Canadian Jobs amid a Pandemic: Examining the Relationship between Professional Industry and Salary to Regional Key Performance Indicators

Rahul Anilkumar†, Benjamin Melone†, Michael Patsula†, Christophe Tran†, Christopher Wang†,
Kevin Dick, Hoda Khalil, Gabriel Wainer
Systems & Computer Engineering
Carleton University
{kevin.dick, hoda.khalil}@carleton.ca, gwainer@sce.carleton.ca
†equal contribution

Abstract*—The COVID-19 pandemic has contributed to massive rates of unemployment and greater uncertainty in the job market. There is a growing need for data-driven tools and analyses to better inform the public on trends within the job market. In particular, obtaining a "snapshot" of available employment opportunities mid-pandemic promises insights to inform policy and support retraining programs. In this work, we combine data scraped from the Canadian Job Bank and Numbeo globally crowd-sourced repository to explore the relationship between job postings during a global pandemic and Key Performance Indicators (*e.g.* quality of life index, cost of living) for major cities across Canada. This analysis aims to help Canadians make informed career decisions, collect a "snapshot" of the Canadian employment opportunities amid a pandemic, and inform job seekers in identifying the correct fit between the desired lifestyle of a city and their career. We collected a new high-quality dataset of job postings from jobbank.gc.ca obtained with the use of ethical web scraping and performed exploratory data analysis on this dataset to identify job opportunity trends. When optimizing for average salary of job openings with quality of life, affordability, cost of living, and traffic indices, it was found that Edmonton, AB consistently scores higher than the mean, and is therefore an attractive place to move. Furthermore, we identified optimal provinces to relocate to with respect to individual skill levels. It was determined that Ajax, Marathon, and Chapleau, ON are each attractive cities for IT professionals, construction workers, and healthcare workers respectively when maximizing average salary. Finally, we publicly release our scraped dataset as a mid-pandemic snapshot of Canadian employment opportunities and present a public web application that provides an interactive visual interface that summarizes our findings for the general public and the broader research community.

Keywords Job Market · COVID-19 · Key Performance Indicators · Employment · Geolocation

I. Introduction

The COVID-19 pandemic has disrupted personal lives and economies worldwide. As the result of quarantining measures and locally mandated lockdowns, many small- to mid-sized businesses suffered revenue losses leading either to closures or restructuring initiatives. Studies on Canadian businesses have reported that the initial impacts of the pandemic resulted in a 15% percent decline in employment [1], [2]. The industries and occupations most affected were public-facing occupations in accommodation and food services [2]. In lieu of these uncertainties, many Canadians are pressured to find new employment opportunities. Providing data-driven tools to help identify new opportunities makes the job search a more informed process and can inform governmental policy for retraining programs.

Given the growing need for data-driven tools to better understand trends within the job market, obtaining a "snapshot" of available employment opportunities mid-pandemic promised insights to inform policy and support retraining programs. We addressed this problem by analyzing data from the Canadian Job Bank and various Key Performance Indicators (KPIs) such as the quality of life and traffic indices. These KPIs were extracted from Numbeo [3], a crowd-sourced website which gathers, aggregates, and publishes statistics from cities around the world. By considering factors such as the availability of specific jobs, their average salary, skill levels, and geolocation, we could provide insights into the mid-pandemic Canadian employment landscape and investigate how these correlate to city-level KPIs. We assumed that the current market trends were representative of ongoing mid-pandemic performance although as a caveat, they are expected to vary as subjected to the uncertainty from emergent trends of this global pandemic. Furthermore, we assumed that the jobs listed on the Canadian Job Bank are representative of the greater job market (*i.e.* jobs listed are an unbiased sample of the national job market).

In this paper, we present a mid-pandemic "snapshot" of available employment opportunities and produce data-driven tools and analyses on the Canadian job market due to the COVID-19 pandemic. In section I, we provide a summary of related work, define the KPIs considered, and define the problem of unemployment due to the COVID-19 pandemic by describing both the data sources and approaches used to collect the Canadian job market data. In section II, we describe the data collecting, cleaning, and analyses methodology and outline potential risks and ethical concerns. We also list exploratory questions related to the KPIs and the broader Canadian job market. In section III, we present and discuss our findings to the hypotheses presented in this work. Finally, in section IV, we describe the interactive web platform that was created to enable users to interact with the datasets generated in this work to enable further analyses. We conclude with a summary of the insights obtained from these datasets along with derived recommendations of various locations within Canada considered as "ideal" places of work based on a selected set of diverse employment types including the construction, health, and information technology (IT) professions.

In the following sections we provide a summary of related papers, provide a description of the KPIs that were used in the data analysis stage, and formulate the problem description considered in this work.

A. Related Literature

Foremost, a study by Akkermans *et al.* demonstrated that both positive and negative career shocks will trigger individual job search processes [4]. As expected, in the context of the COVID-19 pandemic, the career shock of unexpected layoffs has impacted many individuals and for those seeking new employment opportunities insights into the current pandemic job market have become critical in supporting the newly unemployed to make more informed career choices. Web scraping applied to job postings has been utilized as a data collection mechanism in various studies. For example, it has been used to determine skill set requirements for medical and white collar occupations [5], [6], and to inform curriculum decisions for computer science courses [7]. Amid the ongoing pandemic, web scraping approaches have been applied to job postings of various countries to study the impacts of COVID-19 on the job market across the United States of America, Sweden and Mexico [1], [8]–[11].

Dias *et al.* acquired historical and live job postings from Find A Job (an American public website operated by the Department for Work and Pensions) through a combination of web scraping and Freedom of Information requests to determine trends in job vacancies across industries. They noted an overall 70% drop in new vacancies from 2019 to 2020 across industries [1]. Forsythe *et al.* analyzed job posting data from Burning Glass Technologies as well as unemployment statistics from the US Bureau of Labor Statistics during the beginning of the COVID-19 pandemic. They found a sharp decline in job postings during the initial wave of the pandemic (average 44% decrease across the first 1.5 months of the pandemic) [8].

