

# **Web Apps and Visualizations**

Big Data for Public Policy

**Christoph Goessmann**

Law, Economics, and Data Science Group (Prof. Elliott Ash)

Course: 860-0033-00L Big Data for Public Policy, Spring 2023

6 April 2023

# Student Projects

Paper or web app, 1-4 students per project

## Timeline (see syllabus):

- **Topic (0%), due ~~March 31st~~ April 9th.** Please discuss with the instructors to select and confirm a topic, dataset, and approach.
- **Outline (5%), due April 30th.** 1/2 page outline of the motivation, related literature, data, and approach.
- **Presentation (20%), On June 1st.** Students will give a short presentation about their project toward the end of the course. It should include some preliminary analysis. 5 minutes, +2 minutes for each additional student in the group.
- **Deliverable (60%), due July 21st.** Paper OR web app.
  - A **paper** reporting on the project's analysis and results (Intro, Lit Review, Data, Methods, Results, and Conclusion). 2500+ words, +500 words for each additional student in the group.
  - **Web app.** The web app should visualize data, topics, model predictions... Points for creativity. The web app should be accompanied by a 1-page document explaining why this web app is of interest and what were the main challenges.
  - **Replication package (15%).**

# Repetition

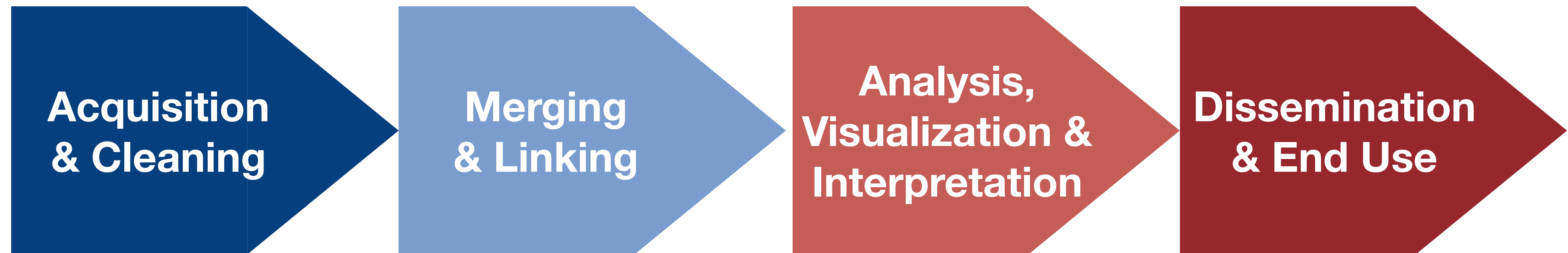
## What did we do last time?

- What is machine learning?
  - “**performance** on **task** improves with **experience/data**”
- Machine learning vs. **econometrics**
  - dependent var. → **label**
  - independent var. → **feature**
  - causal inference → **inference**
- **Supervised** vs. **unsupervised** ML
- Performance metrics, choice highly dependent on **context**
- Regularization: modified loss function to **prevent overfitting**
- **Scraper template** for rate-limited and robust API calls

# Tentative Outline for Today

- **Data value chain**
- Effective **visualizations**
- **Web apps**
  - **SDG Monitor** (how you could embed a web app in a research project)
  - **Hands on** (program own web app)

# Data Value Chain



# Effective Visualizations

- Choose **appropriate chart type**
- **Declutter & simplify**
- Choose the **“right colors”** (no **red-green**)
  - discrete vs. continuous
  - sequential vs. diverging (vs. qualitative)
  - ~ 9% (men) and 1% (women) are **color blind**

## If possible

- Interactive
- Hovering for additional information

## Tools

- **colorbrewer 2.0**
- **viridis** and friends

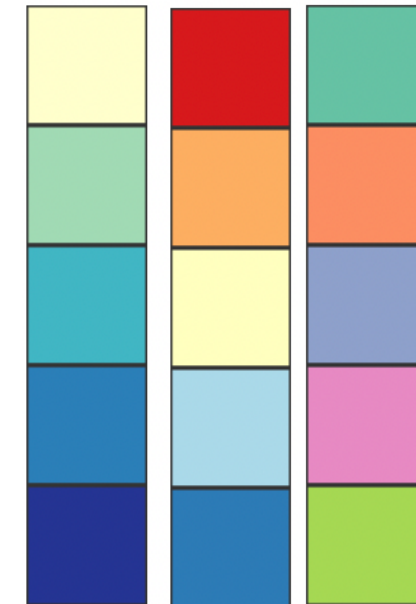


- color blind-friendly
- perceptually uniform
- also work when converted to grayscale
- wide cover of perceptible range
- **Coblis – Color Blindness Simulator**

