Visualization Design and Redesign

Maneesh Agrawala

CS 448B: Visualization
Fall 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: =, ≠

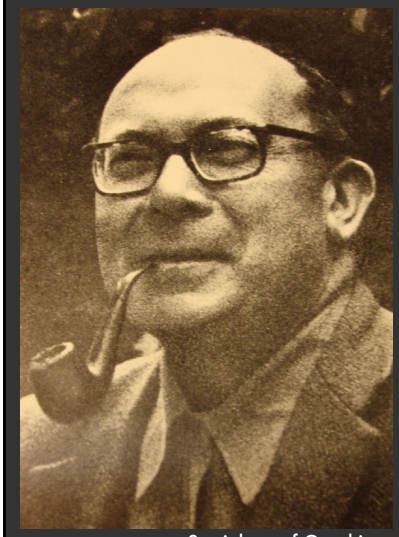
O - Ordered
Quality of meat: Grade A, AA, AAA
Operations: =, ≠, <, >

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

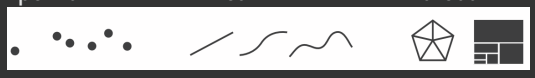
Marks and Visual Variables



Semiology of Graphics
J. Bertin, 1967

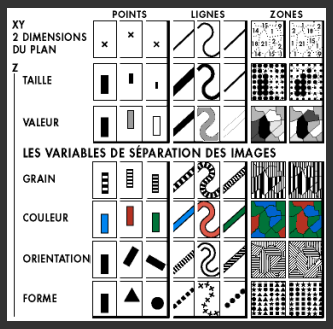
Marks: geometric primitives

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		

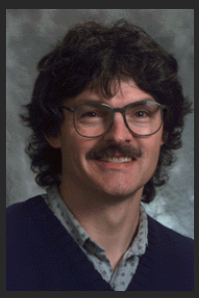
N Nominal
O Ordered
Q Quantitative

Note: Q < O < N

7

Automated design

Jock Mackinlay's APT 86



8

Combinatorics of encodings

Challenge:

- Assume 8 visual encodings and n data fields
- Pick the best encoding from the exponential number of possibilities $(n+1)^8$

9

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

10

Mackinlay's expressiveness criteria

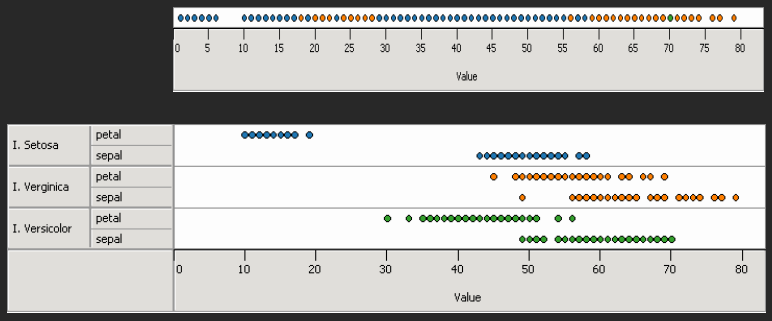
Expressiveness

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.

11

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



12

Expresses facts not in the data

A length is interpreted as a quantitative value;

∴ Length of bar says something untrue about N data

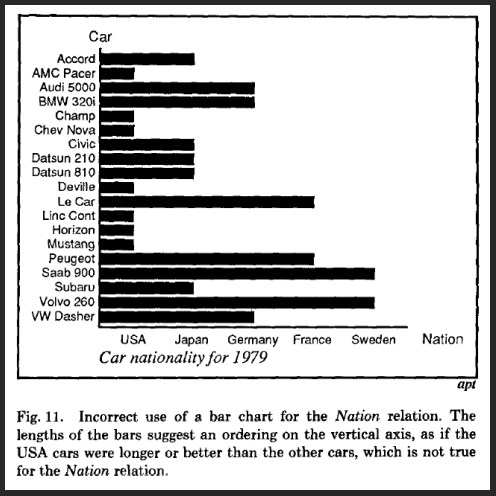


Fig. 11. Incorrect use of a bar chart for the *Nation* relation. The lengths of the bars suggest an ordering on the vertical axis, as if the USA cars were longer or better than the other cars, which is not true for the *Nation* relation.

[Mackinlay, APT, 1986]

13

Mackinlay's effectiveness criteria

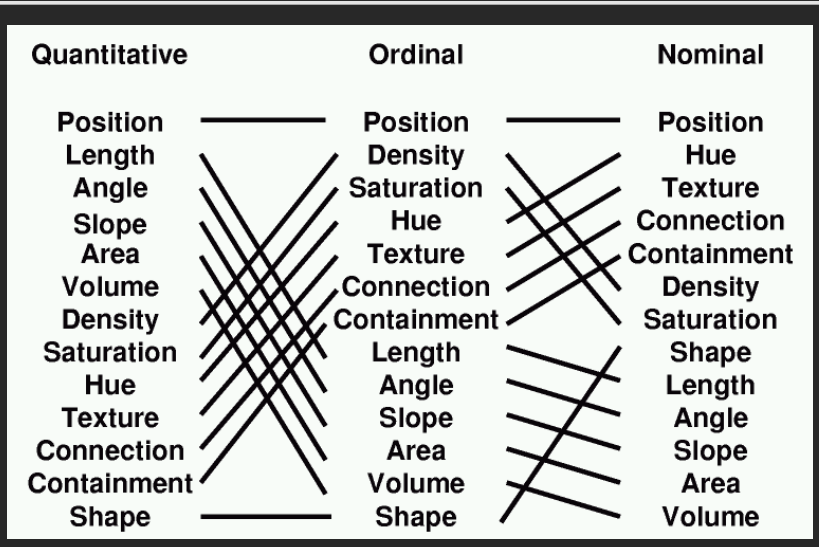
Effectiveness

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

14

Mackinlay's ranking



Conjectured *effectiveness* of the encoding

15

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

17

Automatic chart construction

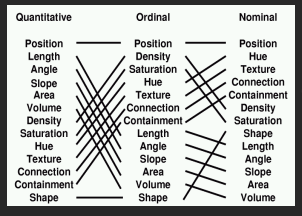
Automating the design of graphical presentation of relational information
J. Mackinlay, 1986

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

Price	Mileage	Weight
13,500	22	3000
7,200	31	1500
11,300	12	4200
...	...	...

