A 40-Year Analysis of Canada’s Labor Market

Gabrielle Friedman, Yang Gao, Ivan
Misic, Lydia Tesconi, Mike Wang

gfriedman@falcon.Bentley.edu, yanggao@falcon.Bentley.edu, imisic@falcon.bentley.edu,
mwang@falcon.Bentley.edu, ltesconi@falcon.Bentley.edu

MA 611, Fall 2024

Canadian Employment Rate

Literature Review

Canadian Labour Market and Skills Researcher Network (2015)

- Various analyses of economic shifts & demographic trends in employment over time

Statistics Canada (2020)

- Short-term and long-term analyses of employment
- Insights into Covid-19's impact on Canadian economy

"Canadian City Unemployment Rates" (1997)

- A study highlighting disparities between Canadian provinces

"Seasonality of Labour Markets" (2000)

- Employs data from 1976 to 1997 on employment seasonality in Canada
- Compares seasonality in Canada to the United States

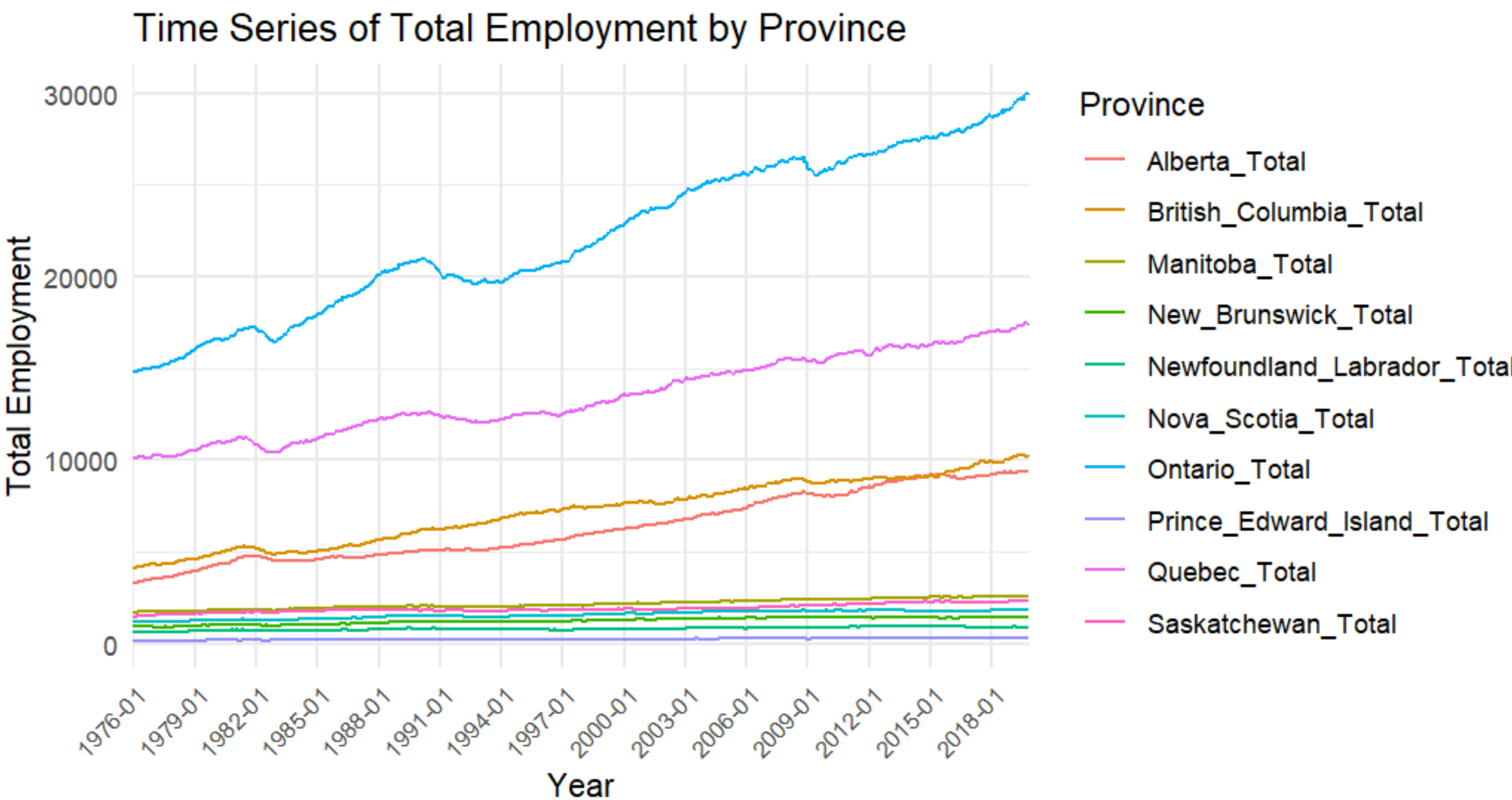
Roadmap

1. Exploratory Data Analysis
2. Materials and Methods
3. Data Analysis
4. Model Findings
5. Future Research

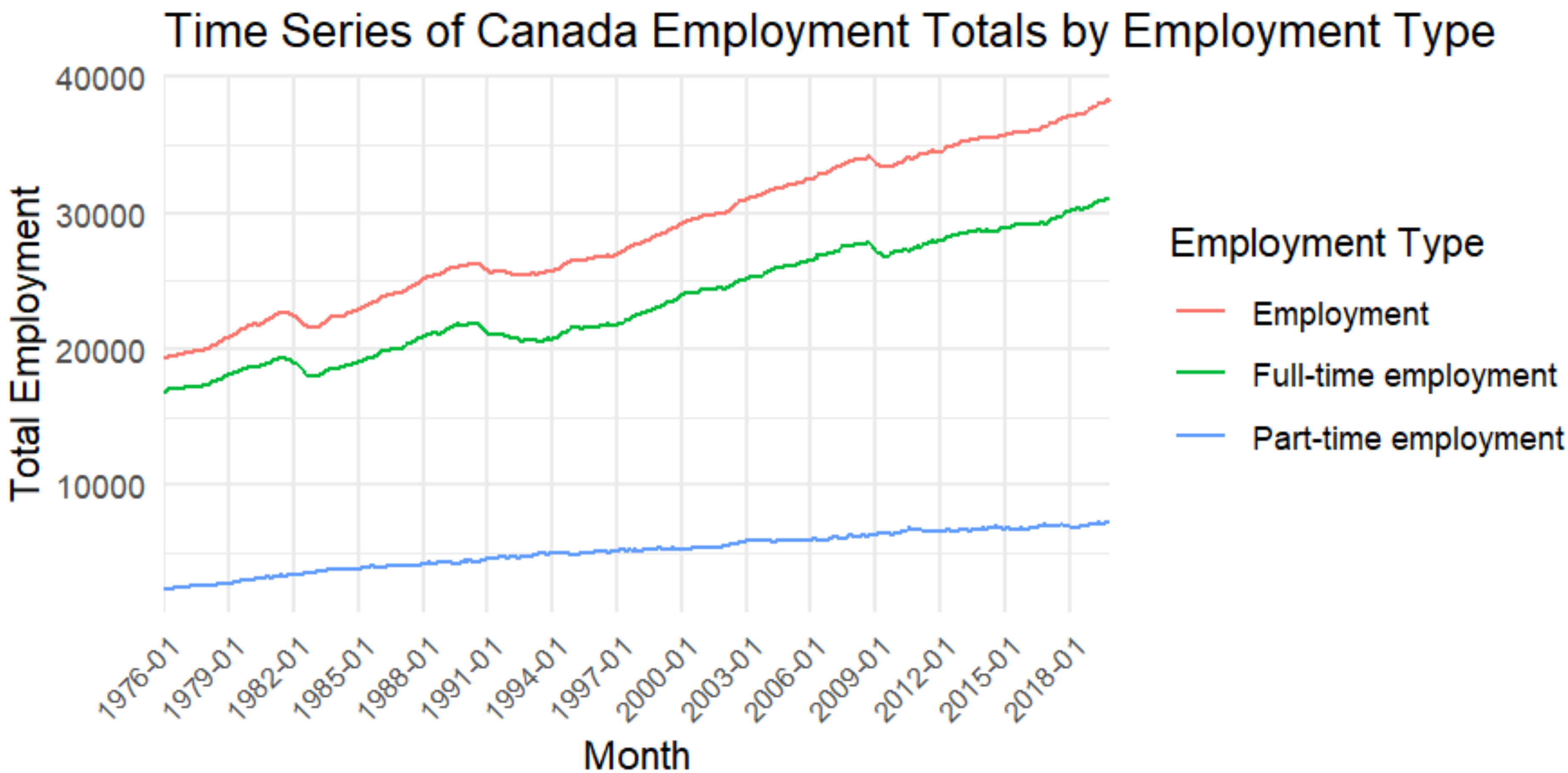
Exploratory Data Analysis

- Dataset: Employment Rate in Canada
- 4,743 observations with 13 variables
- Date range: 1976 to 2019
- No missing values
- Outliers: 17 months between July 2018 and Nov 2019
- Source: Kaggle

