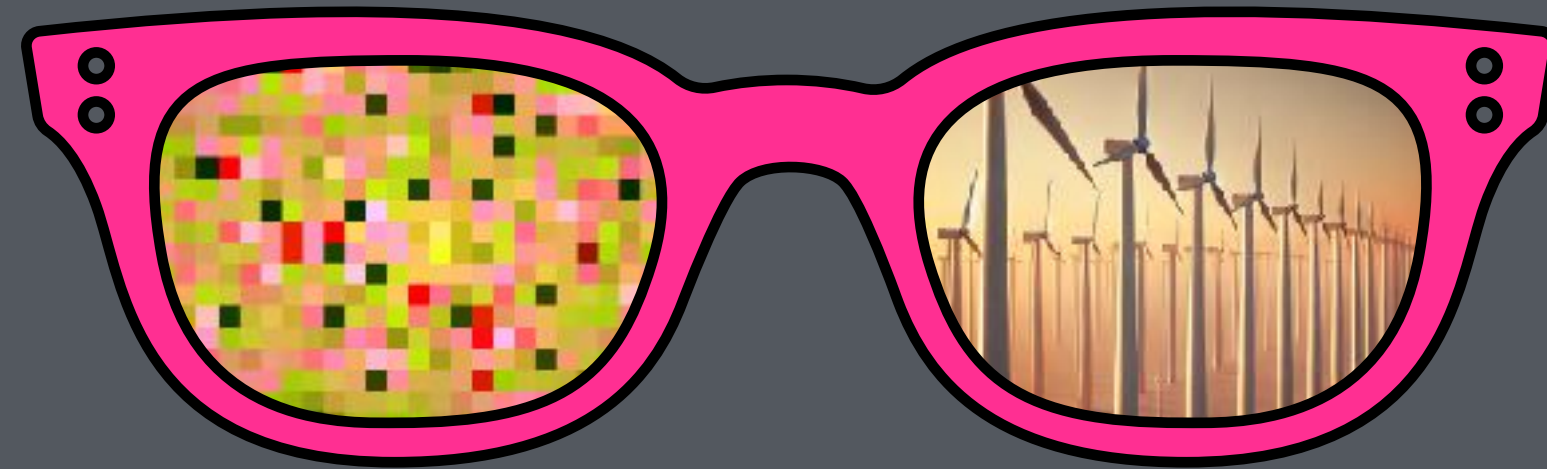


Data and the State

PUBPOL 2130 / INFO 3130



Censuses & tabular data

Lecture 2, Thursday Jan 23

Syllabus announcements

- **Missed class and section policy**

- Material covered in class is fair game for homework and exams
- 10% participation grade heavily depends on section attendance
- If you miss section, there will be an online form with more in-depth questions from notebook and class
- Currently working through accommodations requests, and a few informal notes on needs
- Reading already available in github (github.com/pubpol-2130)
- Notebook for section will be published Thursday nights, HW available after section

