

Forecasting Analytics

Forecasting Unemployment rates in the UK and EU

Project Submission – Group B9

PGID	Name
61710602	Rajesh Manivannan
61710764	Kartick B Muthiah
61710492	Debayan Das
61710353	Devesh Kumar
61710812	Chirag Bhardwaj
61710646	Sreeharsha Konga

Executive Summary

➤ Problem Description:

Unemployment benefits are a huge cost to governments and are highly dependent on projected unemployment rate for the country. Unemployment benefits are social welfare payments to unemployed individuals. The definition of unemployed individual varies from among different governments. Each year Governments allocate a certain percentage of their financial outlay to meet these requirements.

➤ Forecasting Objective:

Forecasting unemployment rates is critical for the budgetary allocations. We forecast monthly unemployment rates for 5 different countries that have income support as part of their social welfare scheme. We choose monthly unemployment rates as generally these payments are monthly.

➤ Client Information:

Our Clients are Ministries of Finance of European and UK governments who have to budget for unemployment benefits as part of their social welfare schemes.

➤ Data Description:

- **Source:** Federal reserve of Economic Data (<https://fred.stlouisfed.org/>)
- **Key Characteristics:** Trend, Level and noise observed for all the data.
- **Countries to be analyzed:** Austria, UK, Ireland, Germany

➤ Forecast period:

The forecast period is chosen as 3 months + 12 months as the client expects us to forecast the results 3 months prior to the start of the Fiscal year.

- Forecasting Horizon -> 15 months
- Seasonality -> 12 months

➤ Final Forecasting Method:

Most of our time series had trend and Seasonality. Hence, we used Holt Winters, MLR and MLR + ARIMA models for forecasting the time series.

➤ Recommendations:

- For most countries we tried many methods of which MLR with lag 1 was highly accurate with MAPE of under 2%
- However, MLR is difficult and more costly to implement. So, one can use Holt Winters to forecast as well.
- We recommend sensitizing the forecasted values using the confidence interval. This should help the government adjust for buffer allowances
- A sample sensitivity analysis presented in below

UK Sensitivity Analysis	Value	Unit
Total People willing to work	30.00	million
Unemployment lower bound	0.75	million
Unemployment Upper bound	0.80	million
Monthly allowance	226.40	pounds
Lower bound monthly budget allocation	168.96	Million Pounds
Upper bound monthly budget allocation	180.09	Million Pounds

The figures shows a sample sensitivity analysis for the month of January.

➤ **Limitations:**

- No external economic Indicators were used.
- The model is highly dependent on the frequency of collection of unemployment data.
- The MLR model uses Lag-1 as an input variable
 - Lag-1 data will not be available ahead of time
 - Allows us to forecast only one month at a time

Alternatives:

- Use lag-13 if Naïve forecast has reasonable MAPE

Technical Summary

This section presents with a generalized methodology for forecasting carried out for each data set

➤ **Data Preparation:** Below are the steps carried out to prepare the data

- **Outliers:** We used various graphical representations for identifying outliers. A regular time series plot was used to identify outliers
- **Missing Values:** No missing values were found in the monthly data we collected. However, for missing values we can replace mean or use lag methods

➤ **Identifying Forecasting Methods:**

For identifying forecasting methods we used graphical representations to check the characteristics of the time series. Below are the details of some graphical representations used for checking characteristics of the time series:

1. **For Trend and Level:** Regular time-series plot with a trend line
2. **For Seasonality:**
 - Yearly: X axis: Years; Y axis: Unemployment Rate
 - Monthly: X Axis: Months; Y axis: Unemployment Rate; Plot each year as a line

Once the characteristics are identified the following methods:

- Trend Only: Exponential and Double Exponential
- Trend & Seasonality: Holt Winters (Multiplicative & Additive), MLR

For MLR we have also used ARIMA to check any lag element in the residuals. Using the ARIMA model we improve our MLR model to arrive at the forecast.

➤ **Data Partitioning:**

For forecasting, it is assumed that the forecast is required by the client 3 months prior to the Fiscal year. This buffer time is assumed because the ministry requires time to incorporate the teams' findings into their budget. Based on this business requirement we have partitioned all the data as follows:

1. Training Period: Entire data set excluding the last 15 months.
2. Validation period: 15 months.

➤ **Holt Winters/Exponential/Double Exponential:**

For these methods, XL miner directly optimizes for alpha (exponential/Double Exp/Holt Winters), beta (Double Exp/Holt Winters) and gamma(Holt Winters).