This finding was mirrored in Sweden where Hensvik *et al.* found a 40% decrease in job postings across four months during the early COVID-19 outbreaks [9]. They analyzed data on a job board called platsbanken.se which is maintained by the Swedish Public Employment Service. However, the authors augmented their analysis with click-through data which was inaccessible by data scraping, finding that the average clicks per user had also decreased during this period.

Finally, Campos-Vazquez *et al.* scraped data from an undisclosed top-five job board in Mexico to analyze the posting trends in Mexico during the early stages of the COVID-19 pandemic. They provided a comprehensive analysis on the job market trends, detailing an increased demand for low-skill workers [10]. Similar analyses on the impacts of the Great Lockdown on the Mexican Job Market was conducted by Hoehn-Velasco *et al.* where it was reported that the overall job market in Mexico contracted by 5% during the first 9 months of the Great Lockdown [11].

Our work differs from these works in three respects: (1) our work considers the job market trends in the Canadian economy; (2) our work considers the trends *mid-pandemic* in complimentarity to the early-pandemic analyses of the three previously cited works; and (3), we consider KPIs in our analysis to motivate actionable analyses by job-seekers, policy-makers, and the broader research community.

B. Key Performance Indicators

When job searching one may be inclined to look beyond the city they currently reside in to capture all opportunities available. When doing so, it is important to consider other factors than just the compensation and duties of a job. For this reason, various quality of life indicators were incorporated into our analyses. KPIs are values that represent a given performance metric [12] and they were collected from Numbeo, the "world's largest cost of living database" [3].

In this work, the KPIs considered are the *quality of life index* (KPI_q; aggregating purchasing power, cost of living, healthcare, safety, *etc.*), the traffic index (KPI_t; aggregating time and CO_2 consumed by commute), the *affordability index* (KPI_a; the inverse of mortgage as percentage of income), and the *cost of living with rent index* (KPI_c; aggregating consumer good prices and rent).

More specifically, the *quality of life index* (KPI_q) was computed as a composite of several other KPIs. Equation 1 defines the calculation used for determining the quality of life index [13]:

$$\begin{aligned}\text{KPI}_q = max(0, \frac{p}{2.5} - h - \frac{c}{10} + \frac{s}{2} + \frac{m}{2.5} - \\ \frac{t}{2} - \frac{2n}{3} + \frac{l}{3} + 100)\end{aligned} \tag{1}$$

where p is the purchasing power index, h is the average house price to income ratio, c is the cost of living index, s is the safety index, m is the health index, t is the traffic index, n is the pollution index and lastly l is the climate index. The constants in Equation 1 were selected empirically by Numbeo [3]. Of note, these constant values are subject to change over time.

The component purchasing power index represents the amount of goods a person can buy with an average salary in a particular region. Numbeo calculates the purchasing power "relative to the purchasing power in New York City" (a purchasing power of 100 represents the purchasing power in New York City) [14]. As defined by Numbeo, the compnent health care index is an "estimation of the overall quality of the health care system" [15]. This estimation is determined by using values such as the cost of health care, the health care equipment available and the experience of the health care professionals within the analyzed region [15]. The component climate index is computed by using historical temperature and humidity data to determine whether the weather conditions in a certain region is liked by most people. The higher the value, the better one can expect the climate to be [16]. Finally, the component pollution index uses air pollution data to represent the overall pollution in a particular region. The pollution index also "incorporates water pollution and other types of pollution" to arrive at the value [17].

The *traffic index* (KPI_t) is used to provide insight into the traffic-related impact one can expect within a certain city. The larger the index value, the worse the impact of traffic is within the considered region [18]. Equation 2 was leveraged to determine the *traffic index* in a given city [18]:

$$\text{KPI}_t = t + \sqrt{t + (t-25)^e} + \sqrt{c} + \sqrt{i} \tag{2}$$

where t is an index of the average time (in minutes) spent in compute per day, c is the CO_2 emission index, and i represents an index encompassing the inefficiencies of traffic. The value 25 in Equation 2 was selected empirically by Numbeo. This value represents an assumed 25 minute commute time before an individual begins to be dissatisfied with the commute [18].

The *affordability index* (KPI_a) refers to the "inverse of mortgage as percentage of income" [19]. The *cost of living plus rent index* (KPI_c) is an indicator of the consumer goods and rent prices. The index value is relative to the cost of living in New York City and replicated from the work in [14].

C. Problem Definition

With the increasing unemployment, cost of living, and transition towards a remote workforce, a trend of the COVID-19 pandemic has seen a mass exodus of individuals moving away from city centers [20]. Consequentially, this work seeks to identify those Canadian cities or provinces best suited to individuals in specific professions, based on the mid-pandemic Canadian job market while accounting for regional factors as well as various KPIs.

The current job market was analyzed by scraping job posting data from the Canadian Job Bank [21]. This is a service provided by Employment and Social Development Canada to help Canadians find employment opportunities across the country. Each job posting contained information such as the job title, compensation, location, and importantly the National Occupation Classification (NOC) number.

The NOC number are Canada's national system for describing occupations and identifying the skill levels of a given occupation [22]. A dataset containing all NOC numbers, related occupation class, skill levels, and description was collected from the Open Canada online database [22], [23]. Since the job board collaborates with provincial governments and external job boards to gather job postings, we considered the Job Bank to be a representative sample of the current hiring ecosystem of the mid-pandemic Canadian job market.

II. Data & Methodology

The following section outlines all of the data sources considered;, their means of acquisition, aggregation, transformation, and cleaning; and defines the ethical concerns and reproducibility of this work. A conceptual overview of this processing pipeline is illustrated in Fig. 1.

A. Approach

Our work sought to explore employment-based trends that could be used to provide insights on provincial, territorial, and city characteristics that might inform individuals seeking relocation or skill-retraining. We base our findings on an outlier-detection search while considering specific KPIs of interest in combination with employment titles/profession and reported hiring salaries.

