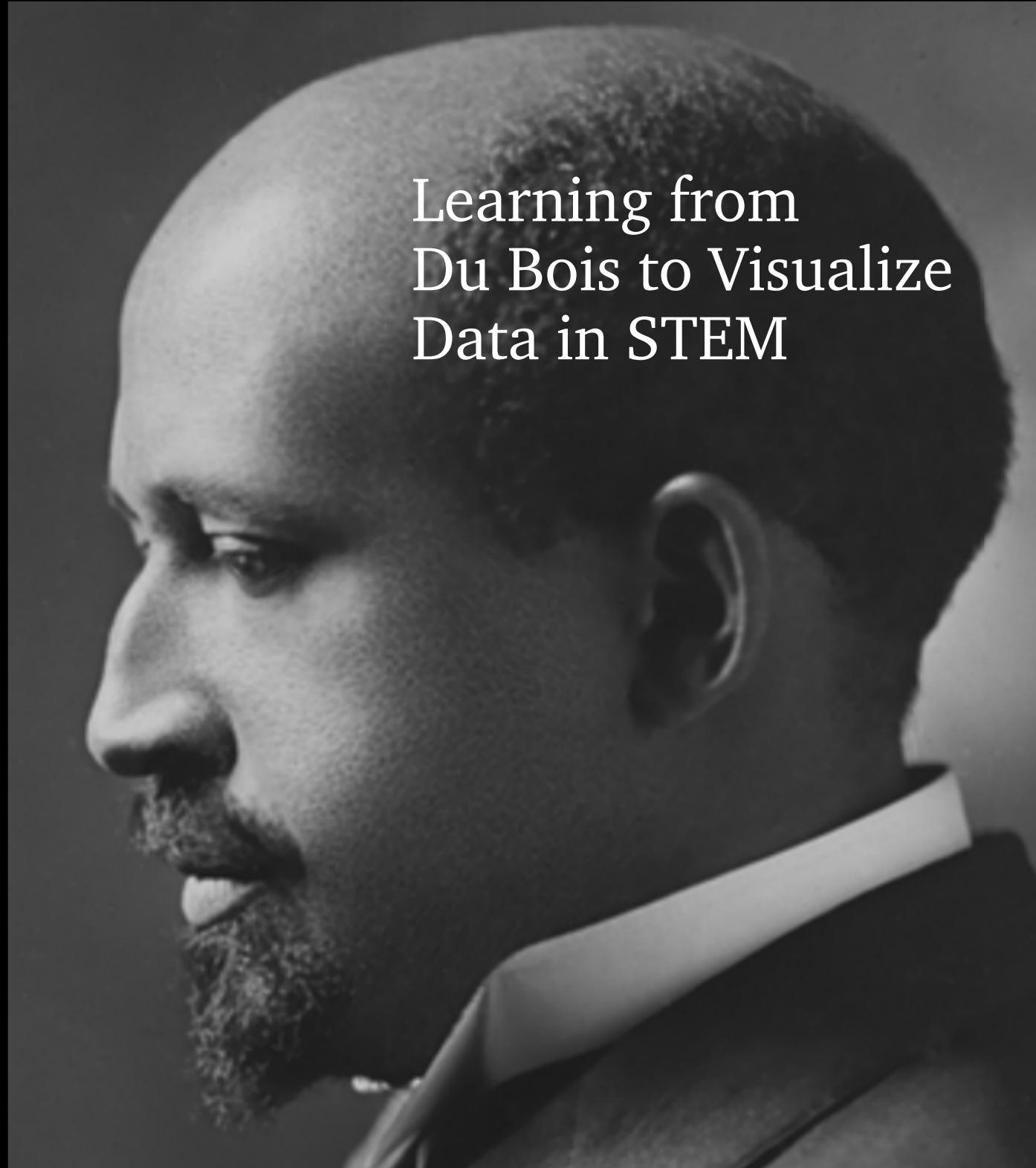




Learning from Du Bois to Visualize Data in STEM





Introduction

The goal of this module is for students to learn how to understand and create scientific data visualizations.

The module explores visualizations created by W.E.B. Du Bois in 1900 and the research methods they employed.

Through deconstruction and re-creation—we will see how Du Bois used visualization to answer research questions regarding racial equality that disproved biased race theories.

Data visualizations remain important in STEM education and careers, and for society as a whole.



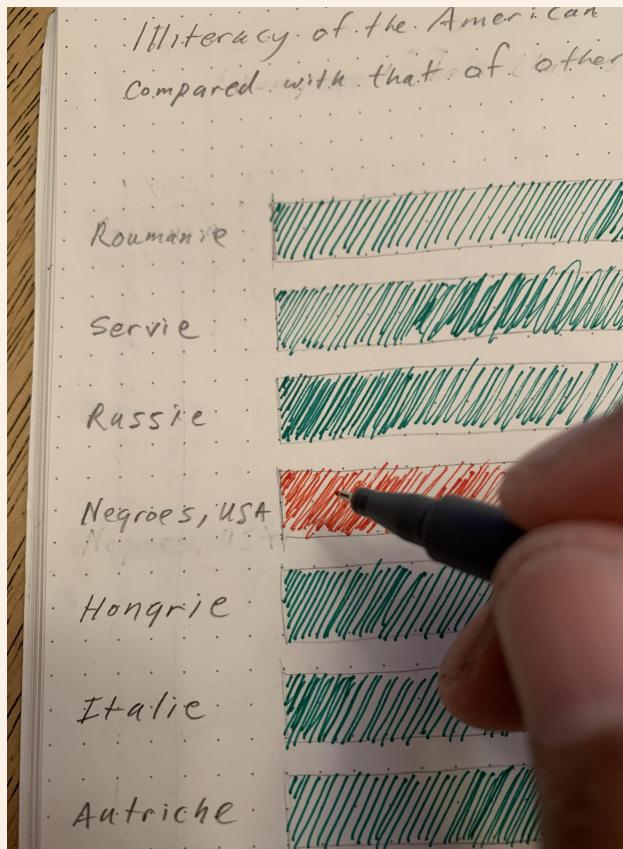
Module Outline

Context



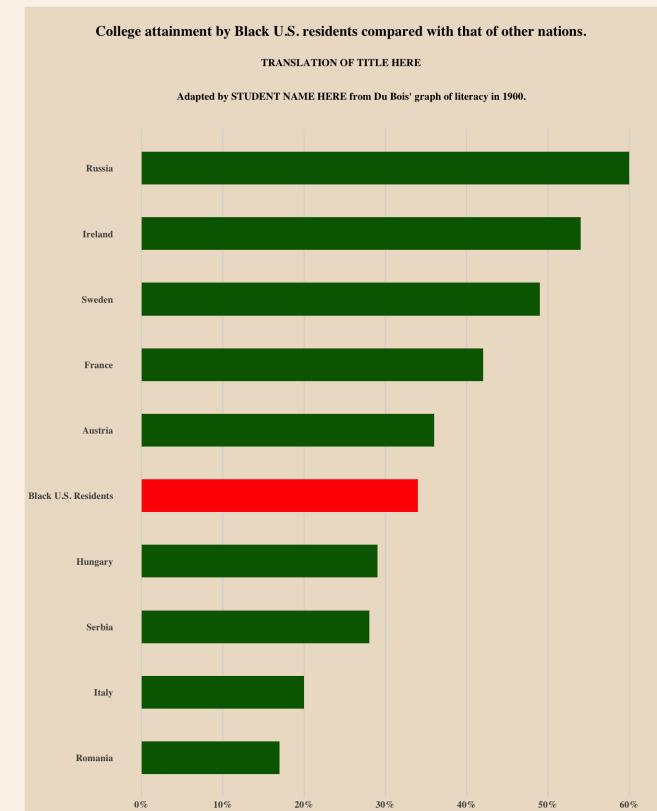
Explain the research questions behind the Du Bois visuals in a historical context.

Explore



An analog exercise to analyze and re-create Du Bois visuals using hand-drawn methods.

Implement



Re-create Du Bois style visuals using digital tools.



Context





Context

The Context section provides background on the conception, motivation and messaging of the visuals. Also discussed is the venue where the visuals were first shown, the Exhibition of the American Negro, within the 1900 Paris Exposition.

To better understand the times when the visuals were created, influential events leading to the Exposition are discussed.



Context: Background



TJ Calloway



Du Bois in Paris



Atlanta U. Students

The motivation for creating the visuals designed and created by Du Bois and his collaborators was to tell, on a world stage, the state and progress of Black Americans 35 years after emancipation from slavery.

Thomas J Calloway, organized the “Exhibition of the American Negro” to be shown during the 1900 Paris Exposition, and commissioned Du Bois, his Fisk University classmate, to create a set charts, graphs, photographs and other artifacts for the world stage.

Du Bois, along with collaborators at Atlanta University spent four months creating the materials the exhibit.



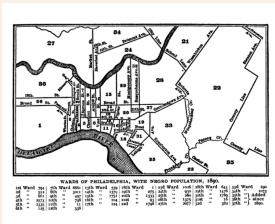
Context: Five years before Paris



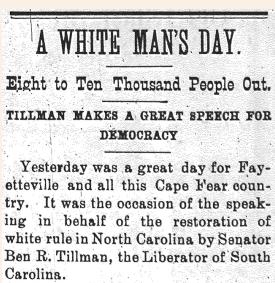
During the summer of 1895, in a Brooklyn park, there was a cotton plantation complete with five hundred Black workers reenacting slavery.



In 1896 the landmark Plessy v. Ferguson Supreme Court decision was handed down, legitimizing racial segregation, leading to generations of Jim Crow laws.



Du Bois created the Philadelphia Negro in 1897 which sought to ascertain something of the geographical distribution of his race, their organizations, and their relation to their million white fellow-citizens.



In 1898 the duly elected people in Wilmington, NC was violently overthrown. The coup occurred after the state's Southern Democrats conspired and led a mob of 2,000 white men to overthrow the legitimately elected local Fusionist government.



During the 1880s and 1890s, instances of lethal mob violence increased steadily, peaking in 1899 when twenty-seven Georgians fell victim to lynch mobs. Included in this number was the horrific lynching of Sam Hose.

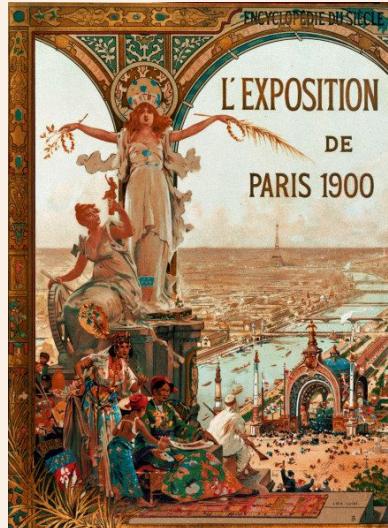
Context: The 1900 Paris Exposition



The Exposition Universelle of 1900 was the venue for Du Bois to tell the story of Black Americans on an international stage.

Fifty-six nations participated in the expo, designed to celebrate the achievements of the 19th century, while looking forward to the developments in the 20th.

The Du Bois visuals were part of the “Exhibition of the American Negro”, seen by over 50 million people. The exhibition also included an African-American bibliography containing 1,400 titles along with 500 photographs.



Exhibition Poster



1900 Paris



Exhibition Venue



Context: Why Visualize Data?

“A picture is worth a thousand words”

It is difficult for people analyze data involving many complex relationships and hundreds (or billions!) of data points.

Visualizing data with graphs can help a scientist answer research questions involving such complex relationships and large amounts of data.

Graphs can also help scientists explain their research findings to other scientists and non-scientists.



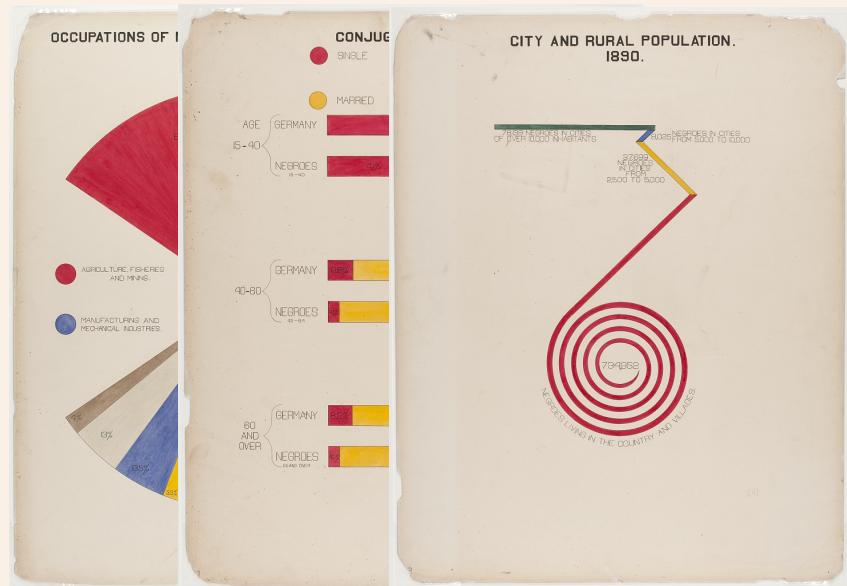
Context: Why Visualize Data?

The Negro has never formed a very large percent of the population of the city, as this diagram shows :



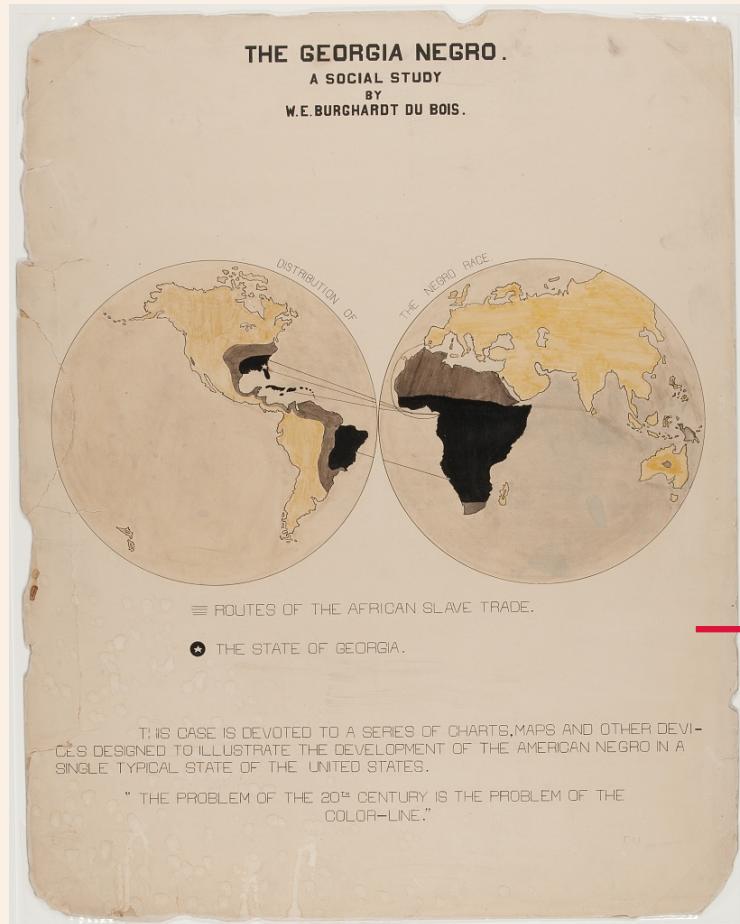
Du Bois used graphs to answer research questions about the social causes of racial inequality in his groundbreaking book *The Philadelphia Negro*.

By combining different types of graphs for different types of data, Du Bois could tell a coherent, scientific story about the social factors supporting and constraining Black Americans pursuit of equality.





Context: The Visuals

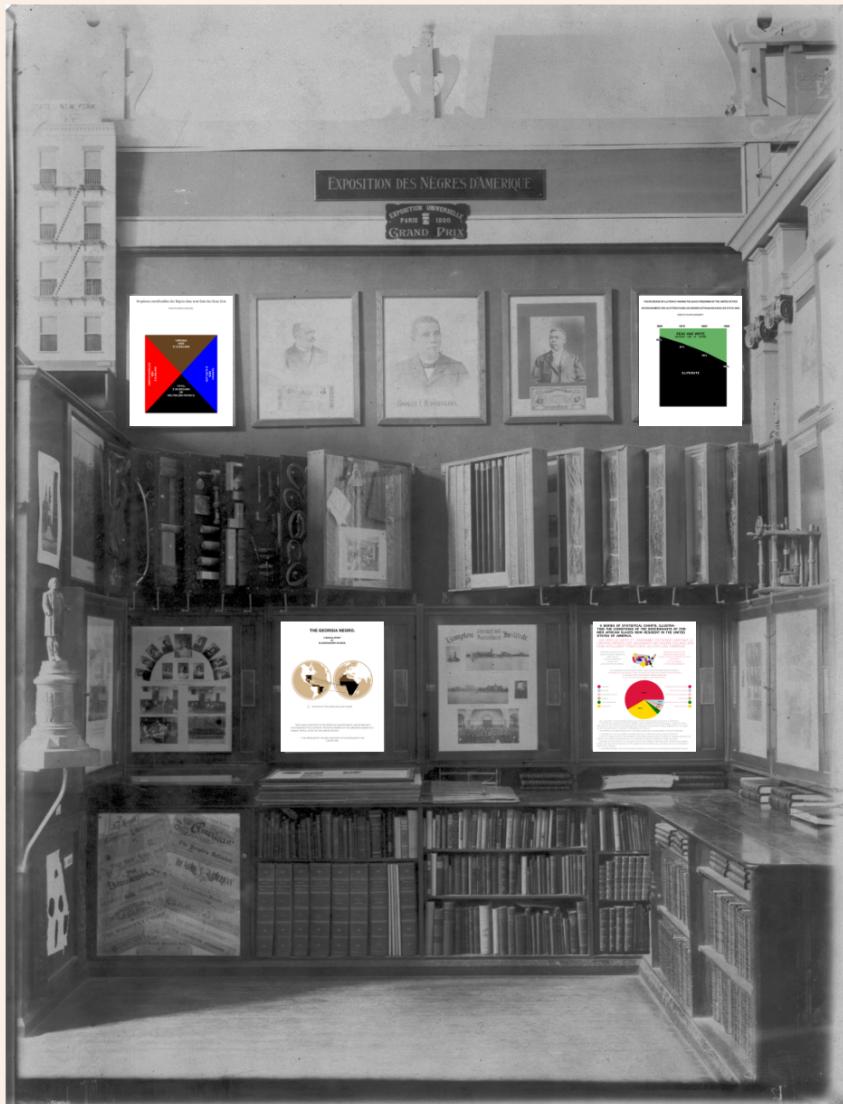


The visuals were hand-drawn using pen, ink and paint on 22x28 in. poster board

The venue shows the visuals as posters to be viewed and other materials may be "flipped through".



Context: The Questions for Discussion



Why do you think Du Bois created a series of graphs and data visualizations of Black life for the exposition?"

Why visualizations instead of a written report?"

What effect did the venue have on the design of the visuals?



Context: References

Paris Exposition of 1900 (Exposition Universelle)

[https://en.wikipedia.org/wiki/Exposition_Universelle_\(1900\)](https://en.wikipedia.org/wiki/Exposition_Universelle_(1900))

Black America, 1895

<https://publicdomainreview.org/essay/black-america-1895>

Plessy v. Ferguson

<https://www.britannica.com/event/Plessy-v-Ferguson-1896>

The Philadelphia Negro

https://www.google.com/books/edition/_/sqwJAAAAIAAJ

Wilmington Insurrection of 1898

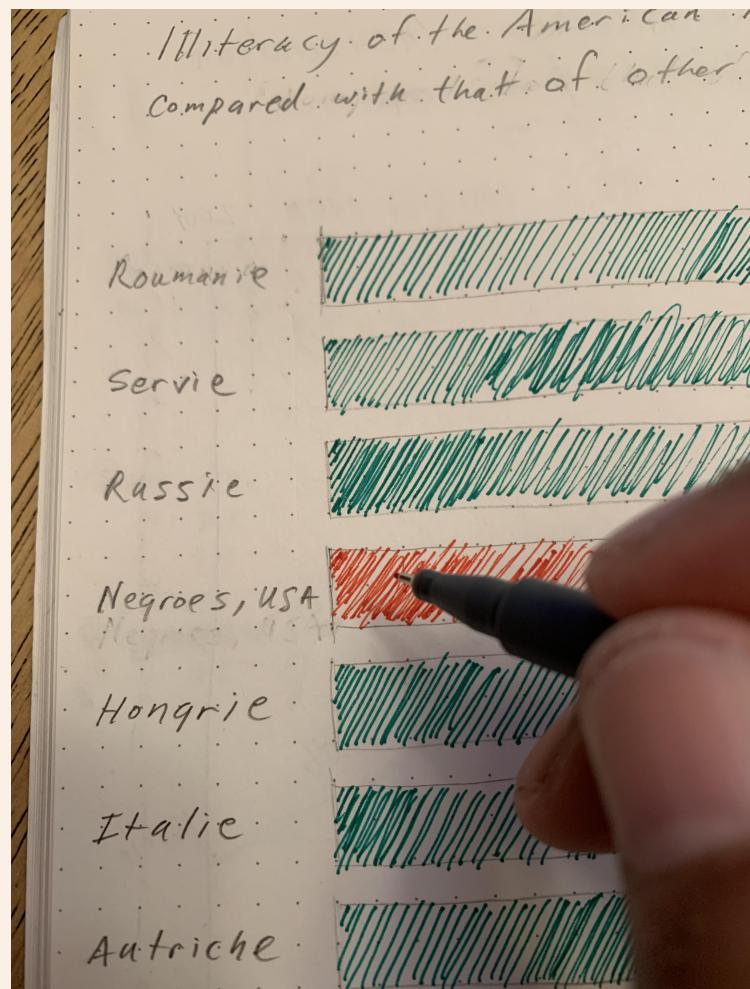
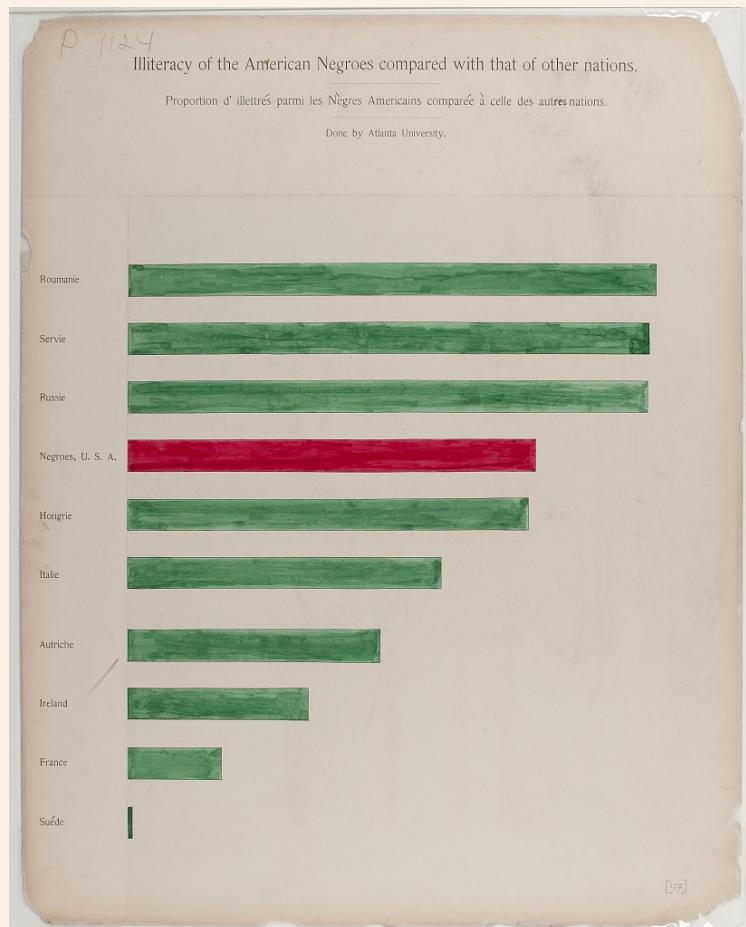
https://en.wikipedia.org/wiki/Wilmington_insurrection_of_1898

The Lynching of Sam Hose

https://en.wikipedia.org/wiki/Lynching_of_Sam_Hose



Explore





Explore

This section is a hands-on session that takes sample Du Bois visualizations, and reconstructs and analyzes various aspects of the visual.

