**Introduction to Time Series Analysis**

**(Assignment for IITP –BSE Course)**

**1. Make separate python file (.ipynb) for each dataset and read the datasets in each file using pandas**

**and store in dataframe, print first 5 rows of data including the header row.**

| **Monthly Car Sales**  **Month** | **Sales** |
| --- | --- |
| **0** | 1960-01 | 6550 |
| **1** | 1960-02 | 8728 |
| **2** | 1960-03 | 12026 |
| **3** | 1960-04 | 14395 |
| **4** | 1960-05 | 14587 |

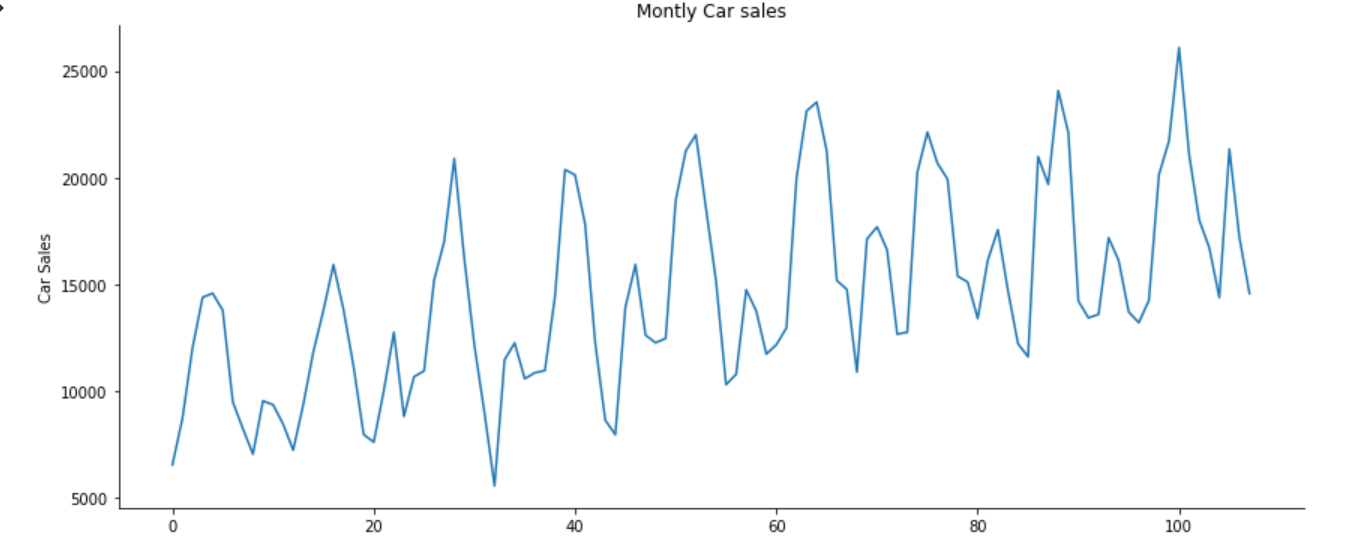
| **Monthly sunspots**  **Month** | **Sunspots** |
| --- | --- |
| **0** | 1749-01 | 58.0 |
| **1** | 1749-02 | 62.6 |
| **2** | 1749-03 | 70.0 |
| **3** | 1749-04 | 55.7 |
| **4** | 1749-05 | 85.0 |

**Daily Female births**

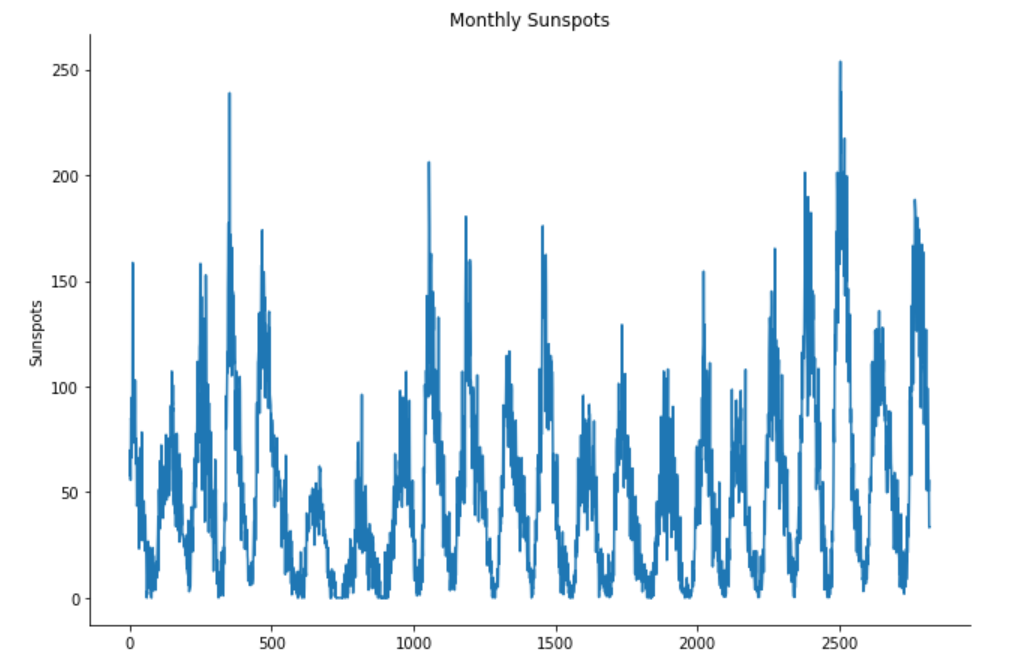
| **Date** | **Births** |
| --- | --- |
| **0** | 1959-01-01 | 35 |
| **1** | 1959-01-02 | 32 |
| **2** | 1959-01-03 | 30 |
| **3** | 1959-01-04 | 31 |
| **4** | 1959-01-05 | 44 |

