Predicting Death Claim Seasonality in Life Insurance - Historic Claims Training Data v Population Deaths Training Data

Jennifer Loftus, Acorn Life, Ireland

5th European
Congress of Actuaries
www.eca2024.org

About the speaker

- **Jennifer Loftus –** *Executive Director & Group CFO, Acorn Life*

Jennifer is an actuary and an accountant with over 20 years' experience in the life insurance industry. She is a member of the Society of Actuaries in Ireland's Data Science Committee and the Institute of Actuaries (UK) Actuarial Data Science Working Party. She is also an Independent Non-Executive Director of VHI, the Irish state-owned private health insurance company.

Acorn Life is based in Galway, Ireland and offers life insurance, pensions and savings solutions to Irish customers. The company was founded in 1989 and currentlyhas under management of over €1.1 billion.

Acorn Life

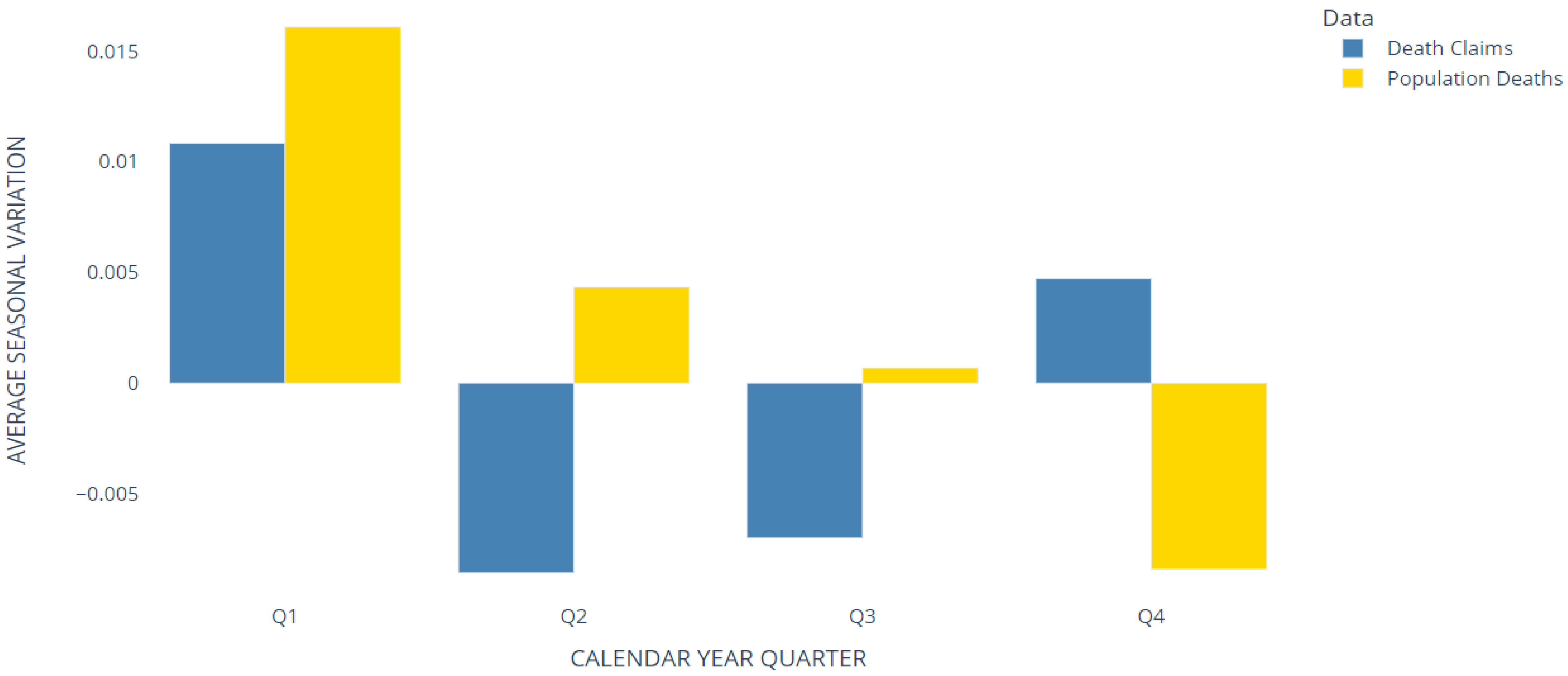
Objectives

1. Compare effectiveness of using life insurer's historic claims data with Irish population data in predicting life insurer's future death claim seasonality
2. Compare effectiveness of machine learning time series models with traditional statistical methods for this purpose

