**Index**

## Exploring The Problem Statement:

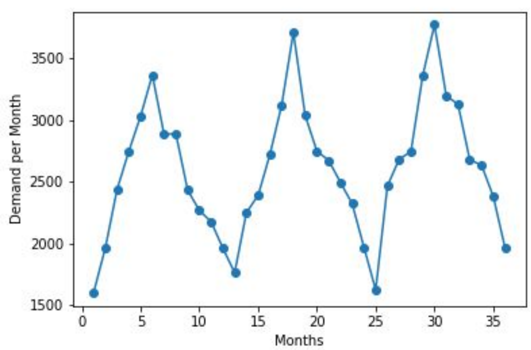
- The Problem Statement is about estimating various factors of ‘Type-A’ medicine out of three given types of medicines: ‘Type-A’, ‘Type-B’ & ‘Type-C’.
- We are currently exploring the scenario of a leading global hospital Thomson & Cook which is trying to manage the stock of medicines taking into consideration all the fluctuating parameters like Flow of Patients, Uncertainty in Demand of Medicines & Lead time to receive medicine from suppliers & simultaneously minimize the cost of managing the inventory.
- We have been asked to simulate the scenario taking into consideration all the parameters & coming up with an effective solution to help the hospital to manage the current situation.
- The Demand we are provided with is based on no. of indoor patients & their severity of condition, consulting doctor treating the admitted patients & right now it has been estimated based on 3 years monthly consumption data.
- The tasks we have been asked to simulate & compute are as follows:
    - Do you find maintaining a constant safety stock to be justifiable? If no, why? State reasons to support your opinion. To make it data analytics driven - what should be your approach?
    - To forecast the demand of medicine Type A for next 4 months.
    - For “Type A” medicine, calculate
        - EOQ
        - The number of orders per year
        - Reorder level
        - The total annual ordering & carrying costs
        - Maximum Inventory level for ‘Type-A’ Medicine for maintaining the Total inventory optimally
- The hospital is planning to follow FMEA (Failure Mode & Effective Analysis) as its risk mitigation plan & in the event of any preferred vendor (effectiveness considering quality, lead time & cost) not able to supply the medicine on time or it expresses its inability to supply for any reason, then another set of listed vendors will be given with the order for a

substitute medicine & these orders will be at a much higher price mostly having a fixed lead time. We have been provided with the corresponding data regarding vendor's assigned for emergency situation.

- We have to work as a high risk taking Hospital Manager who prefers maintaining less amount of safety stock in order to reduce the inventory holding cost & manage the out-of-stock situation by ordering medicines from vendors at higher cost in emergency. We have been asked to simulate the scenarios and investigate/estimate the trade-off between Risk & Money in case of 'Type-A' Medicine.

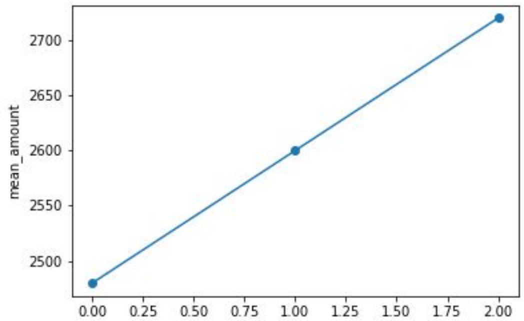
## Studying the Properties of Data Set:

- We have been given the data regarding the Demand of Medicine of 'Type-A' from January 2015 to December 2017 (i.e. total 36 data points are given).
- The following plot represents Demand for 'Type-A' Medicine vs Months which takes January 2015 as '1' and December 2017 as '36'.

