CIVIC DIGITAL FELLOWSHIP

Linking Census with Third Party Data: The Case of Federal Research Funding

Andrea Chen

Informatics, University of Washington

Data Scientist, Civic Digital Fellow
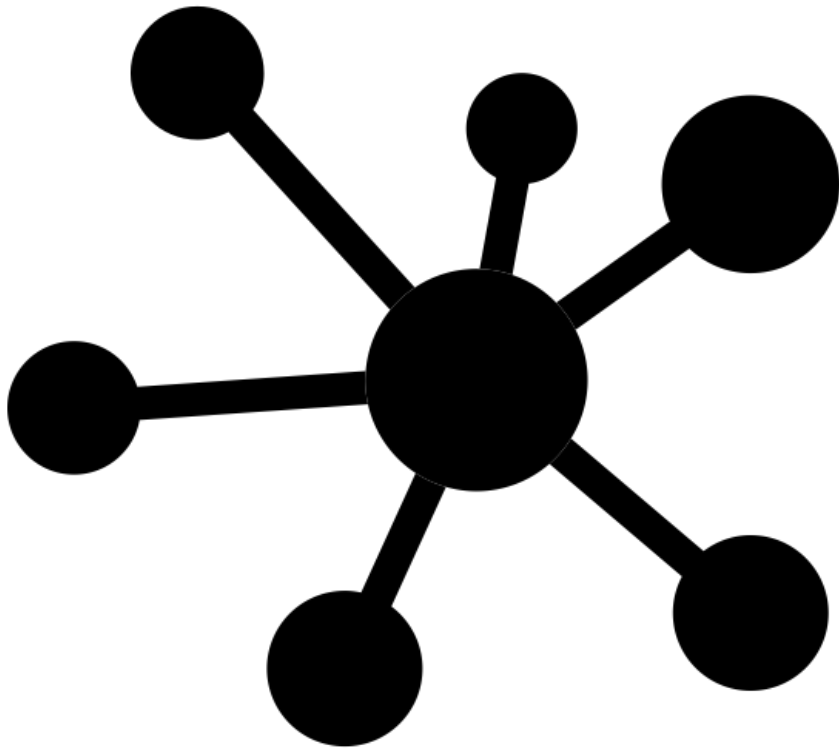
United States Census Bureau

Mentor: John Cuffe, CBDRA

Project Overview

How can the Census use data from external sources?
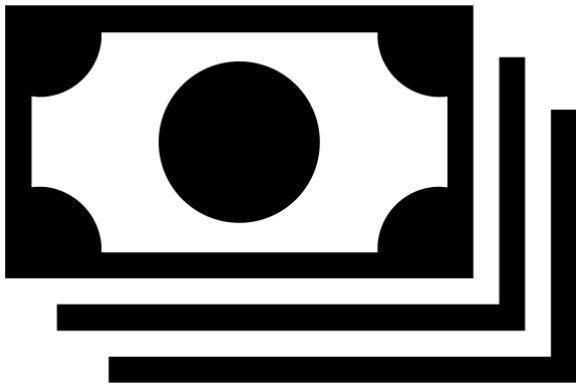
Project Overview

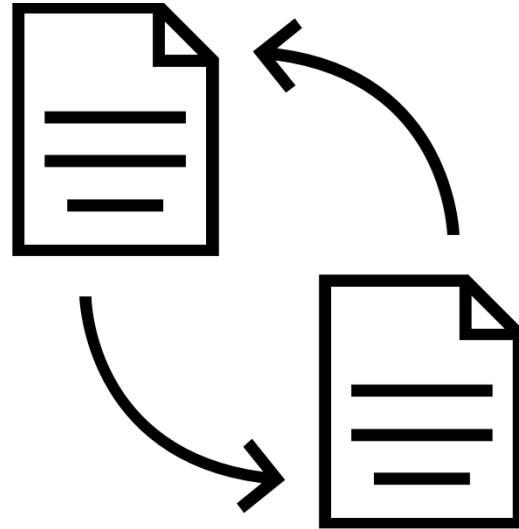Census Center for Big Data Research and Applications