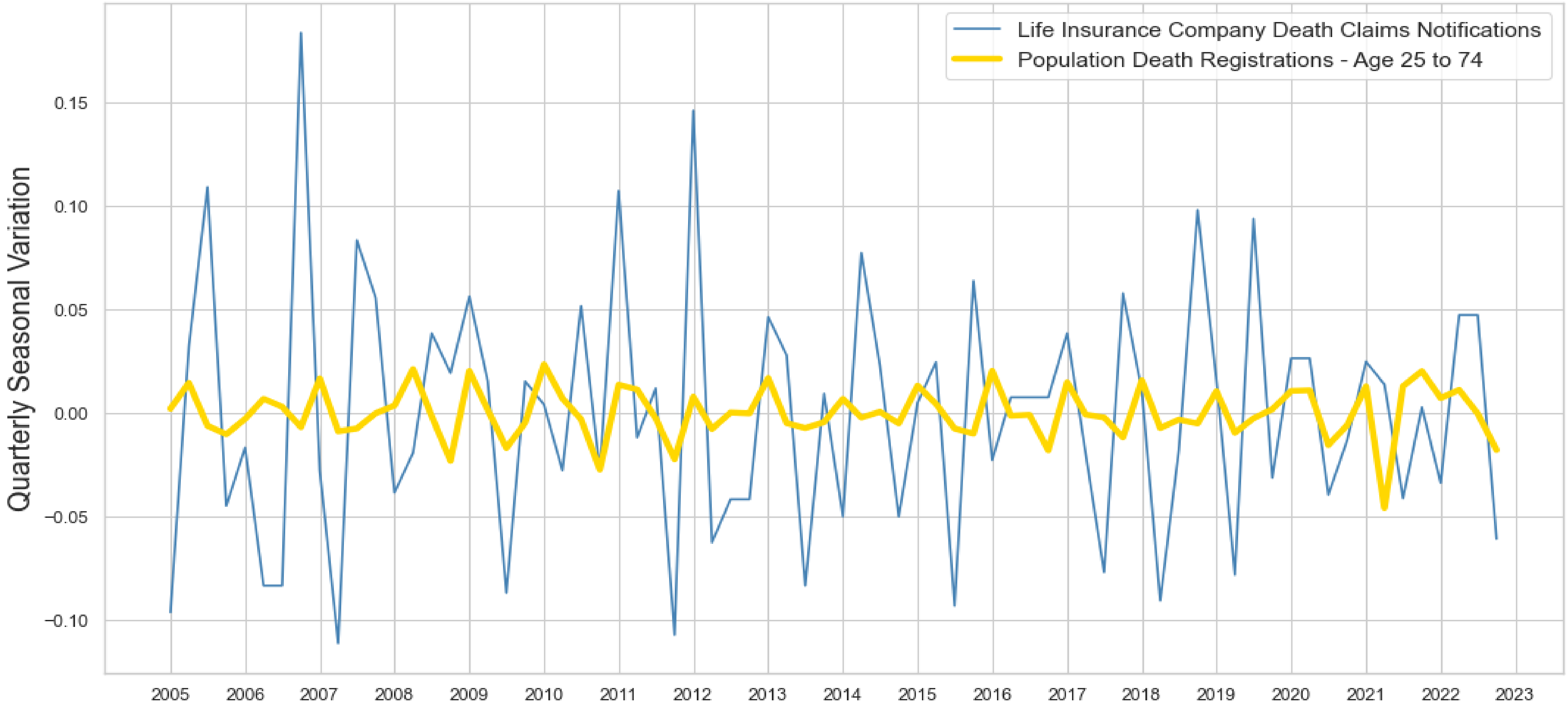
Overview

- Data
  - Irish-based Life Insurance Company Historical Death Claims – 1,117 Deaths
    - 19-year observation period (2005-2023)
    - Covid-19 deaths removed
    - Quarterly Seasonal Variation = Proportion of annual death claims notified per calendar year quarter less 0.25
    - Sum of each calendar year's seasonal variation = 1
    - Claim size ignored
  - Irish Population Deaths data for 25-75 year-old cohort – 95,295 Deaths
    - Same observation period as Death Claims
    - Quarterly Seasonal Variation = Proportion of annual deaths registered per calendar year quarter less 0.25
    - Sum of each calendar year's seasonal variation = 1
- Trend component was removed from both datasets = stationary time series

Average Seasonal Variation per Quarter



Seasonal Variation: Observation Period



Modelling

- 5 types of Time Series Model trained on **both** Death Claims Data and Irish Population Deaths data using Python libraries
  - 3 Machine Learning Regression Models
    - Deep Learning Model x3
  - Classical Time Series Models (SARIMA)
  - A simple Historical Average Model
- Baseline Model
  - No Seasonality
  - Time Series Model Performance compared against Baseline