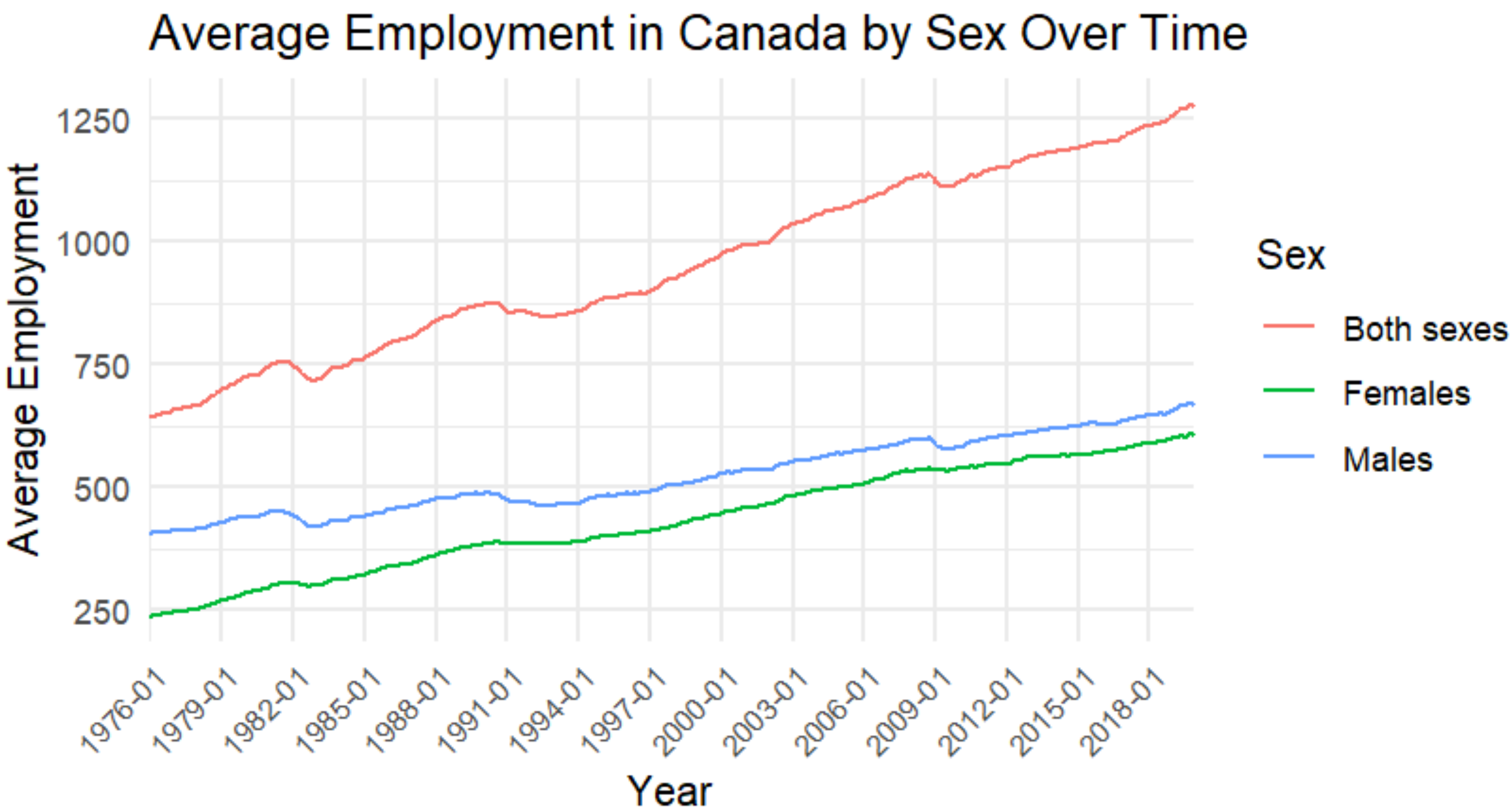
Exploratory Data Analysis



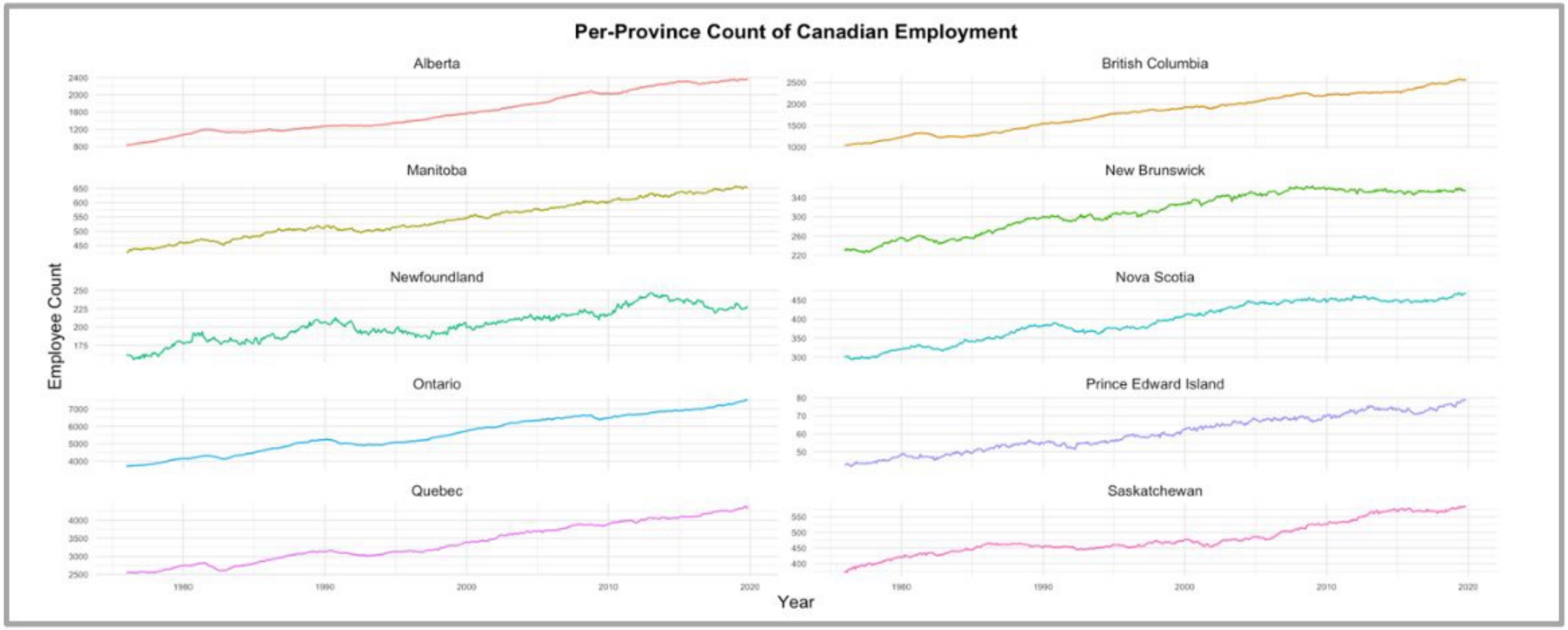
Exploratory Data Analysis

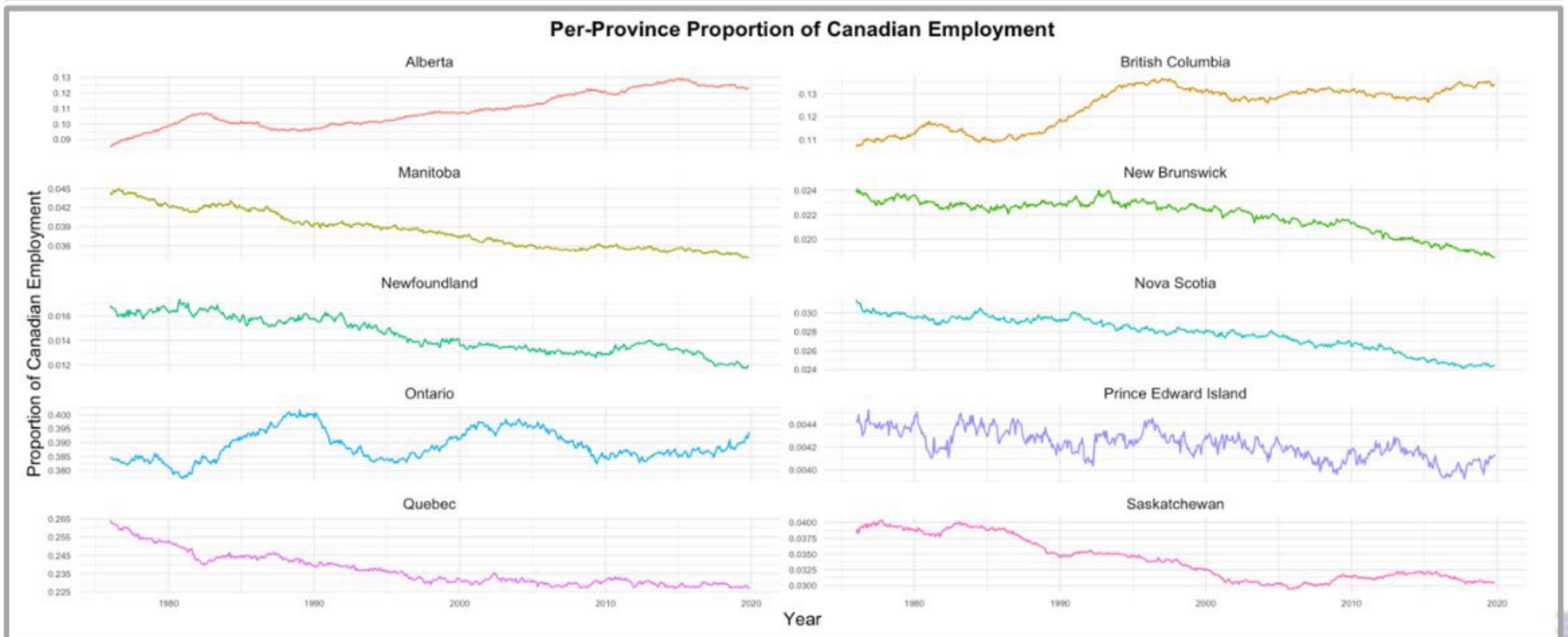


Exploratory Data Analysis

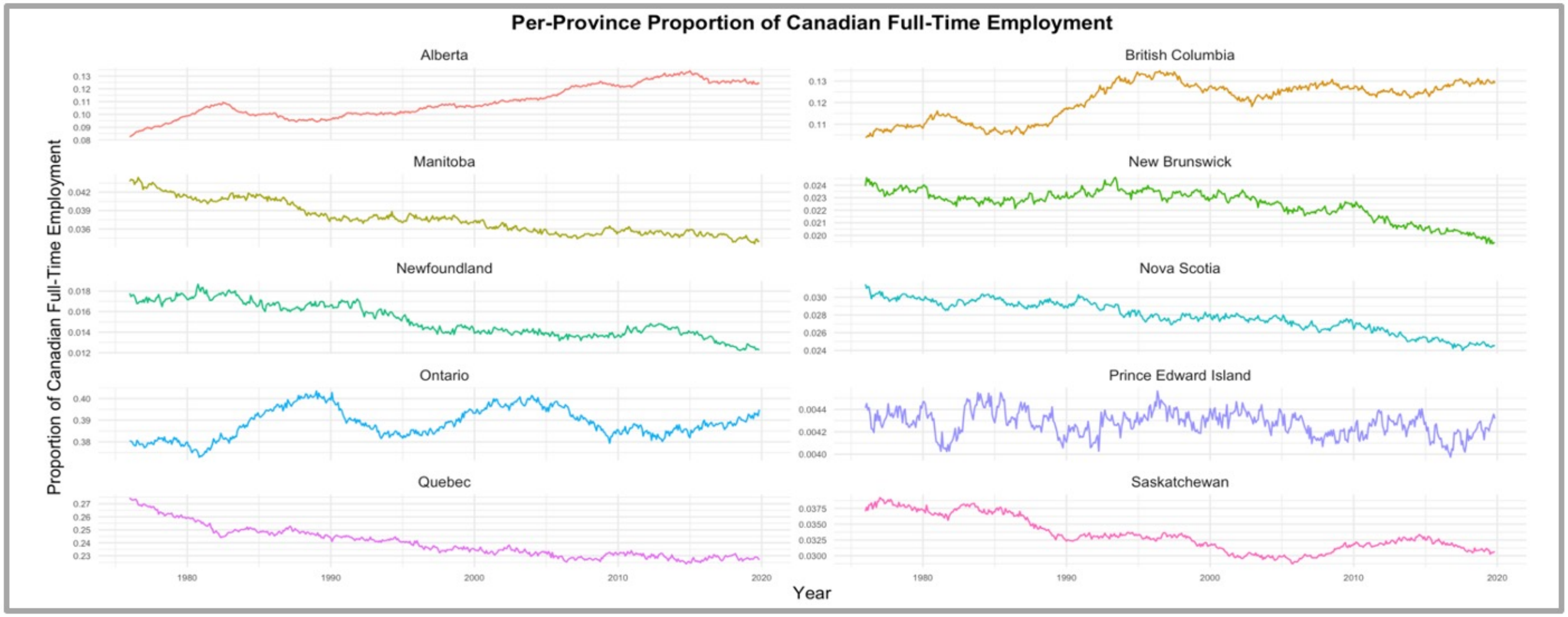


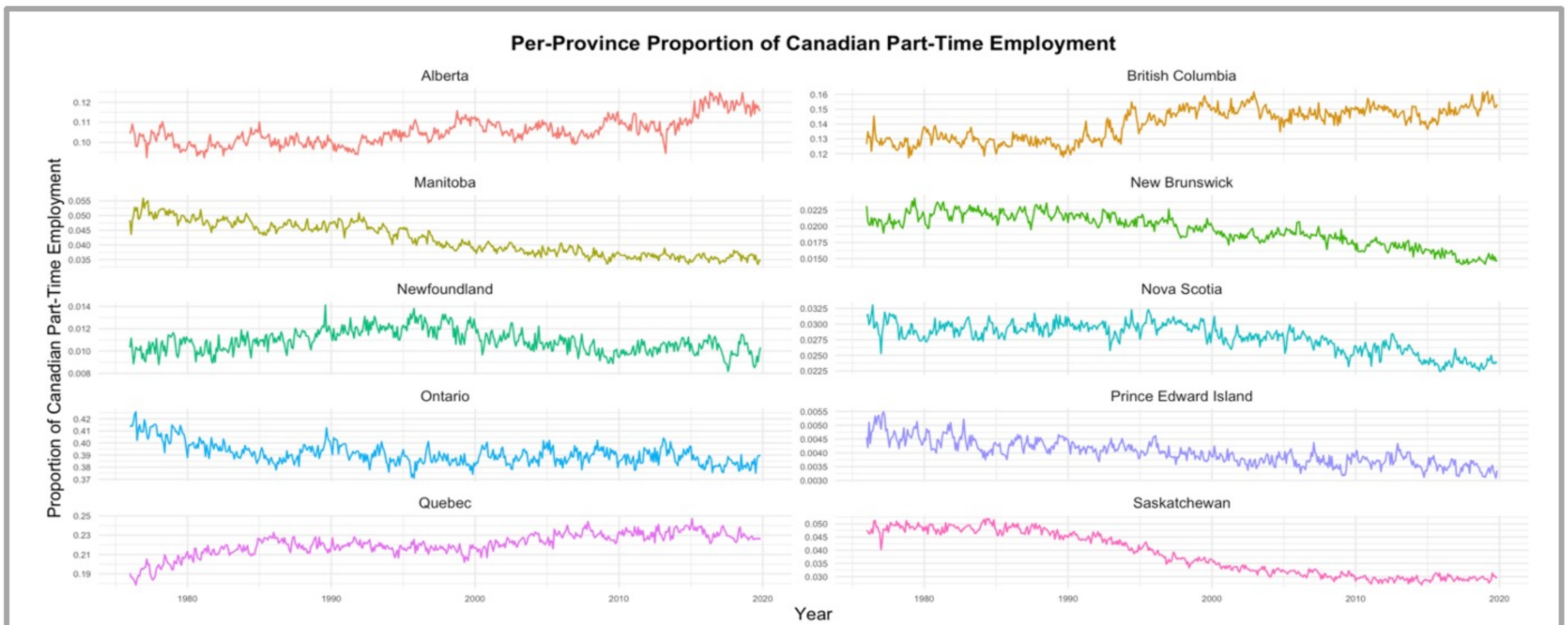
EDA – Counts & Proportions



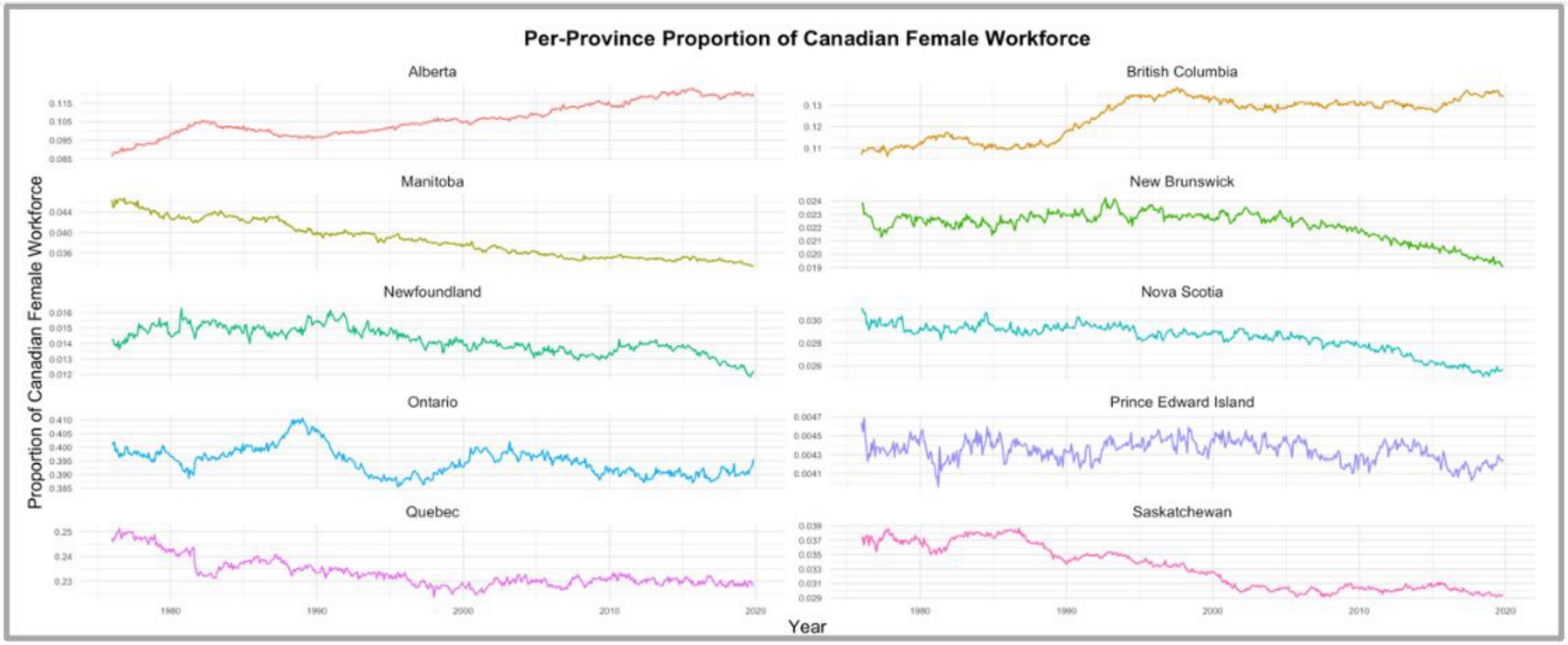


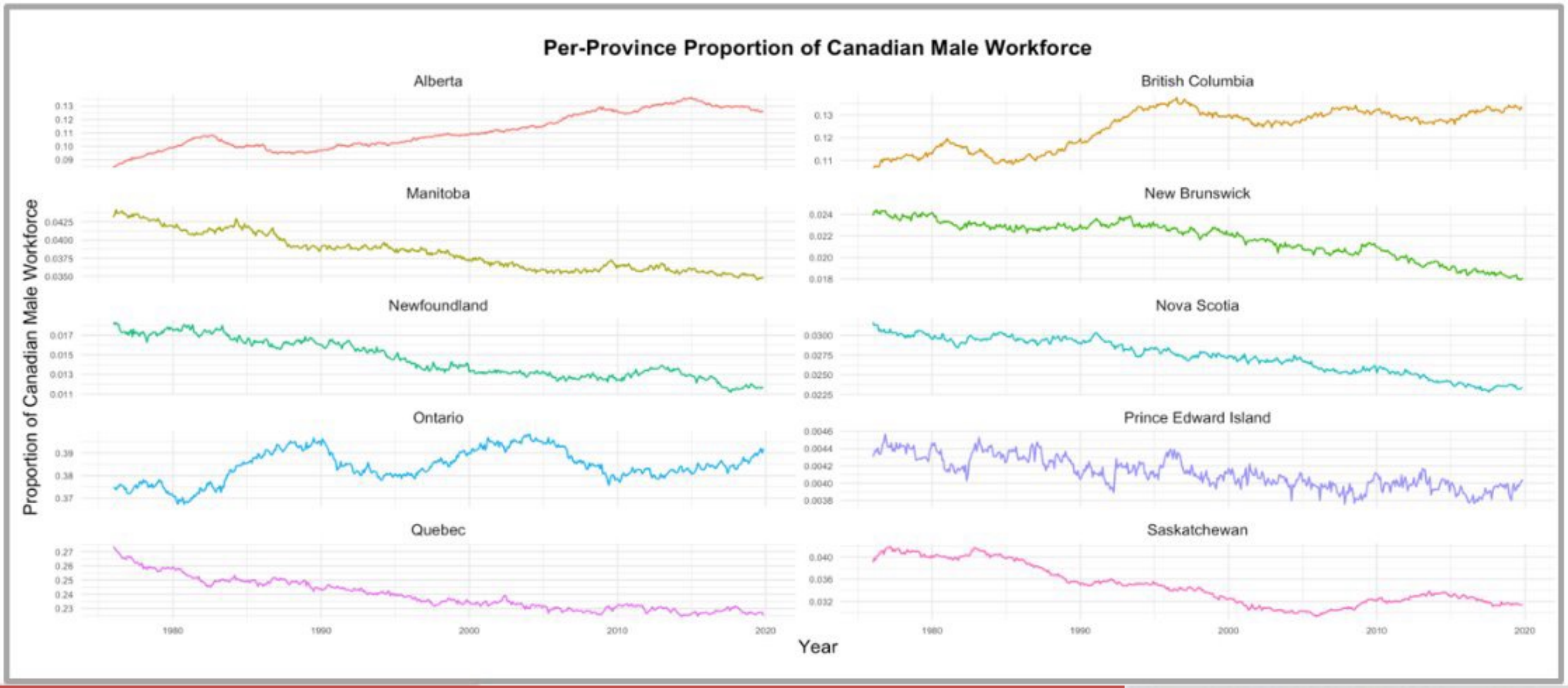
EDA – Proportions by Employment Type



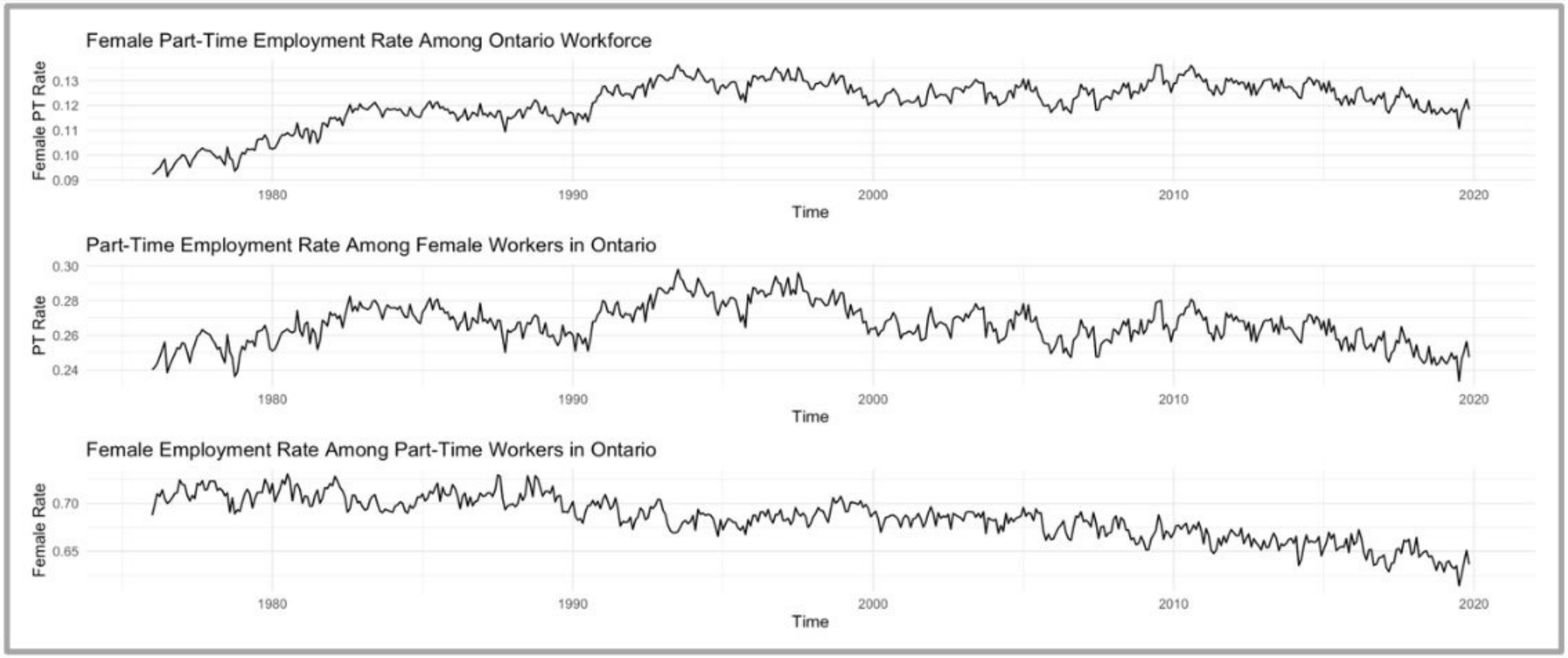


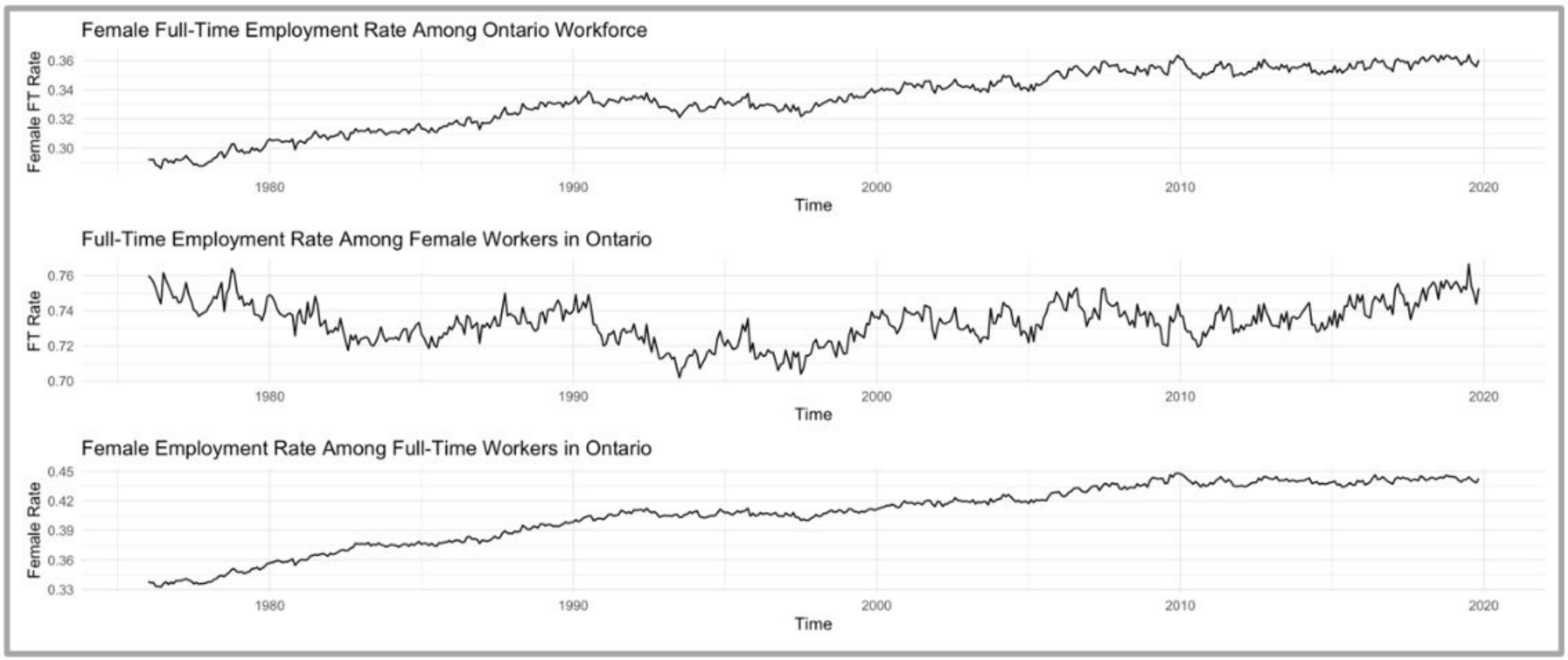
EDA – Proportions by Sex





EDA – Granular Proportions





