



Developing Analytical Capacity and Techniques for Large-Scale Workforce Data: An Annotated Bibliography

Prepared for the
Employment and Training Administration
Office of Policy Development and Research
Division of Research and Evaluation
Edited by Margaret Patterson, PhD, Safal Partners, LLCs

Acknowledgements

This annotated bibliography is a collaboration of the United States Department of Labor (DOL)'s Employment and Training Administration, Office of Policy Development and Research, Division of Research and Evaluation (DRE), and Safal Partners. This effort has received the support of the DRE Director, Wayne Gordon; and staff Neil Ridley, Gloria Salas-Kos, and Jessica Hale. At Safal Partners, Michelle Carson served as the project manager and reviewed the various iterations of the annotated bibliography. Special thanks go to Placido Gomez for proofing and formatting the bibliography.

This document was produced by Safal Partners under US Department of Labor contract # 1630DC-18-A-0005; task order #1630DC-19-F-00024.

Disclaimer: Although every effort is made to ensure that the material within this document is accurate and timely, we make no warranties or representations as to the accuracy or completeness of the contents, whether the contents are current, or free from changes caused by third parties. All information is provided "as is" without warranty of any kind and not intended to be Federal guidance. No information provided in this document may be considered legal advice and it is the responsibility of each user of this annotated bibliography to ensure that their materials meet all Federal, state, and local requirements. Use of this annotated bibliography does not imply compliance with ETA requirements.

Table of Contents

Overview	1
Applications for Advanced Analytical Techniques in Workforce Development	1
Developing Analytical Capacity for Workforce Agencies	2
Annotated Resources	3
Analytical Capacity Development	3
Cluster Analyses	4
Market Basket Analyses.....	6
Predictive Analytics	6
Text Mining and Topic Modeling	7
References	8

Overview

State and local workforce agencies produce, collect, and store large volumes of complex data. These data have the potential to transform performance management and evaluation within public workforce system setting by uncovering new insights into the underlying mechanisms that drive system outcomes and offering predictive capabilities for more efficient and effective management. However, the transformative potential of large-scale data is limited by an agency's ability to process the large volumes of data that they house (Anastasopoulos & Whitford, 2019). Artificial intelligence (AI) and machine learning (ML) offer a set of tools that agencies can use to realize the transformative potential of large-scale data (Anastasopoulos & Whitford, 2019) by developing advanced analytical capacity and techniques. To provide the context for this annotated bibliography, the following discussion provides insights about the use of AI and ML as tools to realize this potential transformation.

AI focuses on the design of systems that mimic features of human intelligence, such as understanding human language or learning (Lee & Shin, 2020). Machine learning, a sub-field of artificial intelligence, focuses more specifically on techniques that enable machines to learn from data (Lee & Shin, 2020). At its core, machine learning relies on algorithms (i.e., computer programs) that apply statistical models (Panch et al., 2018), such as text mining and topic modeling, to examine patterns or predictive algorithms to detect outcomes. However, machine learning relies on insights from statistics and computer science to develop a broader statistical toolkit that can handle large-scale and complex data, including data types like narrative text and images (Anastasopoulos & Whitford, 2019). Taken together, artificial intelligence and machine learning provide an impressive toolkit for translating large-scale data into actionable insights:

Artificial intelligence (AI) promises to transform how government agencies do their work. Rapid developments in AI have the potential to reduce the cost of core governance functions, improve the quality of decisions, and unleash the power of administrative data, thereby making government performance more efficient and effective (Engstrom et al., 2020, p. 6).

