

Framed!

Reproducing 150 Year Old Charts

Susan VanderPlas
Ryan Gorluch
Heike Hofmann

IOWA STATE UNIVERSITY



STATISTICAL ATLAS
OF THE
UNITED STATES
BASED ON THE RESULTS OF THE
NINTH CENSUS 1870

WITH CONTRIBUTIONS FROM MANY EMINENT MEN OF SCIENCE
AND SEVERAL DEPARTMENTS OF THE GOVERNMENT.

COMPILED UNDER AUTHORITY OF CONGRESS
BY
FRANCIS A. WALKER, M.A.

SUPERINTENDENT OF THE 9TH CENSUS.

PROFESSOR OF POLITICAL ECONOMY AND HISTORY,
SHEFFIELD SCIENTIFIC SCHOOL OF YALE COLLEGE.

Slides

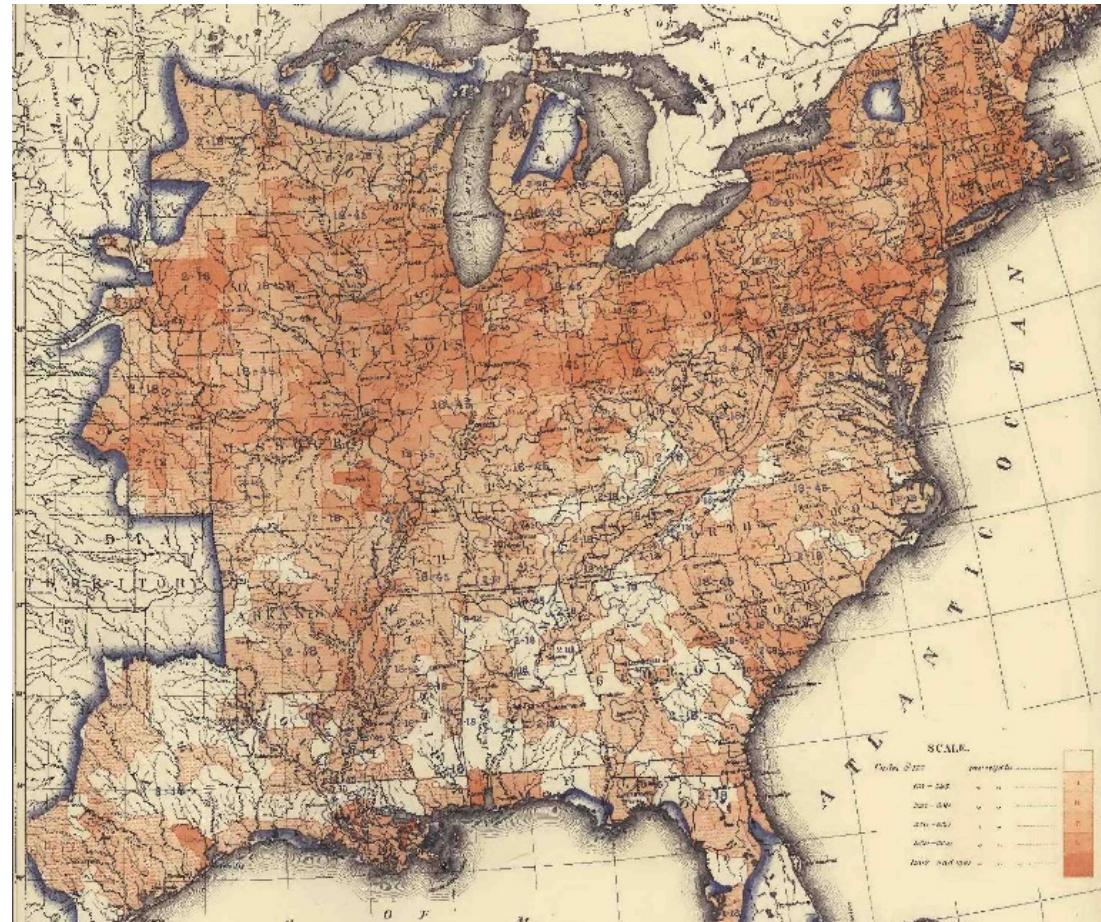
<http://rvanderplas.github.io/Presentations/JSM2018/Framed.html>

Project Repository

https://github.com/srvanderplas/Statistical_Atlas

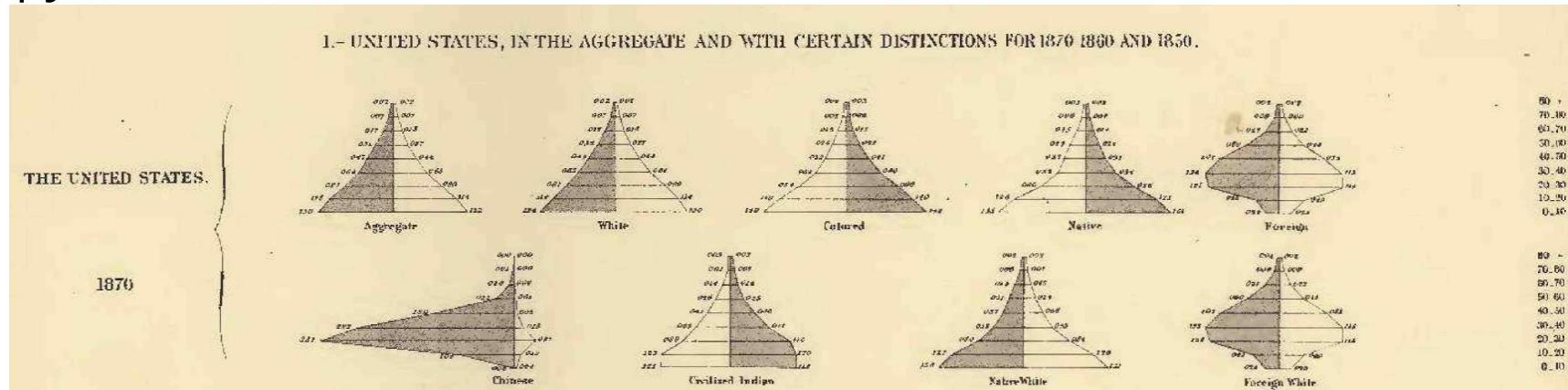
Introduction

- Statistical Atlas published for the 9th, 10th, and 11th census (1870, 1880, 1890)
- Charts created by hand and reproduced using lithography
- Goal: Show the composition of the country
 - Maps



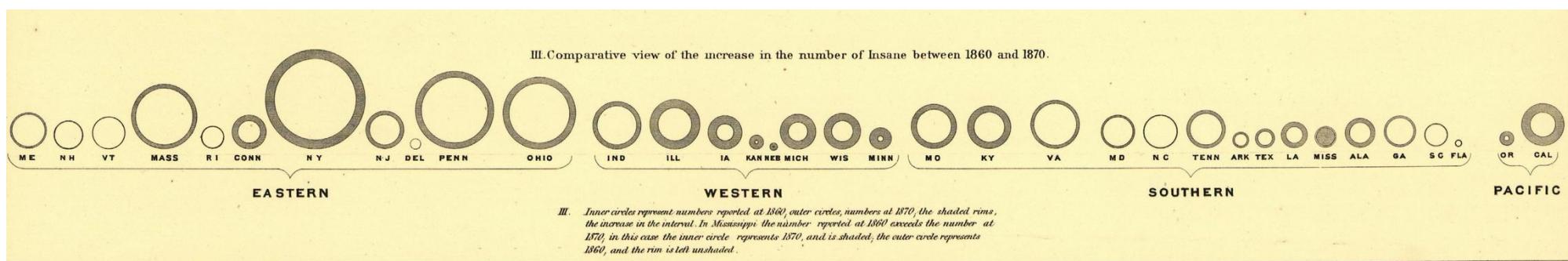
Introduction

- Statistical Atlas published for the 9th, 10th, and 11th census (1870, 1880, 1890)
- Charts created by hand and reproduced using lithography
- Goal: Show the composition of the country
 - Abstract Charts:
 - age pyramids



Introduction

- Statistical Atlas published for the 9th, 10th, and 11th census (1870, 1880, 1890)
- Charts created by hand and reproduced using lithography
- Goal: Show the composition of the country
 - Abstract Charts:
 - age pyramids
 - ring charts



Introduction

- Statistical Atlas published for the 9th, 10th, and 11th census (1870, 1880, 1890)
- Charts created by hand and reproduced using lithography
- Goal: Show the composition of the country
 - Abstract Charts:
 - age pyramids
 - ring charts
 - framed mosaic and spine plots

**CHART
SHOWING THE RATIO OF
CHURCH ACCOMMODATION**

TO THE TOTAL POPULATION OVER 10 YEARS OF AGE

WITH THE PROPORTION OF SUCH CHURCH ACCOMMODATION FURNISHED BY EACH OF
THE LARGEST FOUR DENOMINATIONS WITHIN EACH STATE AND BY EACH OF
THE LARGEST EIGHT DENOMINATIONS WITHIN THE UNITED STATES.