In this exercise you will recreate sample visuals using hand-drawn methods similar to those used by Du Bois and his collaborators (pens, paper, markers, etc.), to immerse yourself in the act of creation, while thinking about the the research question and audiences Du Bois engaged.

Rapid iteration and peer review may be used to gauge the effectiveness of the visual. In the final part of the exercise, you will draw your own visualization with modern data.

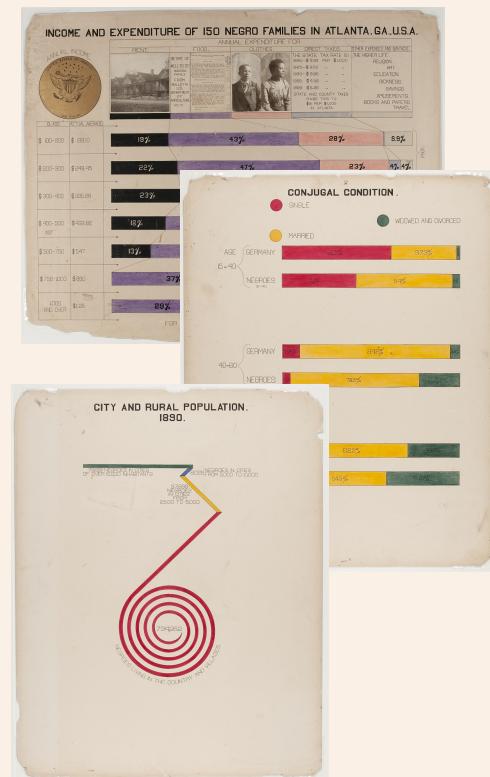
Explore: Outline



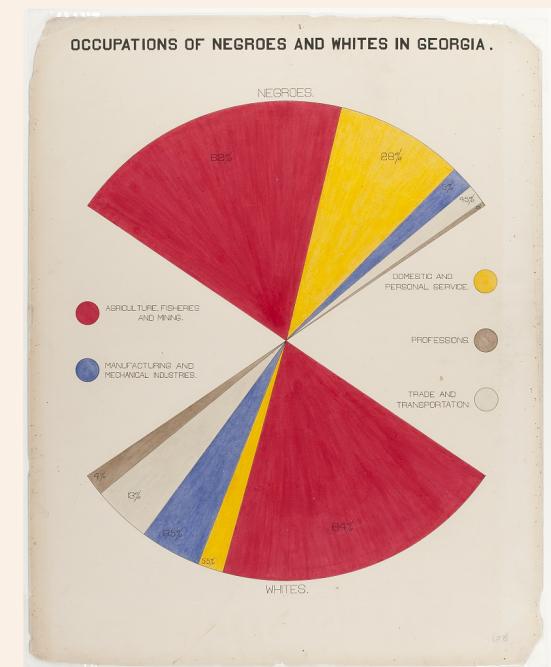
Research Questions and Audience



Analytic methods: Type of graphs and data



Aesthetics, Design and Re-creation





Explore: Research Question and Audience

Data visualization begins with you defining your research question and the audience for your research.

Visualization can help you answer research questions for yourself, but scientific research is social.

We can often gain a better understanding of our findings for ourselves by using visualization to explain our findings to other scientists or non-scientists.



Explore: Research Question and Audience

Research Question

What is the relationship between Black education levels and emancipation from slavery

Audience

International attendees of the Paris Expo, social scientists, and others influenced by Social Darwinism.

Du Bois chose data and visualizations that could answer these questions in a way that these audiences could understand.

Explore: Analytic Methods



We choose data based on our research question: The type of data we choose can vary by:

**Unit of observation
for example:**

- People in social science
- Biological organisms in natural science
- Materials in engineering

**Level of observation
may be:**

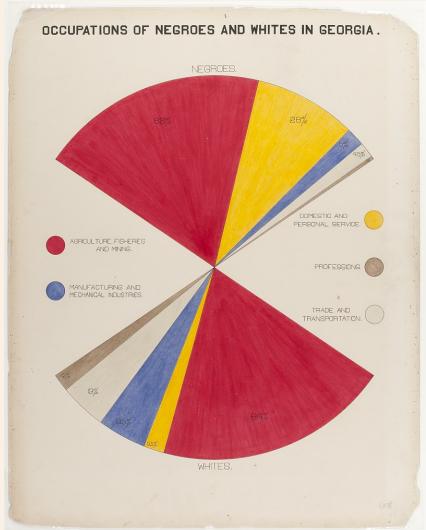
- Categorical/nominal
- Ordinal
- Continuous interval/ratio



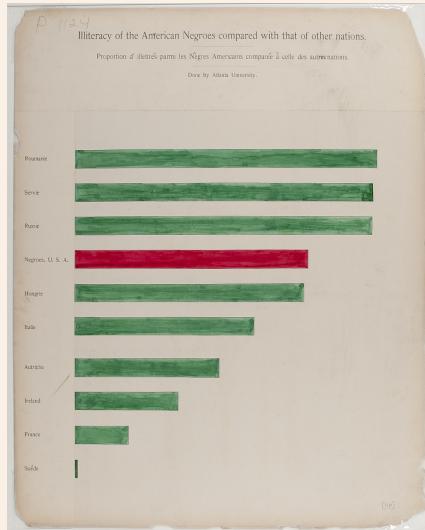
Explore: Chart Types

We use different types of graphs based on the types of data and relationships we are analyzing. Du Bois used variants of most of the major graph types that are still used today:

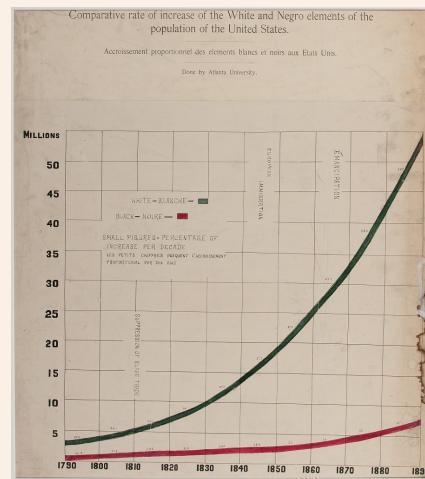
Pie Charts



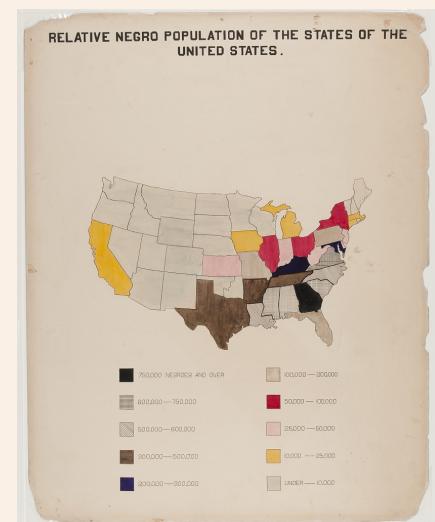
Bar Charts



Line Charts



Statistical Maps

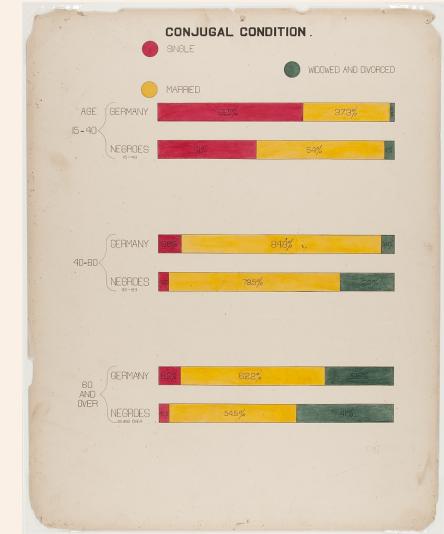
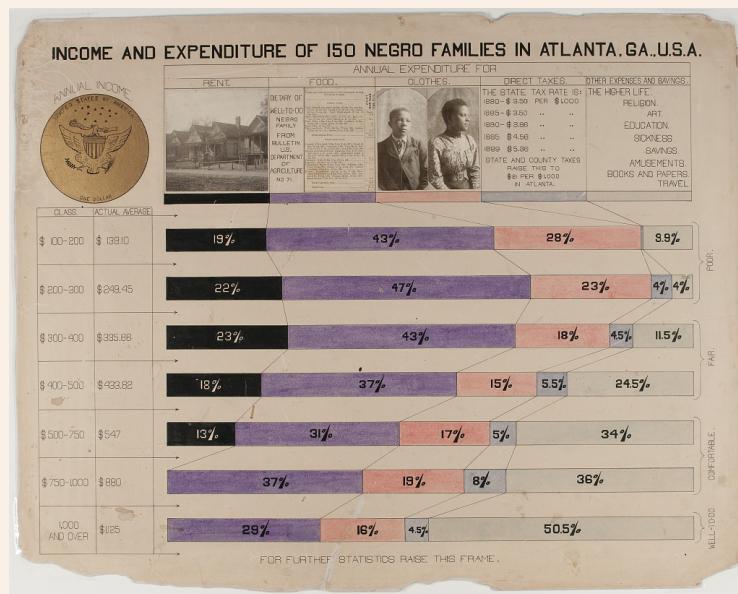
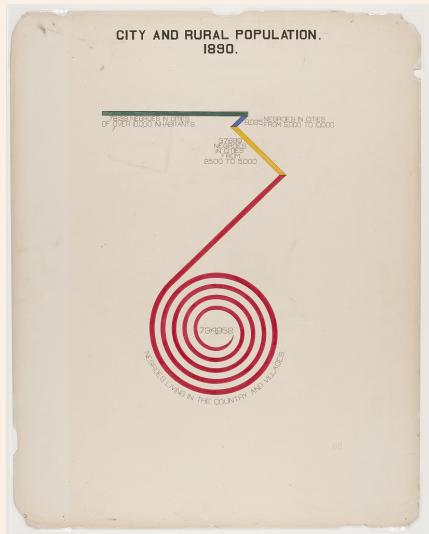




Explore: Chart Types

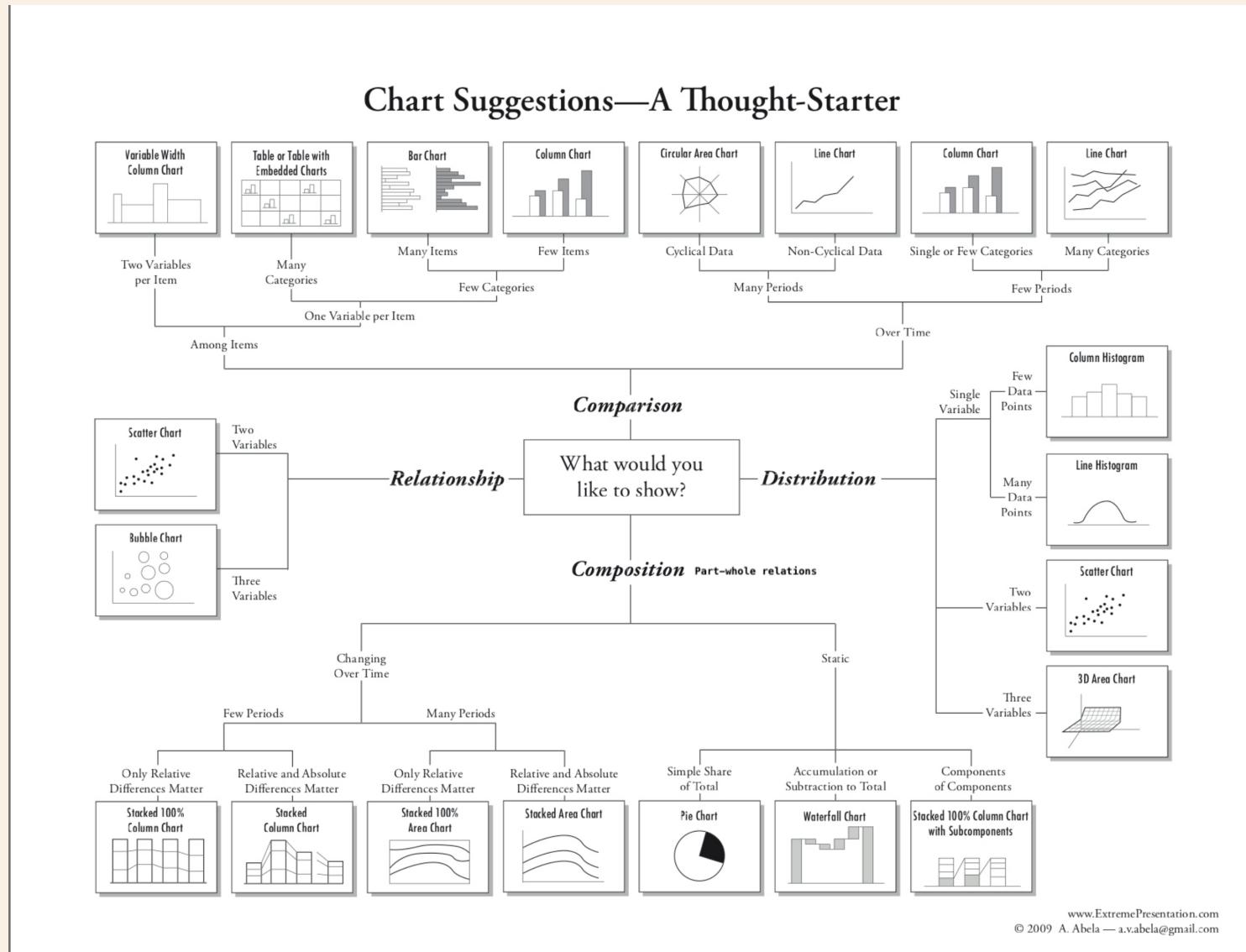
In this lesson, you will learn about the four core types of charts employed by Du Bois and by scientists today.

You can also explore more complex adaptations of these charts using the Du Bois Resources repository for this lesson*
The types include the fanciful Du Bois spiral, stacked bar graphs, and integrated photographs.





Explore: Chart Types



www.ExtremePresentation.com
© 2009 A. Abela — a.v.abela@gmail.com

The Resources page also provides this diagram about many variations of chart types for different types of data and analyses

A Note on Accessibility (a11y)



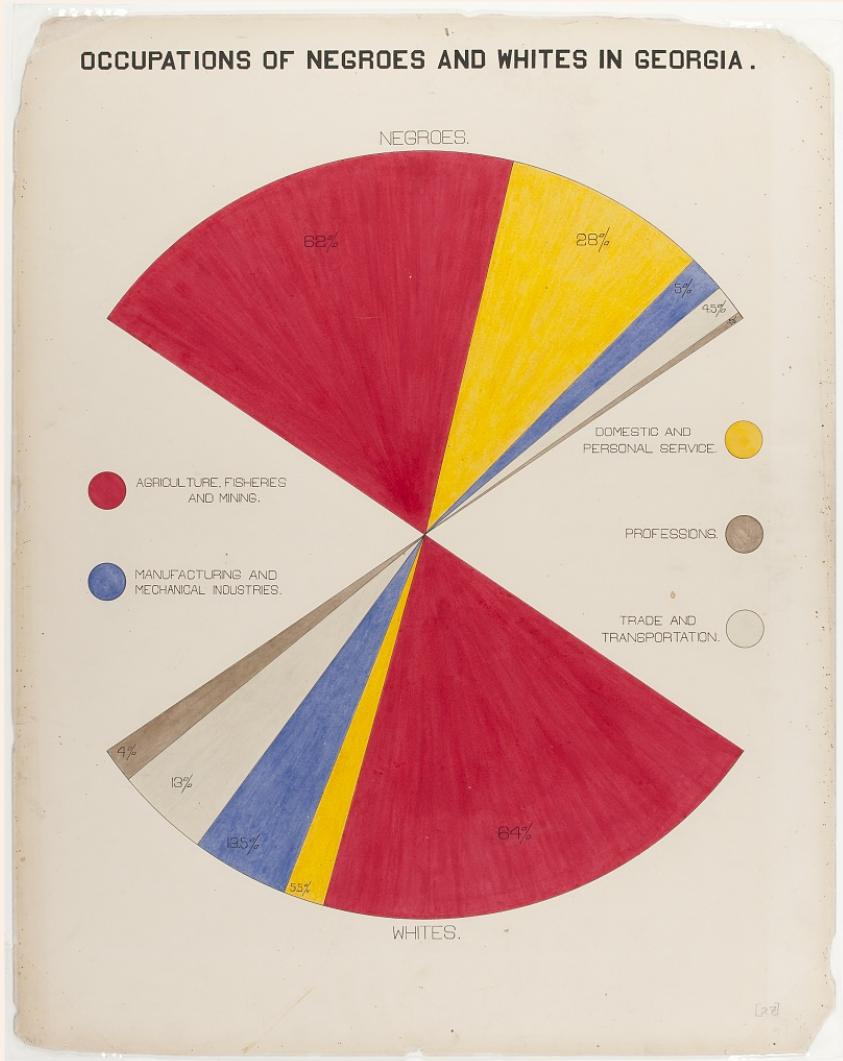
Accessibility involves implementing various design principles and technologies when creating information, to accommodate people with visual, cognitive, auditory, or motor impairments.

Considerations include:

- Proper use of color and contrast
- Use of alternative text
- Using keyboard navigation
- Use descriptive titles and labels
- Avoiding over-reliance on color
- Offering both visual and non-visual formats
- Including executive summaries



Explore: Chart Types: Pie Charts

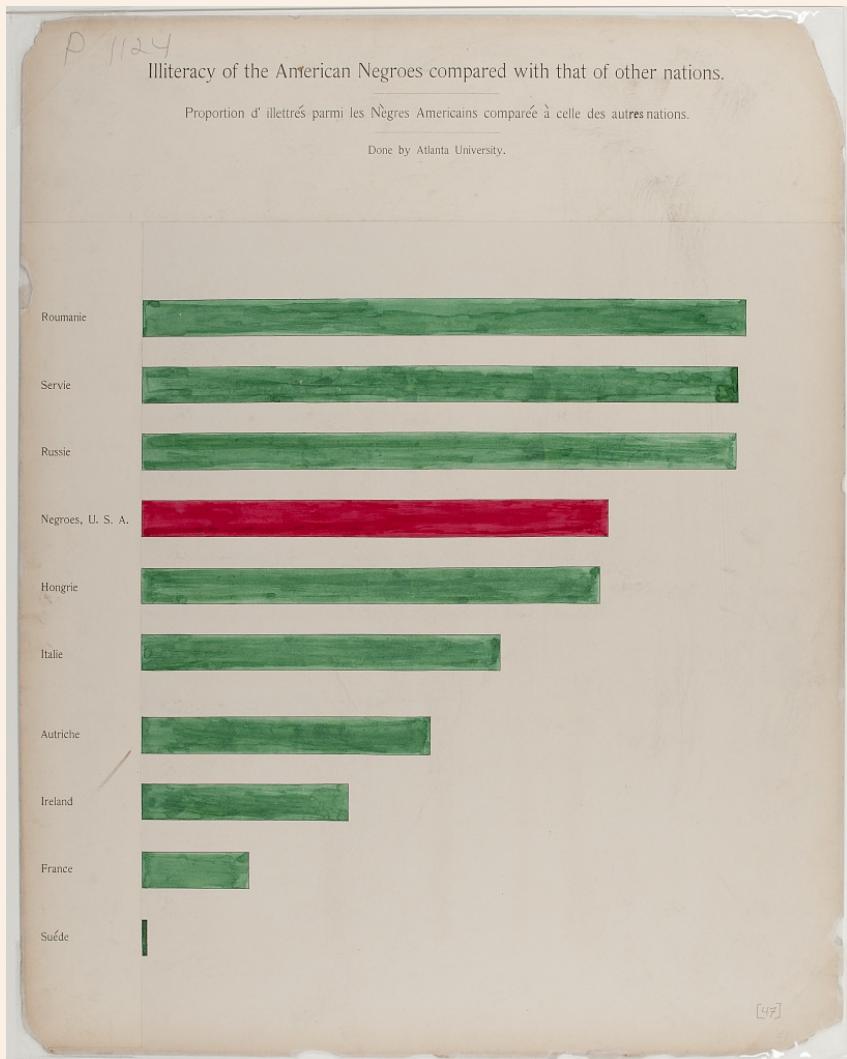


Pie graphs illustrate the percentages of categories (like occupations) within a larger unit (like a population) where all the percentages add up to 100%.

This analyzes a one-dimensional distribution across one categorical variable.



Explore: Chart Types: Bar Charts

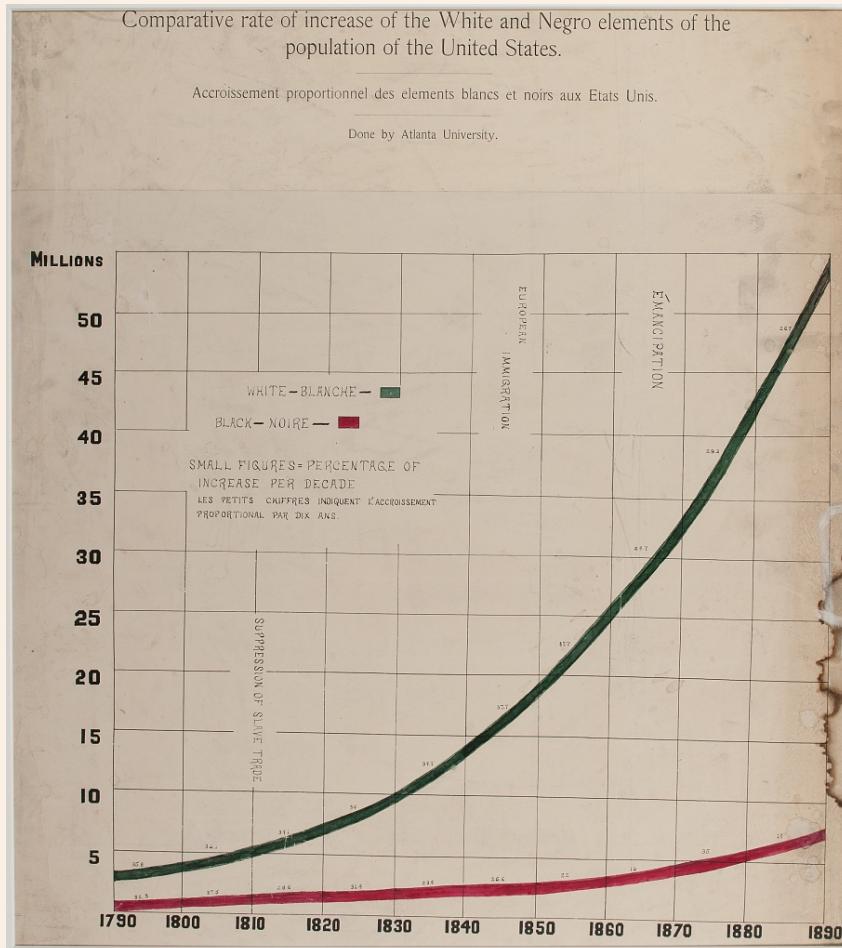


Bar graphs compare frequencies or percentages of one category (like literacy) among other categories (like race or nation).

This helps us analyze two-dimensional relationships, typically between two categorical variables.