→

1. Price (Q)
2. Mileage (Q)
3. Weight (Q)



→

mark: lines
Price → y-pos (Q)
Mileage → x-pos (Q)
Weight → size (Q)

18

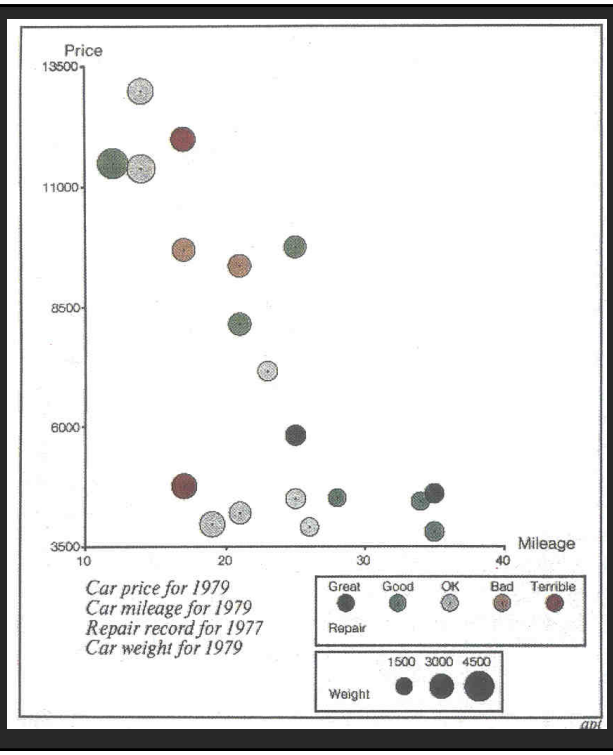
Cars Data

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]



19

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

20

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**

21

Announcements

22

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-fa20

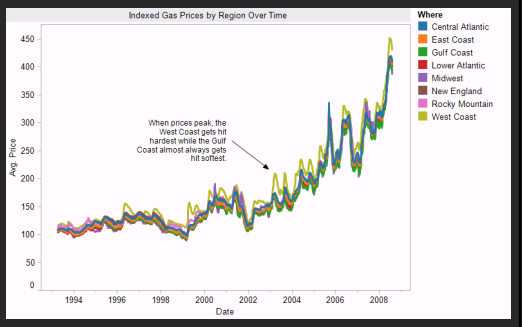
23

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 Oct 6, 2020

24

A1 Review

25

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

27

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

28

Participating in the discussion

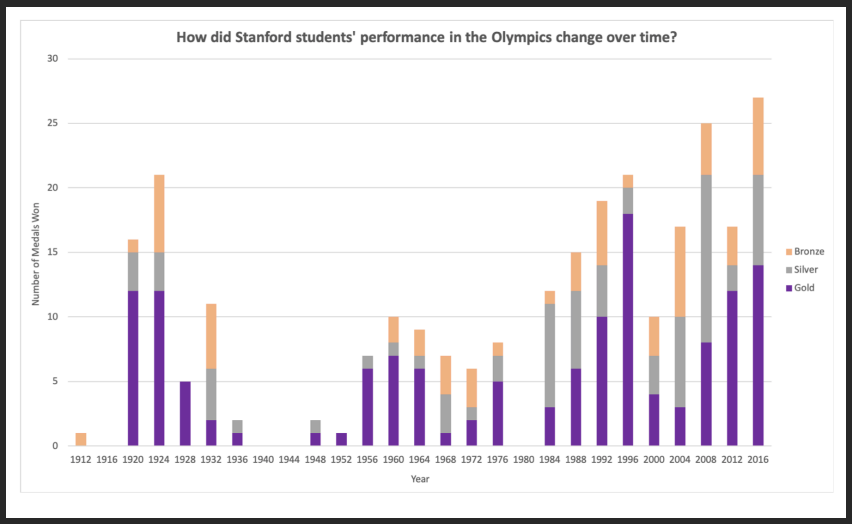
Let's try this

- Everyone unmute
- Speak raise hand in zoom and I'll call on you
- Turn on video if you are comfortable