Compiled from the Social Statistics of the Ninth Census 1870

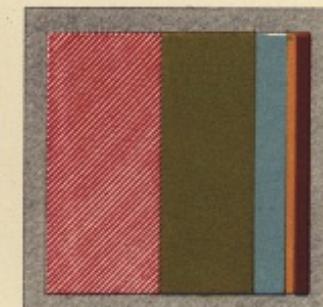
BY
FRANCIS A. WALKER.

*Note: The interior squares represent the proportion of the population which is provided for by the aggregate
sittings in the churches of all denominations. The shaded interval between the inner and outer squares
represents the population for which no church accommodation is provided. Where the aggregate church accommodation
equals or exceeds the population over 10 years of age the shaded interval disappears.*

Methodist	Red
Baptist	Dark Green
Presbyterian	Light Green
Roman Catholic	Blue
Congregational	Yellow
Episcopal	Orange
Lutheran	Dark Red
Christian	Light Red
Dutch Reformed	Dark Green
Universalist	Light Green
Mormon	Blue
All other Denomi- nations	Dark Red



THE UNITED STATES



ALABAMA



ARKANSAS



CALIFORNIA



CONNECTICUT



DELAWARE



FLORIDA



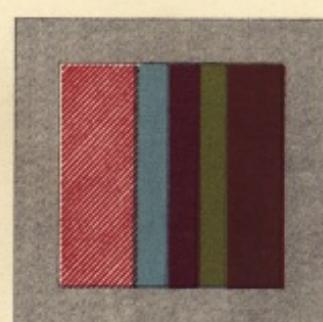
GEORGIA



ILLINOIS



INDIANA



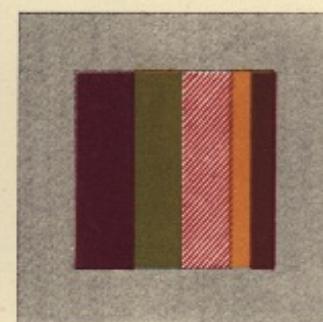
IOWA



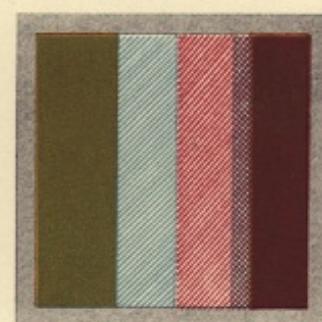
KANSAS



KENTUCKY



LOUISIANA



MAINE



MARYLAND



MASSACHUSETTS





CHART
SHOWING THE RATIO OF
CHURCH ACCOMMODATION
TO THE TOTAL POPULATION OVER 10 YEARS OF AGE
WITH THE PROPORTION OF SUCH CHURCH ACCOMMODATION FURNISHED BY EACH OF
THE LARGEST FOUR DENOMINATIONS WITHIN EACH STATE AND BY EACH OF
THE LARGEST EIGHT DENOMINATIONS WITHIN THE UNITED STATES.

Compiled from the Social Statistics of the Ninth Census 1870
BY
FRANCIS A. WALKER.

Note: The interior squares represent the proportion of the population which is provided for by the aggregate settings in the churches of all denominations. The shaded interval between the inner and outer squares represents the population for which no church accommodation is provided. Where the aggregate church accommodation equals or exceeds the population over 10 years of age the shaded interval disappears.

Methodist
Baptist
Presbyterian
Roman Catholic
Congregational
Episcopal



Lutheran
Christian
Dutch Reformed
Universalist
Mormon
*All other Denomi-
nations*



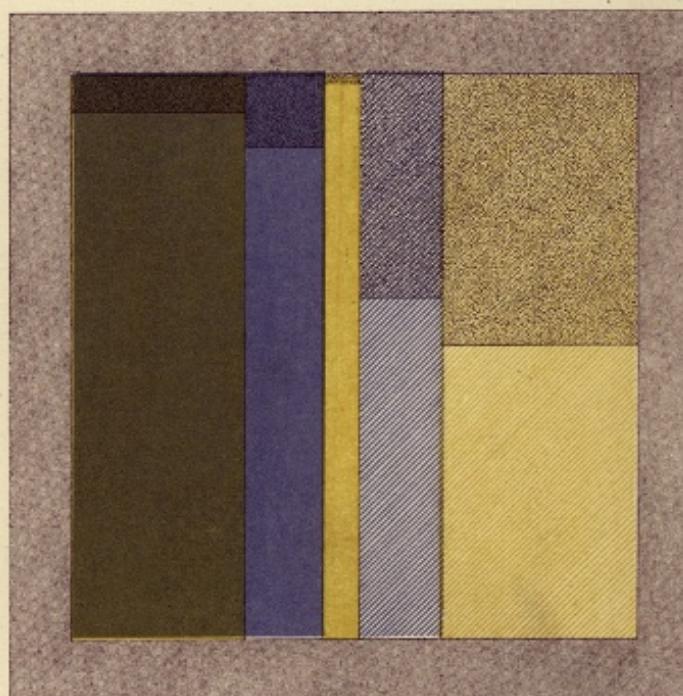
CHART
 SHOWING FOR THE UNITED STATES AND FOR EACH STATE,
 WITH DISTINCTION OF SEX, THE RATIO
 BETWEEN THE TOTAL POPULATION OVER 10 YEARS OF AGE
 AND THE NUMBER OF PERSONS REPORTED AS ENGAGED IN EACH PRINCIPAL CLASS OF

GAINFUL OCCUPATIONS
AND ALSO AS
ATTENDING SCHOOL.

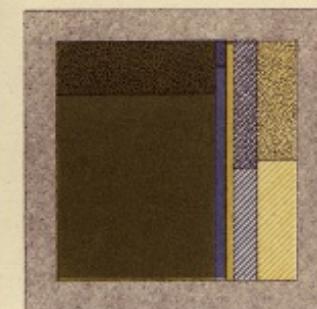
Compiled from the Returns of Population at the Ninth Census 1870
 BY
 FRANCIS A. WALKER.

Note. The interior squares represent the proportion of the population which is accounted for as engaged in gainful occupations or as attending school. The shaded intervals between the inner and outer squares represent the proportion of the population not so accounted for.

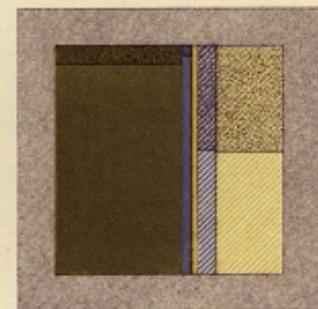
Males Females
 Engaged in Agriculture.
 .. Manufactures & Mining
 .. Trade & Transportation
 .. Prof. & Profess. Services
 Attending School.