Explore: Chart Types: Line Graphs



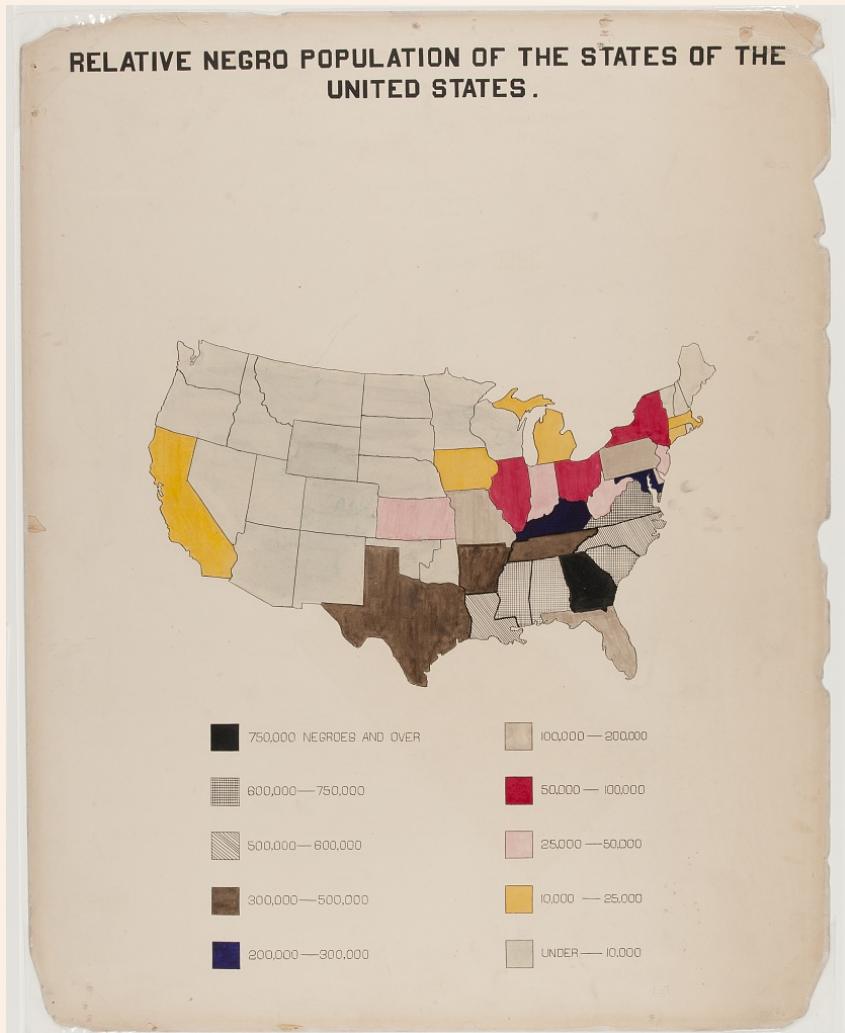
Line graphs plot frequencies or percentages of a continuous interval-ratio variable (like total population) on a y-axis among categories represented by different lines (like racial groups) over a third category of another ordinal or interval ratio variable on an x-axis (like year).

This analyzes three dimensional relationships between three different variables, including interval ratio variables.

Time series line graphs, with time on the x-axis, are the most common type of line graph



Explore: Chart Types: Statistical Maps



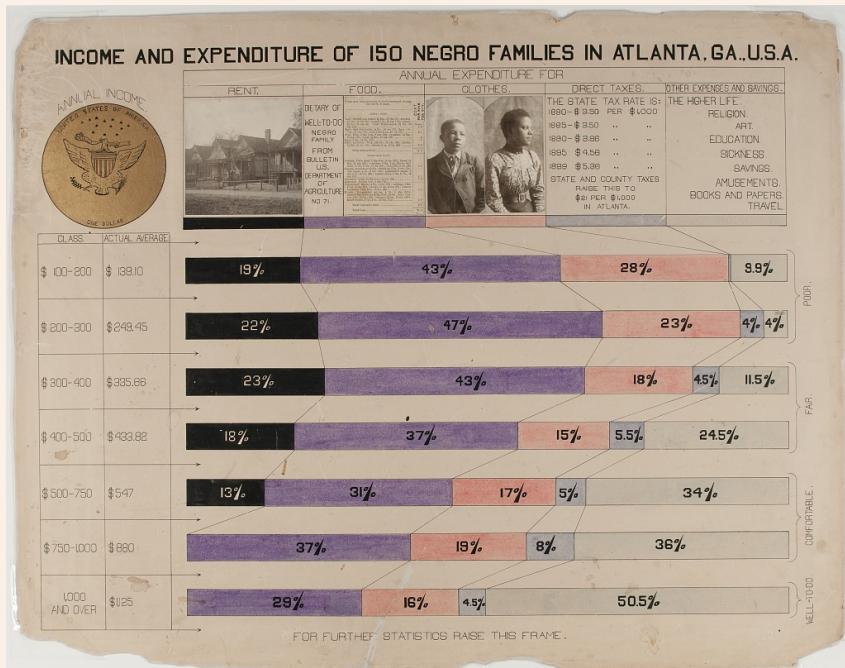
Statistical maps graph geo-spatial distributions of continuous interval-ratio variables (like the Black population of the U.S.).

This allows us to graph a multi-dimensional relationship including two-dimensional spatial locations.



Chart Types: Multivariate Stacked Bar

Multivariate Stacked Bar Graph by Continuous Covariate Brackets with Photographic and Other Data Element Details



Analyze this data visualization and write down your observations about the spending habits of the Black families in Atlanta, GA.

How impactful are the photographs included in this visualization? Are the photographs adding to the information being presented in any way?

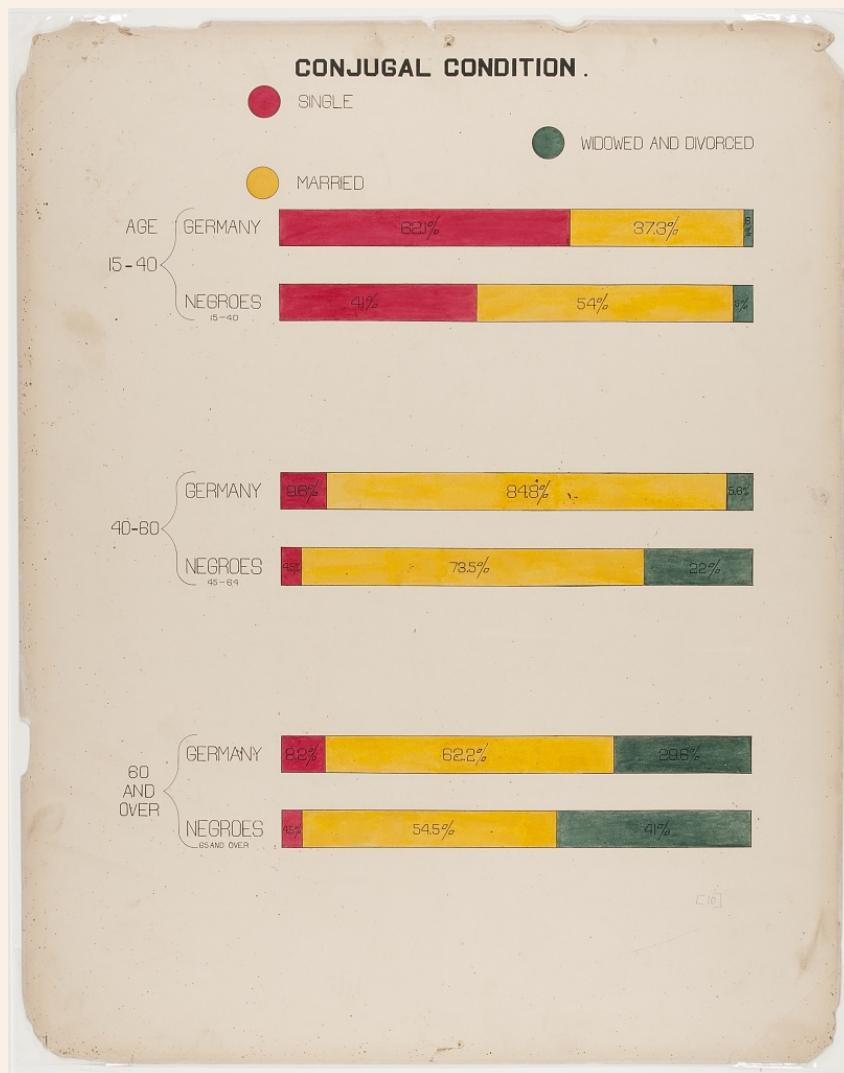
How is this visualization similar to charts you've seen before? How is it different?

What can you observe about families making \$750 and greater annually?



Chart Types: Partial Table Bar Graph

Bivariate Categorical Relationship
(Marriage Status by Racial / National Group) Broken Out by Control Variable (Age)



How are the graphs below similar to ones that you have seen before? How are they different?

What story is being told in the below graphs?

Why do you think the percentage of Black people ages 15 - 40 is higher than German people as displayed in this chart?

Why do you think the widowed and divorced percentages are higher for Black people than German people across all three age groups?

Does the use of horizontal bars make this graph easier to read versus using vertical bars?



Analytic Methods: Progressive Disclosure

Each of these core types of graphs can be adapted or combined to analyze additional dimensions, scales, or components of data and research questions.

Du Bois' use of multiple charts together is sometimes referred to as progressive disclosure.

Progressive disclosure reveals additional, but related information as an audience moves between graphs. This is analogous to zooming in with a telescope to see the qualities of leaves on a tree; and then zooming out to see a whole forest comprised of many trees.

Analytic Methods: Data Stories



Data stories provide a narrative that help us think about the complex relationships that we visualize with multiple graphs.

While the human brain can process images faster than words, we often comprehend and remember more complex ideas through narrative and stories.

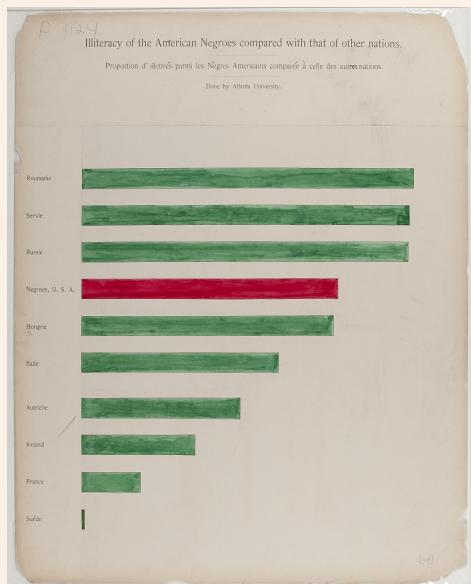
Du Bois's Paris exhibition connected multiple graphs with a scientifically rigorous data story of Black achievement after emancipation, contradicting social darwinism.



Explore: Examples

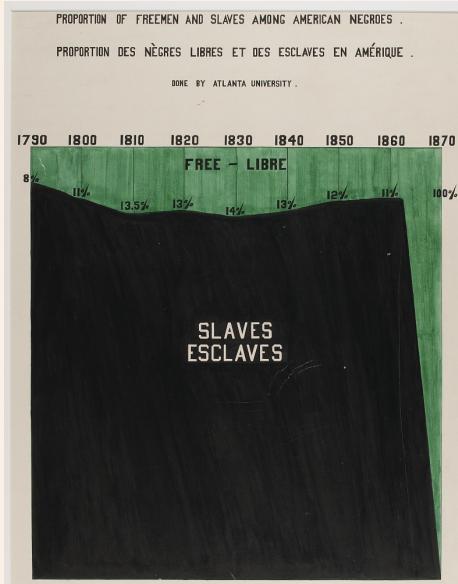
Below are four examples of Du Bois visuals for you to explore. The examples are shown in terms of increasing complexity, and show a variety of chart types that are good exemplars of the Du Bois style.

Example 1



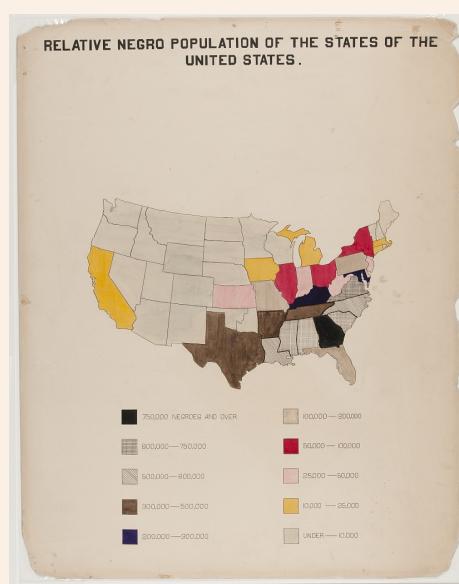
Illiteracy of American Negroes compared with that of other nations

Example 2



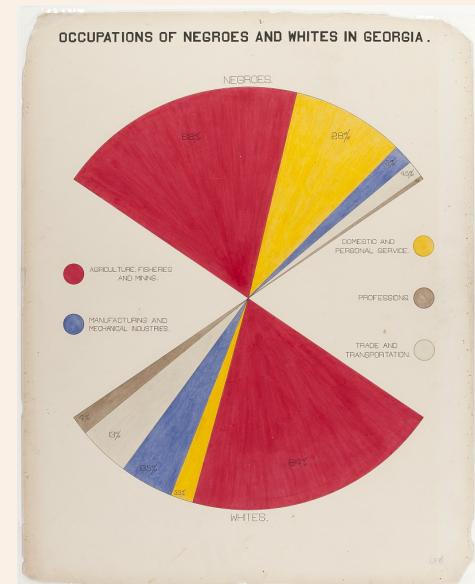
Proportion of Freemen and Slaves Among American Negroes

Example 3



Relative Negro Population of the United States

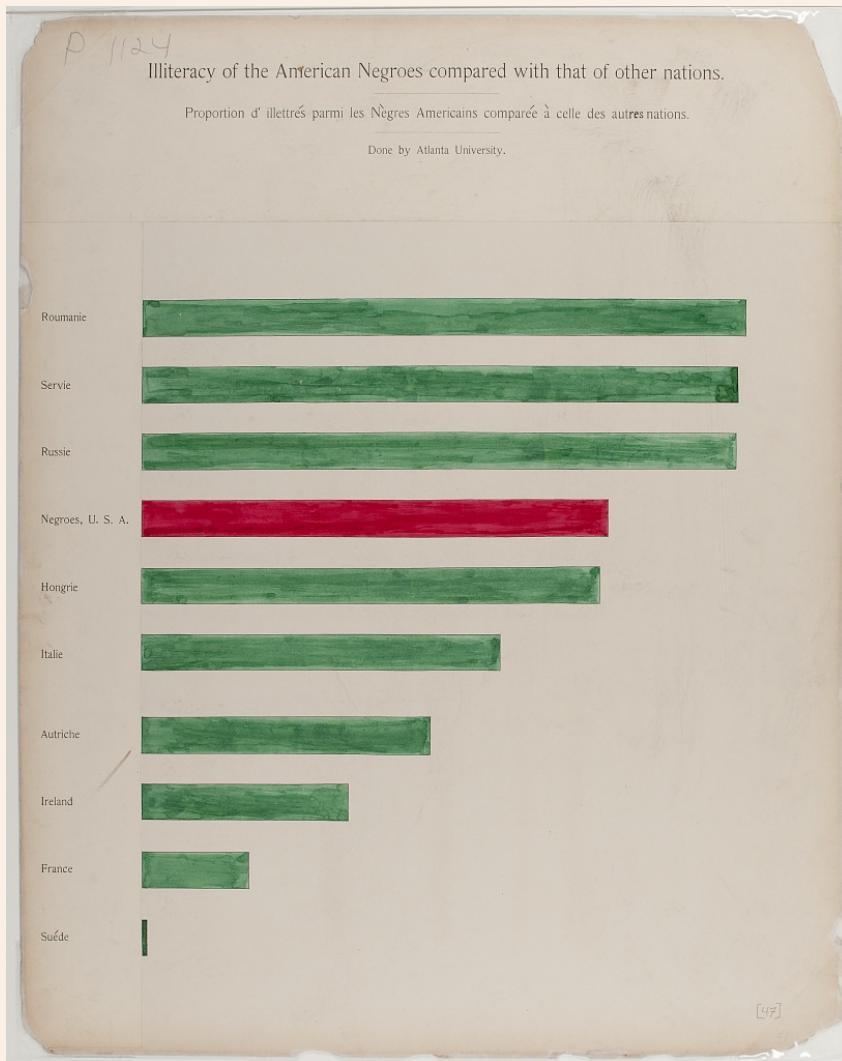
Example 4



Occupations of Negroes and Whites in Georgia



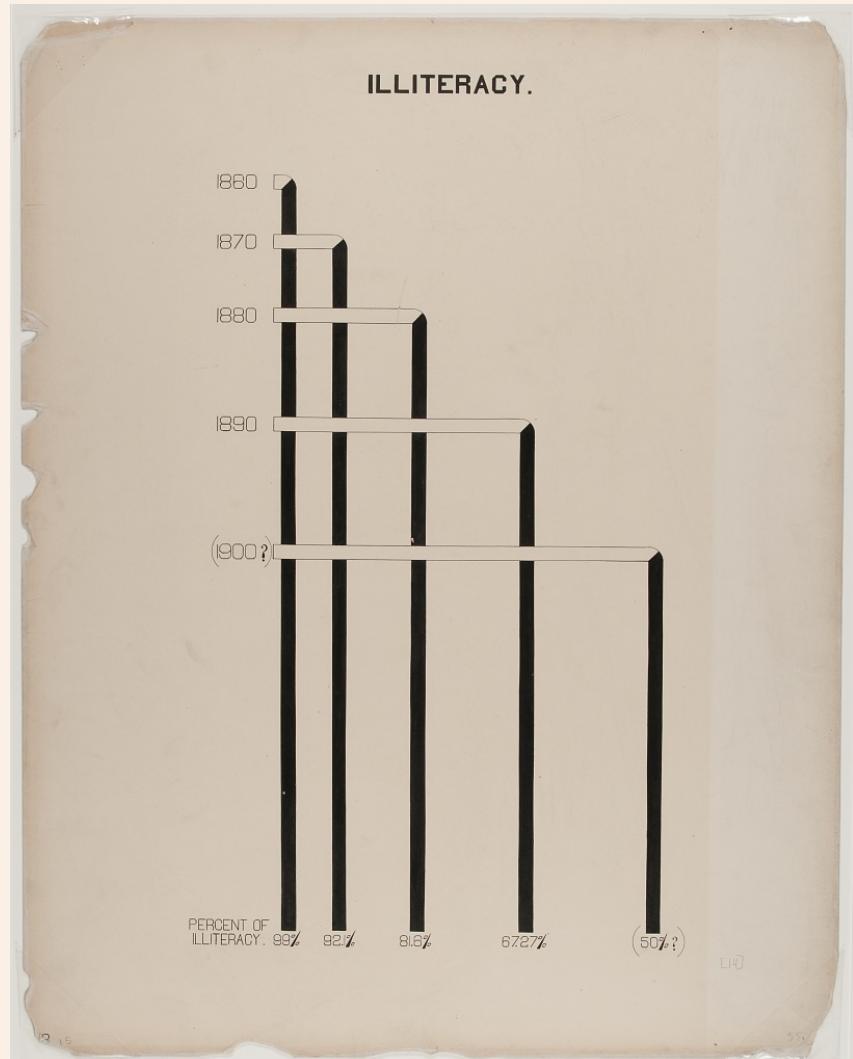
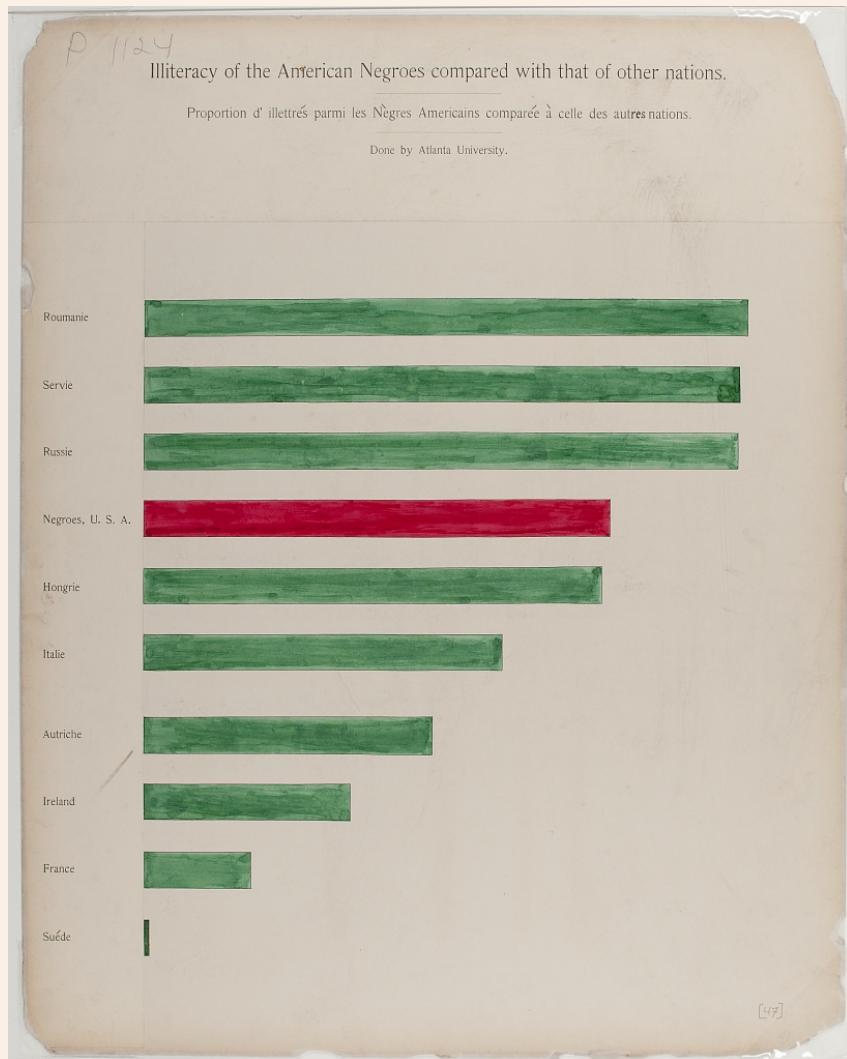
Explore Example 1



Illiteracy of American Negroes compared with that of other nations



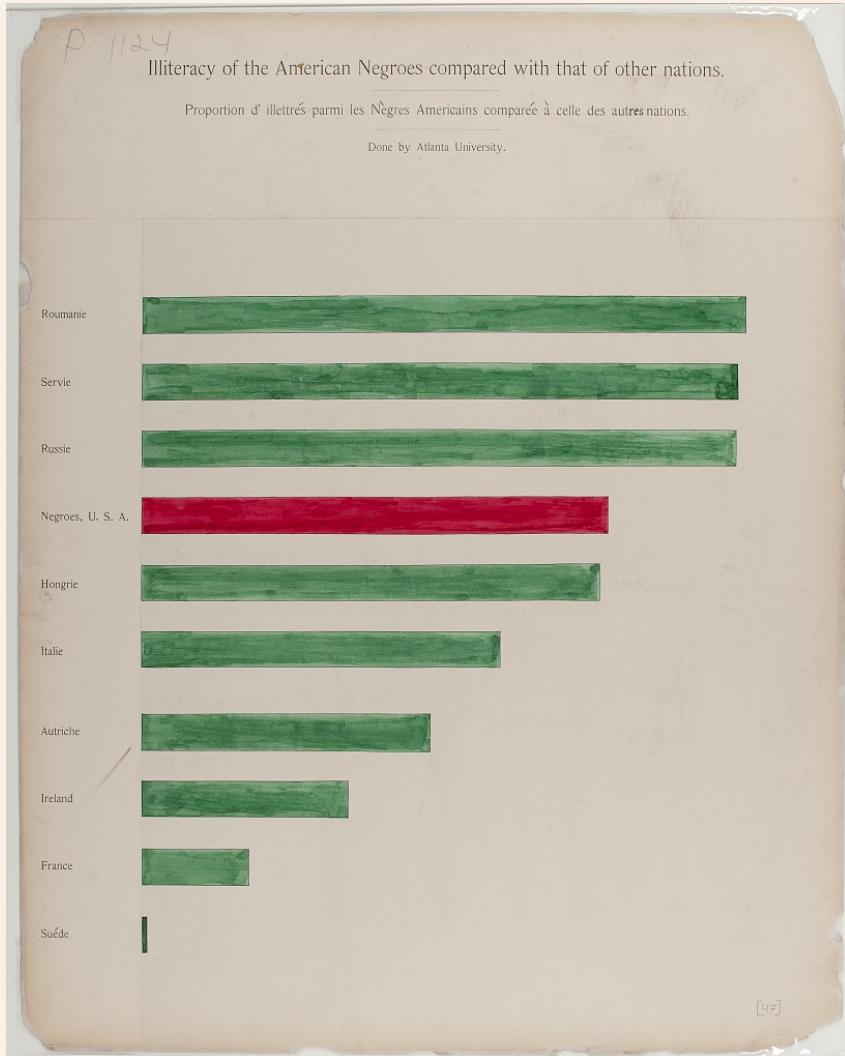
Explore Example 1: Context



An important context of Du Bois's graph of Black illiteracy is the “illegality of literacy” for enslaved people. After emancipation, illiteracy declined rapidly as Black Americans sought to empower themselves through education, as shown on the chart on the right.



