**DRUG SHORTAGE**

**INDUSTRY:** Pharmaceutical.

**DEPARTMENT:** Crucial Pain Points in Organization.

**PROBLEM STATEMENT:** Prediction of Drugs Shortage.

**1. Pharmaceutical Industry:**

Pharmaceutical industry, the discovery, development, and manufacture of drugs and medications (pharmaceuticals) by public and private organizations.

**2. Pain Points in Pharmaceutical:**

* Drug Shortages.
* Prescription Adherence.
* Managing Inventory.
* Reimbursement and Cost Management.
* Managing Workflow.

**Reference:**

<https://blog.shelvingdesignsystems.com/5-most-common-pain-points-for-pharmacies>

<https://info.dicksondata.com/blog/the-biggest-pain-points-of-modern-medicine-and-how-to-treat-them>

**3.Working on:** We are Working on Drug Shortages.

**4. Reasons for Drug Shortages:**

Drug shortage is a global issue affecting low, middle, and high-income countries. Many countries have developed various strategies to overcome the problem, while the problem is accelerating, affecting the whole world. All types of drugs, such as essential life-saving drugs, oncology medicines, antimicrobial drugs, analgesics, opioids, cardiovascular drugs, radiopharmaceutical, and parenteral products, are liable to the shortage. Among all pharmaceutical dosage forms, sterile injectable products have a higher risk of shortage than other forms. The causes of shortage are multifactorial, including supply issues, demand issues, and regulatory issues. Supply issues consist of manufacturing problems, unavailability of raw materials, logistic problems, and business problems. In contrast, demand issues include just-in-time inventory, higher demand for a product, seasonal demand, and unpredictable demand. For regulatory issues, one important factor is the lack of a unified definition of drug shortage. Drug shortage affects all stakeholders from economic, clinical, and humanistic aspects. WHO established global mitigation strategies from four levels to overcome drug shortages globally. It includes a workaround to tackle the current shortage, operational improvements to reduce the shortage risk and achieve early warning, changes in governmental policies, and education and training of all health professionals about managing shortages.

**Reference:**

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8299364/>

**5. Feature Engineering:**

**5.1 Feature Creation:**

1. Manufacturing.
2. Quality Problems.
3. Delays.
4. Discontinuations.
5. Supply issues.
6. Demand issues.
7. Regularity issues.
8. Manufacturing Problems.
9. Lack of raw materials.
10. Logistic Problems.
11. Seasonal demand.
12. Unpredictable demand.
13. Health Emergency.
14. Usage.
15. Industry Consolidation.
16. Marketing.
17. Procurement.

**Reference:**

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8299364/>

**5.2 Feature Selection:**

1. Manufacturing Problem.
2. Delays.
3. Discontinuations.
4. Supply issues.
5. Demand issues.
6. Regularity issues.
7. Lack of raw materials.
8. Seasonal demand.
9. Unpredictable demand.
10. Health Emergency.

## 5.3 Feature Transformation:

1. Manufacturing Problem.(1-50)
2. Lack of raw materials.(1 or 2)
3. Delivered on time( 0 or 1)
4. Demand Rate.(1-100)
5. Health Emergency.(0 or 1)
6. Seasonal demand (1-2)
   1. **Feature Extraction:**
7. Manufacturing Problem.
8. Lack of raw materials.
9. Delivered on time
10. Demand Rate.
11. Health Emergency.
12. Seasonal demand.
13. **Dataset Creation:**
    1. **Dataset Creation Code:**

import random

import pandas as pd

[w1,w2,w3,w4,w5,w6]=[0.4964,3.028,0.719,0.7107,-2.0217,-0.7157]

vals=[]

for i in range(10000):

x1=random.randint(1,50)

x2=random.randint(1,2)

x3=random.randint(0,1)

x4=random.randint(1,100)

x5=random.randint(0,1)

x6=random.randint(1,4)

