Local, timely, and actionable: A plan for democratizing education and workforce data

Julia Lane
NYU and RTI

With many thanks to Angie Tombari, Jessica Cunningham, Anna-Carolina Haensch, Jared Dubin, Xiangyu Ren, Jody Williams, Frauke Kreuter and many others
Overview

• Context
• Vision
• Approach
• Next steps
FACT SHEET: CHIPS and Science Act Will Lower Costs, Create Jobs, Strengthen Supply Chains, and Counter China

On July 28, the President and the bipartisan group announced agreement on the details of a once-in-a-generation investment in our infrastructure, which was immediately taken up in the Senate for consideration. The legislation includes around $550 billion in new federal investment in America’s roads and bridges, water infrastructure, resilience, Internet, and more. The bipartisan Infrastructure Investment and Jobs Act will grow the economy, enhance our competitiveness, create good jobs, and make our economy more sustainable, resilient, and just.

The Inflation Reduction Act Provides Pathways to High-Quality Jobs

Incorporating labor and workforce training standards into this historic climate policy will ensure that the transition to a clean energy economy is built with good jobs.
6 job quality metrics every company should know

WORKER WELLBEING

JOB QUALITY
- Percentage of workers that earn a living wage and have healthcare
- Number of new good jobs created each year

ECONOMIC MOBILITY
- Access to steppingstone jobs: percentage of workers that transition from earning low wages to high wages
- Percentage of workers that cross the one year mark in the company

JOB EQUITY
- Workforce demographic gaps in the company's high-wage occupations
- Gaps in wage mobility across demographic groups
Context

State Postsecondary Data

SHEEO is the higher education community's primary source of information regarding the capacity and effective use of state postsecondary data systems.
Overview

• Context
• Vision
• Approach
• Next steps
Reimagining Labor Market Information
A NATIONAL COLLABORATIVE FOR LOCAL WORKFORCE INFORMATION
MARCH 2023
Julia Lane

Vision
Vision
Overview

• Context
• Vision
• **Approach**
• Next steps
Build local, actionable measures

NEW EXECUTIVE CERTIFICATE IN DATA LITERACY AND EVIDENCE BUILDING

New York University’s Robert F. Wagner Graduate School of Public Service, the University of Maryland’s Social Data Science Center, and Accenture are delighted to announce a new non-credit Executive Certificate in Data Literacy and Evidence Building in conjunction with the Kentucky Center for Statistics and the Coleridge Initiative. The new first-of-its-kind Certificate draws on almost a decade’s successful experience to train participants in how to work with synthesized real-world data drawn from different sources to produce new insights and products.

PROGRAM DETAILS

Dates: September 6, 2023 to November 9, 2023
Lecture: 12:00–4:00 pm ET
That answer local questions

- What are the earnings for graduates with nursing degrees from my local school?
- How long does a typical first job for a business major last in this area?
- What is the difference in earnings for someone who completes a STEM credential here vs. someone who doesn’t?
- How many people who start a two year degree in computer science at this community college complete in three years?
- How many people in the community are just two or three credits short of a job-qualifying credential?
What managers said they want
What participants said they want
Project – products – practice
Project – products – practice
Class content

Class participants will develop the key data analytics skill sets necessary to scope a real world project using real world data, as well as develop an analytical product and apply that product to practice. It trains analysts how to recognize and deal with some of the common pitfalls in terms of making sense of data, identifying and minimizing bias, and communicating clear operational information to elected or appointed officials and other policy-makers. It is explicitly designed to respond to the recommendations of the Commission on Evidence-based Policymaking, the Foundations of Evidence-based Policymaking Act of 2018, and the recommendations of the Advisory Committee on Data for Evidence Building.
The goal of this notebook is to provide you with the code to produce summary statistics for different cohorts of students in the Syntucky dataset. The results will be an overview of how many people have jobs and the associated earnings in a particular cohort, including a count of the number of missing observations.

We will take a look at the number of the following descriptive statistics:

- **count**: the total number of non-null entries in each column.
- **mean**: the average of all entries in each column.
- **std**: the standard deviation of each column, which is a measure of the spread of values.
- **min**: the smallest value in each column.
- 25% (also known as the first quartile or Q1): the value below which 25% of the values fall.
- 50% (also known as the second quartile, median or Q2): the mid-value.
- 75% (also known as the third quartile or Q3): the value below which 75% of the values fall.
- **max**: the largest value in each column.

**Coding tools**

This notebook also identifies some tools that will make our journey smoother and analysis: Pandas and NumPy.

**Pandas**: Pandas is a Python library specializing in making data manipulation and analysis easier. It is designed to handle data in different formats using a table-like format called DataFrames.

**NumPy**: NumPy is a bit like a high-powered calculator for Python. It is another library you need to perform mathematical operations on a lot of numbers simultaneously.

```python
In [1]: import pandas as pd
```

**Measurement**

Note: The “Syntucky” data, which is synthetic in nature, is exclusively designed for training exercises. It is not intended to derive meaningful insights or make determinations about real-world populations.

The focus in this notebook is the heart of data analysis: accurate measurement.

The primary goal of this notebook is to explore methods for identifying and categorizing individuals in the dataset based on their educational attainment. Three distinct categories will be defined in this notebook: ‘completers,’ ‘non-completers,’ and ‘degree pursuers.’

The secondary goal is to measure the employment outcomes of individuals in each group - particularly focusing on the quality of their jobs.

**Import Packages**

As before, we will again use the numpy and pandas packages.

```python
In [ ]: import pandas as pd
In [ ]: import numpy as np
```

**Import Data**

Today, we will use the 2016 completion data.
Ed/workforce agency staff – build their own

Choose the degree of interest
- Nursing
- Education,
- Computer Science
- Business Majors.....

Choose the population of interest
- Completers
- Non completers
- Degree pursuers

Construct the job measure of interest
- Any job
- Earnings above threshold $y$,
- Earnings in at least 2, 3, or 4 quarters;
- Earnings growth

Choose how to visualize
- Comparison by economic status
- Trends over time
- Map
Overview

- Context
- Vision
- Approach
- Next steps
Focus
1. Critical thinking in the context of the underlying data structure (schemas)
2. The opportunities (structuring cohorts, and generating measures of duration)
3. Pitfalls (missing values and systematic coverage challenges)

Enrollments: 52 in class – TX, NJ, AR, OH, PA, KY, VA, MI, UT, MO, CO

Build their own dashboard!
Data Literacy and Evidence Building for the public sector.

Panelists
- Sherry Glied, Dean, Public Policy, NYU
- Keith Marzullo, Dean, Information Studies, UMD
- Josh Hawley, Director, Ohio Education Education Research Center, OSU
- Adam Leonard, Chief Analytics Officer, Texas Workforce Commission
- Matt Berry, Director, KY Stats
- Christina Whitfield, Senior VP, SHEEO
Your role – reach out to your state institutions
Questions?