Explore: Question and Audience



The research question behind this visualization is tied to Du Bois' larger research agenda and the audience at the Paris Exposition:

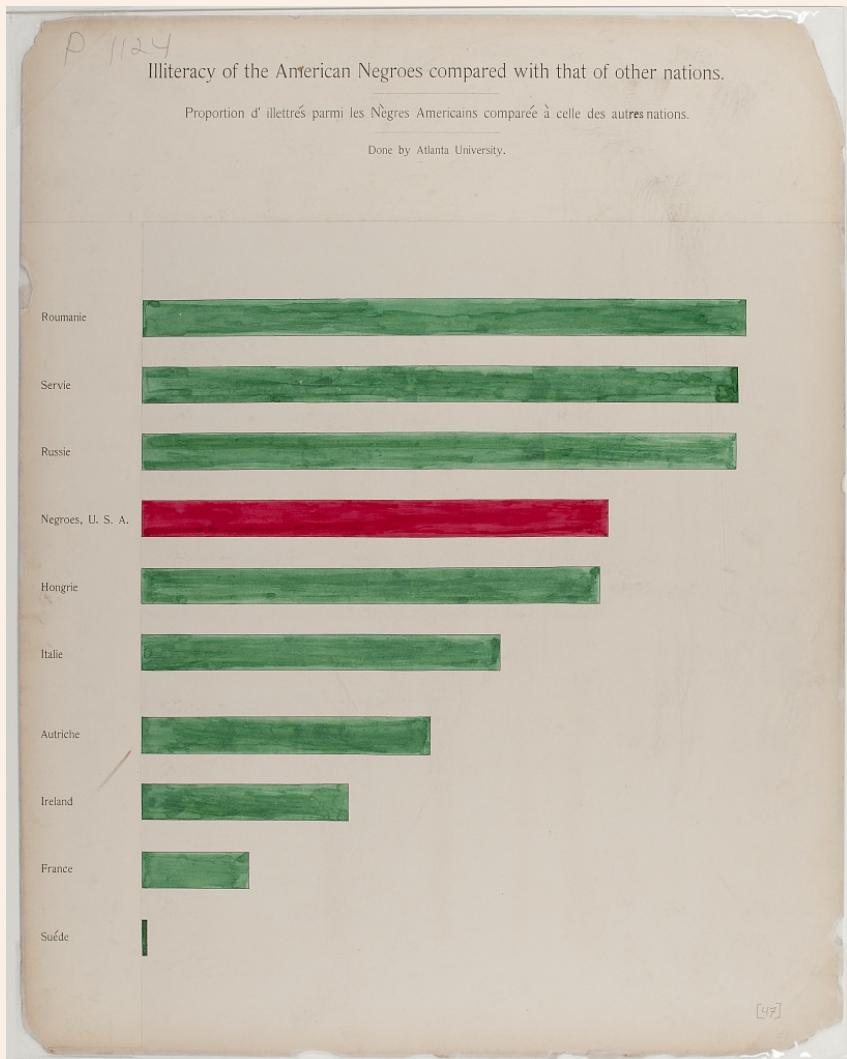
What have Black Americans accomplished since emancipation that defied racist view of the time?

Speaking to the expo's global audience, this graph answers this question by comparing illiteracy of Black Americans (outlined in red) to that of other countries.

The Black illiteracy rate is higher than countries like France, but better than others like Russia where serfdom still denied literacy to peasants.



Explore: Analytic Methods



Consider Du Bois' analytic methods:

What type of graph is this?

How many variables are analyzed?

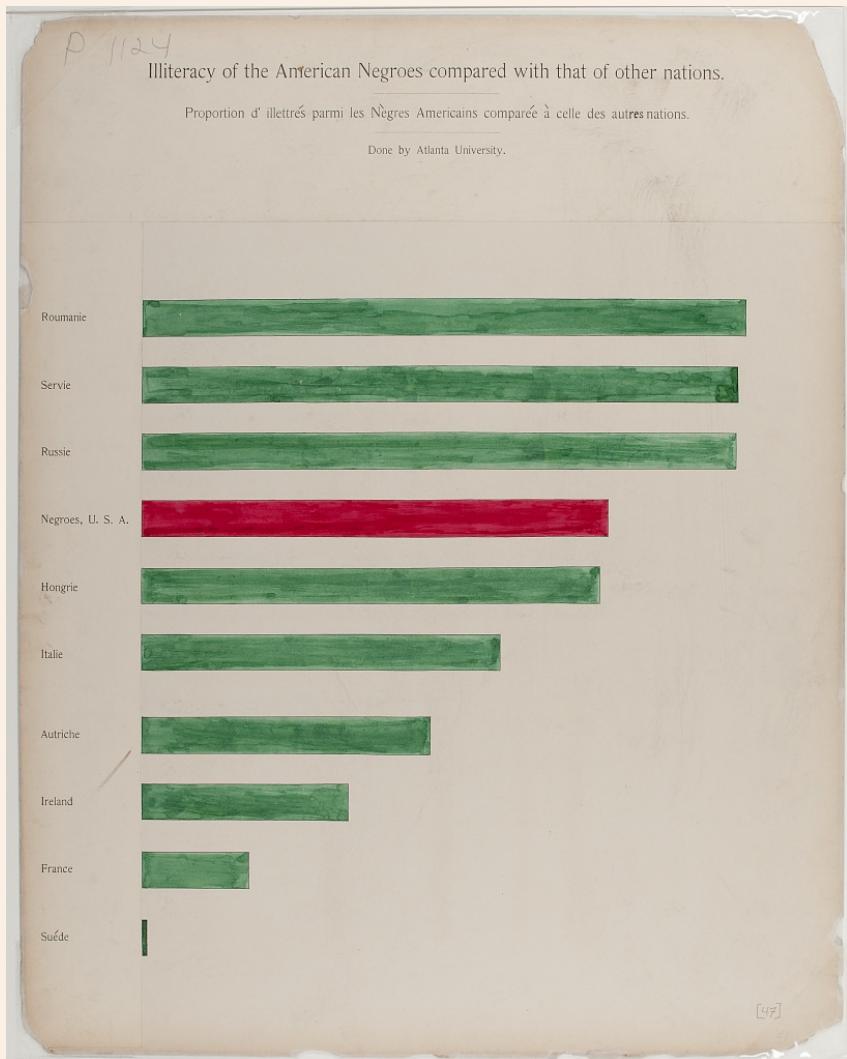
What statistics are plotted?

Can you make clear comparisons across the categories or dimensions of the graph? What are they?

What answer to the research question do you see in the visual?



Explore: Aesthetics and Design



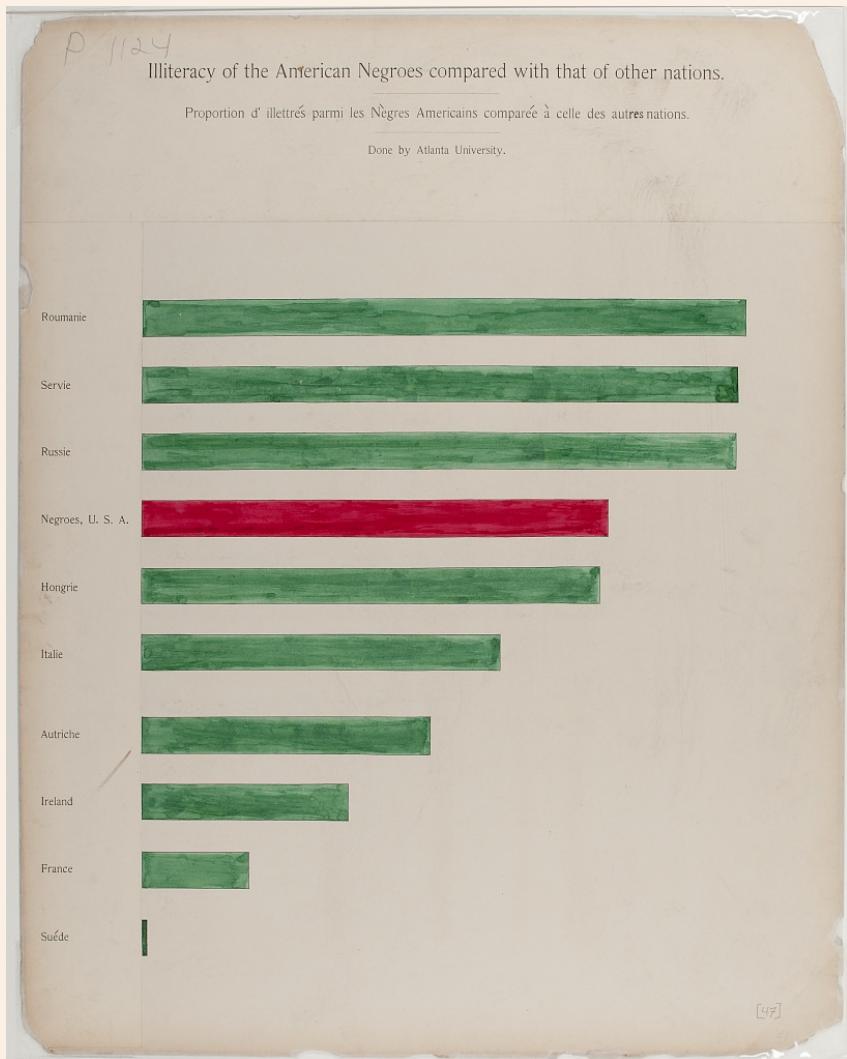
This visual, a conventional bar graph, uses spot color to highlight the data for Black Americans compared to other countries, showing the illiteracy rate to be at the midpoint compared to other nations.

The chart portion is a large percentage of the canvas, simply showing the message.

Note the bilingual labels and titles (a nod to the venue)



Explore: Reflection



Do you think this chart type or another chart type is best for answering Du Bois' question with this data? Why?

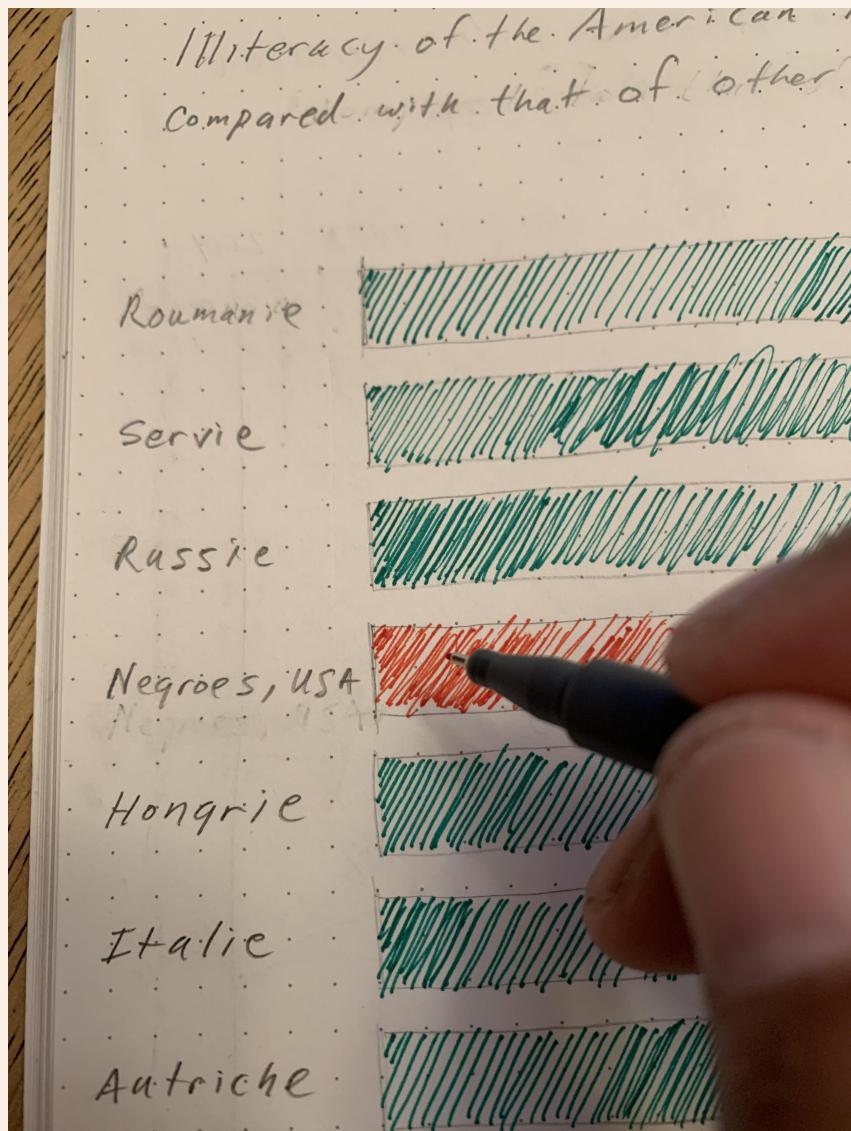
Would you change any of the aesthetic choices and why? Consider dimensions, color selection, titles, data labels, legend, notes etc.

What additional questions, data, or charts does this graph make you think of?

What tools would you use to make or improve on this chart yourself?



Explore: Re-Create

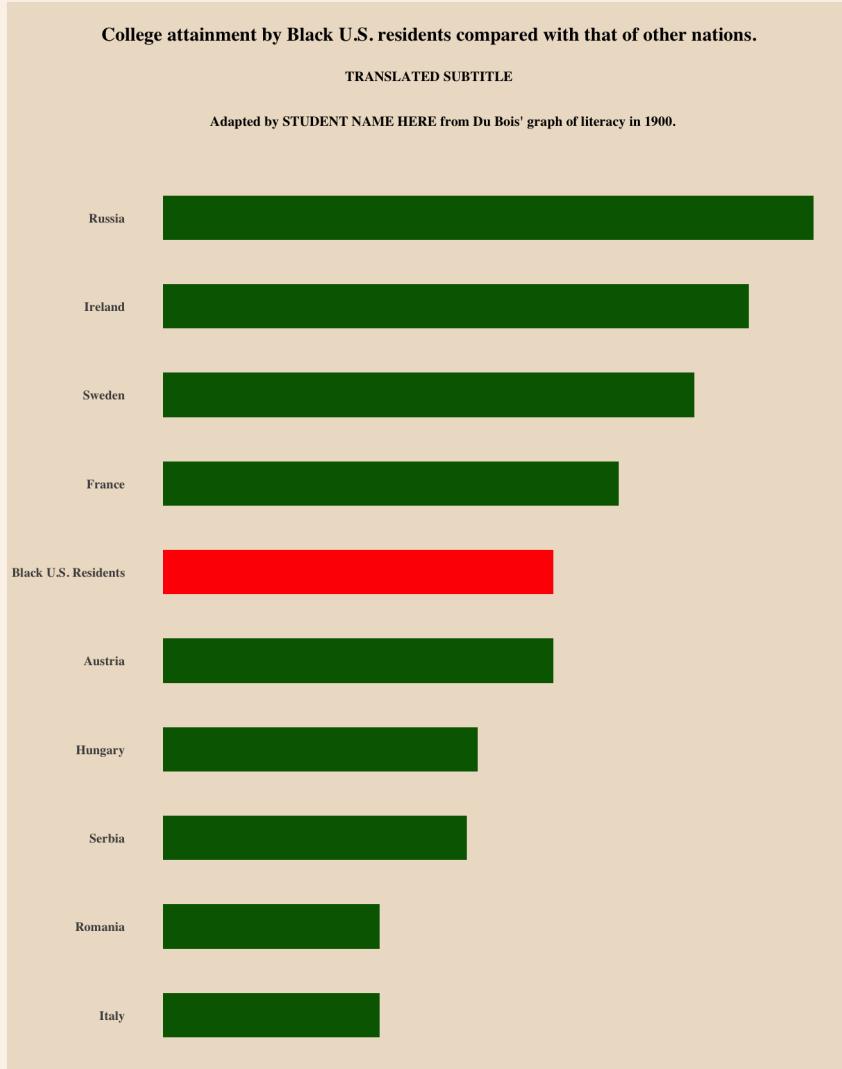


Re-create the visual using pens, paper and other analog materials like markers, colored pencils, rulers, and protractors.

Strive for rapid iteration, and do not focus on precision, but rather exploring the various components of the chart, and how they are conveying the message and information. Students may work together or individually.

At the end of the session, the re-creations may be shared and critiqued as a group.

Explore: Re-Create with Modern Data



Country, Percent attainment

Romania, 20

Serbia, 28

Russia, 60

Black U.S. Residents, 36

Hungary, 29

Italy, 20

Austria, 36

Ireland, 54

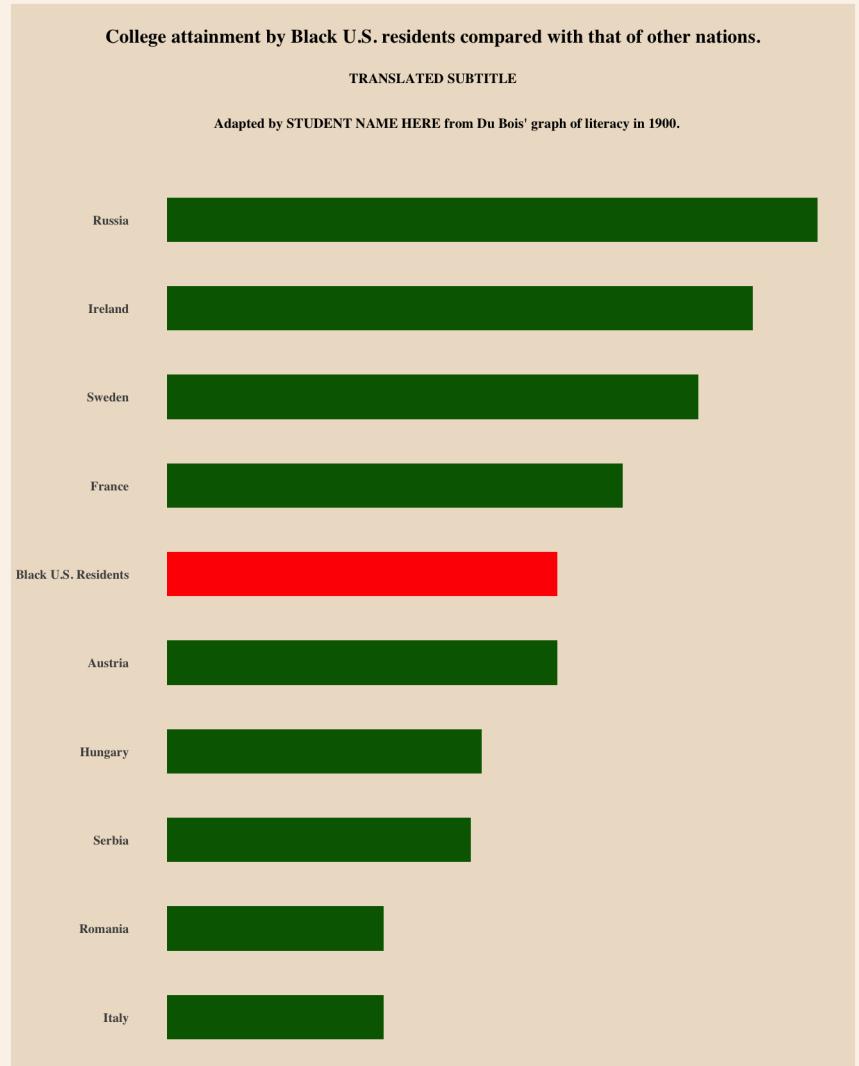
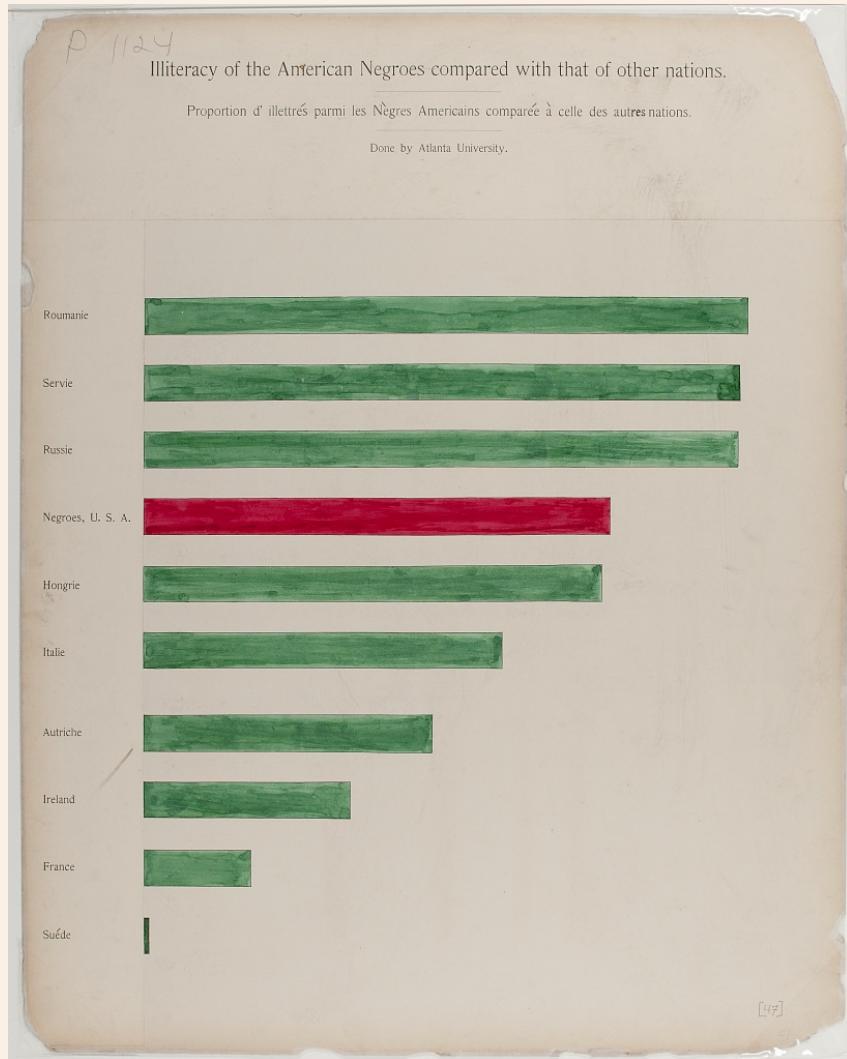
France, 42

Sweden, 49

Using modern data on college attainment, comparing other countries with US Black students.*

*US Census https://www.luminafoundation.org/stronger-nation/report/#/progress/racial_equity
OECD <https://www.oecd.org/en/topics/sub-issues/education-attainment.html>