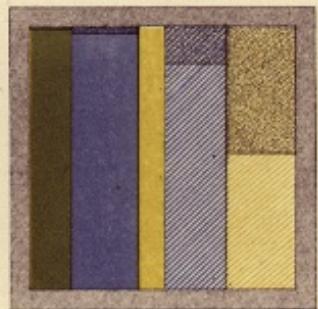
THE UNITED STATES



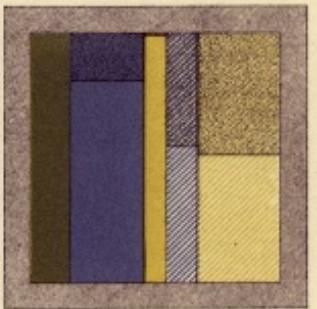
ALABAMA



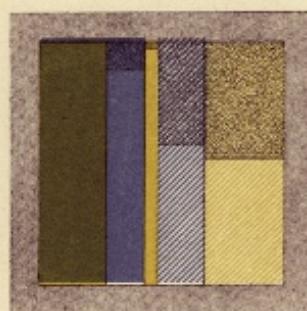
ARKANSAS



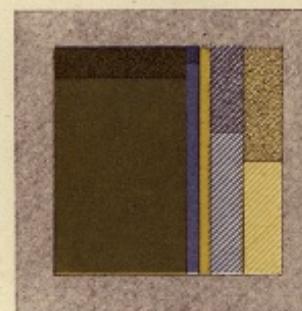
CALIFORNIA



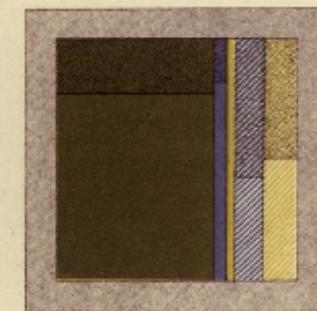
CONNECTICUT



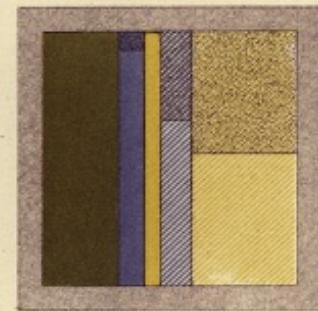
DELAWARE



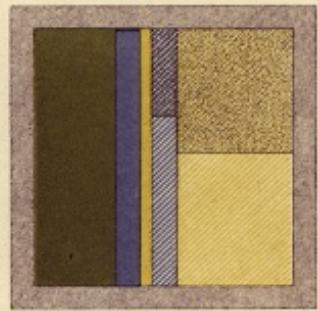
FLORIDA



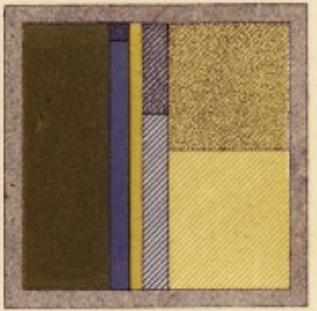
GEORGIA



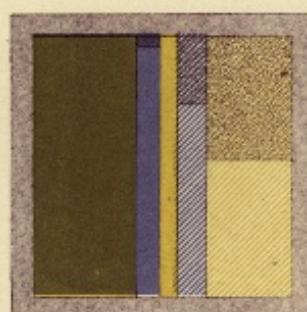
ILLINOIS



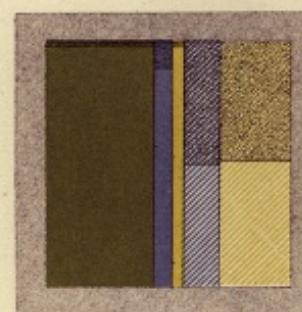
INDIANA



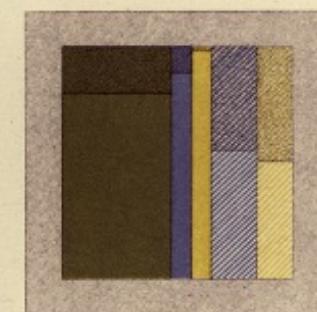
IOWA



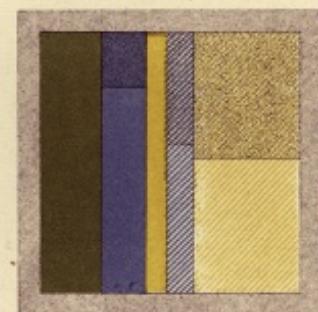
KANSAS



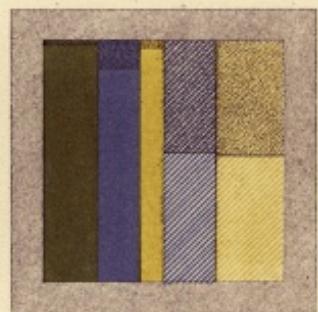
KENTUCKY



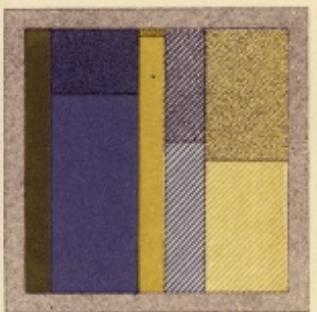
LOUISIANA



MAINE

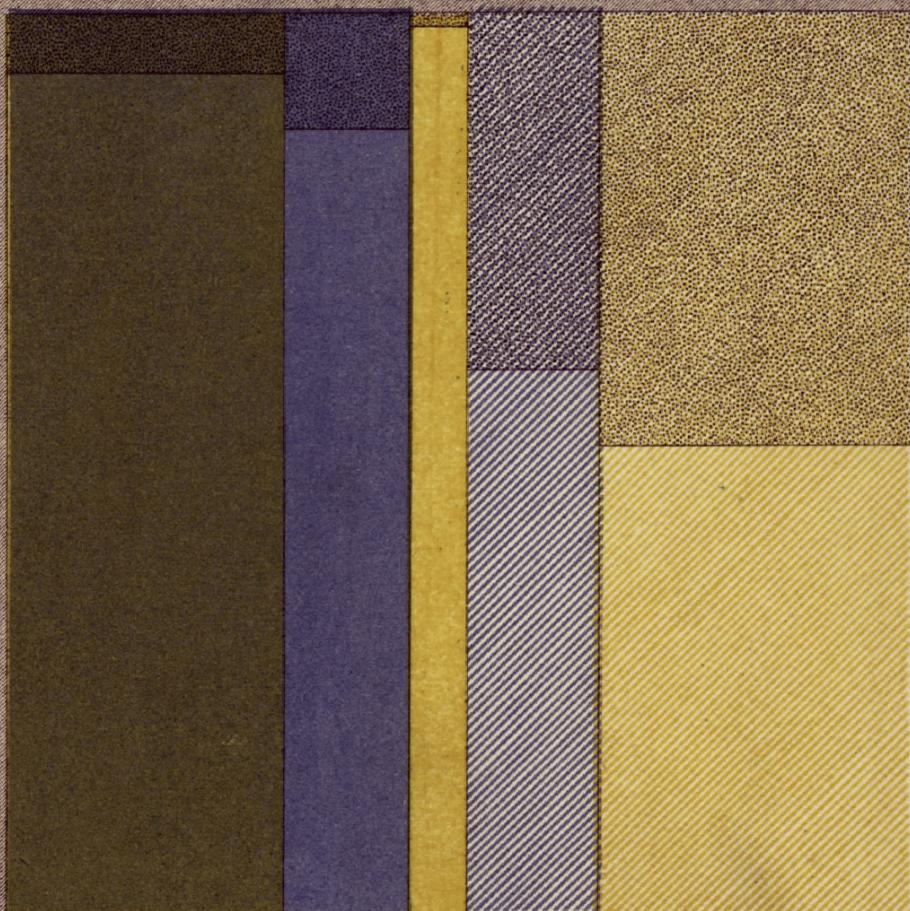


MARYLAND



MASSACHUSETTS





THE UNITED STATES

**CHART
SHOWING FOR THE UNITED STATES AND FOR EACH STATE,
WITH DISTINCTION OF SEX, THE RATIO
BETWEEN THE TOTAL POPULATION OVER 10 YEARS OF AGE
AND THE NUMBER OF PERSONS REPORTED AS ENGAGED IN EACH PRINCIPAL CLA
GAINFUL OCCUPATIONS
AND ALSO AS
ATTENDING SCHOOL.**

Compiled from the Returns of Population at the Ninth Census 1870
BY
FRANCIS A. WALKER.

Note. The interior squares represent the proportion of the population which is accounted for as engaged in gainful occupations or as attending school. The shaded intervals between the inner and outer squares represent the proportion of the population not so accounted for.

Male, Female.

Engaged in Agriculture.

,, ,*Manufactures & Mining.*

,, ,*Trade & Transportation.*

,, ,*Pers. & Profess. Services.*

Attending School.



