Catalogue no. 11-522-x
ISSN: 1709-8211

Proceedings of Statistics Canada Symposium 2021
Adopting Data Science in Official Statistics to Meet Society's Emerging Needs

Bridging the gap between the displaced and in-demand occupations

by Vishal Subramanian Balashankar, Badri Venkataraman and Chris Astle

Release date: October 15, 2021

Statistics Canada Statistique Canada

Canada

Proceedings of Statistics Canada Symposium 2021
Adopting Data Science in Official Statistics to Meet Society's Emerging Needs

Bridging the gap between the displaced and in-demand occupations

Vishal Subramanian Balashankar, Badri Venkataraman, and Chris Astle[1]

Abstract

With labour market uncertainty increasing across Canada, there is a need for innovative ways to help displaced workers to re-skill/up-skill and potentially pivot to in-demand occupations. In our study, we present a unique approach to bridge the gap between the displaced and in-demand occupations and provide a machine learning framework that may be able to forecast employment by NAICS for 6 months. We have combined the monthly employment data from Statistics Canada's Survey of Employment and Payroll Hours, and the monthly job ads counts from Burning Glass to achieve our goal.

Our approach consists of the following three steps:

1. Finding the displaced occupations in Alberta over the last 7 years based on the integrated actual employment and job ads count data. Validation is performed to establish the correlation between the two data sets in this step.

2. Using the list of displaced occupations, a unique pivot graph is developed to map a displaced occupation to a list of in-demand occupations which have skills similar to the chosen displaced occupation. To establish the similarity between occupations, a similarity score (from Burning Glass) is used. Once a prospective in-demand occupation is selected, the skill gap is computed and presented to the user.

3. Applying SARIMA and SARIMAX models to forecast employment for 6 months. The models have a mean absolute percentage error of 1.4 % and 10.76 % in the test sets in 2019 and 2020, respectively across all the NAICS sectors. The monthly predictions have errors less than 0.5 %.

The above approaches are aimed at assisting public policy and planning.

Key Words: Employment; Labour Market; Job Ads; Skills; Time Series Analysis; Forecasting.

1. Introduction

OECD's *Back to Work Canada* report highlights that an alarming 2.2 % of the Canadian workforce, holding at least 1 year of tenure, is involuntarily displaced from their current jobs every year (OECD 2015). This type of displacement typically happens when a firm scales down or files bankruptcy. The displaced workers are challenged with extended periods of unemployment and layoffs resulting in a huge socio-economic cost to them. The ongoing COVID-19 pandemic, coupled with the crude oil price crash made the re-employment prospects of displaced workers worse in Alberta (Business Council of Alberta 2021). The long-term unemployment rate of a displaced worker stood between 2 and 3 % in April – July 2021 for the province of Alberta (Business Council of Alberta 2021, Labour Force Survey 2021). This percentage number is significant and would roughly translate to the population of St. Albert (Alberta Municipal Affairs 2019). Displaced workers are often forced to settle down for part-time or temporary jobs that are not as well-paying while also not being similar in skillsets when compared to their previous roles. While Alberta Learning Information Services (ALIS), an online platform from the Government of Alberta (GoA), provides end-to-end information about career planning, labour market trends, and workplace-related information, there is no tool aimed specifically at helping displaced workers in Alberta, to the best of our knowledge (ALIS 2021). In this paper, a unique data-science framework is demonstrated that provides key insights to a displacing worker regarding in-demand occupations, skill gaps, and future trends. Multiple data sources were leveraged, and a prototype known as "AB- Next

[1]Vishal Subramanian Balashankar, Cybera, Suite 400, 10055 – 106 Street, Edmonton, Alberta, Canada, T5J 2Y2 (vishal.balashankar@cybera.ca); Badri Venkataraman, Cybera, Suite 400, 10055 – 106 Street, Edmonton, Alberta, Canada, T5J 2Y2 (badri.venkataraman@cybera.ca); Chris Astle, CastlePoint Consulting Inc, 9651 Ottewell Rd NW, Edmonton, AB T6B 2E3 (chris.astle@castlepoint.ca)

Step Pivot Tool" was built. While this tool was built considering the Alberta labour market, in principle the current framework could be extended to any province in Canada. This tool is meant to aid governments with policy and planning for developing support programs and policies addressing the needs of the displaced workers.