**2. Plot all three dataset to analyse the series.**

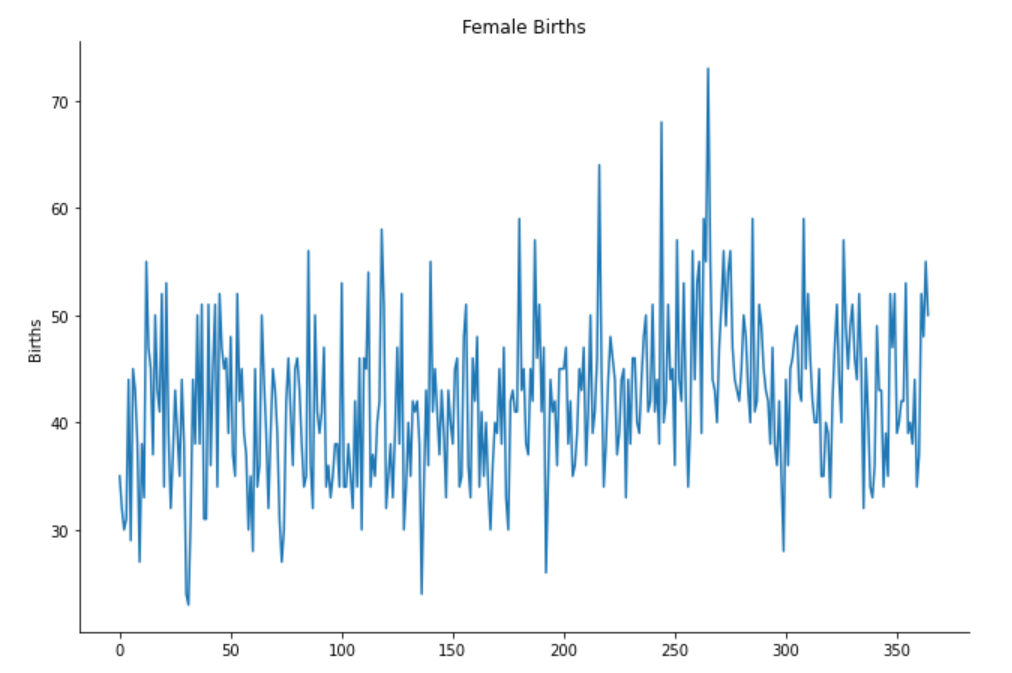
**Monthly Car Sales**



**Monthly sunspots**



**Daily Female births**



**3. Set “Date” column of the dataset as index. Please take the correct name of the column.**

**Monthly Car Sales**

DatetimeIndex(['1960-01-01', '1960-02-01', '1960-03-01', '1960-04-01',

'1960-05-01', '1960-06-01', '1960-07-01', '1960-08-01',

'1960-09-01', '1960-10-01',

...

'1968-03-01', '1968-04-01', '1968-05-01', '1968-06-01',

'1968-07-01', '1968-08-01', '1968-09-01', '1968-10-01',

'1968-11-01', '1968-12-01'],

dtype='datetime64[ns]', name='Month', length=108, freq=None)

**Monthly sunspots**

DatetimeIndex(['1749-01-01', '1749-02-01', '1749-03-01', '1749-04-01',

'1749-05-01', '1749-06-01', '1749-07-01', '1749-08-01',

'1749-09-01', '1749-10-01',

...

'1983-03-01', '1983-04-01', '1983-05-01', '1983-06-01',

'1983-07-01', '1983-08-01', '1983-09-01', '1983-10-01',

'1983-11-01', '1983-12-01'],

dtype='datetime64[ns]', name='Month', length=2820, freq=None)

**Daily Female births**

DatetimeIndex(['1959-01-01', '1959-01-02', '1959-01-03', '1959-01-04',

'1959-01-05', '1959-01-06', '1959-01-07', '1959-01-08',

'1959-01-09', '1959-01-10',

...