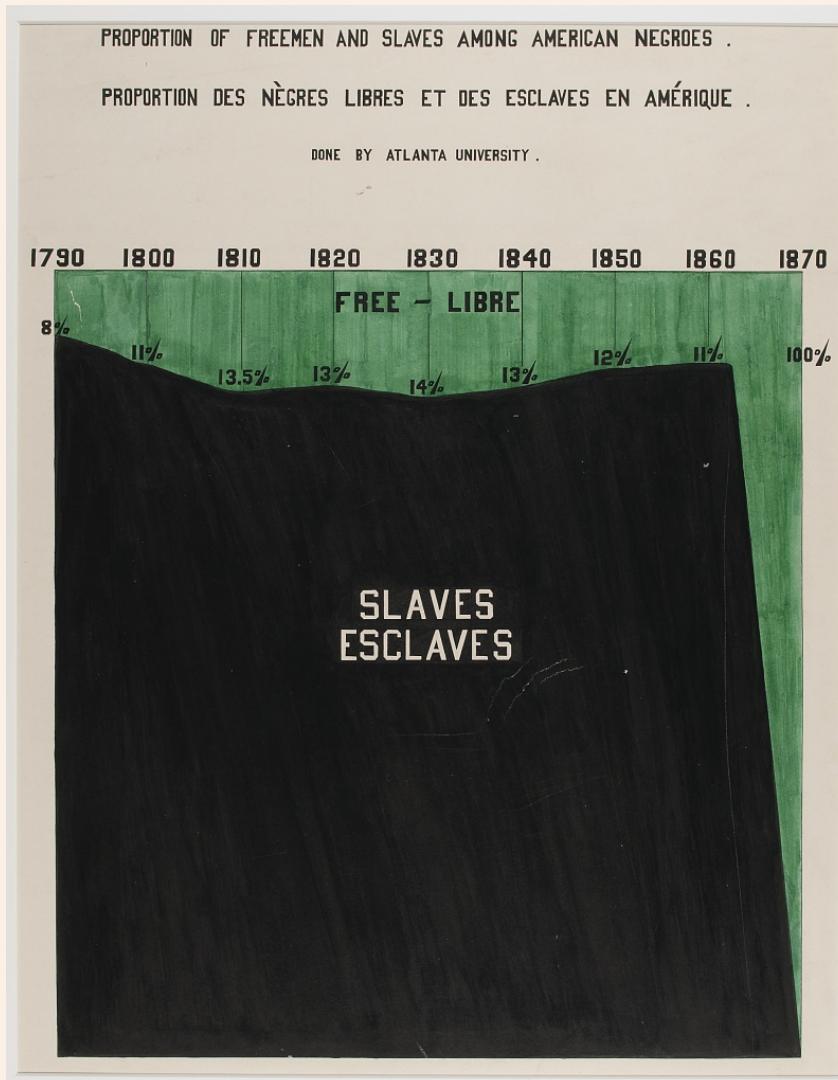
Explore: Re-Create with Modern Data



Original and modern-data versions compared



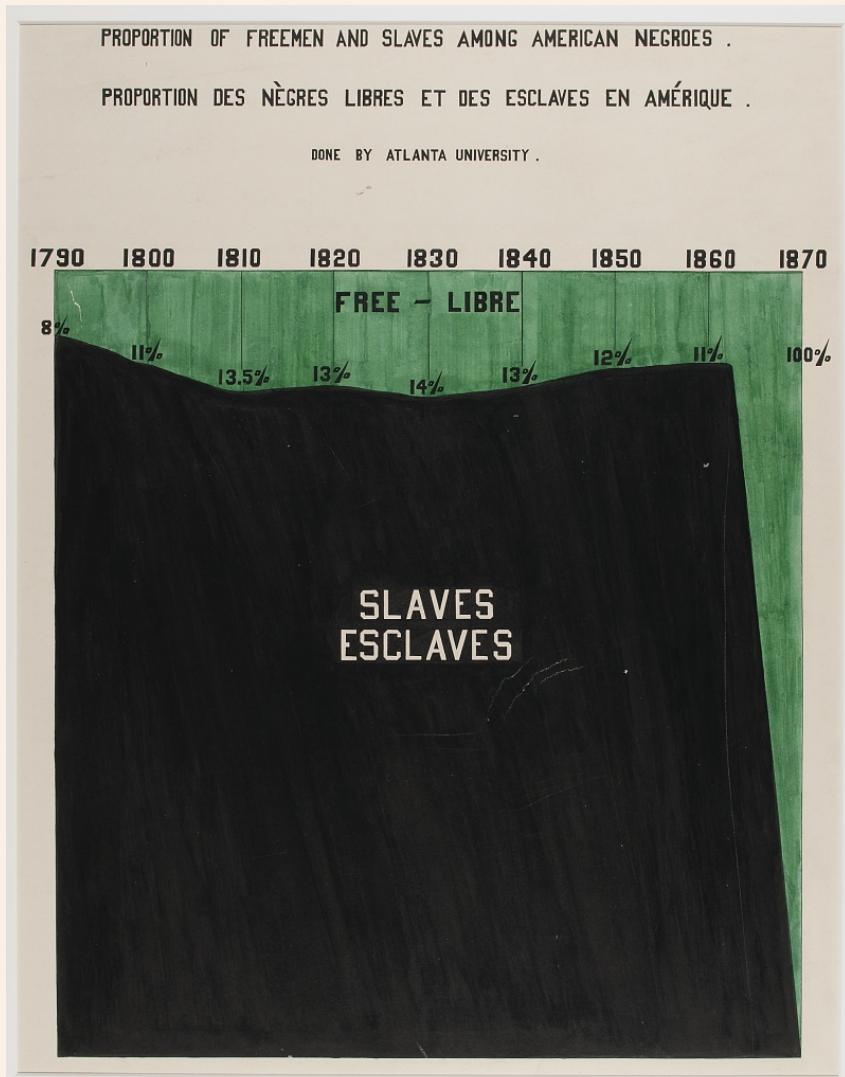
Explore Example 2



Proportion of Freemen and Slaves Among American Negroes



Explore: Message



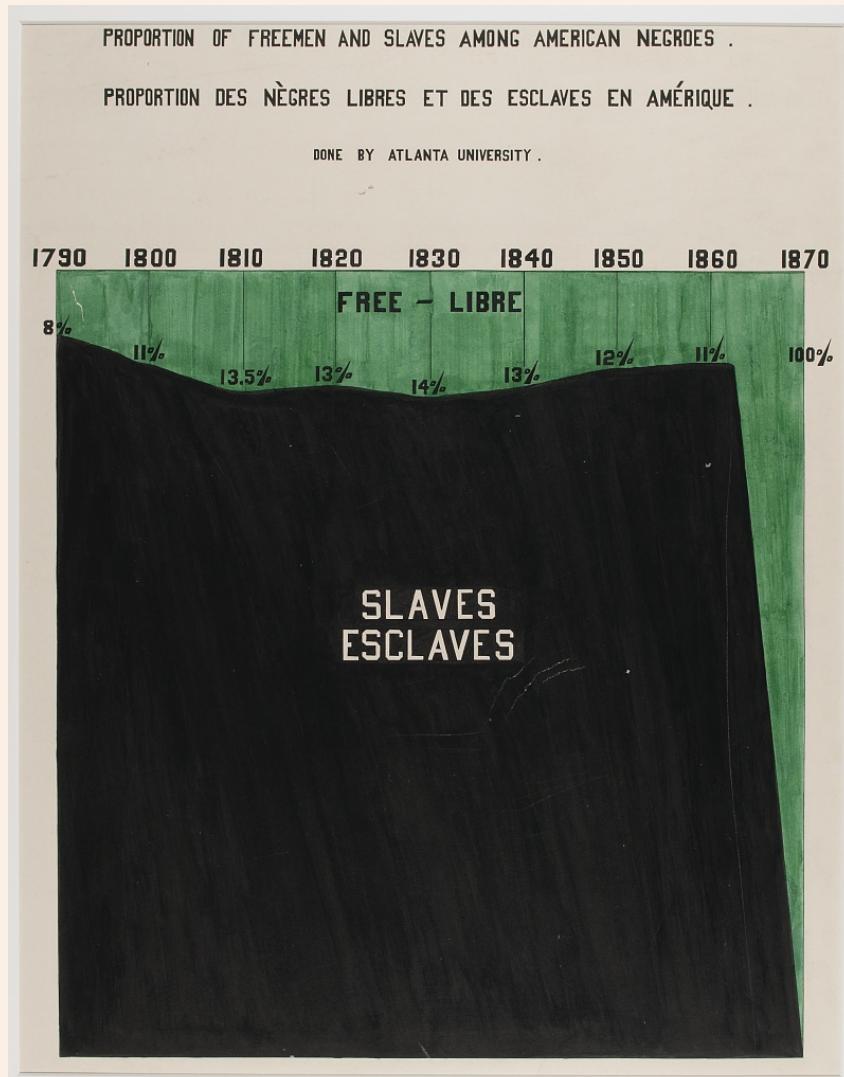
The message of the visual is tied to the overall message from Du Bois at the Paris Exposition:

Describe the status and progress of Black Americans since emancipation, and provide an alternative narrative to racist views of the time.

The visual compares free and slave populations from 1790 to 1870, using an area chart with two primary colors representing freedom and slavery.



Explore: Analytic Methods



Consider Du Bois' analytic methods:

This is a variation of a time series line graph called a stacked area graph. How would a traditional line graph look different?

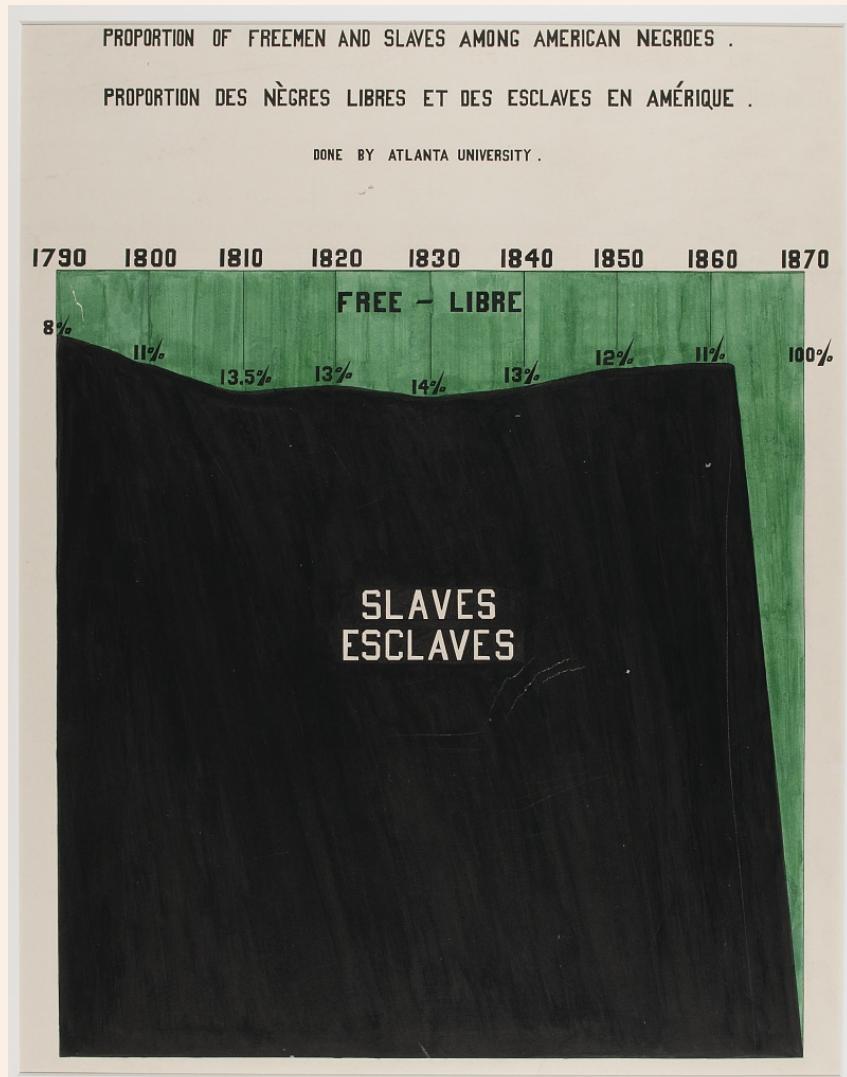
How many variables are analyzed?
What statistics are plotted?

Can you make clear comparisons across the categories or dimensions of the graph? What are they?

What answer to the research question do you see in the visual?



Explore: Aesthetics and Design



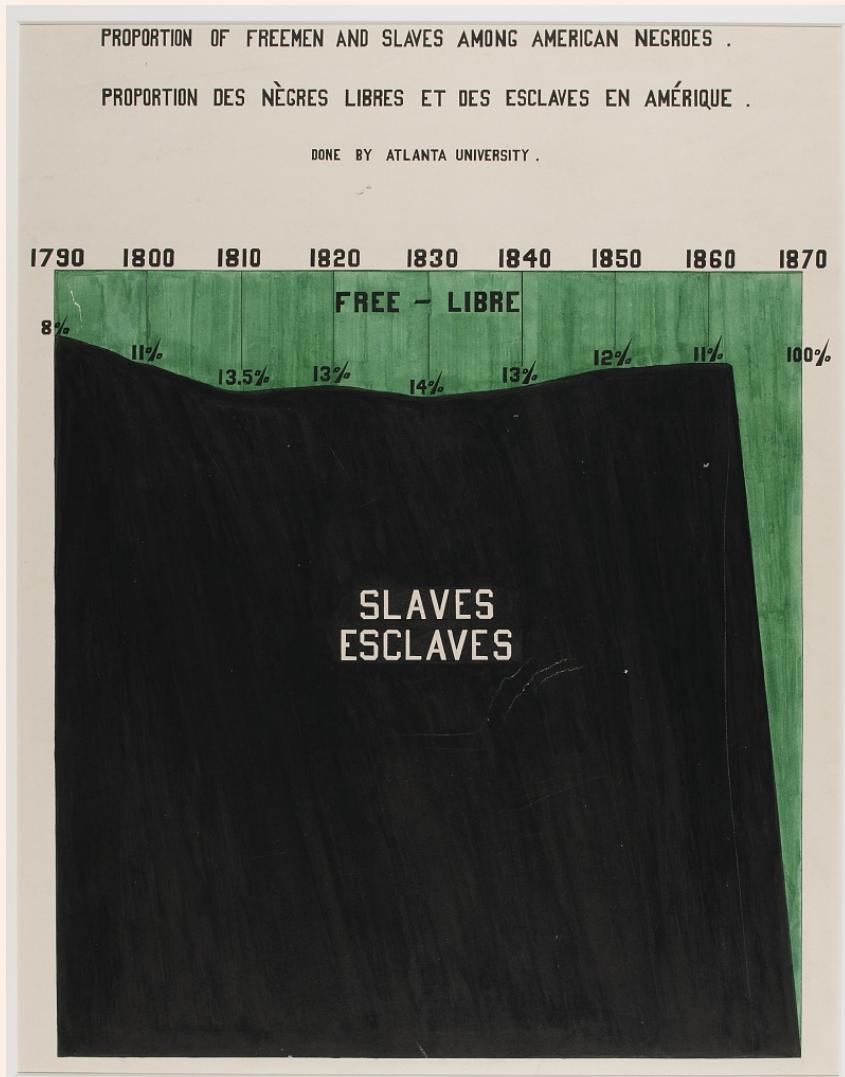
This visual makes uses a bold color contrast (black for slavery, green for freedom) to show the proportion of enslaved and free people.

The area is dominated by black, with contrasting large text, yet the eye is drawn to the sudden jolt of green freedom in 1865. Note that increasing data goes down instead of up.

The titles are bilingual (a nod to the venue), and the X-axis is labeled at the top, contrary to today's convention. Instead of a labeled Y-axis, each individual data point is labeled.



Explore: Reflection



Why do you think Du Bois chose a stacked area graph instead of a simple line graph? What does the stacked area express that a line graph would not?

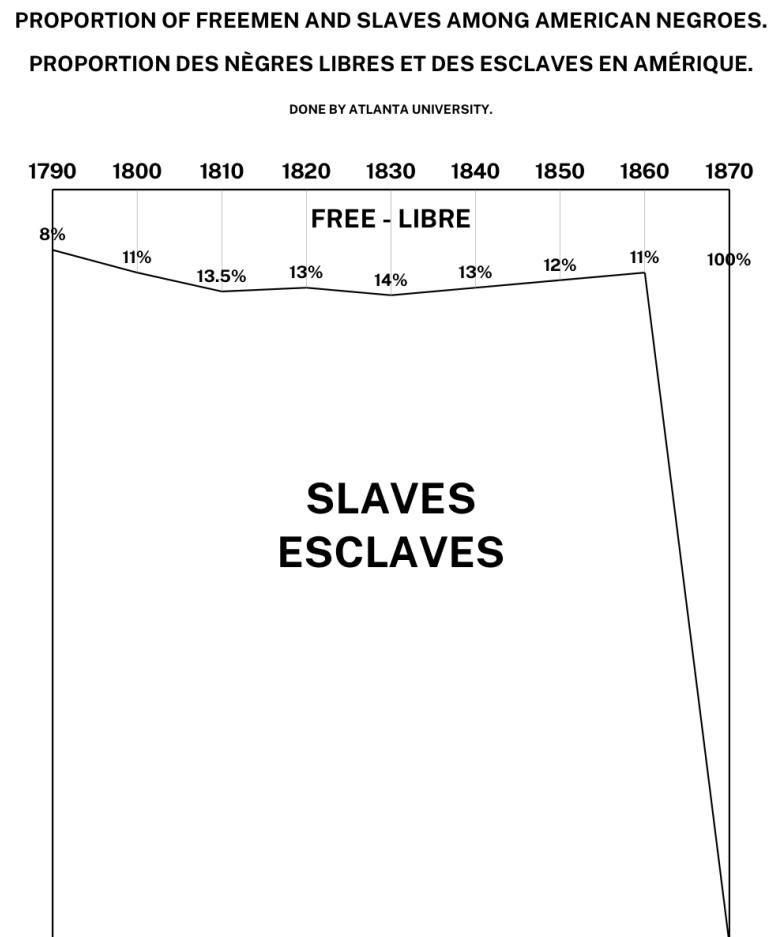
Would you change any of the aesthetic choices and why? Consider dimensions, color selection, titles, data labels, legend, notes etc.

What additional questions, data, or charts does this graph make you think of?

What tools would you use to make or improve on this chart yourself?



Explore: Re-create



Re-create the visual using pens, paper and other analog materials like markers, colored pencils, rulers, and protractors.

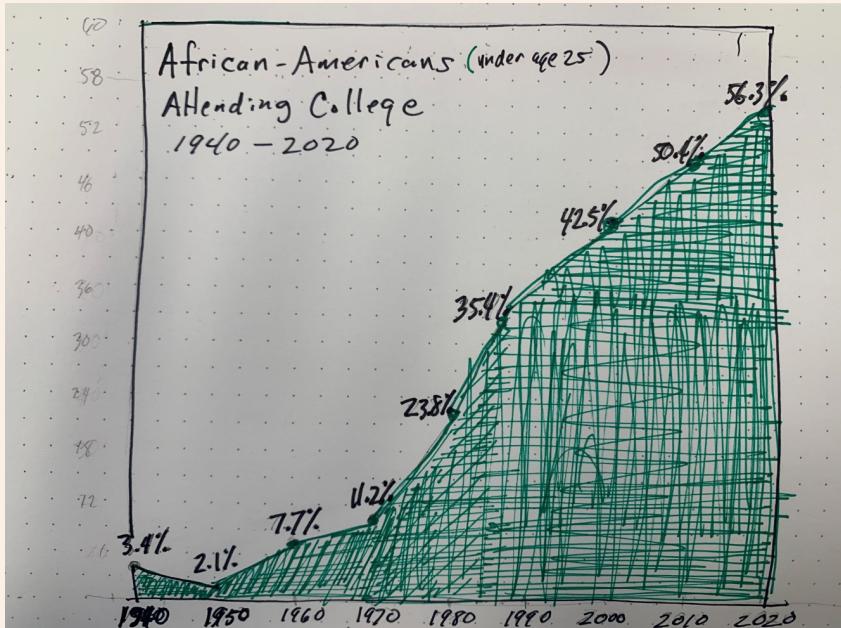
Strive for rapid iteration, and do not focus on precision, but rather exploring the various components of the chart, and how they are conveying the message and information. Students may work together or individually.

At the end of the session, the re-creations may be shared and critiqued as a group.



Explore: Modernize

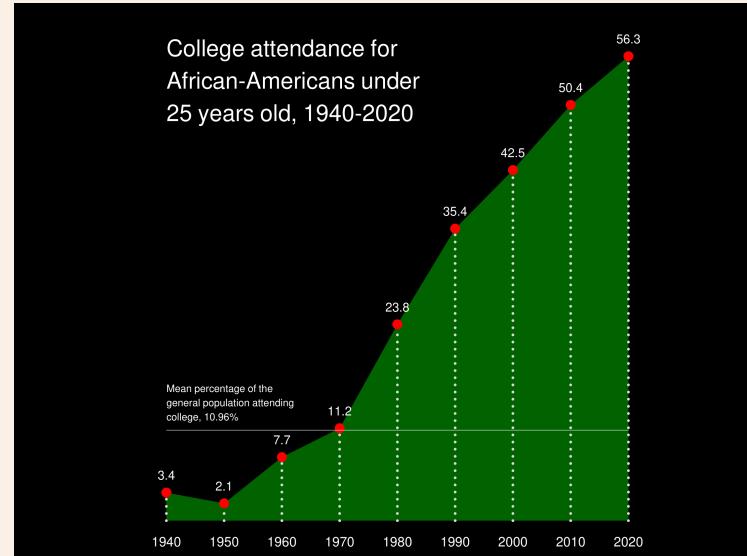
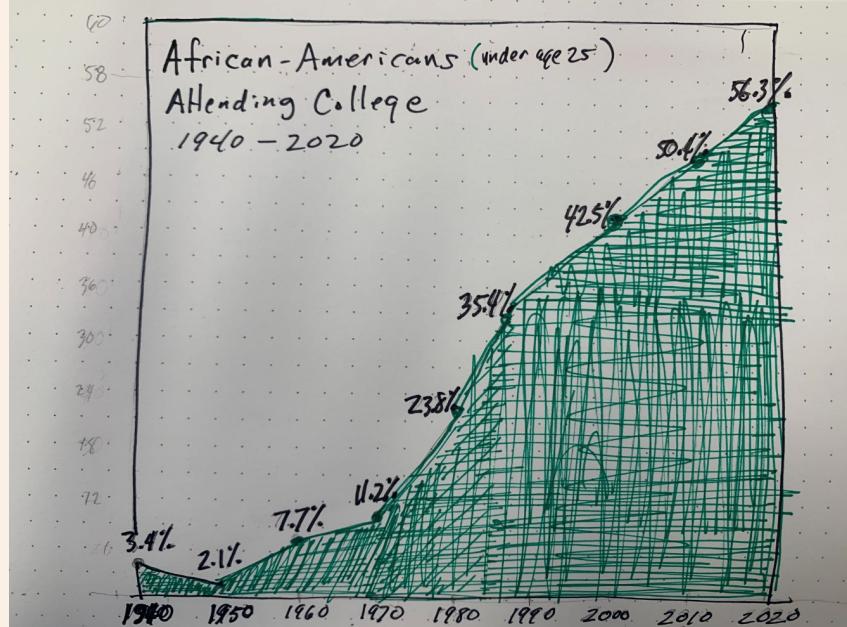
Date	25 and up	25-34	%Increase
1940	3.4	4.2	23.5%
1950	2.1	2.9	38.1%
1960	7.7	11.5	49.4%
1970	11.2	17.1	52.7%
1980	23.8	37.6	58.0%
1990	35.4	44.2	24.9%
2000	42.5	48.4	13.9%
2010	50.4	56.4	11.9%
2020	56.3	59.7	6.0%



In this section, apply modern data to the Du Bois Style. This involves data gathering, selection of appropriate chart type, and defining the elements to apply the Du Bois style (for example color palette, labeling methods) This example shows African American college attendance rates between 1940 and 2020.