A first step to employing AI may include an evidence portfolio used to gather evidence-based resources – this annotated bibliography offers a starting point. Reviewing the evidence not only adds to the collection of available analytical techniques but may also inform the next steps of evaluation and research plans. This annotated bibliography compiles resources for state and local workforce agencies with an interest in leveraging advanced analytical techniques, such as ML, in research and evaluation. It was created in response to a state workforce agency's request during the 2019 Evaluation Peer Learning Cohort, a technical assistance initiative funded by the U.S. Department of Labor, Employment and Training Administration. The following two sections provide an overview of advanced analytical techniques, such as ML, and suggest how they may be applied to build capacity for advanced analytical techniques within workforce agencies. The annotated resources that follow summarize current articles and reports that discuss AI and ML for performance management and evaluation.

Applications for Advanced Analytical Techniques in Workforce Development

Federal, state, and local government agencies have employed AI and ML for diverse tasks, ranging from regulatory enforcement and citizen engagement (Engstrom et al., 2020) to enhancing data security and

improving program operations (Cucuel, 2018). AI, particularly ML, has applications that may be relevant to research and evaluation in workforce development, including cluster analysis, market basket analysis, predictive analytics, and text mining and topic modeling. By applying these types of advanced techniques in evaluation, state agencies have the potential to drive service delivery, evaluate systems, and ensure programs' quality improvement.

ML techniques are often classified as supervised or unsupervised (Anastasopoulos & Whitford, 2019), and both types of techniques have applications for evaluation in workforce development. Unsupervised techniques uncover patterns in large volumes of complex data, such as groups of similar observations, without attention to a specific outcome (Anastasopoulos & Whitford, 2019). Cluster analysis, for example, has been used to sort through large datasets to identify clusters of unemployed young adults who are similar to one another, revealing key shared characteristics within clusters, such as education level, household characteristics, and geographic location (Ross & Holmes, 2019). Understanding the nature of these shared characteristics provides insights into the development of targeted interventions for these groups and the regions in which they reside (Ross & Holmes, 2019). With text mining and topic modeling, ML can uncover patterns in narrative text. For example, Gao and Eldin (2014) use text mining to find job qualifications listed in construction job advertisements on popular job search engines and then apply an unsupervised topic modeling approach to identify groups of skills that are important to employers.

Supervised ML techniques predict future outcomes by training computer programs to learn the drivers of the outcomes with existing data (Panch et al., 2018). For example, Eberts et al. (2002) develop algorithms that use data on the characteristics and outcomes for recent one-stop clients to predict employability and services likely to improve outcomes for incoming clients. These predictive tools are offered to workforce development counselors through accessible computer modules, providing decision support to frontline professionals. One such tool is the Frontline Decision Support System (FDSS) prototype. Prototype materials are available on their website: <https://www.upjohn.org/about/news-events/frontline-decision-support-system-fdss-prototype-deployed>. More examples of predictive analytics are provided in the annotated resources below.

Developing Analytical Capacity for Workforce Development Agencies

Employing advanced analytical techniques in a state workforce agency implies that the agency has, or can develop, analytical capacity. Agencies requiring further development may build analytical capacity by developing their own technical infrastructure or human capital. Alternatively, they may build capacity by identifying researchers or evaluators, such as subject matter experts (SMEs) in AI techniques, or vendors offering AI and ML tools.

Baronia et al. (2020) argue that building in-house capabilities will better enable agencies to respond to their regulatory, legal, and organizational environments. Public agencies often operate with attention to factors that introduce additional layers of complexity to AI and ML tasks, such as transparency, fairness, and accountability to diverse stakeholders (Desouza et al., 2020). Baronia et al. (2020) suggest four key components to internal capacity building—1) upgrading technical systems and data capacity to house both the complex software and large volumes of data needed for AI; 2) ensuring that the agency has in-house expertise for deploying AI and ML in ways that are sensitive to the public context; 3) providing a

margin of error for testing and developing tools; and 4) designing accountability into the system with testing and evaluation. One benefit that the public sector enjoys in developing AI and ML is the transparency of tools deployed by other public agencies (Desouza et al., 2020). Agencies may find open-source tools or seek information on tools developed by their public sector counterparts, such as the FDSS prototype described above and pilot tested in two Georgia Career Centers in early 2001.

