Data Visualization Report

Wages By Occupation Across Canada
Submitted to Dr. Sanaz Mohammad Jafari

Sonal Bisla **4/14/23** **DS8007**

TABLE OF CONTENTS

1. INTRODUCTION

Occupations play a vital role in the economy, shaping the labor market and influencing the overall well-being of workers and employers alike.

In Canada, the National Occupational Classification (NOC) system is a standardized framework that classifies and describes occupations. The NOC system is maintained by the Government of Canada and is used by various federal departments and agencies, provincial and territorial governments, and other organizations to collect and analyze labor market data.

The NOC system groups occupations based on the type of work performed and the skills, education, and training required. It assigns a unique four-digit code to each occupation, which facilitates consistent data collection and analysis across different sources.

The "Wages by Occupation across Canada" project aims to address the lack of clarity and transparency in wage information for various occupations throughout the country. By providing an accessible and user-friendly source of wage data, this project empowers workers and employers to make more informed decisions about career paths, salary negotiation, and job satisfaction.

The primary focus of the project is to create data visualizations that enable users to explore wage data and understand the patterns, trends, and anomalies if any present in the information.

2. DATA DESCRIPTION

The data used for this project has been sourced from publicly available resources such as Dataset Search, the Government of Canada Open Data portal, and Supporting Documents from Statistics Canada.

The dataset is comprehensive, containing information on wages for different occupations, NOC, type of work, year, provinces, age groups and genders etc. This rich dataset allows for a deep analysis of wage trends and disparities across various dimensions, offering valuable insights into the labor market in Canada.

The primary data sources URLs:

- **Dataset Search:** ***https://datasetsearch.research.google.com/search?src=2&query=Employee%20wages%20by%20occupation%2C%20annual&docid=L2cvMTF0eGhnejVzYw%3D%3D***

- **Government of Canada Open Data portal:** ***https://open.canada.ca/data/en/dataset/f0f63701-d4bd-416b-8ed2-7a09f74abc6e***

- **Supporting Documents from Statistics Canada:** ***https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1410041701***

3. EXPLANATORY DATA ANALYSIS (EDA)

Exploratory Data Analysis (EDA) is a crucial step in any data-driven project as it helps analysts understand the data's structure, characteristics, and relationships between variables. By uncovering patterns, trends, and anomalies, EDA guides data cleaning and preprocessing, ensuring the data is reliable for further analysis.

It also assists in formulating hypotheses, selecting appropriate statistical methods, and performing feature engineering to improve model performance. Moreover, EDA enables effective communication of insights through visualizations, empowering decision-makers to take informed actions based on the data.

The EDA process for this project comprises several steps that help transform the raw data into meaningful insights.

3.1 DATA LOADING AND PROCESSING

I created a list called chunks to store the processed chunks, and then use pd. c oncat() to concatenate all chunks into a single Data Frame. This approach is -e fficient because it avoids the deprecation warnings, and it is easy to read the bi -g data size of around 1 GB.

3.2 DATA FILTERING

Filtered the data for the years from 2018 to 2022, for full-time employees only and for the median hourly wage rates.

3.3 DATA AGGREGATION

Grouped the data by the National Occupational Classification (NOC) and calculated the mean wage for each occupation.

3.4 DATA SORTING

Sorted the average wages in descending order, which will help identify the occupations with the highest and lowest wages.

3.5 FEATURE ENGINEERING

Created a new column for wage ranges (Low, Medium, High, and Very High) by defining a function called `wage_range`.

3.6 DATA VISUALIZATION

Utilized diverse interactive visualizations, including charts and plots, to efficiently examine patterns and trends, allowing for a thorough investigation of the research questions.

Graphs Used:

- **Nested Treemaps**
- **Horizontal histograms**
- **Choropleth map**
- **Pie Chart**
- **Bar chart**
- **Line charts**
- **Parallel Coordinate Plot**

4. FINDINGS

Based on the EDA, the following findings were observed:

- **Average wages across different occupations in Canada vary significantly with respect to the categories of the wage range starting from low to medium and high to the highest.**

- **Wages vary by province, with some provinces having higher average wages than the others. Like Alberta being the province with the highest average wage while Prince Edward Island being the lowest province with respect to the average wage.**

- **There are clearly high-wage and low-wage occupations. Legislative and senior management occupations [00] (NOC) being the highest paid occupations while the Sales and service support occupations [65] are the lowest paid occupations.**

- **Wages across different occupations have changed over the last 5 years, with some occupations experiencing a higher rate of wage growth while fluctuation is observed for few others.**

- **Gender pays gap exist across different occupations in Canada, with men generally earning more than women for most of the occupations while for few others, the wage has been equalized for both the genders over the recent years.**

5. RESULTS

The results of this project include a series of data visualizations, such as interactive maps, parallel coordinate plots, and bar charts, etc. that help users explore wage data across different occupations in Canada. These visualizations provide insights into trends and patterns in the data.