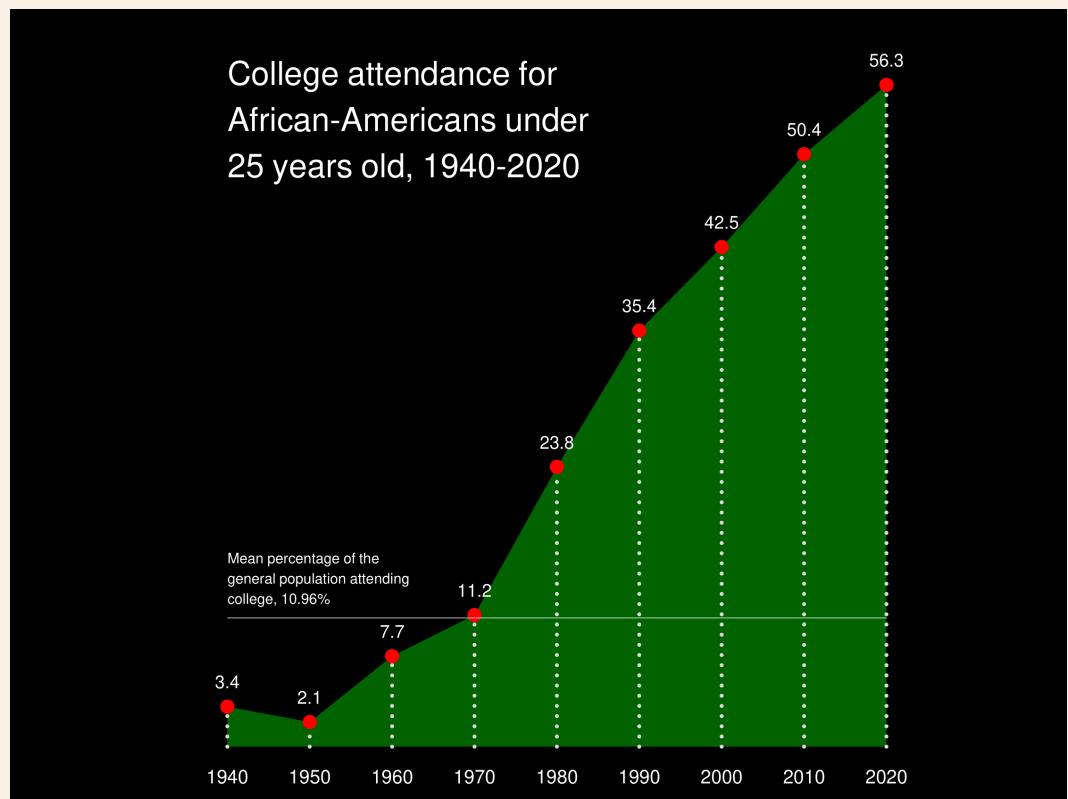
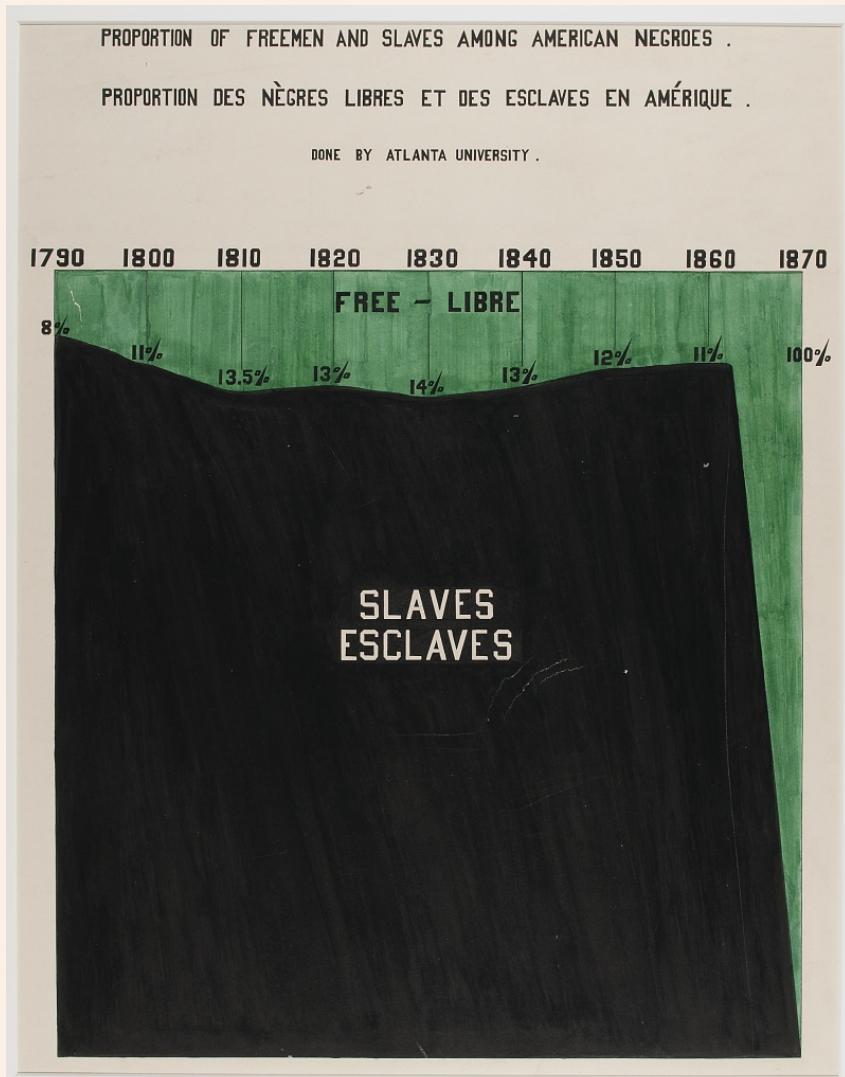
Explore: Modernize



This example updates the hand-drawn version to a digital format, using a color scheme similar to the original, and adding red as a highlight color (note the Pan-African black-red-green theme)



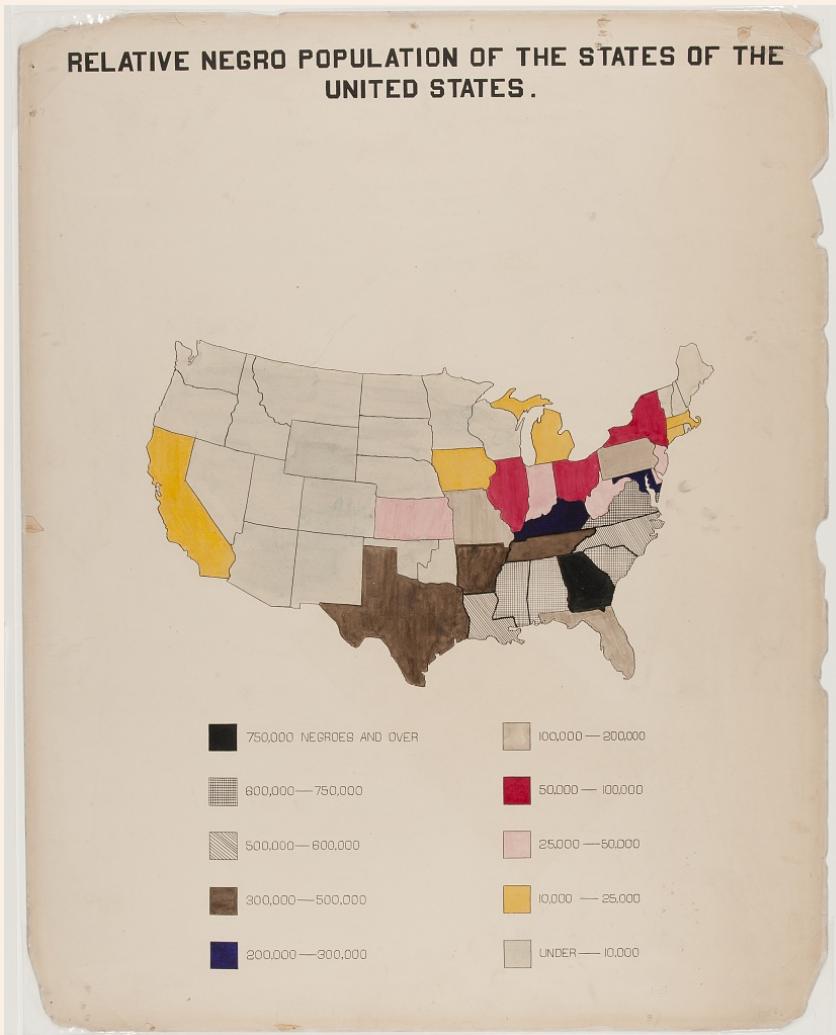
Explore: Modernize



Original and Modern versions compared



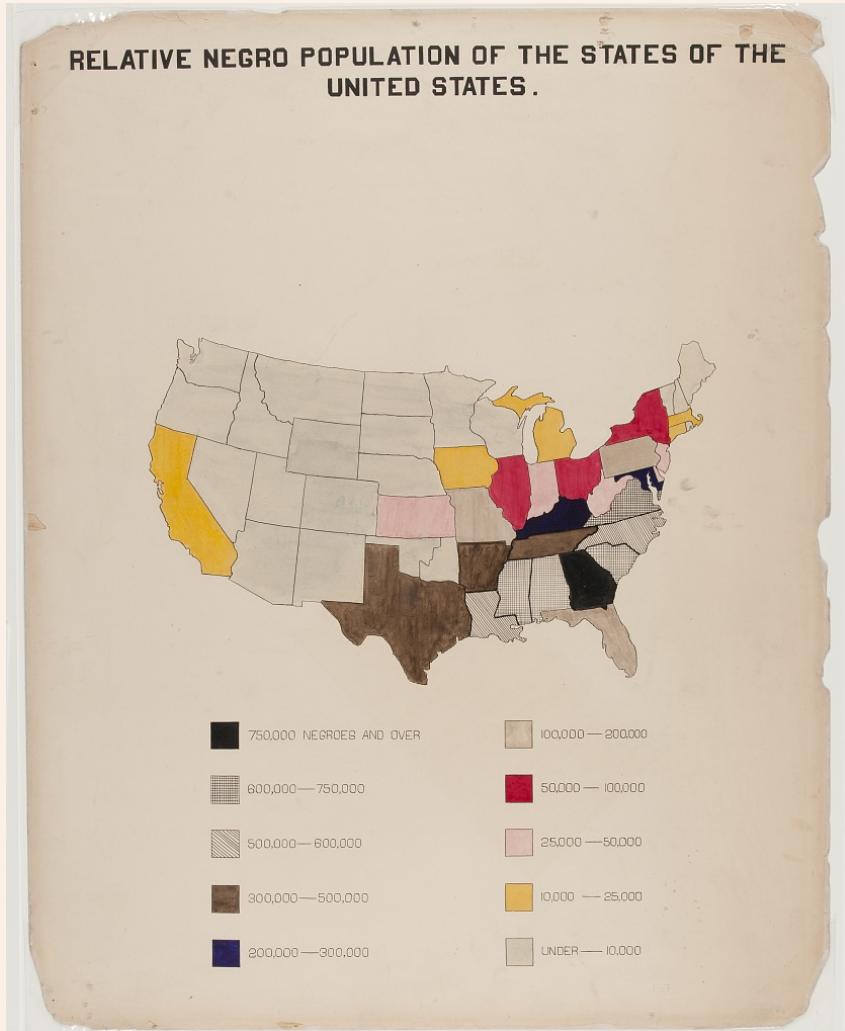
Explore Example 3



Relative Negro Population of the States of the United States



Explore: Message



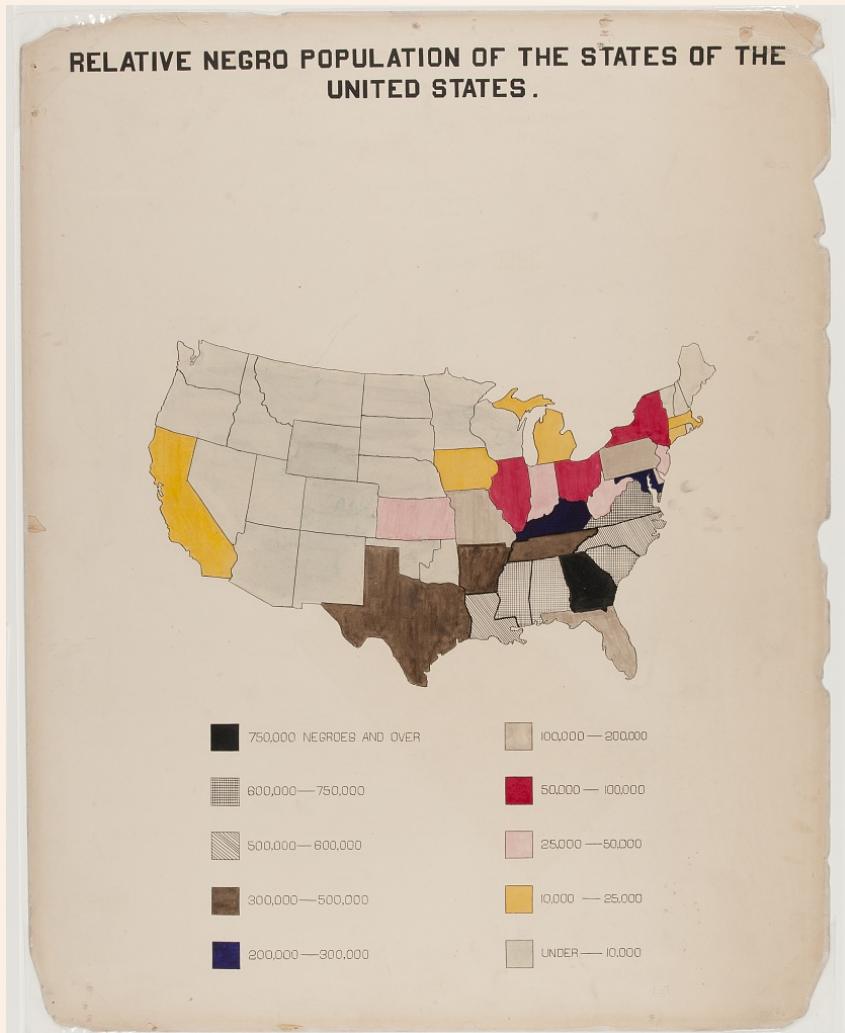
The message of the visual is tied to the overall message from Du Bois at the Paris Exposition:

Describe the status and progress of Black Americans since emancipation, showing the world their place in the US.

The visual shows the population of Black Americans across the United States, with larger populations concentrated in the South, while still showing the distribution throughout the states



Explore: Analytic Methods



This section addresses the effectiveness and analytical methods used in the visual.

Is the choropleth an effective method of invoking the message?

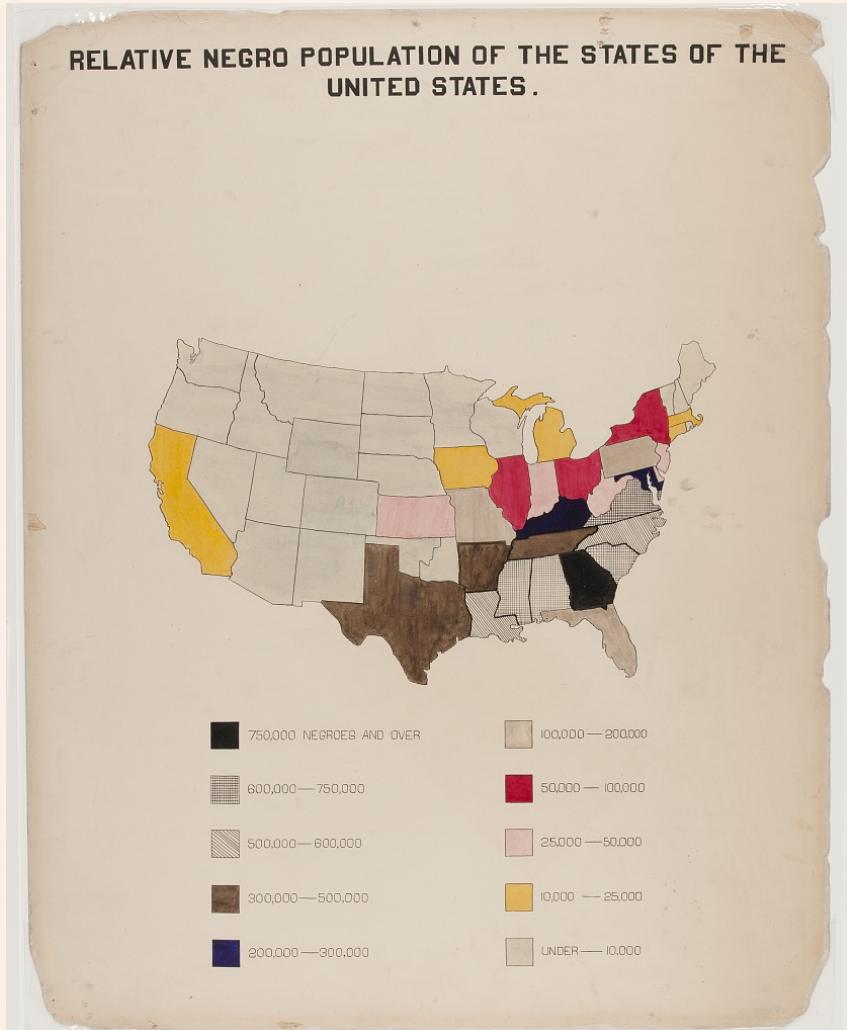
Can the viewer make appropriate comparisons?

What conclusions are drawn from the visual?

What additional questions are raised?



Explore: Aesthetics and Design



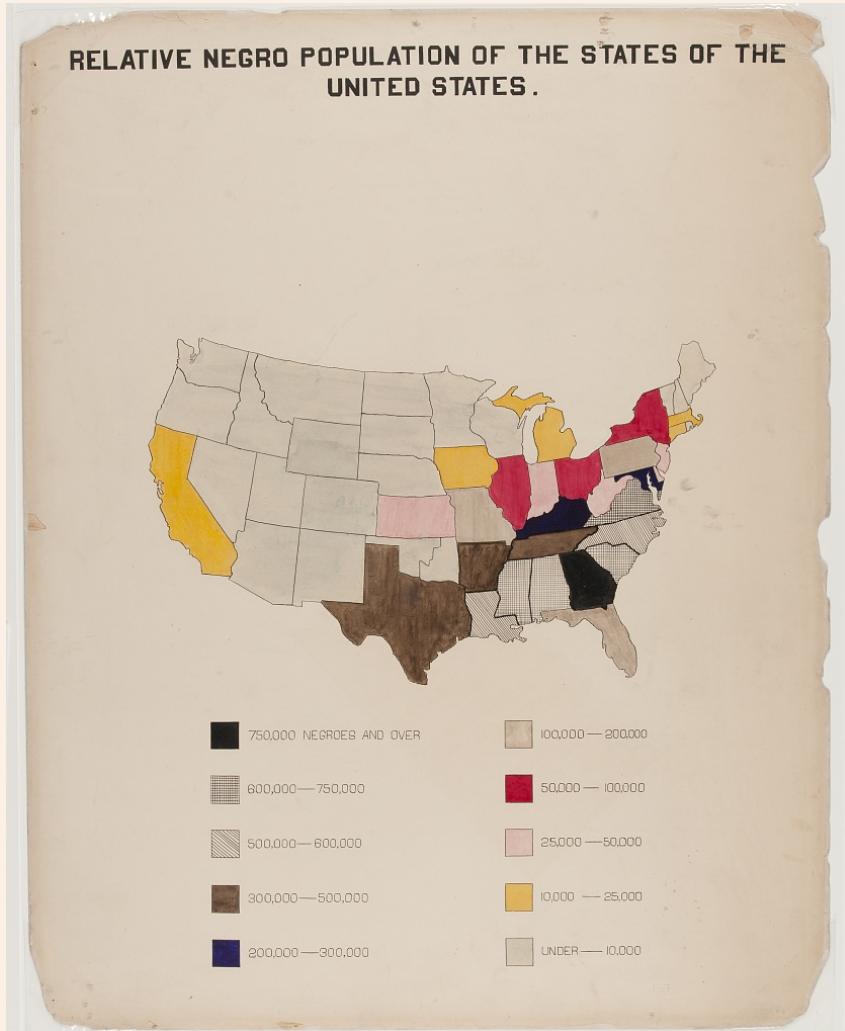
Title

Choropleth Map

Legend



Explore: Aesthetics and Design



This visual has three sections: title, choropleth map, and legend, uses a combination of color (black, brown, tan, red, pink, gold) and cross-hatching to encode ten categories of population data.

The solid and cross-hatched black is used to represent the higher populations based in the South, red and pink encode the mid-level populations, gold used for lower populations. A neutral color is used to states with a very small Black population.

Note the use of ample white space to allow the eye to focus on the map.



Explore: Re-Create

RELATIVE NEGRO POPULATION OF THE STATES OF THE UNITED STATES.



- | | |
|---|--|
| <input type="checkbox"/> 750,000 NEGROES AND OVER | <input type="checkbox"/> 100,000 - 200,000 |
| <input type="checkbox"/> 600,000 - 750,000 | <input type="checkbox"/> 50,000 - 100,000 |
| <input type="checkbox"/> 500,000 - 600,000 | <input type="checkbox"/> 25,000 - 50,000 |
| <input type="checkbox"/> 300,000 - 500,000 | <input type="checkbox"/> 10,000 - 25,000 |
| <input type="checkbox"/> 200,000 - 300,000 | <input type="checkbox"/> UNDER - 10,000 |

Re-create the visual using pens, paper and other analog materials like markers, colored pencils, rulers, and protractors.

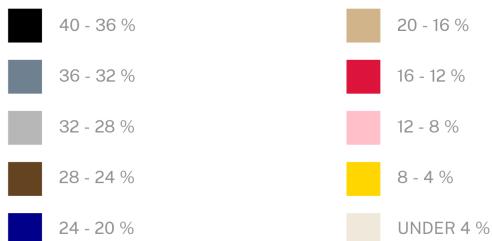
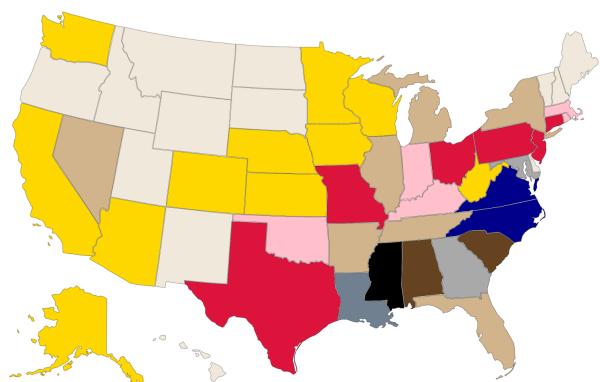
Strive for rapid iteration, and do not focus on precision, but rather exploring the various components of the chart, and how they are conveying the message and information. Students may work together or individually.

At the end of the session, the re-creations may be shared and critiqued as a group.