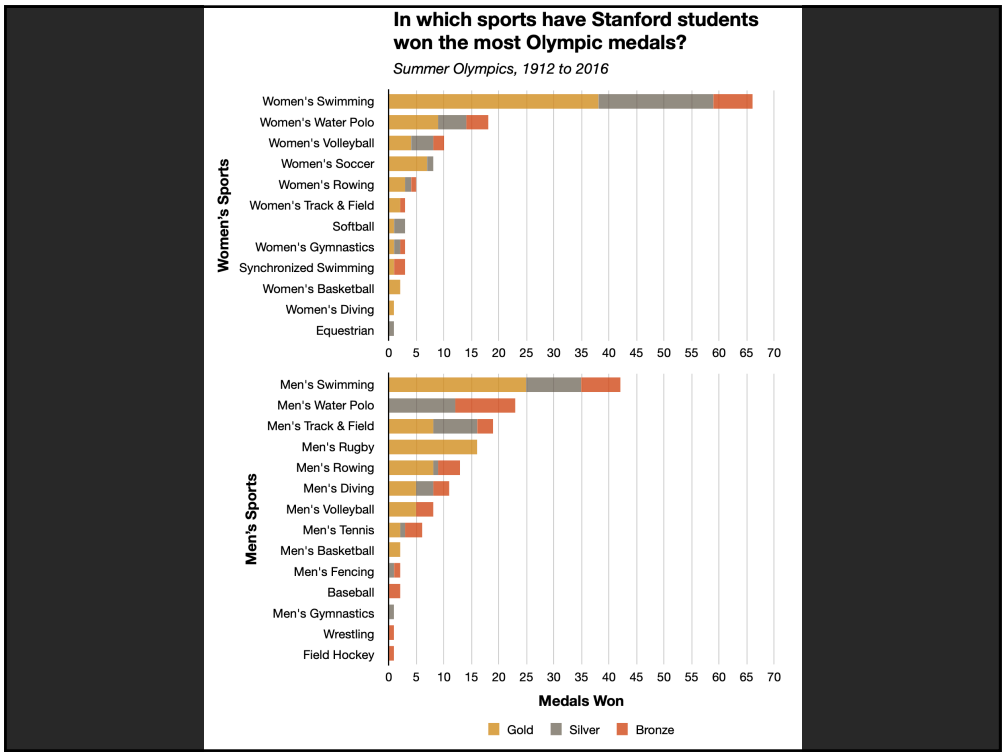
29

Stacked Bar Charts
(most common)

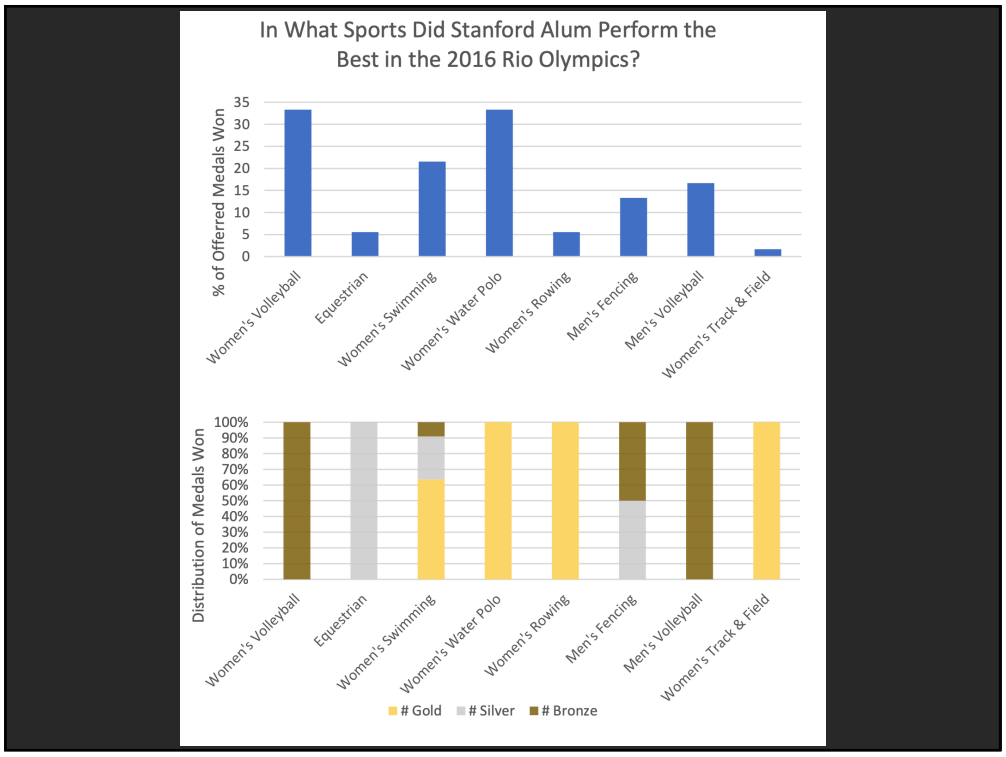
34



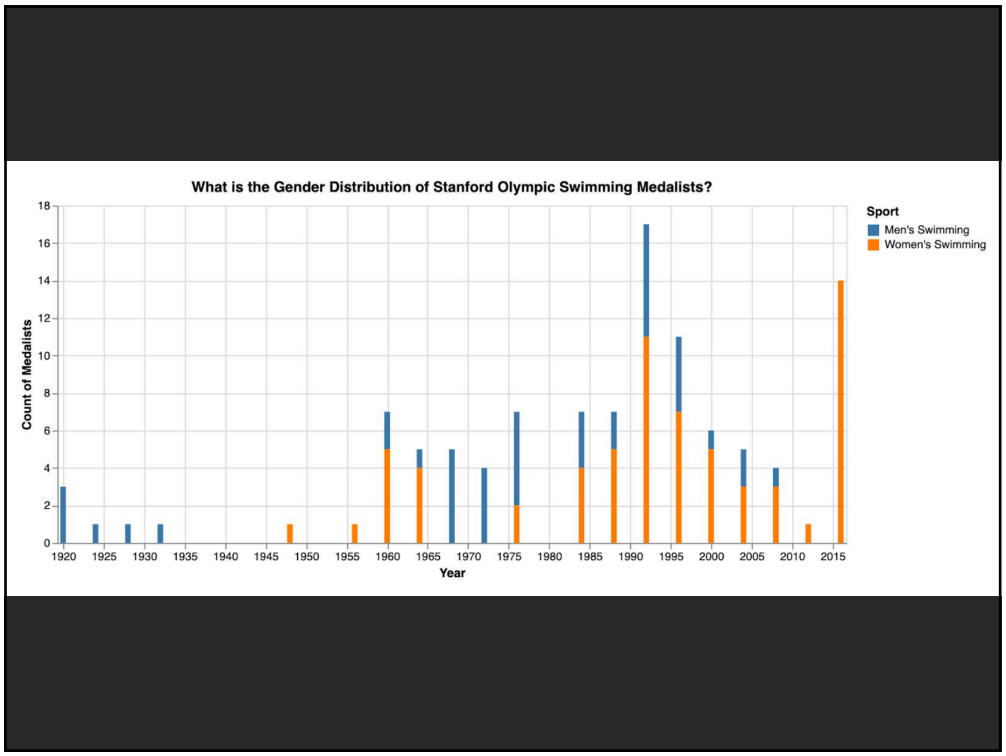
36


In which sports have Stanford students won the most Olympic medals?
Summer Olympics, 1912 to 2016
Women's Sports
Women's Swimming
Women's Water Polo
Women's Volleyball
Women's Soccer
Women's Rowing
Women's Track & Field
Softball
Women's Gymnastics
Synchronized Swimming
Women's Basketball
Women's Diving
Equestrian
Men's Sports
Men's Swimming
Men's Water Polo
Men's Track & Field
Men's Rugby
Men's Rowing
Men's Diving
Men's Volleyball
Men's Tennis
Men's Basketball
Men's Fencing
Baseball
Men's Gymnastics
Wrestling
Field Hockey
0 5 10 15 20 25 30 35 40 45 50 55 60 65 70
Medals Won
Gold
Silver
Bronze