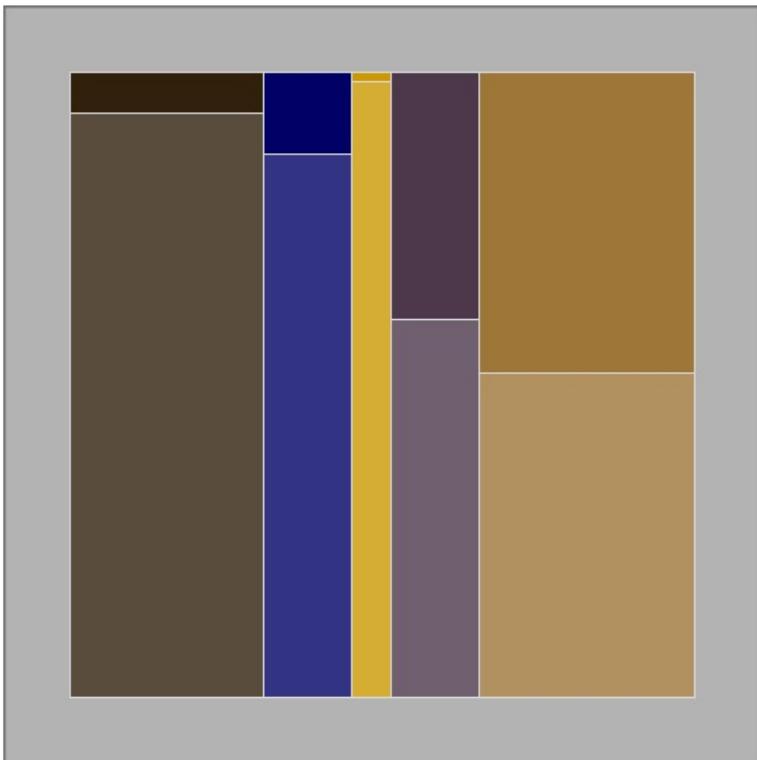
Statistical Archaeology

Reproduce the plots with modern methods

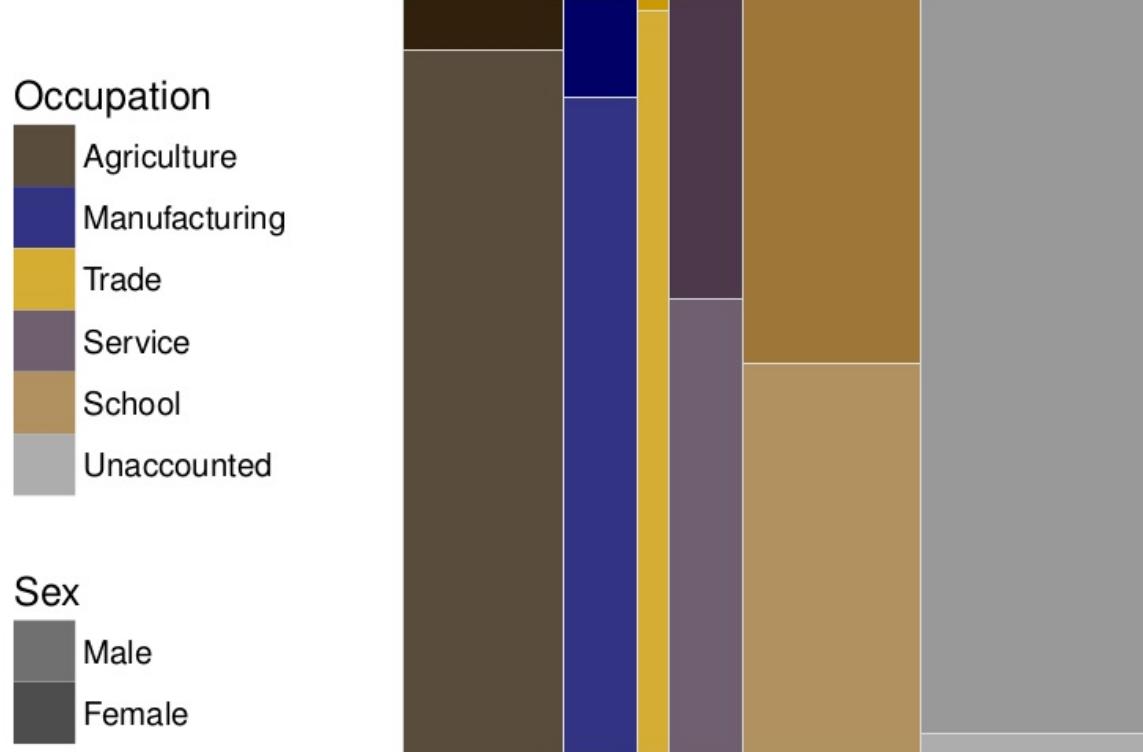
Data Sources:

- National Historical Geographic Information System
- Original 1870 Census tables
- Unaccommodated population estimates using 1% microsample data from Integrated Public Use Microdata Series (IPUMS).
- Pixel measurements (by hand) of the area of each chart (for comparison)

Occupation



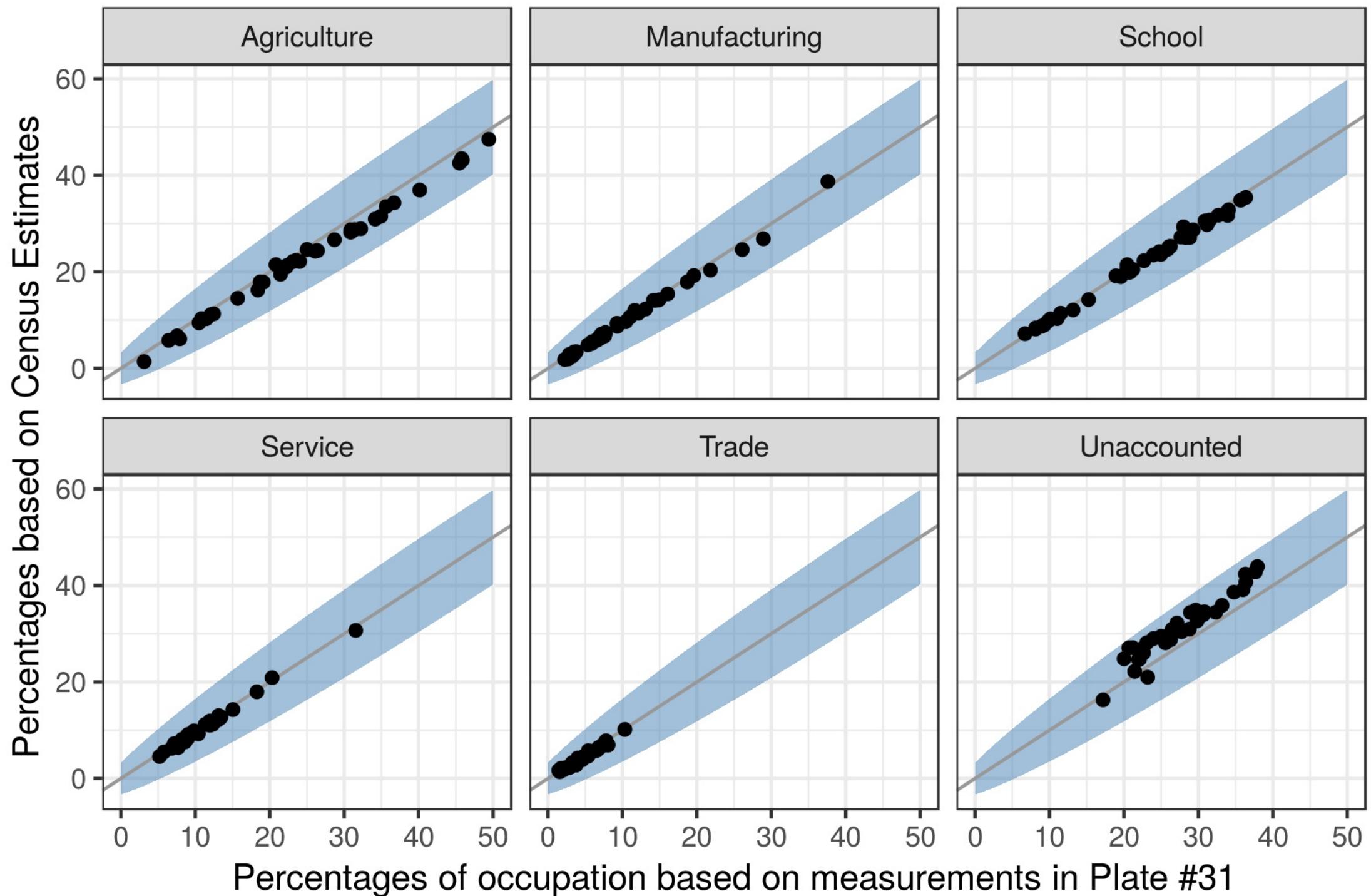
The United States



The United States

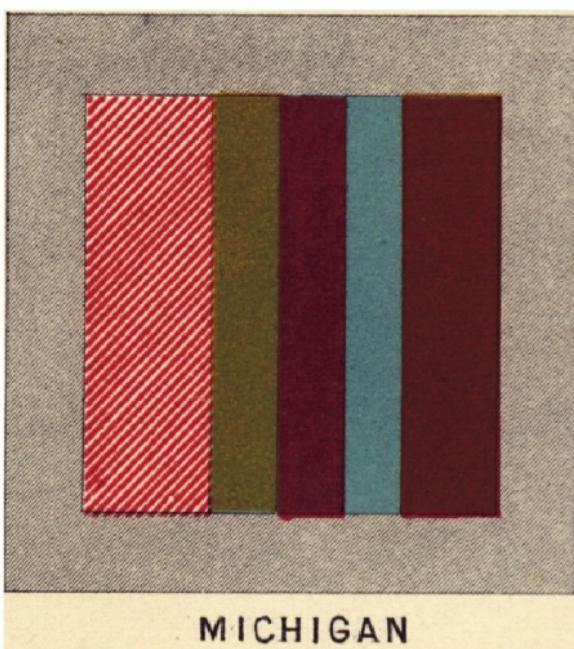
- Frame cuts the uncounted population into quarters visually
- No segmentation of the uncounted population - 97% of uncounted population is female