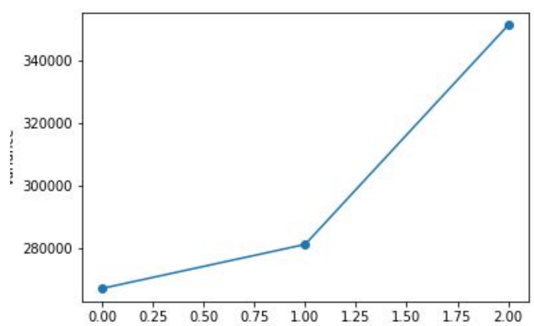



- From the plot we can observe that there is an increase in the Demand in the first-half of the year, & a decrease trend in it as we go further towards December every year.
- This indicates that the data is Seasonal Dependent in an year.
- Therefore, we are assuming that the Annual Demand will follow a Normal Distribution or a Triangular Distribution.
- **Fitting A Normal Distribution:**
    - Here we are trying to fit a Normal Distribution for all the 3 years independently.
    - Normal Distribution requires the Mean & Variance of the data in a year.
    - Formula used for computing Mean & Variance gives a confidence Level ≥ 95% for respective intervals of mean & variance.
    - Mean of all data (Demand per month) in a year = Average Demand in a year.
    - Therefore, mean in a particular year (***μ***) = (Sum of all Demands in a year)/12

- ○ The following plot represents Calculated Mean Demand Value vs Year. The numbers 0,1 and 2 on X-axis represents 2015, 2016 and 2017 respectively.



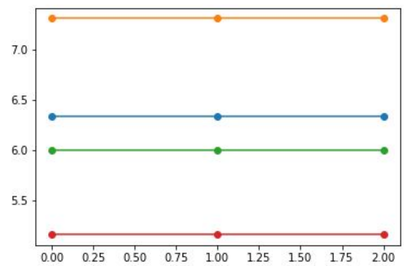

- ○ From the plot we can infer that the mean demands follow a linear polynomial with a slope of 120 units demand per year.
- ○ Mean Demand for the years 2015, 2016 & 2017 are 2480.0 units, 2600.0 units & 2720.0 units respectively.
- ○ Similarly Variance of Demand over a year $= \left(\frac{1}{11}\right) \times \Sigma\,(\text{Demand per Month - Mean for a particular year})^2$
- ○ The following plot represents Calculated Variance of all demand values vs Year. Here again, the numbers 0,1 and 2 on X-axis represents 2015, 2016 and 2017 respectively.




- ○ From the plot, we are unable to observe any known kind of relation between Variance and the years.
- ○ Variance of all Demands for the years 2015, 2016 & 2017 are 267110.1818181819, 281220.54545454547 & 351510.1818181819 respectively.
- ○ Since we did not find any relation between Variance & Years, so we cannot conclude that Demands in a Year are following Normal Distribution.

- **Fitting a Triangular Distribution:**
  - ○ Here we try to fit a Triangular Distribution for all the three years independently.

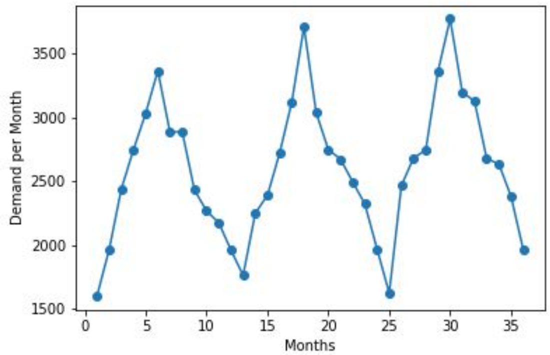
- In the working that follows, variable values of January to December are scaled as '1' to '12'.
- Mean of all months = (A + B + C)/3
- Median of all months = B - $\sqrt{(B-C)*(C-A)*0.5}$
- Mode of all months = C
- Variance of all months = ($A^2$ + $B^2$ + $C^2$ - AB - BC - CA)/18
- 'A' represents the starting Month of the Year ('1' means January).
- 'B' represents the ending Month of the Year ('12' means December).
- 'C' represents the Month which has the Highest Demand in a Year.
- The following graph represents all variables of a triangular distribution vs years. The numbers 0,1 and 2 on X-axis represents 2015, 2016 and 2017 respectively.




- 'Orange colored line' represents Median of All Months in a Year.
- 'Blue colored line' represents Mean of All Months in a Year .
- 'Red colored line' represents Variance of All Months in a Year .
- 'Green colored line' represents Mode of All Months in a Year .
- From the above graph we can conclude that the value of 'C' is constant (= 6) , that means In every Year Maximum Demand for 'Type-A' Medicine occurs in the month 'JUNE' .

## Building the model for dataset:

- **Qualitative analysis of dataset for model selection** - As we can see the dataset has a periodicity of 1 year, a trend of the average values & some error. So if we try to model this using linear regression we won't be able to predict the seasonality and trend described later.




