UNLV | UNIVERSITY LIBRARIES

2018

Datalog for Non-Profit Demographic Analysis

Jordan Mulcahey
University of Nevada, Las Vegas, snhjordy@gmail.com

Patricia Payton
University of Nevada, Las Vegas, payton@unlv.nevada.edu

Ignacio Regalado
University of Nevada, Las Vegas, regalado@unlv.nevada.edu

Repository Citation

Datalog for Non-Profit Demographic Analysis

By: Jordan Mulcahey, Patricia Payton, Ignacio Regalado

foundation for an independent tomorrow

Datalog is a useful tool in organizing data. With it, we can achieve a better in-depth analysis on how the FIT program is performing.

1. What is F.I.T?

FIT, the Foundation for an Independent Tomorrow, is a non-profit organization that provides resources and training for those who want employment.

2. The Problem

- Nonprofits like FIT often store their data in simple programs like Excel.
- This can make statistical data analysis difficult.
- Our team decided to design a data storage and analysis system using Datalog.
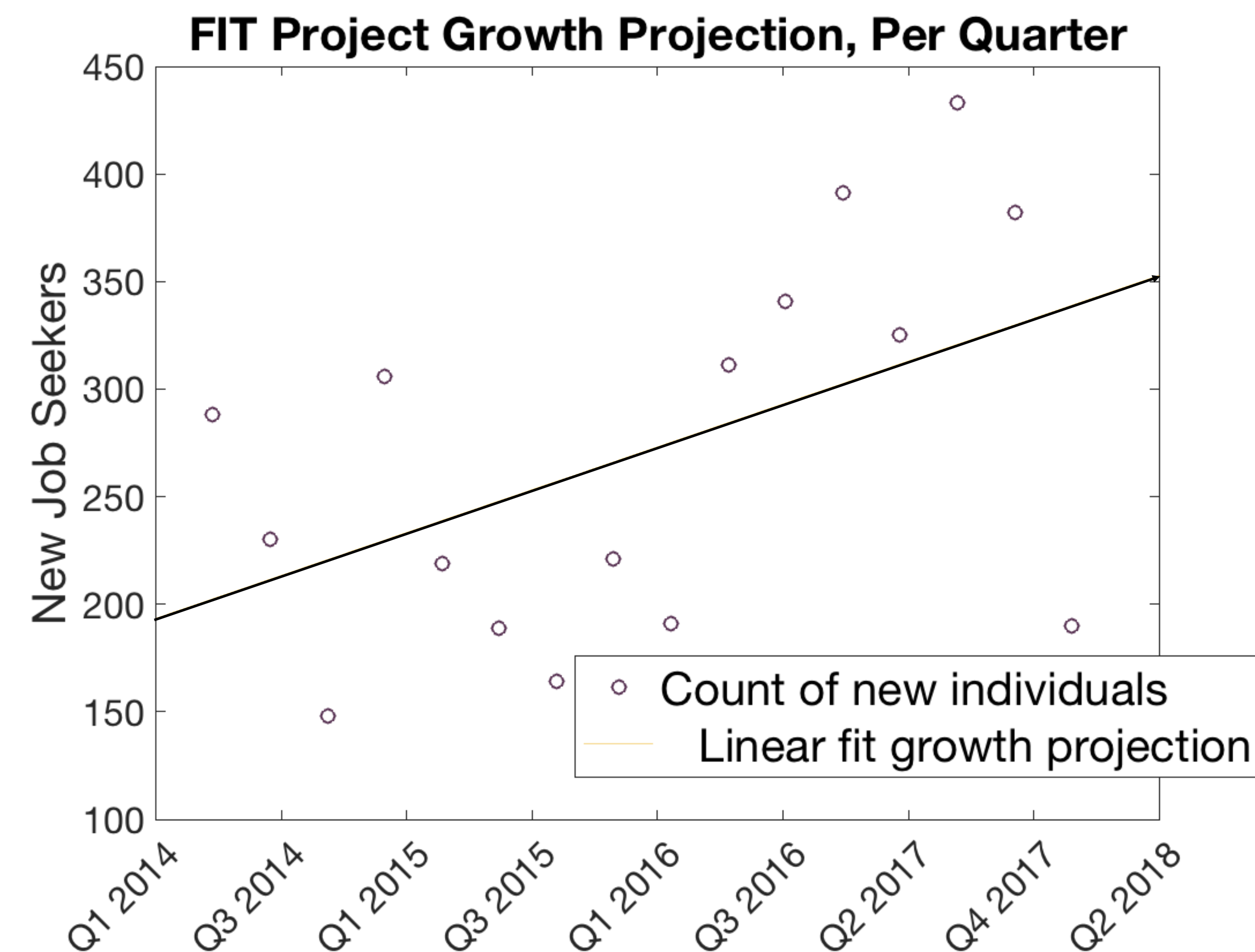
3. The Methodology