Geospatial tools

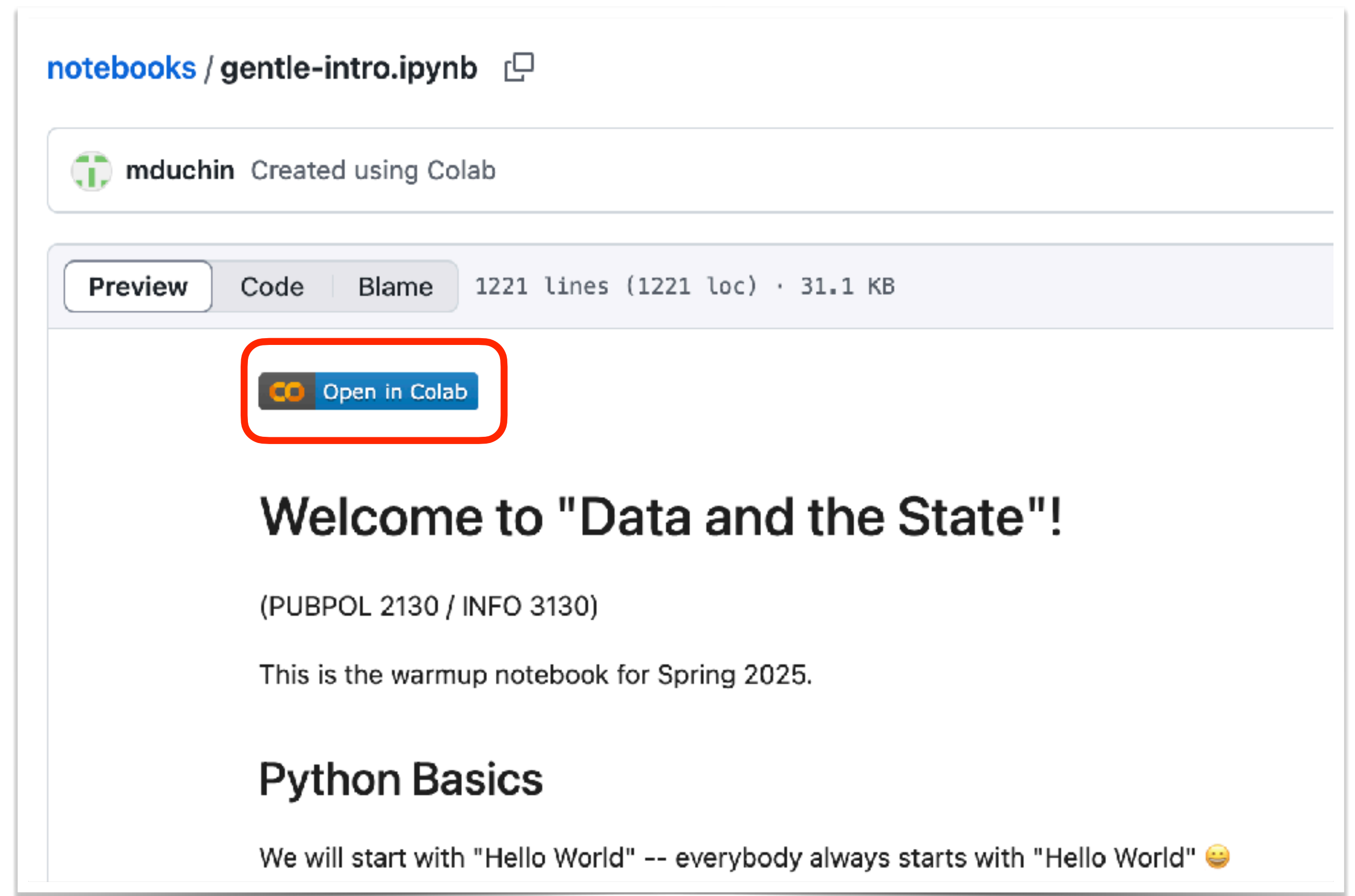
- **GIS** (geographic information systems)
 - vector/raster
- **Python** — human-readable language whose user community has positioned it to be central for applicable data science
- **Pandas** — de facto API for tabular data
- **Geopandas** — expects geometry column, supports map projection etc.
- Google **Colab** is a browser-based environment for executing code, e.g., Python
 - no fussing with installations