From the data we can see that it represents a combination of trend, seasonality (in this case periodicity), & some random term with some probability distribution. So this data changes with time such that its trend increases over time, its seasonality (periodicity) is fixed for one year & the data is having its random variation about the value predicted by the trend & seasonality. So the best model to fit this type of data set should be having all these properties addressed.

- **Description of the model collection** -
  - Linear regression and polynomial regression - Both the models mentioned here are for general prediction of trend only. We have a univariate dataset so we cannot play with different combinations of variables (features). So linear regression will not be useful at all. If we put higher order terms in polynomial it will surely overfit the dataset & is supposed to have huge validation set error. So the forecast will not be effective. Moreover, the data set is very small so these models cannot be trained efficiently.
  - Neural Networks - Neural network is more of a linear regression on multiple levels & is ideally suited for classification problem. If we remove the activation of the last layer then it can be used to predict regression models also. But it will have the same problem of predicting the trend only & will also probably overfit. We can use advanced NN models but they will be computationally expensive & cannot fit our time frame.
  - ARIMA (AutoRegressive Integrated Moving Average) - This model best fits our requirements. It has been described below.
- **Explanation of model working** - An autoregressive model of order p can be written as -

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t$$

Here we are writing lag terms for ‘p’ previous time instances. Epsilon(t) is the error term which may be randomly distributed about according to a probability distribution. Similarly moving average models are ones based not on previous year values but on errors and can be related as

$$y_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}$$

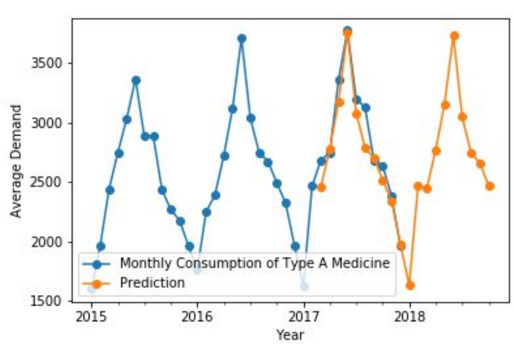
In this case we are using a model which combines both above models as ARIMA model. We have used autoArima from python library statsmodels whose workflow can be shown as

- ○ Plot the data to identify unusual patterns. Then identify unusual observations in the data and try to understand the pattern.
- ○ Use Box-cox transformation to stabilise variance if needed which was not required in our case.
- ○ Plot autocorrelation function and partial autocorrelation function plots of the residuals and check them. If the residuals look like white noise the model should be considered.

- **Conclusion** - The autoArima model fits the dataset almost very closely. The technical details of our implementation are described below. The autoArima used selected ARIMA(0,1,1) which represents "without constant - simple exponential smoothing model" for prediction. Here using the prediction, the trend was increasing, period was of 12 months (which was fixed) and the fitting error was 176.71 units which is acceptable according to the library (discussed later) standards. The forecasted values for next months are in the table below

| Month | Forecast |
|---|---|
| January 2018 | 1630.159 |
| February 2018 | 2467.063 |
| March 2018 | 2449.086 |
| April 2018 | 2767.892 |

We also know from the dataset and the predicted values that monthly maxima occurs in June as per periodicity. Finally a graphical representation of predicted values from March 2017 to October 2018 has been shown below.



## Varying Safety Stock:

- No, maintaining a constant safety stock is not justifiable since safety stock is an additional quantity of an item held in the inventory to reduce the risk that the item will be out of stock.
  So if we keep our safety stock constant (with variable demand), there might be chances of stock-out and we will not be able to meet the customers demand. If we keep our safety stock too low, it might be possible that for some months that the demand is greater than average safety stock leading to a stockout scenario. At the same time , if we keep safety stock too high say to fulfill the peak customer demand, then it might be possible that for the most part of the time period we need less safety stock than the peak value of the demand and rest of our safety stock is left unused leading us to spend a lot on inventory and holding cost for the safety stock.So we can reduce the cost on safety stock by keeping a lower safety stock.

