

Labour and Unemployment Rates in Canada:

A five-decade Analysis using Big Data Tools

Datta S, Fahme F, Ghodousipour F, Kaur J, Rodriguez M and Tang C., 2021 April

Summary

Objectives

A recession occurs when at least 2 consecutive quarters of negative economic growth occur. It is associated with increased unemployment rates. The exponential growth of data, the atypical pattern of the current recession, the availability of increased computing power, and the increased demand for responsive applications make this field a good application for big data analysis. This project aimed to utilize big data techniques in exploring and providing predictive analysis on unemployment rates in Canada. The research questions were:

1. What is the distribution profile of unemployment rates in Canada for the past 4 decades?
2. Can we predict the future trend in unemployment rates in Canada and Ontario?
3. How long does it take to recover from a recession?

Data Preparation

The dataset was obtained from [Statistics Canada](#). It contained the labour force frequencies in Canada from 1976 until 2020. There were 18 attributes and 931,393 observations. After exploratory analyses, targeted analyses were performed on Unemployment rates. Filtering using all industries yielded 35,640 values, of which 68 were blank. These were assumed to be zero, and values were replaced as such.

The platform used was Databricks. Languages used were python, pyspark and sql. The libraries utilized were request, urllib, json, pandas, numpy, matplotlib, seaborn, statsmodels, fbprophet, pyspark.ml and mllib. Predictive analysis techniques included Time series, FBProphet, Linear Regression, Random Forest Regression and Decision Trees.

Definitions from MetaData:

Employed: persons having a job or business

Unemployed: persons without work, are available for work, and are actively seeking work

Labour force = employed + unemployed

Unemployment rate = unemployed / labour force

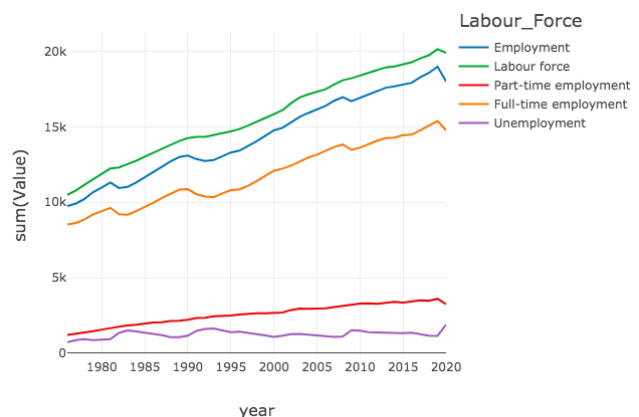


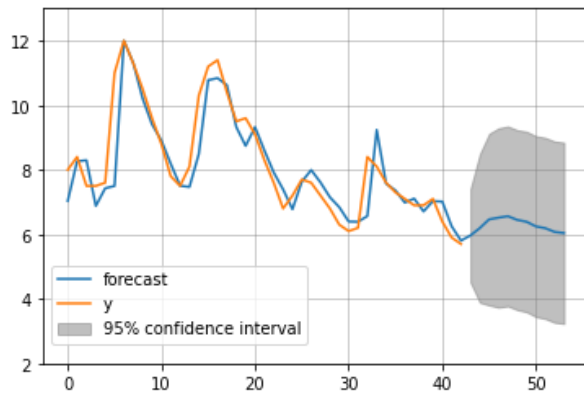
Figure 1. The labour force in Canada from 1976-2020

Analyses

The labour force had an increasing trend with the occasional slow down and recovery (Figure 1). Young adults had a higher proportion and a higher variability in unemployment rates.

Females had higher unemployment rates during 1970s-1980s, however, the proportion shifted

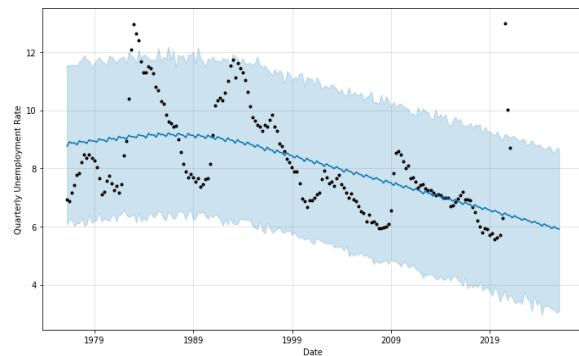
to higher unemployment rates among males from 1990s to present . The unemployment rates in Atlantic Canada were higher than other regions. The goods- and the services producing



industries were dominant. The unemployment rates of goods-producing industries were more volatile, but the service-producing industry showed a steeper increase in unemployment rates for the current recession.

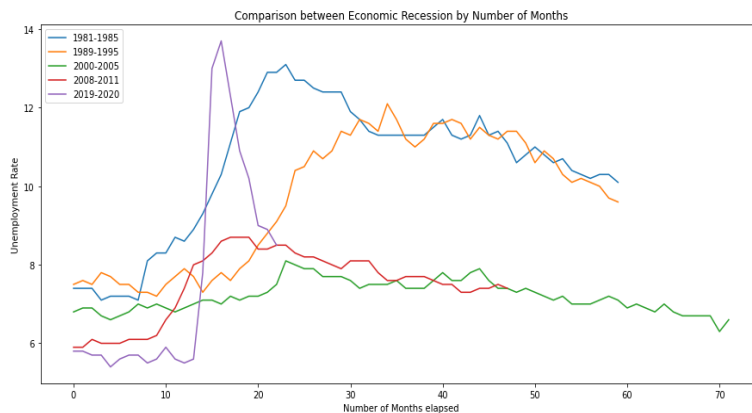
The Autoregressive Integrated Moving Average (ARIMA) model showed that unemployment rates could be a value in the range of 3-9% by 2031 (95% CI) (Figure 2).

Figure 2. ARIMA model



The FBProphet library showed the predicted unemployment rate to be approximately 3-9% by 2030 (Figure 3).

Figure 3. Prediction using FBProphet



The linear regression prediction for the 2021 unemployment rates were 8.8% for Canada and 8.7% for Ontario. The success rates for prediction by the random forest regression was 45%, and 62% for the decision tree model. The average duration of Canadian recessions is 6.3 years (Figure 4).

Figure 4. Unemployment rates across time

Conclusion

The labour force had a general increasing trend, with decreasing

unemployment rates. Unemployment rates were higher among young adults, males, and populations living in Atlantic Canada. The unemployment rates fluctuations were prominent in the top two sectors of goods-producing and service-producing industries. The unemployment rate is predicted to be 8.8% for 2021 and between 3-9% by 2031.

Increasing unemployment rates correspond to recessions, which occur approximately every 10 years. On average, recovery from a recession was achieved within 6 years. However, the extent and direction of the current recession is hard to predict because of its atypical pattern.