Project Proposal: Return to Education and Income in the United States (2019–2023)

a. Research Topic

This project investigates the effect of educational attainment on individual income in the United States, controlling for demographic variables (age, gender, race/ethnicity, nativity, marital status, household size) and county-level economic conditions (number of establishments, unemployment rate, county GDP per capita). The analysis merges individual-level data from the ACS 5-year files (sourced via IPUMS for 2019–2023) with county economic indicators from the County Business Patterns (CBP) datasets during the same period.

b. SMART Questions

- What is the average increase in annual income associated with each additional year of education, after accounting for demographic and county-level economic covariates?

- How does the return to education vary by demographic subgroups and across counties with differing economic conditions?

- To what extent do county-level industrial composition and employment rates moderate the income benefits of education?

c. Data Sources & Observations

- Individual-level microdata: ACS 5-year sample (IPUMS, 2019–2023)

- County-level economic data: County Business Patterns (CBP, 2019–2023)

- Approximate size after cleaning and merging: 15 million observations

d. Team GitHub Repository

https://github.com/constantine-zhang96/DataScience-Project-2

e. Citation

County Business Patterns (CBP):

U.S. Census Bureau. County Business Patterns, 2019–2023. Economic Census Data Tables and Full Datasets, https://www.census.gov/programs-surveys/economic-census/data.html.

IPUMS ACS 5-year microdata:

Steven Ruggles, Sarah Flood, et al. IPUMS USA: Version 1.0 [American Community Survey 5-Year Data, 2019–2023]. Minneapolis, MN: IPUMS, https://usa.ipums.org/usa/index.shtml.