➤ **Multiple Linear Regression(MLR)**

For running MLR we followed the steps below:

- Dummy Variables: We chose dummy variables based the seasonality
- Time Index: This is chosen based fitted trendline
- Lag Variables: Based on ACF plot

➤ **Choosing the best model:**

Once we run these models, the MAPE is calculated for all the models and the best model is selected based on the model that produces lowest MAPE. Once we arrive at the best model, we again calculate MAPE using this model on two different validation periods to check the robustness of the model.

➤ **Choosing the best model:**

Once the best model is chosen, we forecast unemployment rate for the next 15 months.

➤ **Confidence Interval:**

We also present a range for each forecasted values using the confidence interval of the residuals. We take the 5th and the 95th percentile and use it to predict the confidence interval of each forecasted value.

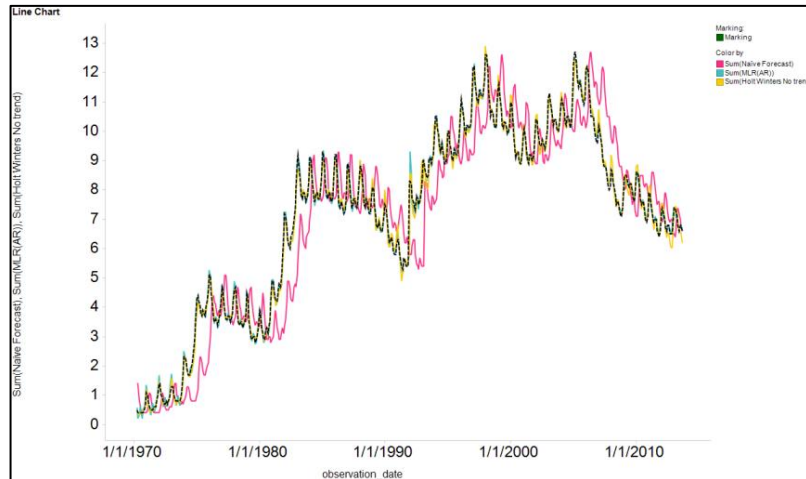
Country Wise Forecasts:

	UK - Log MLR + ARIMA	Austria - MLR + Lag 1	Germany - MLR + ARIMA	Poland - MLR + Lag 1	Netherlands - MLR + Lag 1	Luxembourg - MLR + Lag 1
Input variables	Time	Time	8 Monthly dummy variables	Time	Time	Time
	Time^2	Time^2	lag 2	Lag -1	Lag - 1	Lag -1
	Time^3	11 Monthly dummy variables (excluding September)	8 Monthly dummy variables	11 Monthly dummy variables		11 Monthly dummy variables
	Time^4	Lag -1	lag3, 4, 6, 9, 11, 12			
	11 Monthly dummy variables (excluding September)					
	Lag -1					
	Lag -2					
Output Variables	Log(Unemployment Rate)	Unemployment Rate	Unemployment Rate	Unemployment Rate	Unemployment Rate	Unemployment Rate