Palmer (2018) suggests that the make-or-buy decision depends on the organization's experience with advanced analytics and the presence of leadership, financial, and structural supports to effectively use data. In-house development can reduce costs, but the potential for savings depends upon the knowledge and experience of staff members (Palmer, 2018). If an organization determines that buying AI and ML tools is the more appropriate route, Palmer (2018) offers several considerations to ensure effective oversight of the vendor and implementation of the tools.

Organizations should take steps to determine that the tools will integrate with local systems and meet organizational needs. Contracts should consider the level of technical support offered for implementation. In the contracting process, organizations should also seek information about the algorithms used and their effectiveness. In addition, the organization should plan to conduct its own evaluation of any interventions carried out with the AI/ML tool(s). Finally, the contract should include provisions that consider data ownership, transfer to other vendors, security, and privacy.

Annotated Resources

This annotated bibliography offers three resources for analytical capacity development and 14 resources on analytical approaches that may support interested researchers in building a portfolio of evidence, planning analyses, and conducting evaluations that employ large-scale data. These annotated resources focus on issues associated with approaches to and use of large-scale data, most with potential applications to workforce development research and analyses. Though resources are arranged by topic, analyses will probably require a combination of methods and tools. Thus, examples from more than one topical list may be helpful. For example, use of cluster analysis with large-scale data can lead to applications of predictive analysis.

The citations include the author name(s), year published (with dates from 2002 to 2020), title, and a hyperlink to the online resource. Within the topic areas: analytical capacity development, cluster analysis, predictive analysis, and text mining and topic modeling, the resources are listed alphabetically by authors' last names. Finally, readers may wish to review the reference lists and appendices in each citation to gain further information and potential contacts with analytics experts.

Analytical Capacity Development

Baronia, N., Engstrom, D. E., Ho, D. E., Musgrave, S., & Sharkey, C. (2020). Building internal capacity. In Engstrom, D. E., Ho, D. E., Sharkey, C., & Cuellar, M. (Eds.), *Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies*. Administrative Conference of the United States. <https://law.stanford.edu/publications/government-by-algorithm-artificial-intelligence-in-federal-administrative-agencies/>

As government agencies increasingly turn to AI, “a core challenge for agencies is how to generate the necessary technical capacity—the ability to identify, develop, responsibly use, and maintain complex technical solutions” (p. 71). This chapter focuses on the considerations of building internal capacity in federal agencies. Considerations include building technical infrastructure and data capacity, growing internal staff capacity, investing in AI strategy, and linking capacity to accountability.

Muro, M., Maxim, R., & Whiton, J. (2019). *Automation and artificial Intelligence: How machines are affecting people and places*. Brookings Institution, Metropolitan Policy Program.
<https://www.brookings.edu/research/automation-and-artificial-intelligence-how-machines-affect-people-and-places/>

Considering the impacts of automation over the years 1980 to 2016 and 2016 to 2030, this report offers backward- and forward-looking analyses of trends and their effects on people and communities. The report emphasizes areas of potential occupational change and explores industry, geographical, and demographic variations. The report concludes with a comprehensive response framework for policymakers.

Palmer, I. (2018). *Choosing a predictive analytics vendor: A guide for colleges*. New America.
<https://files.eric.ed.gov/fulltext/ED590742.pdf>

This guide for postsecondary institutions on selecting predictive analytics vendors offers administrators tools to decide on useful algorithms for predicting student success and to ensure that vendors, if selected, use predictive analyses in an ethical way. The guide also shares considerations for deciding on in-house vs. vendor-produced analytics, flexibility and fit of data tools, data transparency, connections with research and evaluation, and privacy and security. An appendix lists institutional contacts with expertise in employing predictive analytics, as well as potential vendors.