API - application
programming interface

wrapper

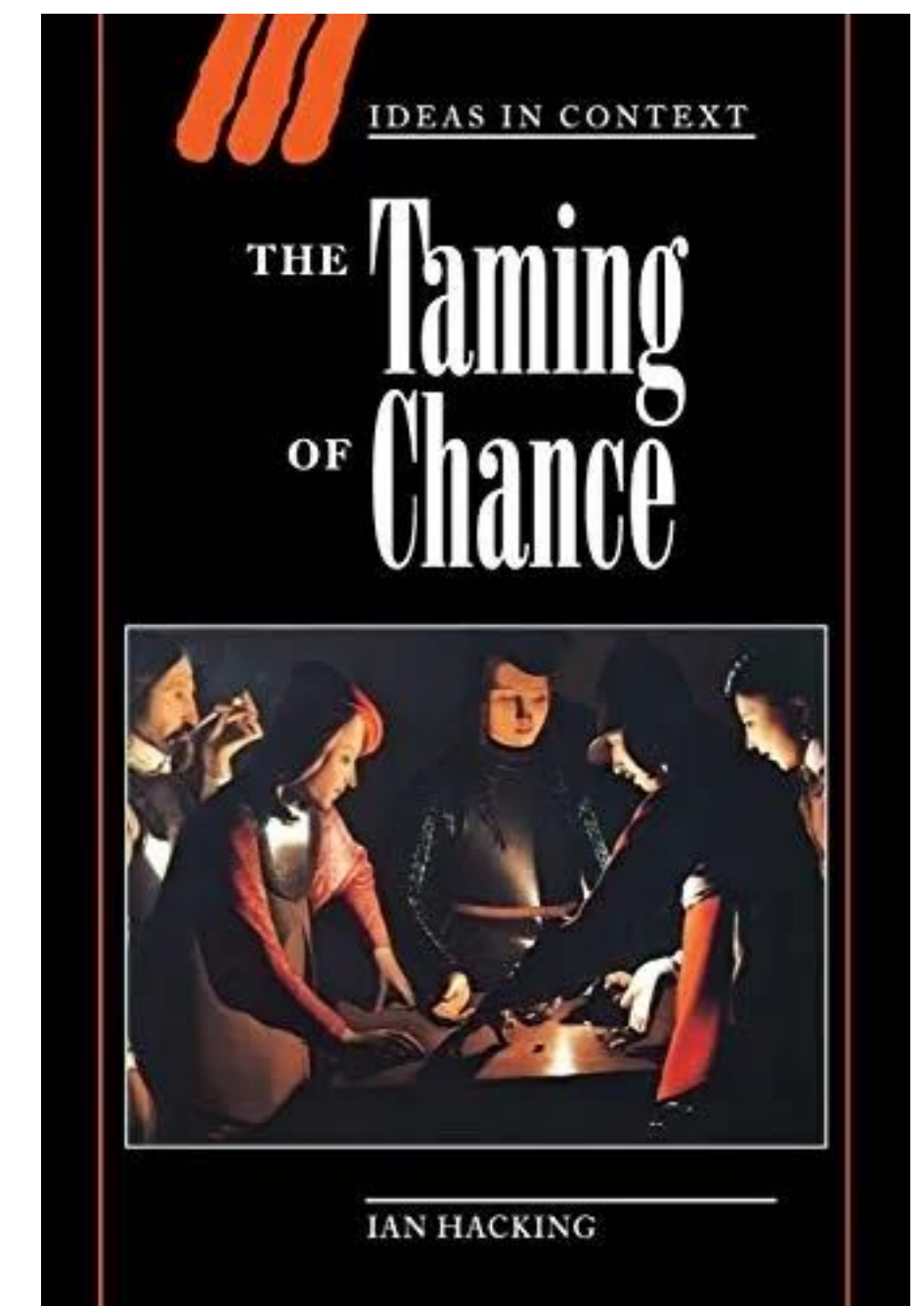
Let's take a look at a notebook

- our main place for stuff is in GitHub:
github.com/pubpol-2130
- folders for published **notebooks** and for **readings**
- each published notebook will have a Colab link
- this takes you to the browser-based environment where things are assembled in a self-contained way
- slow and clunky but highly replicable
- ChatGPT policy



Ian Hacking

- Ian Hacking, an interdisciplinary historian-philosopher, wrote about the history of censuses in his *The Taming of Chance*
- His basic argument: *probability* thinking happened to Europe fast, but *statistics* thinking happened slowly
- Probability thinking — chance, nondeterminism, free will, randomness, stochasticity
- Statistics thinking — human activity viewed in large **aggregates** is very **regular**, even though individuals are stochastic
- Example: marriage



Kant, circa 1800:

“deaths, births, and marriages, considering how much they are separately dependent on the freedom of the human will, should seem to be subject to no law according to which any calculation could be made beforehand of their amount; and yet the yearly registers of these events in great countries prove that they go on with as much conformity to the laws of nature as the oscillations of the weather.”

Normal and deviant

- this produces notions of the **normal** — normal distributions, normalcy, normal people, as against deviation and deviancy
- goals of social control: early European censuses were of **colonies** well before homelands
 - Spain/Peru, France/Quebec, England/Ireland, Europe/Caribbean — 16th/17th c
- Leibniz late 17th c: We need a central office for administration of military, civil, mining, forestry, and police affairs. Must maintain central register of deaths, baptisms, and marriages.



A statistical bureaucracy

- “statistics” as “political arithmetic” — the word statistic was introduced ~1750 by an obscure economist named Achenwall — “**remarkable facts about the state**”
- first collected by amateurs, only later professionalized by states
- “avalanche of printed numbers” after about 1820
- history of collecting official statistics is bound up with questions of objectivity and bureaucracy
 - one guys complaining about another guy in 1860: “I don’t fail to recognize the **diligence and loyalty** with which he has for so many years toiled for the Prussian state, but it would not be right to assign him the role of an **independently thinking** worker.” He is a tool, not a leader.