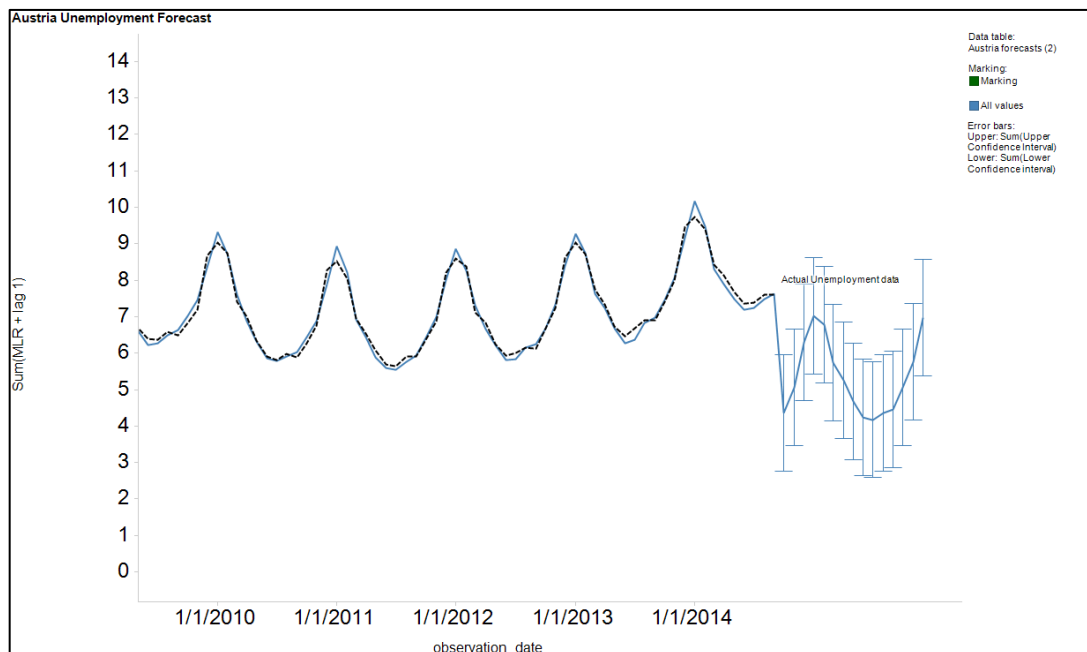
Austria:

We carried out multiple forecasting methods to calculate the **MAPE for validation data**. Below is the summary of the MAPE for various methods. The graph shows Naïve, MLR(AR), Holt winters(No Trend). The dotted line is the actual data.

Method	MAPE
Naïve	7.0%
Exponential smoothing	17.3%
Double exponential	47.2%
MA4	12.7%
HW Mul	5.1%
HW Add	9.2%
HW no trend	9.0%
MLR	27.3%
MLR + lag 1	2.0%



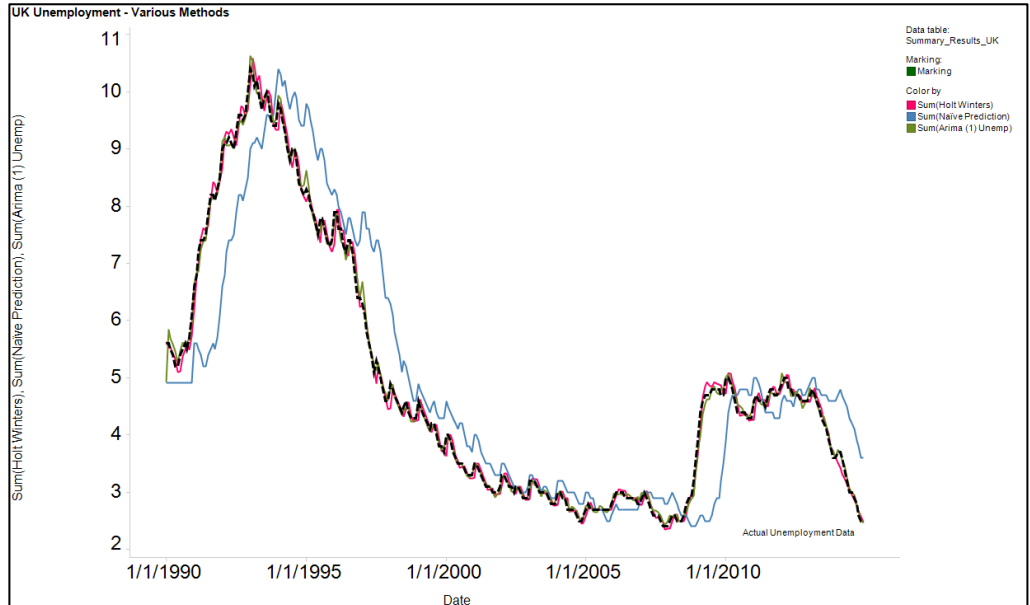
We used the MLR + lag1 model to Forecast the unemployment data. Below is the plot of the forecast.