Cluster Analyses

Bahr, P. R. (2010). The bird’s eye view of community colleges: A behavioral typology of first-time students based on cluster analytic classification. *Research in Higher Education*, 51(8), 724-749. www.jstor.org/stable/40927277

This paper presents a typology of students in community colleges based on clusters of behavior. The author expands on prior classification models using large-scale data from the California Community College System and tests how the cluster solution predicts students’ long-term outcomes. The paper employs K-means cluster analyses to identify an optimal six clusters.

Accompanying the original paper are two methodology guides:

Bahr, P. R., Booth, K., & Willett, T. (2012). *Segmentation model for assessing course-taking patterns: Research methodology and discussion guide*. [Supplemental material].

<https://rpgroup.org/Portals/0/Documents/CompletionInquiryGuideandResearchTools/Segmentation-Model-for-Assessing-Course-Taking-Patterns.pdf>

and *Updated Replication Rules (December 2012)* [Supplemental material].

<https://rpgroup.org/Portals/0/Documents/CompletionInquiryGuideandResearchTools/Segmentation-Model-for-Assessing-Course-Taking-Patterns.pdf>

Houssemand, C., Pignault, A., & Meyers, R. (2014). A psychological typology of newly unemployed people for profiling and counselling. *Current Psychology*, 33(3), 301-320.

<https://doi.org/10.1007/s12144-014-9214-9>

Employing hierarchical cluster analysis, this paper identifies a psychological typology of five groups of unemployed adults based on psychological questionnaire responses at intake. The typology is used to predict employment status after a year.

Ross, M., & Holmes, N. (2019). *Meet the millions of young adults who are out of work: Local profiles of jobless young adults and strategies to connect them to employment*. Metropolitan Policy Program at Brookings.

https://www.brookings.edu/wp-content/uploads/2019/04/2019.04.09_Brookings-metro_Out-of-Work_Youth_Ross-Holmes.pdf

The authors “used cluster analysis to segment out-of-work young adults into five groups... likely to benefit from similar types of employment and educational assistance” in 119 urban cities and counties (p. 4). Clusters are described based on age and educational attainment but rely on multiple characteristics of unemployed young adults.

Accompanying the original paper is a methodology appendix:

Ross, M., & Holmes, N. (2019). *Methods and data sources*. [Supplemental material].

https://www.brookings.edu/wp-content/uploads/2019/04/2019.04.09_Brookings-metro_Out-of-Work_Youth_Methodology-Appendix_Ross-Holmes.pdf

This appendix describes the methods the authors followed. In the main report, clusters rely on 10 sets of young adults’ characteristics. Unemployed young adults were first clustered within the city or county and then sorted across all sampled jurisdictions. Authors employed complete-linkage hierarchical clustering, initially identifying 557 jurisdiction-level clusters, and aggregating them to five major groups.

U.S. Economic Development Administration. (2009). *Crossing the next regional frontier: Information and analytics linking regional competitiveness to investment in a knowledge-based economy*.

https://www.researchgate.net/publication/277142742_Crossing_the_Next_Regional_Frontier_Information_and_Analytics_Linking_Regional_Competitiveness_to_Investment_in_a_Knowledge-Based_Economy

In 2009 the Economic Development Administration commissioned a research project to develop four “data and decision-support tools that (a) facilitate civic dialogue about the new economy,

(b) help local leaders and other stakeholders gain an understanding of their knowledge and skill assets and gaps, and (c) show how public investments need to be aligned or realigned to be supportive of a knowledge-based innovation economy” (p. 20). The four tools employ industry clusters, occupation clusters, an innovation index, and a regional strategy and investment framework. One project tool relied on Ward’s agglomerative hierarchical clustering algorithm to “identify and categorize occupations into clusters based on O*NET-SOC knowledge classifications and measurements” (p. 47). Other algorithms relying on cluster analysis are also described.