Data Cleaning & Set-Up

1. Aggregation & Transformation
   - Canadian Totals Dataset
   - Provinces (10) Dataset
   - Employment Type & Sexes Dataset
   - Employment Rate
2. Training and Testing Splits
   - 80 – 20 split
   - Last 8 years held for testing set
3. Nonlinearity and Residual Checks

Fitted Models

Fitted Model 1: Canadian Totals Time Series

- ARIMA(2, 1, 1)(0, 0, 2)[12] with drift

Fitted Model 2: Province Time Series

- Province Time Series (multiple models)

Fitted Model 3: Sex Time Series

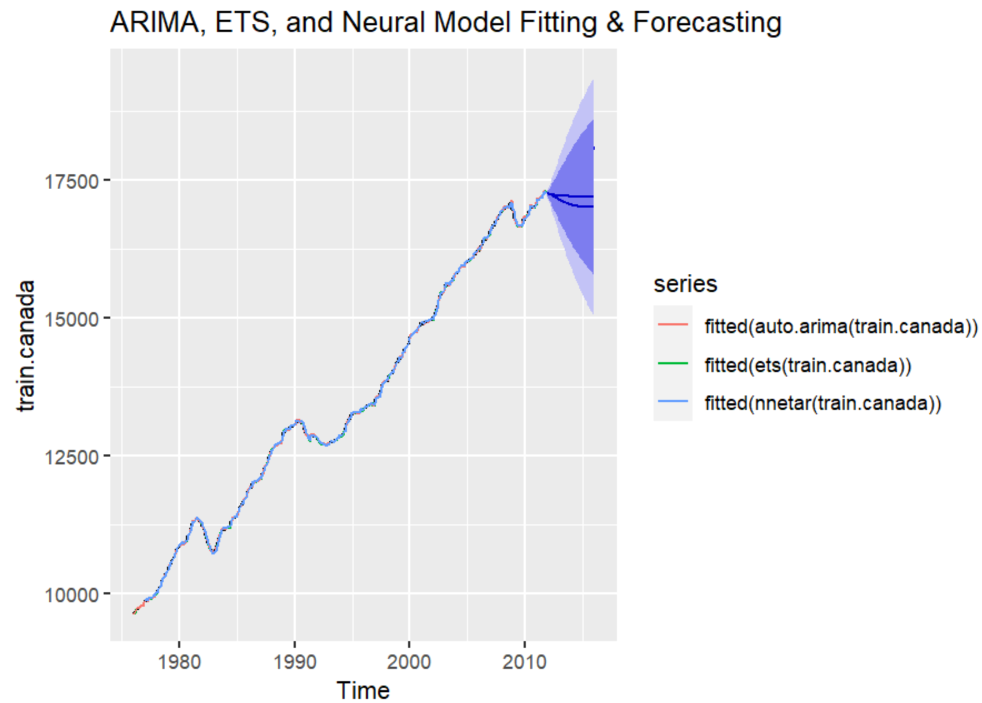
- ARIMA(1, 1, 2)(0, 0, 2)[12] with drift, both sexes