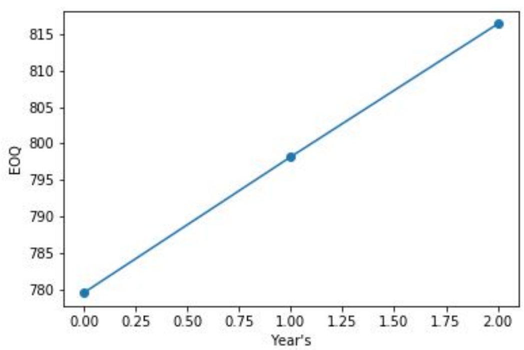
- So we can keep our safety stock proportional to the customer demand over two orders. Suppose we have high demand during two orders (say in our case the demand during June is higher than other demands). So during that period, we can keep more safety stock as compared to others where keeping high safety stock would be a waste of money.

## Calculating Economic Order Quantity:

- Economic Order Quantity (EOQ) is a decision tool used in cost accounting. It's a formula that allows you to calculate the ideal inventory to order for a given product. This calculation allows to minimize the Ordering and Carrying Cost.
- EOQ Formula is based on the following assumptions:
  - **Reorder point:** Reorder point is the time at which next order should be placed. EOQ assumes that we order same quantity at each reorder point.
  - **Demand, relevant ordering cost, & relevant carrying cost:** Customer demand for the product is known. Also, the ordering & carrying costs are certain. A relevant cost refers to a cost you need to consider when you make a decision.
  - **Purchase order lead time:** The lead time is the time period from placing the order to order delivery. EOQ assumes that the lead time is known.
  - **Purchasing cost per unit:** The cost per unit doesn't change with the amount ordered. This removes any consideration of quantity discounts. Assume you'll pay the same amount per unit, regardless of the order size.

- ○ **Stockouts:** No stockouts occur. You maintain enough inventory to avoid a stockout cost. That means you monitor your customer demand & inventory levels carefully.
- ○ **Quality costs:** EOQ generally ignores quality costs.

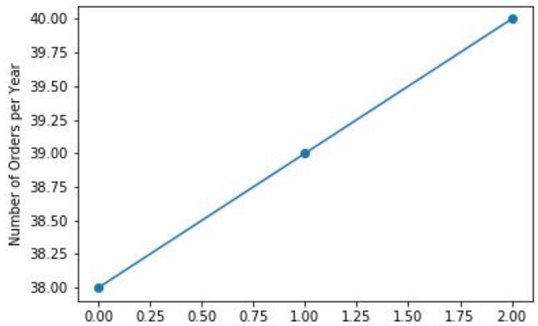
- Formulated, EOQ = $\sqrt{(2 \times D \times C) \div H}$
  - ‘D’ represents the Annual Demand in units of the medicine.
  - ‘C’ represents the Total Ordering cost per unit of medicine.
    - Total Ordering Cost per Unit = Cost of a Unit + Purchasing Cost per Unit.
  - ‘H’ represents the Holding cost of the Inventory.
- The following graph represents EOQ vs Year , Y-Axis represents the EOQ value. Here again, the numbers 0,1 and 2 on X-axis represents 2015, 2016 and 2017 respectively.




- Values of EOQ per Year for 2015, 2016 & 2017 are 779.5506397919253, 798.1879477917466 & 816.3999020088133 respectively.
- As per the calculation, EOQ can be Floating Integer, but ordered quantity of the units cannot be a Floating Integer. So, the Ordered quantity per Year for 2015, 2016 & 2017 are 779, 798 & 816 respectively.

## Calculating Number of Orders Per Year:

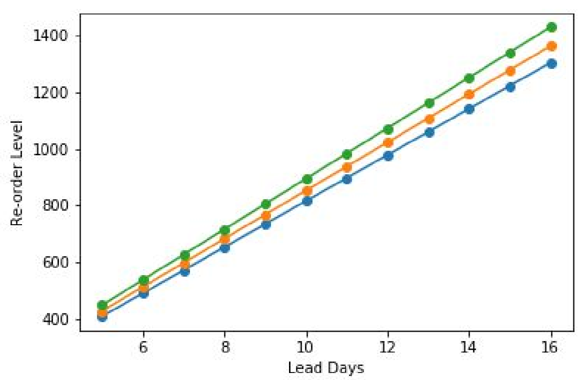
- Number of Orders Per Year = D $\div EOQ$
- Number of Orders Per Year should be Integer.
- The following plot is Number of Orders Per Year vs Year. The numbers 0,1 and 2 on X-axis represents 2015, 2016 and 2017 respectively.