Market Basket Analyses

Ahmad, W., Akhtaruzamman, M., Zahra, U., Ohri, C., & Ramakrishnan, B. (2018). Investigation on the impact of leadership styles using data mining techniques. *IntechOpen*.
<https://doi.org/10.5772/intechopen.78660>

This paper compares use of statistical methods (e.g., regression) and data mining (rule-based approaches and decision tree modelling) to examine association of leadership style with leader outcomes. Data mining “techniques were used to extract hidden trends and patterns in the data to report various ways to increase the employee outcomes” (p. 139).

Luna, J. M., Cano, A., Sakalauskas, V., & Ventura, S. (2016). Discovering useful patterns from multiple instance data. *Information Sciences*, 357, 23-38.
https://www.researchgate.net/publication/299986939_Discovering_Useful_Patterns_from_Multiple_Instance_Data

Authors note that “association rule mining is one of the most common data mining techniques used to identify...relationships between patterns from large datasets” (p. 23). The paper describes multiple-instance association rule mining and an “algorithm to tackle the problem,” employing relationships in unemployment records as an example.

Raorane, A. A., Kulkarni, R. V., & Jitkar, B. D. (2012). Association rule—extracting knowledge using market basket analysis. *Research Journal of Recent Sciences*, 1(2), 19-27.
<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1057.9306&rep=rep1&type=pdf>

This paper provides straightforward explanations of data mining and market basket analysis. It applies association rules to help the reader determine optimal combinations of decisions and understand associated algorithms.

Predictive Analytics

Eberts, R. W., O'Leary, C. J., DeRango, K. J., & Parker, H. (2002). A frontline decision support system for one-stop centers. In R. Eberts, C. O'Leary, and S. Wandner (Eds.), *Targeting Employment Services* (pp. 337-384). W.E. Upjohn Institute for Employment Research.
<https://doi.org/10.17848/9781417524440.ch12>

The W.E. Upjohn Institute for Employment Research has designed a Frontline Decision Support System (FDSS) to integrate advanced decision tools into decision-making. This chapter presents an overview of FDSS and provides examples of analyses that support decision algorithms in FDSS tools. Models “include estimates of the likelihood of finding a job in the industry in which the worker was employed prior to displacement and estimates of the earnings that a displaced worker might expect when looking for reemployment” (p. 339). Other FDSS algorithms identify “occupations that are related to a worker’s occupation held prior to displacement” and “the most effective set of services based on the individual’s characteristics and employment history” (p. 339).

Akers, K., Bond, B., Brandon, B., Renick, T., McKelvey, B., Sullivan, L., Pienaar, D., & Brann, T. (2020, September 10). *Predictive analytics to improve student outcomes* [Webinar]. U.S. Department of Education. <https://mediasite.ed.gov/webcast/Play/44bab8fe1b54425792e62f452837a5641d>

In this webinar, data analytics professionals in higher education discuss the use of predictive analytics to improve student outcomes. The webinar includes four presentations that offer insights on the implementation of predictive analytics from different perspectives. The first presentation discusses components necessary to build state longitudinal data systems that meet the needs of system stakeholders. The second presentation discusses one institution’s use of predictive analytics to identify barriers to student success and to develop solutions to help students overcome those barriers. The third presentation discusses the construction and potential applications of a local longitudinal data system. The final presentation focuses on how to incorporate predictive analytics into management processes to drive decision making. The combined slides for the webinar are hosted at:

<https://www2.ed.gov/about/offices/list/oep/predictiveanalyticsslides.pdf>.

Vaithianathan, R., Putnam-Hornstein, E., Stagner, M., & Weigensberg, E. (2019). *Predictive risk modeling for child protection* [Fact sheet]. Mathematica Progress Together. <https://www.mathematica.org/-/media/publications/pdfs/earlychildhood/2019/predictive-risk-modeling-fs.pdf>

“Decision-support tools such as predictive risk models can help” a Child Protective Services worker make decisions about hotline referrals on children at risk based on data (p. 1). This fact sheet describes predictive risk model tools that employ machine learning to train data, draw upon multiple fields of information, and classify results to support social worker decision-making.