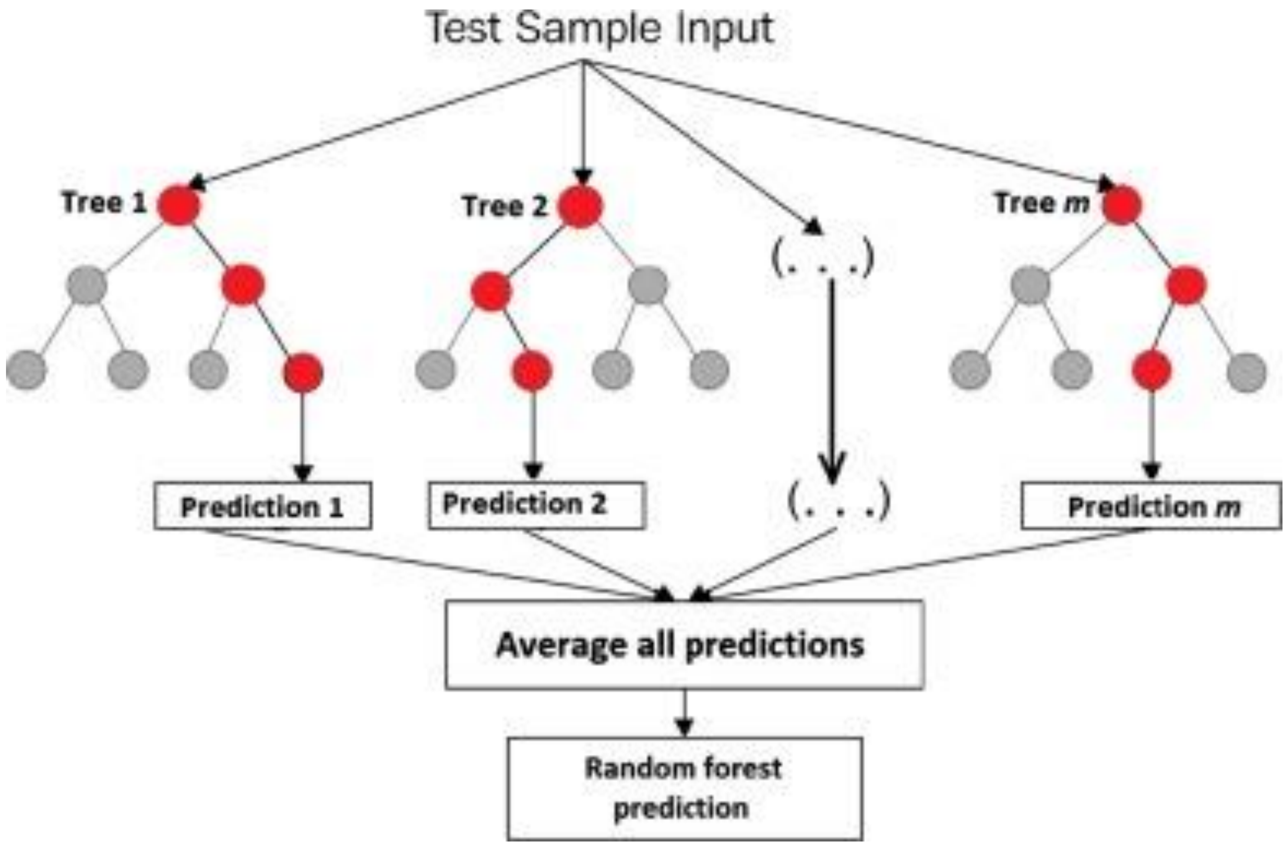
Modelling

- Training Data
  - Death Claims Data
  - Population Deaths Data
    - 2005-2019 (15 years)
- Testing Data
  - Death Claims Data
    - 2020-2023 (4 years)
    - 2020-2022 (3 years)
    - 2020-2021 (2 years)
    - 2020 (1 year)

From 2020 to 2022, Covid-19 altered the typical seasonal pattern of deaths observed in a calendar year. However, as the Training Data period ended immediately prior to 2020 and Covid-19 claims were removed from the Training Data, neither the Training or the Testing Data were impacted by Covid-19.

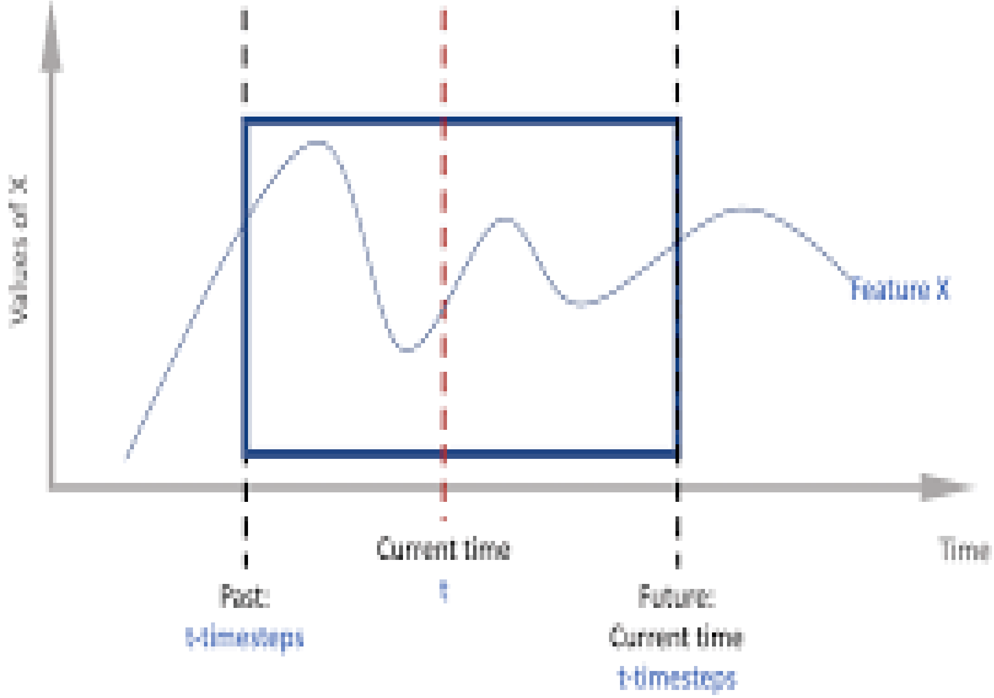
Machine Learning Regression Models

- Random Forest Regressor
  - Collection of Decision Tree Models
  - Predictions made by averaging prediction of each Decision Tree
  - Python library: scikit-learn (RandomForestRegressor)



Machine Learning Regression Models

- LSTM (Long Short-Term Memory)
  - Special type of Recurrent Neural Network (RNN) used for Deep Learning
  - Takes into account current (short-term) information while also remembering important long-term information
  - 4 Timesteps used
    - 15 years Training Data = 60 quarters * 4 timesteps = 240 Datapoints
  - Python libraries: TensorFlow, Keras



