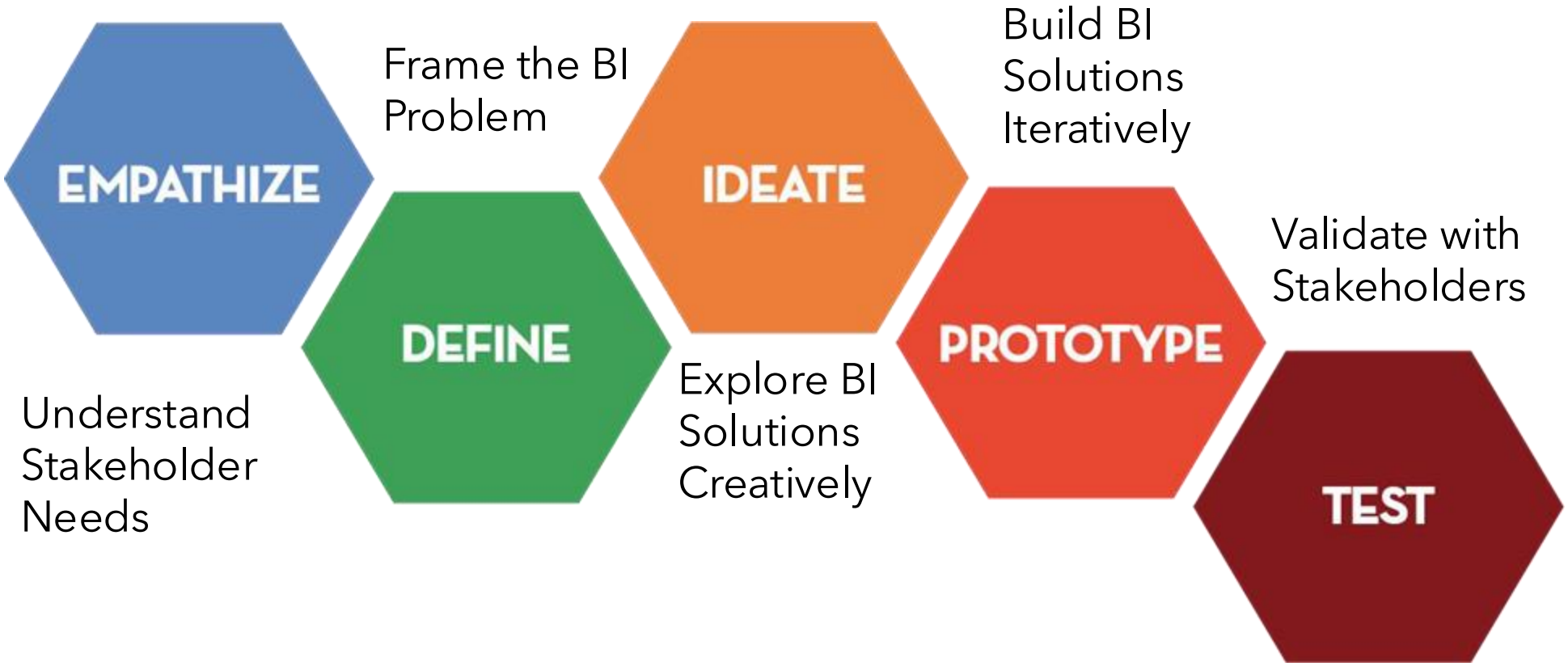
Supply Chain Case Study

CS 459 Business Intelligence

DESIGN THINKING in BI

The framework



Empathize
Understand Stakeholder Needs

Gather deep insights into business problems from stakeholders.

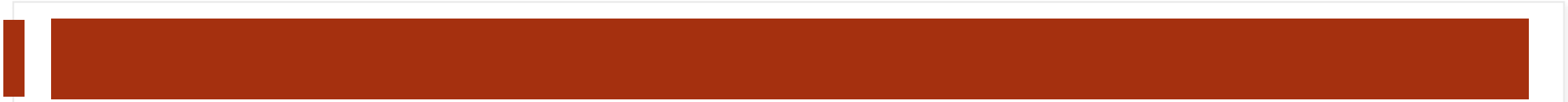
Background Data Knowledge

- Gather knowledge about Supply chain from various resources to understand the dataset and its dynamics.
- It is important to understand how supply chain works before you start to conduct analysis.