Occupation

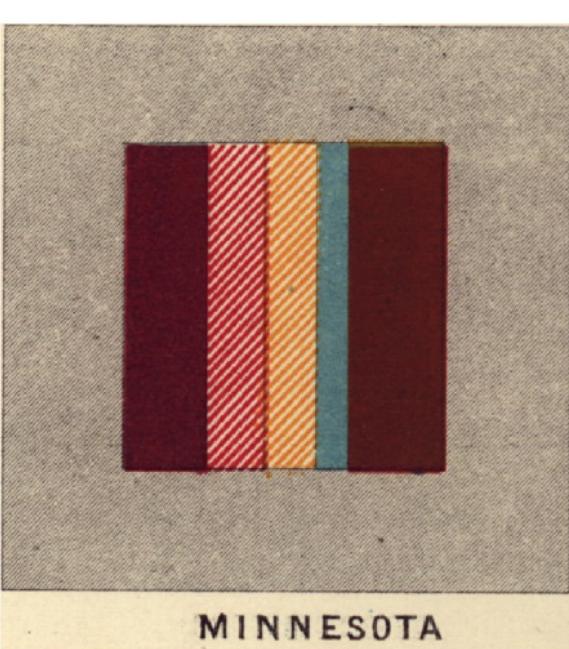


Church Accommodation

Church Accommodation



MICHIGAN

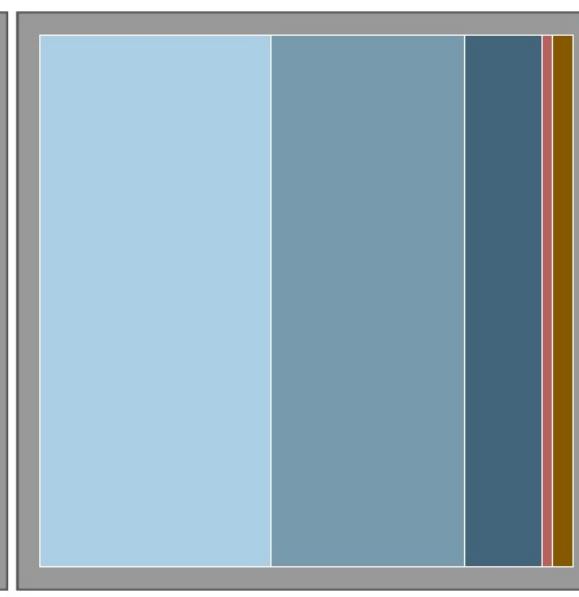
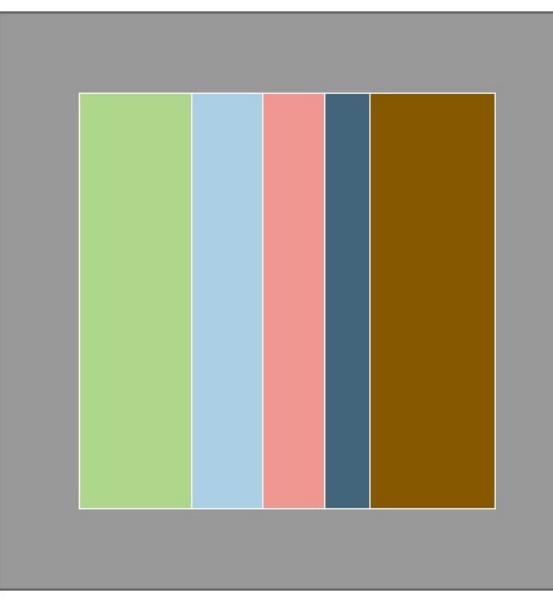
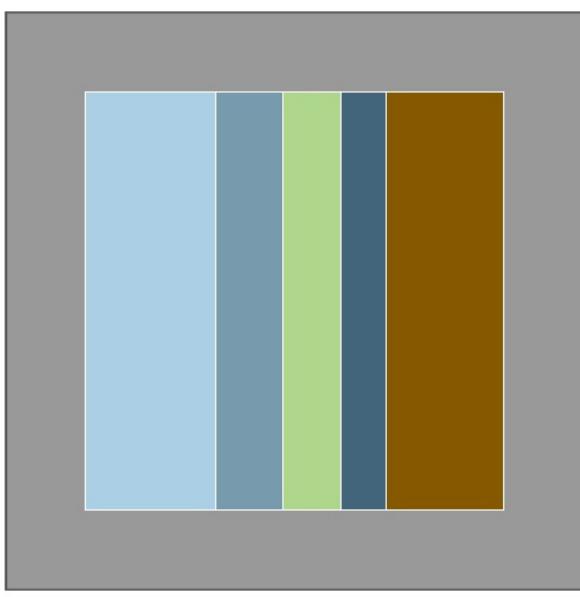
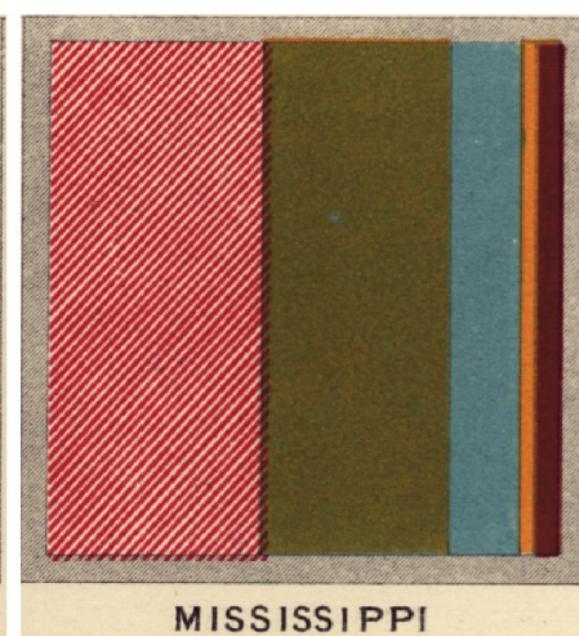
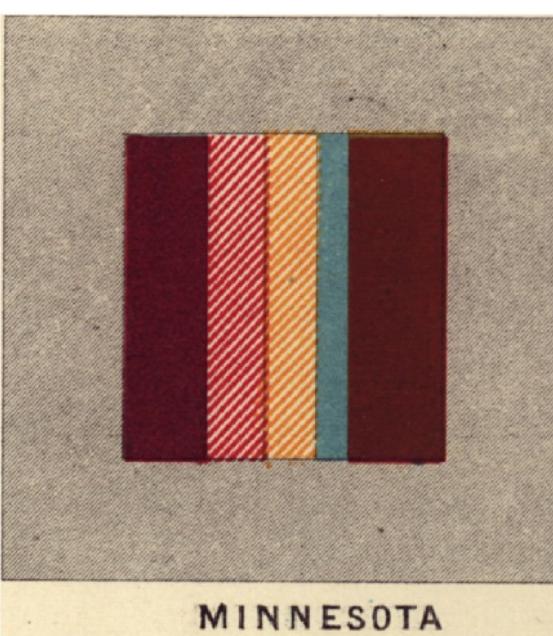
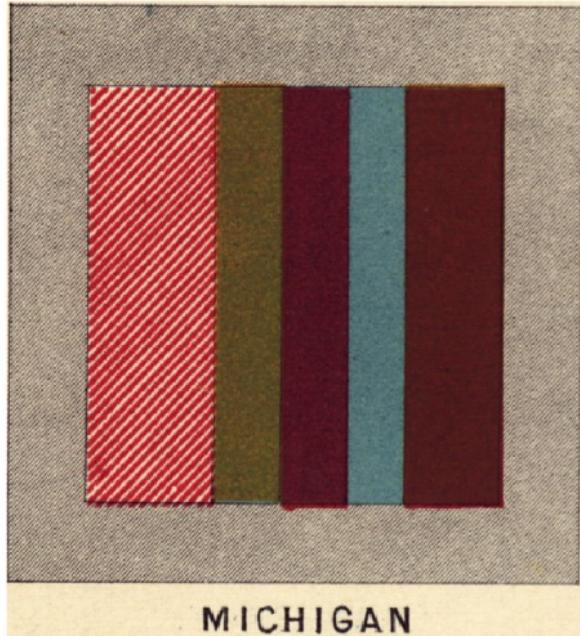


