

# Revisiting Du Bois

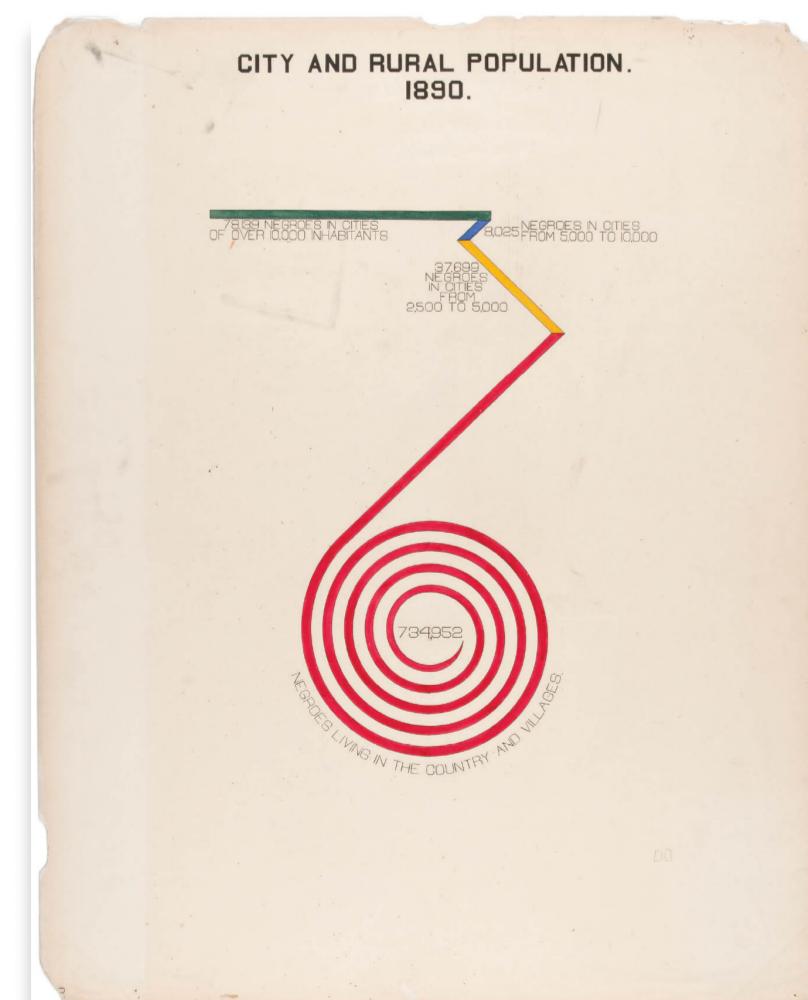
## A Data Visualization Project

### Intro

Prolific and prominent Black historian W.E.B. Du Bois has left a legend in many ways. In addition to being a sociologist who meticulously drew and recorded social statistics for over fifty years, Du Bois was also a Black theorist of knowledge, an educator, a novel writer, and a historian. His groundbreaking contributions to the humanities as a pioneer and radical theorist is undisputed. Less acknowledged are his perspectives on quantitative information and data visualization that were built on the same critical and creative ground as his perspectives in the humanities.

My project seeks to explore his work with a critical and quantitative lens. I hope to bring to Du Bois' portraits the worlds of modern statistics, computing resources, and clearly defined systems of visualization, but also to these fields bring Du Bois' humanist, expressive lens. To do so, I began development on the `ggdubois` graphics library for data visualization in the R programming language. In extending Du Bois' data visualization through practical tools, I hope to also convey his artistic perspective on data visualization.

Du Bois often stretches beyond the “grammar of graphics” and other systems, with multiple coordinate systems, “wrapping” visual items, and inflating geographic areas. These often address practical problems in data visualization, like how to display items that are much larger than others.



Du Bois used shapes like spirals to represent very large and smaller quantities on the same panel.



### The `ggdubois` package

`ggdubois` contributes seven geometries, or “geoms”:

1. `geom_scaledmap`, which represents a variable through the inflated or deflated area of a spatial feature on a map.
2. `geom_spike`, which represents quantities of a variable through rings that are connected by spikes.
3. `geom_spiral`, which represents quantities of a variable through a spiral shape.
4. `geom_pathspiral`, which represents quantities of a variable through a combination of a spiral and connected segments.
5. `geom_wrappedbar`, or a bar chart that wraps quantities onto the next row after they exceed a certain quantity.
6. `geom_wovenbar`, which interweaves two different bar charts on the same plane.
7. `geom_square`, which represents quantities of a variable through a filled square.

As well as two sets of accessories for graphics.

`theme_dubois()`: Handles typography, plot spacing, grid lines, and baseline colors. Two variants of “regular” and the more minimal “canvas” are available.

`dubois_pal`: Provides sequential and divergent color palettes, and interfaces to modify colors on `ggplot2` objects

### Data

This package provides two datasets which can be used to test and create graphics.

The “georgia” dataset: contains county-level information on race and ethnicity, home ownership, educational attainment, and geographic boundaries for Georgia, 1970-2010.

The “demographics” dataset: contains county-level information on air quality, rent burden, income inequality, education, unemployment, and child poverty at the county level for the United States, 2015 only.