Machine Learning Regression Models

- 3 LSTMs trained

LSTM_1: Training Data had 1 Feature – Death Claim Seasonality OR Population Death Seasonality

LSTM_2: Training Data had 2 Features – Death Claim Seasonality OR Population Death Seasonality

and

Calendar Year Quarter Number (1,2,3 or 4)

LSTM_3: Training Data had 3 Features – Death Claim Seasonality

and

Population Death Seasonality

and

Calendar Year Quarter Number (1,2,3 or 4)

Machine Learning Regression Models

- Prophet
  - Developed by Meta’s Data Science time
  - Designed for Time Series forecasting in a business context
  - Specifically captures seasonality
  - Python library: fbprophet

Classical Time Series Model

- SARIMA
  - Seasonal Autoregressive Integrated Moving Average Model
  - Specifically captures seasonality
  - Python library: statsmodels (SARIMAX)

Evaluation Metrics

- MAE (Mean Absolute Error)
  - Mean of the Absolute Value of the Prediction Errors

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i |$$

- RMSE (Root Mean Square Error)
  - Standard Deviation of the Prediction Errors

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$$

- Evaluated for each individual Train/Test period
- Average taken across 4 Train/Test periods

- Python library: scikit-learn

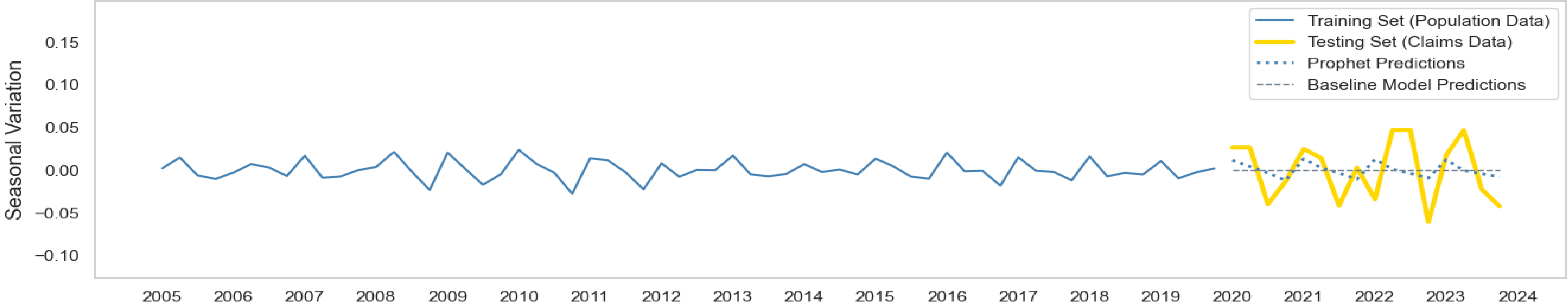
Results (Mean Absolute Error)

