PHY408 Final Report: Employment Time Series Analysis

Mark Buchanan: 1003520745

April 17th, 2020

1 Introduction

My project is analyzing historical Canadian employment data sourced from Statistics Canada using some of the time series analysis techniques discussed throughout the course (Statistics Canada). The dataset includes national employment estimates (in number of people employed) for Canadians aged 15+. For my analysis, I have broken down the data into male and female estimates. The purpose of this project is to understand the significance of any seasonal and annual trends in employment in Canada. Trends longer than this are largely a result of population growth, and trends on the order of weeks or days are essentially noise.

My original plan was to look at unemployment data, but Statistics Canada only offers that data on a rolling 5 month basis. However, the employment data is available all the way back until 1976 from the present, so I decided to use that instead. As my analysis is on the periodic trends of the employment data, there is no practical difference between the employment and unemployment data other than them being shifted along the time axis, which does not change the underlying frequencies of oscillation.

The motivation for this project was based on the soaring unemployment numbers resulting from the COVID-19 pandemic. I initially thought it would be interesting to see the effect that this recent data had on the analysis of the historical data. Once I began analyzing the dataset including the March 2020 drop in employment, I realized that its inclusion resulted in drastically changing the analysis as to essentially mystify the consistent behaviour in employment. As a result, I decided that truncating the data before the COVID-19 pandemic had reached significance allowed for more meaningful and illuminating results to be found from the data. This allowed for annual and seasonal trends to be much more easily isolated and looked into further.

The complete raw historical data from the beginning of the data collection period in January 1976 to March 2020 is shown in Figure 1a. The drastic effect of COVID-19 can be seen in the last data point for males and females, and is more significant than any of the other drops in employment over this recorded period. It is this drastic change which skews the analysis which is why I have truncated it. One of the periods analyzed includes the abrupt drop in employment in the early 1980s which shows a similar type of distortion.

2 Analysis & Results

I decided to analyse the data in chunks of 4 years starting every dozen years starting from 1980 (ie. January 1980 to December 1983, January 1992 to December 1995, etc.). The 4 year duration of time was decided upon by trying to balance the resolution of the signal in frequency space, while also keeping the length of time short enough so that a linear detrend on the employment data would be sufficient for analysis. Since the data is provided in one month intervals, a number of years of signal is beneficial in increasing the resolution in frequency space (4 years time series gives a resolution of 0.25 cycles per month). In order to prepare the data for analysis, I removed the linear trend found in the data over the batches of 4 year intervals. This trend is a result of the population growth of Canada, which hovers around 1% depending on the year (Statistics Canada 2018). This can be approximated quite well as a linear trend over small durations such as 4 years, which is what I have assumed. The detrended data for the 2016-2020 period can be seen in Figure 1b.

After removing the trend in the data, I took the Fourier Transform of the four year period and set all the amplitudes for the frequencies greater than 2 cycles per month to 0 (maintaining symmetry requirements). The purpose of this was to remove the higher frequency oscillations in the data to smooth out the data to see the dominant seasonal and annual behaviours more clearly. We can see that the dominant frequencies in the data are those with $|f| \leq 2$ cycles/month in Figure 1c.

After frequency filtering, I took the Inverse Fourier Transform of the frequency filtered and detrended data to see the desired behaviour. This behaviour can be seen for 2016-2020 in Figure 1d, and is representative of the results from the other four-year samples. We can see that the behaviour is quite consistent year to year, although the behaviour is different for men and women. These results match other analyses, showing that males experience a much higher rate of seasonal employment than women, due to men being overrepresented in industries that experience high seasonal employment (Sharpe, 5). The so-called primary industries, including fishing, construction, agriculture, forestry, mining, etc. are made up of only 20% women, and make up the industries with the largest amount of seasonal employment (Sharpe, 5). This contributes to men having larger swings in employment, which is almost entirely on an annual cycle. They reach the peak of their employment around July, and the trough around February.