eq=w1\*x1+w2\*x2+w3\*x3+w4\*x4+w5\*x5+w6\*x6

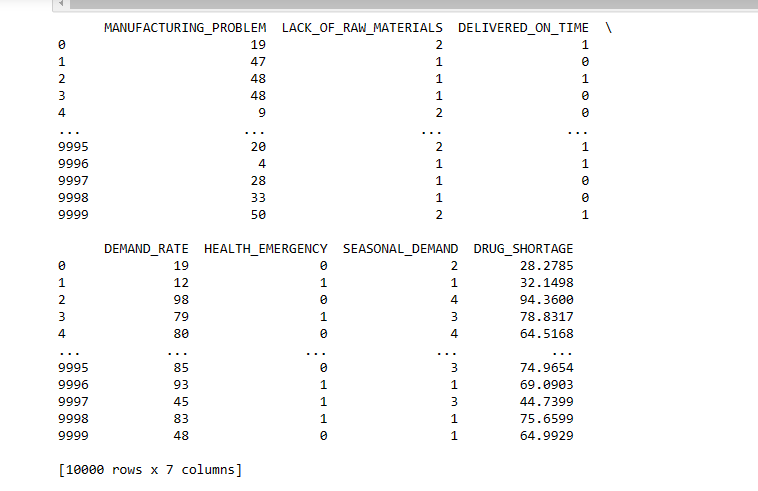
vals.append([x1,x2,x3,x4,x5,x6,eq])

df=pd.DataFrame(vals,columns=['MANUFACTURING\_PROBLEM','LACK\_OF\_RAW\_MATERIALS','DELIVERED\_ON\_TIME','DEMAND\_RATE','HEALTH\_EMERGENCY','SEASONAL\_DEMAND','DRUG\_SHORTAGE'])

print(df)

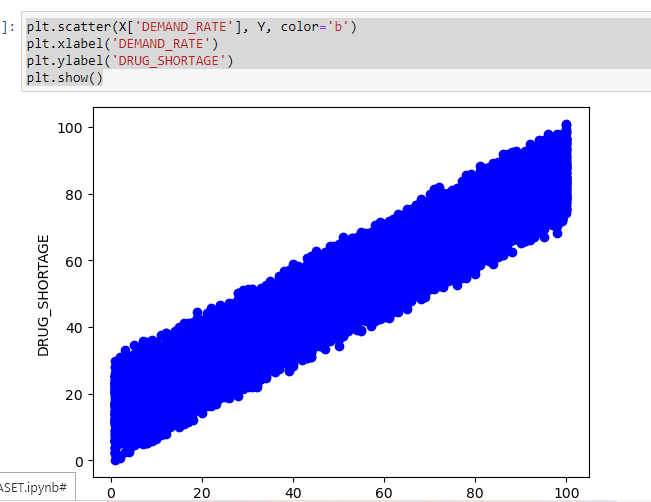
df.to\_csv('Drugshortage.csv',index=False)

