STRATEGIC DATA PROJECT

FELLOWSHIP CAPSTONES

Capstone Impact Story

Fellow: Sue Mukherjee, Shippensburg University, Pennsylvania's State System of Higher Education

Looking Past the Vault: *Leveraging New Ways to Track Graduates to Gain Consumer Insights*

Synopsis

There are multiple purposes of higher education. Preparing students for the labor market[1] is one of them. The higher education sector tracks internal progress of students using numerous metrics around access, completion, and cost. There also exists high demand for a holistic discussion on the value proposition of higher education. With the top motivation for students (entering the higher education enterprise) being work, this study adds to this discussion in light of new technologies in the Pennsylvania context. It offers a replicable process to link student records to their workforce outcomes to answer the question *how does education link to work in the real world?* This study is based on the premise that aggregate wages is one facet of the outcomes discussion. The value of higher education is unique by programs, students, geographies, and institutions. Understanding region-specific career trajectories of graduates is essential to comment on the education-workforce connection. It is inspired by a 2017 conversation with a graduate from Pennsylvania's State System of Higher Education and the way he designed the value of his education – via the career trajectory he dreamed of:

> *'We are lower middle class students. The value I'm getting is the ability to make connections and preparing for a career. I would want to go to graduate school and try to get my PhD in Political Science. Also, I would like to run for public office in order to help the people who have been abandoned by our current government.'*

The Challenge

To understand our learners, who they are, how we serve them, and where they go, the higher education enterprise has data systems in place. They originate internally from student and alumni data for the most part. To understand the value students add to the workforce, survey data are used. This self-reported data has been difficult to collect and maintain, struggled with low response rates, and cannot be benchmarked across systems. Other sources exist. National and regional surveys collect data (from students) about their education and workforce experience[2]. The College Scorecard[3] reports annual earnings of graduates aggregated by post-secondary institutions. But, there is a need for more granular data and the audience for this information is far beyond policy makers. Parents and learners need real time agile access to this information at a program level as education becomes one of their most important financial decisions. Faculty and employers gain valuable insight on the skills alignment (or misalignment) specific to programs and occupations (respectively) only if information is disaggregated by those attributes.

One method to gain such insight is by developing coordinated (longitudinal) data systems that link student and workforce data (from administrative wage records[4]). This data linkage connects individual student records to their employment history and has to be done by state agencies. Linking student

records to their wage record through this administrative database or merging student data (from higher education institutions) with employment information of graduates (from administrative wage records[5]) can illustrate how students are doing in the workforce by institution, program, and geography.

Acknowledging the need for this critical information, over $750 million in federal funds[6] has been distributed to states to build infrastructures to support this data exchange[7]. Developing a functional education-workforce data warehouse and getting access to the linked information is a monumental task for most states. Construction of and access to this bureaucratic vault requires navigation through a multitude of data-security and legal issues (Appendix 1). Additionally, policy, politics, interpretation of state and federal laws, and fear of data linkage are factors that have created a culture where this capital-intensive linkage has been an almost impossible task for many states. And, while promising practices of education and workforce connectivity exist at state and interstate levels[8], challenges exist. Even when states have been successful in linking education and workforce data, the audience has been policy makers for the most part, access has been an issue[9], and key data gaps have been recognized. Translating the linkage into meaningful information has also been a challenge. Concentrating mostly on aggregate wages by industry as an outcome measure (of higher education), this linkage has not been able to shed light on the career-trajectories of graduates by occupation.

Pennsylvania has been the recipient of federal grants to jump start creation of longitudinal data systems connecting higher education and workforce data. Progress in this area has been slow[10] for the commonwealth. With over 390 higher education institutions there is one additional challenge the commonwealth faces with the traditional workforce-data linkage. It cannot track students who leave the state after graduating. This is important since brain-drain is an old problem for Pennsylvania for institutions outside the public sector (17.5 percent as gross brain drain for 2017[11]).

Summing up, higher education today faces a host of challenges from quality to cost. Lack of a data infrastructure that can analyze the connection between instructional programs and the workforce continues to be a major challenge. A solution (to this challenge) is the creation of an agile framework that can be accessed by key stakeholders. An infrastructure that helps convey the linkages between education and occupation in real time.

The Intervention

In light of the challenges and a growing need for real time workforce outcome of alumni, this study presents a technology enabled solution[12] that paves the way for educators to continuously have answer to the question, *how are students and learners doing in the world of work?* It is not meant to replace the traditional data linkage which is an asset for states. It presents a democratic process that can be used by all higher education institutions to understand the workforce outcomes of graduates (irrespective of their destination post-graduation) and shines a bright light on the education-occupation connectivity of higher education programs.

Creating a ***Customized Alumni Database*** outside of State Bureaucracy

*Pennsylvania's State System of Higher Education i*s the public four-year sector for the Keystone State. With fourteen universities spread all across the beautiful state, it provides access to affordable education to Pennsylvanians. With over 100,000 students, most of the students attending the fourteen universities are residents of the state. After graduation most stay in Pennsylvania. Similar to other higher education enterprises, student level data are collected and maintained at the System for students from fourteen universities. *Alumni Insight* is a proprietary database of professional profiles maintained by a private organization, Emsi/Strada. It provides employment outcome data gathered from over 100 million professional profiles and resumes in the United States[13]. Approximately three million profiles exist in this database of profiles referencing Pennsylvania educational institutions and over 400,000 referencing one of the fourteen State System universities. A subset of this, 350,000, were utilized for this project.

To create the ***customized alumni database*** utilized for this study, professional profiles[14] of State System graduates (resumes 'scraped' from professional boards such as LinkedIn, CareerBuilder, Indeed, and GitHub) were linked with student records (from Pennsylvania's State System Student Information System). The main asset of professional boards is the huge number of profiles with detailed data on employer name, job title, job description, and skills. This data is what makes the site rich and useful for analysts. Data scraping refers to collecting topic-specific data for the purpose of research/analysis. While both job boards and data scraping are not new, the use of this big data to create customized databases that link workforce outcomes with student profiles at an institution/system level is. Appendix 2 shows examples of scraped professional profiles. The linkage resulted in a ***customized alumni database***[15] with education data (student, geographic, and instructional program profile) and workforce data (employer, geographic, occupation, job title, and skills profile). Required legal standards were adhered to link data and analyze the data. This unique ***customized alumni database*** demonstrated the capacity to comment on the labor market outcomes of State System alumni. It produced a process to link student records to their workforce outcomes outside of state bureaucracy and revealed meaningful ways to answer the question: *how does education link to work (occupation and skills) specific to an institution, program, and geography*? Most importantly, it demonstrated that with the right investment, internal legal review, and data design-framing, higher education institutions have agency to realize the benefits of linked data outside of state bureaucracy.