Text Mining and Topic Modeling

Bishop, A., Dobre, A. M., & Mateos-Garcia, J. C. (2018). Economic complexity and the emergence of new ideas. <https://dx.doi.org/10.2139/ssrn.3242960>

In this study, the authors use topic modeling to “detect ‘topics’ in the websites of businesses and analyse their relation to economic complexity and industrial sectors” (p. 1). They then “measure

emergence by looking for novel words in the websites of businesses in different locations and sectors” (p. 1). By linking web data and industry sector data, the study develops “tools to aid policymakers in understanding the developing specialisation profiles of a local economy in a timely manner” (p. 30).

Gao, L., & Eldin, N. (2014). Employers’ expectations: A probabilistic text mining model. *Procedia Engineering*, 85, 175-182. <https://doi.org/10.1016/j.proeng.2014.10.542>

This paper employs probabilistic text mining to analyze employment data in the construction industry. Topic modeling is an unsupervised, machine-learning technique that detects word and phrase patterns in documents and automatically clusters the most meaningful word groups. The method in this paper examines job qualifications from online employment ads. By identifying needed skillsets, adults entering career pathways can determine how to prepare for and seek future positions.

~*~

References

- Ahmad, W., Akhtaruzamman, M., Zahra, U., Ohri, C., & Ramakrishnan, B. (2018). Investigation on the impact of leadership styles using data mining techniques. *IntechOpen*. <https://doi.org/10.5772/intechopen.78660>
- Akers, K., Bond, B., Brandon, B., Renick, T., McKelvey, B., Sullivan, L., Pienaar, D., & Brann, T. (2020, September 10). *Predictive analytics to improve student outcomes* [Webinar]. U.S. Department of Education. <https://mediasite.ed.gov/webcast/Play/44bab8fe1b54425792e62f452837a5641d>
- Anastasopoulos, L. J., & Whitford, A. B. (2019). Machine learning for public administration research, with an application to organizational reputation. *Journal of Public Administration Research and Theory*, 29(3), 491-510. <https://doi.org/10.1093/jopart/muy060>
- Bahr, P. R., Booth, K., & Willett, T. (2012). *Segmentation model for assessing course-taking patterns: Research methodology and discussion guide*. [Supplemental material]. <https://rpgroup.org/Portals/0/Documents/CompletionInquiryGuideandResearchTools/Segmentation-Model-for-Assessing-Course-Taking-Patterns.pdf>
- Bahr, P. R. (2010). The bird’s eye view of community colleges: A behavioral typology of first-time students based on cluster analytic classification. *Research in Higher Education*, 51(8), 724-749. www.jstor.org/stable/40927277
- Baronia, N., Engstrom, D. E., Ho, D. E., Musgrave, S., & Sharkey, C. (2020). Building internal capacity. In Engstrom, D. E., Ho, D. E., Sharkey, C., & Cuellar, M. (Eds.), *Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies*. Administrative Conference of the United States. <https://law.stanford.edu/publications/government-by-algorithm-artificial-intelligence-in-federal-administrative-agencies/>
- Bishop, A., Dobre, A. M., & Mateos-Garcia, J. C. (2018). Economic complexity and the emergence of new ideas. <https://dx.doi.org/10.2139/ssrn.3242960>
- Cucuel, K. (2018, March 6). How state and local governments can maximize output with machine learning. *StateTech*. <https://statetechmagazine.com/article/2018/03/how-state-and-local-governments-can-maximize-output-machine-learning>
- Desousa, K. C., Dawson, G. S., & Chenok, D. (2020). Designing, developing, and deploying artificial intelligence systems: Lessons from and for the public sector. *Business Horizons*, 63, 205-213. <https://doi.org/10.1016/j.bushor.2019.11.004>