Innovation Measurement Initiative
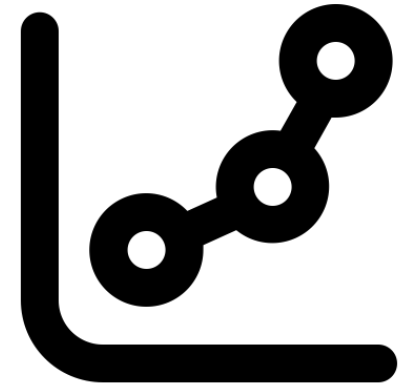
Project Impact

Reduce costs of re-producing existing statistics
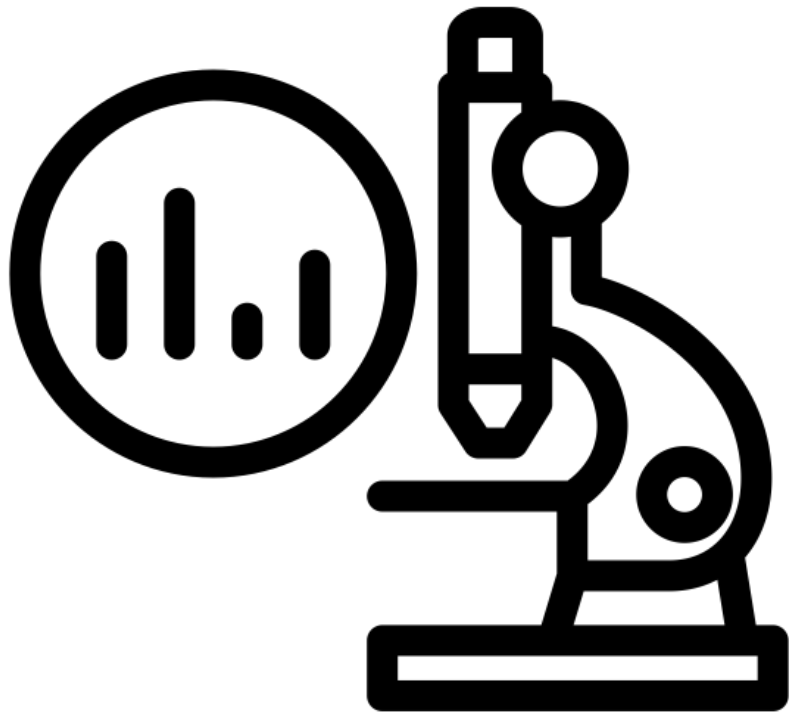
Increase accuracy of the data

Produce new statistics and economic indicators

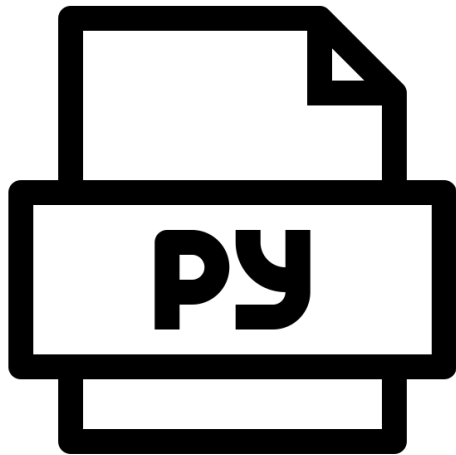# $18 billion

Research Question

What is the economic impact of research spending?

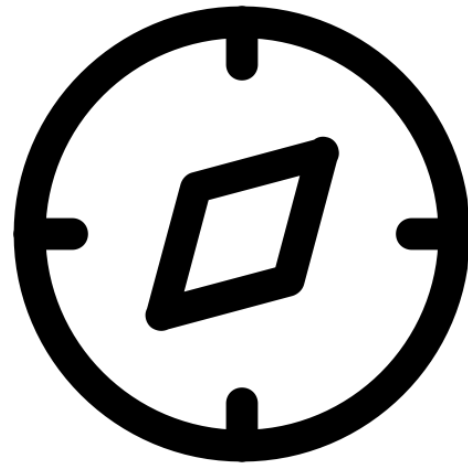What are the career outcomes of students paid by research grants?