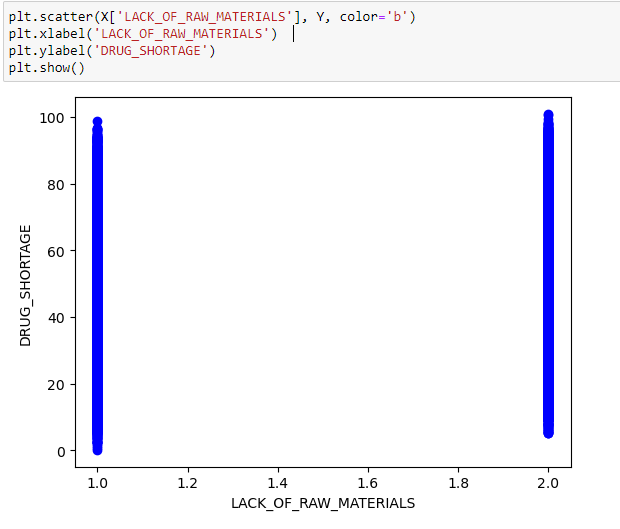
* 1. **Final Dataset:**

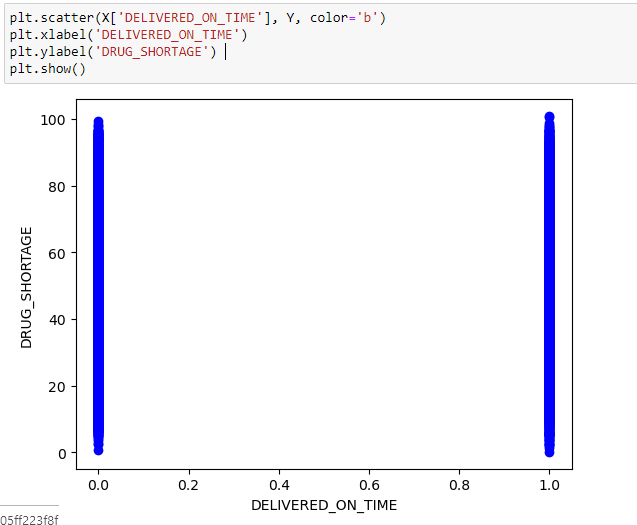


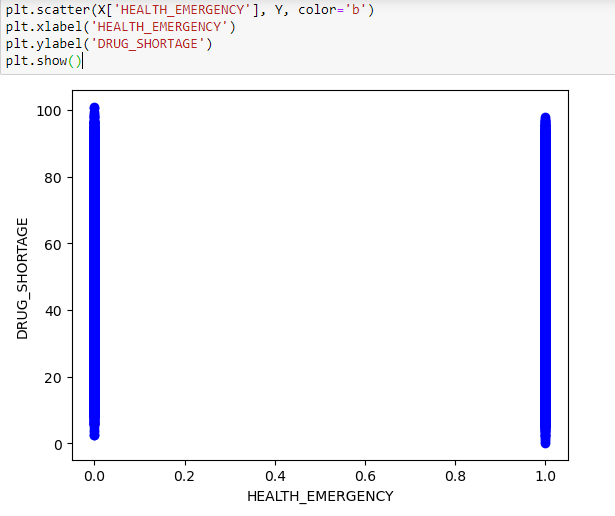
**8 Analysis of Scatterplot :**

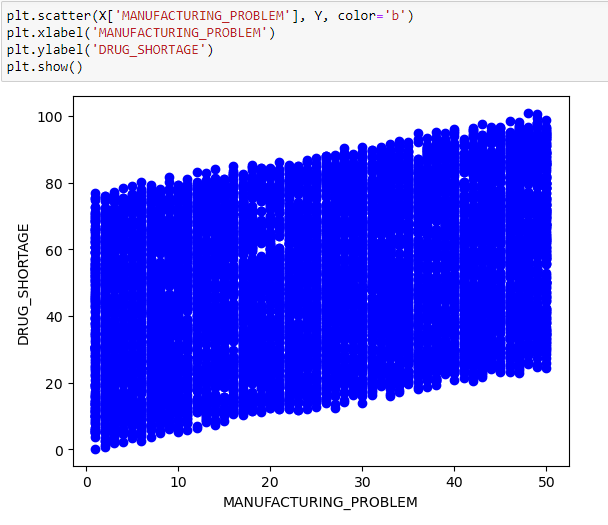
From the scatterplot we can say that the shortage of drugs increases if the demand rate of the drug increases.