Result[16] from the *Customized Alumni Database*

Three promising areas emerged from the data linkage[17]:

1. **Institution-specific Classification of Instructional Program and Standard Occupation Classification (CIP-SOC) Crosswalk**

The CIP to SOC crosswalk is a tool to relate programs to data about occupations and labor market. A crosswalk is maintained at the federal level[18] utilized by the higher education enterprise. Some states (like Florida and Pennsylvania), have created systems to update the crosswalk to their regional economy.

With the ***customized alumni database***, it was possible to create a crosswalk specific to universities driven by the actual outcomes of their graduates. It was possible to create CIP-to-SOC maps at varying scales e.g. for programs areas (see example on Environmental Studies, Appendix 2), or the institution as a whole. The project generated custom CIP to SOC crosswalks for all fourteen System universities (refer Appendix 3 for an example). An institution specific CIP-SOC crosswalk is a key tool for viewing the real time connections of education programs with the labor market. It is especially important for liberal arts and humanities. The federally maintained taxonomy from degrees to jobs (CIP to SOC) works reasonably well for academic programs that have a direct connection to the workforce (such as STEM programs). General programs have historically suffered creating national dialog on limited job opportunities for Bachelor's degrees in Liberal Arts, General Studies. Empowering the higher education enterprise (especially small colleges and universities) with the ability to develop a taxonomy that is true to their landscape is valuable.

2. **Real-time Career Trajectories**

Utilizing the ***customized alumni database***, it was possible to identify career trajectories of alumni. An example of the career trajectory specific to Computer and Information Science is illustrated in Appendix 4. This enhancement helps faculty and staff view the diverse career paths the education (they provide) are creating specific to their institutions. Additionally, it offers the opportunity to make course corrections to strengthen and better align them to the labor market.

3. Competitive Advantage (Wage) by Program[19]

Location Quotient calculation is a technique for gauging the relative concentration or specialization and measures the competitive advantage of a sector. This concept was utilized (by this study) to measure the competitive advantage of instructional programs by the wages earned by graduates. The focus of Location Quotient calculations is primarily the identification of those program clusters that have an LQ exceeding 1.00. Any figure above 1.00 implies a competitive advantage. The ***customized alumni database*** produced detailed education program level data connected to occupations. Linking those occupations (at the detailed program level) to wage data, it was possible to measure the competitive advantage of programs specific to universities. Appendix 5 provides details on the use of the ***customized alumni database*** to gain insights on program areas that are performing well and not-so-well in terms of wage based competitive advantage. Use of wage data to illustrate the comparative advantage of an education program (as done in this study) mitigates issues associated with use of absolute wage data (in the traditional linkage).

Impact

In late 2012, the federal-state Workforce Information Council established the Administrative Wage Record Enhancement Study Group to examine the feasibility of adding variables to wage record reports that employers submit as part of state Unemployment Insurance (UI) Programs. This came from the recognition that the traditional workforce-education connection established through wage records lacked a key body of information***, occupations,*** concentrating solely on wages and industry attachments of graduates[20]. Implementing the design recommended by the Study Group required significant investment[21]. As a result, there was little appetite at the state level to increase employer work-load by adding variables to the wage record. As efforts continue at the national and state level to create meaningful longitudinal data systems and enhance them with occupational information, leveraging strategic partnerships to create ***customized alumni databases*** can significantly increase the capacity of the higher education enterprise (both public and private) to view career trajectories of their alumni and use it to create a feedback loop of consumer insight in real time. Designing such a ***customized alumni database*** has costs and requires data-governance but, at a much smaller scale (as compared to interagency memorandums of understanding). And, the process adds missing dimensions to the traditional workforce education data linkage that has (so far) been dominated by absolute wage outcomes. It can be replicated by other higher education sectors and institutions to understand learner outcomes and can also be utilized to refine the federally maintained education-occupation taxonomy.