'1959-12-22', '1959-12-23', '1959-12-24', '1959-12-25',

'1959-12-26', '1959-12-27', '1959-12-28', '1959-12-29',

'1959-12-30', '1959-12-31'],

dtype='datetime64[ns]', name='Date', length=365, freq=None)

**4. Try fetching the data using index i.e. for a specific date, specific month and specific year.**

**Monthly Car Sales**

Sales 7049

Name: 1960-09-01 00:00:00, dtype: int64

**Monthly sunspots**

Sunspots 66.3

Name: 1749-08-01 00:00:00, dtype: float64

**Daily Female births**

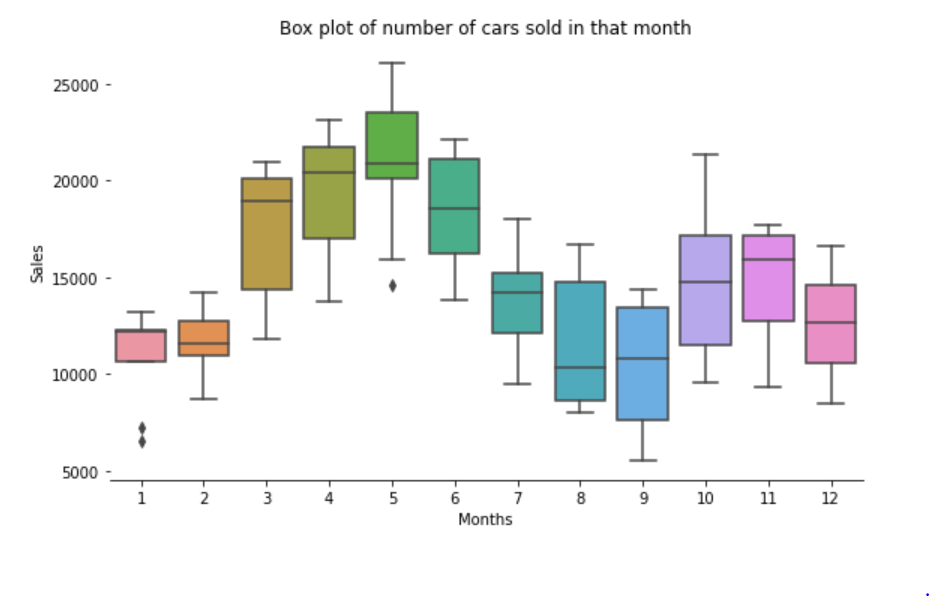
Births 35

Name: 1959-01-01 00:00:00, dtype: int64

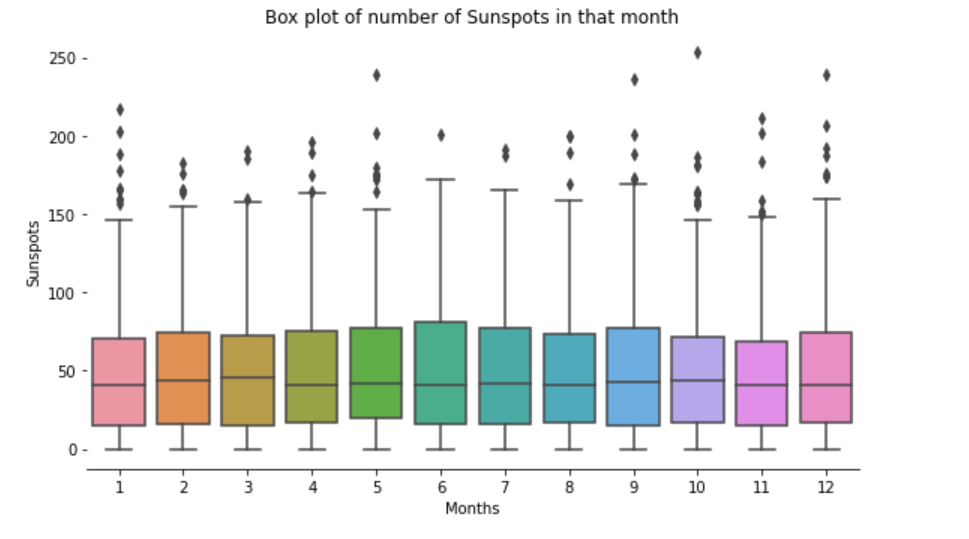
**5. Plot the boxplot using daily data for daily-total-female-births dataset and monthly data for**

