Determining Interesting Stories

Improving Efficiency for Analysts via Automation

Summer Long

08-14-2023

Agenda

- Introduction of DIPS

- Motivation for Project

- Process

- Program Demonstration

Introduction of DIPS

■ The Division of Industry Productivity Studies produces annual measures for detailed industries, including but not limited to:

  ▶ Productivity

  ▶ Output

  ▶ Hours

  ▶ Labor compensation

■ Organized by groups of industries and released once a year

(Credit: Eugene Yum)

BLS

Motivation for Project

- Analysis packets are prepared for the Associate Commissioner
  - ▶ However, they prove to be laborious on analysts
- Potential solution: automate part of the materials in the decision process

Process

1. Reviewing Existing Analysis Packets

- Common visualizations from analysis packets were identified

- Select visualizations were chosen for automation

  ▶ An interest score was devised to emulate analyst selection which is a weighted score including output index, hours index, output per hour index, and employment

- End goal of generation: a table displaying commonly used measures, and sorted labor productivity graphs

2. Familiarizing With Database and Measures

- Database was accessed through SQLAlchemy

- Crash-coursed in commonly leveraged measures, such as:
  - Output index
  - Hours index
  - Output per hour
  - Unit labor cost

3. Figure Generation

- Figures were generated with Matplotlib

- Description of parameters and figures were imported on Word for easy use of analysts using docx package

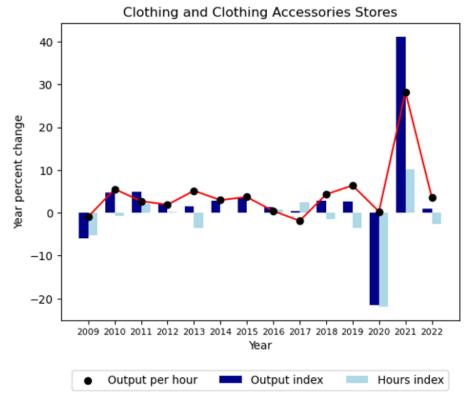- 508 compliance was prioritized to incorporate into generation
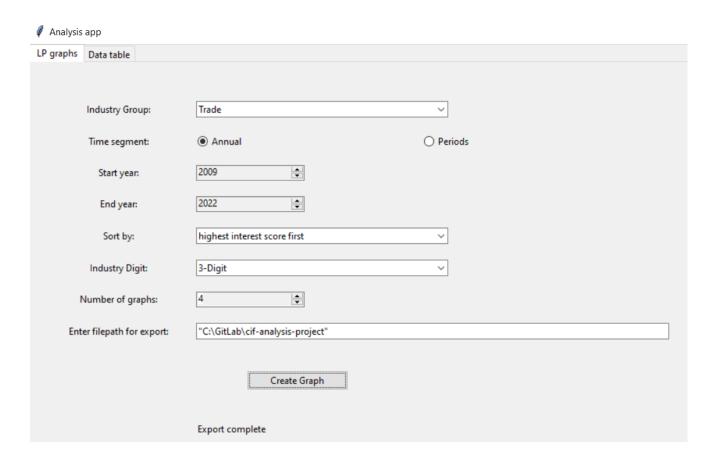
4. User Interface Creation

■ Tkinter was leveraged to create a user interface with user selected parameters

▶ More user-friendly than running code from the command line

■ Currently accessed by running main_window.py on an IDE, but will be packaged into a program that is ran in the future

Clothing and Clothing Accessories Stores

	Output index	Hours index	Output per hour
2009	-6.0	-5.2	-0.9
2010	4.8	-0.8	5.5
2011	4.8	2.0	2.7
2012	2.1	0.2	1.9
2013	1.5	-3.5	5.1
2014	2.9	-0.1	3.0
2015	3.7	-0.0	3.7
2016	1.3	0.8	0.4
2017	0.5	2.4	-1.9
2018	2.8	-1.5	4.3
2019	2.7	-3.5	6.4
2020	-21.6	-21.9	0.3
2021	41.2	10.1	28.2
2022	0.9	-2.6	3.6