Fitted model 4: Granular (Rate) Time Series

- Bagged Model with 15 Bootstrapped Time Series

Model 1 Findings

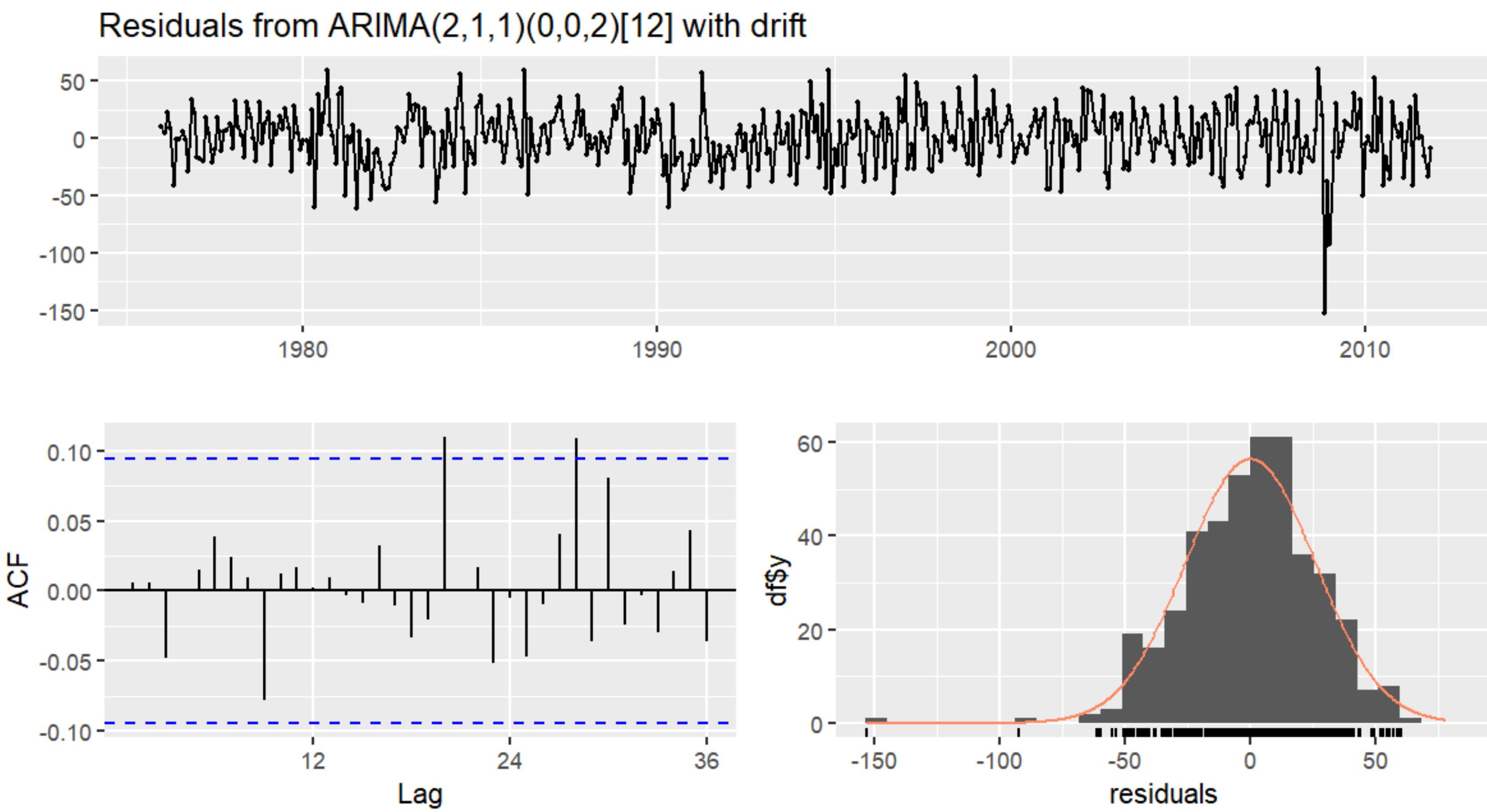
- ETS, ARIMA, and Neural Models fitted to Canadian Totals Time Series
- ARIMA model best performance on test set
- Some concerns with nonlinearity tests




```
> accuracy(forecast(canada_ets),canada_ts)
                     ME      RMSE      MAE        MPE      MAPE       MASE        ACF1 Theil's U
Training set   3.215036  27.24804 21.36251 0.02515652 0.1604159 0.07721579 -0.01917318        NA
Test set     320.756314 360.04233 320.75631 1.82013149 1.8201315 1.15938851  0.87438332   9.65013
> accuracy(forecast(canada_arima),canada_ts)
                      ME    RMSE       MAE          MPE      MAPE       MASE        ACF1 Theil's U
Training set  -0.005107199  25.982  20.29626 0.0002223291 0.1529079 0.07336176 0.006343202        NA
Test set     123.214084968 134.562 125.40070 0.7002276183 0.7128936 0.45326661 0.730253656   3.61694
> accuracy(forecast(canada_neural),canada_ts)
                      ME      RMSE       MAE           MPE      MAPE       MASE      ACF1 Theil's U
Training set  -0.005297073  28.82198  22.60697 -0.0007472166 0.1709067 0.08171393 0.1987748        NA
Test set     402.290368092 457.76038 402.29037  2.2821741223 2.2821741 1.45409711 0.8830664  12.26512
```

Model 1 Residual Check

```
> checkresiduals(canada_arima)

        Ljung-Box test

data:  Residuals from ARIMA(2,1,1)(0,0,2)[12] with drift
Q* = 13.068, df = 19, p-value = 0.8351

Model df: 5.   Total lags used: 24
```



Model 2 - Findings with Modeling on Each Province

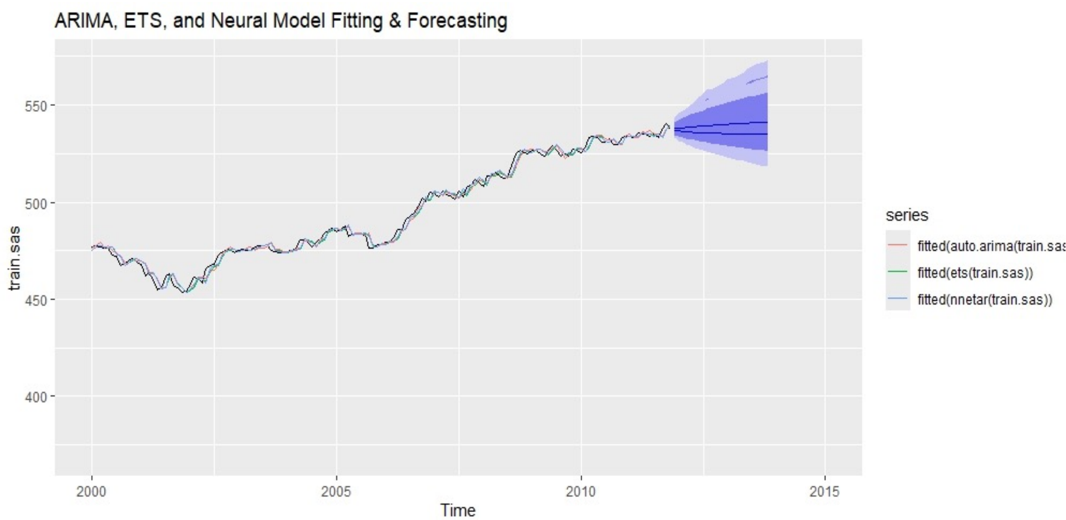
- ETS, ARIMA, and Neural Network Models are fitted to each province employment total counts data regardless of employment type and sex.
- Large Province favors ARIMA
- Smaller Province favors ETS & NN

Province	Population	Best Model
Ontario	14,223,942	ARIMA (1,1,3) (0,0,2) [12]
Quebec	8,501,833	ARIMA (2,1,1) (2,0,1) [12]
British Columbia	5,000,879	ETS (M, Ad, N)
Alberta	4,262,635	ETS (M, Ad, N)
Manitoba	1,342,153	ARIMA (2,1,2) (2,0,1) [12]
Saskatchewan*	1,132,505	ARIMA (3,1,2) (2,0,1) [12]
Nova Scotia	969,383	ETS (M, A, N)
New Brunswick	775,610	NNAR (1,1,2) [12]
Newfoundland and Labrador	510,550	NNAR (3,1,2) [12]
Prince Edward Island	154,331	ETS (M, A, N)

Model 2 - Findings with Modeling on A Province

Saskatchewan

- 6th largest province by population




```
** Teraesvirta's neural network test  **
Null hypothesis: Linearity in "mean"
X-squared =  9.807699  df =  2  p-value =  0.007417972

** White neural network test  **
Null hypothesis: Linearity in "mean"
X-squared =  7.752501  df =  2  p-value =  0.0207284

** Keenan's one-degree test for nonlinearity  **
Null hypothesis: The time series follows some AR process
F-stat =  6.333221  p-value =  0.01221518

** McLeod-Li test  **
Null hypothesis: The time series follows some ARIMA process
Maximum p-value =  0

** Tsay's Test for nonlinearity **
Null hypothesis: The time series follows some AR process
F-stat =  6.649384  p-value =  0.01025284

** Likelihood ratio test for threshold nonlinearity **
Null hypothesis: The time series follows some AR process
Alternativce hypothesis: The time series follows some TAR process
X-squared =  9.178531  p-value =  0.1038953
```

```
—— Accuracy Comparison for Saskatchewan ——
> accuracy(forecast(sas_ets),sas_ts)
                   ME     RMSE      MAE        MPE      MAPE      MASE         ACF1 Theil's U
Training set  0.15522  2.47116  1.90960 0.02943557 0.4191763 0.2582873 -0.001376089        NA
Test set     16.23674 18.31158 16.23674 2.89511291 2.8951129 2.1961367  0.852198964  4.946555
> accuracy(forecast(sas_arima),sas_ts)
                     ME      RMSE       MAE          MPE      MAPE      MASE          ACF1 Theil's U
Training set -0.03139195  2.372782  1.834633 -0.006978111 0.4033773 0.2481474 -0.0004380786        NA
Test set     11.20032278 12.857600 11.278419  1.996077577 2.0105668 1.5254883  0.8113482939  3.472713
> accuracy(forecast(sas_neural),sas_ts)
                     ME      RMSE       MAE          MPE      MAPE      MASE        ACF1 Theil's U
Training set  0.00134703  2.368207  1.839662 -0.002423596 0.4002354 0.2488276 -0.09031669        NA
Test set     20.67639493 22.959604 20.676395  3.689153176 3.6891532 2.7966329  0.86444942  6.204486
```

Model 3 Sex Residual Checks

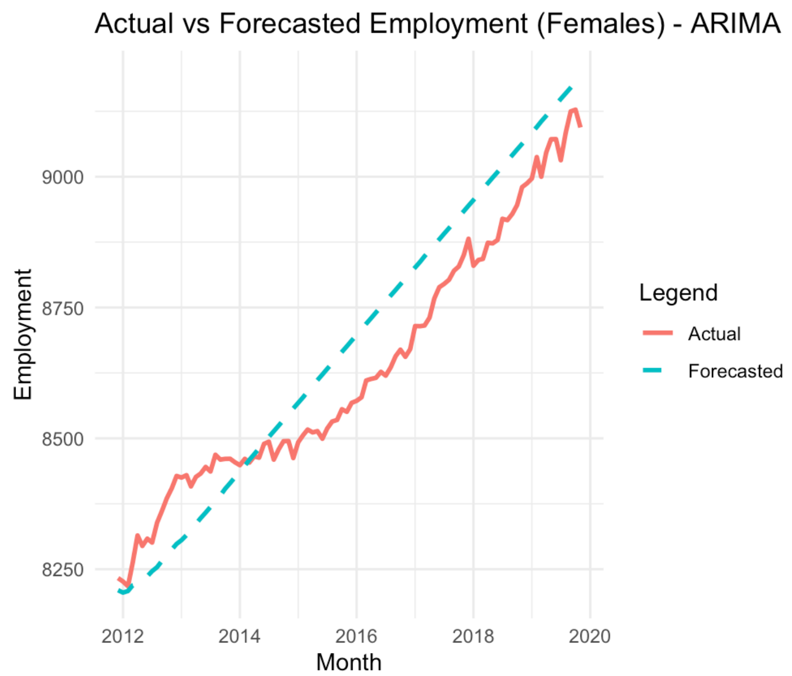
```
> checkresiduals(male_ets_model)

        Ljung-Box test

data:  Residuals from ETS(M,Ad,N)
Q* = 34.22, df = 24, p-value = 0.08084

Model df: 0.   Total lags used: 24
```

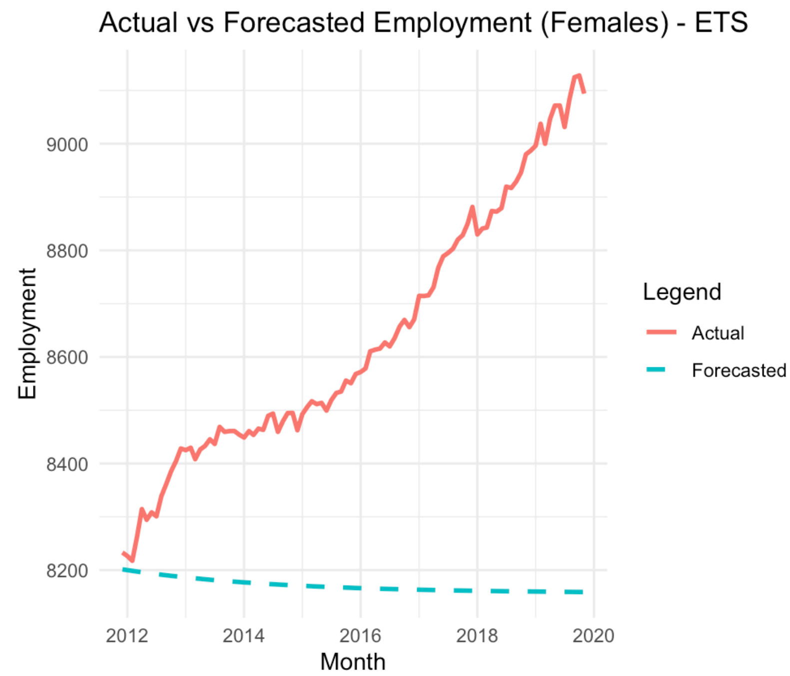
```
> checkresiduals(female_arima_model)

        Ljung-Box test

data:  Residuals from ARIMA(1,1,2)(0,0,2)[12] with drift
Q* = 10.21, df = 19, p-value = 0.9476

Model df: 5.   Total lags used: 24
```

```
> checkresiduals(male_arima_model)

        Ljung-Box test

data:  Residuals from ARIMA(1,1,2)(0,0,2)[12] with drift
Q* = 13.105, df = 19, p-value = 0.8331

Model df: 5.   Total lags used: 24
```

```
> checkresiduals(female_ets_model)

        Ljung-Box test

data:  Residuals from ETS(M,Ad,N)
Q* = 26.376, df = 24, p-value = 0.3344

Model df: 0.   Total lags used: 24
```