Explore: Modernize

PERCENT BLACK POPULATION OF THE UNITED STATES. (2021)

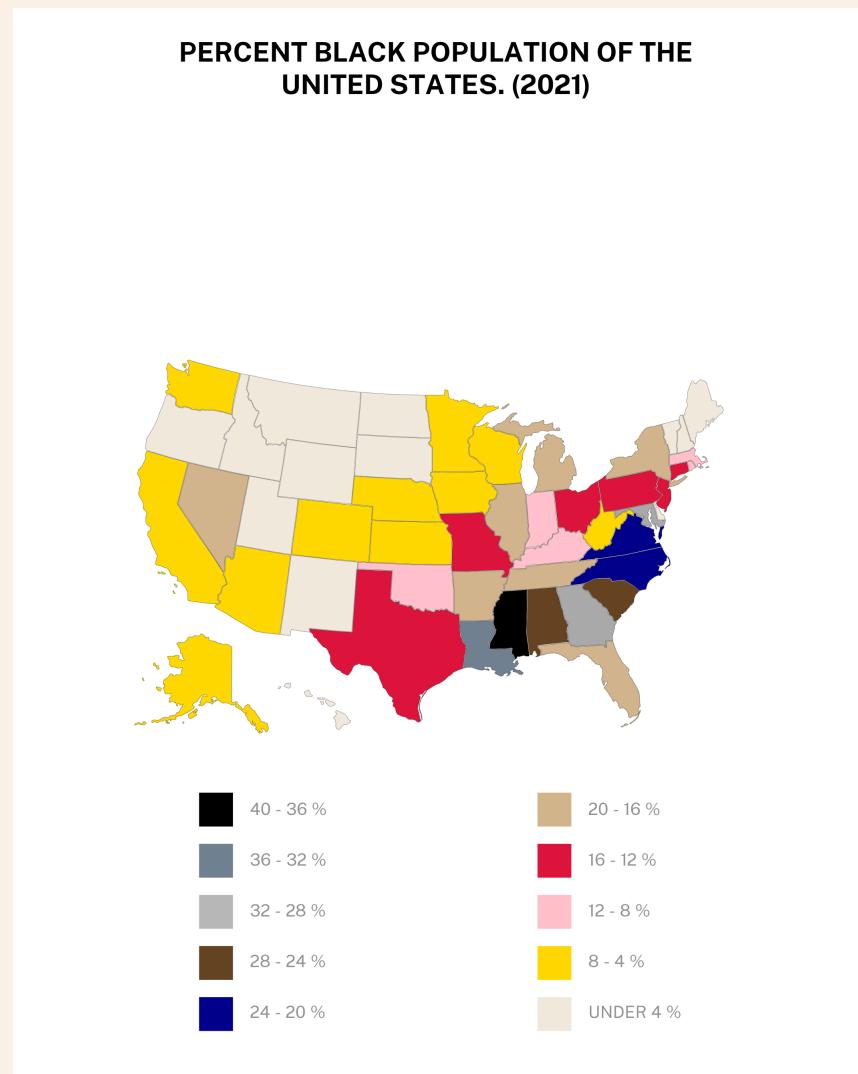
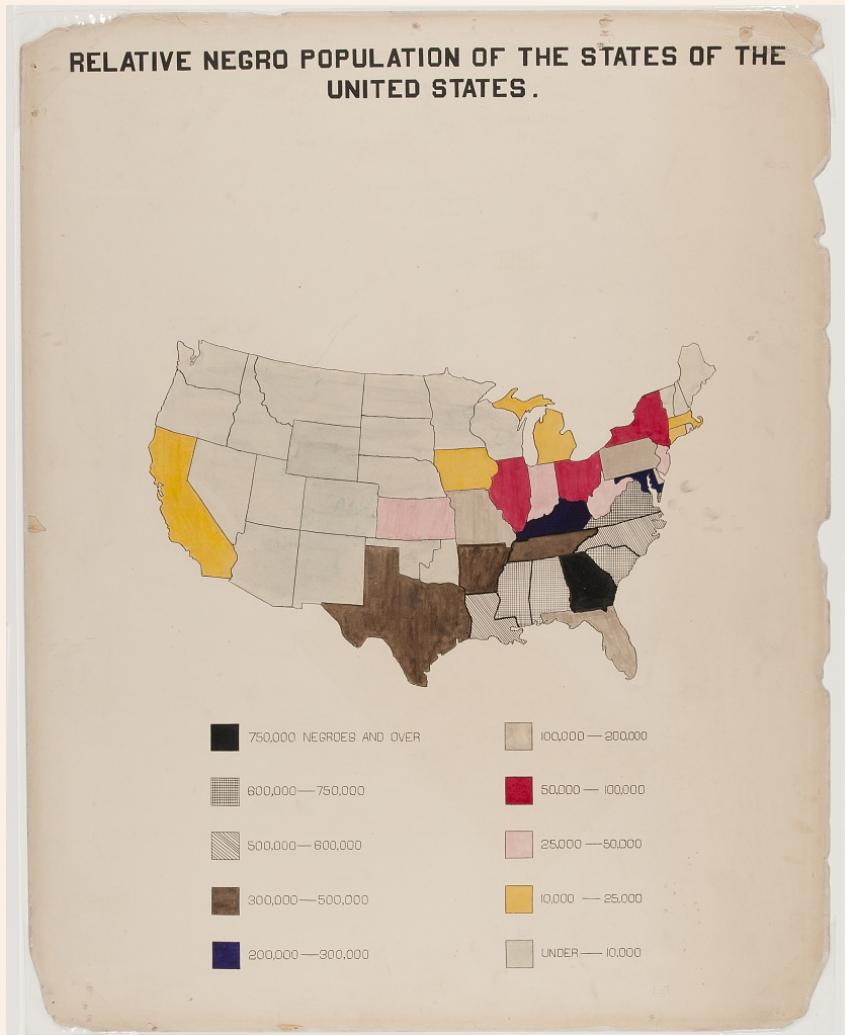


This example uses modern data to update the population of Black Americans to 2021.

It mimics the layout and color scheme of the original, but uses uses percentages instead of numbers, retaining the ten categories.



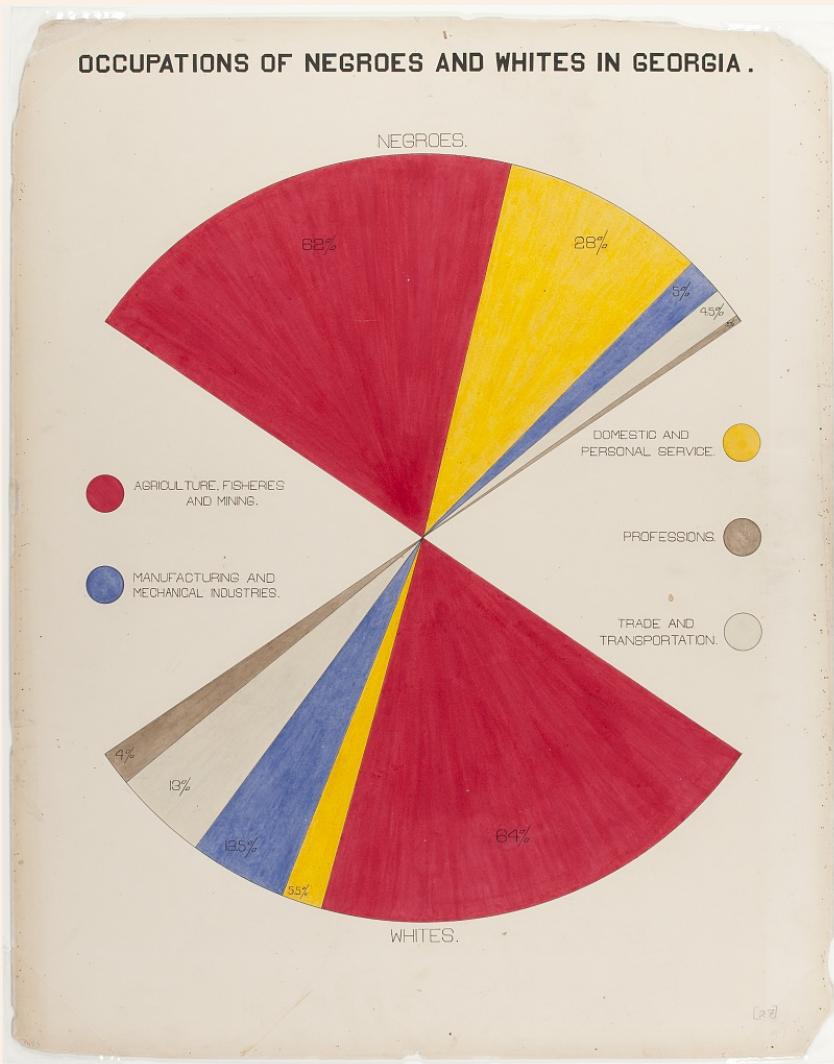
Explore: Modernize



Original and Modern Data Versions Compared



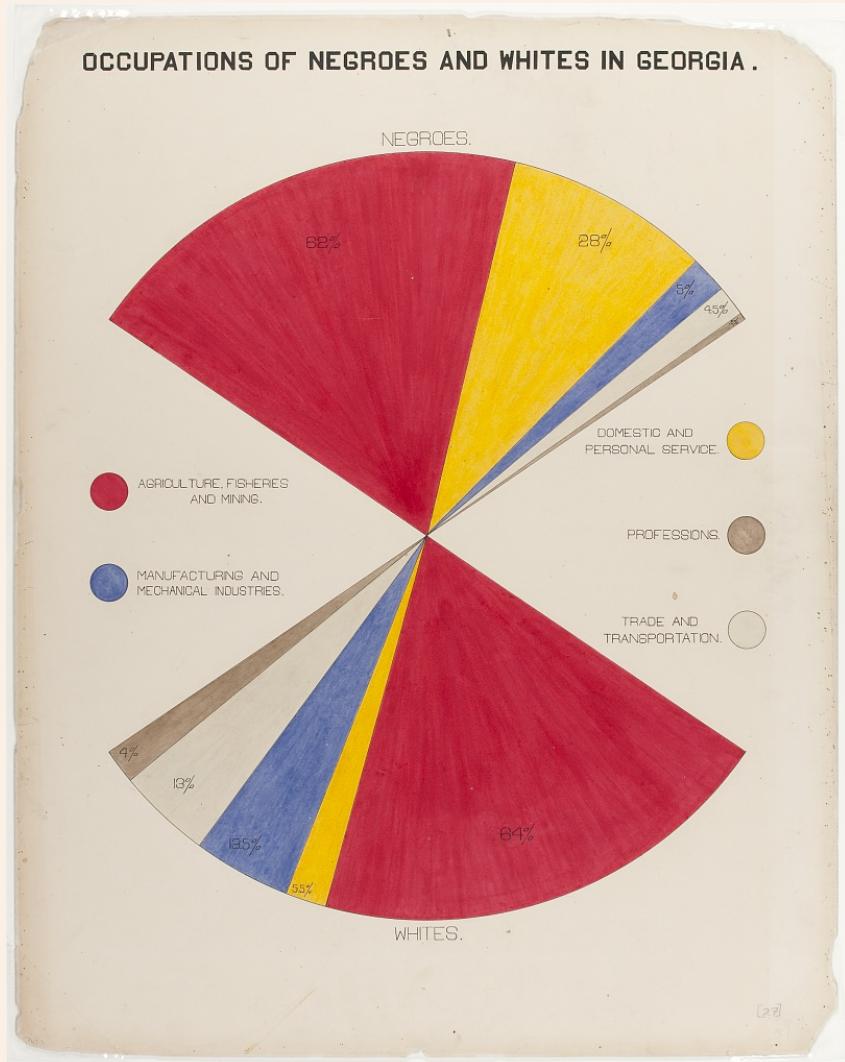
Explore Example 4



Occupations of Negroes and Whites in Georgia



Explore: Message



The message of the visual is tied to the overall message from Du Bois at the Paris Exposition:

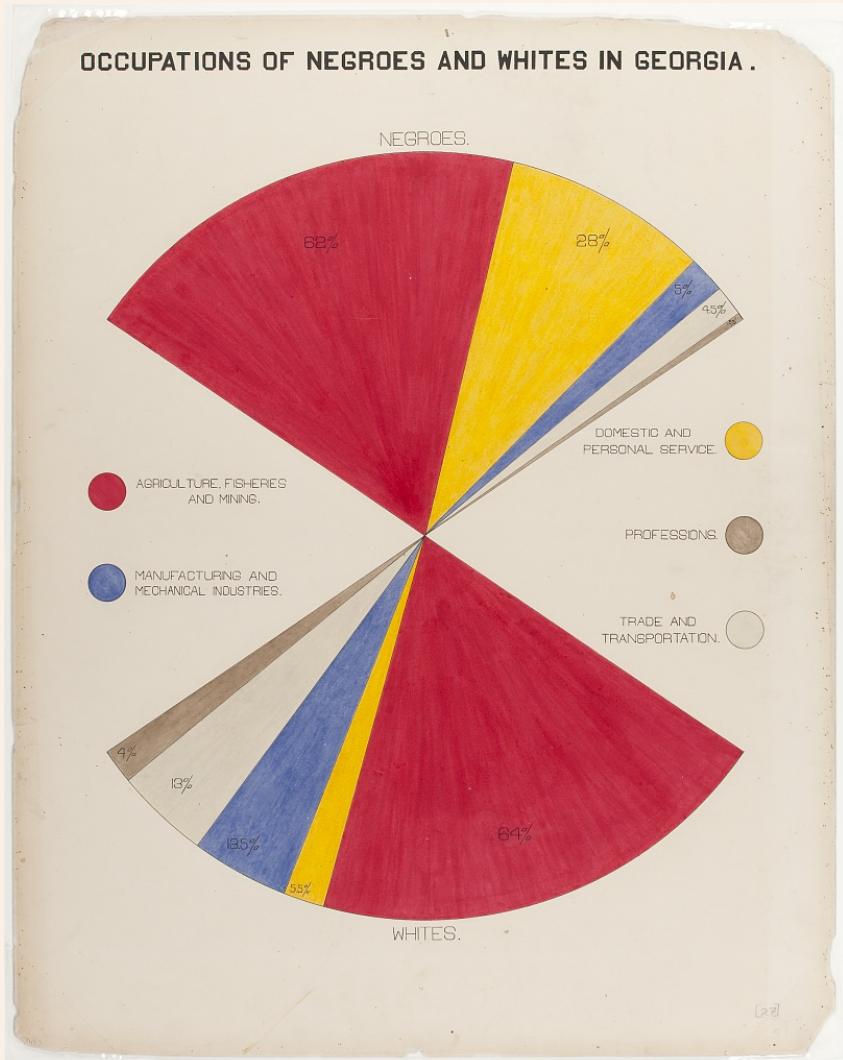
Describe the status and progress of Black Americans since emancipation.

The visual compares the occupations of Black and white populations of Georgia, showing parity for some occupations, while highlighting differences for others.

Activity: Discuss the message that the visual is attempting to convey



Explore: Analytical Methods

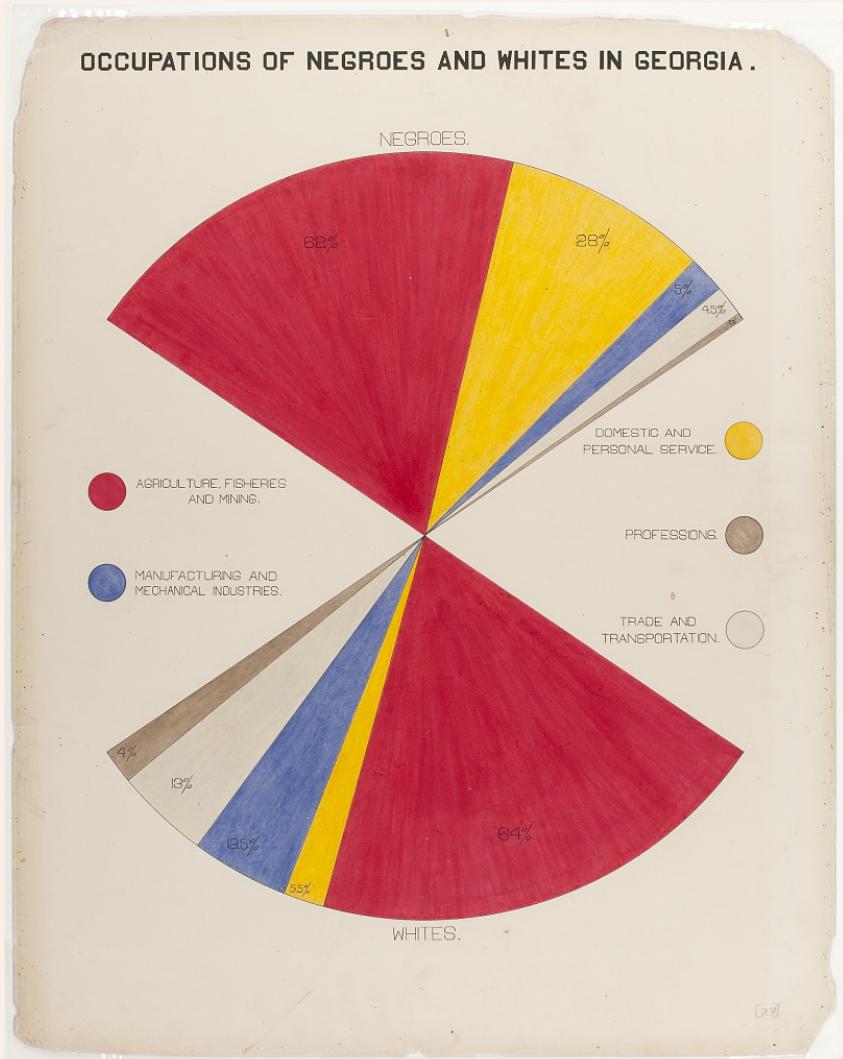


This section address the effectiveness and analytical methods used in the visual.

- Is the “fan chart” an effective?
- Can the viewer make comparisons?
- What conclusions are drawn?
- What additional questions are raised?



Explore: Aesthetics and Design



The “Fan-Chart” visual uses proportionally sized wedges to encode the data. The populations are arranged, in sections, top (Black), and bottom (white) such that the wedges may be easily compared.

Also note the legends arranged on the left and right, placed in the gaps between the wedges.

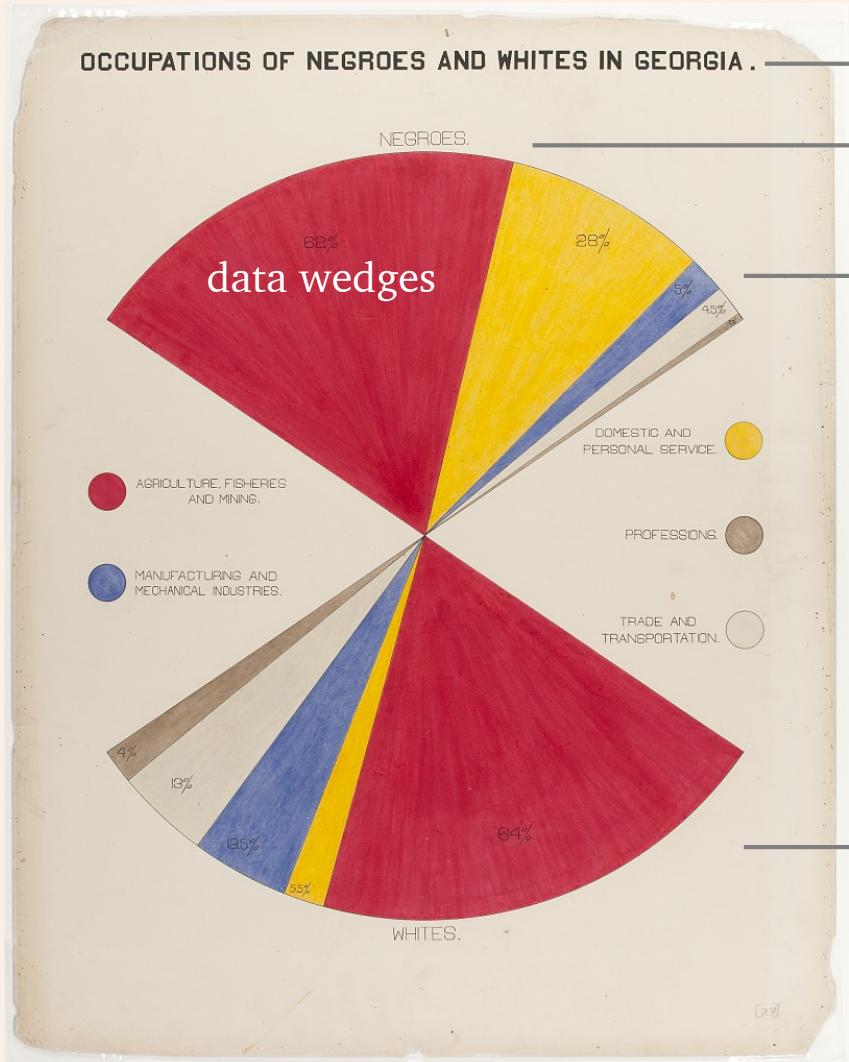
The visual uses the typical Du Bois color palette.





Explore: Aesthetics and Design

22x28 inch portrait mode



Title

Labels

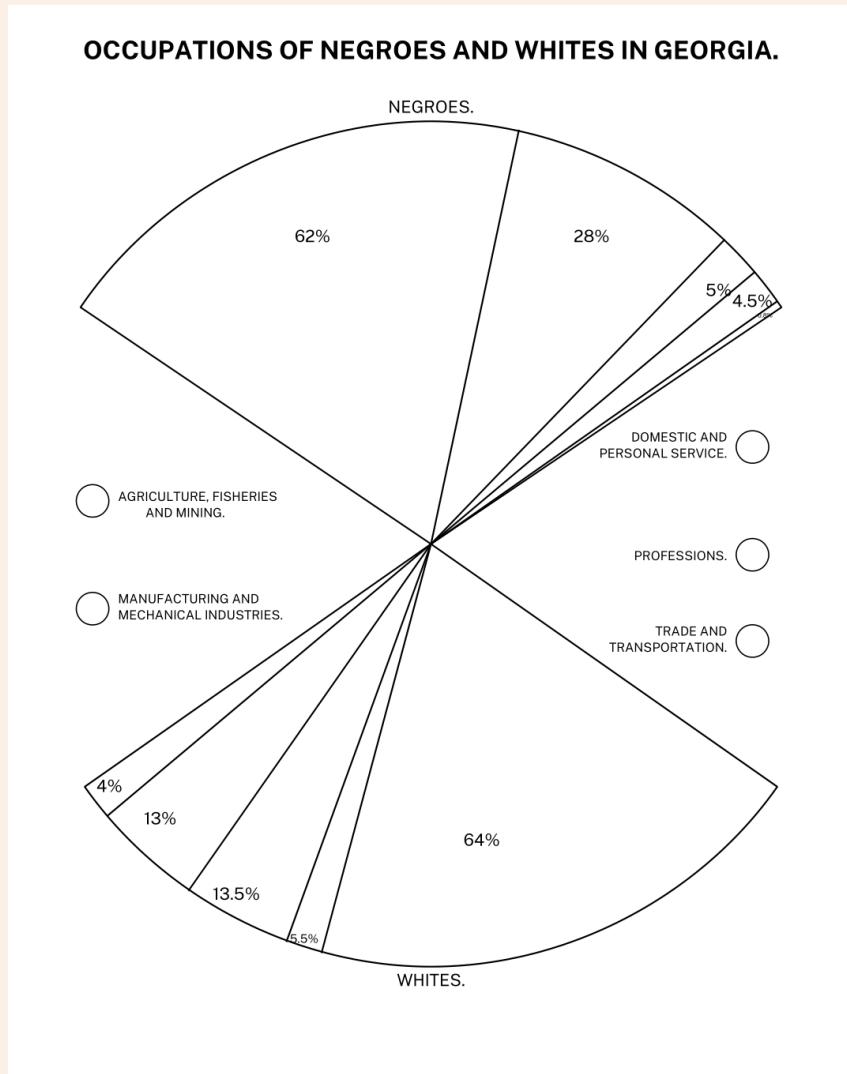
Black Georgians

Legend

White Georgians



Explore: Recreate



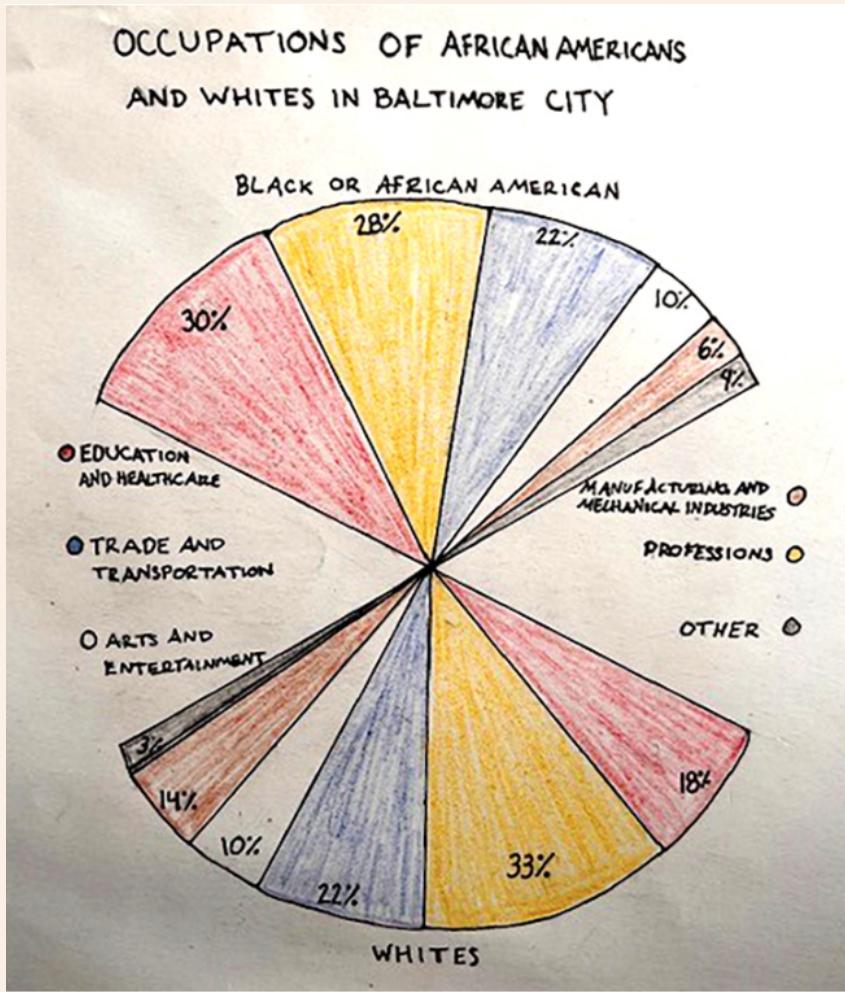
Re-create the visual using pens, paper and other analog materials like markers, colored pencils, rulers, and protractors.