- Number of Orders Per Year for 2015, 2016 & 2017 are 38, 39 & 40 respectively.

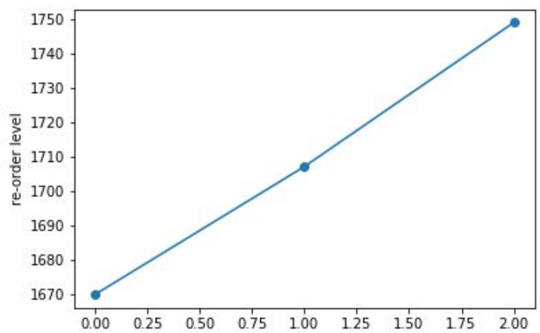
## Estimating Re-Order Level:

- Re-Order Level is the Inventory Level at which we need to make an order, so that we will not be Out of Stock.
- Re-Order Level = Lead days x Daily Demand of product
  - Mean of Lead days is given as 10
  - Standard Deviation of Lead Days = 3
  - So, we will be varying the lead days according to the accuracy which means we will get the Re-order Level with some probability
  - Daily Demand of Product is Total Annual Demand divided by the Number of Days in the Year
- The following plot is Re-order Level Vs Lead Days for all years.




- Blue, Orange & Green lines represents 2015, 2016 & 2017 respectively.
- Data for the above graph is available in this file ‘Reorder and maximum inventory level Vs lead day(Venerators).csv’
- If we don’t take Mean & Standard deviation of the Lead Days into consideration then the constant safety stock comes into consideration.

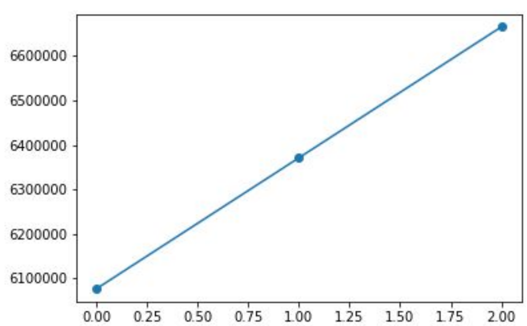
- The above statement implies that, Re-order level = Lead days mean x Daily Demand + Constant Safety Stock.
- Value of Constant Safety Stock = 854.8 Units
- The following plot represents Re-order Level vs Year, for the above mentioned formula. Here again, the numbers 0,1 and 2 on X-axis represents 2015, 2016 and 2017 respectively.



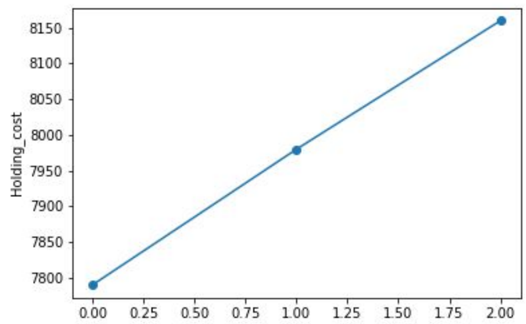

- Re-order level for 2015, 2016 & 2017 are 1670.1424657534246, 1707.2590163934426 & 1749.0465753424658 respectively.

## Calculating Total Ordering Cost and Holding Cost:

- Total Ordering Cost = Total Annual Demand x (Purchasing cost per unit + price of a unit)
- The following plot represents Total ordering cost vs Year. The numbers 0,1 and 2 on X-axis represents 2015, 2016 and 2017 respectively.