MINNESOTA

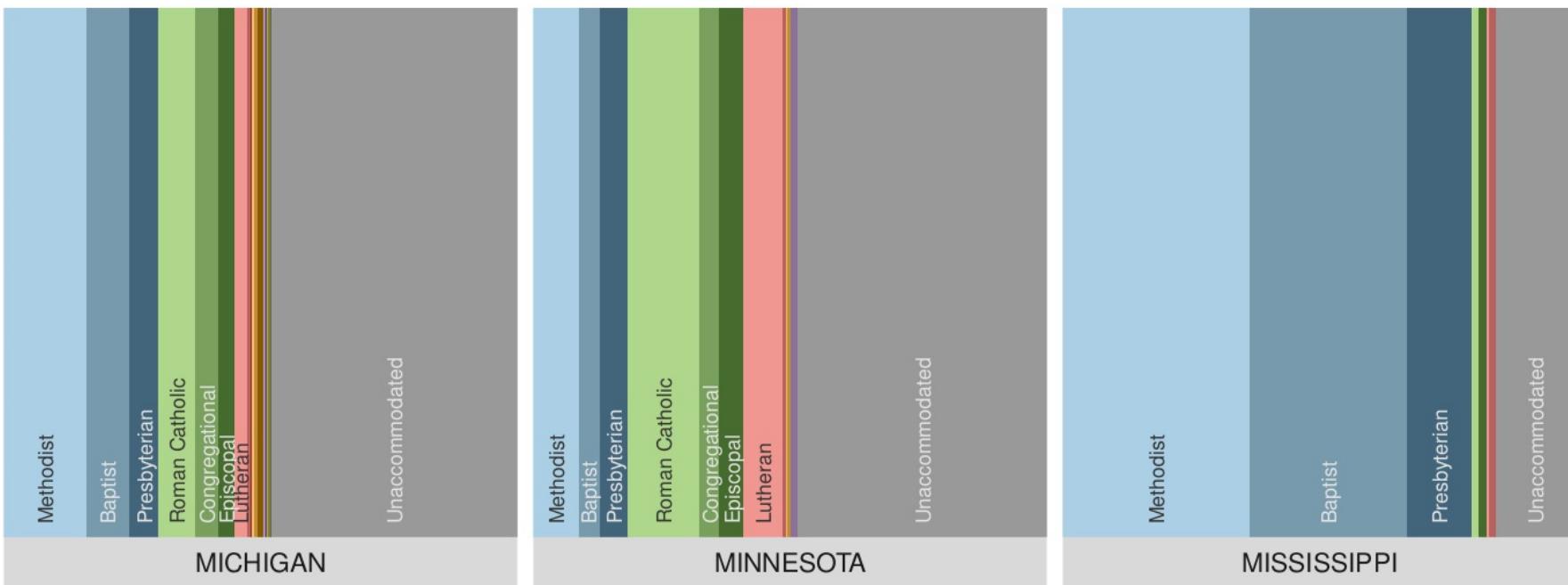
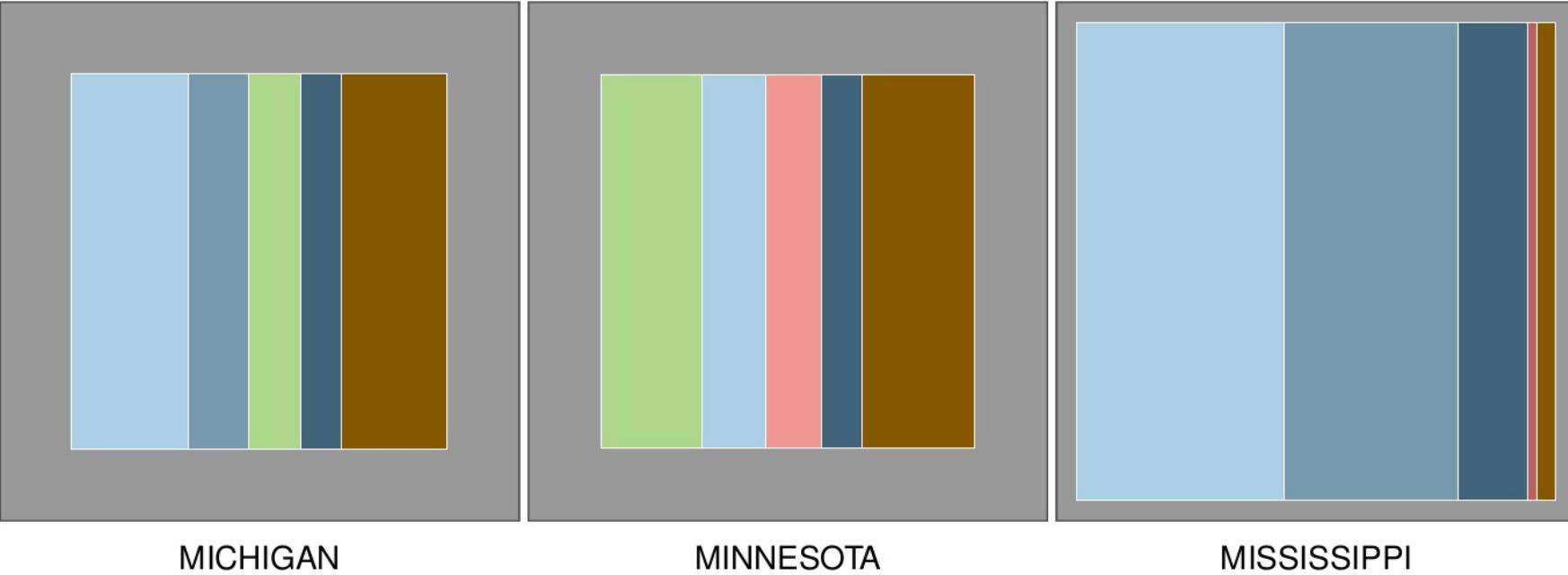