- In our study, we use tuple relational calculus and Codd's theorem to show the massive expressive power of Datalog.
- We designed a custom, easy to use Datalog system, and applied it to the FIT dataset.
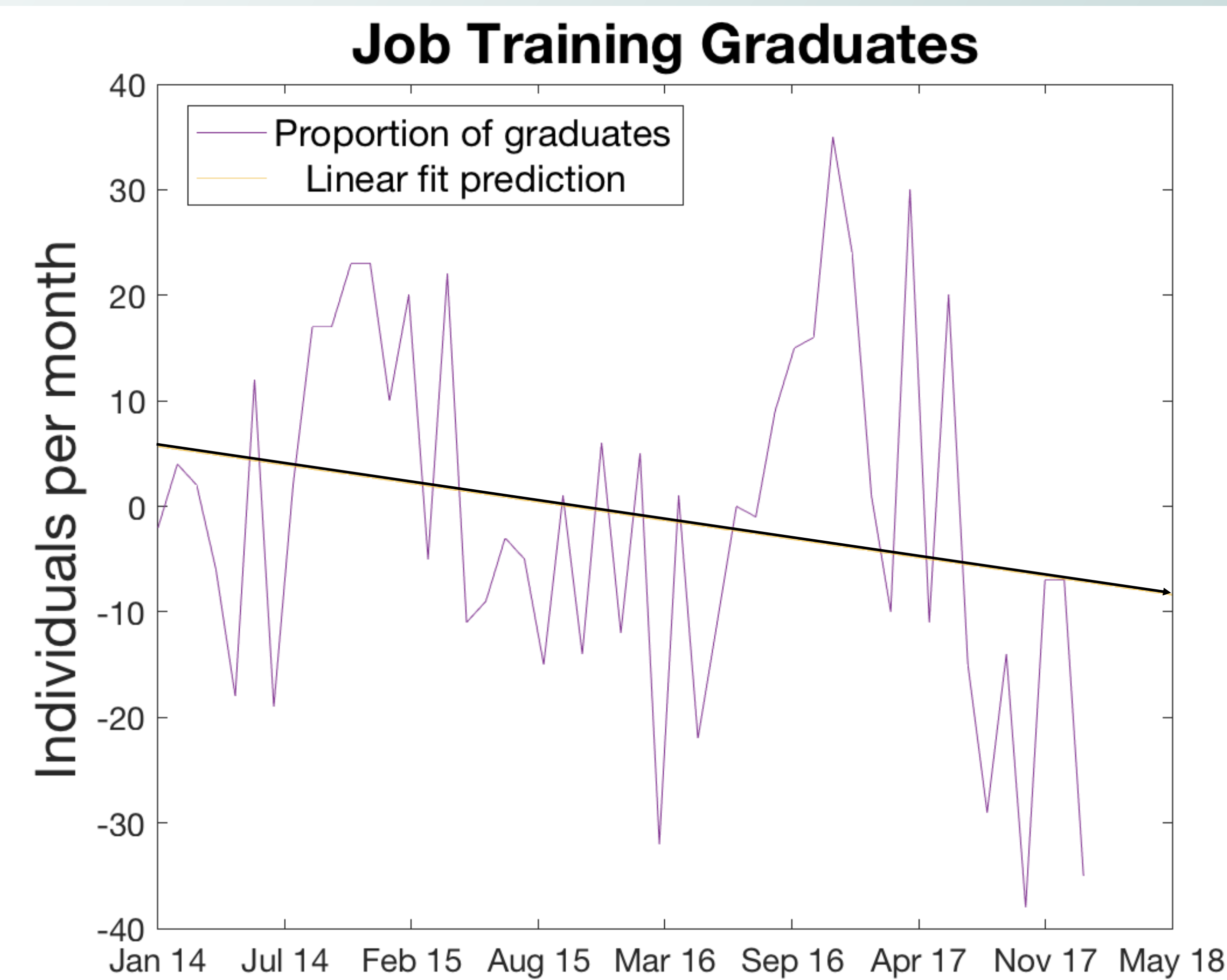- Linear fit was used on the correlated data to predict future project performance.
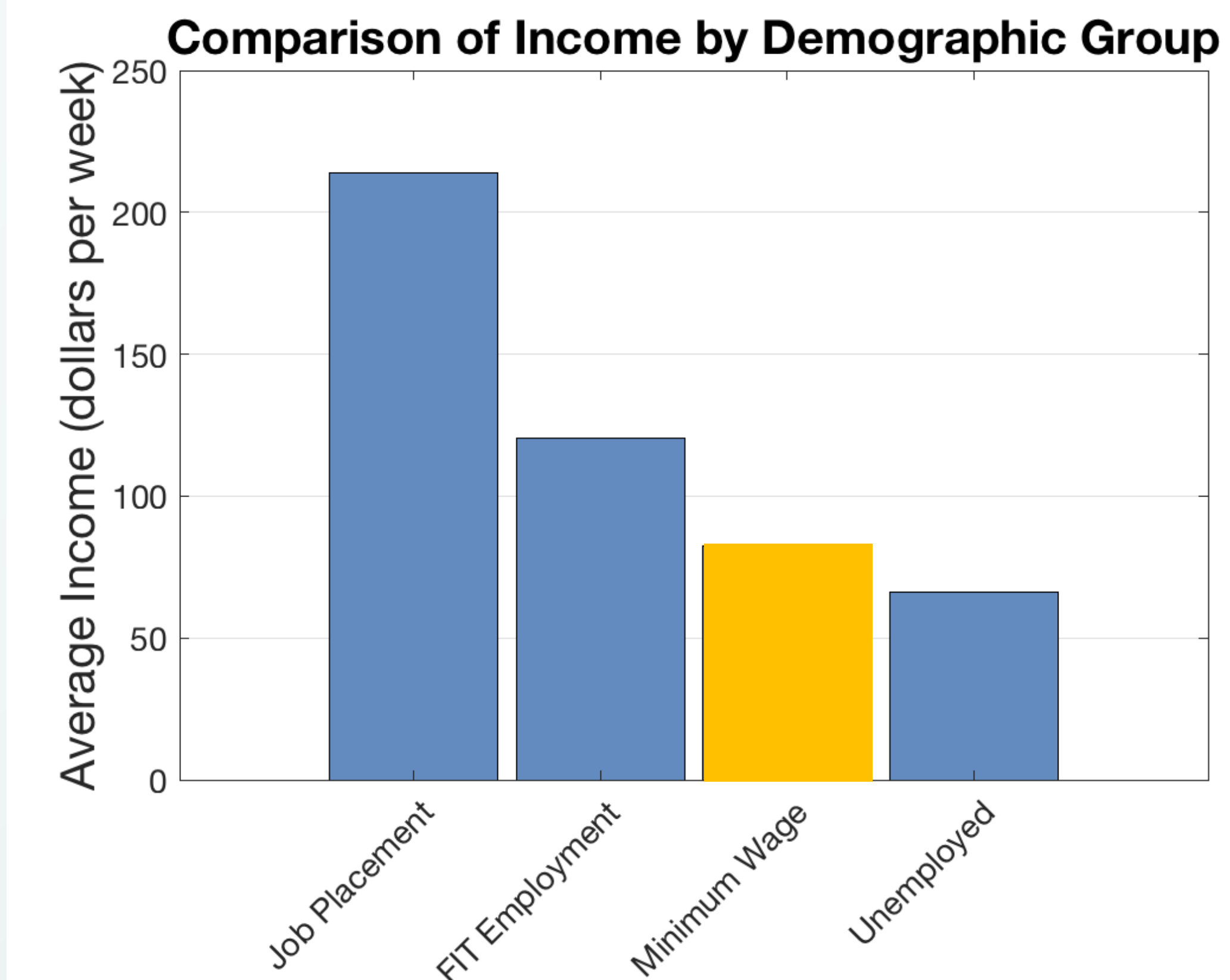
4. The Results