Research Question-1 Charts

Fig 1.1

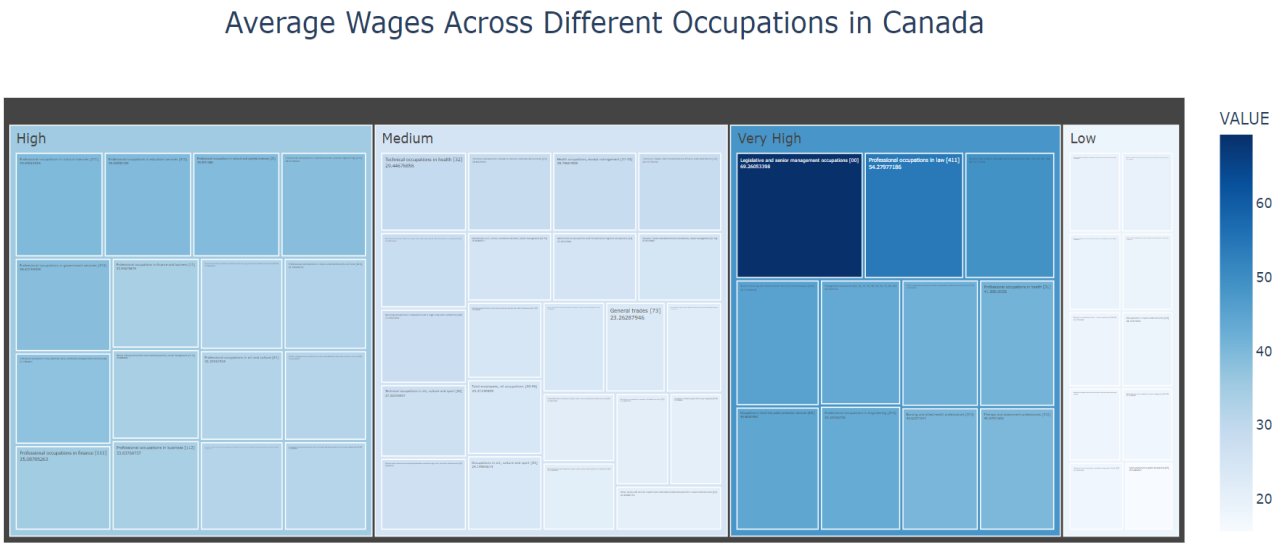
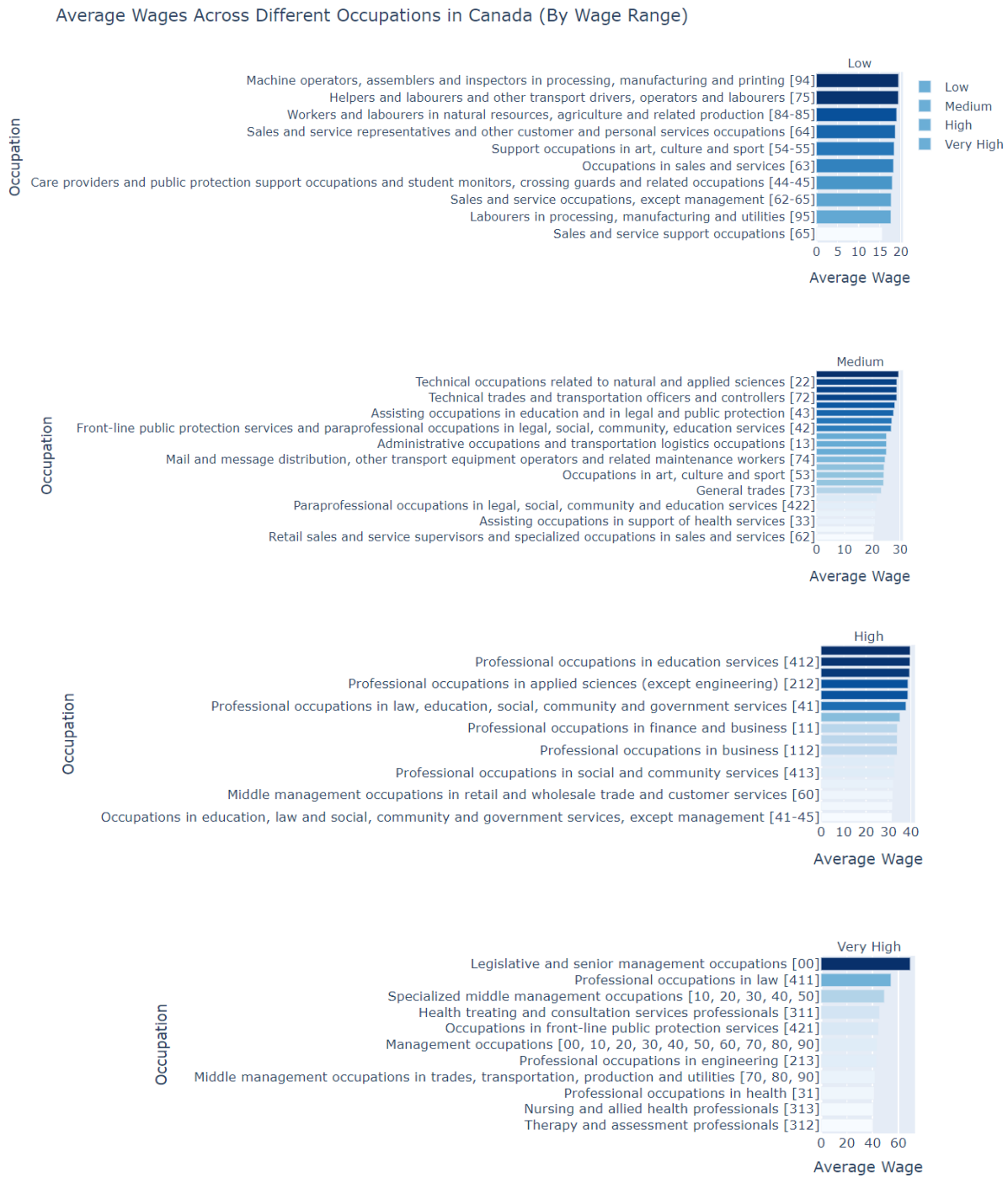