The data sources used in this work were the annual & monthly employment data offered by the Survey of Employment and Payroll Hours (SEPH) from Statistics Canada and the job ads data supplied by Emsi Burning Glass Technologies (BGT) (Statistics Canada 2020). EMSI – BGT, an analytics software company based in Boston, is a leading provider of real-time labour market insights leveraging the online job ads information using their powerful AI/ML frameworks. In this work, we have combined both the above-stated data sources and described a framework through the following three core components: i) finding the displacing occupations, ii) bridging the displaced and in-demand occupations, and iii) presenting the forecast attempt to date. The timeframe of our framework is from 2013 to 2020 as the job ads data is only available from 2013 on. The prototype was developed in Python leveraging the cloud computing service offered by Cybera.

2. Method to find displaced occupations

The following twofold approach was devised to find the list of displacing occupations in Alberta:

2.1 Finding the list of top displacing NAICS

First, a list of displacing industrial sectors is found using North American Classification System (NAICS 2 - digit) annual employment data by SEPH (Statistics Canada 2020). NAICS is an industrial classification system that provides the nomenclature for industrial structures across North America. In our work, we have used the Canada 2017 version 3.0 NAICS, and focussed on the two-digit classification level (approx. 20 different sectors). While we also performed the analysis for four-digit NAICS, the BGT data has a lower quality of NAICS to occupation mapping, given the granularity of the industrial structure definition, and is not detailed in this paper.

Methodology: A reference and target year are chosen from 2013 to 2020 and the corresponding annual employment data for all NAICS sectors are recorded. The constraint here is that the chosen target year should always be ahead of the reference year. Then, a difference in actual employment is computed using StatsCan data. This differenced number corresponds to the magnitude of jobs lost and the list is sorted by descending order to find the top areas of displacement Table 2.1-1 highlights the five highest NAICS areas of displacement between 2013 (reference year) and 2019 (target year) in Alberta. Construction [23], Mining, quarrying, & oil & gas extraction [21] and Manufacturing [31-33] sectors accounted for approximately 80,000 displaced workers between 2013 and 2019. Please note that the magnitude and the order of jobs lost very much depend on the chosen combination of the reference and target year and the number of NAICS sectors one wishes to explore. These three quantities (reference year, target year, and number of NAICS sectors (< 10)) are left as variables in our prototype with the user having the flexibility to set them based on their respective use cases. We are considering exploring the complex relationship between these variables and how they might be optimized as future work.

Table 2.1-1
Top 5 displacement NAICS (2-digit) between 2013 (reference year) and 2019 (target year) in Alberta.

NAICS (2-digit)	Employment (2013)	Employment (2019)	Jobs Lost
Construction [23]	206,384	171,322	35,062
Mining, quarrying, & oil & gas extraction [21]	131,689	100,203	31,486
Manufacturing [31-33]	134,543	121,975	12,568
Other services (except public administration) [81]	80,246	75,576	4,670
Wholesale Trade [41]	103,114	99,688	3,426