MISSISSIPPI

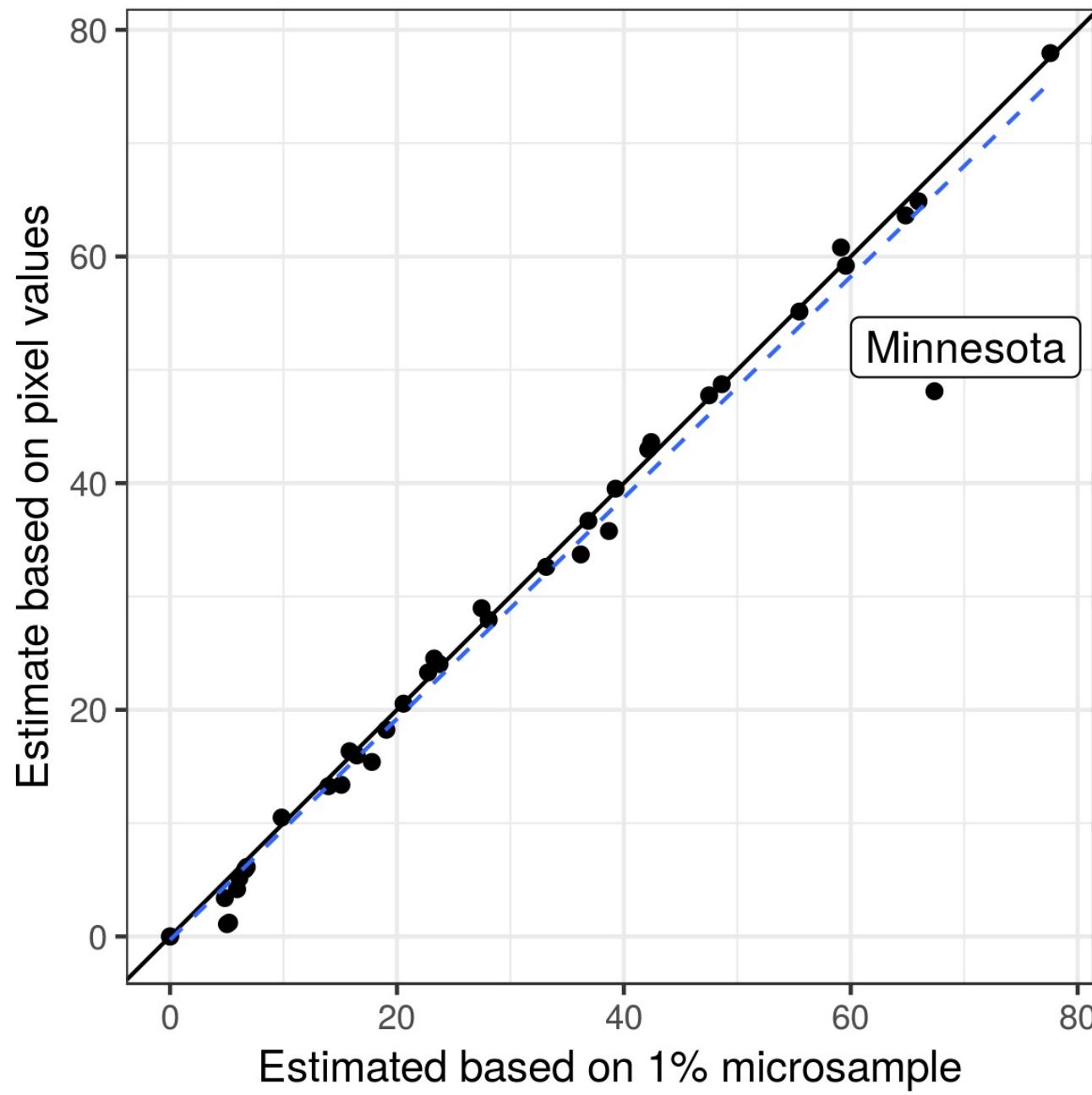
Church Accommodation



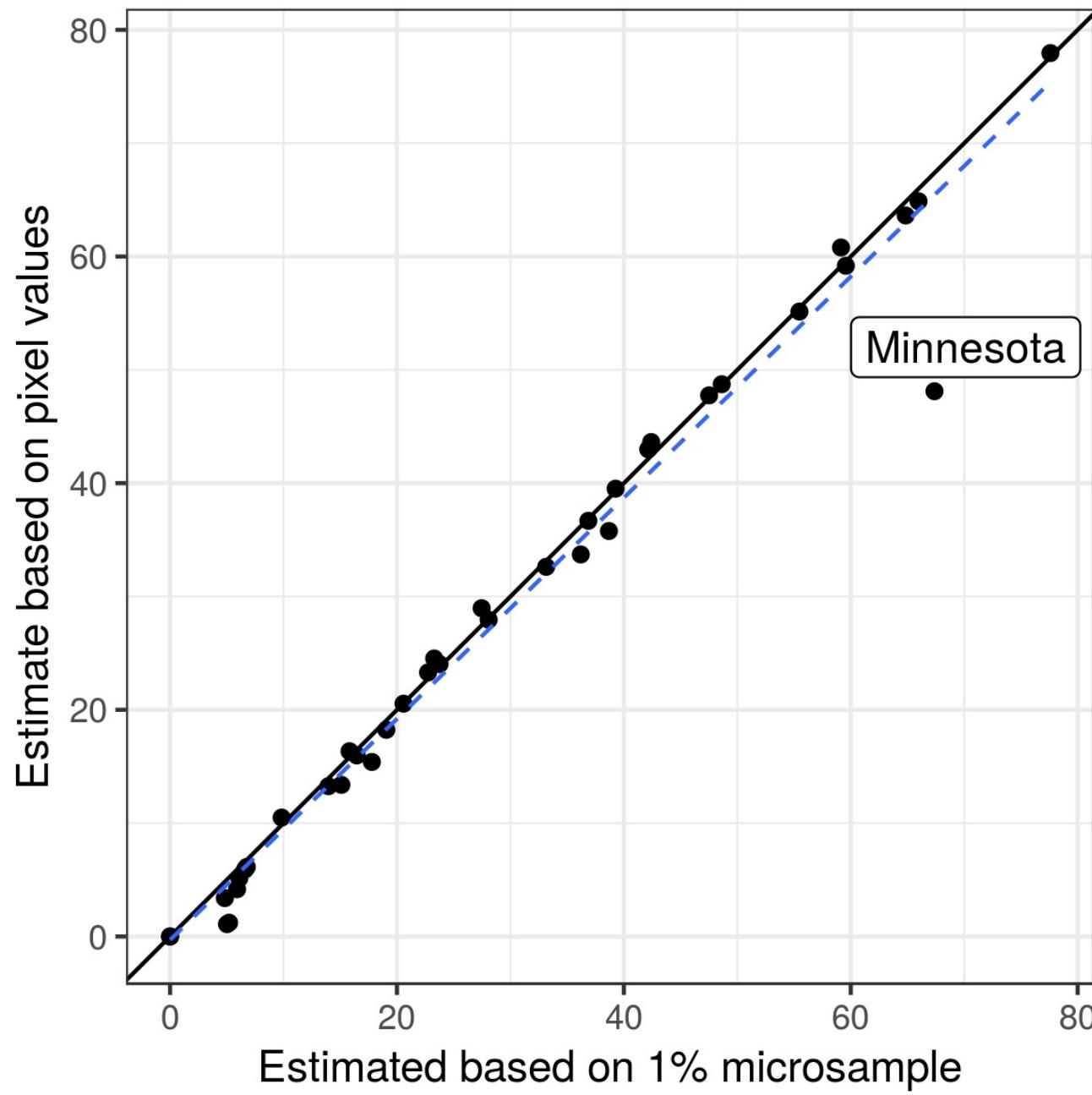
Church Accommodation



Church Accommodation



Church Accommodation



Whoops. The lithographer screwed up.

Perception of Framed Plots

- Frame cuts the unaccommodated population into quarters - hard to estimate the size

Perception of Framed Plots

- Frame cuts the unaccommodated population into quarters - hard to estimate the size
- Difficult to compare between states, as there's no aligned axis due to the frame

Perception of Framed Plots

- Frame cuts the unaccommodated population into quarters - hard to estimate the size
- Difficult to compare between states, as there's no aligned axis due to the frame
- Estimation requires a two-stage process:
 1. Area of the main plot that is accommodated
 2. Area of the center region devoted to a specific category

Perception of Framed Plots

- Frame cuts the unaccommodated population into quarters - hard to estimate the size
- Difficult to compare between states, as there's no aligned axis due to the frame
- Estimation requires a two-stage process:
 1. Area of the main plot that is accommodated
 2. Area of the center region devoted to a specific category
- Primary comparisons are on the x-axis; the frame disrupts the ease of these comparisons

Frames in Practice

- 98 Participants (32 from reddit.com/r/samplesize, 63 from Amazon Mechanical Turk)

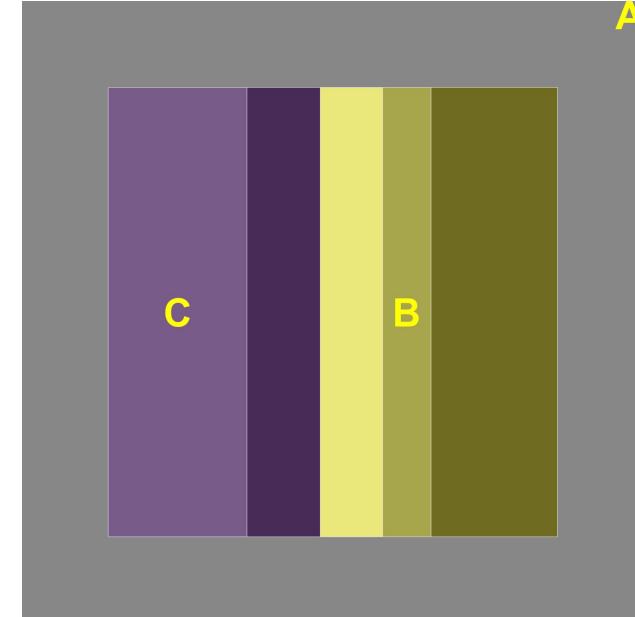
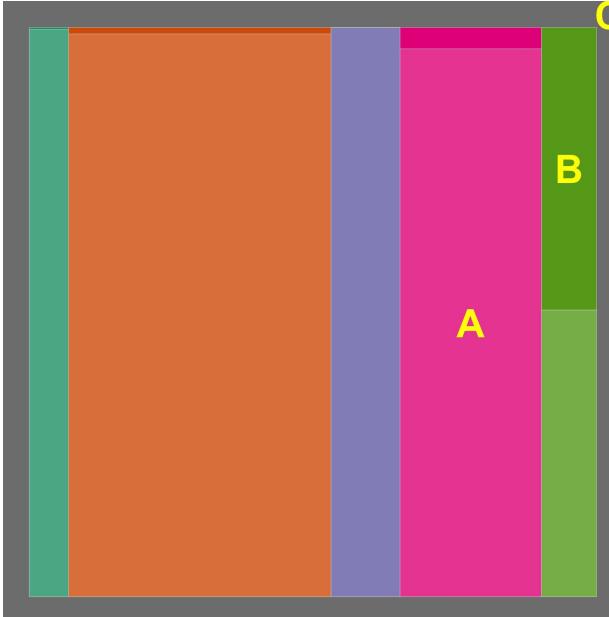
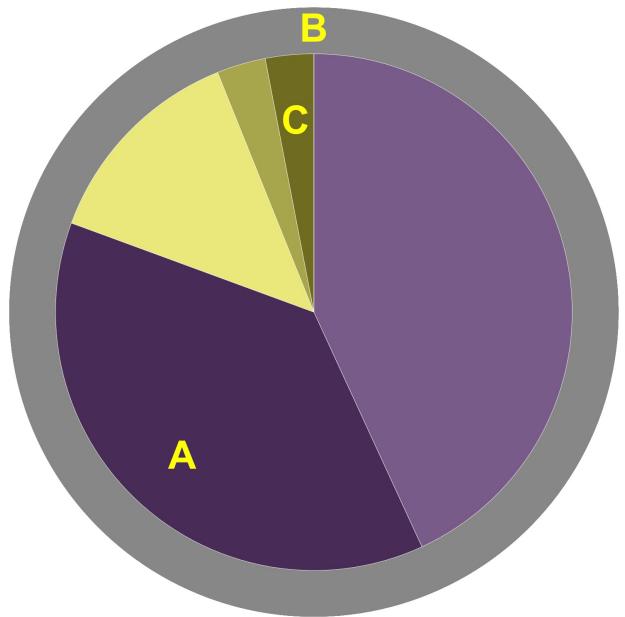
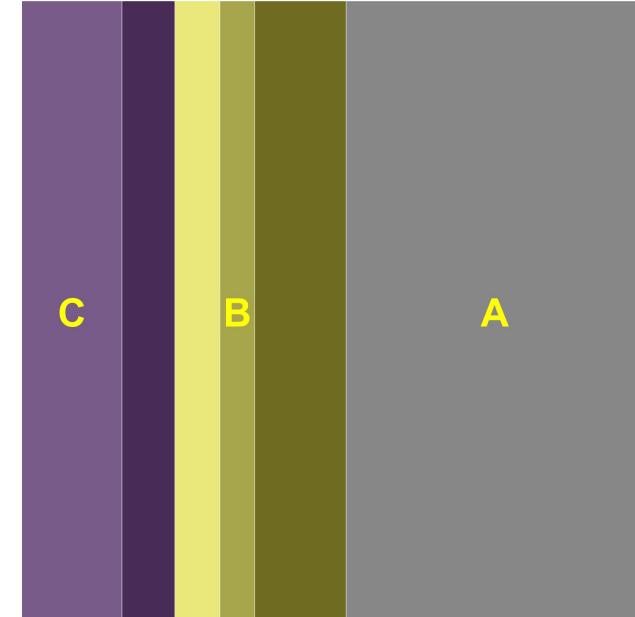
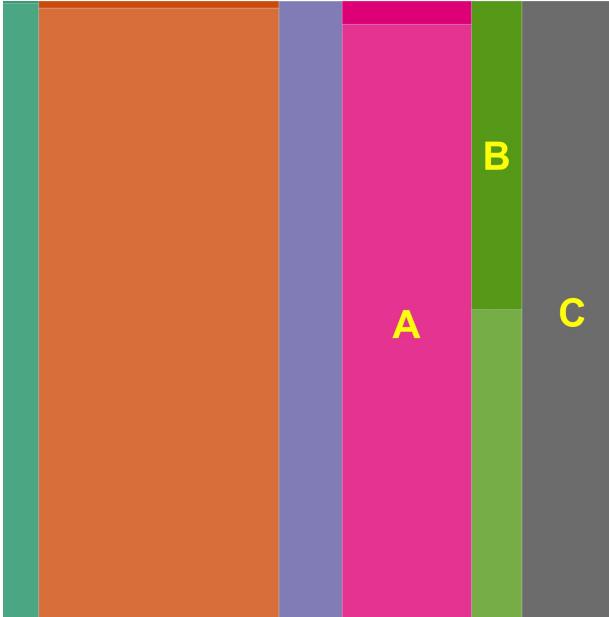
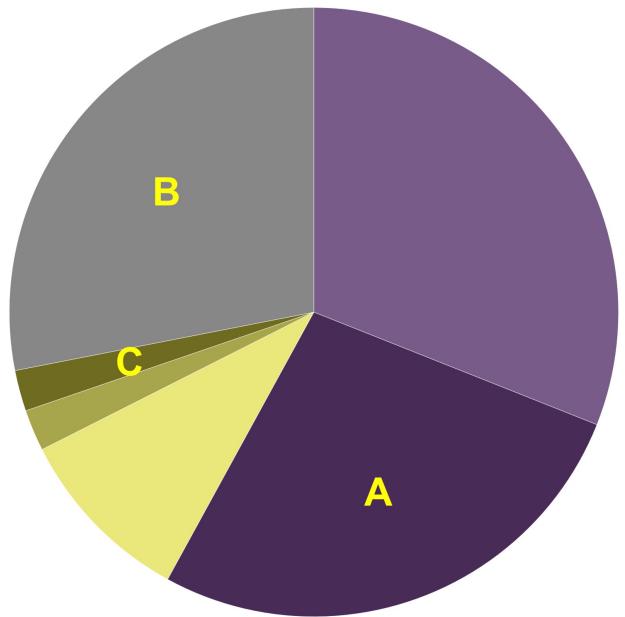
Frames in Practice