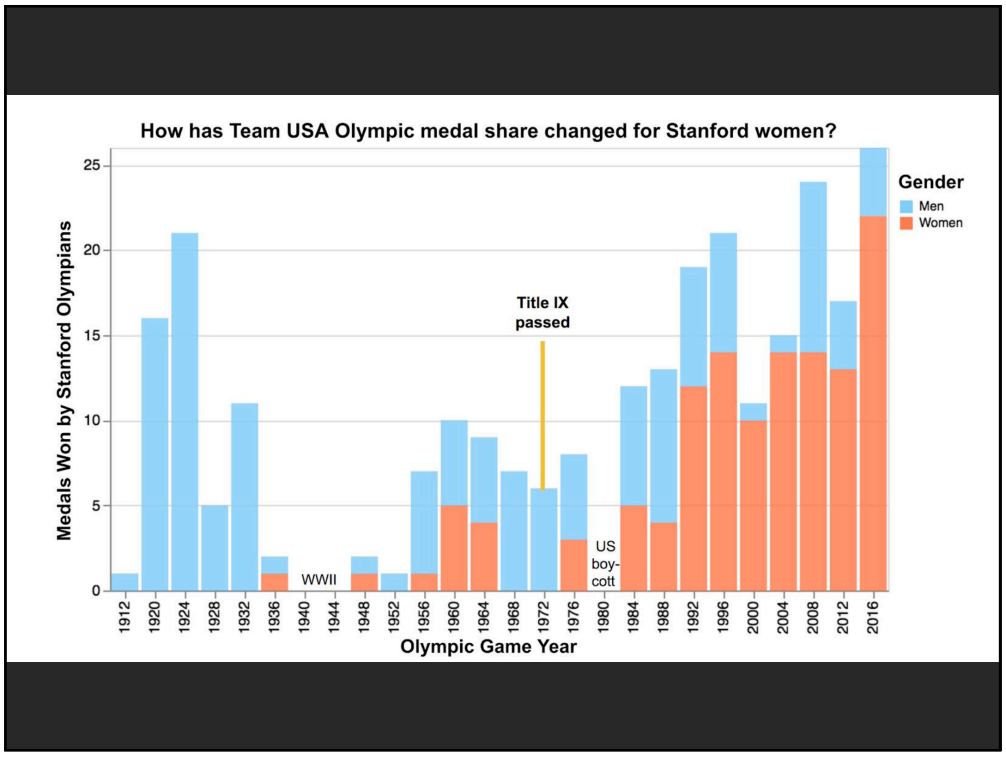
37


In What Sports Did Stanford Alum Perform the Best in the 2016 Rio Olympics?
% of Offerred Medals Won
35
30
25
20
15
10
5
0
Women's Volleyball
Equestrian
Women's Swimming
Women's Water Polo
Women's Rowing
Men's Fencing
Men's Volleyball
Women's Track & Field
Distribution of Medals Won
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Women's Volleyball
Equestrian
Women's Swimming
Women's Water Polo
Women's Rowing
Men's Fencing
Men's Volleyball
Women's Track & Field
Gold
Silver
Bronze

38


What is the Gender Distribution of Stanford Olympic Swimming Medalists?
Count of Medalists
Year
Sport
Men's Swimming
Women's Swimming

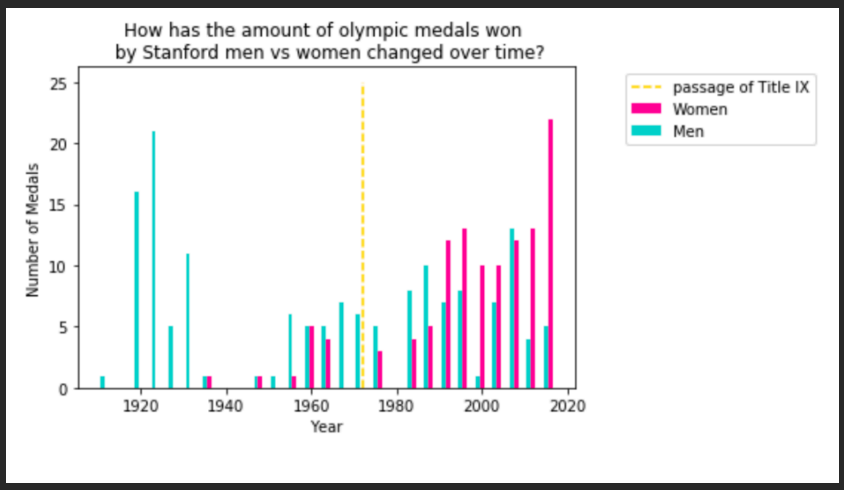
39


How has Team USA Olympic medal share changed for Stanford women?
Medals Won by Stanford Olympians
Olympic Game Year
Gender
Men
Women
Title IX passed
WWII
US boy-cott