Understand what is a Supply Chain? - Different perspectives

- *"Conjures thoughts of purchasing and procurement needed to acquire raw materials for production".*
- *"Warehousing and distribution, transportation and retail channels".*
- *"Sources of capital or human resource"*

Recall EdPuzzle Video and Questions



What is Supply Chain Management? - Standard Definitions

- Design and management of value-added process across organizational boundaries to meet customer needs

Institute for Supply Management

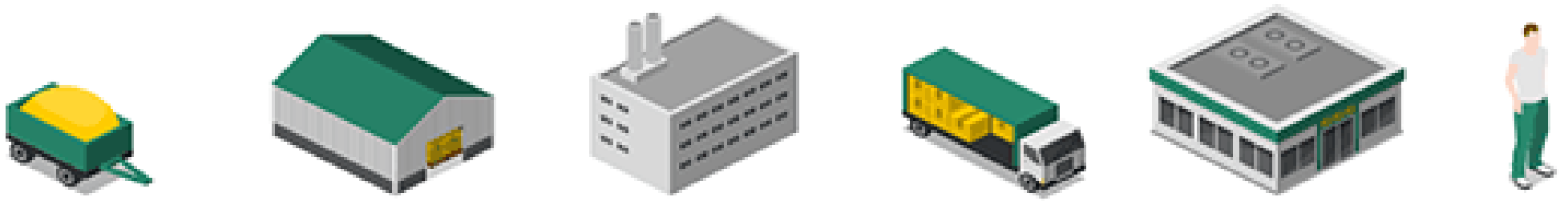
- Managing supply and demand, sourcing raw materials and parts, manufacturing and assembly, warehousing and inventory tracking, order entry and order management, distribution across all channels, and delivery to the customer

The Supply Chain Council

What is a Supply Chain?

- *All stages involved*, directly or indirectly, in fulfilling a customer request
- Includes manufacturers, suppliers, transporters, warehouses, retailers, and customers
- Within each company, the supply chain *includes all functions involved* in fulfilling a customer request (product development, marketing, operations, distribution, finance, customer service)

Supply Chain Diagram



Raw Materials → Supplier → Manufacturer → Distributor → Retailer → Consumer

Fundamentals of the Pharmaceutical Supply Chain

- The pharmaceutical supply chain is complex - address common challenges to provide patients their needed medications efficiently.

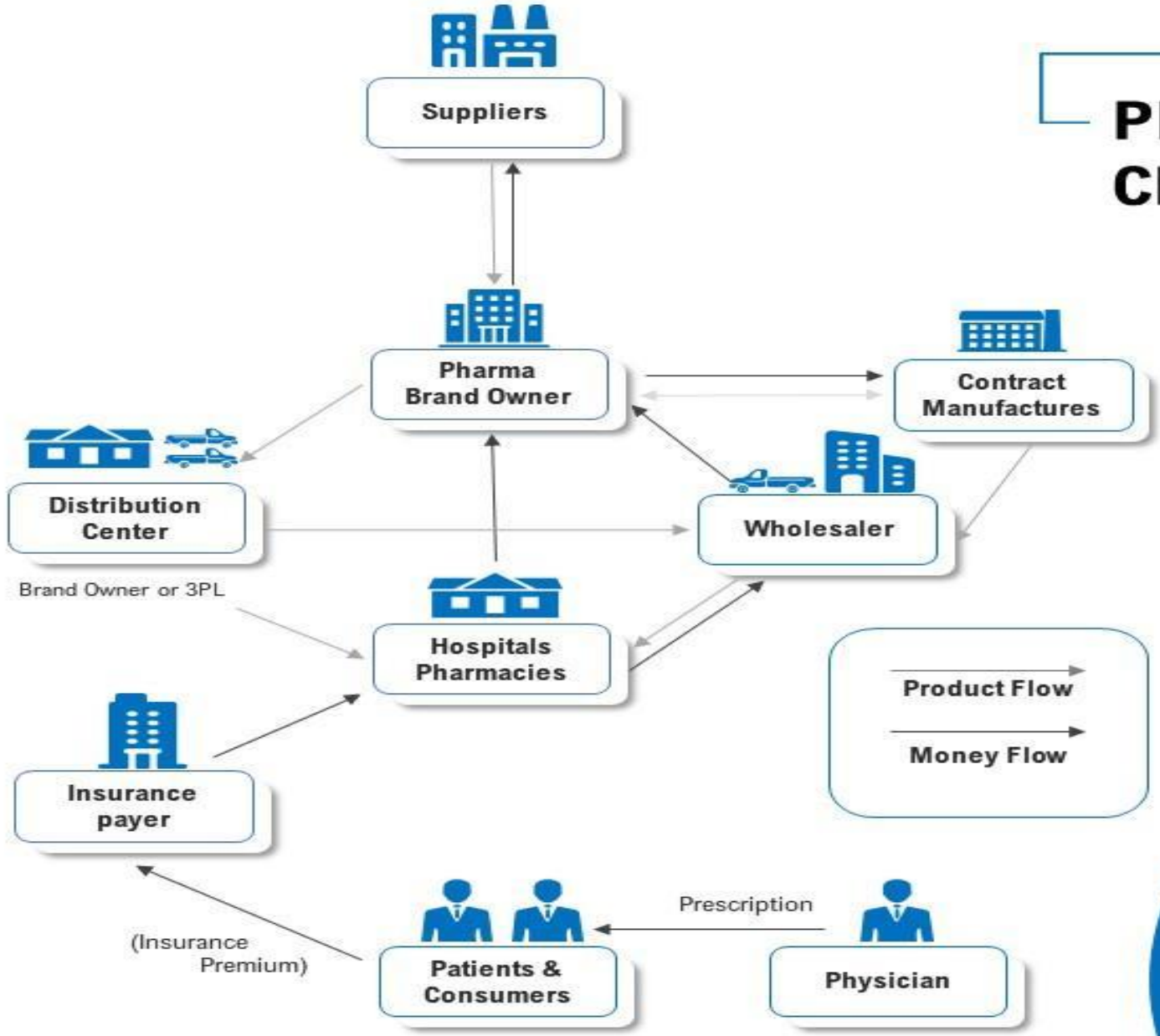
Fundamentals of the Pharmaceutical SupplyChain(pharmanewsintel.com)