Are undergrads who receive research grants more or less likely to work at young firms?
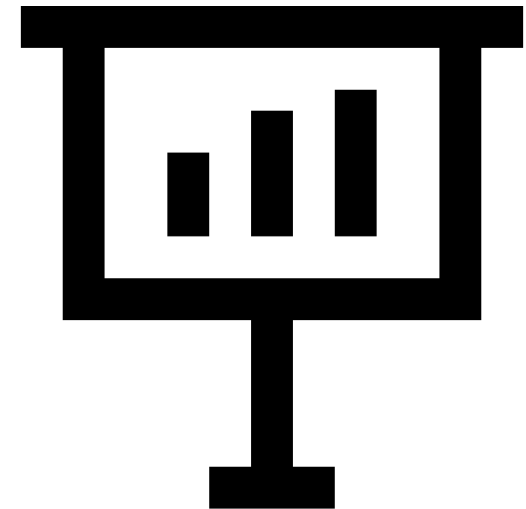
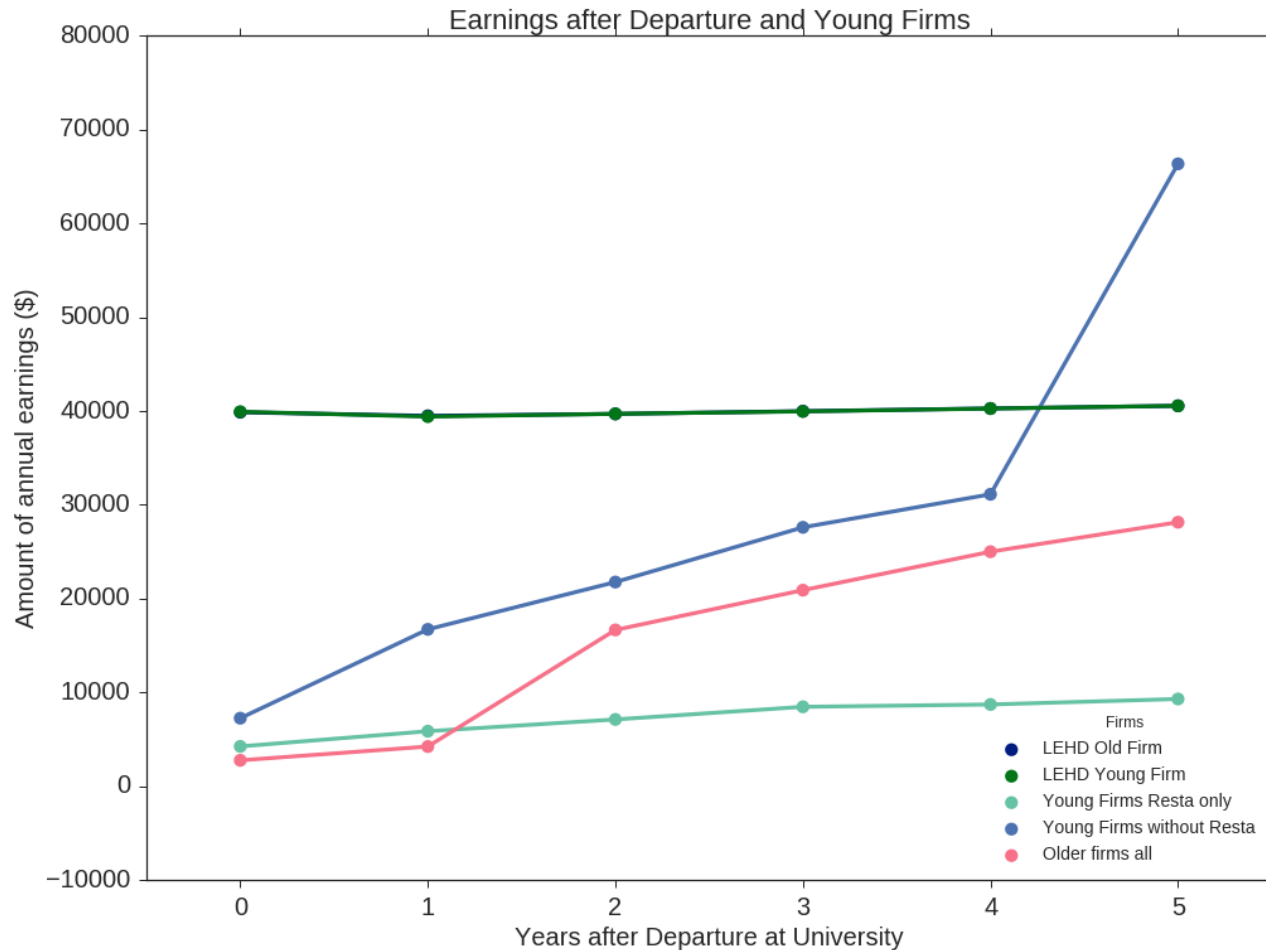# Technologies Incorporated