The most granular U.S. Census data

- It's the Redistricting Data File, called the PL94-171, or the Decennial Census data
- counts of people by race/ethnicity (126 histogram bins)
- voting age (≥ 18) or not
- This is a massive simplification of what's in the short form
- Important to distinguish the PL data from all other Census Bureau data products, such as the American Community Survey (ACS)

- P1. – Race
- P2. – Hispanic or Latino, and not Hispanic or Latino by Race
- P3. – Race for the Population 18 Years and Over
- P4. – Hispanic or Latino, and not Hispanic or Latino by Race for the Population 18 Years and Over
- P5. – Group Quarters Population by Major Group Quarters Type
- H1. – Occupancy Status (Housing)

Censuses: Why we count, who counts

- Essentially all census data is self-reported, but the questions shape the data product
- In the U.S., we are instructed to count for the purposes of political **apportionment**
- The very same sentence contains the “three-fifths compromise”
- This raises the issue that is often called “virtual representation”

Constitution of the United States

Article I

Article I Explained

Section 1

All legislative Powers herein granted shall be vested in a Congress of the United States, which shall consist of a Senate and House of Representatives.

Section 2

The House of Representatives shall be composed of Members chosen every second Year by the People of the several States, and the Electors in each State shall have the Qualifications requisite for Electors of the most numerous Branch of the State Legislature.

No Person shall be a Representative who shall not have attained to the Age of twenty five Years, and been seven Years a Citizen of the United States, and who shall not, when elected, be an Inhabitant of that State in which he shall be chosen.

Representatives and direct Taxes shall be apportioned among the several States which may be included within this Union, according to their respective Numbers, which shall be determined by adding to the whole Number of free Persons, including those bound to Service for a Term of Years, and excluding Indians not taxed, three fifths of all other Persons. The actual Enumeration shall be made within three Years after the first Meeting of the Congress of the United States, and within every subsequent Term of ten Years, in such Manner as they shall by Law direct. The Number of Representatives shall not exceed one for every thirty Thousand, but each State shall have at Least one Representative; and until such enumeration shall be made, the State of New Hampshire shall be entitled to chuse three, Massachusetts eight, Rhode Island and Providence Plantations one, Connecticut five, New-York six, New Jersey four, Pennsylvania eight, Delaware one, Maryland six, Virginia ten, North Carolina five, South Carolina five, and Georgia three.

Politics of who counts

- 1960s Supreme Court jurisprudence says “One Person, One Vote”
- The first Trump administration made a push to restrict the apportionment count to **citizens**
- JD Vance argued that **children** should be factored into voting power



These children are the future of this country and yet the parents who have them actually have no advantage in our democratic process. They have a smaller voice in some ways — in very many cases — than the people who don't have any children at all. The children who come from these families have no real representative in our democracy. Why don't we change that? Now some people will say this is radical and this is crazy. The Democrats are talking about giving the vote to 16-year-olds but let's do this instead: **Let's give votes to all children in this country but let's give control over those votes to the parents of those children.**

Politics of what is visible



PRESIDENTIAL ACTIONS

DEFENDING WOMEN FROM GENDER IDEOLOGY EXTREMISM AND RESTORING BIOLOGICAL TRUTH TO THE FEDERAL GOVERNMENT

EXECUTIVE ORDER

January 20, 2025

Sec. 2. Policy and Definitions. It is the policy of the United States to recognize two sexes, male and female. These sexes are not changeable and are grounded in fundamental and incontrovertible reality. Under my direction, the Executive Branch will enforce all sex-protective laws to promote this reality, and the following definitions shall govern all Executive interpretation of and application of Federal law and administration policy:

(a) “Sex” shall refer to an individual’s immutable biological classification as either male or female. “Sex” is not a synonym for and does not include the concept of “gender identity.”

(b) “Women” or “woman” and “girls” or “girl” shall mean adult and juvenile human females, respectively.

(c) “Men” or “man” and “boys” or “boy” shall mean adult and juvenile human males, respectively.

(d) “Female” means a person belonging, at conception, to the sex that produces the large reproductive cell.

(e) “Male” means a person belonging, at conception, to the sex that produces the small reproductive cell.