Strive for rapid iteration, and do not focus on precision, but rather exploring the various components of the chart, and how they are conveying the message and information. Students may work together or individually.

At the end of the session, the re-creations may be shared and critiqued as a group.



Explore: Modernize



In this section, students may apply modern data to the Du Bois Style. This involves data gathering, selection of appropriate chart type, and defining the elements to apply the Du Bois style to (for example color palette, labeling methods)

This example shows modern data on the occupations of African-Americans and Whites in the city of Baltimore.



Explore: References

Du Bois Visualizations for Consideration in STEM Education

<https://github.com/ajstarks/dubois-data-portraits/tree/master/dubois-stem>

Du Bois Visualization Style Guide

<https://github.com/ajstarks/dubois-data-portraits/blob/master/dubois-style.pdf>

Recreating the Du Bois Data Portraits

<https://speakerdeck.com/ajstarks/recreating-the-dubois-data-portraits>

Du Bois Visualizations Catalog (originals)

<https://speakerdeck.com/ajstarks/du-bois-visualizations-originals>

African American Photographs Assembled for 1900 Paris Exposition

https://www.loc.gov/collections/african-american-photographs-1900-paris-exposition/?st=grid&sb=date_desc

W.E.B. Du Bois Data Portraits, Visualizing Black America

<https://papress.com/products/w-e-b-du-boiss-data-portraits-visualizing-black-america>

Black Lives 1900, W.E.B. Du Bois at the Paris Exposition

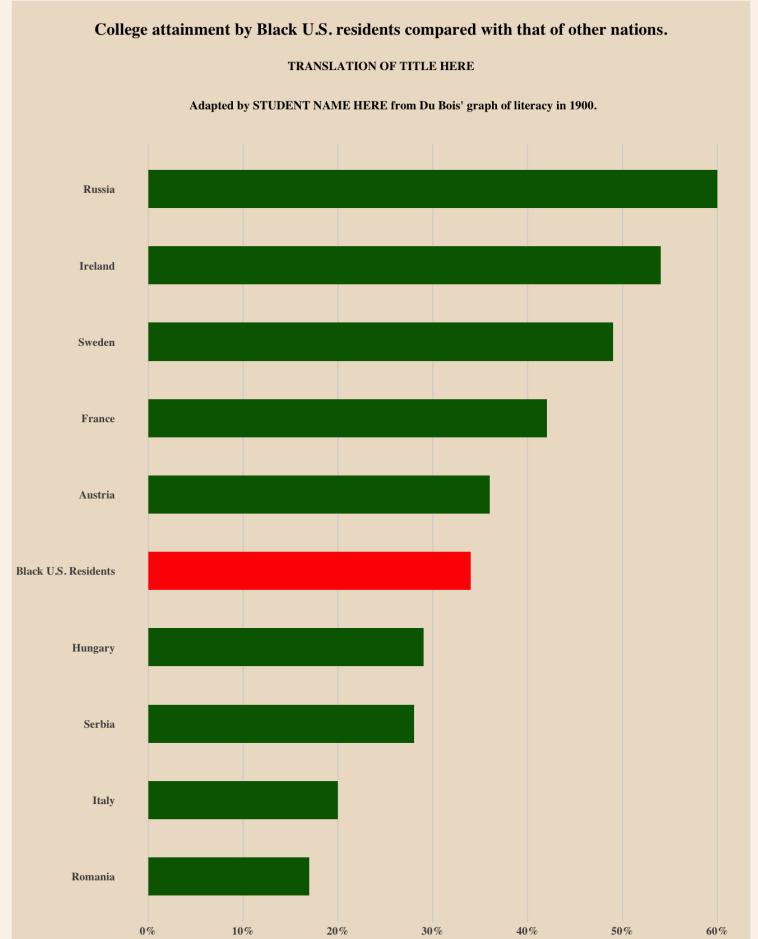
<https://www.theredstoneshop.com/products/black-lives-1900>



Implement

```
# below is code you learned above for reading in the data and setting up ggplot and graphing options
d_college_country<- read.csv("data/d_college_country.csv")
library(ggplot2)
options(repr.plot.width=22/2, repr.plot.height=28/2)
source("https://raw.githubusercontent.com/HigherEdData/Du-Bois-STEM/refs/heads/main/theme_dubois.R")

# below is the code you learned above
ggplot(d_college_country, aes(
  x = college,
  y = reorder(country, college),
  fill = country == "Black U.S. Residents" # this is the fill statement
)) +
  geom_col(width = .5) +
  theme_dubois() +
  theme(text = element_text('serif')) +
  scale_fill_manual(values = c("TRUE" = "red", "FALSE" = "darkgreen")) +
  labs(
    title = "\nCollege attainment by Black U.S. residents compared with that of other nations.\n",
    # fill in the blank below to translate the title in the language of your choice
    subtitle = "TRANSLATION OF TITLE HERE\n\n"
    Adapted by STUDENT NAME HERE from Du Bois' graph of literacy in 1900.\n\n"
  ) +
  scale_x_continuous(
    breaks = seq(0, 60, by = 10), # Set tick positions every 10 units
    labels = function(x) paste0(x, "%") # Add a "%" symbol to each label
  ) +
  theme(
    axis.text.x = element_text(size = 12),
    panel.grid.major.x = element_line(color = "lightgray")
  )
```





Implement

The Implement section builds on the Context and Explore sections and guides you through building a Du Bois-style visualization using digital tools.

The example shown here uses R and Jupyter notebook, which provides an interactive environment to explore the visual, starting from a basic graph, to applying the Du Bois style.

Note that participants of the Du Bois Challenge, an annual online exercise that re-creates a curated set of Du Bois visuals, provides examples of digital re-creations using tools such as R, python/matplotlib, and Tableau.

Implement: Advantages of Digital



Scalability:

The ability to work with increasing amounts of data,

Reproducibility and Adaptability:

you can reproduce and adapt graphs with less work than when we draw by hand,

Reach:

works are available to millions on the Internet; others are free to adapt, share and remix without having to reinvent the wheel.



Implement: creating a bar chart

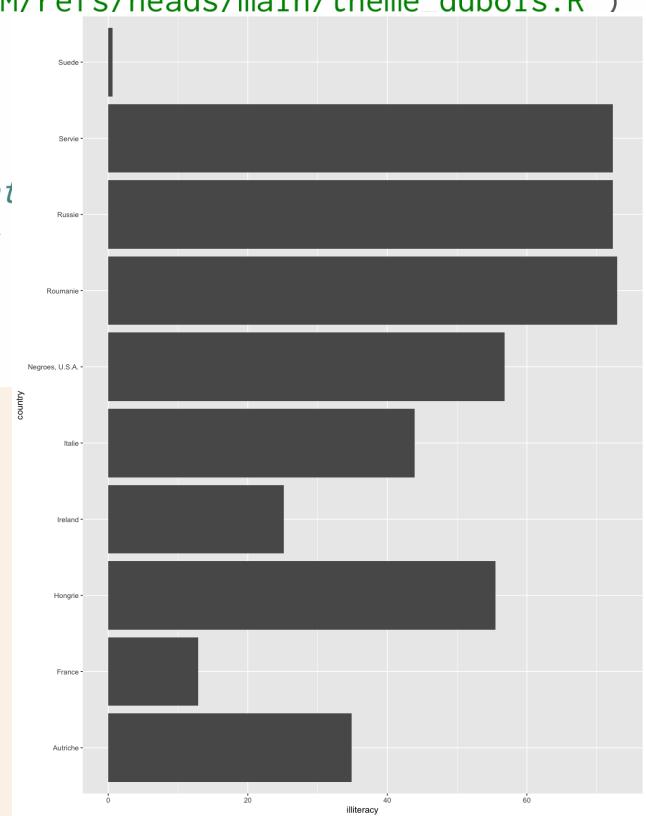
```
# Read in the data from a CSV file into a dataframe called d_literacy_country
d_literacy_country <- read.csv("data/d_literacy_country.csv")

# Load the ggplot2 package for creating visualizations
library(ggplot2)

# Set the dimensions of the plot output (width and height in inches)
options(repr.plot.width = 22/2, repr.plot.height = 28/2)

# Source a custom ggplot theme from a remote URL to style the plot (theme_dubois)
source("https://raw.githubusercontent.com/HigherEdData/Du-Bois-STEM/refs/heads/main/theme_dubois.R")

# Begin the ggplot call with the data and aesthetic mappings
ggplot(d_literacy_country, aes(
  x = illiteracy, # replace _____ with the variable name that
  y = country)) + # replace _____ with the variable name that
  # graph a bar chart based on the x and y mappings above
  geom_col()
```

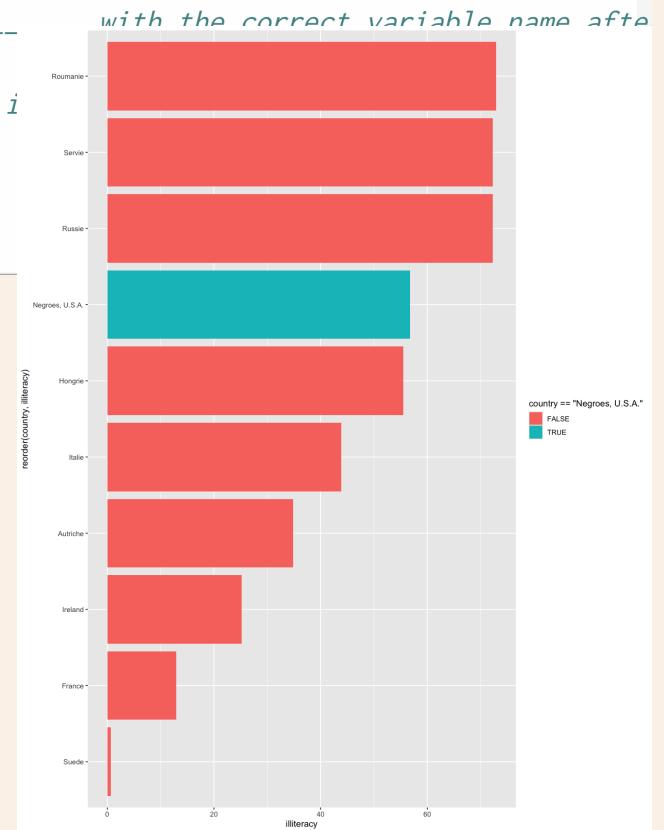




Implement: ordering bars, use spot color

```
# below is code you learned above for reading in the data and setting up ggplot and graphing options
d_literacy_country <- read.csv("data/d_literacy_country.csv")
library(ggplot2)
options(repr.plot.width=22/2, repr.plot.height=28/2)
source("https://raw.githubusercontent.com/HigherEdData/Du-Bois-STEM/refs/heads/main/theme_dubois.R")

# below is the code you learned above to graph Du Bois data
ggplot(d_literacy_country, aes(
  x = illiteracy,
  # then by adding "reorder" below, you reorder the countries by filling in _____
  y = reorder(country, illiteracy),
  # then you can tell R which country to fill with a different color by filling in _____
  fill = country == "Negroes, U.S.A."
)) +
  geom_col()
```

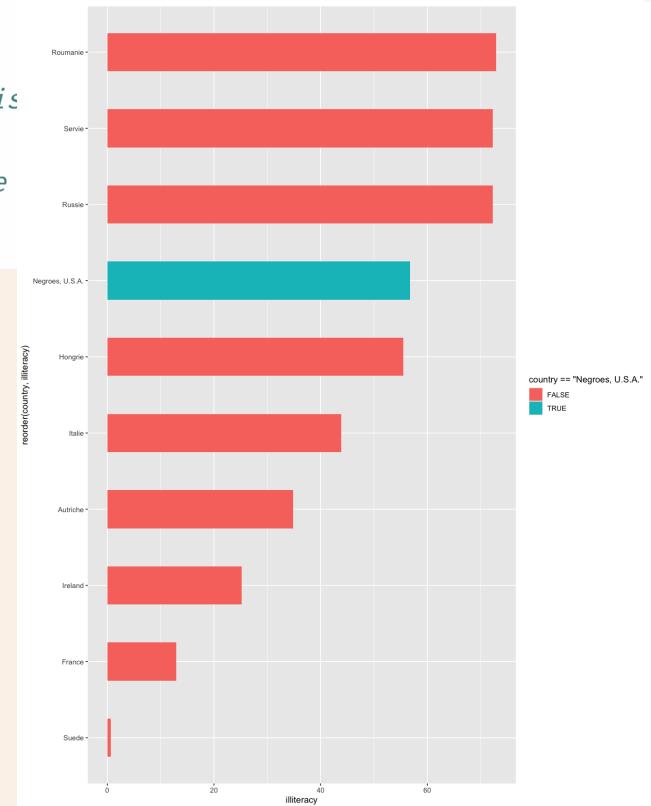


Implement: use the Du Bois theme



```
# below is code you learned above for reading in the data and setting up ggplot and graphing options
d_literacy_country <- read.csv("data/d_literacy_country.csv")
library(ggplot2)
options(repr.plot.width=22/2, repr.plot.height=28/2)
source("https://raw.githubusercontent.com/HigherEdData/Du-Bois-STEM/refs/heads/main/theme_dubois.R")

# below is the code you learned above to graph Du Bois data with sorting and different colors
ggplot(d_literacy_country, aes(
  x = illiteracy,
  y = reorder(country, illiteracy),
  fill = country == "Negroes, U.S.A."
)) +
# fill in the ____ below to adjust the bar widths to be more similar to Du Bois
  geom_col(width = .5)
# add a plus sign above this comment to add a new line of code. Then write the
```

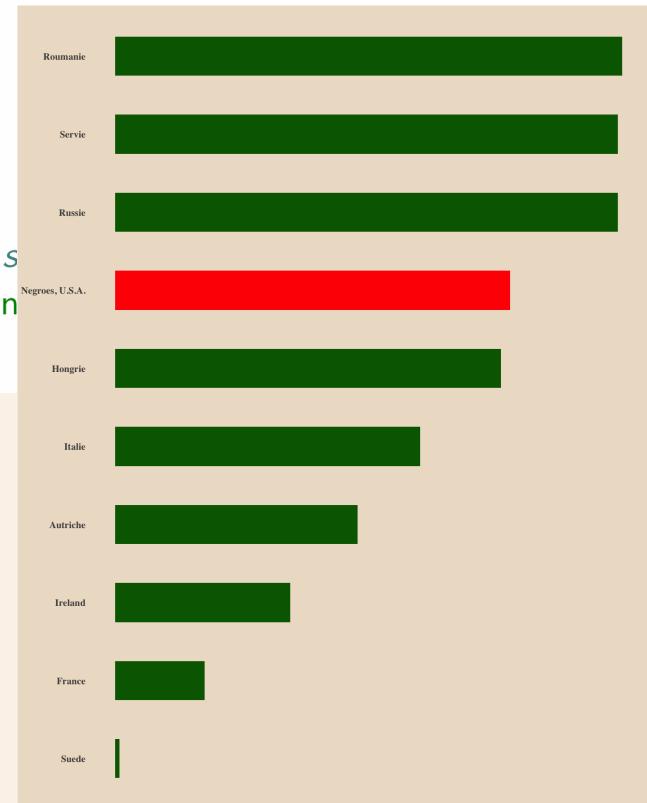




Implement: change text font and color bar

```
# below is code you learned above for reading in the data and setting up ggplot and graphing options
d_literacy_country <- read.csv("data/d_literacy_country.csv")
library(ggplot2)
options(repr.plot.width=22/2, repr.plot.height=28/2)
source("https://raw.githubusercontent.com/HigherEdData/Du-Bois-STEM/refs/heads/main/theme_dubois.R")

# below is the code you learned above
ggplot(d_literacy_country, aes(
  x = illiteracy,
  y = reorder(country, illiteracy),
  fill = country == "Negroes, U.S.A." # this is the fill statement
)) +
  geom_col(width = .5) +
  theme_dubois() +
#change the themes's font by filling in the blank
  theme(text = element_text('serif')) +
# fill in the blank for which color the bar should be when the fill s
  scale_fill_manual(values = c("TRUE" = "red", "FALSE" = "darkgreen"))
```

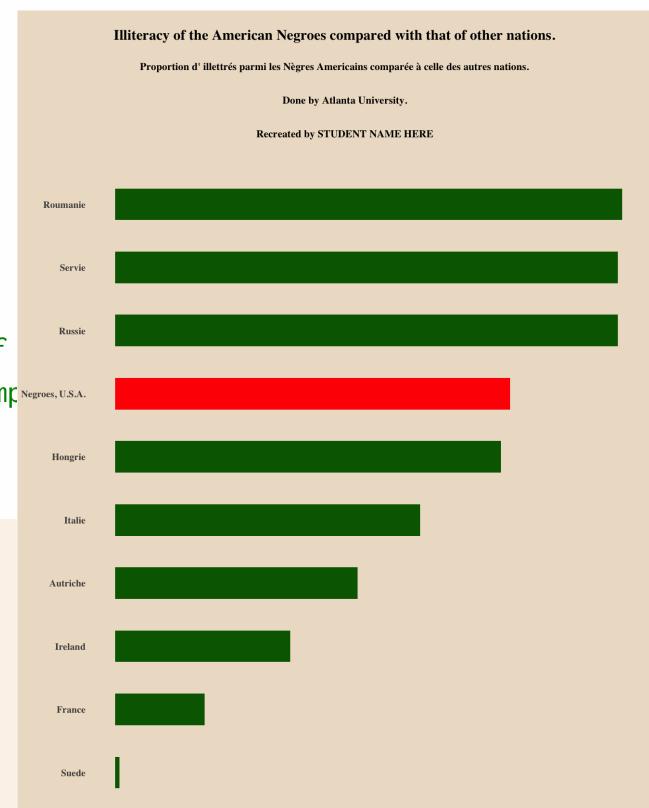




Implement: titles and subtitles, your name

```
# below is code you learned above for reading in the data and setting up ggplot and graphing options
d_literacy_country <- read.csv("data/d_literacy_country.csv")
library(ggplot2)
options(repr.plot.width=22/2, repr.plot.height=28/2)
source("https://raw.githubusercontent.com/HigherEdData/Du-Bois-STEM/refs/heads/main/theme_dubois.R")

# below is the code you learned above
ggplot(d_literacy_country, aes(
  x = illiteracy,
  y = reorder(country, illiteracy),
  fill = country == "Negroes, U.S.A." # this is the fill statement
)) +
  geom_col(width = .5) +
  theme_dubois() +
  theme(text = element_text('serif')) +
  scale_fill_manual(values = c("TRUE" = "red", "FALSE" = "darkgreen")) +
# YOU LEARNED ABOVE ABOUT ALL THE CODE ABOVE THIS LINE
# the code below adds a title and a subtitle
  labs(
    title = "\nIlliteracy of the American Negroes compared with that of",
    subtitle = "Proportion d' illettrés parmi les Nègres Americains comp",
    Done by Atlanta University.\n\n",
    Recreated by STUDENT NAME HERE\n\n"
  )
```

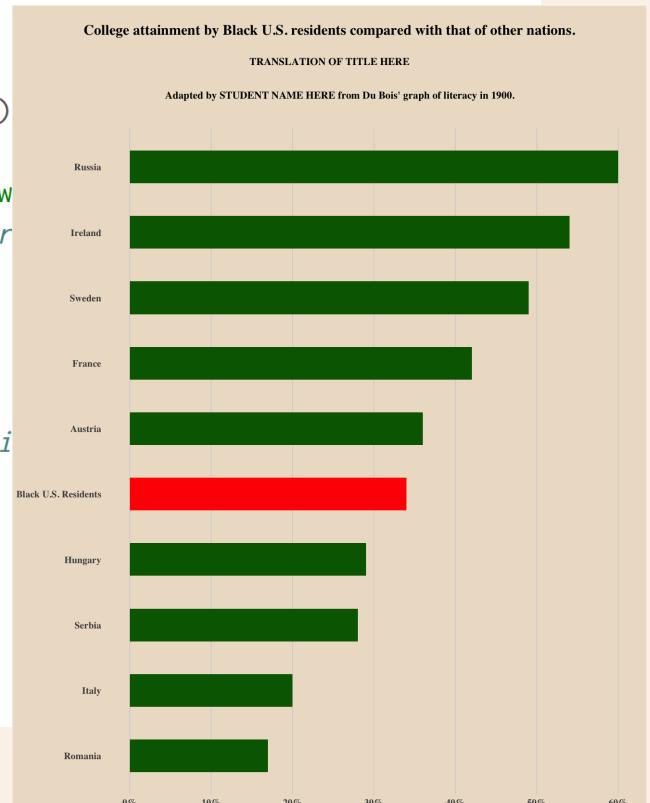




Implement: use modern data, update legend

```
# below is code you learned above for reading in the data and setting up ggplot and graphing options
d_college_country<- read.csv("data/d_college_country.csv")
library(ggplot2)
options(repr.plot.width=22/2, repr.plot.height=28/2)
source("https://raw.githubusercontent.com/HigherEdData/Du-Bois-STEM/refs/heads/main/theme_dubois.R")

# below is the code you learned above
ggplot(d_college_country, aes(
  x = college,
  y = reorder(country, college),
  fill = country == "Black U.S. Residents" # this is the fill statement
)) +
  geom_col(width = .5) +
  theme_dubois() +
  theme(text = element_text('serif')) +
  scale_fill_manual(values = c("TRUE" = "red", "FALSE" = "darkgreen"))
  labs(
    title = "\nCollege attainment by Black U.S. residents compared w
# fill in the blank below to translate the title in the language of your
    subtitle = "TRANSLATION OF TITLE HERE\n\n"
    Adapted by STUDENT NAME HERE from Du Bois' graph of literacy in
  ) +
  scale_x_continuous(
    breaks = seq(0, 60, by = 10), # Set tick positions every 10 uni
    labels = function(x) paste0(x, "%") # Add a "%" symbol to each
  ) +
  theme(
    axis.text.x = element_text(size = 12),
    panel.grid.major.x = element_line(color = "lightgray")
  )
```





Implement: References

Github Repository for the #DuBoisChallenge2024

<https://github.com/ajstarks/dubois-data-portraits/blob/master/challenge/2024/README.md>

Du Bois Challenge 2024 Recap

<https://speakerdeck.com/ajstarks/du-bois-challenge-2024-recap>

2024 Du Bois Challenge using R Programming.

<https://medium.com/illumination/2024-du-bois-challenge-using-r-programming-02af8afa5626>

Developing Du Bois's Data Portraits with Python and Matplotlib

<https://www.edriessen.com/2024/02/07/developing-du-boiss-data-portraits-with-python-and-matplotlib/>

Three Tricks I Learned In The Du Bois Data Visualization Challenge

<https://nightingaledvs.com/recreating-historical-dataviz-three-tricks-i-learned-in-the-du-bois-data-visualization-challenge/>

Molly Kuhs Du Bois Challenge repo

<https://github.com/makuhs/DuBoisChallenge>

#DuBoisChallenge2024 using Python and Matplotlib

<https://github.com/edriessen/dubois24-python-matplotlib>

#DuBoisChallenge2024 using R

<https://github.com/sndaba/2024DuBoisChallengeInRstats/tree/main>

#DuBoisChallenge2024 using Tableau

<https://public.tableau.com/app/profile/camaal.moten7357/vizzes>