- Eberts, R. W., O'Leary, C. J., DeRango, K. J., & Parker, H. (2002). A frontline decision support system for one-stop centers. In R. Eberts, C. O'Leary, and S. Wandner (Eds.), *Targeting Employment Services* (pp. 337-384). W.E. Upjohn Institute for Employment Research.
<https://doi.org/10.17848/9781417524440.ch12>
- Engstrom, D. E., Ho, D. E., Sharkey, C., & Cuellar, M. (Eds.) (2020). *Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies*. Administrative Conference of the United States.
<https://law.stanford.edu/publications/government-by-algorithm-artificial-intelligence-in-federal-administrative-agencies/>
- Gao, L., & Eldin, N. (2014). Employers' expectations: A probabilistic text mining model. *Procedia Engineering*, 85, 175-182. <https://doi.org/10.1016/j.proeng.2014.10.542>
- Houssemand, C., Pignault, A., & Meyers, R. (2014). A psychological typology of newly unemployed people for profiling and counselling. *Current Psychology*, 33(3), 301-320.
<https://doi.org/10.1007/s12144-014-9214-9>
- Lee, I., & Shin, Y. J. (2020). Machine learning for enterprises: Applications, algorithm selection, and challenges. *Business Horizons*, 63, 157-170. <https://doi.org/10.1016/j.bushor.2019.10.005>
- Luna, J. M., Cano, A., Sakalauskas, V., & Ventura, S. (2016). Discovering useful patterns from multiple instance data. *Information Sciences*, 357, 23-38.
https://www.researchgate.net/publication/299986939_Discovering_Useful_Patterns_from_Multiple_Instance_Data
- Muro, M., Maxim, R., & Whiton, J. (2019). *Automation and artificial Intelligence: How machines are affecting people and places*. Brookings Institution, Metropolitan Policy Program.
<https://www.brookings.edu/research/automation-and-artificial-intelligence-how-machines-affect-people-and-places/>
- Palmer, I. (2018). *Choosing a predictive analytics vendor: A guide for colleges*. New America.
<https://files.eric.ed.gov/fulltext/ED590742.pdf>
- Panch, T., Szolovits, P., & Atun, R. (2018). Artificial intelligence, machine learning, and health systems. *Journal of Global Health*, 8(2), 1-8. <https://doi.org/10.7189/jogh.08.020303>
- Raorane, A. A., Kulkarni, R. V., & Jitkar, B. D. (2012). Association rule-extracting knowledge using market basket analysis. *Research Journal of Recent Sciences*, 1(2), 19-27.
<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1057.9306&rep=rep1&type=pdf>
- Ross, M., & Holmes, N. (2019). *Meet the millions of young adults who are out of work: Local profiles of jobless young adults and strategies to connect them to employment*. Metropolitan Policy Program at Brookings.
https://www.brookings.edu/wp-content/uploads/2019/04/2019.04.09_Brookings-metro_Out-of-Work_Youth_Ross-Holmes.pdf
- Ross, M., & Holmes, N. (2019). *Methods and data sources*. [Supplemental material].
https://www.brookings.edu/wp-content/uploads/2019/04/2019.04.09_Brookings-metro_Out-of-Work_Youth_Methodology-Appendix_Ross-Holmes.pdf
- U.S. Economic Development Administration. (2009). *Crossing the next regional frontier: Information and analytics linking regional competitiveness to investment in a knowledge-based economy*.
https://www.researchgate.net/publication/277142742_Crossing_the_Next_Regional_Frontier_Information_and_Analytics_Linking_Regional_Competitiveness_to_Investment_in_a_Knowledge-Based_Economy
- Vaithianathan, R., Putnam-Hornstein, E., Stagner, M., & Weigensberg, E. (2019). *Predictive risk modeling for child protection* [Fact sheet]. Mathematica Progress Together.
<https://www.mathematica.org/-/media/publications/pdfs/earlychildhood/2019/predictive-risk-modeling-fs.pdf>