Women on the other hand have somewhat more sustained employment numbers with less significant swings in employment, but they experience two different oscillations per year. Their larger peak is in June and have a secondary peak in November, with the more significant trough occurring in February and a secondary trough in September. The main peak and trough for the women is roughly aligned with the men and is caused by participation in the same industries, albeit to a lesser extent, that causes the men's annual oscillation. The secondary oscillation is more interesting because it illuminates more complex behaviour.

However, this behaviour of the secondary oscillation in the women's employment data has not always existed. The plots of the filtered four-year periods have been overlaid on top of each other for both male and female values can be seen in Figures 1e and 1f respectively. Based on inspection, it seems this secondary oscillation reaches prominence somewhere around 1993 and grows in significance up until current times. I was not able to find a paper that discussed this trend, which is partially explained by most of the analyses I found on this topic being published in the mid 1990s to mid 2000s. I thought it may have been a mistake, but the secondary variation is visible in the raw data as well. I do not know what could be causing such a significant change on the national scale that is exclusive to women.

Also, it is worth acknowledging the anomaly in the 1980-1984 data, which can be described by the significant abrupt drop in employment during this period which is visible in Figure 1a. Even through this drastic change, the overall seasonal and annual trends in the data is quite consistent with the other periods analyzed.

The overlaid frequency filtered curves for the various periods in Figures 1e and 1f are helpful, but they make it difficult to compare each year with every other year. For example the first year of the 1992-1996 data can and should be compared with the third year of the 2016-2020 data, not just all the other first years of the periods. To address this, I also calculated cross (and auto) correlations for each period with itself and the other periods. These correlations were done on the detrended frequency filtered data shown in Figures 1e and 1f and are again split into male and female. The correlations for each period with all the periods are roughly the same, so I have only included the 2016-2020 period cross/auto correlated with the others. This can be seen for males and females in Figures 2a and 2b respectively.

The plot for the men, Figure 2a, shows what we would expect based on the previous plots. We have that each of the datasets is strongly correlated at integer year offsets and anti-correlated in between these. This means that the employment changes consistently follow an annual cycle throughout the analyzed periods. The damped oscillation type behaviour of the correlations is due to the finite sampling interval, where more of the signals get cut off from each other as the offset deviates further from a zero offset.

Switching to the plot of the women, Figure 2b, we do not have such clean correlations. This is also expected based on previous plots, as we noted the change from women having one dominant annual oscillation in the past, to having two significant oscillations per year starting at some point in the mid-1990s. Similar to the men, the women show strong correlations for an integer number of years offset which suggests there is an annual cycle. However, it also shows that for half integer year offsets there is a clear trend of increasing correlation as time progresses. The datasets were initially anti-correlated at these offsets when cross correlating 2016-2020 with 1980-1984, but after a dozen years, the anti-correlation gets weaker until the auto-correlation of the 2016-2020 data shows a small correlation. This suggests a continued development of a secondary oscillation in women's annual employment cycle, which is offset from the main oscillation by about 6 months.

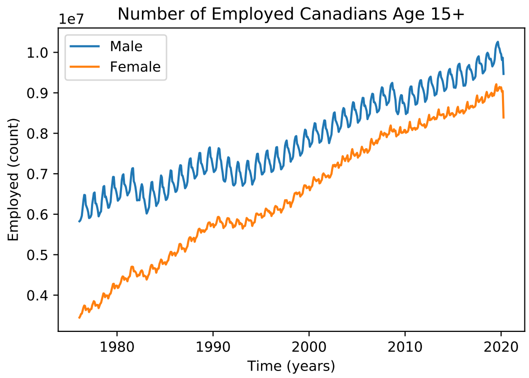
A report from Statistics Canada found that from 1976 to 1997, men and women a 16% and 26% drop in seasonal employments respectively (Marshall, 20). Assuming this trend continued, it is difficult to tell whether it is visible in Figures 1e and 1f. The men's data seems to follow this drop, but due to less uniformity in the women's data, it is difficult to see if this is followed.