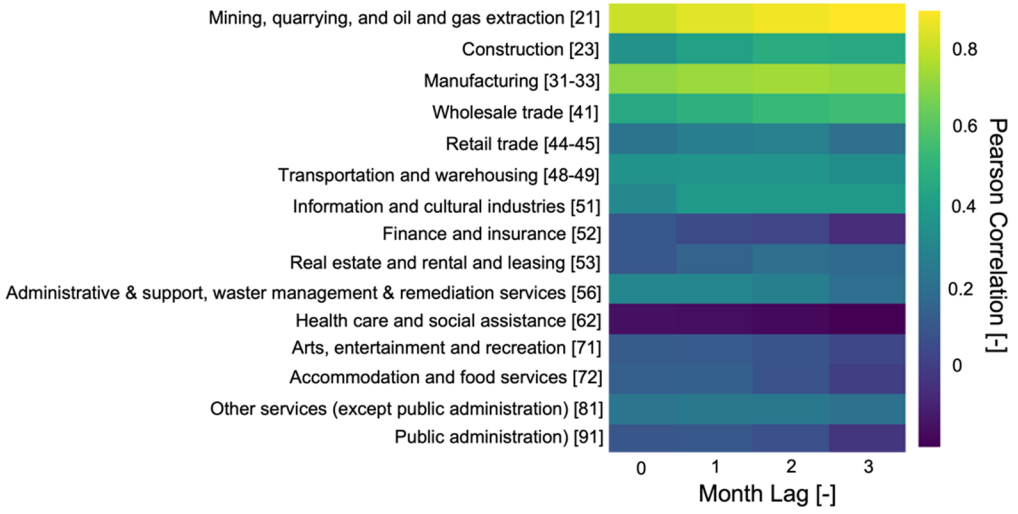
While the above table provides a high-level view of the sectors that are facing displacement, it does not provide deeper insights into the specific occupations that are getting displaced. As a potential method to add occupation-level data to the more general NAICS data, integrating EMSI-BGT job ads data was explored. The online job ads data, unlike the traditional economic data, provides a real-time snapshot of labour market dynamics. It was hoped that adding online

job ads data would bring a different, and valuable dimension to the work. Support for this approach can be found in a study from Georgetown University that found a good correlation between the EMSI-BGT job ads count and the American Job Openings and Labour Turnover Survey (JOLTS) data during 2006 - 2013 (Carnevale 2014). This further reinforced our confidence in using the BGT job ads data to predict labour insights in our work.

2.2 Finding the list of top BG occupations

BGT provided us with the monthly count of job ads by the NAICS sector from Jan – 2013 to Dec - 2020. Our hypothesis here is to check if the monthly time series job ads count trend from BGT correlates with the monthly time series employment trend from SEPH, Statistics Canada. While the potential for success was not considered high from the outset, the analysis turned out to be interesting and eventually insightful. A series of time lags (in steps of 1 month up to 3 months) was applied to the BGT monthly job ads count data while the corresponding Pearson correlation values of all the NAICS sectors against the monthly SEPH data were found. A heat map of the Pearson Correlation values resulting from this process is constructed and described in Figure 2.2-1. The lighter shades of yellow correspond to a higher correlation while the dark shades correspond to poorer values as seen. On one hand, it is found that the volatile sectors like Mining, quarrying, oil and gas extraction [21], Construction [23], and Manufacturing [31-33] in Alberta have higher correspondence between the BGT and Statistics Canada data. This proves that job ads data could in principle be used as a proxy to predict the employment for these sectors. On the other hand, the sectors like Health care and Social Assistance [62], Finance and Insurance [52] had the lowest correlation values, indicating that BGT job ads counts may not be useful in capturing the employment trends in these areas. Each of the sectors had a unique lag interval that maximized the correlation value, as shown on the X-axis of Figure 2.2-1. For instance, a two-month lag produced the highest correlation for the Construction sector [23], potentially indicating that jobs posted on-line in this sector translate into actual employment two months down the line. It is important to acknowledge here that the correlation does not necessarily mean causation. We still made this bold attempt to unearth interesting combinations of data sources while getting the best out of both data sources.

Figure 2.2-1
Heat map describing the correlation between the time series of monthly BGT job ads count data and the SEPH, Statistics Canada employment data across all the NAICS sectors considered in our study.