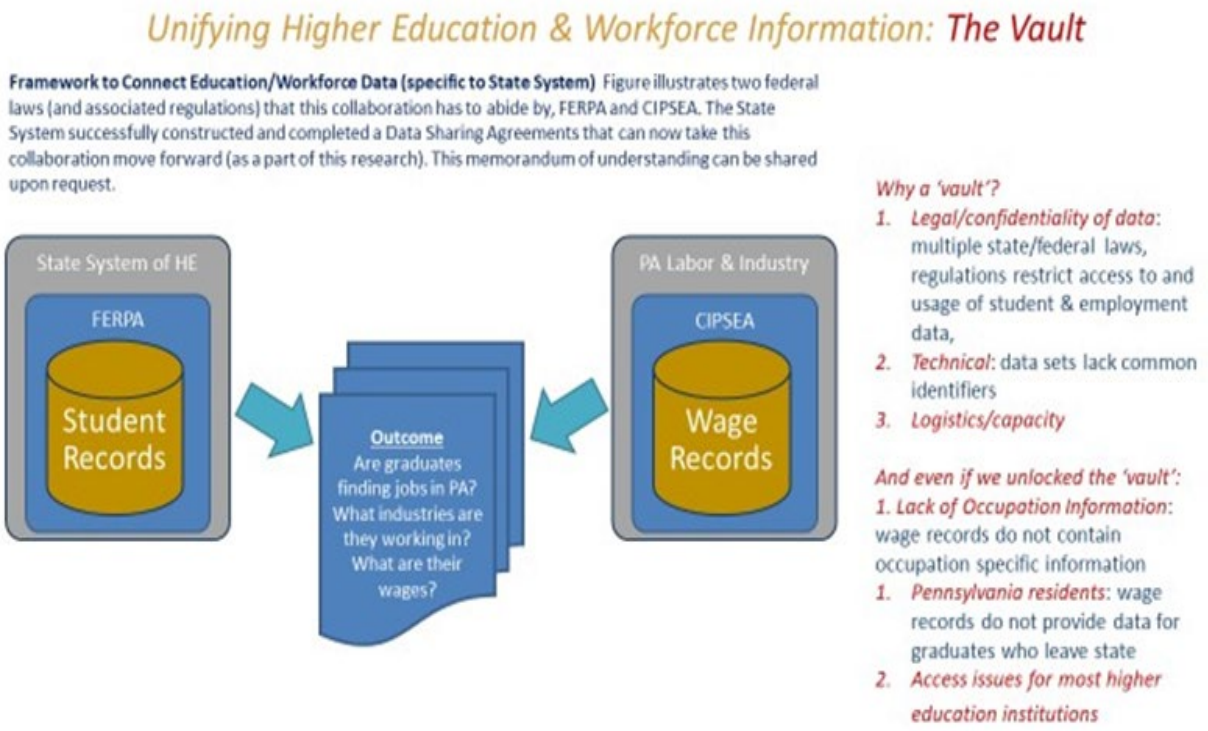
Lessons Learned

My familiarity with the intersection of policy, technology, and data-linkage stems from a career spent in designing strategic use of data for both workforce and higher education. I have been an active participant in various data innovations in Pennsylvania and the nation specific to creation of longitudinal data systems including the 2012 Kauffman Study Group [22] and have experienced challenges in creating, maintaining, and providing access to the traditional education-workforce linked database. For the first time in a lengthy career spent 'unlocking the vault', it was exhilarating to have the agency to design-frame an infrastructure that had capacity to analyze career trajectories, skills gap, education-bundles, mobility patterns, and job trajectories[23] of alumni. It is heartening to engage in labor market conversations with faculty and provide data that they can use strategically to take action. As we prepare to analyze the education-occupation connectivity in the world of artificial intelligence, on the developmental side, data accessibility will pose significant hurdles. Commercially viable data that have societal applications will be key. This study illustrates one example of such an application. An application that offers a common framework of language for key stakeholders in the current competency-skills gap narrative, educators and employers, and empowers them to be active partners contributing to learner success and regional stewardship.

Appendices and Endnotes

Appendix 1a and 1b: Creating Education-Workforce Connectivity, the two ways

1a. Traditional Education-Workforce Linkage, an Overview of the Framework and Challenges



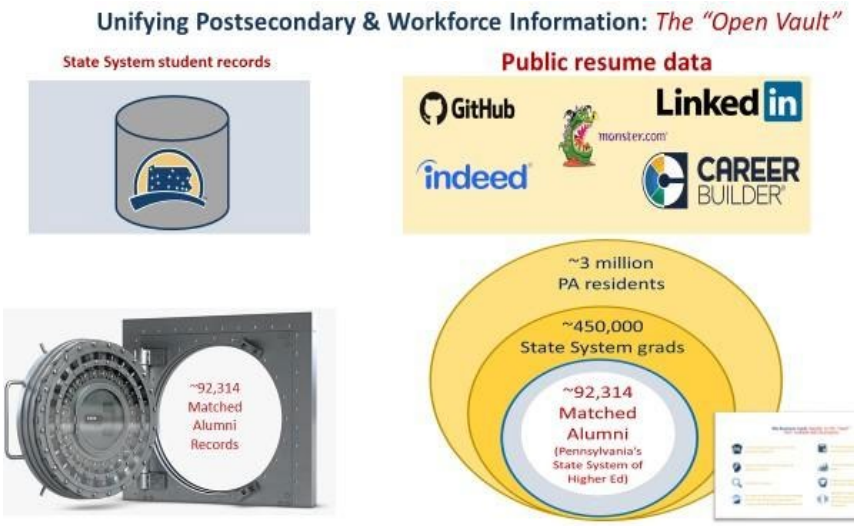
CIPSEA or the Confidential Information Protection and Statistical Efficiency Act, is a United States federal law that establishes confidentiality protections for information collected for statistical purposes by U.S. statistical agencies such as wage records.

FERPA or the Family Educational Rights and Privacy Act is a federal law that establishes confidentially protections in sharing of student data and affords parents the right to have access to their children's education records, the right to seek to have the records amended, and the right to have some control over the disclosure of personally identifiable information from the education records.

The table (below) illustrates an example of how this data linkage could look at the data element level.

University	Instructional Program	Number of Graduates	1 Year After Graduation		10 Years After Graduation	
			Number Employed	Average Wages	Number Employed	Average Wages
xxxxxxxx	Agriculture	50	40	$51,234	46	$55,678
	Architecture	75	62	$56,789	69	$59,123
xxxxxxxx	Agriculture	80	65	$51,234	71	$55,678
	Architecture	100	83	$56,789	89	$59,123

1b.Overview of the Customized Alumni Database



Note: of the 450,000 State System records, approximately 350,000 full records were used to match with State System's Student Information System to arrive at a ***customized alumni database*** of linked records (92,314). This linked database contained both education and workforce records of the alumni.

Data extraction and linkage was conducted in November 2017. Analysis was conducted in phases through August 2018. From August 2018 to April 2019, analysis was restricted to one of the State System's 14 university, Shippensburg University of Pennsylvania. All examples and the data linkage can be replicated.

Appendix 2: Example from Alumni Insight, Professional Profiles Scraped from Pennsylvania Profiles (*Professional Name and Employer Name Redacted)*