3 Discussion

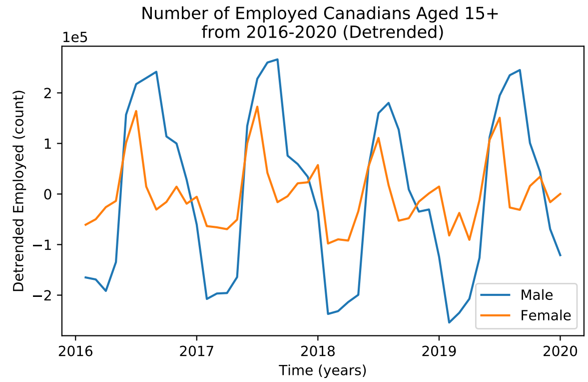
The analysis done here is quite coarse and could be done more rigorously through more thorough pre-processing of the data (ie. analyse periods of more than 4 years at a time with better detrending of the data over a simple linear best fit removal). Another aspect that would have been beneficial would be to break down the data more than just into male and female. Employment seasonal changes depends largely on age and also area; people in the 15-24 age groups having much higher seasonal employment than average due to largely being students and Canada as a whole having many regions of varied employment trends (Sharpe, 3-4). It would be useful to do more thorough analysis including all these varied factors. In my research, I found that the Atlantic provinces have a much higher rate of seasonal employment changes due to many of their jobs coming from industries that rely on specific times of the year such as agriculture and fishing (Sharpe, 8). It would also be interesting to compare the results of different filtering techniques in addition to the direct frequency filtering utilized, such as Z-transform filters. Taking all of this into consideration, I think my rudimentary analysis was still illuminating since it seems to match up nicely with published results from Statistics Canada among others.

This type of analysis is useful for governments, both federal and provincial. For example, they can use this type of information to understand when to expect more claims for unemployment and other social assistance programs. This could also be useful in economic planning, where governments could monitor how well policy changes are working based on the change from the standard expected employment variations throughout the year.

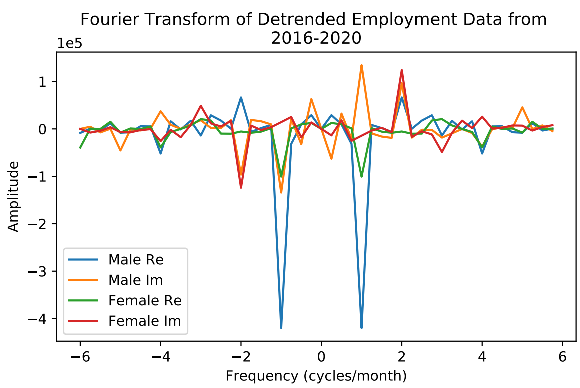
4 Figures



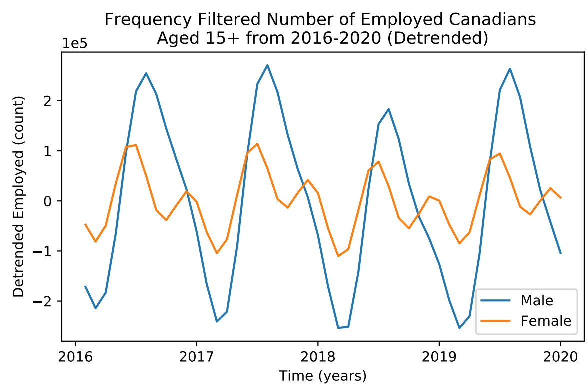
(a) Raw historical Canadian employment data from January 1976 to March 2020