- At the most basic level, there are five-steps in the pharmaceutical supply chain to ensure that drug inventory is readily available for distribution to providers and patients

Steps in Pharma Supply Chain

1. Pharmaceuticals originate in manufacturing sites
2. Are transferred to wholesale distributors
3. Stocked at retail, mail-order, and other types of pharmacies
4. Subject to price negotiations and processed through quality and utilization management screens by pharmacy benefit management companies
5. Dispensed by pharmacies; and ultimately delivered to and taken by patients

Pharmaceutical Supply Chain Flow



Further background knowledge

- **Competitor Analysis** – **Know Your Competitor**
- **Interviews** – **Get into the Manager's shoes – Know Your Stakeholders**
- Excel Analysis
- SQL Analysis

Complete your background knowledge before proceeding.

Record each step on a provided template.

Data Knowledge Analysis

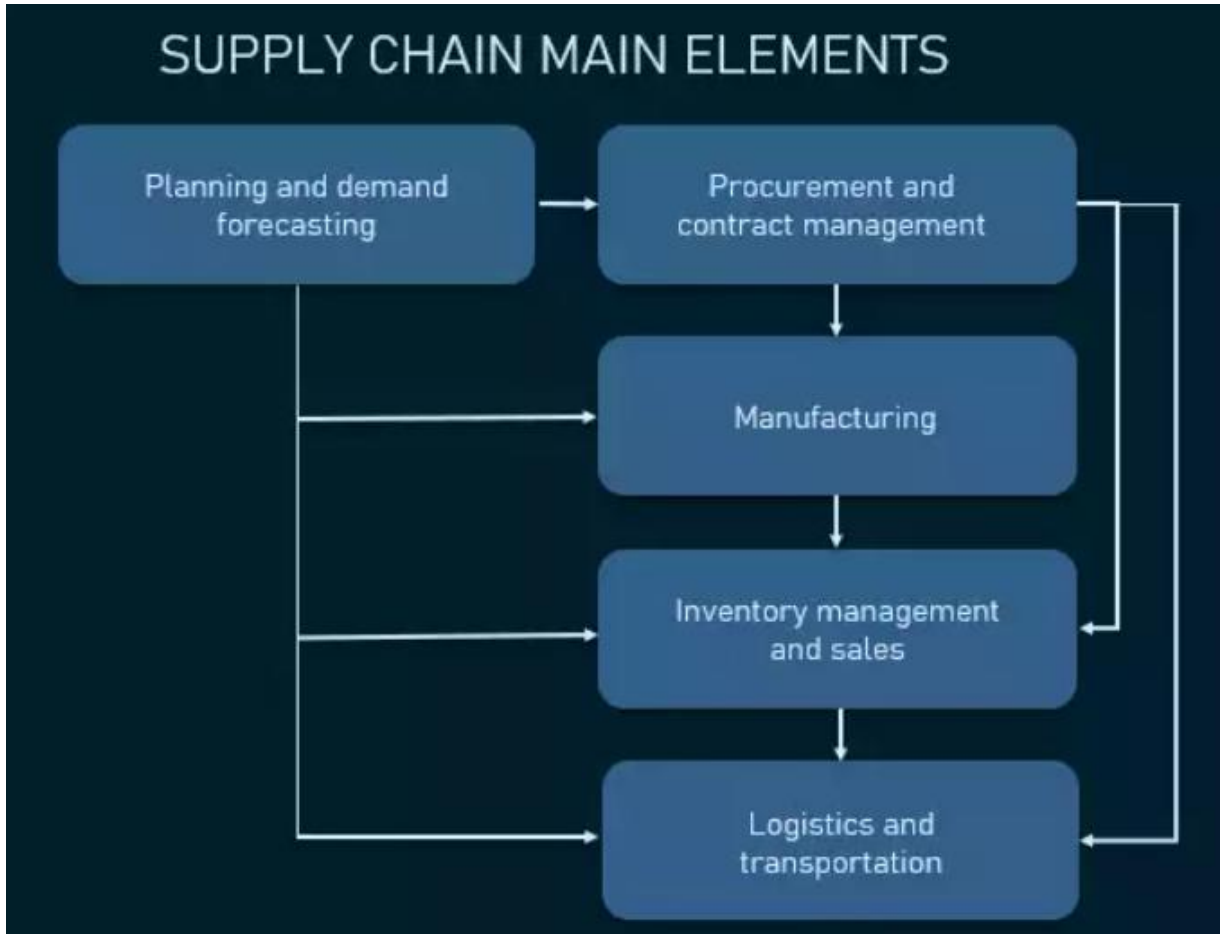
- Now that we know how the supply chain works, we will proceed to analyze that information
- The following to be quantified by problem being solved and the performance metrics:

1. *What are we doing?*
2. *What the competitors are doing?*
3. *What is the global perspective?*

- Gauge the performance of the company based on the above and identify the gaps

Elements of Supply Chain

- **Planning** mostly concerns demand forecasting and resource planning.
- **Procurement** is a set of operations related to choosing vendors, negotiating the terms of cooperation, and buying supplies needed for your business.
- **Manufacturing** deals with production and capacity management.
- **Inventory management** is focused on keeping the optimal stock balance, sales, and warehousing operations.
- **Logistics management** covers order fulfillment and all delivery activities.



Optimization Opportunities in Supply chain for use of Analytics in general

Define