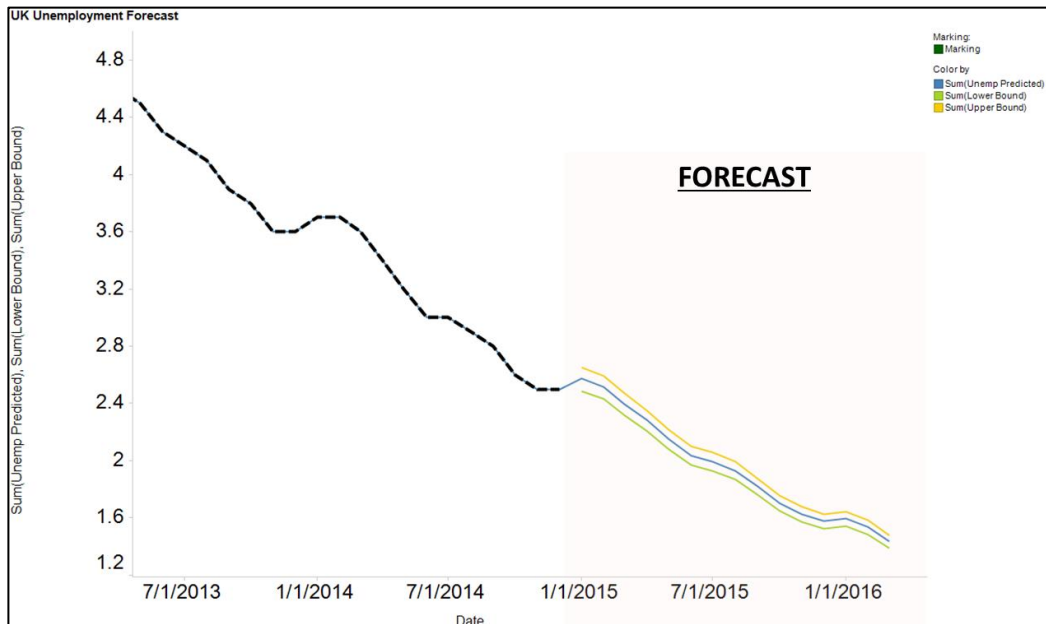
United Kingdom:

Below is the summary of the MAPE for various methods. The graph shows Naïve, MLR(AR), Holt winters. The dotted line is the actual data.

Methods	MAPE
Naïve	35.9%
Exponential	24.7%
Double Exponential	18.8%
Moving Average (W=2)	27.9%
Moving Average (W=3)	30.0%
Holt-Winters Multiplicative	2.8%
Holt-Winters Additive	2.9%
MLR(Normal)	7.1%
MLR+ARIMA(1)	1.9%
MLR+ARIMA(2)	1.93%
LogMLR(Normal)	31%
LogMLR + ARIMA(1)	1.47%
LogMLR + ARIMA(2)	1.62%



We used the Log MLR + ARIMA(1) model to Forecast the unemployment data. Below is the plot of the forecast.