**monthly-car-sales.csv, monthly-sunspots.csv, daily-total-female-births datasets.**

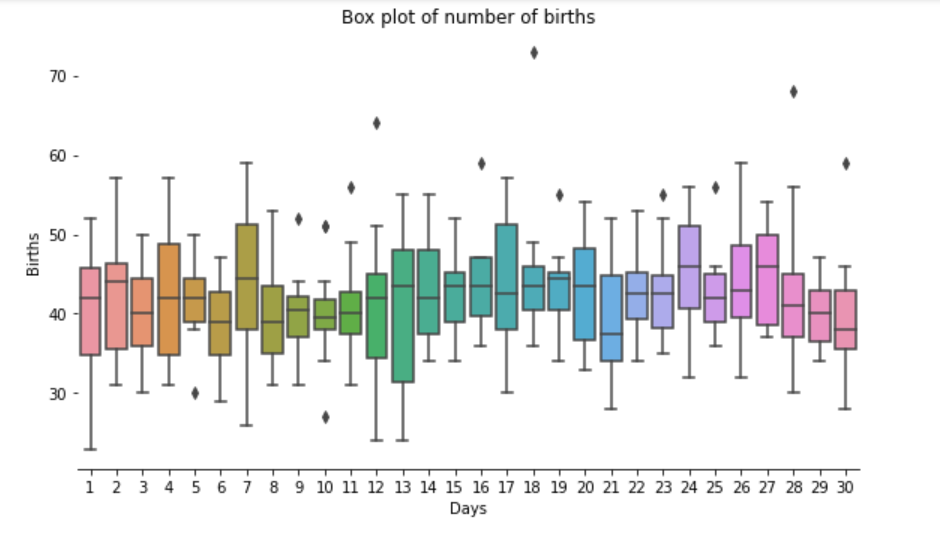
**Monthly Car Sales**



**Monthly sunspots**

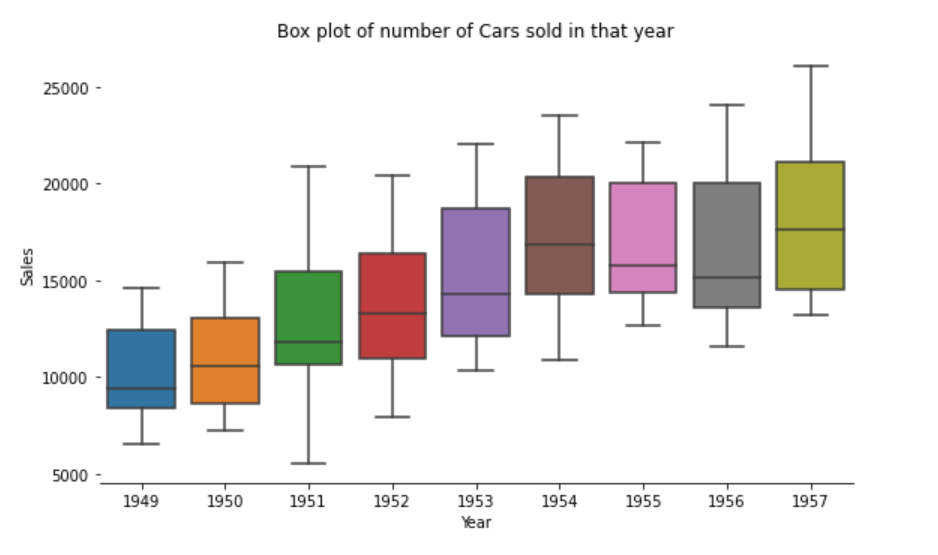


**Daily Female births**

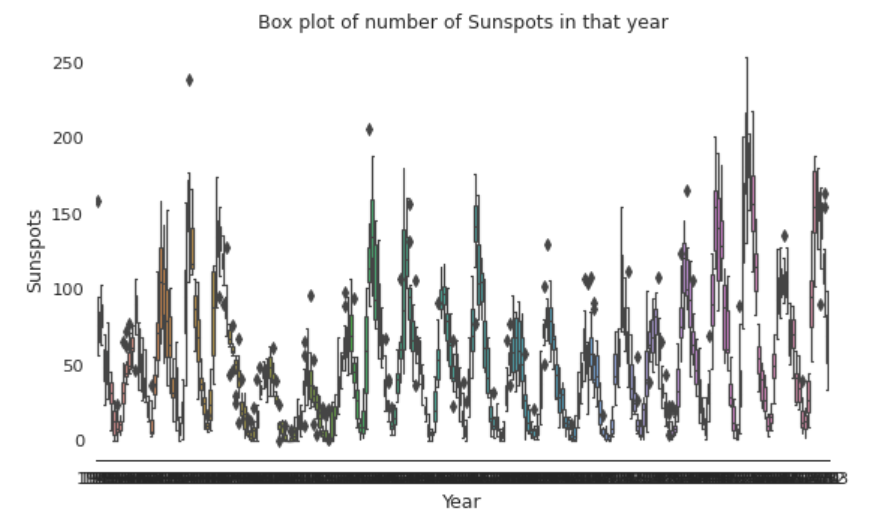


**6. Plot the boxplot using yearly data for (i) and (ii).**

**Monthly Car Sales**

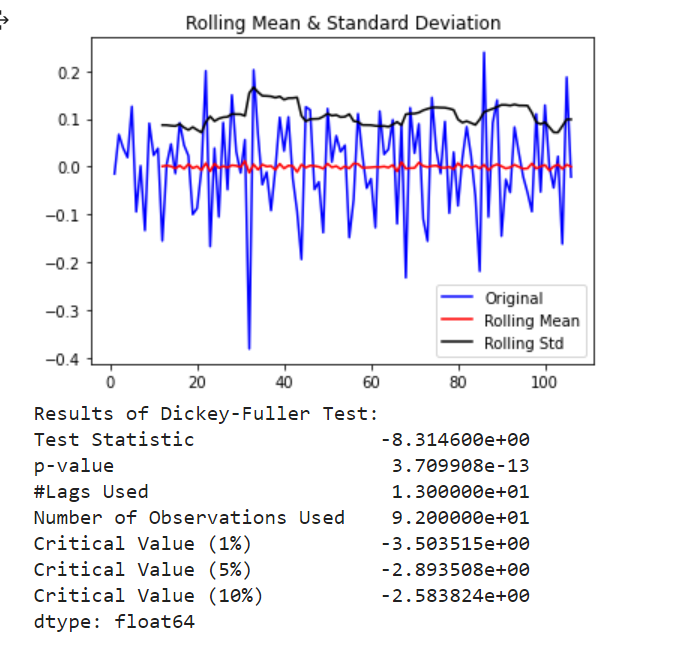