The above-described exercise validates the use of monthly job ad counts as a proxy for monthly employment across sectors with higher correlation values (> 0.3). The challenge then became finding a procedure to map occupations across the 20 NAICS sectors. This was accomplished using the additional data connections provided in the BGT data including mapping job postings to employers, and further mapping employers to the NAICS (two-digit). Thus, 20 NAICS sectors are now mapped to 660 occupations across 18 broad career areas. For instance, between 2013 and 2020, of the total job postings attributed to the Construction Manager occupation, the NAICS breakdown was 60 % Construction [23], 8 % Real Estate and Leasing [53] and 5 % Finance [52]. With the NAICS to occupation mapping, a list of occupations is found for the top displacing NAICS sector, as shown in Table 2.1-1. Corresponding job ads

count for the occupations are then found for the chosen reference and target years. The difference in job ads count is then computed between the target and reference year. The difference corresponds to the total job ads lost, and the list is then sorted in descending order. Table 2.2-1 describes the top 10 displacing occupations between 2013 (reference year) and 2019 (target year). Electrician tops the list with a reduction of 533 job postings (~82 % drop) followed by Mechanical Engineer with 461 lost job ads (~65 % drop). While we used the job ads count as a proxy for employment, it is important to note the magnitude of difference between the job ads lost (Table 2.2-1) and actual jobs lost (Table 2.1-1).

Table 2.2-1
List of top 10 displacing occupations based on EMSI-BGT job ads data between 2013 (reference year) and 2019 (target year)

BG Occupations	Job Ads Count (2013)	Job Ads Count (2019)	Job Ads Lost
Electrician	656	123	533
Mechanical Engineer	705	244	461
Construction Manager	832	441	391
Welder / Solder	593	202	391
Estimator	455	134	321
Millwright	480	206	274
Reservoir / Petroleum Engineer	302	38	264
Civil Engineer	363	131	232
Quality Inspector / Technician	324	98	226
Project Manager	402	190	212

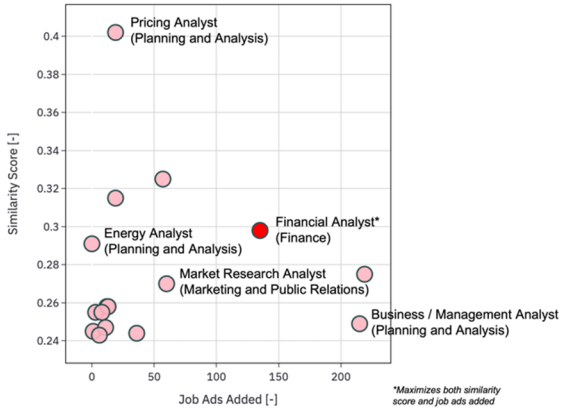
3. Bridging the displaced and in-demand occupations

The above section dealt in detail with the methodologies to find the list of displacing NAICS sectors and further the displacing occupations. This section will now describe approaches to find the in-demand potential occupations for each displaced occupation that the displaced workers can pivot to, and further provide the skill gap between them.

3.1 Pivot Graph

With the same choice of reference and target year from the above section, a list of BGT occupations that had an increase in job ads count (as opposed to reduction as in Section 2.2) is found. The increase in job ads count is taken to indicate an increase in demand for that occupation over the chosen time interval. Another important criterion to consider while identifying potential occupations for a displaced worker to pivot into is skills match, in order to decrease the amount of time and effort needed for re-training, and maximize the use of existing skills. In order to quantify this, the similarity score offered by the BGT is used. The in-demand occupation with a higher similarity score would mean a closer correspondence to a displaced occupation. Please note that the similarity score is computed by the BGT's powerful NLP algorithm.