Fig 1.2


Average Wages Across Different Occupations in Canada (By Wage Range)
Low
Machine operators, assemblers and inspectors in processing, manufacturing and printing [94]
Helpers and labourers and other transport drivers, operators and labourers [75]
Workers and labourers in natural resources, agriculture and related production [84-85]
Sales and service representatives and other customer and personal services occupations [64]
Support occupations in art, culture and sport [54-55]
Occupations in sales and services [63]
Care providers and public protection support occupations and student monitors, crossing guards and related occupations [44-45]
Sales and service occupations, except management [62-65]
Labourers in processing, manufacturing and utilities [95]
Sales and service support occupations [65]
0 5 10 15 20
Average Wage
Occupation
Low
Medium
High
Very High
Medium
Technical occupations related to natural and applied sciences [22]
Technical trades and transportation officers and controllers [72]
Assisting occupations in education and in legal and public protection [43]
Front-line public protection services and paraprofessional occupations in legal, social, community, education services [42]
Administrative occupations and transportation logistics occupations [13]
Mail and message distribution, other transport equipment operators and related maintenance workers [74]
Occupations in art, culture and sport [53]
General trades [73]
Paraprofessional occupations in legal, social, community and education services [422]
Assisting occupations in support of health services [33]
Retail sales and service supervisors and specialized occupations in sales and services [62]
0 10 20 30
Average Wage
Occupation
High
Professional occupations in education services [412]
Professional occupations in applied sciences (except engineering) [212]
Professional occupations in law, education, social, community and government services [41]
Professional occupations in finance and business [11]
Professional occupations in business [112]
Professional occupations in social and community services [413]
Middle management occupations in retail and wholesale trade and customer services [60]
Occupations in education, law and social, community and government services, except management [41-45]
0 10 20 30 40
Average Wage
Occupation
Very High
Legislative and senior management occupations [00]
Professional occupations in law [411]
Specialized middle management occupations [10, 20, 30, 40, 50]
Health treating and consultation services professionals [311]
Occupations in front-line public protection services [421]
Management occupations [00, 10, 20, 30, 40, 50, 60, 70, 80, 90]
Professional occupations in engineering [213]
Middle management occupations in trades, transportation, production and utilities [70, 80, 90]
Professional occupations in health [31]
Nursing and allied health professionals [313]
Therapy and assessment professionals [312]
0 20 40 60
Average Wage