**Monthly sunspots**

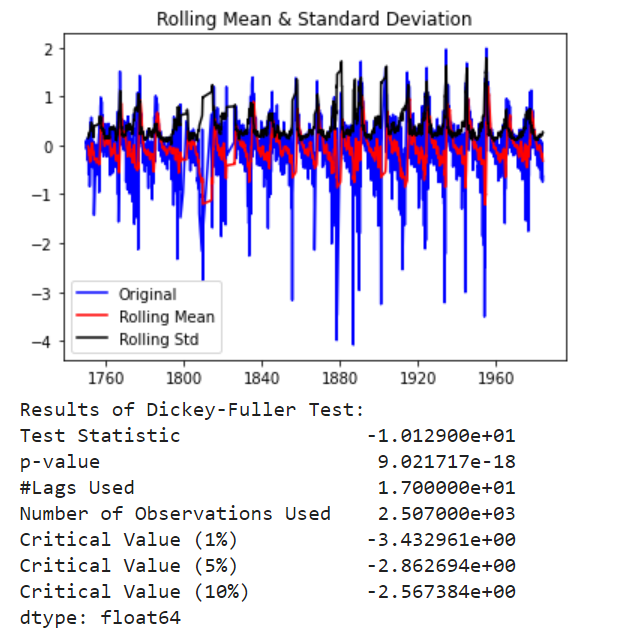


**7. Calculate moving average of the data and test stationarity (rolling mean, standrad deviation)**

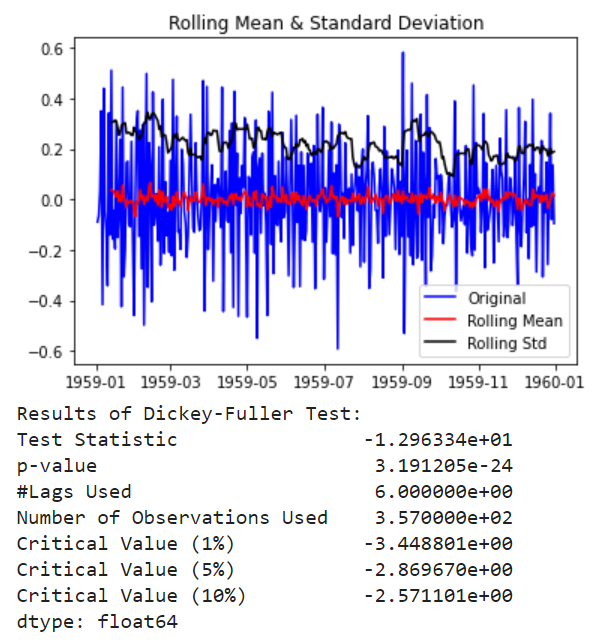
**Monthly Car Sales**



**Monthly sunspots**



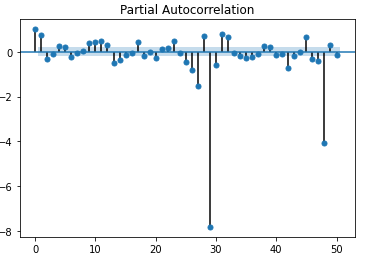
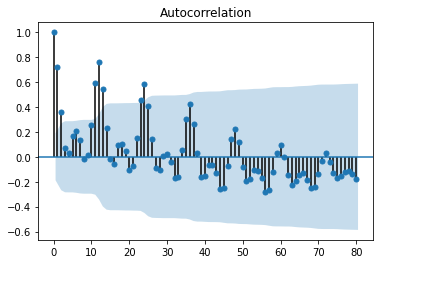
**Daily Female births**