Name Employer					
Given Job Title	district sales manager	director, human resources	executive director of public relations	purchasing agent	account manager
Job Description					
Standard Job Title	Sales Manager (Management)	Human Resources (HR) Director	Director of Public Relations (Management)	Purchasing Agent	Account Manager (Sales and Related)
Alum Of	Shippensburg University of Pennsylvania	Shippensburg University of Pennsylvania	Shippensburg University of Pennsylvania, Hagerstown Community College	Shippensburg Universit	Shippensburg University of Pennsylvania
Skills	Goal Oriented, Sales Management, Business Intelligence Tools, New Business Development, Microsoft Office, Data Intelligence, Managerial Economics, Logistics Management, Leadership Development, Strategic Planning, Internal Communications, Social Entrepreneurship, Event Management, Event Planning, Business Development, Sales, Public Relations, Team Building, Management, Public Speaking, Leadership, Communications, Tournaments, Customer Service, Operations, Entrepreneurship, Economics, Microsoft Word, Data Analysis, Microsoft Excel, Wholesaling, Business Intelligence	Industrial Relations, Workforce Planning, Compensation Management, Human Capital, Labor Relations, Strategic Planning, Organizational Effectiveness, Project Management Software, Resource Management, SAP Business Suiteing, SAP NetWeaver, SAP Human Resource Management Software, Human Resources Information System (HRIS), Team Building, Customer Relationship Management, Warehousing, Information Systems, SAP NetWeaver Business Warehouse, Negotiation, Human Resource Management, Management, Certified Urologic Physician's Assistant, Data Warehousing, Project Management	Strategic Communication, Microsoft Office, Marketing Communications, Video Editing, Press Releases, Order Fulfillment, Office Assistant, Marketing Strategies, Marketing Operations, Final Cut Pro, Communications Training, Event Planning, Digital Photography, Creative Writing, Copywriting, Journalism, Public Relations, Copy Editing, AP Stylebook, Adobe InDesign, Editing, Writing, Adobe Creative Suite, Public Speaking, Adobe Photoshop, Social Networks, Annual Reports, Presentations.	Continuous Improvement Process, Management Process, Microsoft Office, Inventory Turnover, Operations Management, PCB Assembly Services, Planning Tools, Displacement (Fluid), Procurement Management, Product Management, Resource Planning, Centrifugal Pumps, Business Process Improvement, Building Services Engineering, Standards Organization, Strategic Planning, Material Handling, Machining, Six Sigma Methodology, Forging, Procurement, New Product Development, Molding, Purchasing, Welding, Fabrication, Supply Chain Management, International Traffic In	Critical Thinking, Event Planning, Inventory Management, Microsoft Office, Time Management, Management, Sales, Social Networks, Customer Service, Leadership, Sanitation, Food Safety, Data Entry, Research

Appendix 3: Comparing Federal and the Customized Alumni Database for a program of study, Environmental Studies at Shippensburg University, Pennsylvania's State System of Higher Education

Consider the CIP title *Environmental Studies* below for a university within the State System. The national CIP- SOC federally maintained taxonomy illustrates two occupational outcome for Environmental Studies generally in U.S. Utilizing the ***customized alumni database*** provided over 35 connections of real time career trajectories of graduates from Environmental Studies for this university.

Federal CIP SOC Crosswalk

CIP	SOC	Education Attainment
Environmental Studies.	Environmental Science Teachers, Postsecondary	Requires a Bachelor's or Higher
Environmental Studies.	Environmental Scientists and Specialists, Including Health	Requires a Bachelor's or Higher

CIP SOC Created from ***Customized Alumni Database***

Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	29	29-9011	Occupational Health and Safety
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	37	37-3011	Landscaping and Groundskeepi
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	43	43-6014	Secretaries and Administrative
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	47	47-1011	First-Line Supervisors of Constr
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	47	47-2061	Construction Laborers
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	11	11-9021	Construction Managers
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	11	11-1021	General and Operations Manag
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	13	13-1199	Business Operations Specialists
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	15	15-1132	Software Developers, Applicati
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	17	17-3022	Civil Engineering Technicians
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	17	17-3025	Environmental Engineering Tec
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	19	19-1031	Conservation Scientists
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	39	39-9032	Recreation Workers
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	41	41-4011	Sales Representatives, Wholesa
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	41	41-4012	Sales Representatives, Wholesa
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	41	41-9022	Real Estate Sales Agents
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	43	43-5081	Stock Clerks and Order Fillers
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	51	51-9111	Packaging and Filling Machine
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	11	11-2011	Advertising and Promotions Ma
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	11	11-3071	Transportation, Storage, and Di
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	11	11-3121	Human Resources Managers
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	13	13-1051	Cost Estimators
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	13	13-1151	Training and Development Spe
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	15	15-1134	Web Developers
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	15	15-2031	Operations Research Analysts
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	15	15-2041	Statisticians
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	17	17-2051	Civil Engineers
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	17	17-2111	Health and Safety Engineers, Ex
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	17	17-2199	Engineers, All Other
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	17	17-3023	Electrical and Electronics Engin
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	17	17-3026	Industrial Engineering Technicia
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	17	17-3031	Surveying and Mapping Technic
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	19	19-1021	Biochemists and Biophysicists
Bachelor's	03	03.0103	ENVIRONMENTAL STUDIES	19	19-3051	Urban and Regional Planners

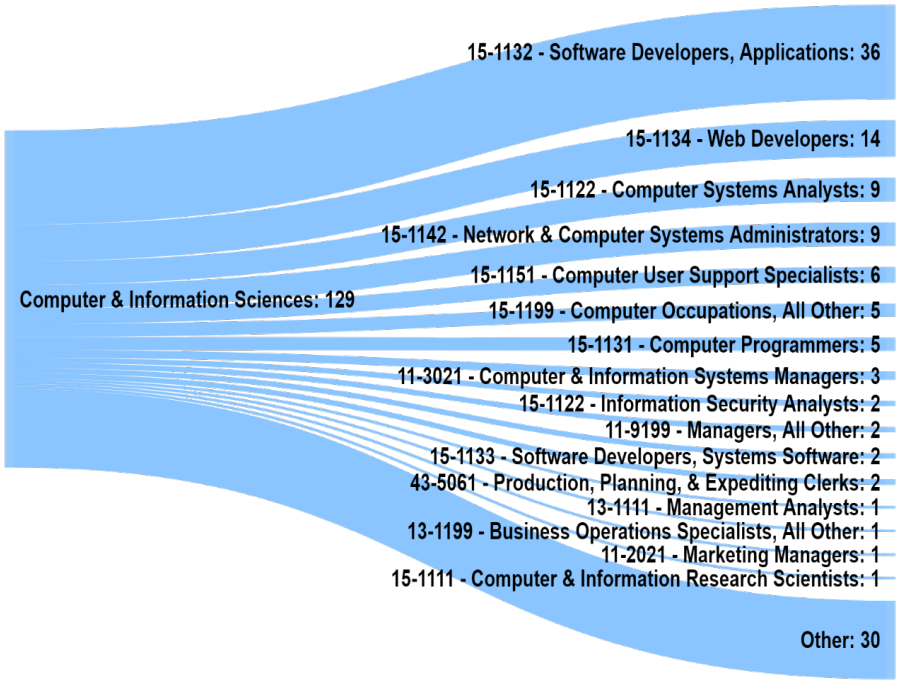
Appendix 4: Comparing Federal and the *Customized Alumni Database* for a program of study Computer and Information Sciences, at Shippensburg University, Pennsylvania's State System of Higher Education

The first table shows the education-occupation crosswalk from the federal government's taxonomy while the Sanky chart traces the actual career trajectory of recent Shippensburg graduates.