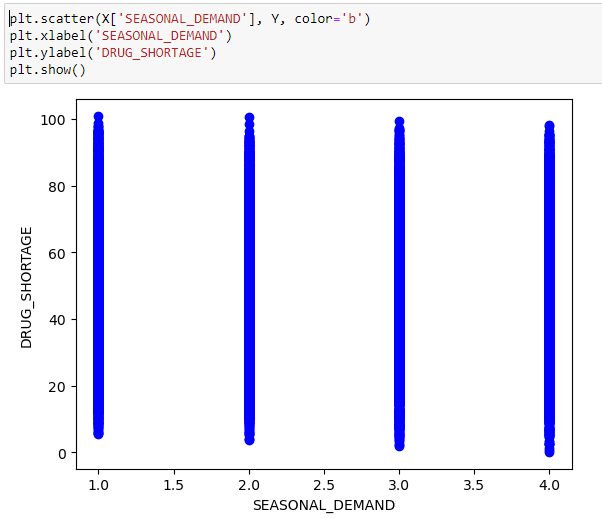
****

****

****

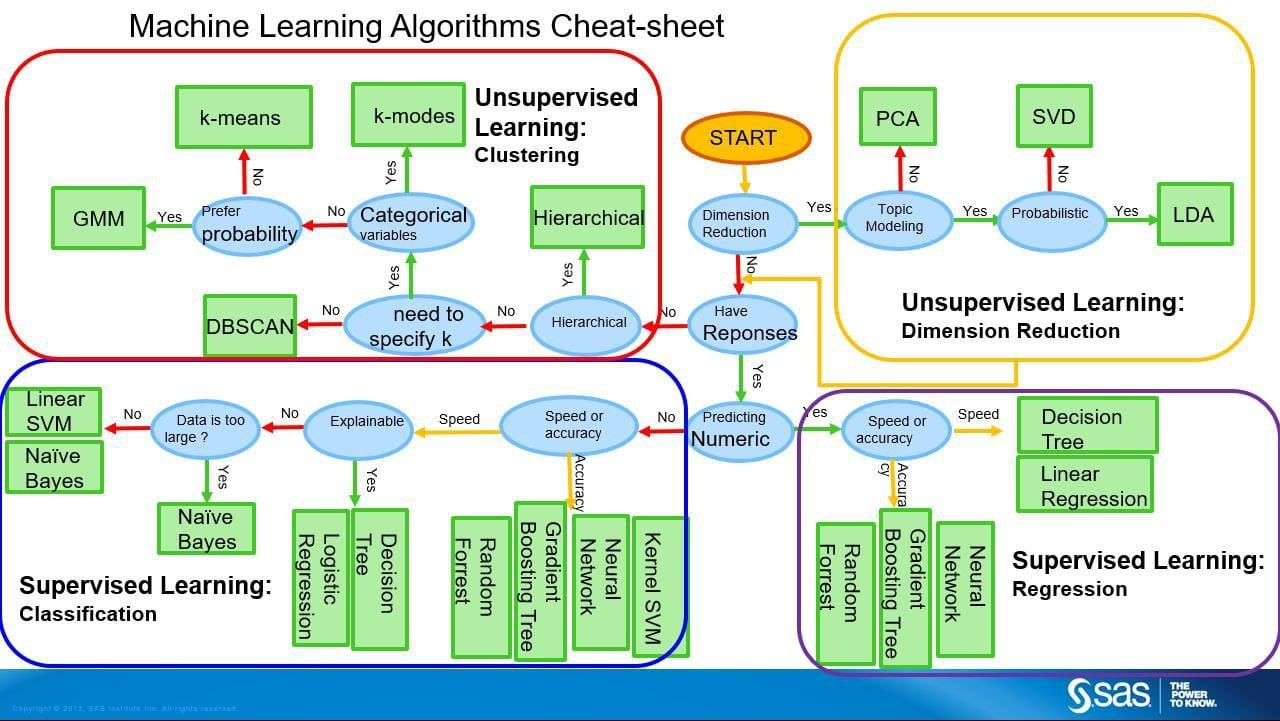
****

****

****

**8.Model selection:**

As we have intended to predict numerical values we preferred linear regression as first choice but due to existence of multiple independent variables we have chosen Multiple regression but data is overfitted . From given Cheat sheet we selected Random Forest as the best model for our requirement based on accuracy.