Frame the BI Problem

Problem Statement Generation

Supply Chain Case Study

Problems in Supply Chain Dataset

Efficiency of supply chain
PQ First Sent to Client Date
PO Sent to Vendor Date
Scheduled Delivery Date
Delivered to Client Date
Delivery Recorded Date
Using the dates to compute new variables that tell about the supply chain efficiency.
Can drill down to find orders that took longer than necessary (comparing to a benchmark) - identify delays in delivery
Crucial to see which regions or areas have an underlying problem
Productivity is the volume of output over a set period.
Efficiency refers to the quality of work and how well allocated resources are utilised.

Problems in Supply Chain Dataset

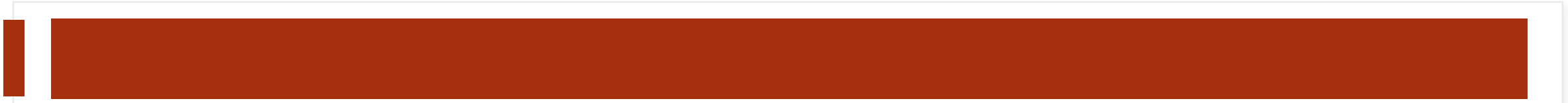
Analysis of Productivity in supply chain
Identify the productivity of VENDORS by generating volume of products delivered by them. (weekly, monthly, annual productivity)
Analysis by Project code and manufacturing site.
Other variables can be used to do the same.
Freight Cost Analysis - Goal could be to minimize costs
Identify areas of high costs (in this case we have Freight costs only)
Understand which products are causing these costs
Sales Trends - Demand Forecasting
Understand the sales trends
Time Series Analysis and Forecasting demand in near future

Problem Statement 1 - Analysis of the Efficiency of Supply Chain

Efficiency of supply chain
PQ First Sent to Client Date
PO Sent to Vendor Date
Scheduled Delivery Date
Delivered to Client Date
Delivery Recorded Date
Using the dates to compute new variables that tell about the supply chain efficiency.
Can drill down to find orders that took longer than necessary (comparing to a benchmark) - identify delays in delivery
Crucial to see which regions or areas have an underlying problem
Productivity is the volume of output over a set period.
Efficiency refers to the quality of work and how well allocated resources are utilised.