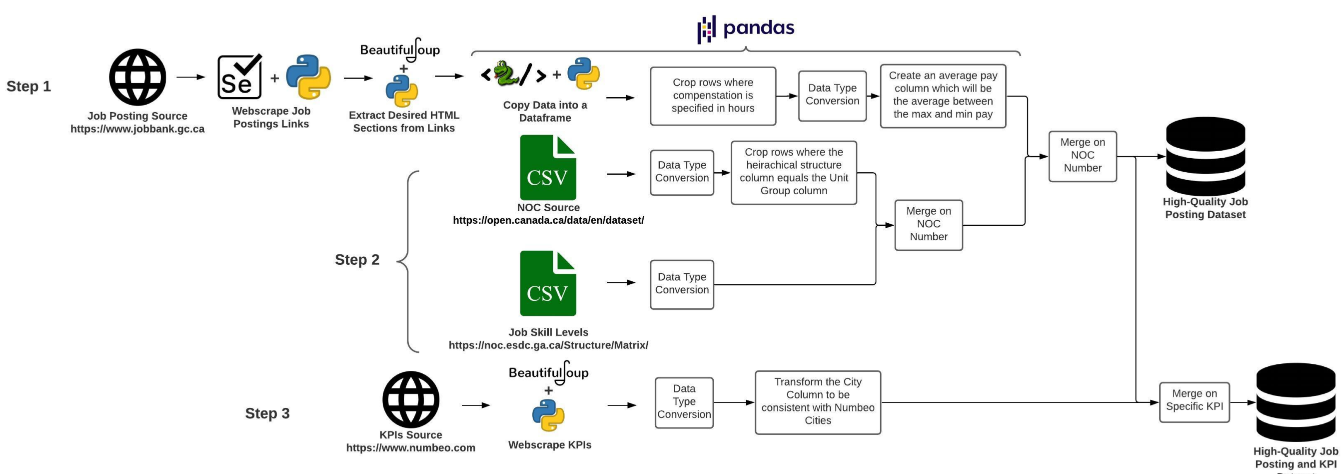

Fig. 1. **Overview of the data Acquisition, Cleaning and Transformation Pipeline.** Two datasets were created, one which merges data from job postings from the Canadian Job Bank with NOC-related fields and another which incorporates data from the first dataset and a specific KPI found sourced from Numbeo. The job postings and KPIs from Numbeo were web-scrapped using Beautifulsoup4 and/or Selenium. Python was used throughout the pipeline to scrape, clean, and transform the data into a high-quality dataset.

The specific questions that we considered were:

(A) What is the *distribution of the number of jobs* with respect to their *job levels across provinces*?
(B) What is the *distribution of the average salary* of jobs with respect to their *job levels across provinces*?
(C) How is the *quality of life index* related to the *average hiring salary*?
(D) How is the *affordability index* related to the *average hiring salary*?
(E) How is the *cost of living index* related to the *average hiring salary*?
(F) How is the *traffic index* related to the *average hiring salary*?
(G) What *cities are optimal* with regards to *hourly wage for specific professions* (*e.g.* construction, healthcare, IT)?

Our aim was to combine the KPI data with the job posting data to form a high-quality dataset that could be used to identify correlations and trends to gain insight on these questions. Our analysis is unique as it considers the current hiring salaries (how much will one be earning) and therefore is future-facing. In contrast, other analyses typically use current or historical salary information which might not be relevant for making future career and relocation decisions.

B. Data Collection

The steps for Data Collection and Data Cleaning are depicted in Fig. 1. This figure illustrates how data was obtained from the web and transformed into a cleaned and high-quality dataset.

The first step in the data collection process was to scrape the job details for each listing on the job board. A variety of details were gathered including the job title, pay, and geolocation. This process was broken down into three steps:

1) Executing a python script which web-scraped and collected 91,251 hyperlinks of Job Bank job posts and stored them in a compressed format. This process used the Selenium library which required a browser-specific webdriver. In this work, we leveraged the *geckodriver* for Firefox.
2) Running a Python script which loaded the HTML for each of the collected hyperlinks and extracted relevant HTML components which are stored in a separate set of compressed files. To ensure graceful recovery, we executed every 1,000 links and confirmed successful acquisition upon restarts.
3) Finally, we used a Python script to extract relevant details from the extracted HTML and transformed it into a CSV format. Posts were dropped if they were no longer available on the job board, leaving a total of 64,945 rows.

The second step was to obtain the NOC dataset from Open Canada online database [23]. This dataset was used to match with the NOC numbers for each job posting in order to identify the job classification title. A Python script was executed to map a NOC designation to its job skill level according to the National Occupational Classification Matrix [23].

The third step for the data collection process was to web-scrape Numbeo in order to extract the KPI values for each specific city in Canada. This involved executing a Python scraper to collect the data for each KPI category (*e.g.* quality of life, traffic, health, *etc.*). To reproduce our study and facilitate similar snapshot acquisition and analyses in a terminal-pandemic and post-pandemic future, our code is open-sourced and available on GitHub [24].

C. Data Cleaning

After collecting all of the raw data, the data was cleaned to produce a high-quality aggregate dataset to simplify data exploration and subsequently analyses. An overview of the complete processing pipeline is illustrated in Fig. 1. Initially, our four CSV files with data coming from various sources (*e.g.* job postings, Numbeo, NOC) were each imported into a separate dataframe and then either type-cast into numerical representations or stripped of whitespace if a text-field. As depicted in step 1 of Fig. 1, the job postings lacking hourly compensation fields were then removed. Where posted salaries reported a salary range, a new column comprising the average between the minimum and maximum compensation was added; otherwise the field was conservatively set to the minimum reported salary. For the NOC dataframe (step 2 in Fig. 1), all rows where the *Hierarchical Structure* did not equal the *NOC Unit group* were deleted. Each NOC number belongs to a specific hierarchical category. Each category generalizes what field of work is accomplished. There are four levels of NOC numbers and the Unit group is the most precise NOC number. Only the NOC numbers which represent a Unit group were kept because only the Unit group codes represented a full NOC code for an actual occupation title.