**9.Code:**

from sklearn.ensemble import RandomForestRegressor

mdl = RandomForestRegressor(n\_estimators=100,max\_depth=6)

mdl.fit(X, Y)

pred = mdl.predict(inps)

print("Predicted value (RFR): ",pred[0])

print("Accuracy (RFR): ",mdl.score(X[:100], Y[:100])\*100)

plt.scatter(X['DEMAND\_RATE'], Y, color='b')

plt.plot(X['DEMAND\_RATE'], mdl.predict(X),color='black',linewidth=3)

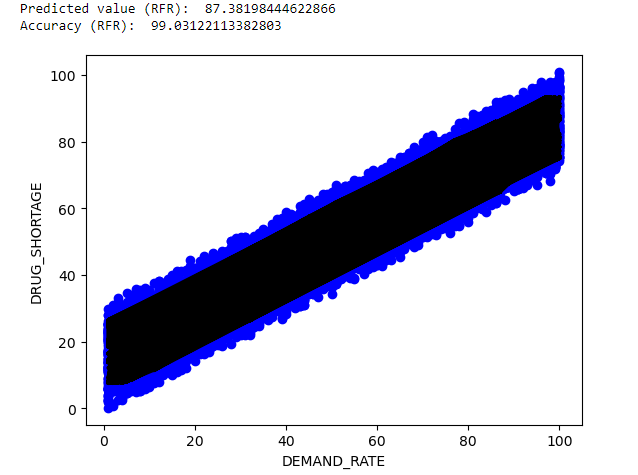
plt.xlabel('DEMAND\_RATE')

plt.ylabel('DRUG\_SHORTAGE')

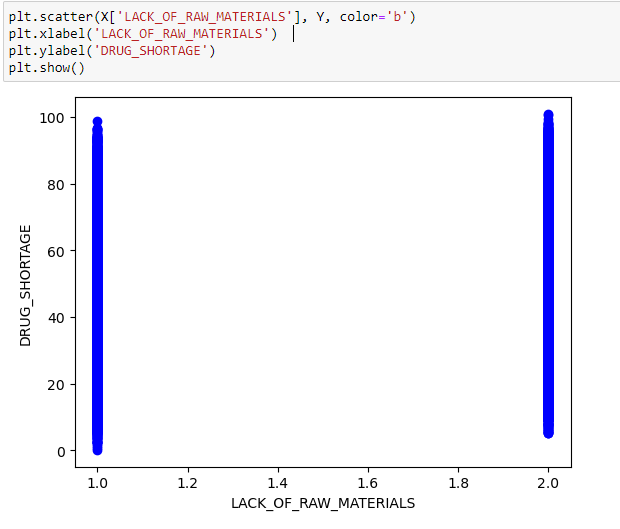
plt.show()

**9.1 Observation:**

* From the scatterplot, between drugs shortage and demand rate, we can say that the shortage of drugs increases if the demand rate of the drug increases

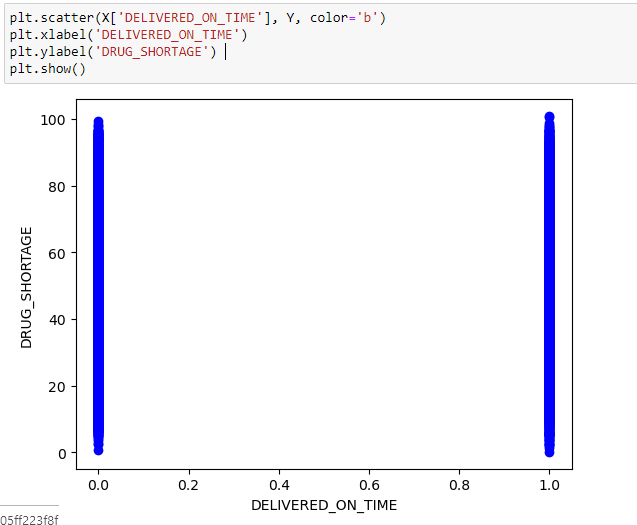


* From the scatterplot between Drugs Shortage and Lack\_of\_Rawmaterials,,we observed that points are at 1 and 2 since we take values of Lack\_of\_Rawmaterials as 1 and 2