Research Question-2 Charts

Fig 2.1

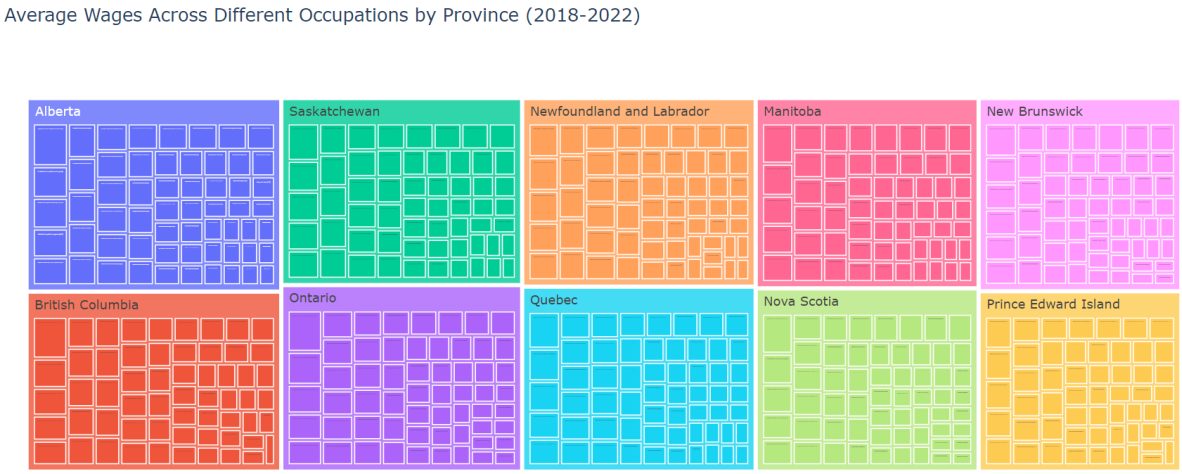


Fig 2.2

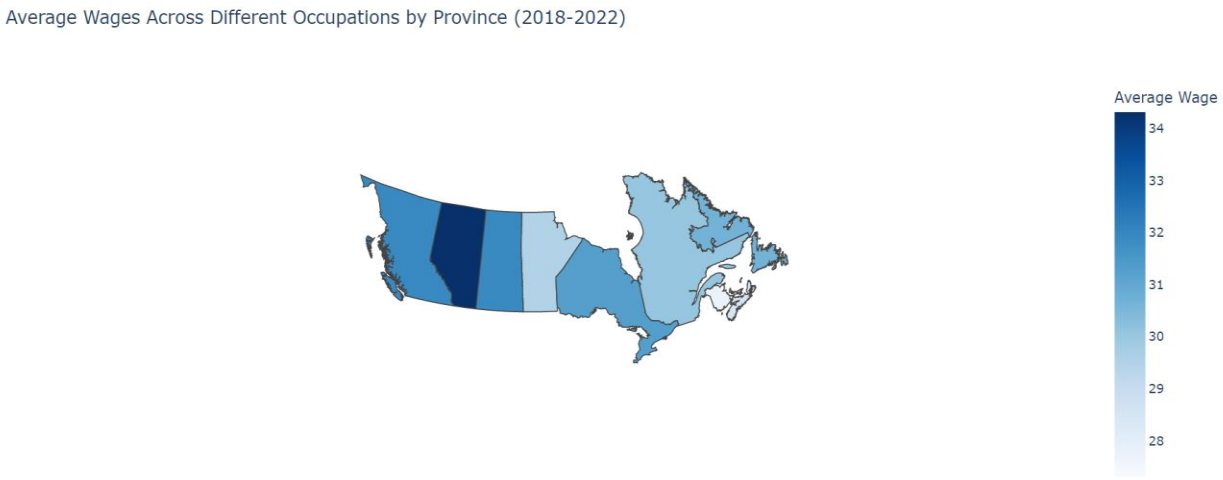