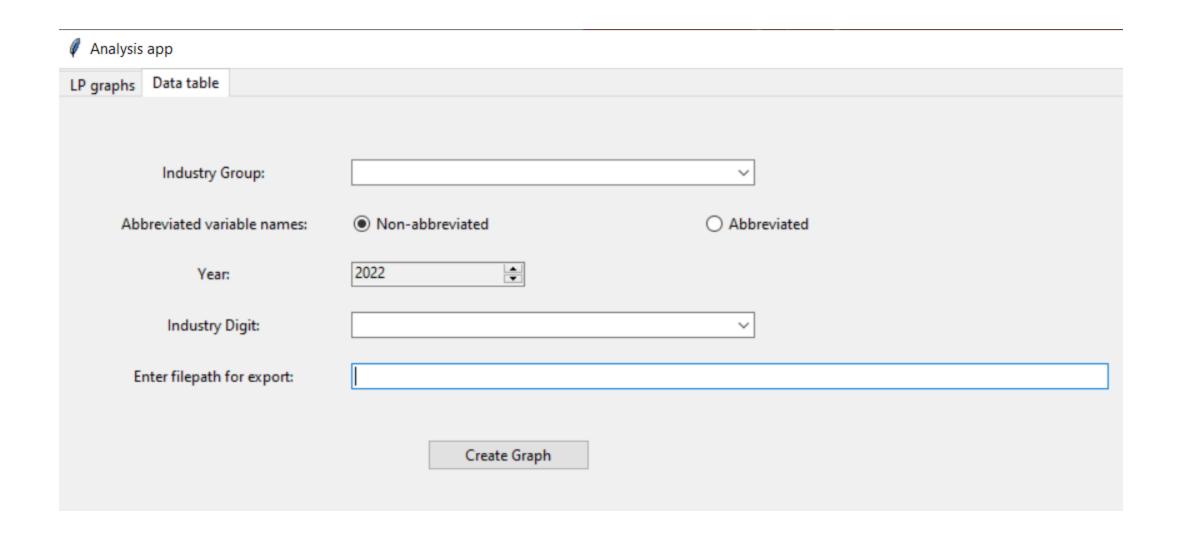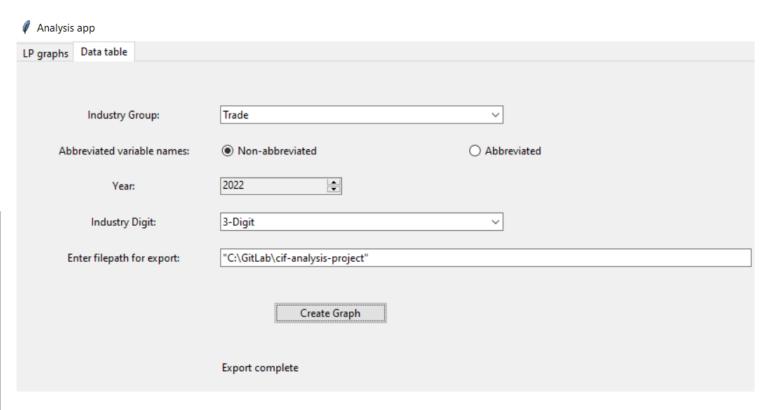"Most interesting chart" generated from the input displayed above

Table of Trade 3-Digit Industries Percent Change Metrics

NAICS	Industry	Output	Hours	Output per hour	Value of production	Implicit price deflator	Unit labor cost
424	Merchant Wholesalers, Nondurable Goods	-0.42	1.45	-1.85	17.77	18.28	10.67
425	Wholesale Electronic Markets and Agents and Brokers	-2.01	1.39	-3.35	14.8	17.15	14.87
423	Merchant Wholesalers, Durable Goods	4.13	4.06	0.07	10.56	6.18	8.74

Abridged sample of table generation from input displayed above

Program Demonstration

Contact Information

Summer Long
Data Science Intern
University of Chicago B.A. '23, M.S. '24
https://www.linkedin.com/in/summer-long-
salong@uchicago.edu

Appendix: Interest Score Calculation

$(|.2x| + |.2y| + |.6z|) * b$

- x is the difference between the percent change in the output index for the 2 most recent datapoints recorded

- y is the difference between the percent change in the hours index for the 2 most recent datapoints recorded

- z is the difference between the percent change in the output per hour index for the 2 most recent datapoints recorded,

- b is the number of employees in the most recent datapoint recorded