Finally, the job-level dataframe and NOC dataframe were merged with the job posting dataframe on the NOC column resulting in our finalized high-quality job posting dataset. To summarize, the available fields within this dataset are:

NOC Class title, NOC Class definition, NOC Number, Job Level, Job Title, Raw Pay, Minimum Pay, Maximum Pay, Pay Unit, Industry Average Pay, Work Hours, Start Date, Employment Terms, Full or Part Time, Vacancies, Special Commitment, Benefits, Median Wage, Date Modified, Hiring Organization, Job Source, NOC Title, Education Requirements, Qualifications, Experience Requirements, Responsibilities, Skills, Language, Posting ID, HREF, Address, City, Postal Code,

To create the job posting and KPI dataset (refer to step 3 in Fig. 1) the city name in the Numbeo dataframe had to be transformed to a city name and province format so that they could be joined with the job posting dataset. Following this transformation, a specific KPI could be joined with the job posting dataset to create the high-quality job posting and KPI dataset.

D. Ethical Concerns

Websites may implement a variety of solutions in order to safeguard and protect against potential web scraping and unauthorized data access [25]. Techniques such as defining a terms of service policy can restrict the actions that users can take on the platform [26]. Due to this, the relevant user policies were analyzed before any web scraping activities were performed.

The first website that was scraped was the Job Bank job board which is offered by the Canadian Government. Non-commercial reproduction of the data is allowed without charge or further permission assuming that due diligence is ensured. In addition, the terms and conditions document requires that the original source is referenced in the reproduction of the data [27].

The second website that was scraped was Numbeo. Numbeo provides data outlining KPI's for cities across the globe. The metrics for Canadian cities were scraped. The terms and service state that use of Numbeo data for academic purposes is allowed so long as a link back to the Numbeo website is included [28].

In accordance with ethical scraping practices [25], we performed due diligence and ensured a sufficient sleep-time delay between subsequent requests to avoid overloading the web servers [29]. This ensured that the amount of traffic sent to the websites was sent in a controlled manner.

The remaining datasets were acquired through open-source resources and the primary ethical concern had to do with potential biases within the job data collected. Certain skill levels of jobs might be over-represented or under-represented. For example, as a limitation to this work, if we assume a uniform distribution of jobs by skill-level, our data has an over-representation of NOC Level B jobs. This could potentially lead us to incorrect recommendations to our readership. In addition to this skill level over-representation, some provinces have a proportionally higher number of jobs. As an example, there are very few job postings from the territories and East-coast provinces compared to those in larger provinces such as Ontario. A future improvement to this paper could be to extract job postings from other Canadian job boards to reduce potential biases in job-level postings and using a frequency normalized evaluation.

E. *Reproducibility*

As previously mentioned, this work has open-sourced all code for reproducibility [24] and the following summarizes the high-level steps required to reproduce this paper:

1) Web scrape the job postings in the Canadian Job Bank.
2) Web scrape Numbeo for the KPIs (Canada).
3) Download the NOC data sets.
4) Merge the NOC, job posting data set, and KPI datasets.
5) Clean the merged data sets.
6) Leverage Python libraries to analyze and visualize the cleaned datasets.

The detailed steps, code, and datasets used to develop this paper can be found at [24].

III. Results & Discussion

At the mid-point of the COVID-19 pandemic, it is of critical importance to explore and understand the employment landscape of various nations to better inform governmental policy and improve re-skilling and unemployment programs. The insights of this work are limited to the snapshot of employment opportunities within Canada during the first half of 2021. Specifically, these data were collected between 2021-03-25 and 2021-04-04.

At the outset, it was insightful to determine the distribution of all wages on the job board. In Fig. 2 we illustrate a histogram to visualize this distribution. A modal 20CAD/hr

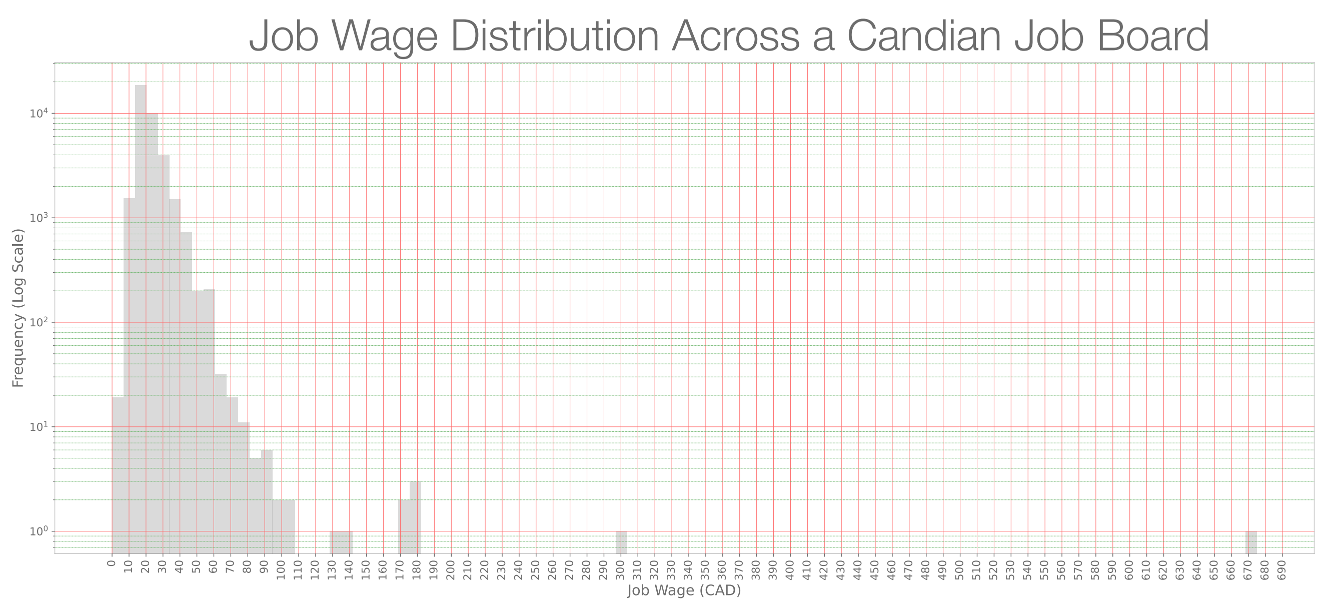

Fig. 2. **Histogram of Hourly Job Wage from Job Board.** The graph displays the distribution of the job wages proposed by the job postings collected. The job wages are measured in Canadian dollars per hour. The modal wage is 20 dollars per hour and most wages distribute around this wage.