Federal Crosswalk

Computer and Information Sciences, General	
11-3021.00	Computer and Information Systems Managers
15-1111.00	Computer and Information Research Scientists
15-1121.01	Informatics Nurse Specialists
15-1141.00	Database Administrators
15-1142.00	Network and Computer Systems Administrators
15-1143.00	Computer Network Architects
15-1143.01	Telecommunications Engineering Specialists
15-1199.01	Software Quality Assurance Engineers and Testers
15-1199.03	Web Administrators
25-1021.00	Computer Science Teachers, Postsecondary

Crosswalk from ***Customized Alumni Database*** for graduates from Shippensburg University, Pennsylvania's State System of Higher Education

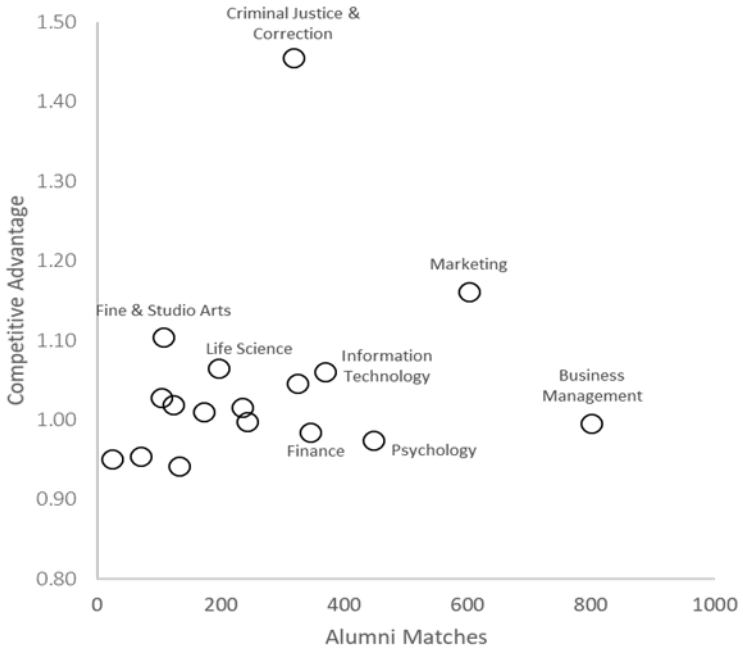


Appendix 4: Using the Customized Alumni Database to Calculate Competitive Advantage

For the purpose of this project (and to limit scope), occupations and instructional programs were clustered into "Program Areas" and into "Career Areas." Outcomes were quantified in terms of earnings, based on occupations and their corresponding Occupational Employment Statistics wage estimates. The OES program from Bureau of Labor Statistics collects data on wage and salary workers in order to produce employment and wage estimates for about 800 occupations.

The OES survey is a federal-state cooperative program between the Bureau of Labor Statistics (BLS) and State Workforce Agencies (SWAs). Using the competitive advantages calculated through benchmarking in concert with the underlying career trajectories (CIP-to-SOC maps), deeper analysis of programmatic and institutional outcomes become possible. This includes such insights such program areas performing well and not-so-well in terms of competitive advantage; the proportion of alumni earning a minimum salary, or the average earnings for all alumni; and the top career outcomes for alumni, in terms of their occupations and associated earnings. An example follows - analysis that utilized linked data of professional Shippensburg profiles with their individual student records and the OES wage data.

Shippensburg University, Pennsylvania's State System of Higher Education, Programs Outcomes



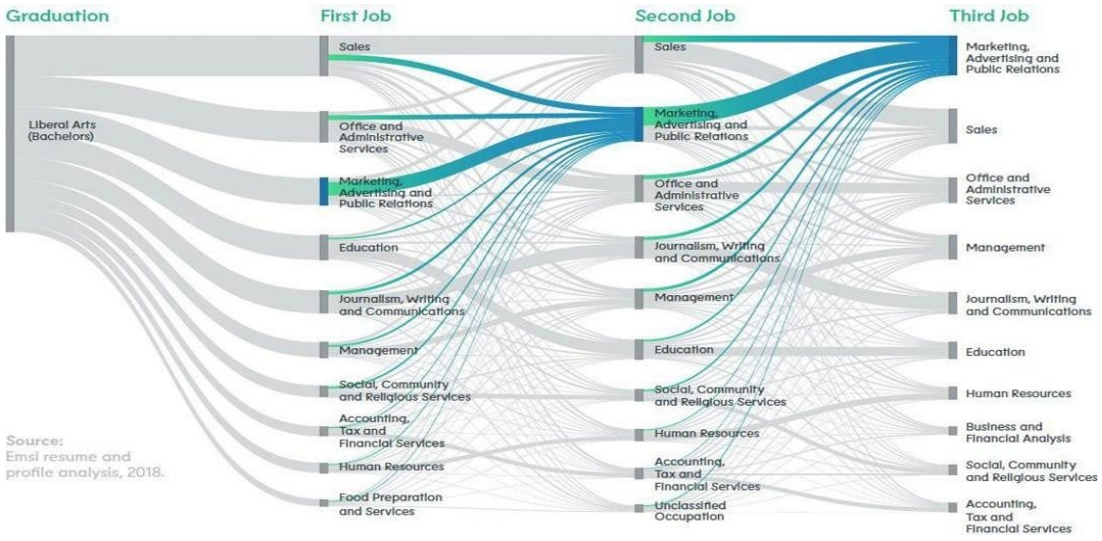
Programmatic Insights: e.g. Information Technology