Download at <http://github.com/18kimm/ggdubois/data>

by Nathan Kim  
Advised by  
Professor Elisa Celis

### Examples

#### geom\_spiral

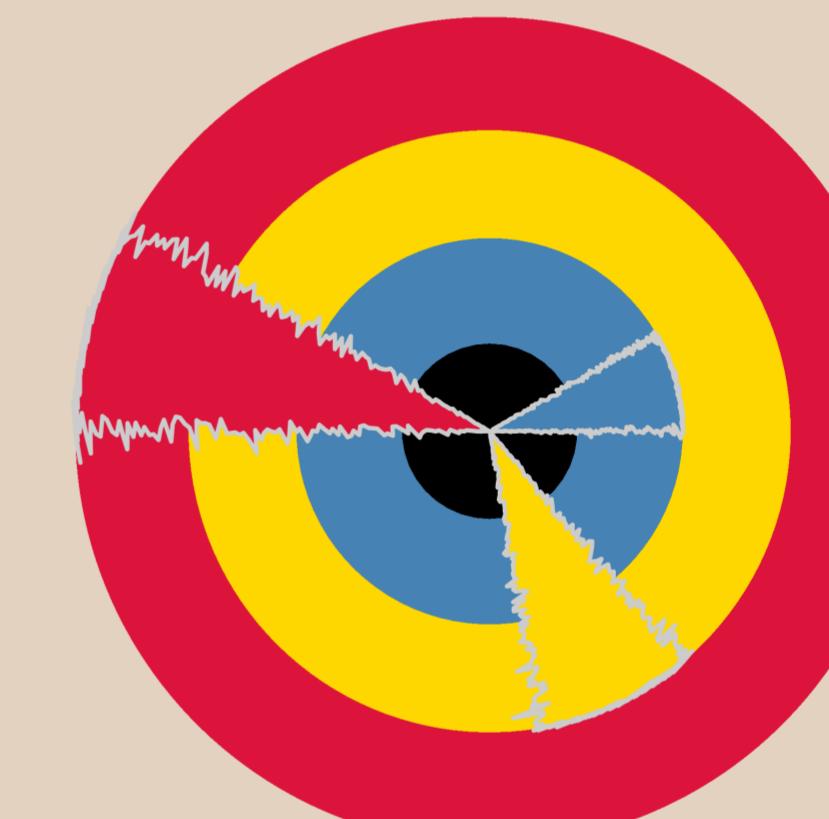
Median income in Georgia, 1980-2000.



year  
— 1980  
— 1990  
— 2000

#### geom\_spike

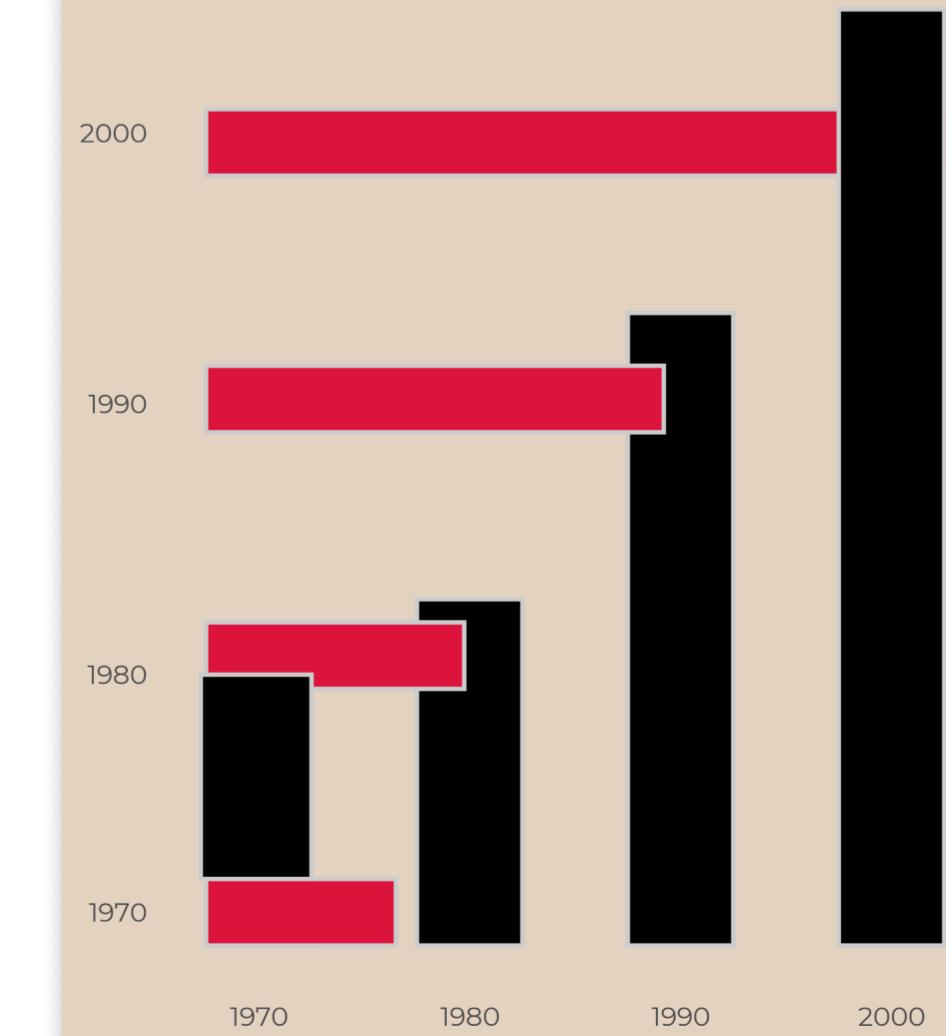
Homeownership in Georgia steadily increases



year  
— 1970  
— 1980  
— 1990  
— 2000

#### geom\_wovenbar

Income and payroll in Georgia, 1970 - 2000  
Length of bars is proportional to value



income  
payroll

Du Bois used shapes like spirals to represent very large and smaller quantities on the same panel.