Model 3 Female Findings

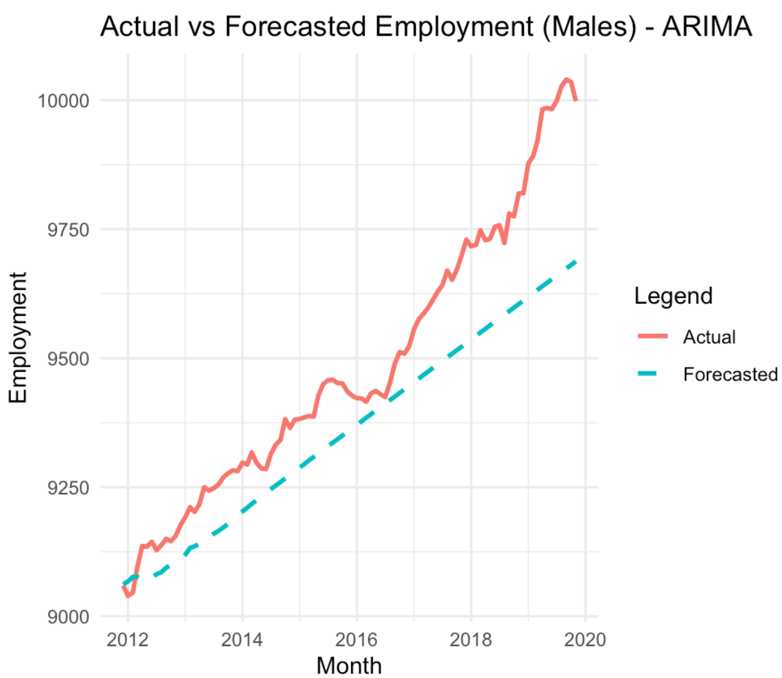


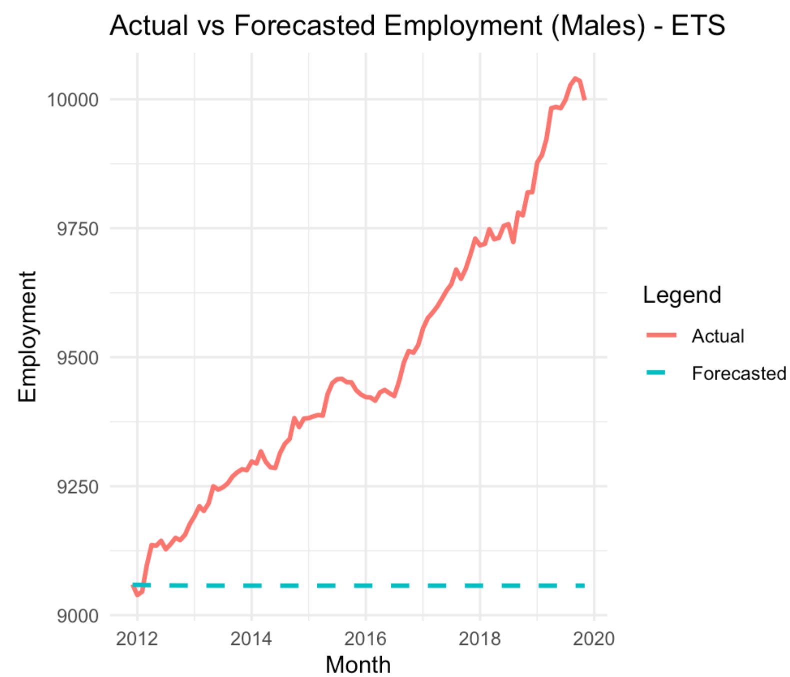



```
> accuracy(forecast(female_arima_forecast), female_actual)
                      ME     RMSE      MAE          MPE      MAPE      MASE        ACF1
Training set -0.002176451 15.58266 12.33345  0.002832798 0.2081054 0.7629782 -0.00670282
Test set    -48.320153970 95.03659 87.34957 -0.544108868 1.0090637 5.4036623          NA
```

```
> accuracy(forecast(female_ets_forecast), female_actual)
                    ME      RMSE       MAE        MPE      MAPE       MASE        ACF1
Training set  1.287655  16.21521  12.88019 0.02404398 0.2170915  0.7968007 -0.03989791
Test set    462.305974 525.84156 462.30597 5.27907781 5.2790778 28.5993999          NA
```

Model 3 Male Findings

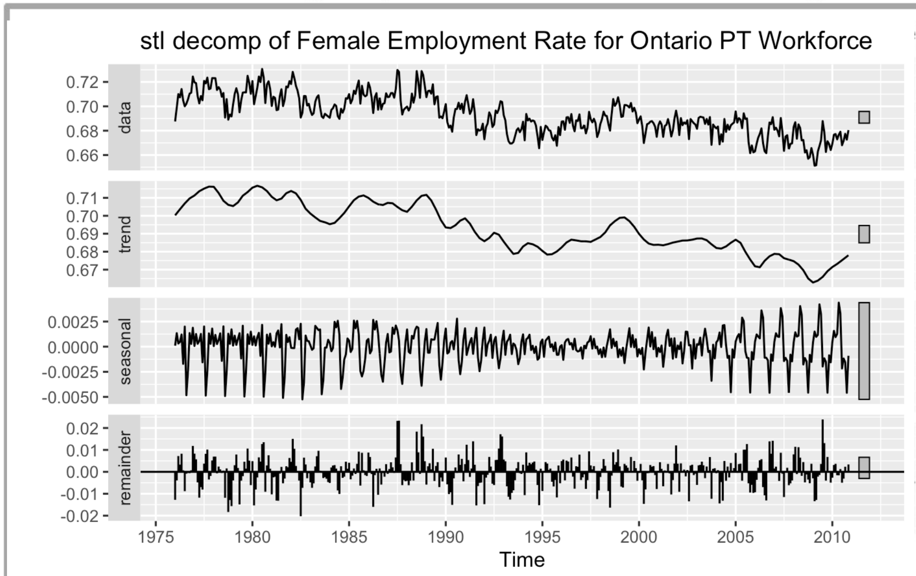





```
> accuracy(forecast(male_arima_forecast), male_actual)
                      ME      RMSE       MAE           MPE      MAPE      MASE         ACF1
Training set  -0.01178028  18.82155  14.52158 -0.0008980298 0.1960787 0.8269145 -0.001795361
Test set     121.96574548 150.83106 123.22652  1.2639983510 1.2779401 7.0169906           NA
> accuracy(forecast(male_ets_forecast), male_actual)
                    ME      RMSE       MAE       MPE      MAPE       MASE         ACF1
Training set  2.165879  19.44039  15.10863 0.0293449 0.2039562  0.8603431 -0.007156636
Test set     425.863263 501.89712 426.53742 4.4166532 4.4241092 24.2886774           NA
```

Model 4 – Methods: Ontario Employment Type x Sex

What proportion of Ontario's part-time workforce has historically been female?



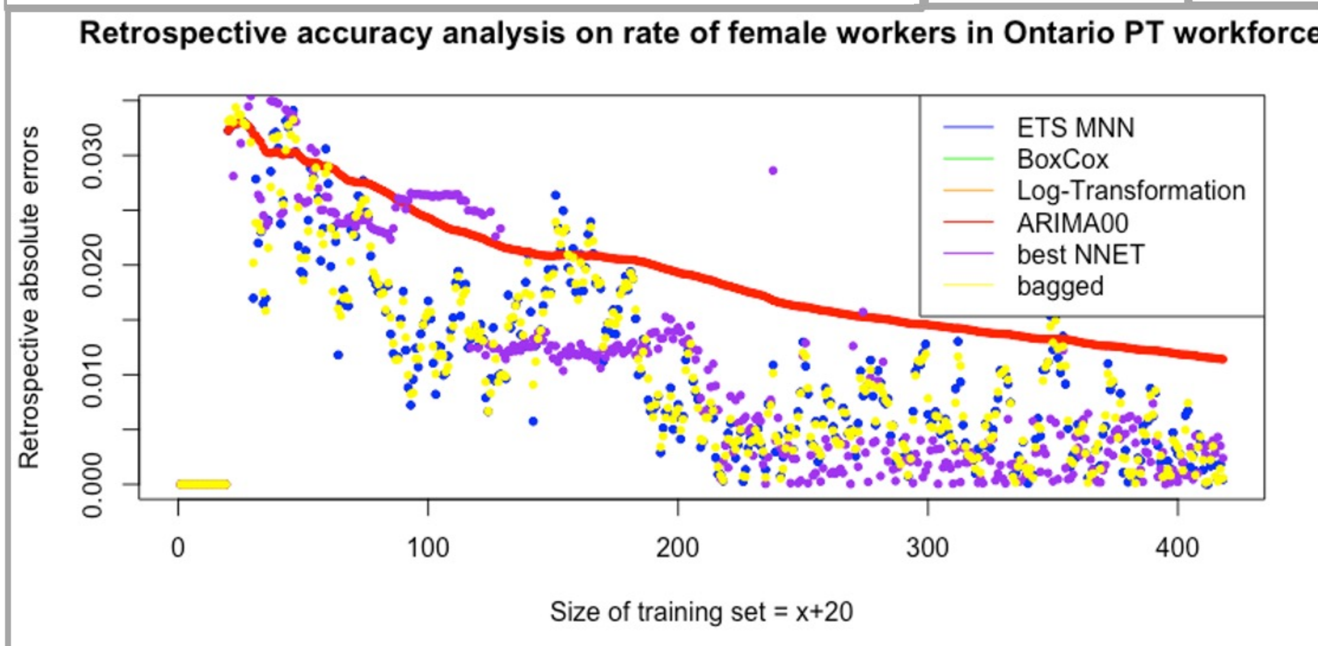
pval.TNN:

0.0357

pval.WNN:

0.0035

	[,1]	[,2]	[,3]
[1,]	"model"	"train MASE"	"test MASE"
[2,]	"naive"	"0.5799"	"1.0281"
[3,]	"randomwalk"	"0.5437"	"0.8084"
[4,]	"best ETS"	"0.5478"	"0.7294"
[5,]	"BoxCox ETS"	"0.5476"	"0.7291"
[6,]	"log-trans ETS"	"0.5479"	"0.729"
[7,]	"best ARIMA"	"1.3209"	"2.3238"
[8,]	"best neural"	"0.033"	"0.9087"
[9,]	"bagged"	"0.5264"	"0.7304"



Entropy: 1.153

F$_T$: 0.8959, **F$_S$**: 0.0113

Best ETS: ETS (M,N,N)

Box Cox: ETS (A,N,N)

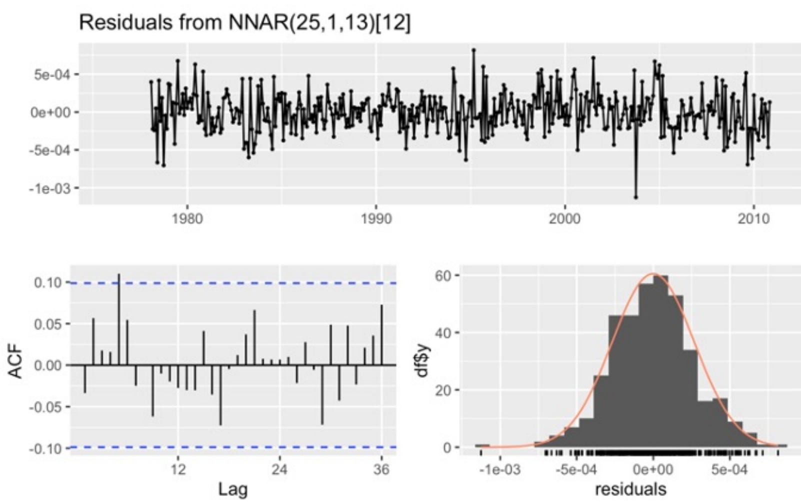
ARIMA: ARI(0,0) nonzero mean

Best neural: NNAR(25,1,13)[12]

Model 4 – Model Choice Rationale

What proportion of Ontario's part-time workforce will likely be female in the future?

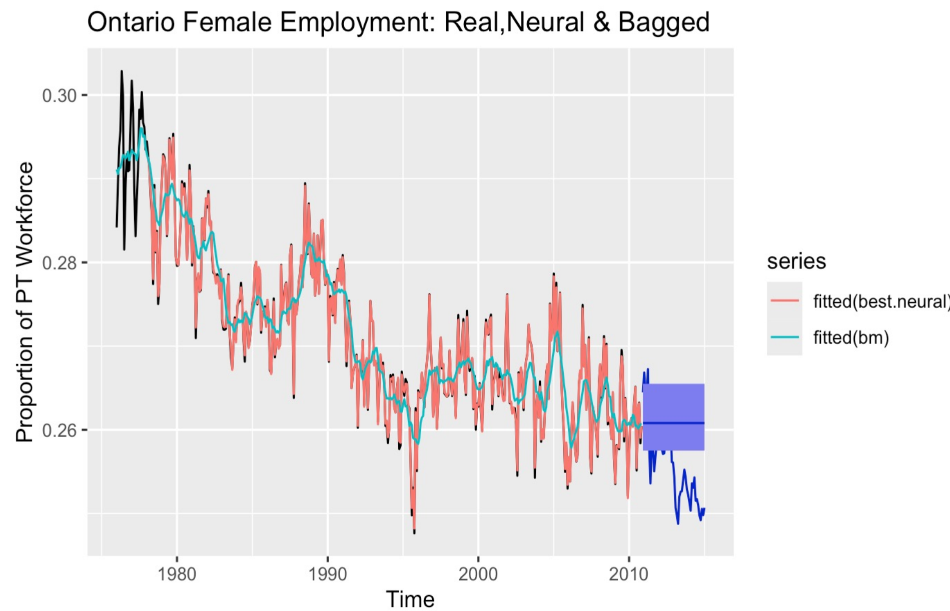
NNET (Model-Dependent)




```
Series: prop_PT.f_ontario.train
Model:  NNAR(25,1,13)[12]
Call:   nnetar(y = prop_PT.f_ontario.train)

Average of 20 networks, each of which is
a 25-13-1 network with 352 weights
options were - linear output units

sigma^2 estimated as 9.533e-08
```

```
        Ljung-Box test

data:  Residuals from NNAR(25,1,13)[12]
Q* = 17.167, df = 24, p-value = 0.8415

Model df: 0.   Total lags used: 24
```

Bagged (Data-Dependent)




```
Series: prop_PT.f_ontario.train
Model:  baggedModel
Call:   baggedModel(y = prop_PT.f_ontario.train, bootstrapped_series
= bld.mbb.bootstrap(prop_PT.f_ontario.train,
    15))
```

```
        Ljung-Box test

data:  Residuals from baggedModel
Q* = 167.67, df = 24, p-value < 2.2e-16

Model df: 0.   Total lags used: 24
```

Conclusions and future work

- Understanding employment trends, seasonality, and forecasts helps the Canadian government publish effective employment policies.

- Understanding the employment trends gives policymakers a “heads-up" before a significant seasonal fluctuation occurs.

- Limitations:
  - Lack of annual population data for each province
  - Lack of more relevant variables for deeper and broader research

- Future work:
  - Collect further data to analyze significant events' impact on employment
  - Study the impact of a significant employment policy on employment rate

References

Economic Insights. Statistics Canada. (2020). https://www150.statcan.gc.ca/n1/en/catalogue/11-626 X#wb-auto-2

Employment Rate. OECD. (2023). https://www.oecd.org/en/data/indicators/employmentrate.html

Guillemette, R., L'Italien, F., & Grey, A. (2000). Seasonality of Labour Markets Comparison of Canada, the U.S. and the Provinces. *Government of Canada*, (R-00-8E). https://doi.org/https://publications.gc.ca/collections/Collection/MP32-29-00-8E.pdf

Immigration to Canada's Provinces and Major Industries. Winny Immigration. (2021). https://winnyimmigration.com/immigration-to-canadas-provinces-and-major-industries

Ortiz-MacLeod, D. (2020). Employment Rate in Canada, Version 1. Retrieved November 21, 2024, from https://www.kaggle.com/datasets/ortizmacleod/employment-rate-canada/data

Suthar, J. (2022). *Employment Rate*. WallStreetMojo. https://www.wallstreetmojo.com/employment-rate/

Tarzwell, G. (1997). Canadian City Unemployment Rates and the Impact of Economic Diversity. *The Canadian Journal of Regional Science*, 20(3), 389-.