- 2nd ranked PA State System Program Area: Adv = 1.11 (10th in matched alumni)
- 2,130 matched alumni from 7 CIPs & 12 institutions
  - Top matched CIP: 11.0101 – Computer & Information Sciences, General (71%)
  - Top matched institutions: Shippensburg (17%)
- 70% alumni earning > $60,000
- Average salary $69,706
- Top earners:
  - 33% in Software Development & Programming ($95,650)
  - 16% in IT Networks & Systems ($94,431)
  - 4% in Business & Operations Analysis ($74,285)
  - 4% in Computer Science ($88,849)
- Top career: 33% in Software Development & Programming ($95,650)

$$\text{Competitive Advantage} = \frac{\text{Shippensburg Earnings}}{\text{Avg US Earnings}}$$

Note: Wages were averaged by the weighted distribution of occupational outcomes. Average program outcomes ($) were calculated for the institution and its constituent programs, and separately, for the nation; Competitive Advantage is defined as the ratio of these earnings. The difference in earnings reflects differences in the career outcomes for an institutional program relative to typical (national) outcomes for a graduate from a similar program.

Appendix 5: Example of a job-trajectory possible with the *customized alumni database* (this was not a part of this study)



Appendix 6: Labor Market Sources used by this study

Traditional labor market information includes data related to unemployment, wages, educational attainment, population and demographics, and information pertaining to workers, industry, and the economy. Technology enabled labor market data include data related to job postings by occupation, employers, region and advertised skills. Alumni Insight is a part of this newer labor market data set and includes professional profiles of graduates. To conduct research associated with this study primarily three data sets were utilized including:

Economic Modeling Specialists, EMSI: EMSI provides proprietary labor market data on the total U.S. workforce. Working through CareerBuilder and Strada, EMSI has access to a large amount of worker profiles and salary data points utilized in its database.

Bureau of Labor Statistics of the U.S. Department of Labor: The **Bureau of Labor Statistics** (BLS) of the U.S. Department of **Labor** is the principal federal agency responsible for measuring **labor** market activity, working conditions, and price changes in the economy.

United States Census Bureau: is a principal agency **of the U.S.** Federal Statistical System, responsible for producing data about the American people **and** economy. A variety of data tools are available through the U.S. Census including American Fact Finder, On the Map, and Local Employment Dynamics. The Local Employment Dynamics provide Quarterly Workforce Indicators (QWI) including employment, turnover, and earnings data by industry.

Appendix 7: A Labor Market Data Brief

A significant portion of my career has been spent on projects that are at the intersection of higher education, labor market information, and economic development. One project entailed the creation of the 'Data Brief' I share next. My hope in ***sharing this forward*** is to help individuals who are exploring the power of this 'intersection'.

This White House Data Brief contains publicly available data sets related to job availability, employment levels, skills needs, education, and other related topics. The data sets come from the Departments of Agriculture, Commerce, Education, Labor, Health and Human Services, and Veterans Affairs.

White House Data Brief

I. Employment data

Department of Labor: Job Openings and Labor Turnover Survey (JOLTS)
http://www.bls.gov/jlt
JOLTS, from the Bureau of Labor Statistics, produces data on job openings, hires, and separations. It provides an assessment of the availability of unfilled jobs and information to help assess the presence or extent of labor shortages in the United States.

Department of Health and Human Services: National Directory of New Hires (NDNH)
http://www.acf.hhs.gov/programs/css/resource/national-directory-of-new-hires-information-for-families
The NDNH is a national database of wage and employment information to assist state child support agencies in locating noncustodial parents to establish their paternity and enforce child support obligations. Information on individuals includes New Hire, Quarterly Wage (Unemployment Insurance Wage records), and Unemployment Insurance benefits.

Department of Labor: National Compensation Survey (NCS)
http://www.bls.gov/ncs
The NCS provides measures of occupational wages, employment cost trends, benefit incidence, and detailed plan provisions. It does this for both metropolitan and non-metropolitan areas, broad geographic regions, and on a national basis.

Department of Labor: Employment, Hours, and Earnings from the Current Employment Statistics Survey (State and Metro Area)
http://www.bls.gov/sae
This monthly survey provides detailed industry data on employment, hours, and earnings of workers on nonfarm payrolls for all 50 states, the District of Columbia, Puerto Rico, the Virgin Islands, and over 300 metropolitan areas and divisions. It includes about 150,000 businesses and government agencies, representing approximately 390,000 individual worksites

Department of Labor: Local Area Unemployment Statistics (LAUS)
http://www.bls.gov/lau
LAUS produces monthly and annual employment, unemployment, and labor force data by place of residence for Census regions and divisions, States, counties, metropolitan areas, and many cities.

Department of Labor: Labor Force Statistics from the Current Population Survey (CPS)
http://data.bls.gov/cgi-bin/surveymost?ln
CPS data come from a monthly survey of households and provides a comprehensive body of data on the labor force, including employment, unemployment, and persons not in the labor force.

Department of Labor: The Employment Projections (EP) Program
http://www.bls.gov/emp
The EP program develops information about the labor market for the Nation as a whole for 10 years in the future.

Department of Labor: Standard Occupational Classification (SOC) System
http://www.bls.gov/soc
The SOC System is used by Federal statistical agencies to classify workers into 820 occupational categories for the purpose of collecting, calculating, or disseminating data.
Department of Labor: Wage Record Interchange System (WRIS)
http://www.doleta.gov/performance/wris.cfm

WRIS helps state workforce program performance agencies track the wages of citizens in workforce investment programs in one state who secure employment in another. This data can be used by states to recognize the effectiveness of their workforce investment programs and by third party developers to create tools to measure the "cost/benefit" of various job training or college degree programs.

Department of Labor: Business Employment Dynamics
http://www.bls.gov/bdm
This data consists of gross job gains and gross job losses statistics from 1992 forward. It provides a picture of the dynamic state of the labor market. Data comes from the Quarterly Census of Employment and Wages, or ES-202, program.

Education and training data

Department of Labor: Education and Training Data
http://www.bls.gov/emp/ep_education_training_system.htm
Education and training data from the Bureau of Labor Statistics has information about education and training requirements for hundreds of occupations, which can be used to help estimate the education and training needs for the labor force as a whole, the level of education achieved by current workers in particular occupations, as well as the outlook for the future education and training needs.