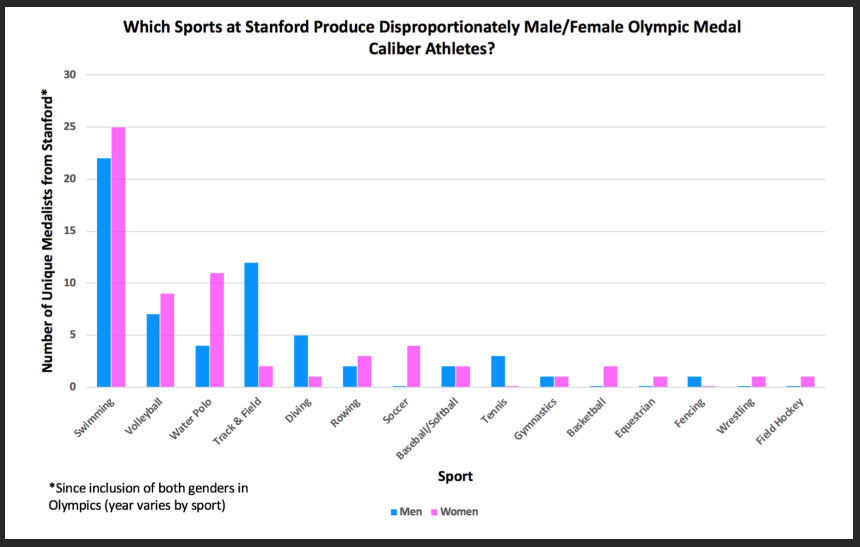
40

Grouped Bar Charts

47



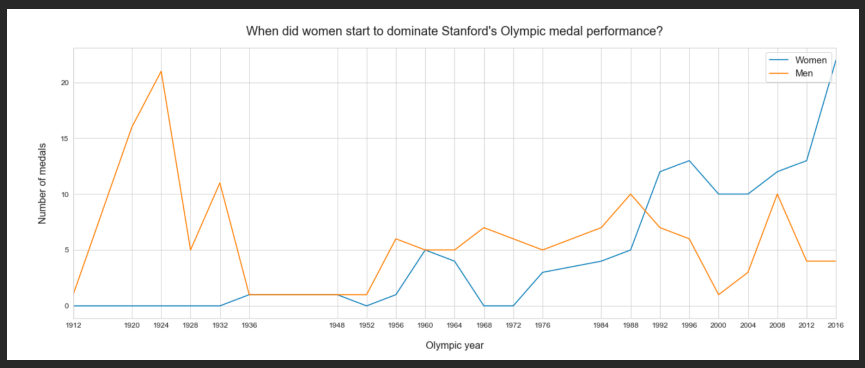
48



49

Line Charts

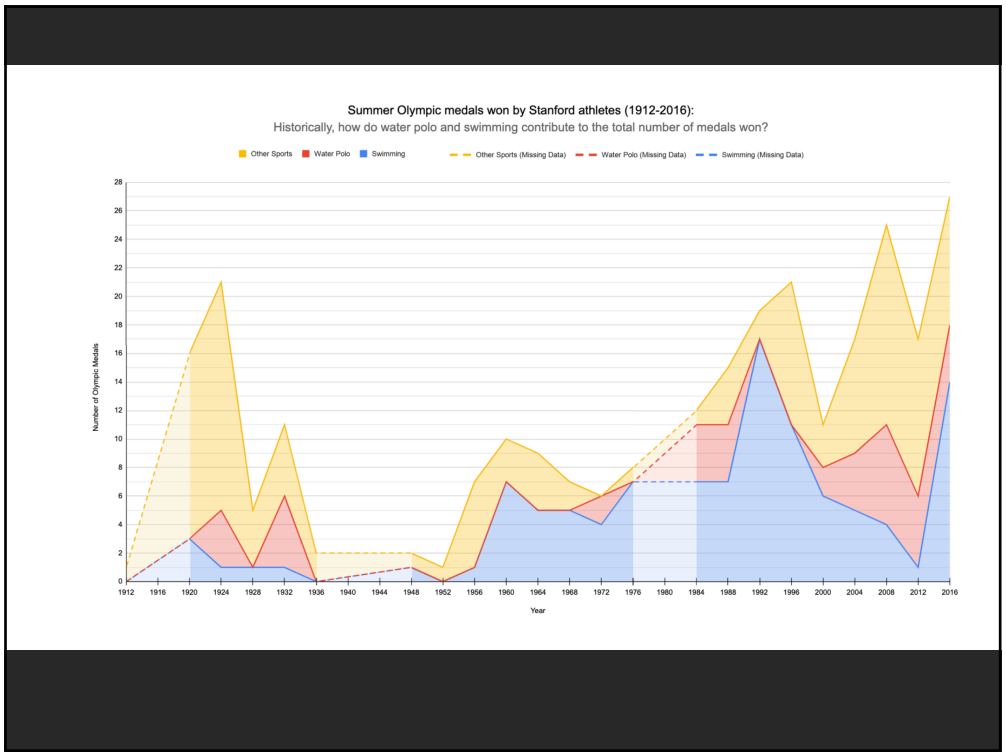
50



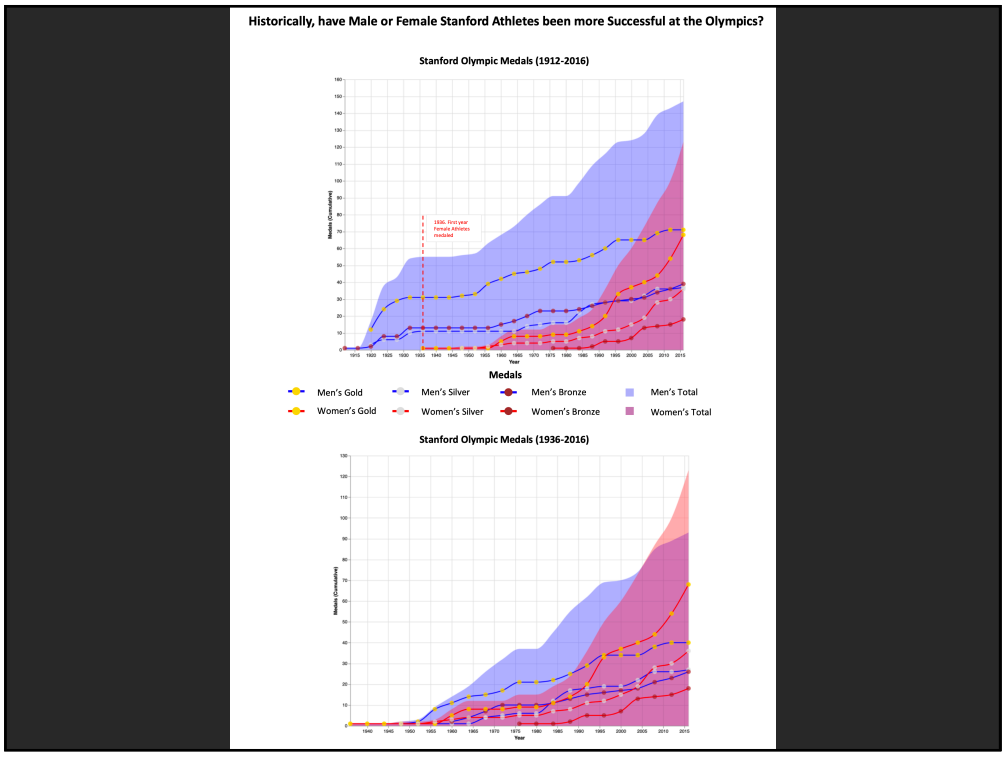
51

Area Charts

54


Summer Olympic medals won by Stanford athletes (1912-2016):
Historically, how do water polo and swimming contribute to the total number of medals won?
Other Sports
Water Polo
Swimming
Other Sports (Missing Data)
