

```
-----  
-- Time Series Analysis and Prediction  
-- Author: Group 2  
-----
```

```
-----  
-- Before building time series analysis let us check max and avg temperature from  
Arizona state 1950 onwards.  
-----
```

```
select max_temp_c, avg_temp_c, record_date  
from `data-225-group-project.climate_dwh.climate_fact` cl_fact inner join  
`data-225-group-project.climate_dwh.location_dim` l_dim on  
cl_fact.stationid=l_dim.station_id  
where state='Arizona' and extract(year from record_date)>=1950  
order by 3;
```

```
-----  
-- Time Series Model: ARIMA model building  
-----
```

```
CREATE OR REPLACE MODEL `data-225-group-  
project.climate_analytics_data.climate_arima_model`  
OPTIONS  
  (model_type = 'ARIMA_PLUS',  
   time_series_timestamp_col = 'record_date',  
   time_series_data_col = 'max_temp_c',  
   auto_arima = TRUE,  
   data_frequency = 'AUTO_FREQUENCY',  
   decompose_time_series = TRUE  
  ) AS  
select record_date, max_temp_c  
  from `data-225-group-project.climate_dwh.climate_fact` cl_fact inner join  
`data-225-group-project.climate_dwh.location_dim` l_dim on  
cl_fact.stationid=l_dim.station_id  
where state='Arizona' and extract(year from record_date)>=1950;
```

```
-----  
-- Evaluation Metrics  
-----
```

```
SELECT  
  *  
FROM  
  ML.ARIMA_EVALUATE(MODEL `data-225-group-  
project.climate_analytics_data.climate_arima_model`);  
-----
```

-- The seasonal_periods column is about the seasonal pattern inside the input time series. The has_holiday_effect, has_spikes_and_dips, and has_step_changes columns are only populated when decompose_time_series=TRUE. They are about the holiday effect, spikes and dips, and step changes inside the input time series, which are not related to the ARIMA modeling. Therefore they are all the same across all output rows [Reference: <https://cloud.google.com/bigquery/docs/arma-single-time-series-forecasting-tutorial>]

-- Coefficient of the Model

```
SELECT
*
FROM
ML.ARIMA_COEFFICIENTS(MODEL `data-225-group-
project.climate_analytics_data.climate_arima_model`);
```

-- The results include the following
columns:ar_coefficients,ma_coefficients,intercept_or_drift

-- Forecasting using ARIMA model

```
SELECT
*
FROM
ML.FORECAST(MODEL `data-225-group-
project.climate_analytics_data.climate_arima_model`,
STRUCT(30 AS horizon, 0.8 AS confidence_level));
```

-- Here STRUCT(30 AS horizon, 0.8 AS confidence_level) clause indicates that the query forecasts 30 future time points from current date present in the dataset, and generates a prediction interval with a 80% confidence level.