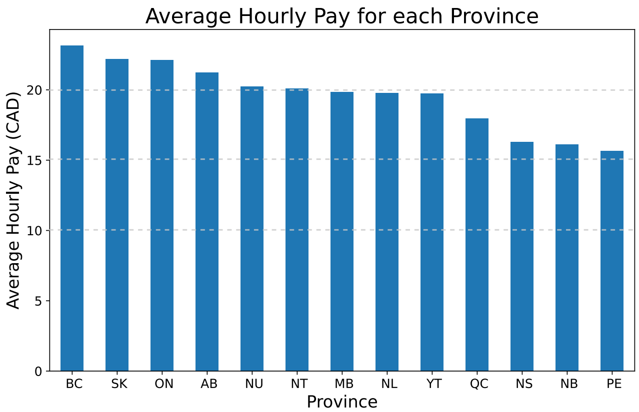

Fig. 3. **Distribution of all wages across all Canadian Provinces and Territories during the Pandemic months.** The data depicted in this graph was collected from job posting on the Canadian Job Bank. The provinces are sorted by the highest average hourly pay.

wage is immediately apparent and it is important to note the logarithmic frequency axis indicative that the vast majority of wages are distributed proximal to the distribution mode (Fig. 2). Notably, the number of wages in the [0,10] range may suggest that our study under-represents minimum wage and/or tip-based employment.

To then understand the distribution of wages across the provinces and territories, Fig. 3 visualizes the distribution of hourly wages, ordered in decreasing average wage. British Columbia, Saskatchewan, and Ontario appeared to have the job postings with the highest average hourly pay, while Prince Edward Island had the lowest. This would make certain provinces appear more attractive than others if the only criteria considered is average pay, a regional proxy for economic prosperity [30].

Conceptualizing an "ideal" region or city of employment is subjective and depending on individual worker's priorities. To further characterize regional employment promise, we pose several questions to better realize the ideality of provinces and cities with respect to employment types while mid-pandemic. The key metrics considered include job compensation, the quality of life index, pollution emitted in the city (pollution index), and the level of health services available (health index). Individuals also need to consider what professions and specific skills are in demand and need to adapt to market trends.

Throughout our analyses, we generated visualizations and each are made available on an interactive web platform, discussed later (section V). Notably, two interactive and qualitative figures lend high-level insights to job characteristics across Canada within this snapshot and the related skills required.

Foremost, a choropleth map of Canada was used to represent the geographical data and to summarize the KPI and job salary data using an intuitive interactive interface. This ensured that the quality of life information corresponding to each province was easily accessible. All key information was visually represented using a color coded scale as well as text boxes for fine grained details. Fig. 4 illustrates the choropleth identifying the average pay and the key metrics for each province.

Secondly, a word cloud was generated to visualize the common words among the text-based data series of the skills listed in each job post. In the context of desired skills for a job, it is useful to quickly determine the skills which are trending in greater demand by employers (Fig. 5).

Finally, and more quantitatively, we then sought to address the individual questions posted in Section II-A.

A. What is the distribution of the number of jobs with respect to their job levels across provinces?

To address this question, we visualize in Fig. 6 the total jobs in a province based on job level. Job levels refer to the skills required to perform a job; these job levels were obtained by mapping the National Occupation Classification (NOC) for each job on the Canada Open Job Bank with the NOC dataset retrieved from Statistics Canada Open Government. The NOC values group jobs into five skill levels:

- **Skill Level 0**: Management jobs
- **Skill Level A**: Professional jobs requiring a university degree
- **Skill Level B**: Technical jobs and skill trades requiring a college diploma or apprentice training
- **Skill Level C**: Intermediate jobs requiring high school diploma or job-specific training
- **Skill Level D**: Labour jobs requiring on-the-job training

From the data displayed within Fig. 6, we can see that depending on the province, the number of total jobs posted appears disproportionate. For instance, the number of jobs posted between provinces such as British Columbia and New Brunswick are vastly different. This difference is expected considering the size of major cities. Provinces such as British Columbia, Alberta, Ontario, and Quebec host many of the major metropolitan hubs within Canada and this may be the reason why there are more positions in these regions. In addition, we can see that British Columbia, Quebec, Ontario, and Alberta have more Technical and Intermediate jobs (levels

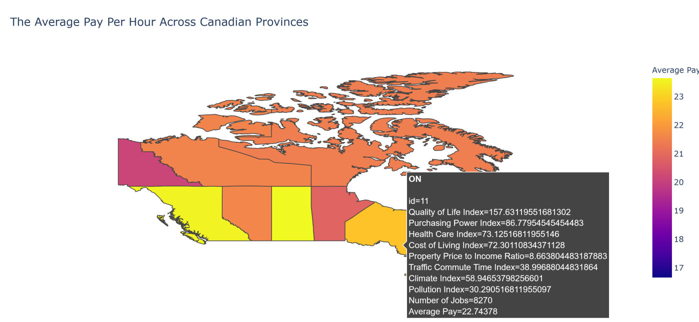

Fig. 4. **Interactive Choropleth map of the Average Pay Per Canadian Province.** When a user hovers their cursor over a certain province additional information is displayed such as various KPIs and the number of job posting analyzed to compute the average pay relating to the province being hovered over.