Water Polo (Missing Data)
Swimming (Missing Data)
Number of Olympic Medals
Year

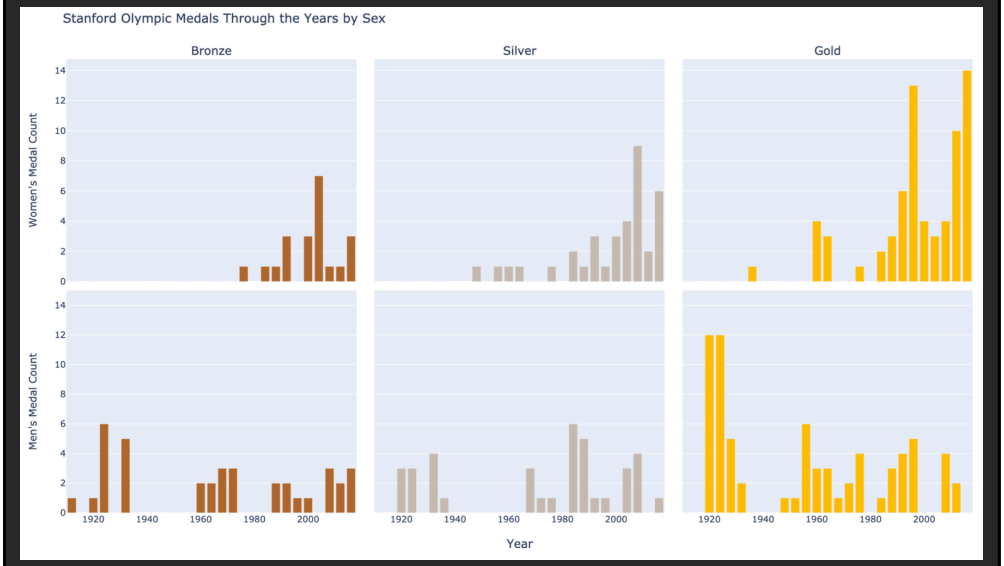
55


Historically, have Male or Female Stanford Athletes been more Successful at the Olympics?
Stanford Olympic Medals (1912-2016)
1936, First year Female Athletes medaled
Year
Medals
Men's Gold
Men's Silver
Men's Bronze
Men's Total
Women's Gold
Women's Silver
Women's Bronze
Women's Total
Stanford Olympic Medals (1936-2016)
Year