Utilized Python and its packages
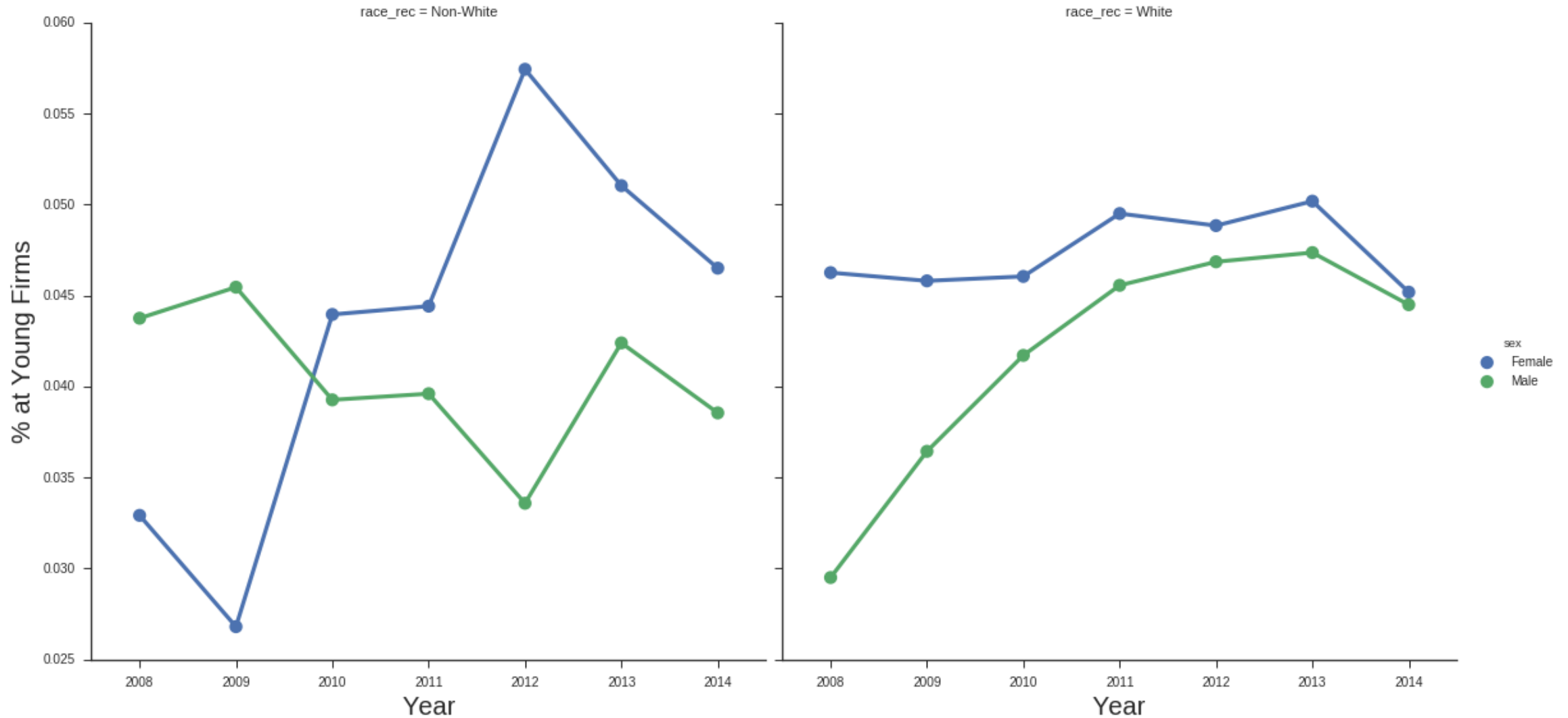
Explored data with pandas and numpy