Using the "How might we…" frame for the problem definition

How might we analyze and improve the supply chain efficiency?



Fill the Background Data Knowledge Template

Gain better understanding of the business, the problem and the data.

Data Gathering and Exploration

Gather Data and Perform EDA

Generating new columns for Analysis

- **Delivery.Delay = Delivered – Scheduled**
  The delays in supply chain in days
  - NEGATIVE: Early delivery
  - ZERO: On-time Delivery
  - POSITIVE: Late Delivery
- **Delivery.Time = Delivered – PO sent**
  How long it took from PO sent to vendor to when it was delivered to client
- **DeliveryDate.Entry.Delay** OR **DD.Entry.Delay = Recorded – Delivered**
  How long did it take for delivery to be recorded after it was delivered.

Data Profile

The dataset contains 4638 rows and 32 columns:

```
#Identifying Rows and Columns in the data
df.shape
```

```
(4638, 32)
```

The dataset contains 9 numeric columns and 23 non-numeric columns.

```
#Exploring Data Type
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4638 entries, 0 to 4637
Data columns (total 32 columns):
 #   Column                      Non-Null Count  Dtype
---  ------                      --------------  -----
 0   Unnamed: 0                  4638 non-null   int64
 1   Project Code                4638 non-null   object
 2   PQ #                        4638 non-null   object
 3   PO / SO #                   4638 non-null   object
 4   Country                     4638 non-null   object
 5   Managed By                  4638 non-null   object
 6   Fulfill Via                 4638 non-null   object
 7   Vendor INCO Term            4638 non-null   object
 8   Shipment Mode               4638 non-null   object
 9   PQ First Sent to Client Date  4638 non-null   object
 10  PO Sent to Vendor Date      4638 non-null   object
 11  Scheduled Delivery Date     4638 non-null   object
 12  Delivered to Client Date    4638 non-null   object
 13  Delivery Recorded Date      4638 non-null   object
 14  Product Group               4638 non-null   object
 15  Sub Classification          4638 non-null   object
 16  Vendor                      4638 non-null   object
 17  Item Description            4638 non-null   object
 18  Molecule/Test Type          4638 non-null   object
 19  Brand                       4638 non-null   object
 20  Dosage                      4638 non-null   object
 21  Dosage Form                 4638 non-null   object
 22  Unit of Measure (Per Pack)  4638 non-null   int64
 23  Line Item Quantity          4638 non-null   int64
 24  Line Item Value             4638 non-null   float64
 25  Pack Price                  4638 non-null   float64
 26  Unit Price                  4638 non-null   float64
 27  Manufacturing Site          4638 non-null   object
 28  First Line Designation      4638 non-null   object
 29  Weight (Kilograms)          4638 non-null   float64
 30  Freight Cost (USD)          4638 non-null   float64
 31  Line Item Insurance (USD)   4638 non-null   float64
dtypes: float64(6), int64(3), object(23)
```

List of numeric columns:

```
#Prints all the numeric columns
numeric_columns = get_num_cols(df)
print(numeric_columns)
```

```
['Unnamed: 0', 'Unit of Measure (Per Pack)', 'Line Item Quantity', 'Line Item Value',
```

List of non-numeric columns:

```
#Prints all the categorical columns
categorical_columns = get_cat_cols(df)
print(categorical_columns)
```