Figure 3.1-1
Pivot Graph showing the potential in-demand occupations for *Estimator* (displaced occupation) between 2013 and 2019 in Alberta

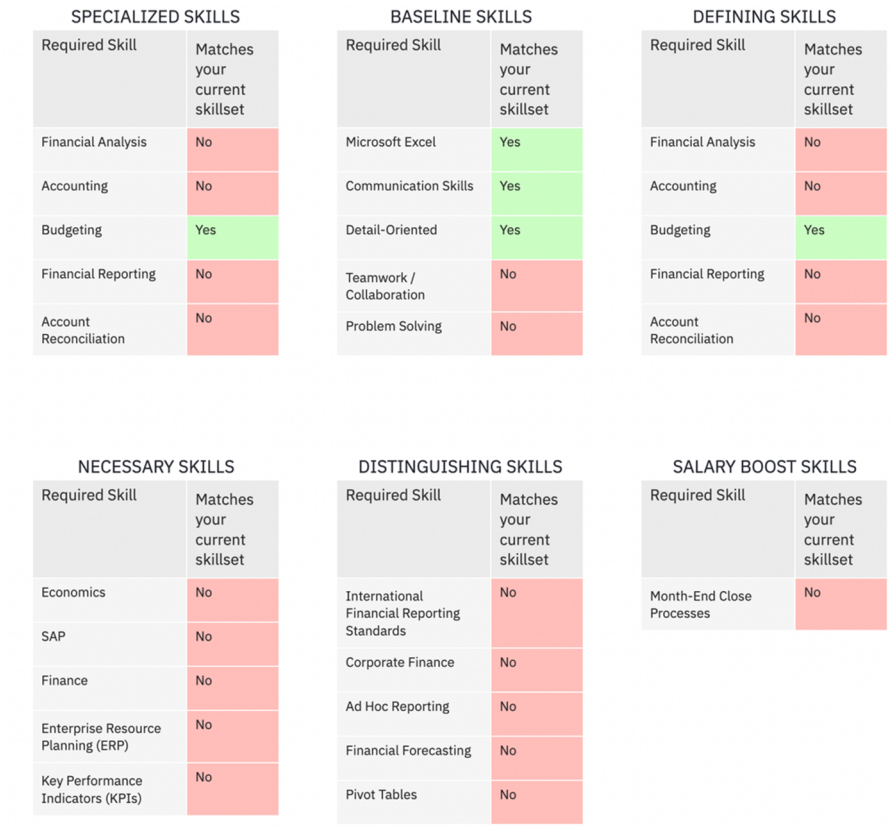


A pivot graph for each displaced occupation is constructed using the net job ads added and similarity score as two variables. Figure 3.1-1 provides an example pivot graph for a displaced occupation of *Estimator* (obtained from Table 2.2-1) in Alberta. The X-axis corresponds to the net job ads added between the reference and target year and the Y-axis quantifies the similarity between the displaced and in-demand occupation in terms of skillsets. Each marker in the figure corresponds to a different potential pivot occupation. While the *Estimator* occupation is listed under the Construction, Extraction, and Architecture career area in the BGT data, a powerful feature of this tool is to exhibit what potential pivot occupations lie beyond a worker's current career area. A potential pivoting occupation should have higher similarity score and net job ads added (top right portion in the plot). Potential pivot occupations including Financial Analyst (from Finance), Business / Marketing Analyst (from Planning and Analysis) and Market Research Analyst (from Marketing and Public Relations) are marked in Figure 3.1-1. The occupation indicated by the red marker has relatively high values for both similarity score and number of job ads added and is worth investigating further Presenting the results in this way provides the user the ability to interact and explore the data depending on personal preference and situation. A unique pivot graph is constructed for each of the displaced occupations across different career areas from Table 2.2-1.

3.2 Skill Gap

The next obvious question would be to see the skill gap between an identified displacement occupation and a recommended in-demand occupation. BGT categorizes skills into 6 different bins for each occupation: Specialized Skills, Baseline, Defining, Necessary, and Salary Boost Skills, with nearly 6000 skills identified overall. Figure 3.2-1 replicates the skill gap feature from our prototype. The red shade denotes that a skill is not typically possessed by the displaced occupation but is identified as useful for the pivot occupation. Green shading indicated that the displaced occupation typically has the skill that is identified as useful for the pivot occupation. This is an important feature as it potentially provides a displaced worker with a roadmap of skill requirements necessary for the transition.