Fig. 5. **Word Cloud for Job Skills.** The word cloud depicts the most common job skills demanded within the job posting collected. From the word cloud it is evident that *first-aid*, *CPR*, *WHMIS knowledge*, and a *driver's license* are among the most common skills demanded from job posting within the Canadian Job Bank.

B and C) than other job levels, while P.E.I. has more Labour jobs (level D) than any other province or territory. These results indicate that P.E.I. has a high demand for Labour jobs relative to its population size.

B. What is the distribution of the average salary of jobs with respect to their job levels across provinces?

From Fig. 6, we obtain some insight on the distribution of jobs by skill level, however, it does not consider the starting salary of the various positions. Fig. 7 plots the distribution of the jobs by skill level for each province against the average starting salary. This distribution demonstrated that for jobs of skill levels B, C, and D the average pay is approximately consistent across the country. From this data it is evident that

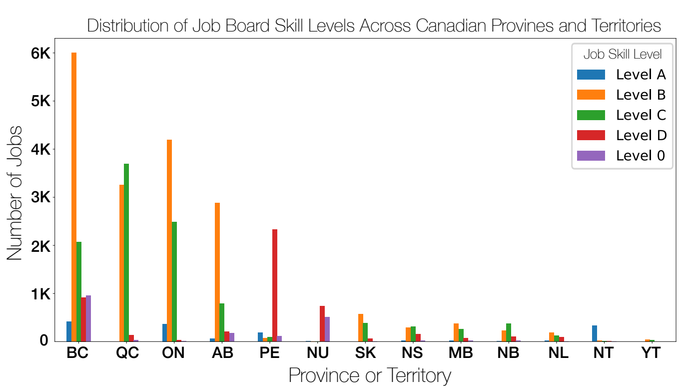

Fig. 6. **Distribution of Job Skill Levels on the Job Board across Canadian Provinces and Territories.** The data used in this graph was collected from the Canadian Job Board within March 2021. The provinces and territories displayed in the graph are sorted by the number of jobs across each job skill levels from highest to lowest.

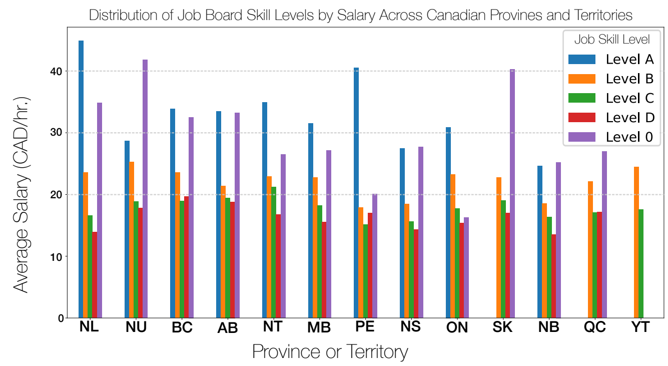

Fig. 7. **Distribution of Salary on Job Board by Job Skill Level across Canadian Provinces.** The data used in this graph was collected from the Canadian Job Board within March 2021. The provinces and territories displayed in the graph are sorted by the average salary across each job skill levels from highest to lowest.

working in a higher skilled position pays more when working in Newfoundland and Labrador, the North West Territories, Nunavut, Prince Edward Island and Saskatchewan. These provinces may provide higher pay for these positions as they need to entice highly educated and trained individuals to move to more rural, northern, or lower-resource areas.

C. How is the quality of life index related to the average hiring salary?

To address this question, Fig. 8A depicts the average pay to quality of life index for various cities in Canada. The plot has a slight positive (though negligible) correlation between the average hiring pay and the quality of life index ($R^2 = 0.027$). The reason for the slight correlation is likely explained by the fact that the quality of life index accounts for the cost of living in the city when computing it's value, and it is expected to be similar to the average hiring pay in the city. From the graph it is evident that Mississauga, ON and Edmonton, AB are outliers with abnormally high average hiring pay to quality of life index value. Therefore, if one wants to maximize their average pay and quality of life, it is recommended that one moves to Mississauga, ON or Edmonton, AB.

D. How is the affordability index related to the average hiring salary?

An ideal location to move should have a housing market that is affordable given the average job compensation for a city. Based on the data displayed within 8B, there is a negative correlation between the average starting salary and affordability index ($R^2 = 0.343$). From the data it is evident that Saskatoon, SK, Edmonton and Calgary, AB are the ideal places to live considering only the average pay of the jobs within the respective city and its affordability index. These cities are ideal due to the relatively high affordability index in relation to the average salary. This ultimately means one can financially afford a better lifestyle working in these cities more than any other city in Canada.

E. How is the cost of living index related to the average hiring salary?

Fig. 8C displays average pay versus cost of living plus rent index. The data shows that there is a positive correlation between the average pay and the cost of living plus rent index ($R^2 = 0.235$). This correlation makes sense because as the cost of living increases, the average pay will also increase. From the data, Saskatoon, SK and Edmonton, AB appear to have higher than average pay, while having a low cost of living index plus rent index. This means those cities are ideal to live in if one wishes to have a high income while having low cost of living.

F. How is the traffic index related to the average hiring salary?

To address this question, Fig. 8D displays the traffic index per city and the average pay. There is a positive correlation between the traffic index and the average pay ($R^2 = 0.283$). This is likely a result of potentially high correlations between commute time and wages against the increased number of workers in larger cities [31]. Cities that deviate from the mean in a beneficial way include Mississauga, ON and Edmonton, AB. This suggests that these cities are optimal if one's goal is to maximize average salary while minimizing commute time.

G. What cities are ideal with respect to the hourly wage for specific professions (Construction, Healthcare, IT)?

Cities generally attract certain industries. For instance, Silicon valley is famously known to have a developed, high-paying tech industry. For this reason, those looking for a given type of job would be drawn to particular provinces and/or cities. This is visually depicted over three general types of professions in Fig. 9. We selected three professions of focus: the low-skilled construction profession, the pandemic-related healthcare profession, and the high-skilled IT profession.