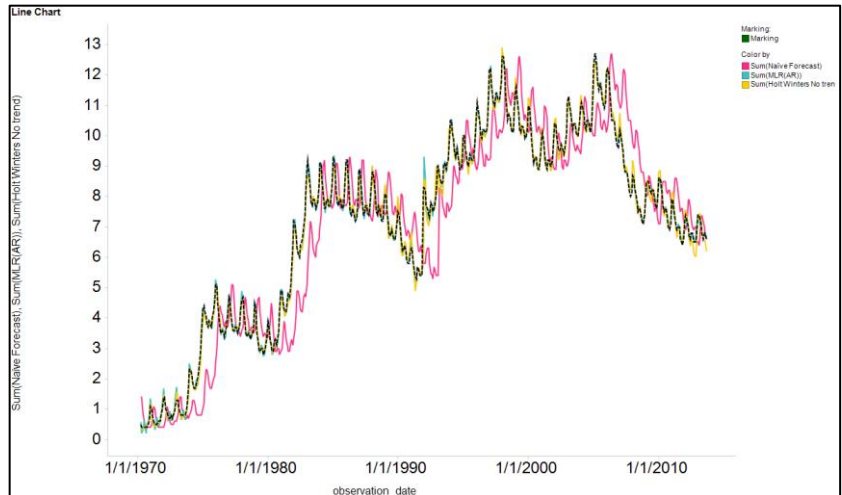


APPENDIX

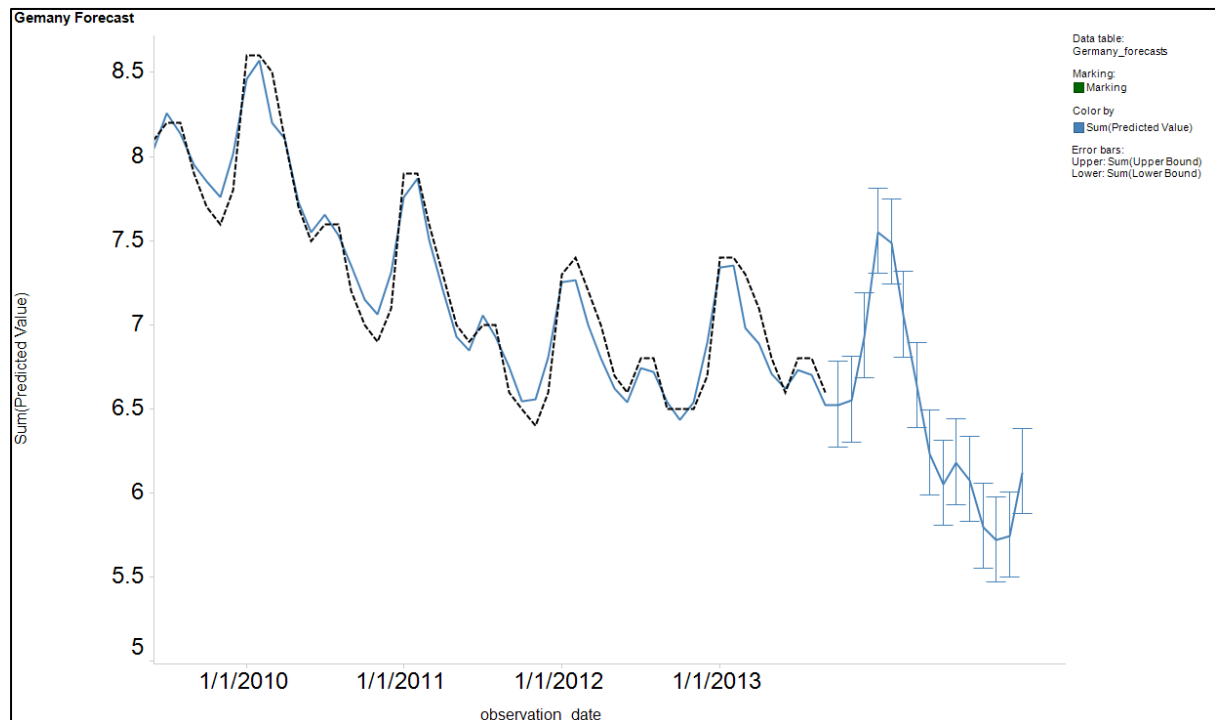
Germany:

We carried out multiple forecasting methods to calculate the **MAPE for validation data**. Below is the summary of the MAPE for various methods. The graph shows Naïve, MLR(AR), Holt winters(No Trend). The dotted line is the actual data.

Method	MAPE
Naïve	6.2%
Exponential	3.9%
Double Exponential	8.0%
Moving Average	3.7%
Holt Winters Mult	5.6%
Holt Winters Add	5.8%
Holt Winters No trend	3.7%
MLR	5.1%
MLR + ARIMA	1.4%
MLR(LN)	21.1%
MLR(sqrt)	41.4%
MLR(inverse)	293.5%