Fig 2.3

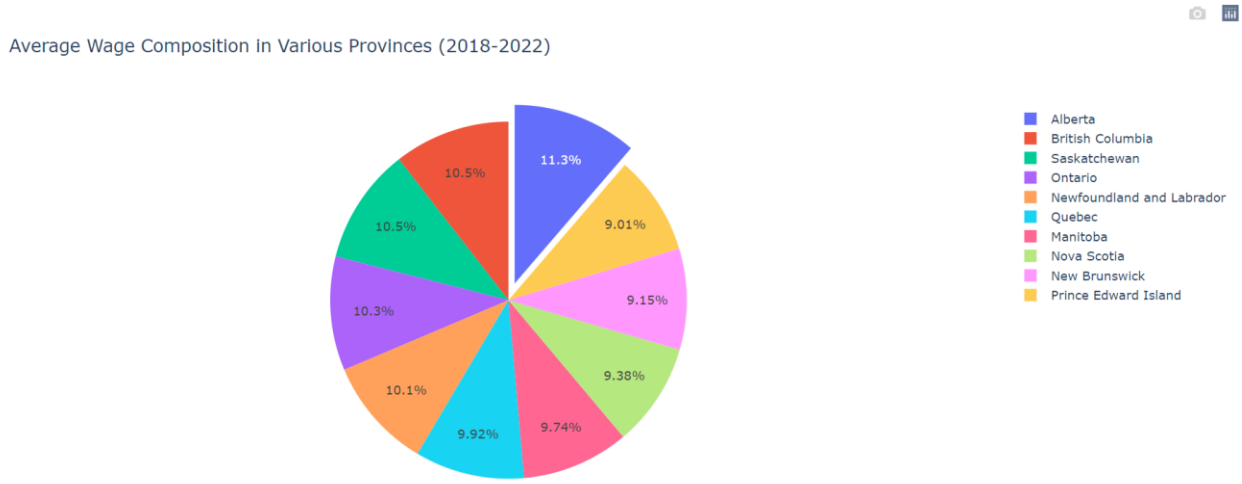
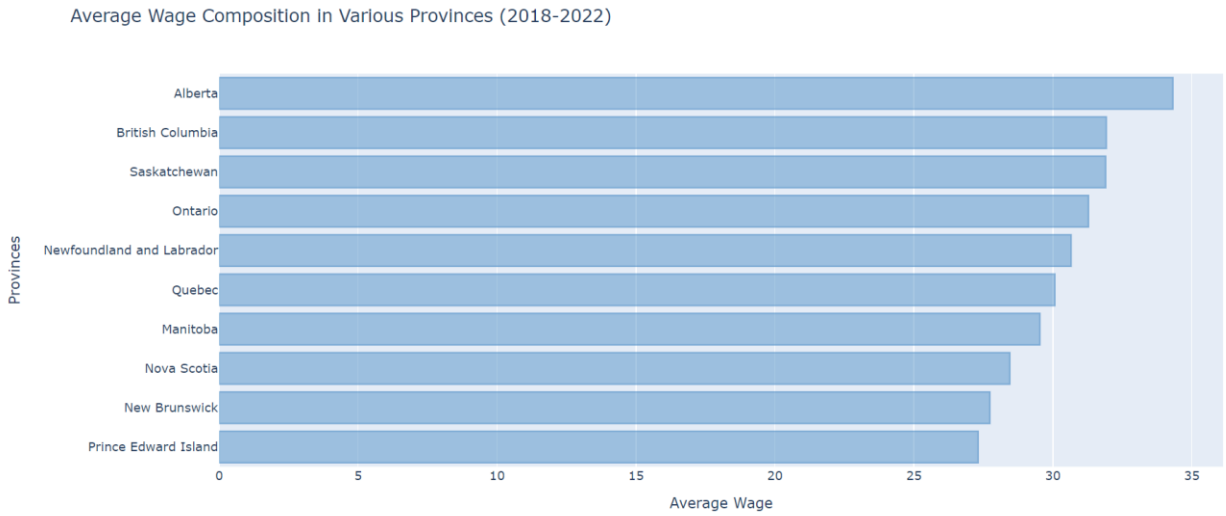
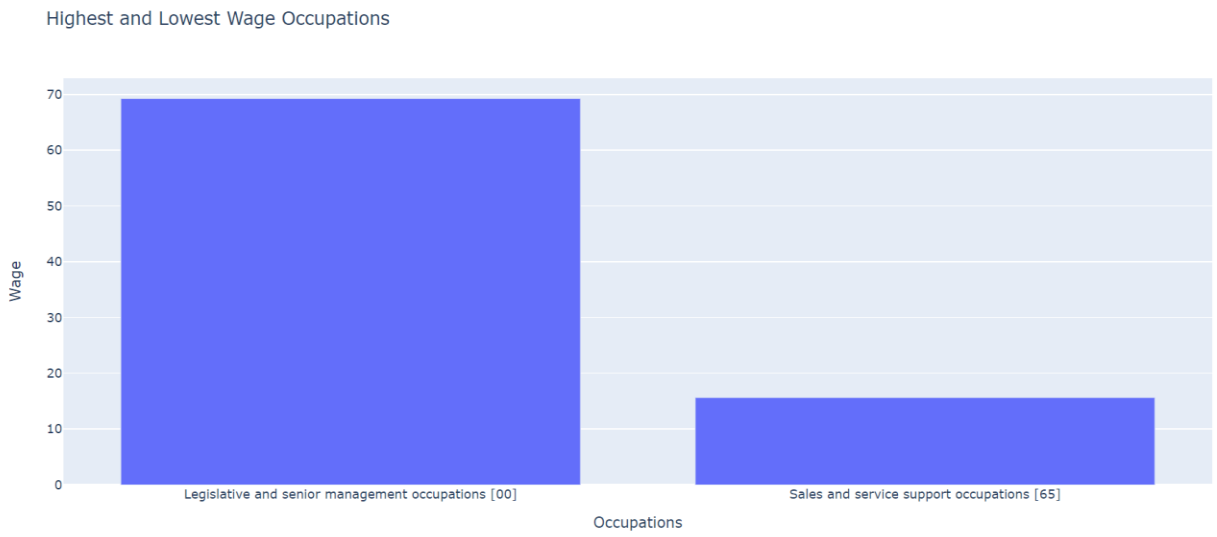