Fig. 9A ranks the top-20 cities by the highest mean hourly wage for the construction industry. The data set was queried using key terms found in the job title. These results suggest that the highest paid construction jobs are found in Marathon, ON.

Fig. 9B ranks the top-20 cities by the highest mean hourly wage for the healthcare industry. From these rankings, it can be observed that the majority of these positions are situated in smaller towns and cities. The highest paid positions are available in Chapleau, ON. In addition there seems to be a greater demand for health industry professionals in the province of Saskatchewan (that is 14/20).

Fig. 9C ranks the top-20 cities by the highest mean hourly wage for jobs that have an IT-related title within the job board. From this graph, it is clear that Ajax, ON may have the highest paying IT jobs. An important consideration when interpreting this graph is that due to the limited data available, some cities might be represented by an insignificant number of job postings. In the future, a frequency-weighted wage averaging may be considered.

Fig. 8. **(A) Average Pay vs the Quality of Life Index for Canadian Cities, (B) Average Pay vs the Affordability Index for Canadian Cities, (C) Average Pay vs the Cost of Living Index for Canadian Cities, (D) average Pay vs the Traffic Index for Canadian Cities.** Each of the graphs compares the average salary proposed within the job postings collected (during the pandemic) for specific cities with a KPI sourced from Numbeo.

Further analysis on other industries can be conducted using the web application provided within this work. For example, users and the broader research community can input industry identifying key terms to automatically query our dataset for further insight into industries of interest.

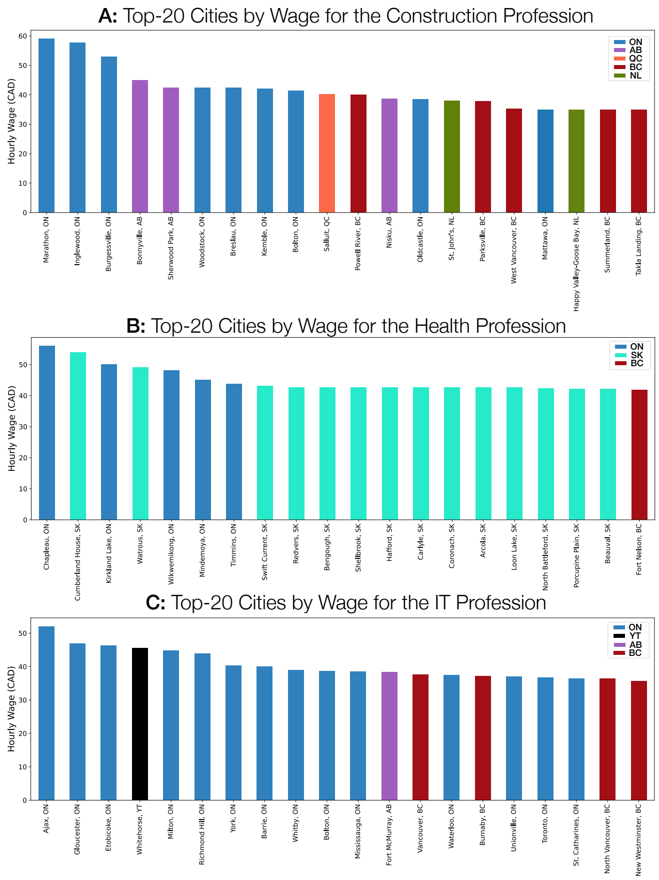

Fig. 9. **Highest Hourly Wage for Construction (A), Health (B), and IT (C) Professionals per Canadian City.** The data used to build each of these graphs was collected from the job posting web scrapped on the Canadian Job Bank during the pandemic. To determine whether a job is related to a certain profession the title of each job posting was queried to see if it contained industry specific keywords. For example, if the title of a job posting contained the words software, computer, information, web or data it was considered an IT related job posting.

IV. Interactive Web Platform

An interactive web platform was developed as a research tool using Dash Plotly. The tool was designed to allow users access to the processed data sets for further analysis and study with regards to a snapshot of the Canadian job market during the COVID-19 pandemic. The web platform can be found through GitHub: .

The platform enables users to analyze trends between various KPIs and the Canadian industry average pay for different locations. Additionally, users are able to utilize the extracted job posting data set to display a word cloud depicting words that appear more frequently in the Canadian job bank for a specific keyword metadata. This may include identifying the most common skills listed in job posts, most common job titles, and most common qualifications required in job posts. Lastly, users can input industry key terms to identify top paying cities related to the industry of interest.

The aim for the web platform is to provide an interactive data set of the Canadian Job market during a pandemic as a resource to the broader research community, policy makers, and ultimately, the general population. Our scrapers, curated data, and interactive interface are expected to prove useful for broader time-series analyses of Canadian employment opportunities post-pandemic by means of methods such as Regression Discontinuity Analysis [32].

V. Conclusion

This paper provides Canadians and policy-makers with mid-pandmic job market insights needed to select a putative work location to when considering a number of factors such as quality of life, traffic, cost of living, and affordability among others.

When maximizing the average hiring salary while considering the KPIs analyzed, it was found that Edmonton, AB was a notable outlier among all categories considered and can thus be considered an ideal city of employment. For construction, healthcare, and IT professionals, when maximizing the agerage hiring salary, Marathon, Chapleau, and Ajax, ON are respectively the cities that rank highest.

An interactive web platform was created to help researchers, policy makers, and the general population access the processed dataset of the Canadian industry during a pandemic. This promises to enable post-pandemic studies and highlight differences in the Canadian employment opportunity landscape.

In the future, this paper could be extended by gathering more job posting data from other external job boards. This would ensure that there is a larger sample of jobs to analyze and would help decrease the bias with regards to the type of industries posting. In addition to existing interactive visualization, more interaction features could be added to help better present the information. To support the replication of this work, all code and our dataset are available open-source: [24].

References

[1] M. C. Dias, A. N. Keiller, F. Postel-Vinay, and X. Xu, "Job vacancies during the covid-19 pandemic," 2020.