Visualized the data with seaborn and matplotlib

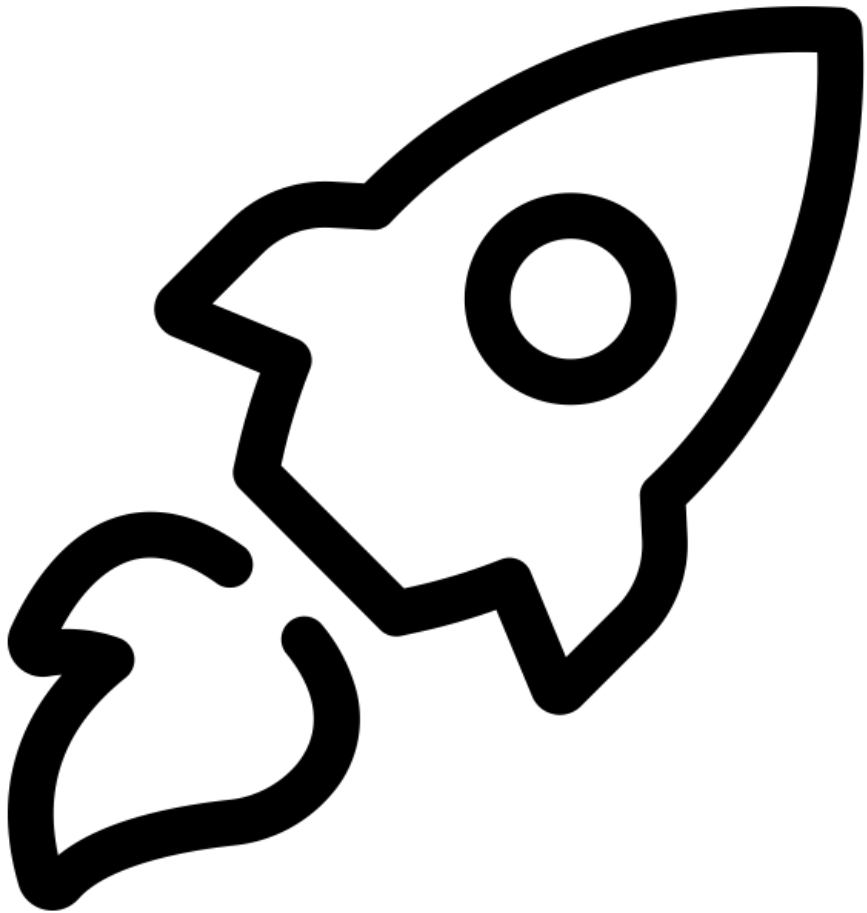# Earnings at a young firm tend to grow at a faster rate.

Earnings after Departure and Young Firms

More women tend to be at young businesses.

Concluding Discoveries

Real data is incredibly complex and messy

Data security was always a top priority throughout my research at the Census

Be bold and ask questions

Thank you!

John Cuffe

Sudip Bhattacharjee

Jeff Meisel

Adria Thomas

Shawn Roberts

Ann Wittenauer

William Wisniewski

Contact Information

Andrea Chen, Civic Digital Fellow

University of Washington

abchen@uw.edu

John Cuffe, Researcher

Center for Big Data Research and Applications

john.cuffe@census.gov