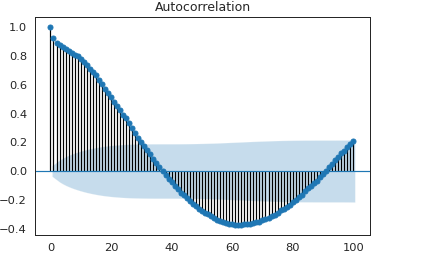
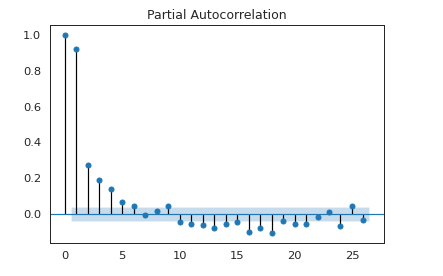
**8. Plot autocorrelation and partial autocorrelation and interpret the results of ACF and PACF plots**

**to find the order of AR, MA, ARMA, ARIMA models**

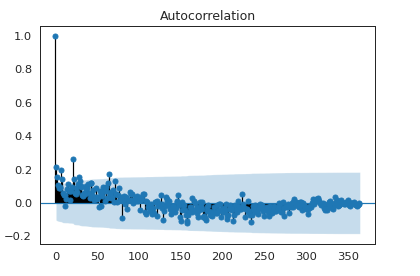
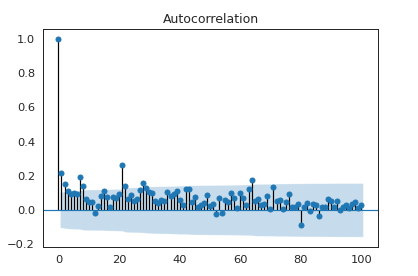
**Monthly Car Sales**



**Monthly sunspots**

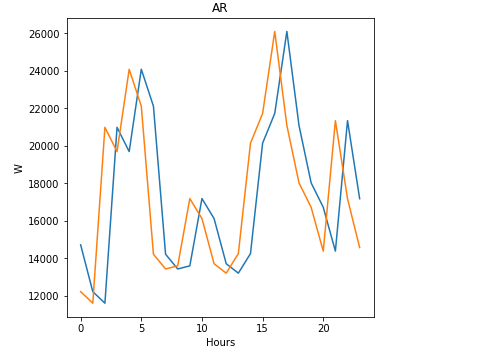
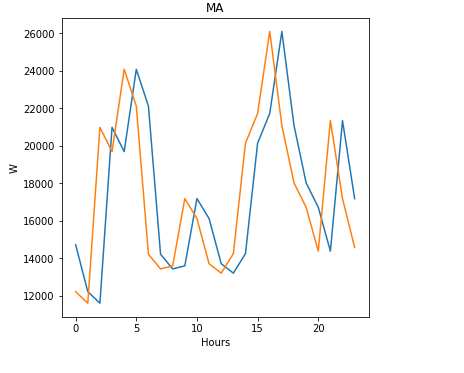
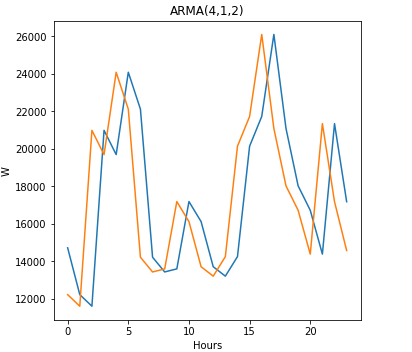
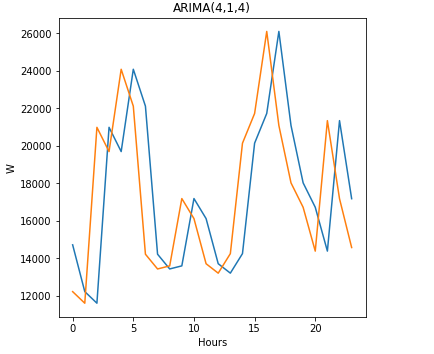
 

**Daily Female births**

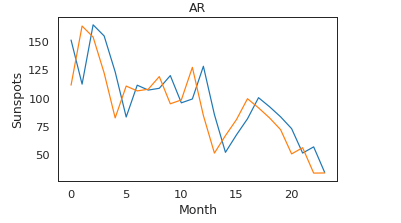
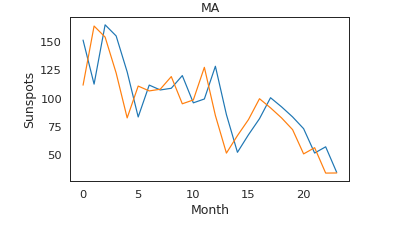
 

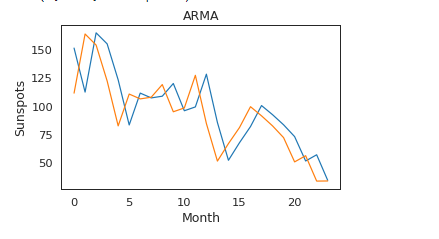
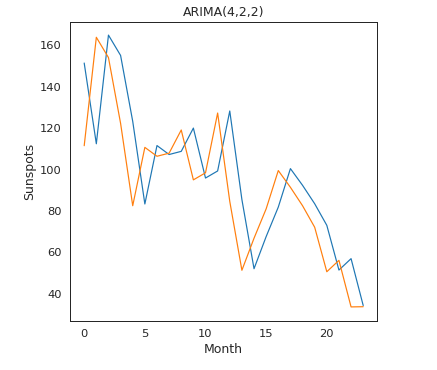
**11. Plot the original versus predicted graph for each of the model prediction.**