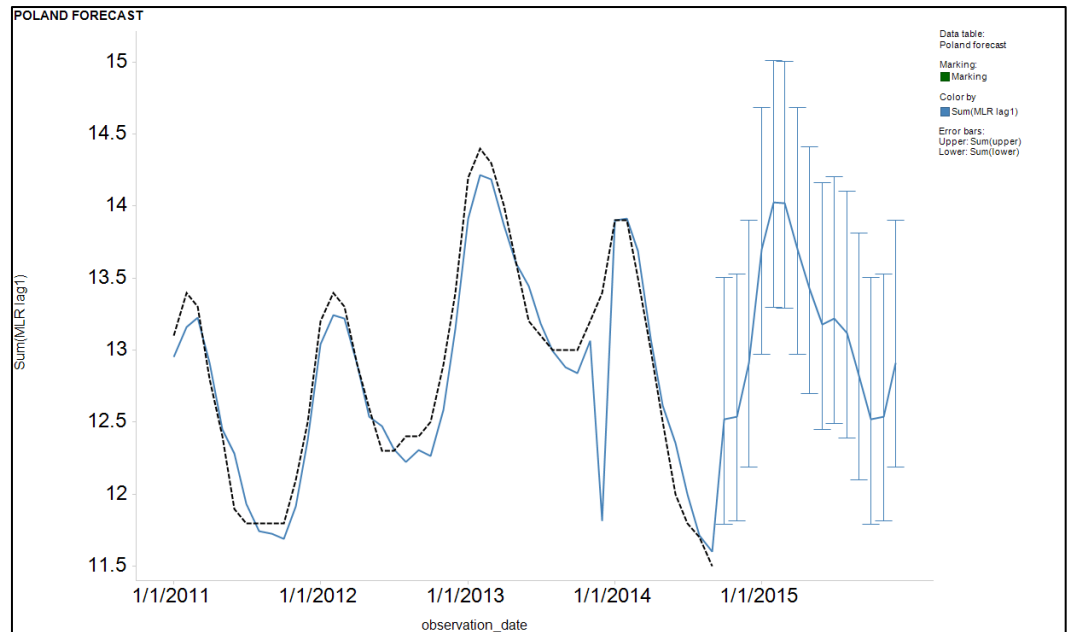


We used the MLR + ARIMA model to Forecast the unemployment data. Below is the plot of the forecast.



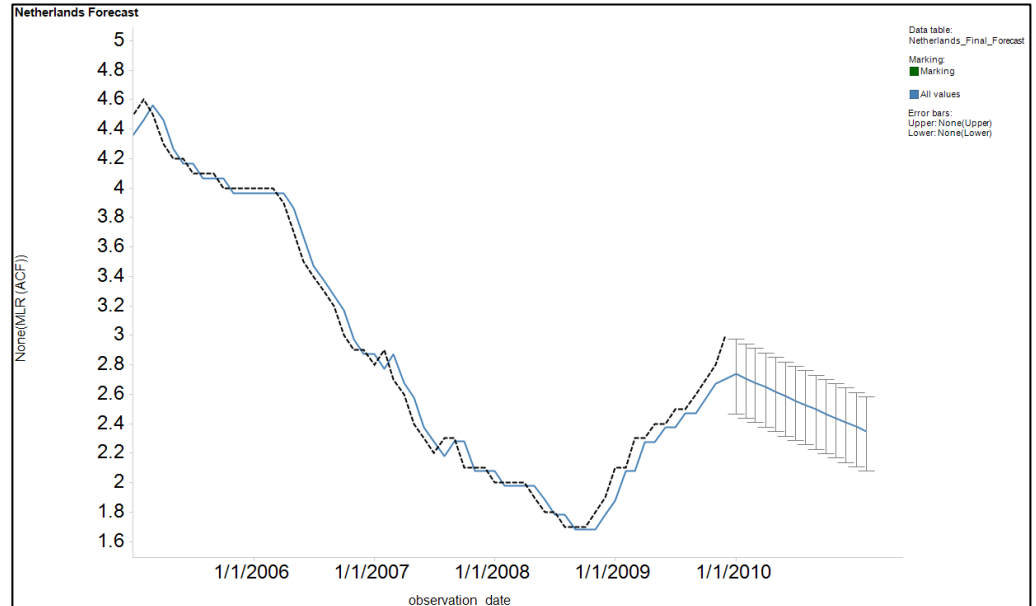
POLAND:

Method	MAPE
Naïve	6.3%
HW no trend	5.3%
HW Add	27.5%
HW Mult	8.7%
MLR	14.98%
MLR lag1	1.62%



NETHERLANDS:

Methods	MAPE
Naïve	18.7%
Exponential	25.5%
Double Exponential	41.7%
Moving Average	25.5%
Holt-Winters Multiplicative	41.21%
Holt-Winters Additive	46.76%
Holt-Winters No Trend	27.75%
MLR	18.79%
MLR (lag1)	4.72%



Luxembourg:

Method	MAPE
Naïve	14.6%
Holt Multiplicative (monthly seasonality)	4.8%
Holt Additive (monthly seasonality)	4.8%
Holt Additive(annual seasonality)	2.3%
MLR	20.29%
MLR+ARIMA(1)	1.94%