Average Wage Composition in Various Provinces (2018-2022)
11.3%
9.01%
9.15%
9.38%
9.74%
9.92%
10.1%
10.3%
10.5%
10.5%
Alberta
British Columbia
Saskatchewan
Ontario
Newfoundland and Labrador
Quebec
Manitoba
Nova Scotia
New Brunswick
Prince Edward Island

Fig 2.4


Average Wage Composition in Various Provinces (2018-2022)
Provinces
Alberta
British Columbia
Saskatchewan
Ontario
Newfoundland and Labrador
Quebec
Manitoba
Nova Scotia
New Brunswick
Prince Edward Island
0
5
10
15
20
25
30
35
Average Wage

Research Question-3 Charts

Fig 3.1



Research Question-4 Charts

Fig 4.1


Average Wages Across Different Occupations (2018-2022)
Therapy and assessment professionals [312]
Average Wage (CAD)
43
42
41
40
39
38
37
2018
2019
2020
2021
2022
Year

Research Question-5 Charts

Fig 5.1

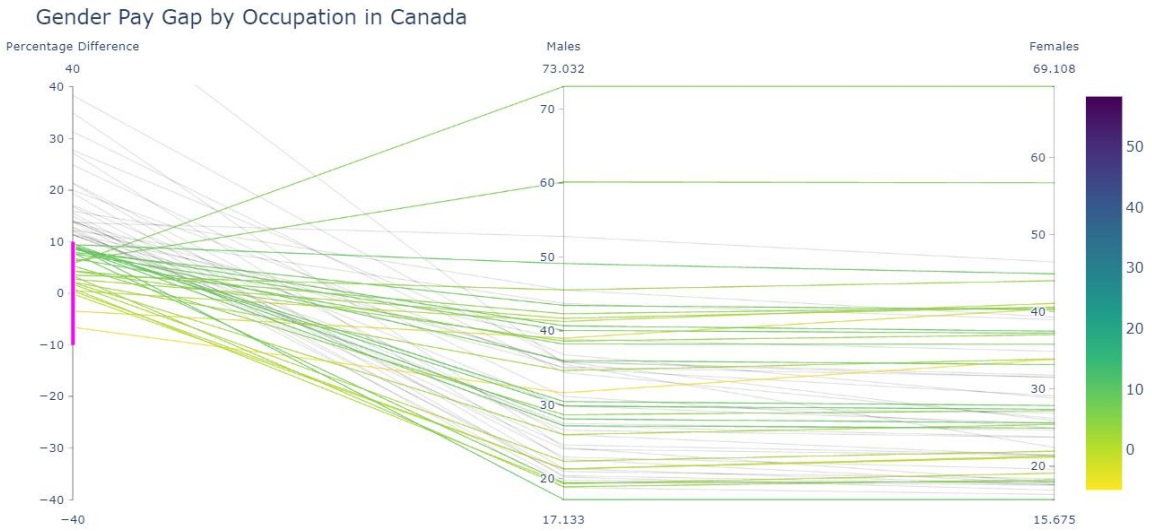
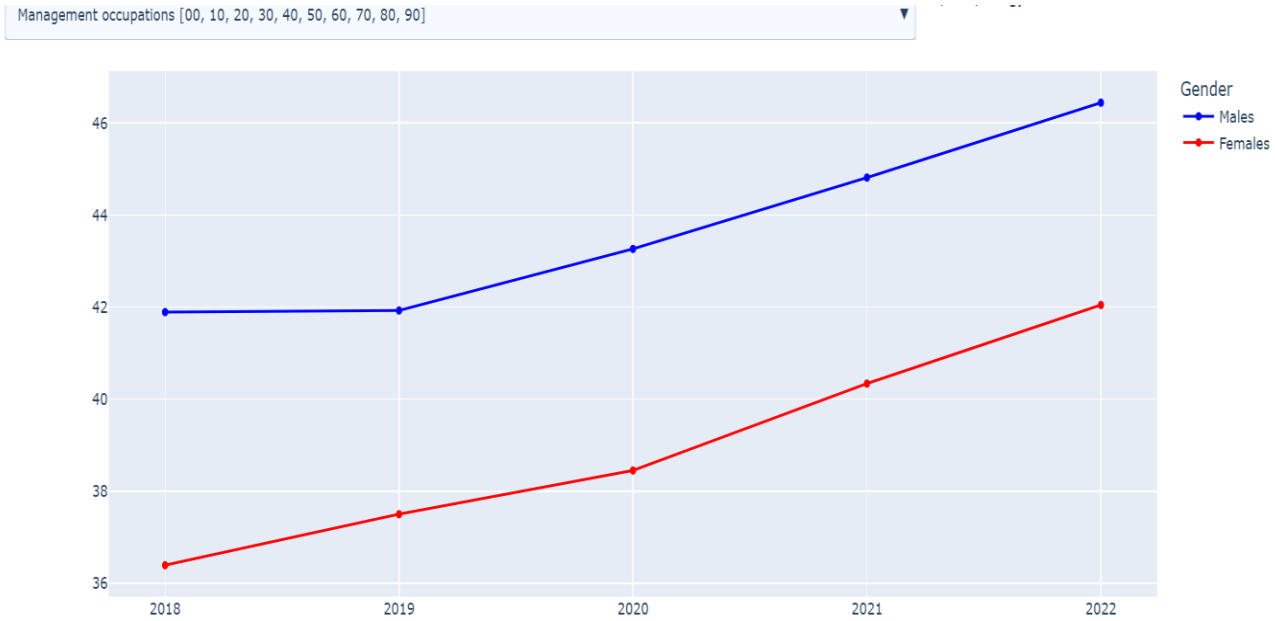


Fig 5.2



6. DISCUSSION

6.1 LEARNINGS & FUTURE WORK

Based on the size of the data, it is advisable to utilize Google Colab due to the availability of free GPU resources which can effectively enhance computation performance for large datasets

It is worth noting that the visualizations that I presented could be generated with the alternate charts while considering additional factors into consideration as well, such as age groups, skill sets, other types of jobs like part time category and experience level etc.

In future investigations, it is recommended to consider the previously stated factors to facilitate the identification of further insights.

6.2 CONCLUSION

The project successfully provides a transparent and accessible source of information on wages across different occupations in Canada.

It empowers workers and employers to make more informed decisions about career paths, salary negotiation, and job satisfaction.

It's important to note that the dataset only provides a snapshot of the current state of wages including provinces only and may not fully capture all possible factors affecting wages.

7. GITHUB REPOSITORY / CODE FILES LINK

The code for this project is available on GitHub.

7.1 MY REPOSITORY LINK:

https://github.com/BislaSonal/Wages-By-Occupation-Across-Canada

The repository contains a README file with detailed instructions on how to run the code and reproduce the results.