- 98 Participants (32 from reddit.com/r/samplesize, 63 from Amazon Mechanical Turk)
- Each participant evaluated 18 plots
 - Region size: small, medium, large
 - Pie, Mosaic, Spine plots
 - Framed or Unframed charts

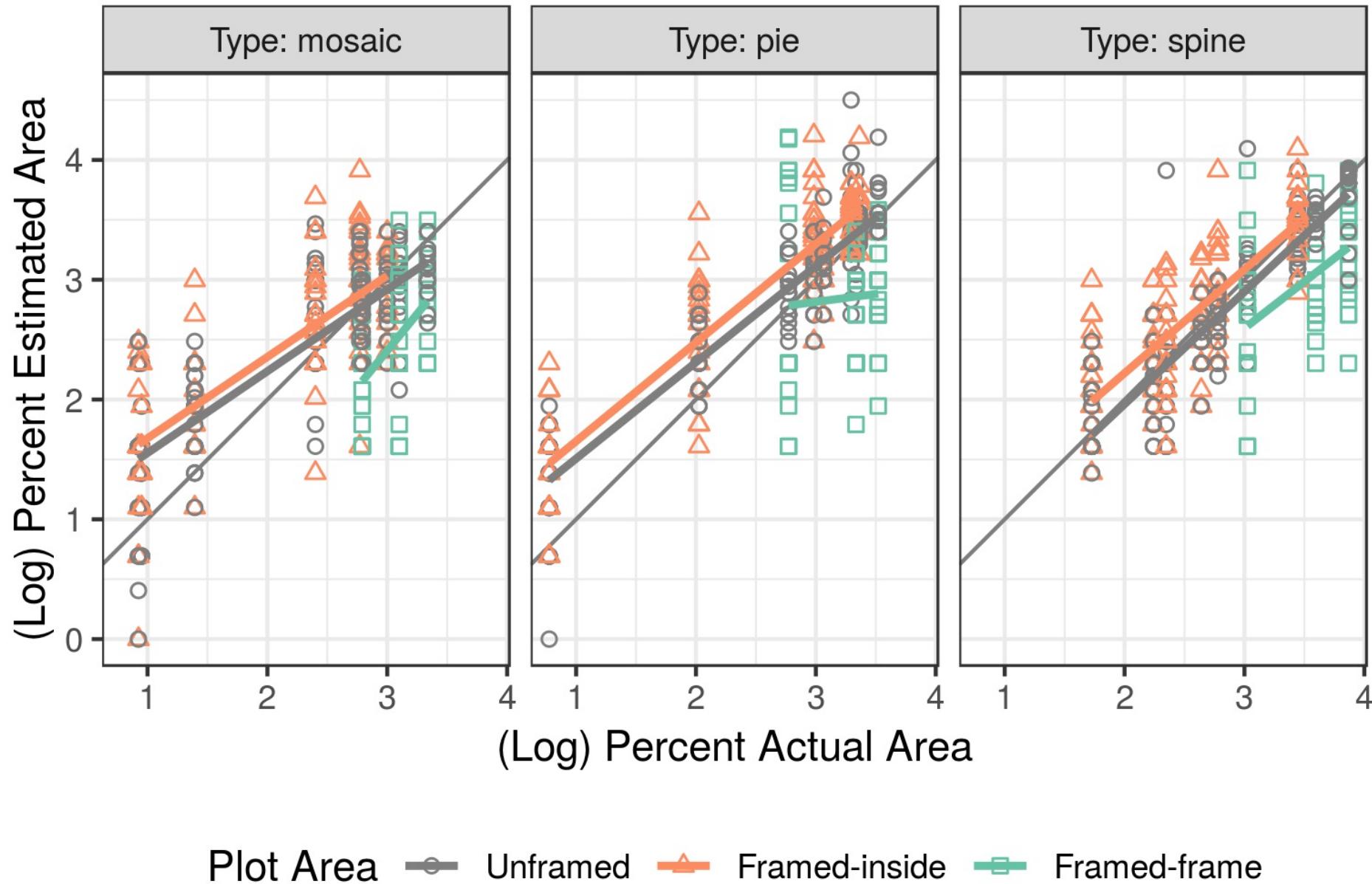
Frames in Practice

- 98 Participants (32 from reddit.com/r/samplesize, 63 from Amazon Mechanical Turk)
- Each participant evaluated 18 plots
 - Region size: small, medium, large
 - Pie, Mosaic, Spine plots
 - Framed or Unframed charts
- 25 seconds per plot

Frames in Practice



Results



Conclusions

- Estimation of the framed portion of plots is significantly worse than the same area in an unframed plot

Conclusions

- Estimation of the framed portion of plots is significantly worse than the same area in an unframed plot
- Framed plots are not great for accurately communicating the data
 - They don't appear in the 1880 or 1890 census
 - Pie charts largely replaced mosaic and spine plots

Conclusions

- Estimation of the framed portion of plots is significantly worse than the same area in an unframed plot
- Framed plots are not great for accurately communicating the data
 - They don't appear in the 1880 or 1890 census
 - Pie charts largely replaced mosaic and spine plots
- The Census Bureau and associated institutions did a great job of preserving the data necessary to reproduce the results and identify flaws in the analysis process

Conclusions

150 years later, we can use completely different technology to get the same charts

Conclusions

150 years later, we can use completely different technology to get the same charts

How's that for reproducibility?