Department of Education: State Dropout and Completion Data
http://nces.ed.gov/ccd/drpcompstatelvl.asp
This data reports the number of dropouts from grades 9 - 12 and the relevant event dropout rates to report the numbers of high school diploma recipients, other high school completers, and the relevant Averaged Freshman Graduation Rate (AFGR).

Department of Education: Vocational Rehabilitation (VR) Programs
https://explore.data.gov/Labor-Force-Employment-and-Earnings/Vocational-Rehabilitation-Programs-Core-Data/ujvb-c756
Core Data allows you to view tables and charts with key data and download basic Vocational Rehabilitation (VR) program tables for each state agency.

Department of Education: Career and Technical Education Act, Carl D. Perkins - Consolidation Annual Reports
State Websites
These reports provide enrollment and performance data under secondary and postsecondary career and technical education administered as part of the Carl D. Perkins Career and Technical Education Act of 2006. Commonly known as "Perkins", these programs serve as the principal source of federal funding to states for the improvement of secondary and postsecondary career and technical education programs.

General economic data at local level

US Department of Commerce: Small Area Income & Poverty Estimates (SAIPE)
http://www.census.gov/did/www/saipe
SAIPE provides the best single-year estimates of median household income and poverty. It provides the most accurate sub-national estimates of poverty for counties and school districts to help determine resource needs.

US Department of Commerce: American Community Survey (ACS)
https://www.census.gov/acs/www/guidance_for_data_users/guidance_main
The American Community Survey (ACS) is an ongoing survey that provides data every year — giving communities the current information they need to plan investments and services. Survey questions include information on demographics, employment, income and benefits, education, veteran status, and others.

US Department of Commerce: County Business Patterns
http://www.census.gov/econ/cbp/download
The County Business Patterns is an annual series that provides economic data by industry and is particularly useful for studying the economic activity of small areas. It can also be used for analyzing economic changes over time. For businesses, is helpful for analyzing market potential, measuring the effectiveness of sales and advertising programs, setting sales quotas, and developing budgets.

US Department of Commerce: American Housing Survey (AHS)
http://www.census.gov/programs-surveys/ahs
AHS is the largest, regular national housing sample survey in the United States and includes data on apartments, single-family homes, mobile homes, vacant homes, family composition, income, housing and neighborhood quality, housing costs, equipment, fuels, size of housing unit, and recent movers.

Department of Agriculture: County-Level Data Sets
http://www.ers.usda.gov/data-products/county-level-data-sets.aspx#.U5m8Q_ldXFk
These County-Level Data Sets include socioeconomic indicators like the poverty rate, population change, unemployment rate, and education levels. It compiles them into maps and a data display/download application that allows users to identify and compare states and counties on these indicators.

Department of Agriculture: USDA Economic Research Service Data Products
http://www.ers.usda.gov/data-products.aspx#.U5m8gfldXFk
These Data Products include information on the rural economy and population, farm economy, and much more.

Existing online jobs/ business tools

Department of Labor: O*NET
http://www.onetonline.org
O*NET has detailed information for use by job seekers, workforce development and HR professionals, students, researchers, and more. O*NET provides detailed descriptions of jobs that link to national and state jobs information, future career possibilities, and salary and wage information to support career exploration and planning.

BusinessUSA: Resources to Help Grow Your Business
http://business.usa.gov
BusinessUSA provides access to training and staff development to help Americans start and grow their businesses. It also provides other resources, like help with exports and access to financing.

Department of Labor: Occupational Outlook Handbook
http://www.bls.gov/ooh
The Occupational Outlook Handbook is a nationally recognized source of career information, designed to provide valuable assistance to individuals making decisions about their future work lives. The Handbook is revised every two years.

Department of Labor: American Job Center
http://jobcenter.usa.gov
On this site, one can search for available jobs with specific sections for youth and veterans. One can quickly locate and get directions to the nearest American Job Center via either Zip code or city. The site also returns information about services available by location. It includes both comprehensive and affiliate American Job Centers' name, location (with map and driving directions), and contact information for veterans representatives when available.

Department of Labor: Get Certification Information
One can search a database of occupational certification information by occupation, industry, or keyword. Results include certification name, occupation, industry, certifying organization's contact information, requirement information, and renewal information.

Department of Labor: Build an Occupational Profile
One can access detailed occupation data for more than 800 occupations. Information is derived from the Occupational Information Network O*NET database and includes:
•knowledge, skills, and abilities by occupation
•tasks and activities by occupation
•common tools and technology used in select occupations

Veterans' data

Veterans Affairs: Labor Force Participation Rates of Veterans
http://www.va.gov/vetdata/Report.asp
This data compares trends in labor force participation rates between veterans and non-veterans using data from 10 years of the Current Population Survey.

Veteran Affairs: Unemployment Rates of Veterans
http://www.va.gov/vetdata/Report.asp
This data compares trends in unemployment rates between veterans and non-veterans using data from 10 years of the Current Population Survey.

Open source tools, resources, and APIs

NLX Jobs Feed (API)
The jobs feed currently stands at over 1.6 million currently available, un-duplicated jobs. The NLX includes jobs from State Job Banks, USAjobs.gov, and from DirectEmployers members and non-members.
NLX Historical Jobs Data- 3 months (Real-Time LMI)
Includes jobs data from last three months: March, April, and May.

Department of Education: National Center for Education Statistics
http://nces.ed.gov/ipeds/datacenter
Provides a great deal of data on educational institutions.