Figure 3.2-1
Skill Gap between Estimator (displaced occupation) and Financial Analyst (in-demand occupation) (screenshot from our AB-Next Step Pivot Prototype Tool)



SPECIALIZED SKILLS

Required Skill	Matches your current skillset
Financial Analysis	No
Accounting	No
Budgeting	Yes
Financial Reporting	No
Account Reconciliation	No

BASELINE SKILLS

Required Skill	Matches your current skillset
Microsoft Excel	Yes
Communication Skills	Yes
Detail-Oriented	Yes
Teamwork / Collaboration	No
Problem Solving	No

DEFINING SKILLS

Required Skill	Matches your current skillset
Financial Analysis	No
Accounting	No
Budgeting	Yes
Financial Reporting	No
Account Reconciliation	No

NECESSARY SKILLS

Required Skill	Matches your current skillset
Economics	No
SAP	No
Finance	No
Enterprise Resource Planning (ERP)	No
Key Performance Indicators (KPIs)	No

DISTINGUISHING SKILLS

Required Skill	Matches your current skillset
International Financial Reporting Standards	No
Corporate Finance	No
Ad Hoc Reporting	No
Financial Forecasting	No
Pivot Tables	No

SALARY BOOST SKILLS

Required Skill	Matches your current skillset
Month-End Close Processes	No

4. Forecasting attempt to date

Now, the focus shifts to forecasting employment by NAICS sectors for 6 months. Two broad prediction models were built for each NAICS sector. First, using exclusively the StatsCan Employment data, Seasonal Auto-Regressive Integrated Moving Average (SARIMA) model is used to predict employment for 6 months. Second, using the combined StatsCan and BGT data, Seasonal Auto-Regressive Integrated Moving Average with eXogenous Variables (SARIMAX) is used to predict the employment using BGT job ads data as the eXogenous variable. This work is currently in progress, with initial results showing monthly predictions from SARIMAX (using StatsCan + BGT data) having prediction errors less than 0.5 %.

5. Conclusions

An exclusive well-rounded prototype to aid the displaced workers in Alberta is described in this paper. A list of displaced NAICS sectors, and a resulting list of displaced occupations are found within a certain time frame. This step is supported by the validation study comparing the monthly SEPH data and online job ads data. A unique pivot graph is then presented showcasing the in-demand (pivot) occupations for each displaced occupation. For each in-demand (pivot) occupation, a skill gap analysis is presented between the displaced and in-demand (pivot) occupation. While the potential of using BGT job ads and augmenting them with Statistics Canada data is shown to derive deeper insights. It's important to acknowledge the job ads data may be biased towards certain sectors and suffer from other limitations. Caution is advised while interpreting this early-stage work.

References

Alberta Municipal Affairs (2019), *2019 Municipal Affairs Population List*, Government of Alberta. https://open.alberta.ca/dataset/daab9fce-c2f6-49d1-a433-375b2b7aee24/resource/61cd908d-e2b9-4837-939b-533848d723b9/download/2019_mapl_web.pdf

ALIS (2021), *Alberta Learning Informational Services*. Retrieved from https://alis.alberta.ca/

Business Council of Alberta (2021), *Silent Crisis: Trends and Challenges Facing Alberta's Displaced Workers*, Background Paper. https://www.businesscouncilab.com/wp-content/uploads/2021/09/BCA_Silent-Crisis-Trends-and-Challenges-Facing-Albertas-Workers-FINAL.pdf.

Carnevale, A. P., Jayasundera, T., & Repnikov, D. (2014). *Understanding online job ads data*. Georgetown University, Center on Education and the Workforce, Technical Report (April). https://cew.georgetown.edu/wp-content/uploads/2014/11/OCLM.Tech_.Web_.pdf

OECD (2015), *Back to Work: Canada: Improving the Re-employment Prospects of Displaced Workers*, Back to Work, OECD Publishing, Paris. http://dx.doi.org/10.1787/9789264233454-en

Statistics Canada (2020), *Table 14-10-0201-01 Employment by industry, monthly, unadjusted for seasonality*.