56

Small Multiples

57

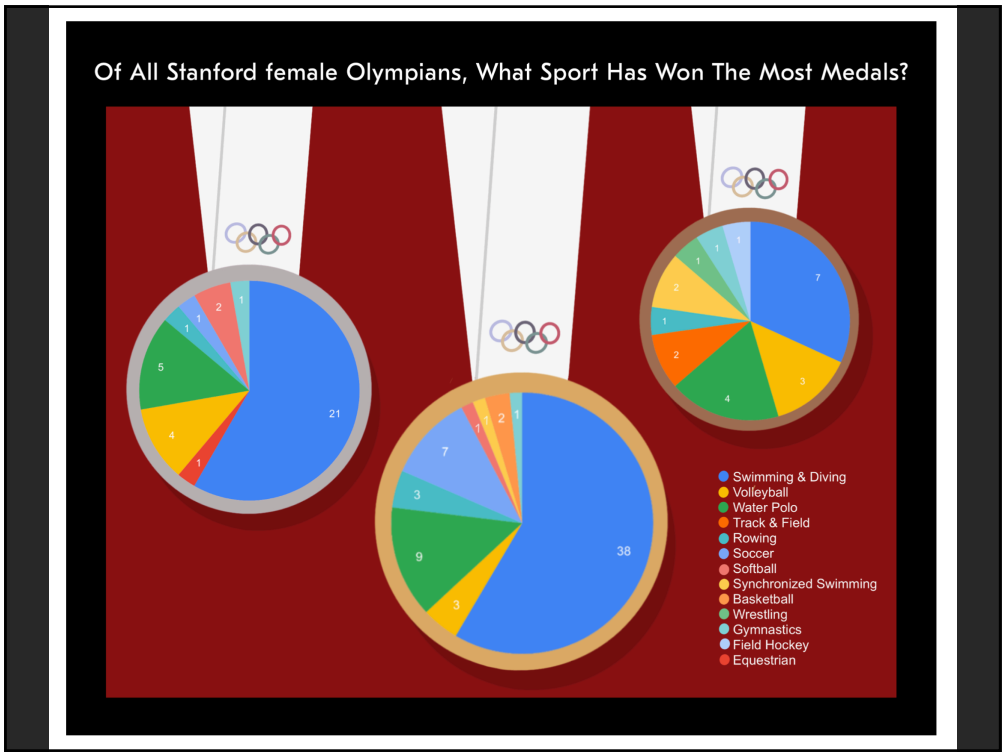


58

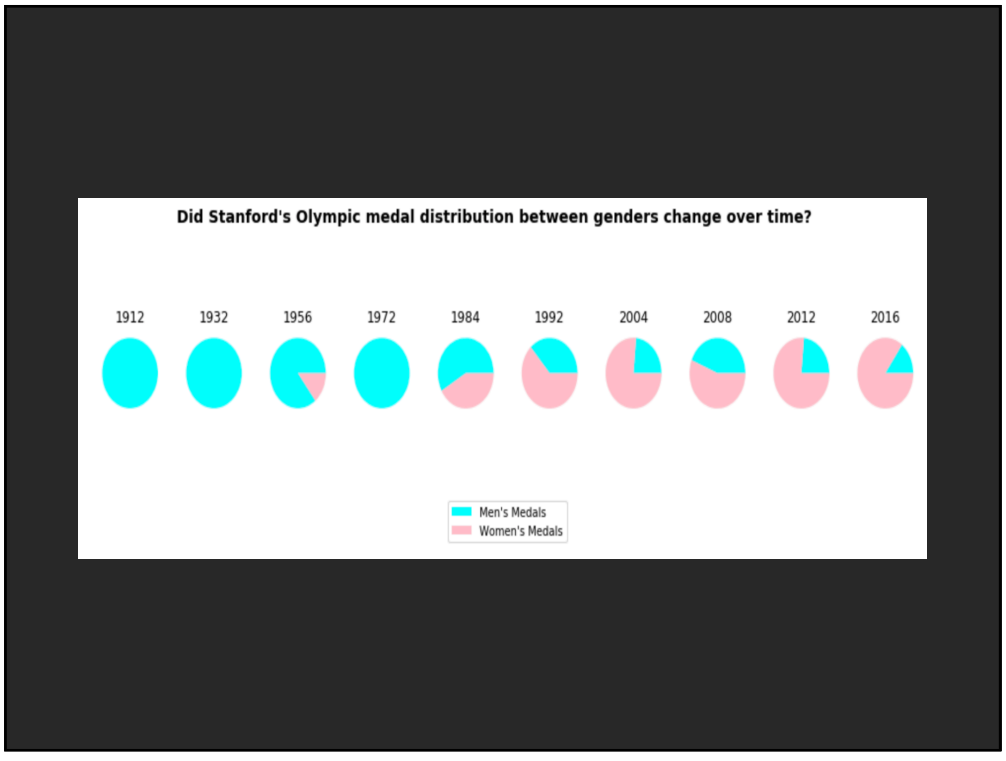
59

Pie Charts

60


Of All Stanford female Olympians, What Sport Has Won The Most Medals?
21
5
4
1
1
1
2
1
38
9
3
7
3
1
1
2
1
7
3
4
2
1
1
1
1
Swimming & Diving
Volleyball
Water Polo
Track & Field
Rowing
Soccer
Softball
Synchronized Swimming
Basketball
Wrestling
Gymnastics
Field Hockey
Equestrian

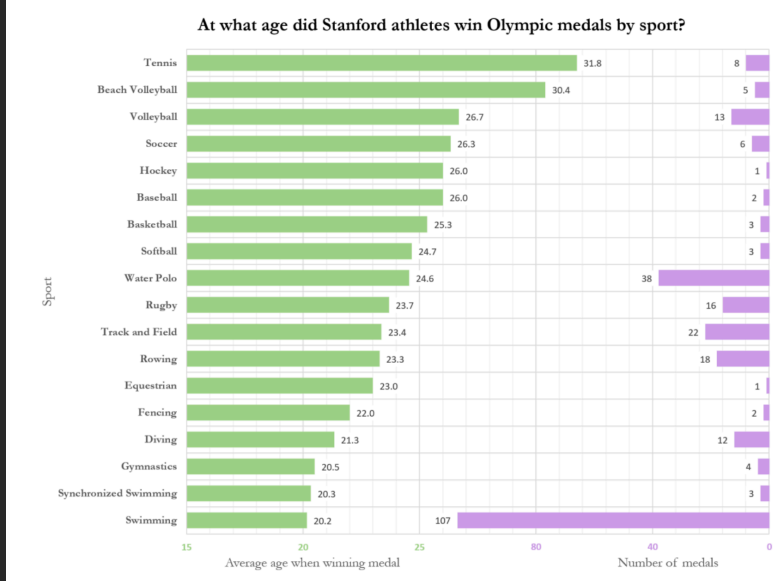
61


Did Stanford's Olympic medal distribution between genders change over time?
1912
1932
1956
1972
1984
1992
2004
2008
2012
2016
Men's Medals
Women's Medals