**Monthly Car Sales**

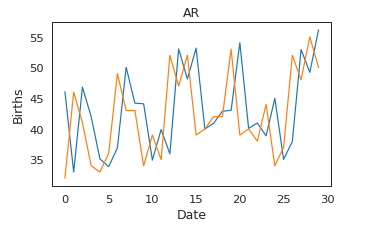
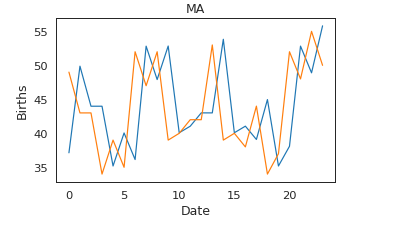
   

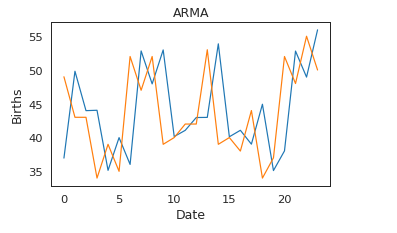
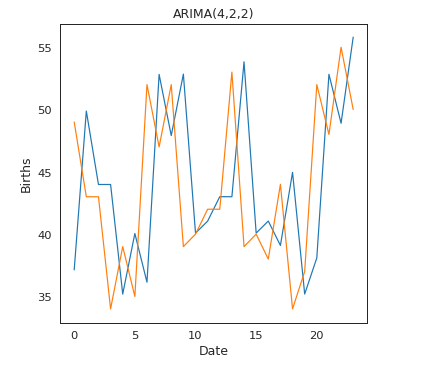
**Monthly sunspots**

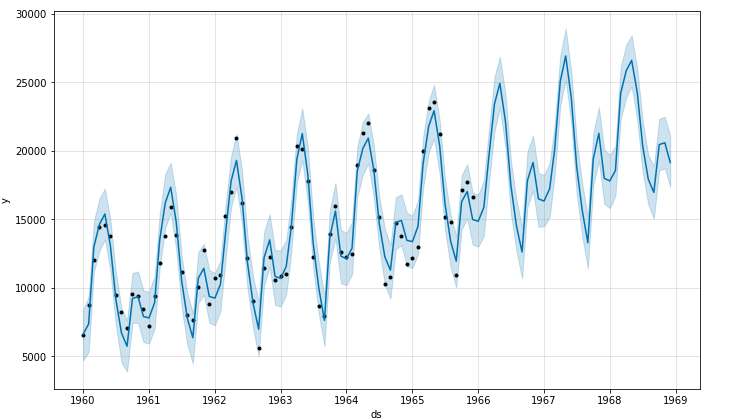
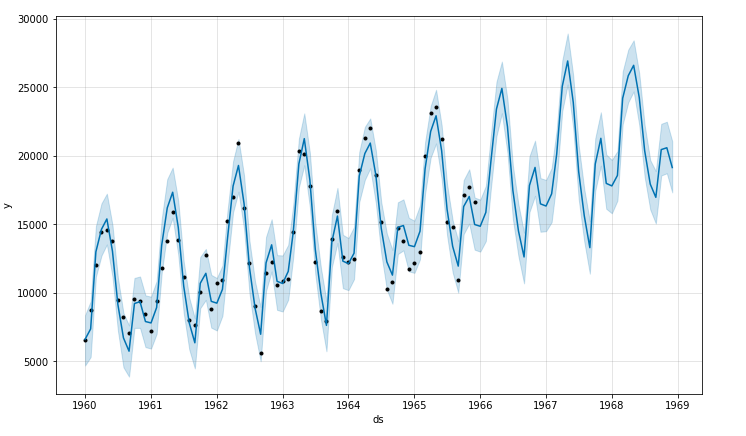
**Daily Female births**