FIT Project Growth Projection, Per Quarter

(X-axis: Q1 2014, Q3 2014, Q1 2015, Q3 2015, Q1 2016, Q3 2016, Q2 2017, Q4 2017, Q2 2018; Y-axis: New Job Seekers, 100–450)

- ○ Count of new individuals
- — Linear fit growth projection

We predict that FIT enrollment will steadily grow.

Job Training Graduates

(X-axis: Jan 14, Jul 14, Feb 15, Aug 15, Mar 16, Sep 16, Apr 17, Nov 17, May 18; Y-axis: Individuals per month, -40 to 40)

- — Proportion of graduates
- — Linear fit prediction

Graduation rates appear to have a declining trend as the project grows.

5. Comparison

Comparison of Income by Demographic Group

(Y-axis: Average Income (dollars per week), 0–250; X-axis: Job Placement, FIT Employment, Minimum Wage, Unemployed)

Compares FIT participants' wages versus unemployed income.

Proportion of Income to Public Assistance Received

(Y-axis: Average Income (dollars per week), 0–250; X-axis: FIT Employed, Job Placement, Unemployed)

- ■ Income
- ■ Public Assistance
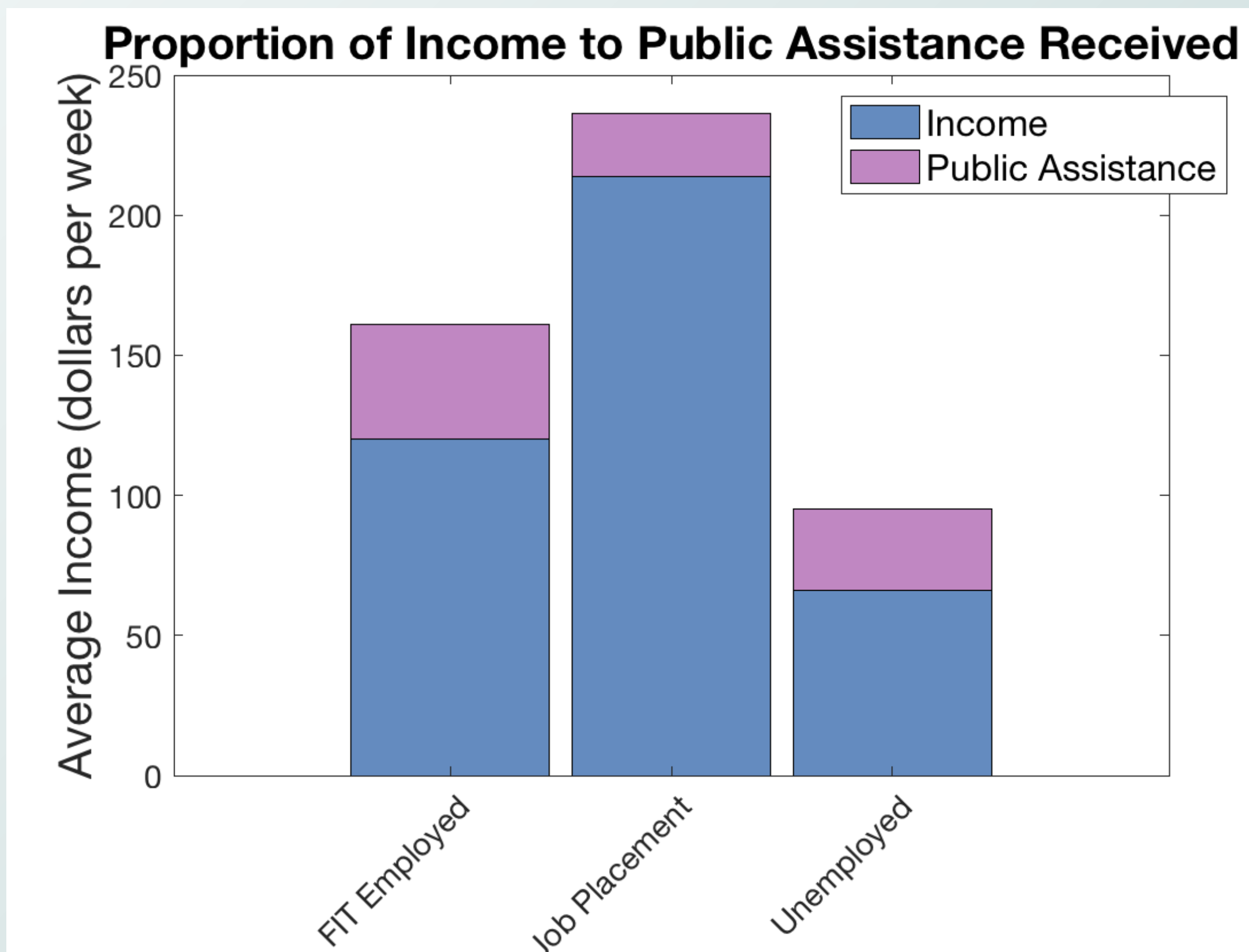
Compares weekly income and their use of public assistance between FIT participants.

The left two bars are FIT participants who were employed directly by FIT, or employed elsewhere.

6. Conclusion

Through data analysis and correlation powered by Datalog, we were able to expose the strengths and weaknesses of FIT's current data collection methods, and reveal several opportunities for greater FIT program efficiency.

Warren, David S. "Programming in Tabled Prolog", 1999, https://www3.cs.stonybrook.edu/~warren/xsbbook/book.html
Ramsdell, John D. "Datalog User Manual", 2004, http://www.ccs.neu.

edu/home/ramsdell/tools/datalog/datalog.html
G. Gottlob, S. Ceri and L. Tanca, "What You Always Wanted to Know About Datalog (And Never Dared to Ask)," in IEEE Transactions on Knowl-

edge & Data Engineering, vol. 1, no. , pp. 146-166, 1989.
Pichler, Reinhard, "Database Theory", 2018, http://www.dbai.tuwien.ac.at/staff/pichler/dbt/slides/dbt03.pdf

Xiong, Li, "Relational Calculus and SQL", http://www.mathcs.emory.edu/~lxiong/cs377_f11/share/slides/08_SQL.pdf
Zaïane, Osmar R, "Database Systems and Structures", 1998, https://

www.cs.sfu.ca/CourseCentral/354/zaiane/material/notes/contents.html

UNLV