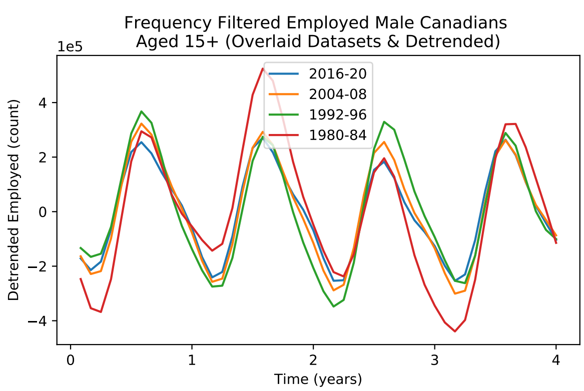
(b) Linearly detrended employment data from January 2016 - December 2019



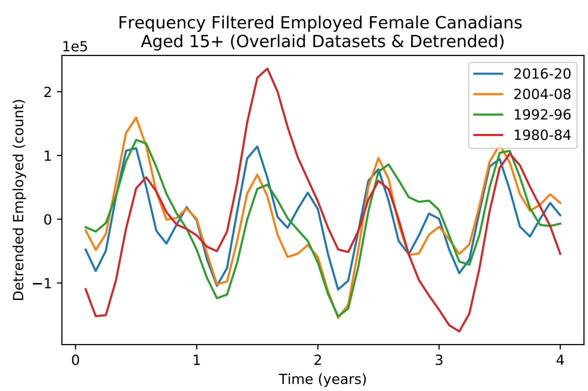
(c) Signal's Fourier Transform before frequency filtering



(d) Signal after frequency filtering out signals with frequencies > 2 cycles/month