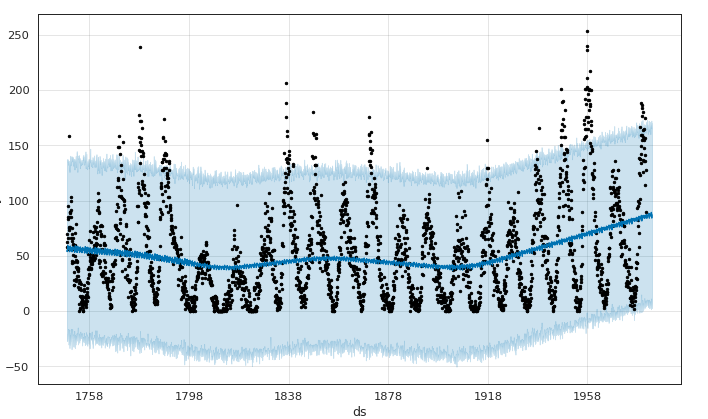
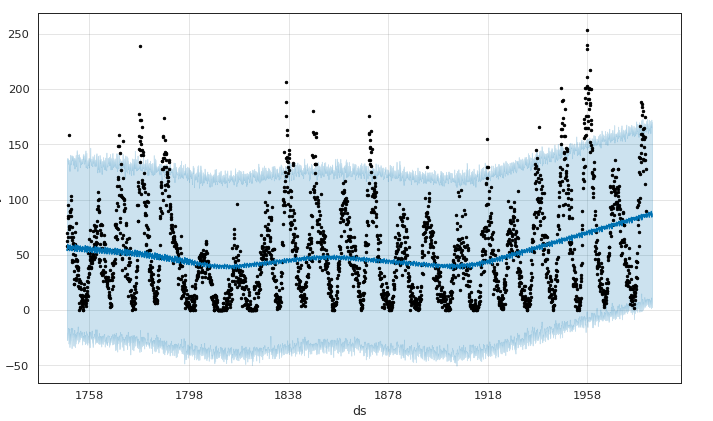
 

**12. Apply Prophet model, calculate the error matrices & plot the original versus predicted.**

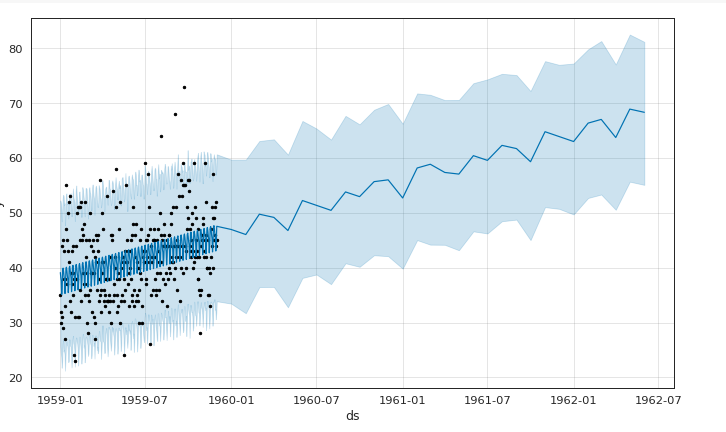
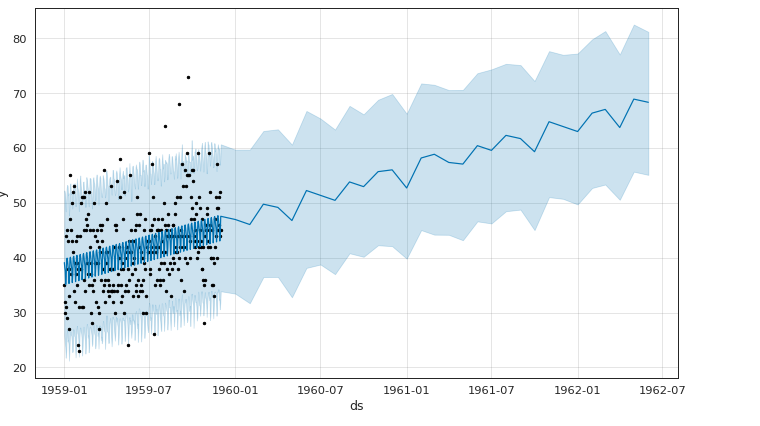
**Monthly Car Sales**

**Monthly sunspots**

**Daily Female births**

**14. Calculate RMSE & MAPE matrices on the prediction of each model and compare all the results**

**in tabular format.**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Monthly Car Sales** | | **Monthly Sunspots** | | **Daily Female births** | |
|  | **RSME** | **MAPE** | **RSME** | **MAPE** | **RSME** | **MAPE** |
| **AR** | 3947.22 | 17.32 | 24.64 | 24.1 | 8.372 | 16.24 |
| **MA** | 3947.22 | 17.32 | 24.63 | 24.1 | 8.372 | 16.24 |
| **ARMA** | 3947.2 | 17.32 | 24.63 | 24.1 | 8.177 | 16.24 |
| **ARIMA** | 3947.14 | 17.32 | 24.64 | 24.1 | 8.137 | 16.24 |
| **ARIMA (min AIC)** | 3647.57 | 17.32 | 24.63 | 24.11 | 7.685 | 16.24 |
| **PROPHET** | 3108.93 |  | 42.61 |  | 16.813 |  |
| **LSTM(Vanila)** | 4193.01 | 21.5 | 26.78 |  | 9.185 | 17.52 |
| **LSTM(Stack)** | 4279.78 | 22.6 | 27.28 |  | 8.094 | 15.69 |
| **LSTM(Bi directional)** | 4310.82 | 23.9 | 25.68 |  | 9.555 | 18.75 |