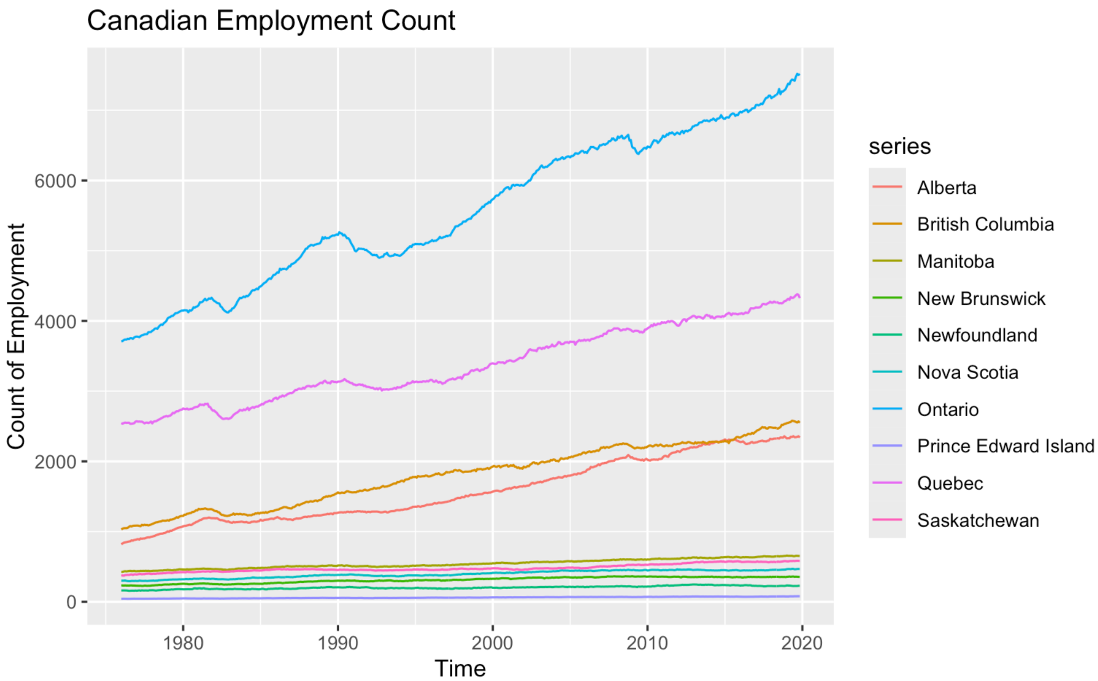
Thank you!

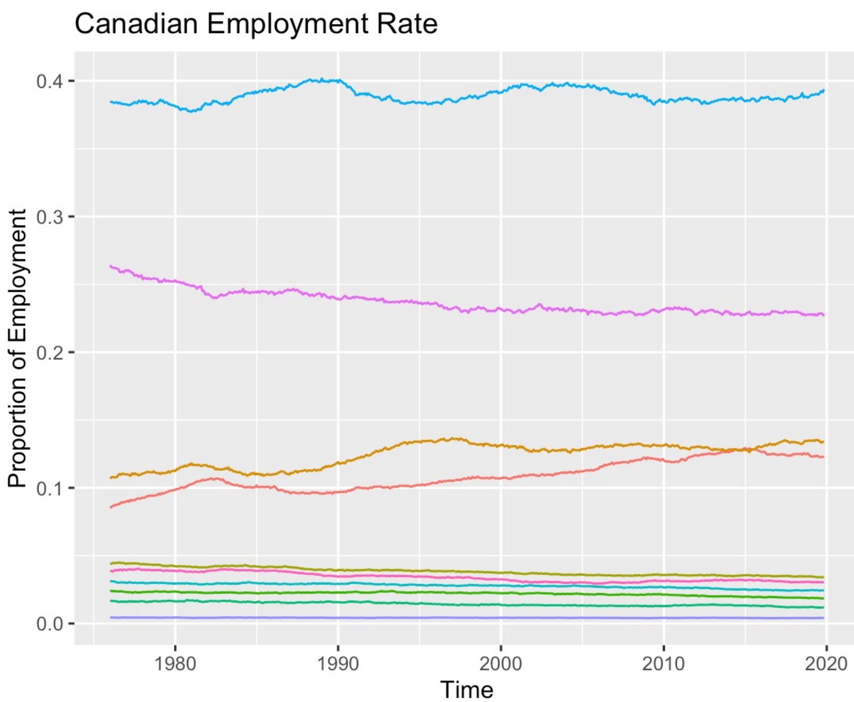
Questions or comments?

Appendix A: Zooming In on Rates EDA

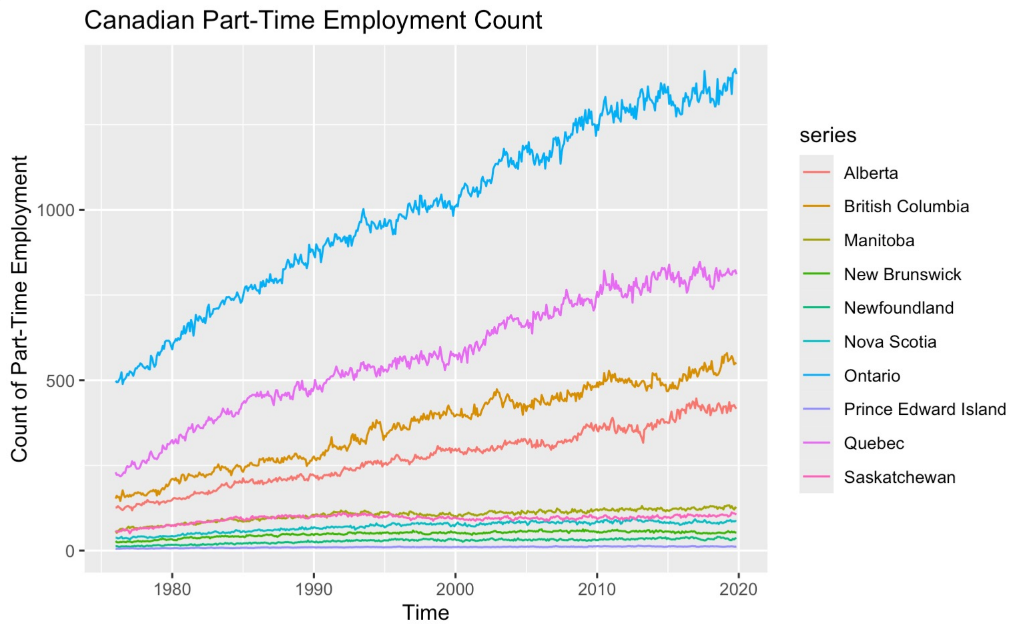
EDA – Employment Rates

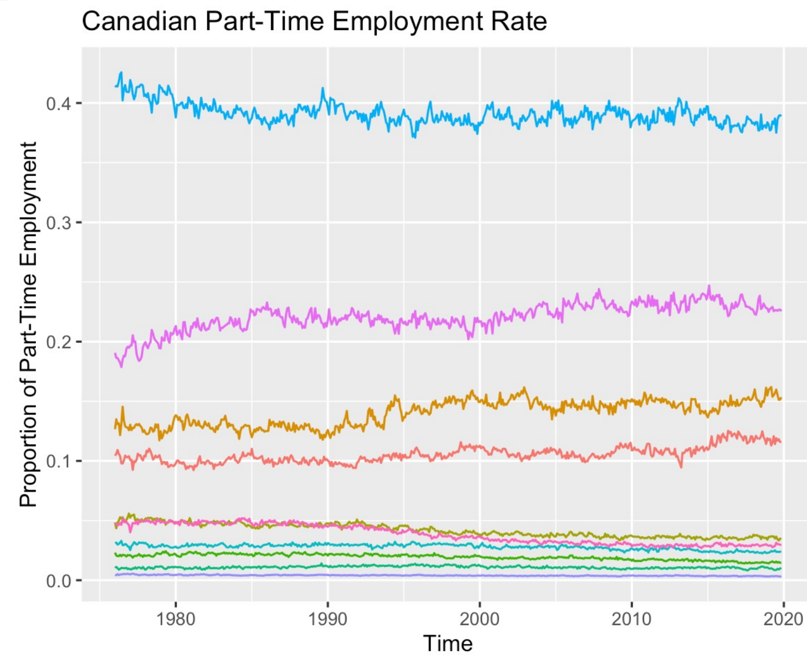
Why Did We Choose to Include Rate Data?