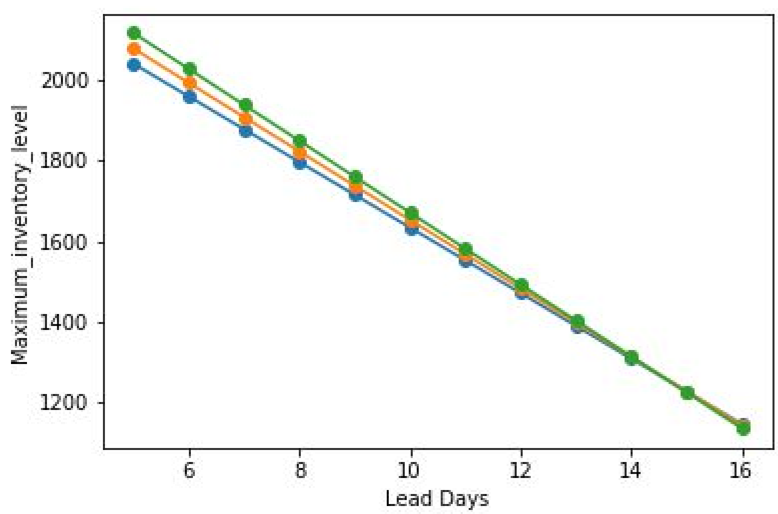

- Total Ordering Cost for 2015, 2016 & 2017 are 6076992.0 (INR), 6371040.0 (INR) & 6665088.0 (INR) respectively.
- Total Holding Cost = EOQ * 0.5 * Inventory Cost.
- The following plot represents Total ordering cost vs Year. Here again, the numbers 0,1 and 2 on X-axis represents 2015, 2016 and 2017 respectively.




- Total Holding Cost for 2015, 2016 & 2017 are 7790.0 (INR), 7980.0 (INR) & 8160.0 (INR).

## Maximum Inventory Level:

- Maximum inventory level = Mean lead time * Daily Demand + Constant Safety Stock - Lead time * Daily Demand + EOQ.
- The following plot represents Maximum inventory vs Year. Here again, the numbers 0,1 and 2 on X-axis represents 2015, 2016 and 2017 respectively.




- Data for the above graph is available in this file 'Reorder and maximum inventory level Vs lead day(Venerators).csv'
- Expected Maximum inventory values for the year using probability of lead days occurrence for 2015, 2016 & 2017 are 1521, 1538 & 1552 respectively.

## Risk vs Money scenario:

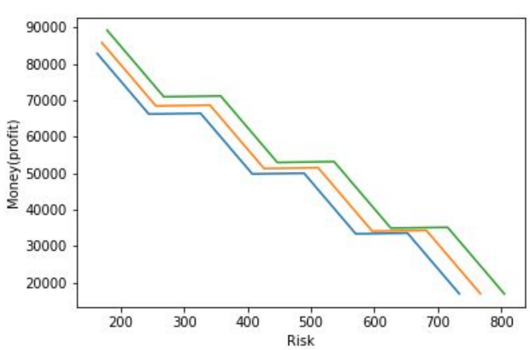
Implementation for Type-A Medicine:

- Let the lead stock in the previous case when we had optimal safety stock to handle emergency case be "L"
- And the lead days when we prefer to take orders at higher costs in case of emergency be " L` ".
- And D be the amount of consumption of medicine per day.
- When the safety stock ranges in between "L*D and L`*D", the hospital inventory can be said to be out of stock and can run into emergency at any time.
- Now the formula to relate between risk and money is given by:

  [10*D-k*D]*204.2+[(Cosant safety stock)*(holding cost)]-[[(10-k)+1]/2 *206.8*D]

  Where k is the no.of lead days which can fluctuate between L` to (L-1)(for the hospital to run out of optimal stock).

- As we will be paying more for each unit of type -A medicine now than in the earlier optimal safety stock case, the loss will increase as the no.of units of purchasing from emergency vendors increases.
- The total profit we got is the difference between safety stock costs in the both cases. And the losses incurred by purchasing units of medicine is deducted from total profit obtained which gives the total money.
- This is the trade-off between risk and money.
- So,Basically As the risk increases,our total money will be decreasing which can be seen from the implementation.
  [here the risk is no.of units of type -A medicine purchased in case of emergency at high prices].
- The following plot represents Maximum_inventory_level Vs Year , Y-Axis represents the Total Holding Cost and X-Axis represents the Year(On the X-Axis 0 , 1 , 2 represents 2015 , 2016 , 2017 respectively) .




- In the above implementation:

- Blue line represents the year 2015
- Orange line represents the year 2016
- Green line represents the year 2017
- From the Plot we can interpret that , if as the risk Increases there is a Decrease of Profit .

## Python Libraries Used :

1. Pandas
2. NumPy
3. StatsModels
4. Matplotlib
5. Pyramid
6. Pyramid.arima