64

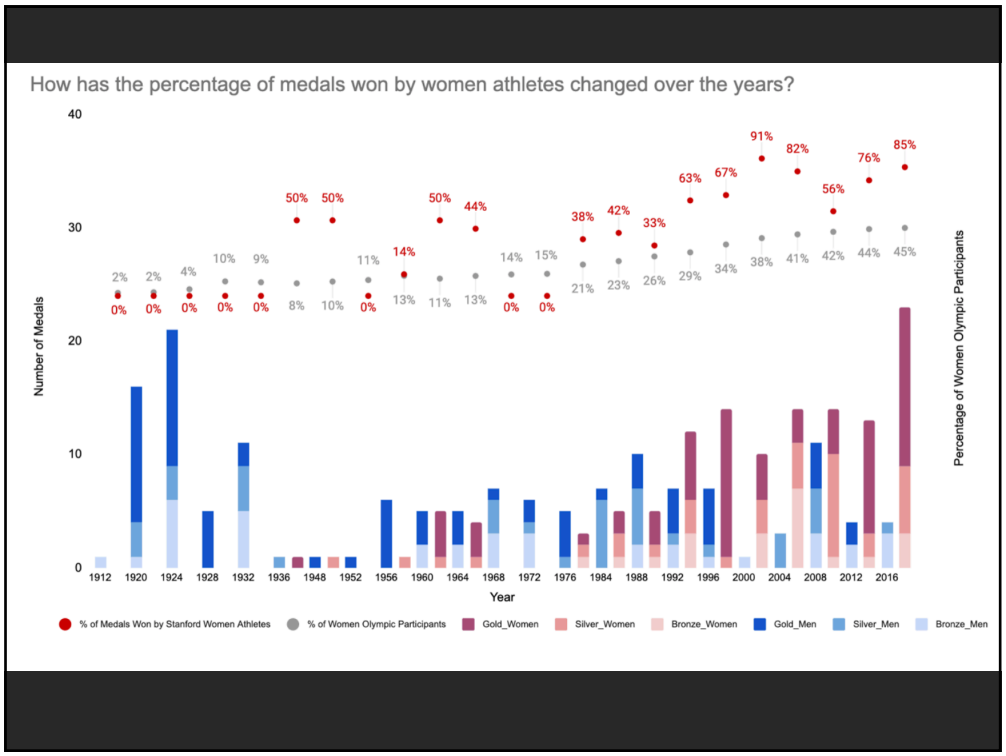
Multiple Encodings

66



age data is from https://www.kaggle.com/heesoo37/120-years-of-olympic-history-athletes-and-results

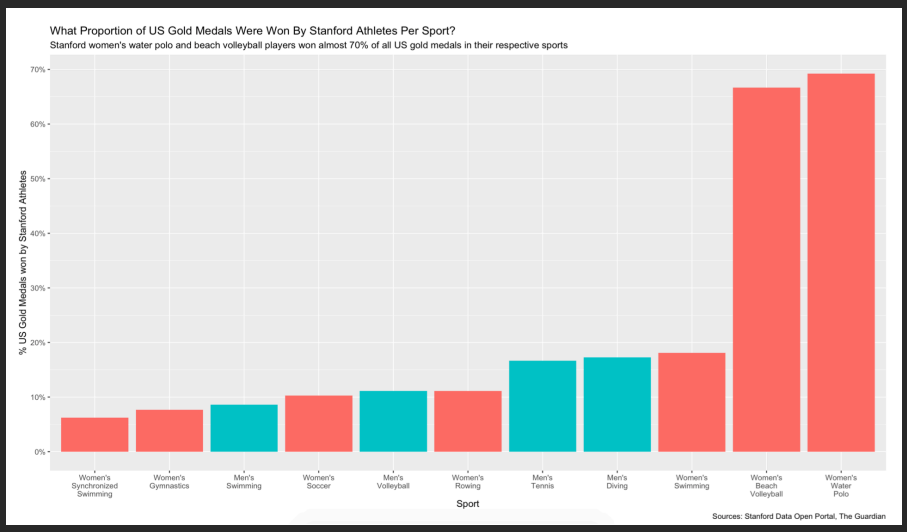
67



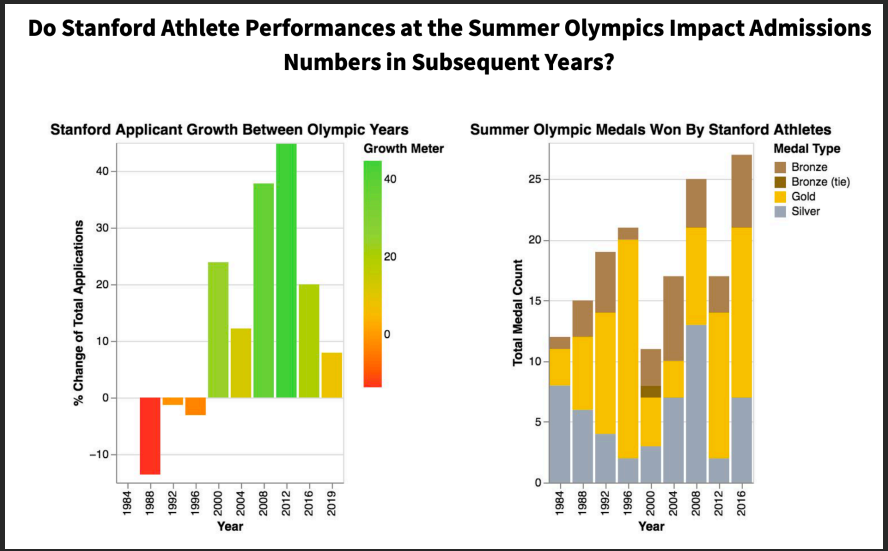
68

Additional Data

70


What Proportion of US Gold Medals Were Won By Stanford Athletes Per Sport?
Stanford women's water polo and beach volleyball players won almost 70% of all US gold medals in their respective sports
% US Gold Medals won by Stanford Athletes
70%
60%
50%
40%
30%
20%
10%
0%
Women's Synchronized Swimming
Women's Gymnastics
Men's Swimming
Women's Soccer
Men's Volleyball
Women's Rowing
Men's Tennis
Men's Diving
Women's Swimming
Women's Beach Volleyball
Women's Water Polo
Sport
Sources: Stanford Data Open Portal, The Guardian

71


Do Stanford Athlete Performances at the Summer Olympics Impact Admissions Numbers in Subsequent Years?
Stanford Applicant Growth Between Olympic Years
% Change of Total Applications
Growth Meter
40
20
0
Year
Summer Olympic Medals Won By Stanford Athletes
Total Medal Count
Medal Type
Bronze
Bronze (tie)
Gold
Silver
Year

72