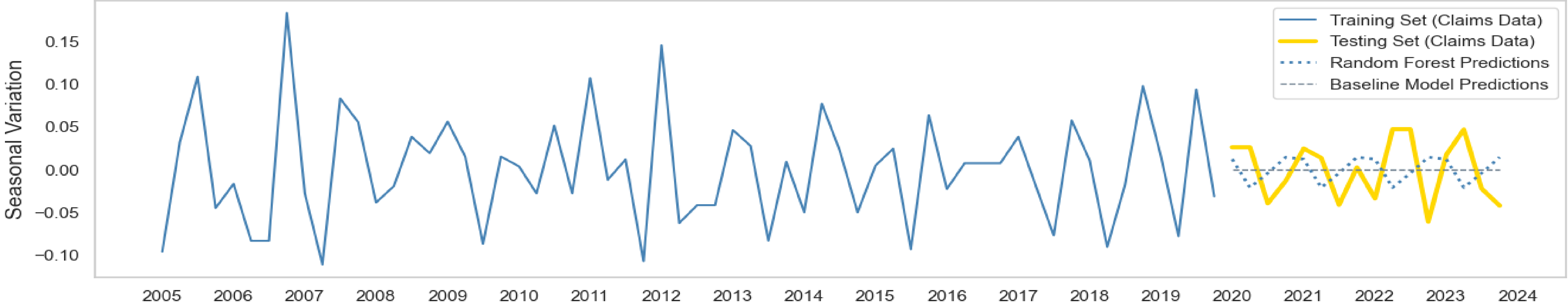
	4 Year Test MAE	3 Year Test MAE	2 Year Test MAE	1 Year Test MAE	Average MAE
Prophet (Population Data)	0.0281	0.0286	0.0185	0.0186	**0.0188**
Random Forest (Population Data)	0.0279	0.0288	0.019	0.0194	**0.019**
Historical Average (Population Data)	0.028	0.0289	0.0191	0.0194	**0.0191**
LSTM_2 (Population Data)	0.0283	0.0296	0.0188	0.0217	**0.0197**
SARIMA (Population Data)	0.0291	0.0301	0.0203	0.0231	**0.0205**
Prophet (Claims Data)	0.0361	0.0332	0.02	0.017	**0.0213**
LSTM_1 (Population Data)	0.032	0.0303	0.0226	0.0245	**0.0219**
LSTM_3 (Combined Data)	0.0297	0.0323	0.0229	0.0255	**0.0221**
Baseline Model	0.0316	0.0314	0.0235	0.0263	**0.0226**
LSTM_2 (Claims Data)	0.0315	0.0317	0.0243	0.0257	**0.0226**
SARIMA (Claims Data)	0.0343	0.0341	0.026	0.0243	**0.0237**
LSTM_1 (Claims Data)	0.0339	0.0364	0.0292	0.0336	**0.0266**
Historical Average (Claims Data)	0.038	0.0384	0.0275	0.031	**0.027**
Random Forest (Claims Data)	0.0381	0.0385	0.0276	0.0311	**0.0271**

Results (Root Mean Squared Error)

	4 Year Test RMSE	3 Year Test RMSE	2 Year Test RMSE	1 Year Test RMSE	Average RMSE
Random Forest (Population Data)	0.0324	0.0334	0.0222	0.0228	**0.0222**
Prophet (Population Data)	0.0328	0.0335	0.022	0.0225	**0.0222**
Historical Average (Population Data)	0.0324	0.0334	0.0222	0.0229	**0.0222**
LSTM_2 (Population Data)	0.0331	0.0341	0.0218	0.0238	**0.0226**
SARIMA (Population Data)	0.0336	0.0345	0.0233	0.0261	**0.0235**
LSTM_3 (Combined Data)	0.0348	0.0366	0.0253	0.0271	**0.0248**
LSTM_2 (Claims Data)	0.0353	0.0355	0.0263	0.0272	**0.0249**
Baseline Model	0.0351	0.0353	0.0265	0.0279	**0.025**
LSTM_1 (Population Data)	0.0377	0.0345	0.0255	0.0275	**0.0251**
Prophet (Claims Data)	0.0441	0.0399	0.0223	0.02	**0.0253**
SARIMA (Claims Data)	0.0374	0.038	0.0292	0.026	**0.0261**
LSTM_1 (Claims Data)	0.0382	0.0408	0.0322	0.0387	**0.03**
Historical Average (Claims Data)	0.0436	0.0431	0.0302	0.0333	**0.0301**
Random Forest (Claims Data)	0.0438	0.0432	0.0303	0.0334	**0.0301**