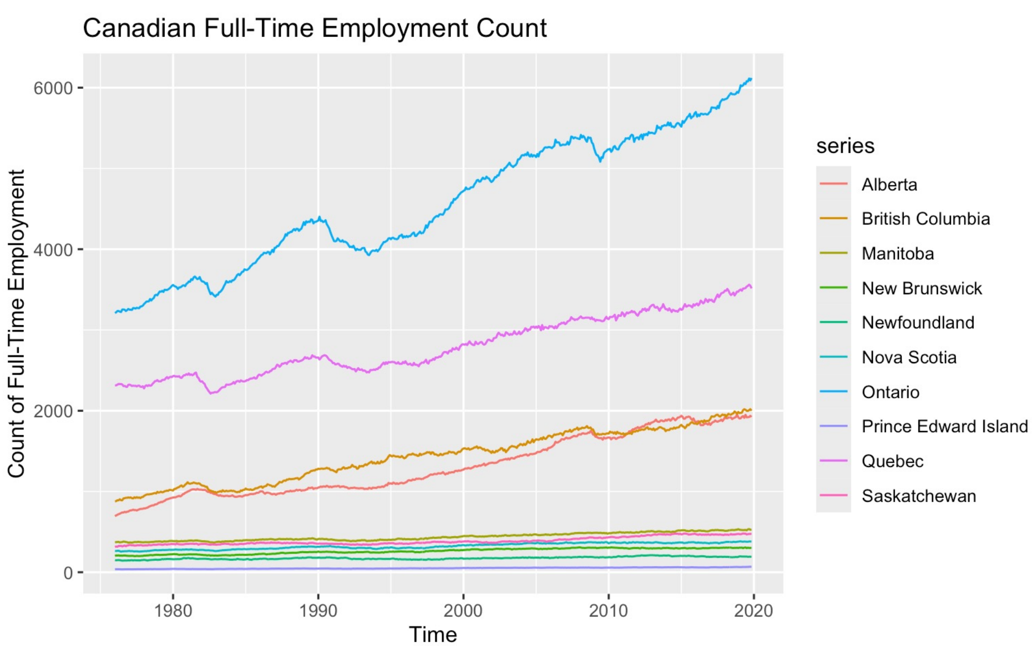


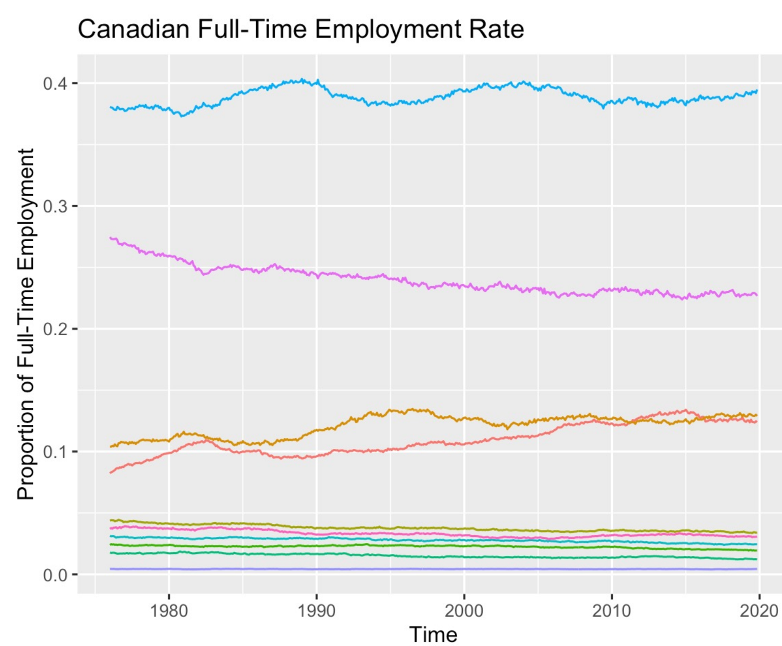


EDA – **Employment Type** per Province : Counts vs Rates



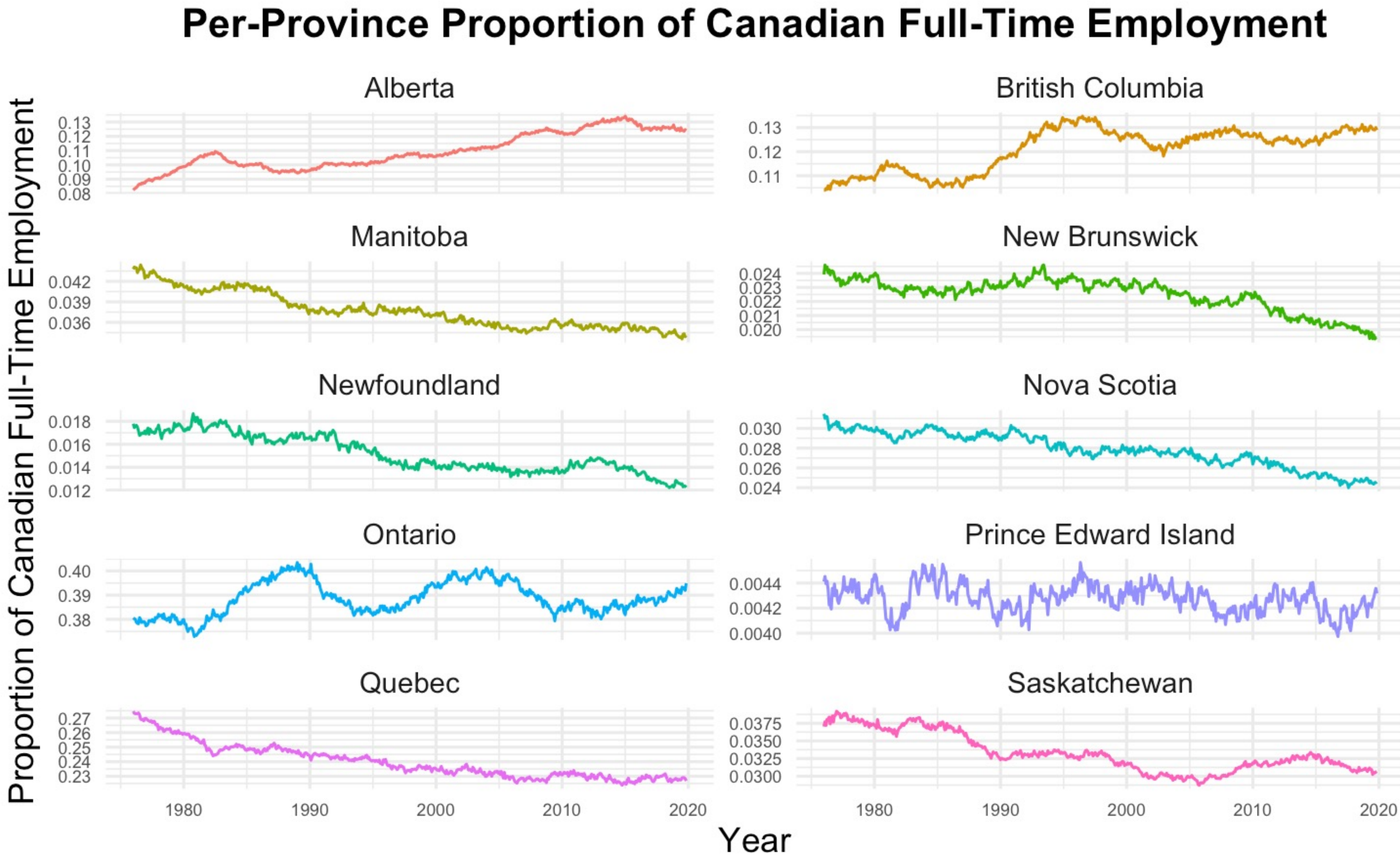






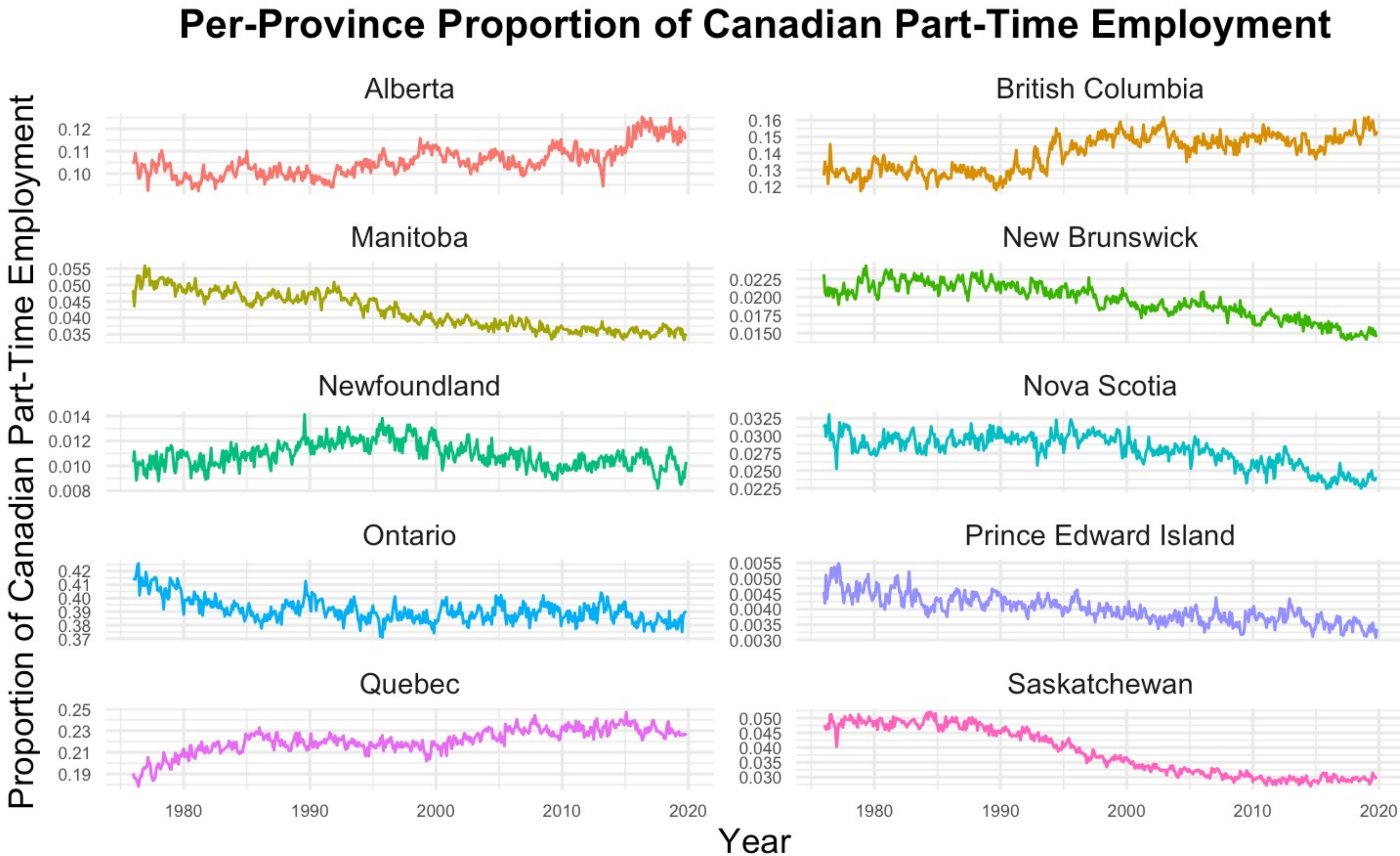
EDA – Rate of **Full-Time Employment** per Province

Total: *All* FULL-TIME workers in Canada (what proportion are working in each province?)

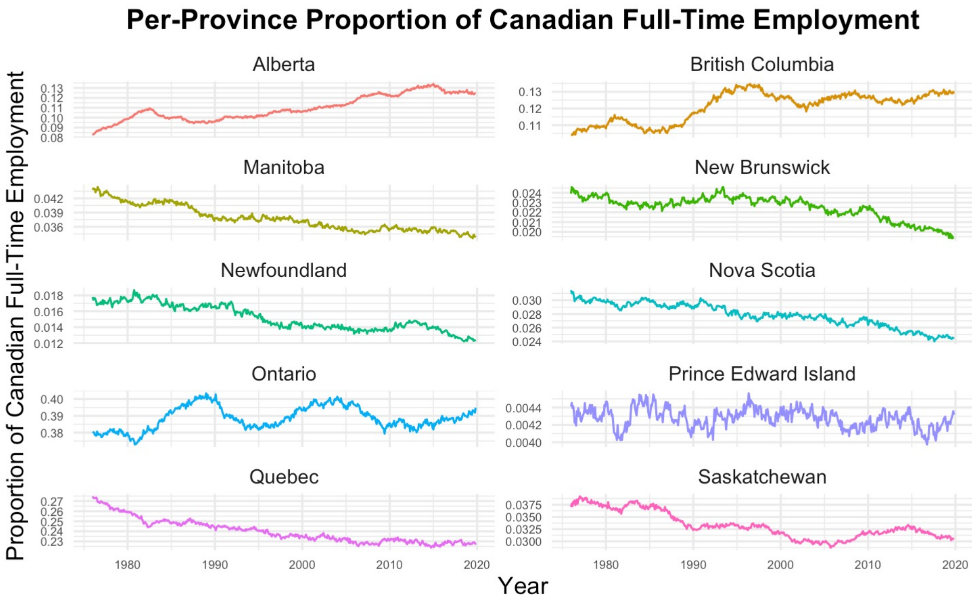


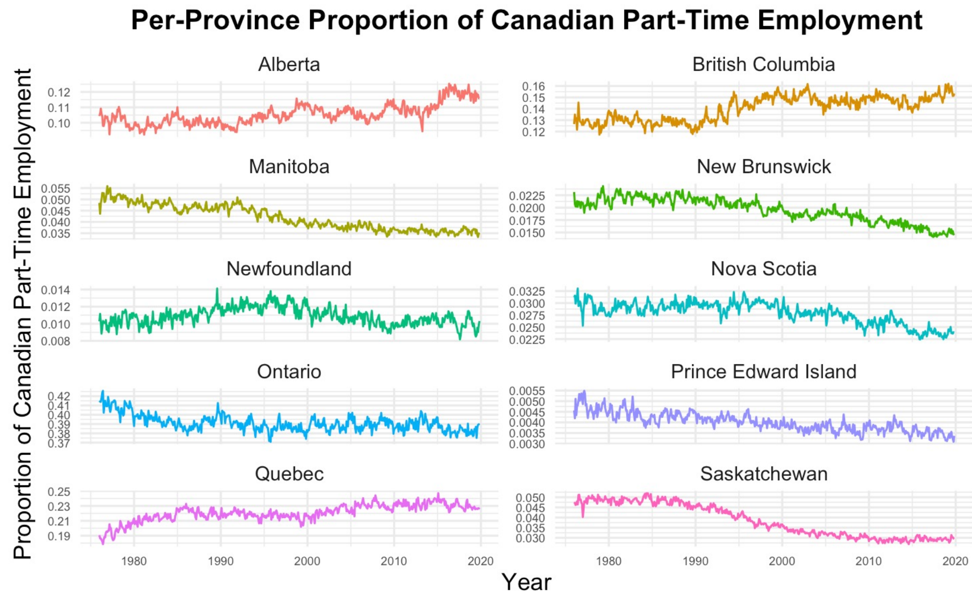
EDA – Rate of **Part-Time Employment** per Province

Total: *All* PART-TIME workers in Canada (what proportion are working in each province?)



EDA – Provincial Rate by Employment Type

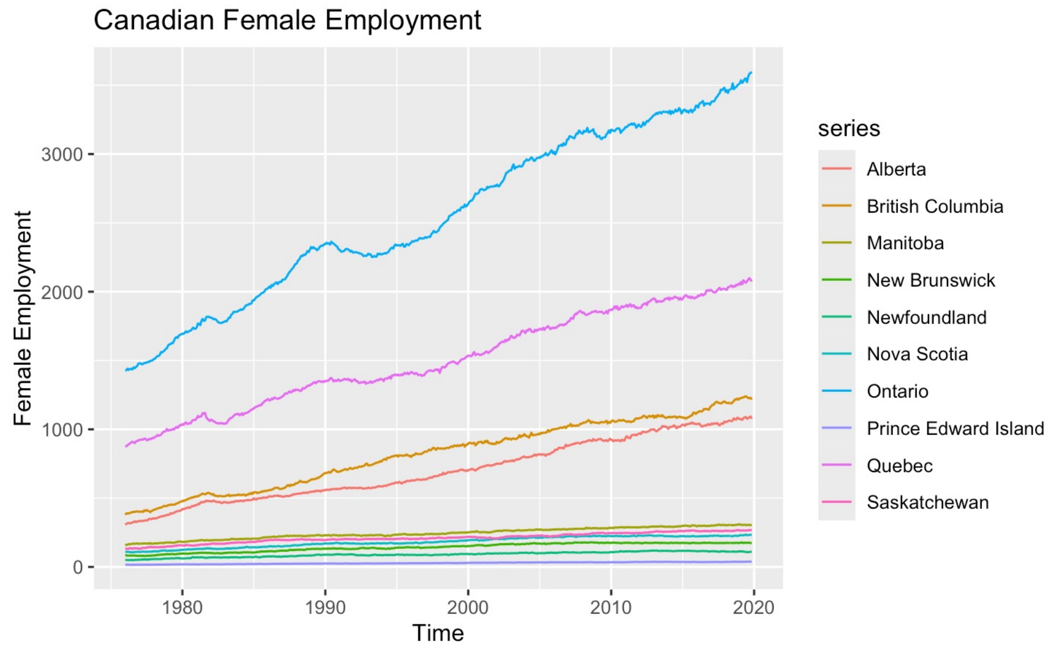


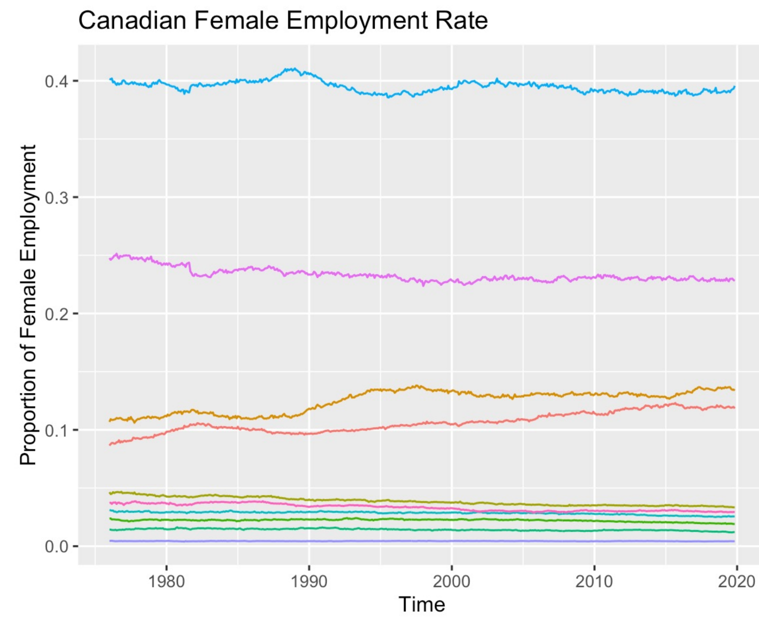


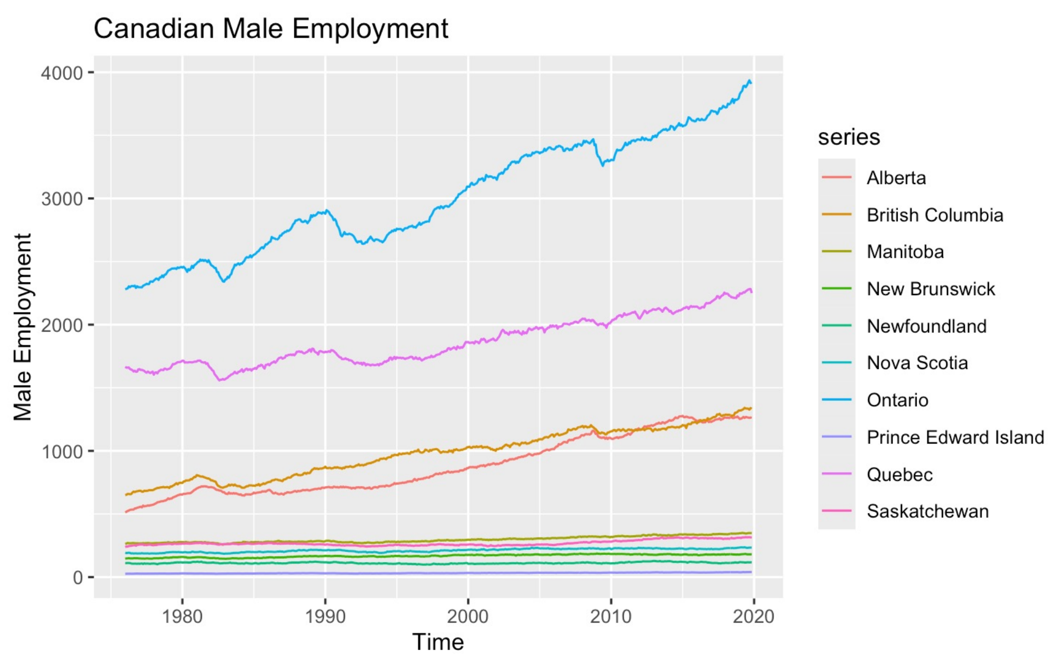
Data-Driven Questions:

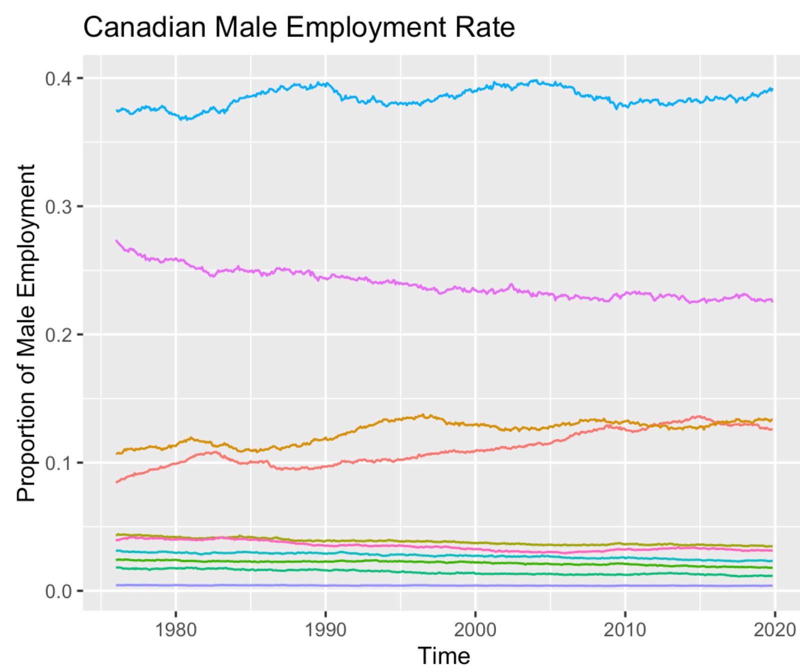
- Why do trend and seasonality across employment type appear *so different* for these provinces:
  - **Prince Edward Island**, **Quebec**, and **Ontario** ?
- In contrast, why do the they appear to *change so little* for these other provinces:
  - **Saskatchewan**, **Manitoba**, and **British Columbia** ?

EDA – **Sex** per Province: Counts vs Rates



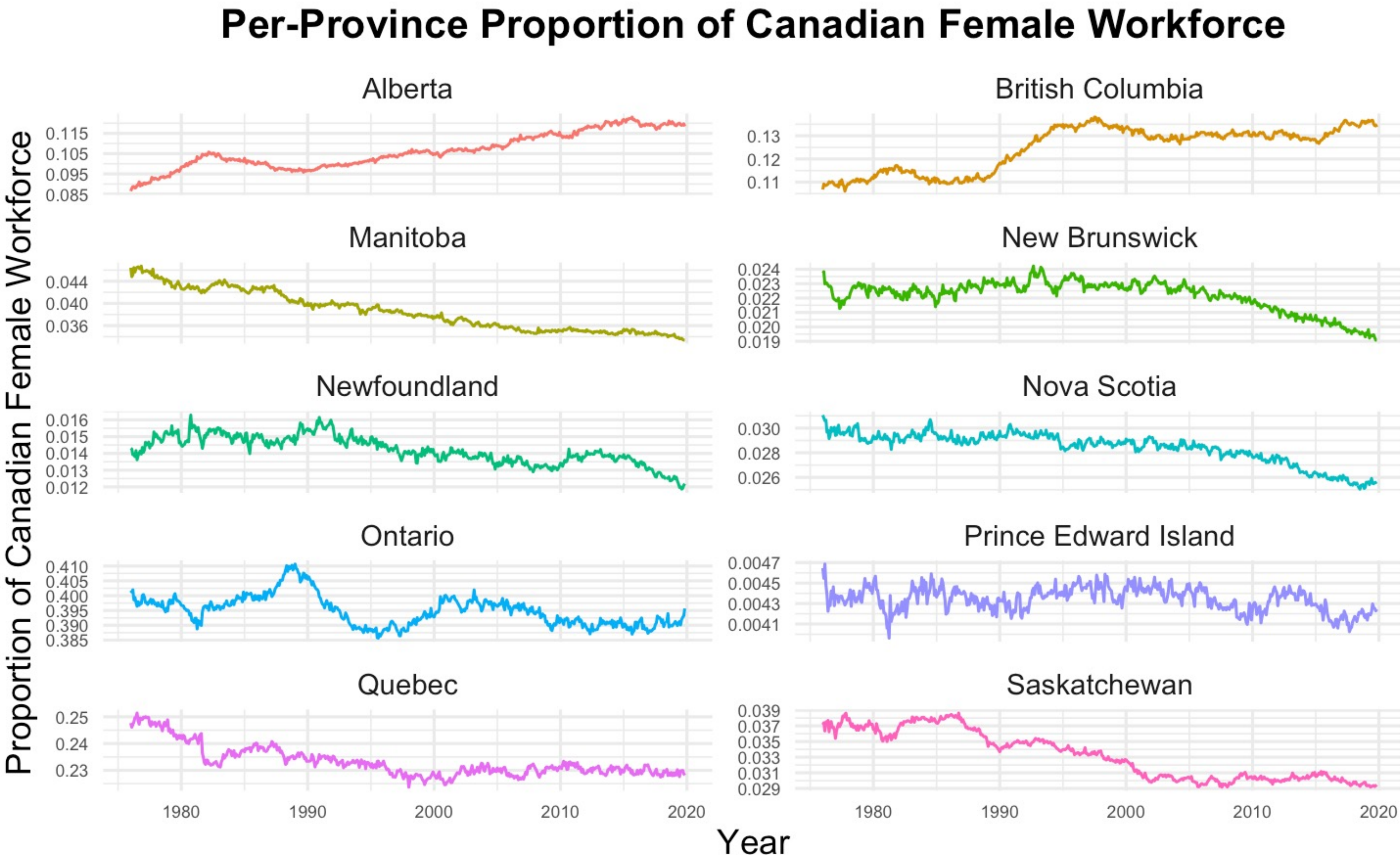






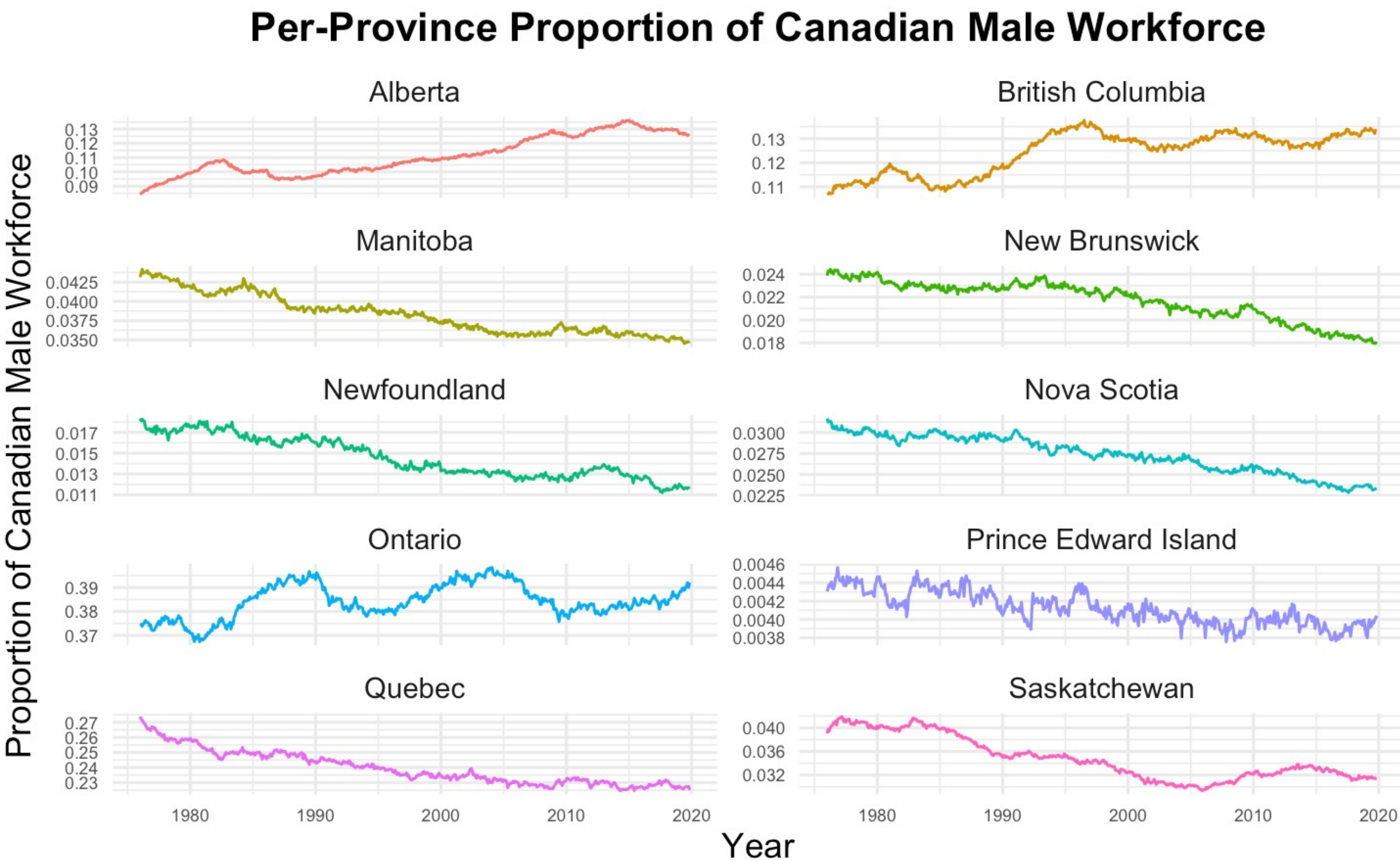
EDA – Rate of **Female** Employment per Province

Total: *All* FEMALE workers in Canada (what proportion are working in each province?)

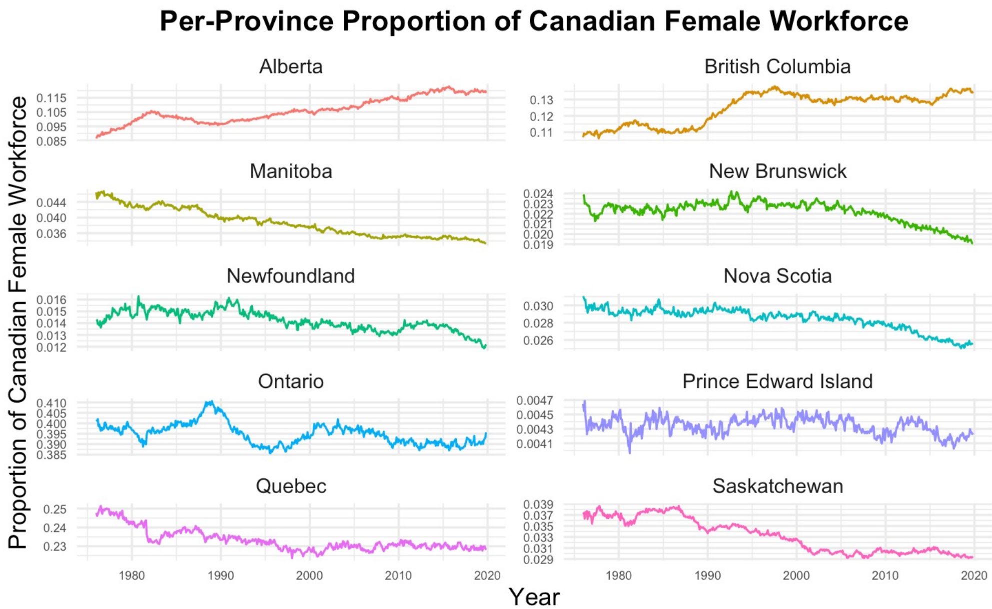


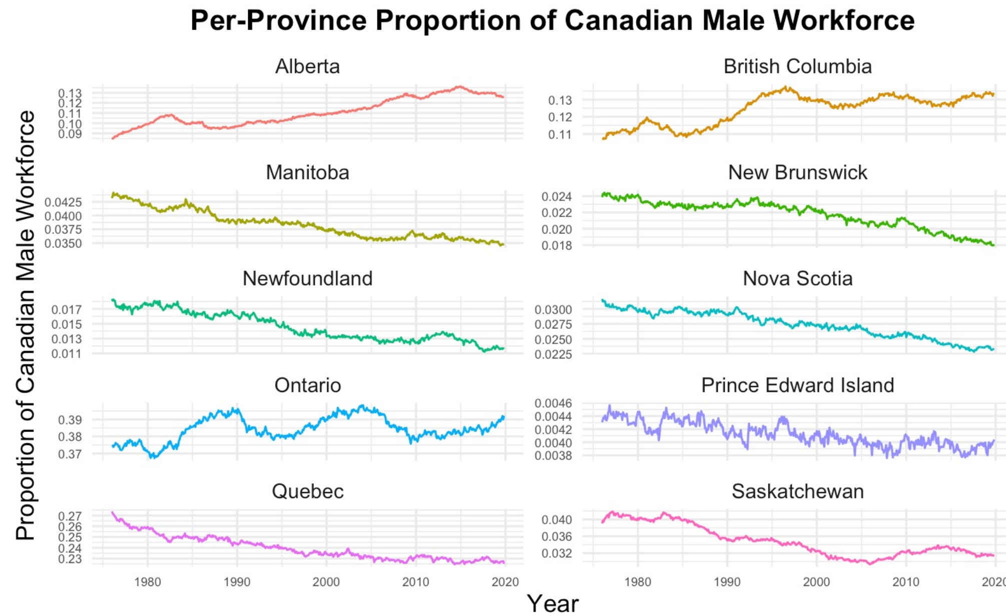
EDA – Rate of **Male** Employment per Province

Total: *All* MALE workers in Canada (what proportion are working in each province?)



EDA – Provincial Employment Rate by Sex





Data-Driven Questions

- Is there *more variance* in female vs. male employment rate in these provinces and why:
  - **Newfoundland**, **New Brunswick**, **Prince Edward Island**, **Quebec**, and **Nova Scotia** ?
- How do *trend and seasonality* in employment rates by sex compare for these provinces and why:
  - **Newfoundland**, **New Brunswick**, **Prince Edward Island**, and **Ontario** ?