Endnotes

[1] Quantitative and qualitative data and analysis related to employment, workforce, and education comprise a set of intelligence loosely termed as labor market data.
[2] Examples include NSSE (National Survey of Student Engagement), Gallup-Strada poll, First Destination Surveys etc.
[3] https://collegescorecard.ed.gov/
[4] State unemployment insurance (UI) data systems is the traditionally used source of data on earnings and employment outcomes. The UI Program is federally mandated and regulated; rules regarding the terms under which UI data can be made available are established separately by each state. Among those states that have chosen to make these data available, students' social security numbers are matched with records from the state UI database after securing data-sharing agreements that include significant layers of protection to ensure that information cannot be used to identify outcomes for specific individuals. Through this match, analyses can show individuals' employment status, industry, and wages prior to and after enrolling in postsecondary education, but only for students who work in the same state in which the college they attended is located.
[5] State unemployment insurance (UI) data systems is the traditionally used source of data on earnings and employment outcomes. The UI Program is federally mandated and regulated; rules regarding the terms under which UI data can be made available are established separately by each state. Among those states that have chosen to make these data available, students' social security numbers are matched with records

from the state UI database after securing data-sharing agreements that include significant layers of protection to ensure that information cannot be used to identify outcomes for specific individuals. Through this match, analyses can show individuals' employment status, industry, and wages prior to and after enrolling in postsecondary education, but only for students who work in the same state in which the college they attended is located.

[6] Student Longitudinal Data System and Workforce Data Quality System grants

[7] It can be realized by merging student data (from higher education institutions) with employment information of graduates (from state administrative wage records). Linking these distinct data-sets across the education and workforce agencies can produce this valuable information.

[8] [8] Florida has built one of the nation's most comprehensive statewide data systems to track school-to-work outcomes. This robust data set is utilized in numerous ways, from creating specialized programs like dropout prevention to guiding legislative initiatives.

[8] A pilot program by the Western Interstate Commission for Higher Education (WICHE) connects education and wage data across states to evaluate program performance and track labor migration across state borders.

[8] The University of Texas System recently made groundbreaking progress by signing a 10-year agreement with the U.S. Census Bureau to provide salary and jobs data of UT graduates across the nation.

[9] Arriving at a memorandum of understanding to acquire granular data (de-identified) between the higher education institution and the agency that houses employment wage records takes significant time even for states that have longitudinal data linkages. For the purpose of this project a MOU was arrived at (signed by all parties) to start exchange of information at an aggregate level. The process took almost nine months.

[10] In 2012, PA received the U.S. Department of Labor (DOL) Workforce Data Quality Initiative funds to develop a state workforce longitudinal administrative database. This system includes information on programs that provide training and employment services and will be linked longitudinally at the individual level to allow for analysis leading to enhanced opportunity for program evaluation and better information for customers and stakeholders of the workforce system. This effort was in support of a parallel and much larger effort, the Statewide Longitudinal Data System (SLDS) grants administered by the U.S. Department of Education. These two programs encourage the development of state education and workforce longitudinal administrative databases. Ultimately, databases developed through WDQI should be linked to education data at the individual level. The Keystone state has 392 higher educational institutions. The state has not been successful in creating a system that is inclusive of the largest players as of writing of this report. Even when created, it is unlikely individual institutions would have agility to access program level data without going through significant effort.

[11] https://www.jec.senate.gov/public/index.cfm/republicans/2019/4/losing-our-minds-brain-drain-across-the-united-states#_edn38

[12] This research reflects a public private partnership with EMSI. Sincere thanks and appreciation for Luke Jankovic, Yustina Saleh, James Prado, and Stuart Muller (EMSI) for brainstorming through this great challenge that started in 2016 with a humble goal of using alumni data to creation of the linked data- base in 2017. Special mention to Stuart for being key research partner

[13] These numbers reflect the data extracted at the start of this project in Fall 2017. The full study was done based on the data extracted at this point. Today, August 2019, Alumni Insight has over 124 million profiles.

[14] Alumni Insight provides essential employment outcome data gathered from more than 100 million professional profiles and resumes in the United States. The database contains nearly three million profiles for Pennsylvania residents. State System institutions were referenced in 450,000 of those professional (alumni) profiles. Pennsylvania's State System of Higher Education is the public four year comprised of 14 individual universities spread across the commonwealth. The State System is the largest producers of knowledge workers in the commonwealth.

[15] For this study, p***rofessional profiles*** (at an individual level) were connected back to the university ***student profiles*** that contained detailed program level data on the individual education experience. Common data elements used for the linkage included student/professional first name and last name. The match rate for this initial project was 35 percent.

[16] This research reflects a public private partnership with EMSI. Sincere thanks and appreciation for Luke Jankovic, Yustina Saleh, James Prado, and Stuart Muller (EMSI) for brainstorming through this great challenge that started in 2016 with a humble goal of using alumni data to creation of the linked data- base

[17] For the purpose of this project, a clustering taxonomy was used to group CIPs into related "Program Areas" and SOCs into "Career Areas." Outcomes were quantified in terms of wages, based on occupations and their corresponding Occupational Employment Statistics wage estimates, averaged by the weighted distribution of occupational outcomes. Average program outcomes ($) are calculated for the institution and its constituent programs, and separately, for the nation; Competitive Advantage is defined as the ratio of these earnings. The difference in earnings reflects differences in the career outcomes for an institutional program relative to typical (national) outcomes for a graduate from a similar program.

[18] By the Bureau of Labor Statistics and the National Center for Education Statistics

[19] For the purpose of this project, a clustering taxonomy was used to group CIPs into related "Program Areas" and SOCs into "Career Areas." Outcomes were quantified in terms of wages, based on occupations and their corresponding Occupational Employment Statistics wage estimates, averaged by the weighted distribution of occupational outcomes. Average program outcomes ($) are calculated for the institution and its constituent programs, and separately, for the nation; Competitive Advantage is defined as the ratio of these earnings. The difference in earnings reflects differences in the career outcomes for an institutional program relative to typical (national) outcomes for a graduate from a similar program.

[20] https://www.bls.gov/advisory/bloc/enhancing-unemployment-insurance-wage-records.pdf

[21] The Study Group explored enhancing administrative records with occupation and geographic information to help state and local education and training program planning and accountability, economic analysis, career planning, and workforce program administration

[22] Examples include membership in the U.S Workforce Information Advisory Council of ETA, Pennsylvania's Workforce Data Quality Initiative team, Pennsylvania SIMS team.

[23] Source: 2019 ACE Presentation Rob Sentz, EMSI. This was a part of a joint presentation by Ben Wildavsky, Senior Vice President, National Engagement, Strada Education Network; Rob Sentz, Chief Innovation Officer, Emsi; Dr.Sue Mukherjee, Chief Strategy Office, Shippensburg University, "How Consumer Insights and Outcomes Data Can Increase Higher Education Value" (March 17, 2019)