Results: Best and Worst Performing Models

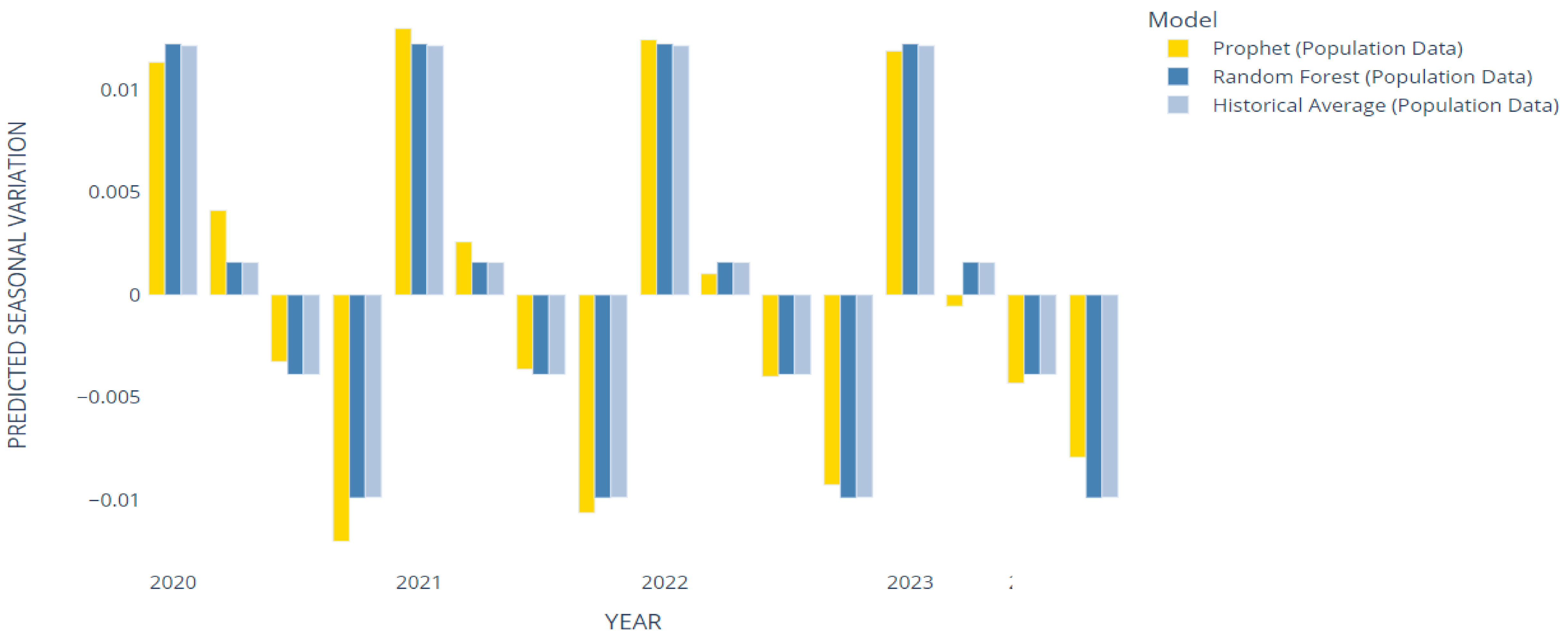




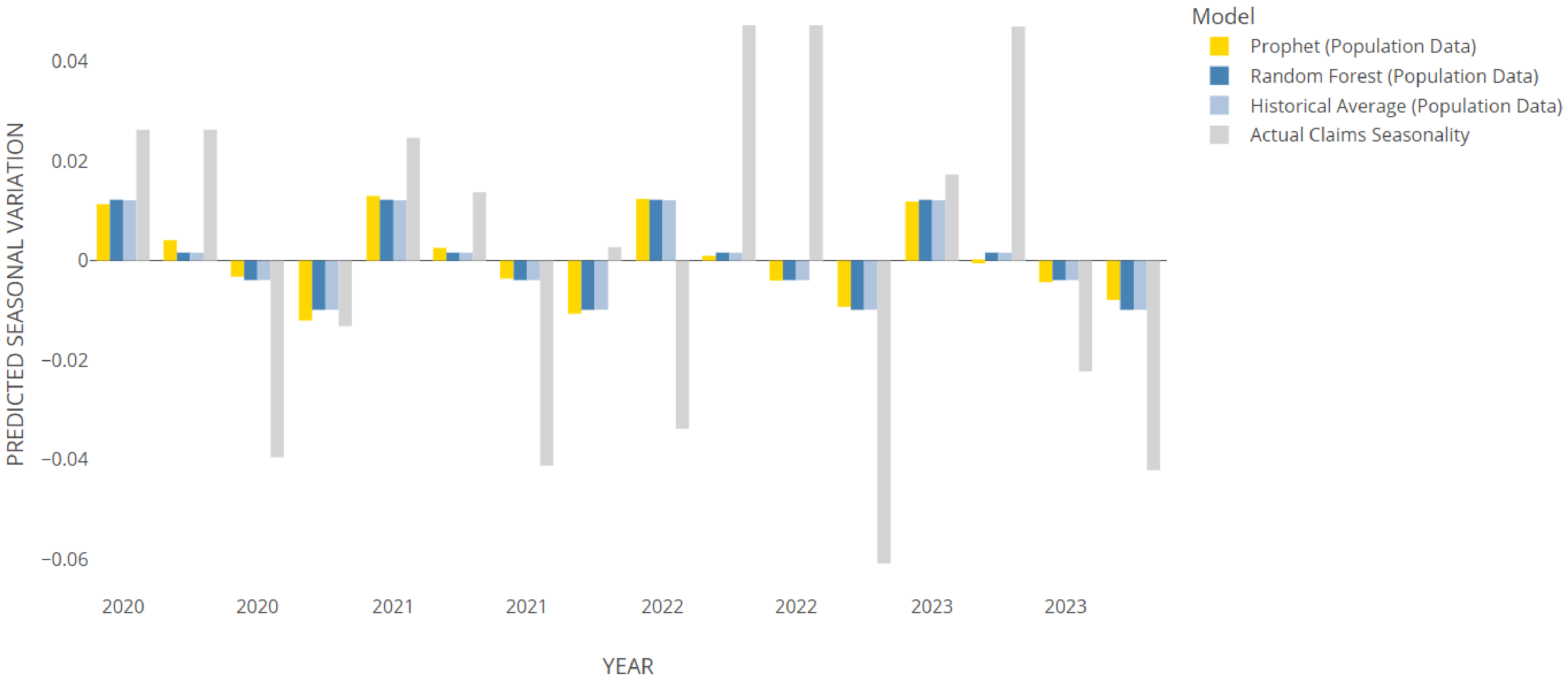
Comparison of Best Performing Model Predictions

NOTIFICATION QUARTER	Prophet (Population Data)	Random Forest (Population Data)	Historical Average (Population Data)
2020-01-01	0.011344	0.012245	0.012151
2020-04-01	0.004120	0.001600	0.001584
2020-07-01	-0.003252	-0.003873	-0.003873
2020-10-01	-0.012017	-0.009890	-0.009861
2021-01-01	0.012995	0.012245	0.012151
2021-04-01	0.002583	0.001600	0.001584
2021-07-01	-0.003620	-0.003873	-0.003873
2021-10-01	-0.010628	-0.009890	-0.009861
2022-01-01	0.012436	0.012245	0.012151
2022-04-01	0.001029	0.001600	0.001584
2022-07-01	-0.003970	-0.003873	-0.003873
2022-10-01	-0.009259	-0.009890	-0.009861
2023-01-01	0.011894	0.012245	0.012151
2023-04-01	-0.000541	0.001600	0.001584
2023-07-01	-0.004301	-0.003873	-0.003873
2023-10-01	-0.007910	-0.009890	-0.009861

Comparison of Best Performing Model Predictions



Best Performing Models v Actual Claims Seasonality



Best Performing Models v Actual Claims Seasonality

Conclusions

- Irish Population deaths data (age 25-75) better predictor of future death claim seasonality than historic claims data
- MAE and RMSE rankings broadly similar
- Random Forest, Prophet and Historical Average = top performers under both MAE and RMSE
- LSTM performance improved when Quarter number incorporated as separate training feature
- LSTM trained on combined Population and Claims data outperformed Baseline Model but not the LSTM trained on Population data only
- Random Forest Regressor = Historical Average
  - Simple statistical model often the best model!!
- Average ranking across 5 Train/Test periods was similar to individual Train/Test period rankings
- Models trained using historic claims data generally performed worse than the Baseline model (which assumed no seasonality at all!)

Thank you

Contact Details: jennifer.loftus@acornlife.ie