[2] T. Lemieux, K. Milligan, T. Schirle, and M. Skuterud, "Initial impacts of the covid-19 pandemic on the canadian labour market," *Canadian Public Policy*, vol. 46, no. S1, pp. S55–S65, 2020. [Online]. Available: https://doi.org/10.3138/cpp.2020-049

[3] "Numbeo," accessed: April 7, 2021. [Online]. Available: https://numbeo.com/common/

[4] J. Akkermans, D. G. Collings, S. P. da Motta Veiga, C. Post, and S. Seibert, "Toward a broader understanding of career shocks: Exploring interdisciplinary connections with research on job search, human resource management, entrepreneurship, and diversity," *Journal of Vocational Behavior*, vol. 126, p. 103563, 2021, 50th Anniversary Issue. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S000187912100035X

[5] J. Schedlbauer, G. Raptis, and B. Ludwig, "Medical informatics labor market analysis using web crawling, web scraping, and text mining," *International Journal of Medical Informatics*, vol. 150, p. 104453, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1386505621000794

[6] M. Dillender and W. E. Upjohn, "White collar technological change: Evidence from job posting data," 2019.
[7] J. Schuessler and D. Jones, "Web scraping to inform curriculum decisions: An exploratory study," 03 2018.
[8] E. Forsythe, L. B. Kahn, F. Lange, and D. Wiczer, "Labor demand in the time of covid-19: Evidence from vacancy postings and ui claims," *Journal of Public Economics*, vol. 189, p. 104238, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S004727272030102X
[9] L. Hensvik, T. Le Barbanchon, and R. Rathelot, "Job search during the covid-19 crisis," *Journal of Public Economics*, vol. 194, p. 104349, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0047272720302139
[10] R. M. Campos-Vazquez, G. Esquivel3, and R. Y. Badillo, "How has labor demand been affected by the covid-19 pandemic? evidence from job ads in mexico," *CEPR Press*, vol. 46, 2020. [Online]. Available: https://www.banxico.org.mx/publications-and-press/articles-and-other-publications/%7B54956C5E-E33B-3ABD-C545-F2FD7550DD64%7D.pdf
[11] L. Hoehn-Velasco, A. Silverio-Murillo, and J. R. Balmori de la Miyar, "The long downturn: The impact of the great lockdown on formal employment," *Journal of Economics and Business*, p. 105983, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0148619521000011
[12] "What is a key performance indicator (kpi)?" accessed: April 7, 2021. [Online]. Available: https://kpi.org/KPI-Basics
[13] "About quality of life indices at this website," accessed: April 7, 2021. [Online]. Available: https://www.numbeo.com/quality-of-life/indices_explained.jsp
[14] "About cost of living indices at this website," accessed: April 7, 2021. [Online]. Available: https://www.numbeo.com/cost-of-living/cpi_explained.jsp
[15] "About health care indices at this website," accessed: April 7, 2021. [Online]. Available: https://www.numbeo.com/health-care/indices_explained.jsp
[16] "About climate indices at this website," accessed: April 7, 2021. [Online]. Available: https://www.numbeo.com/climate/indices_explained.jsp
[17] "About pollution indices at this website," accessed: April 7, 2021. [Online]. Available: https://www.numbeo.com/pollution/indices_explained.jsp
[18] "About traffic indices at this website," accessed: April 7, 2021. [Online]. Available: https://www.numbeo.com/traffic/indices_explained.jsp
[19] "About property value and investment indices at this website," accessed: April 7, 2021. [Online]. Available: https://www.numbeo.com/property-investment/indicators_explained.jsp
[20] L. Ioannou, "Vast migration of over 14 million americans coming due to rise in remote work, study shows," October 2020, accessed: April 7, 2021. [Online]. Available: https://www.cnbc.com/2020/10/28/vast-migration-of-over-14-million-americans-coming-due-to-remote-work.html
[21] "Job bank," March 2020, accessed: April 7, 2021. [Online]. Available: https://www.jobbank.gc.ca/aboutus
[22] "National occupational classification matrix," March 2019, accessed: April 7, 2021. [Online]. Available: https://noc.esdc.gc.ca/Structure/Matrix/23990881df3e417f8cdb9046fb1efe22
[23] "National occupational classification (noc) 2016 version 1.3," accessed: April 7, 2021. [Online]. Available: https://open.canada.ca/data/en/dataset/2192de22-84bf-4672-b0c0-f64c0cb43c94
[24] "Github - canadian-jobs-amid-a-pandemic," accessed: April 7, 2021. [Online]. Available: https://github.com/ChristopheTran/Canadian-Jobs-Amid-a-Pandemic
[25] A. Luscombe, K. Dick, and K. Walby, "Algorithmic thinking in the public interest: navigating technical, legal, and ethical hurdles to web scraping in the social sciences," *Quality & Quantity*, pp. 1–22, 2021.
[26] J. K. Hirschey, "Symbiotic relationships: Pragmatic acceptance of data scraping," *Berkeley Tech. LJ*, vol. 29, p. 897, 2014.
[27] "Terms and conditions," April 2021, accessed: April 7, 2021. [Online]. Available: https://www.canada.ca/en/transparency/terms.html
[28] "Terms of use - numbeo.com," accessed: April 7, 2021. [Online]. Available: https://www.numbeo.com/common/terms_of_use.jsp
[29] V. Krotov and L. Silva, "Legality and ethics of web scraping," 2018.
[30] M. E. Porter, "The economic performance of regions," in *Regional Competitiveness*. Routledge, 2012, pp. 137–166.
[31] D. Timothy and W. C. Wheaton, "Intra-urban wage variation, employment location, and commuting times," *Journal of Urban Economics*, vol. 50, no. 2, pp. 338–366, 2001.
[32] D. L. Thistlethwaite and D. T. Campbell, "Regression-discontinuity analysis: An alternative to the ex post facto experiment." *Journal of Educational psychology*, vol. 51, no. 6, p. 309, 1960.