# Activity



 **Check a visualization you last worked with using Coblis.**



**Diverging?**  
**Sequential?**  
**Qualitative?**

# **Monitoring SDG Data Gaps**

Developing a web app for decision makers,  
Assessing the state of SDG data

**Example for embedding a web app  
in a research project.**



# Hands On

## 0. Take a look at SDG Monitor

## 1. Create a virtual environment

```
$ pip install pipenv
```

```
$ mkdir bdpp_session4
```

```
$ cd bdpp_session4
```

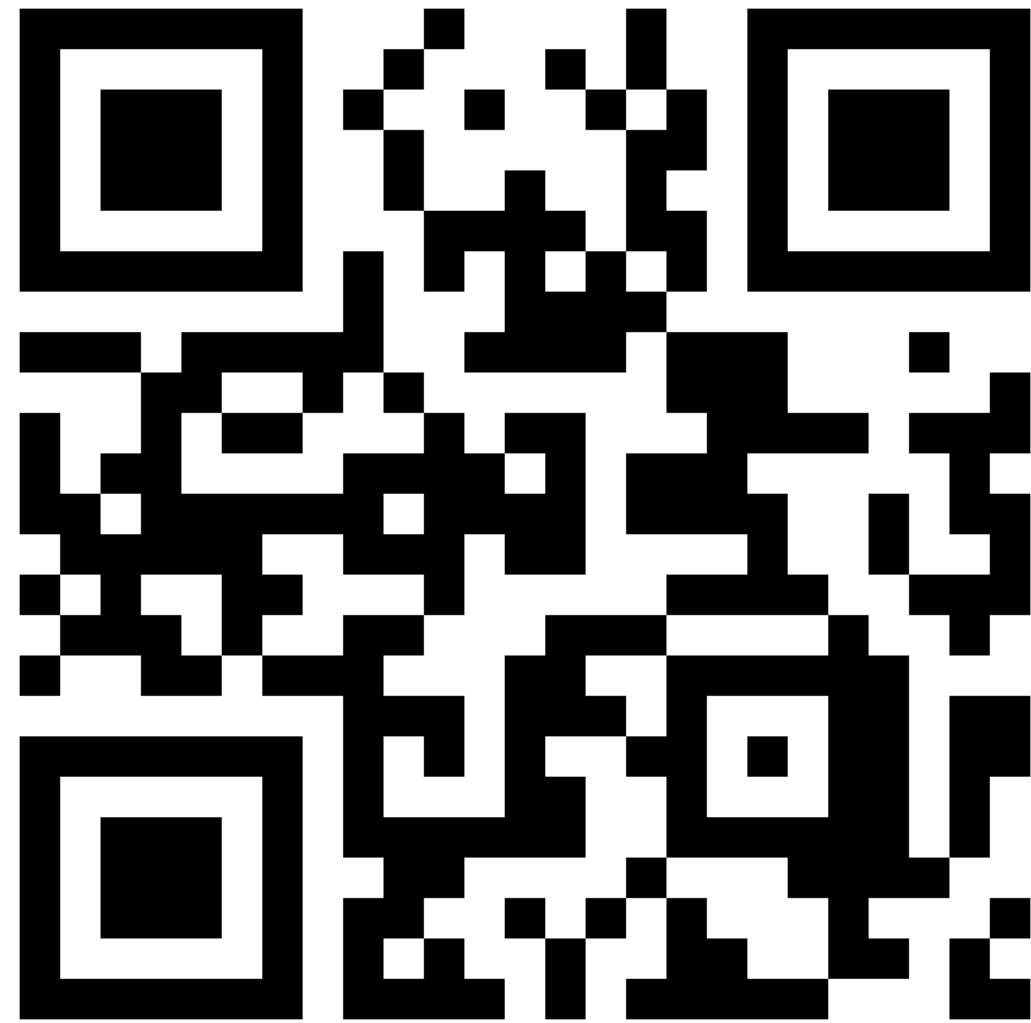
```
$ pipenv install jupyter-dash pandas requests
```

## 2. Open this lecture's jupyter notebook

[Dash documentation](#)

# End-of-Lecture Survey

ETH Edu App



Web app

<https://eduapp-app1.ethz.ch/>



iOS



Android