****

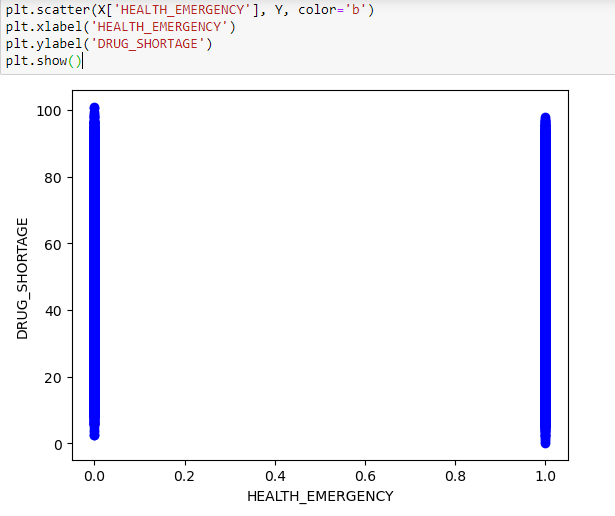
* From the scatterplot between Drugs Shortage and Delivery\_on\_time,we observed that points are at 0 and 1 since we take values of Delivery\_on\_time

as 0 and 1

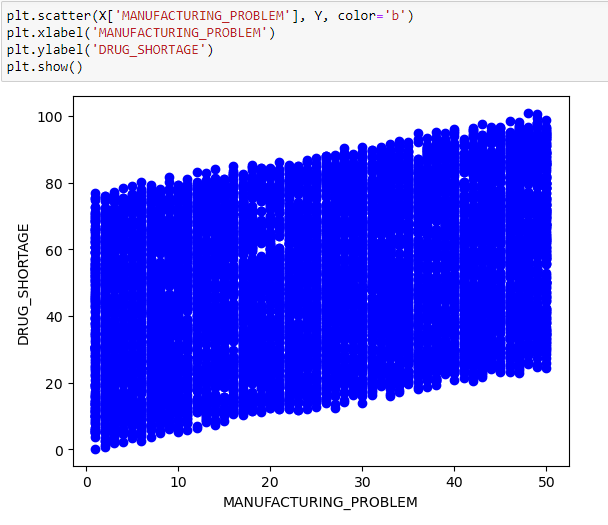


* From the scatterplot between Drugs Shortage and Health Emergency,we observed that points are at 0 and 1 since we take values of Health Emergency

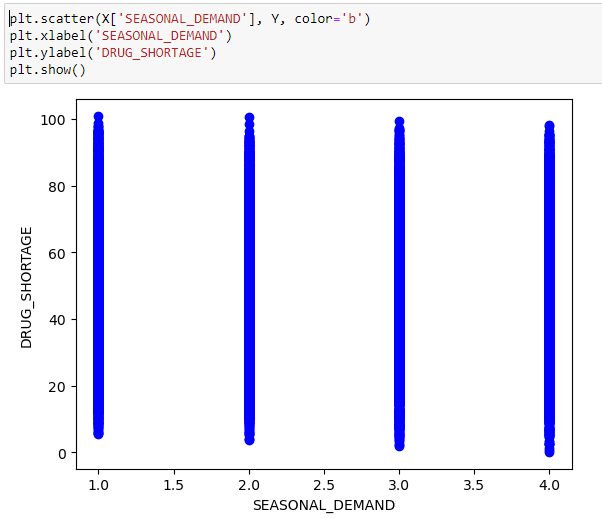
as 0 and 1



* From the scatterplot between Drugs Shortage and Manufacturing\_problem,we observed that Drug\_Shortage increases in increase of Manufacturing\_problem



* From the scatterplot between Drugs Shortage and Seasonal\_Demand,we observed that points are at 4 attributes,since we gave 4 different values for 4 different seasons



Prediction output:

When we take input as 14,2,1,56,0,2 we get output , Drug\_shortage as 49.313402451877664 with accuracy 99.01902348558514

