS)
Music(MA)
Chemical Engineering(MS)
Computer Science(MS)
History(MA)
Statistics(MS)
Symbolic Systems(MS)
Applied and Engineering Physics(MS)
Mathematics(MS)
Computational and Mathematical Engineering(MS)
Materials Science and Engineering(MS)
Electrical Engineering(MS)
Geophysics(MS)
Engineering(MS)

48

Multiple Encodings

49

Question: Are more undergrads opting for co-term in Computer Science?



50

How does the Number of CS Coterm MS Degrees Awarded Change based on Coterm Undergraduate Plan?

by - January 13, 2020

Change in the Number of CS Coterm MS Degrees Awarded based on Undergraduate Plan

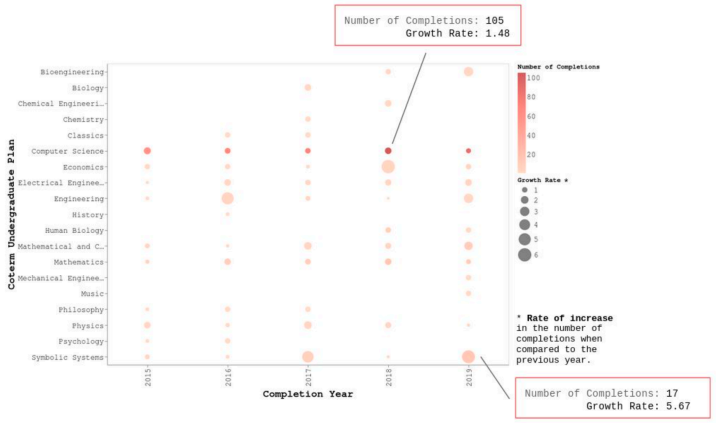


Figure 1. Visualization showing the change in the number of CS coterm MS degrees awarded based on undergraduate plan, from 2015 to 2019.

51