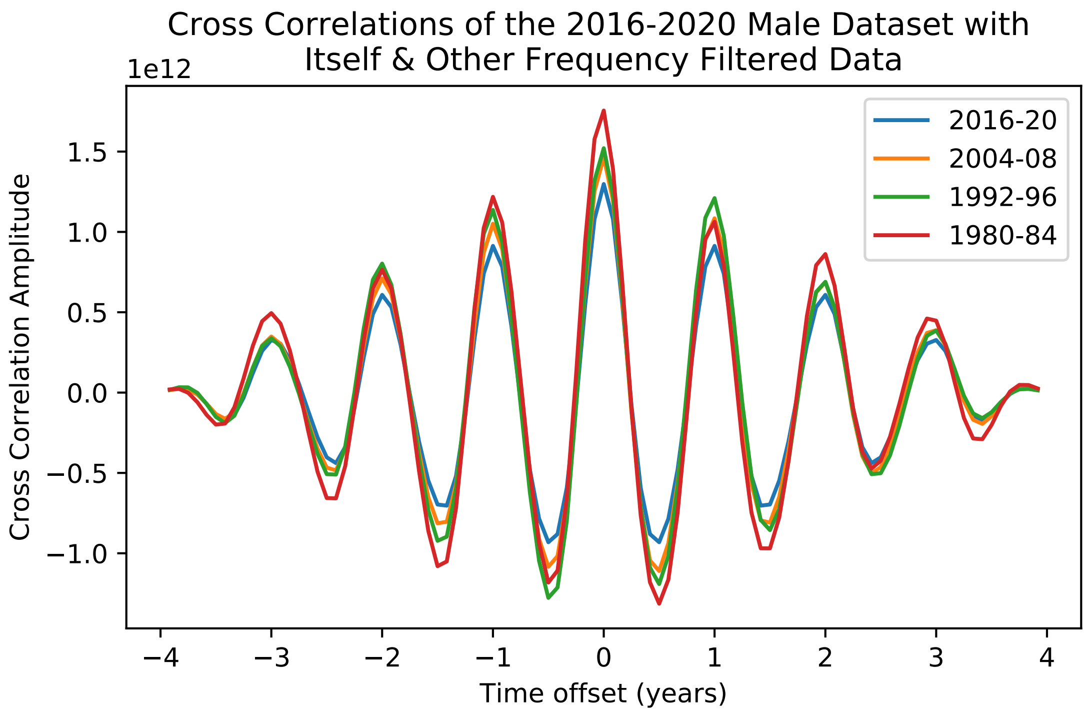


(e) All of the processed male datasets overlaid on each other

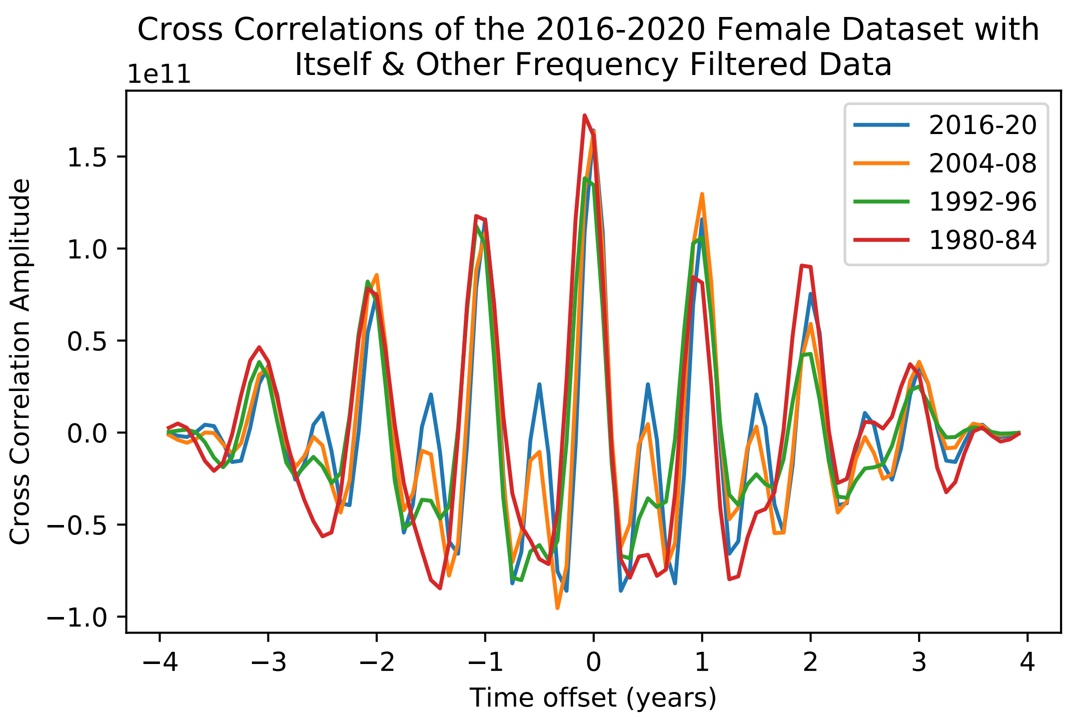


(f) All of the processed female datasets overlaid on each other

Figure 1: Beginning with the raw data and progressing through detrending and frequency filtering the data, broken down by sex.



(a) 2016-2020 auto and cross correlated with other periods, detrended and frequency filtered male data



(b) 2016-2020 auto and cross correlated with other periods, detrended and frequency filtered female data

Figure 2: Auto & cross correlations performed using numpy's correlate with "full" option.

5 References

Statistics Canada. Table 14-10-0017-01 Labour force characteristics by sex and detailed age group, monthly, unadjusted for seasonality (x 1,000)
https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1410001701
DOI: https://doi.org/10.25318/1410001701-eng

Statistics Canada. "Annual Demographic Estimates: Canada, Provinces and Territories, 2018 (Total Population Only) Analysis: Total Population." Analysis: Total Population, 27 Sept. 2018, www150.statcan.gc.ca/n1/pub/91-215-x/2018001/sec1-eng.htm.

Sharpe, Andrew, and Jeremy Smith. "Labour Market Seasonality in Canada: Trends and Policy Implications." Centre for the Study of Living Standards, Feb. 2005, pp. 1–19., http://www.csls.ca/reports/csls2005-01.pdf.

Marshall, Katherine. "Seasonality in Employment." Statistics Canada, 1999, www150.statcan.gc.ca/n1/en/pub/75-001-x/1999001/4408-eng.pdf.