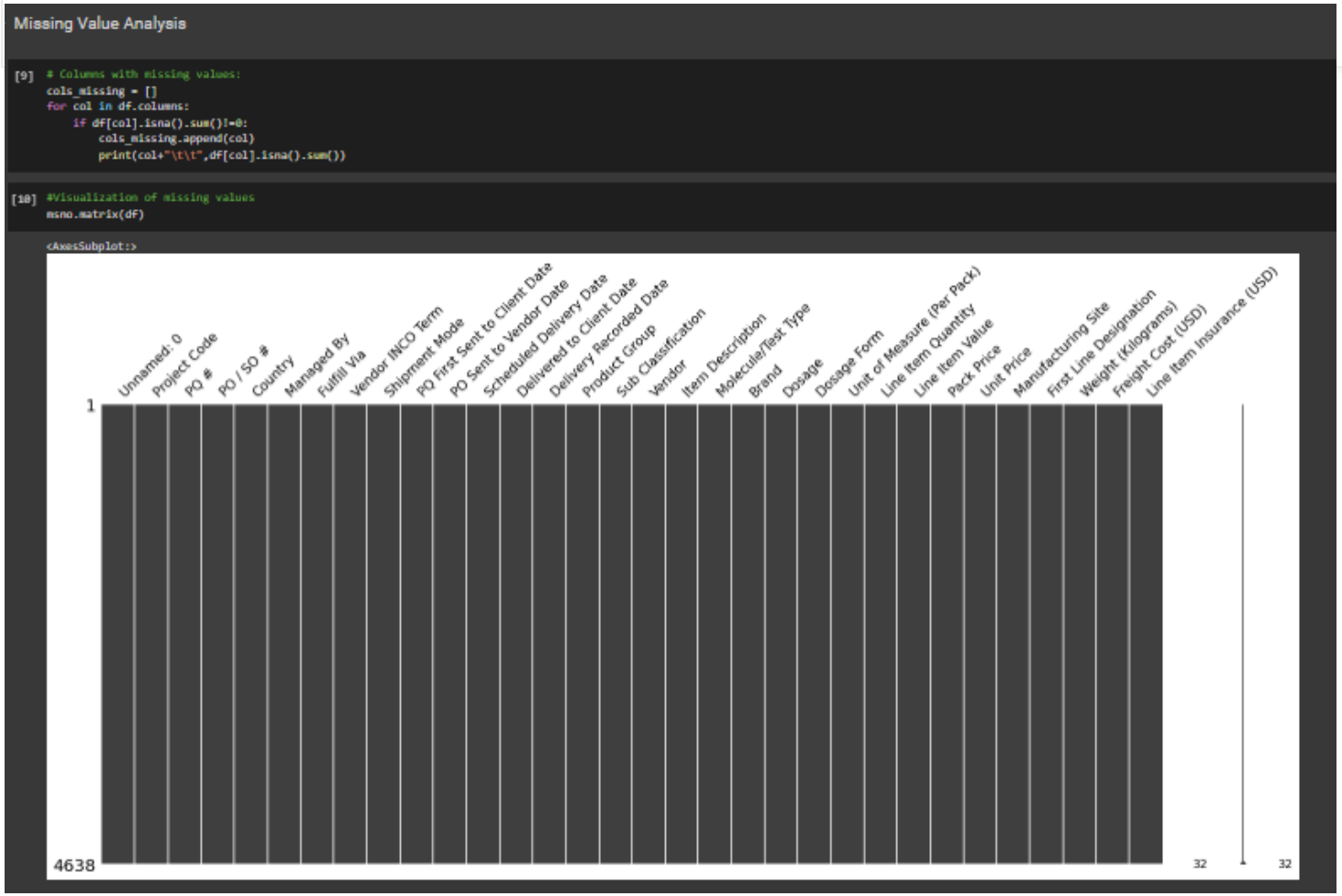
```
['Project Code', 'PQ #', 'PO / SO #', 'Country', 'Managed By', 'Fulfill Via',
'Vendor INCO Term', 'Shipment Mode', 'PQ First Sent to Client Date', 'PO Sent to Vendor Date',
'Scheduled Delivery Date', 'Delivered to Client Date', 'Delivery Recorded Date', 'Product Group',
'Sub Classification', 'Vendor', 'Item Description', 'Molecule/Test Type', 'Brand',
'Dosage', 'Dosage Form', 'Manufacturing Site', 'First Line Designation']
```

Missing Values and Duplicates

Missing Value Analysis

```
[9] # Columns with missing values:
    cols_missing = []
    for col in df.columns:
        if df[col].isna().sum()!=0:
            cols_missing.append(col)
            print(col+"\t\t",df[col].isna().sum())
```

```
[10] #Visualization of missing values
     msno.matrix(df)
```



The dataset seems to contain no missing values

The dataset seems to contain no duplicate records

```
df.duplicated().value_counts()
```

```
False    4638
dtype: int64
```

Descriptive Statistics

Numeric Data:

Summary of Numerical Data

```
df.describe().T
```

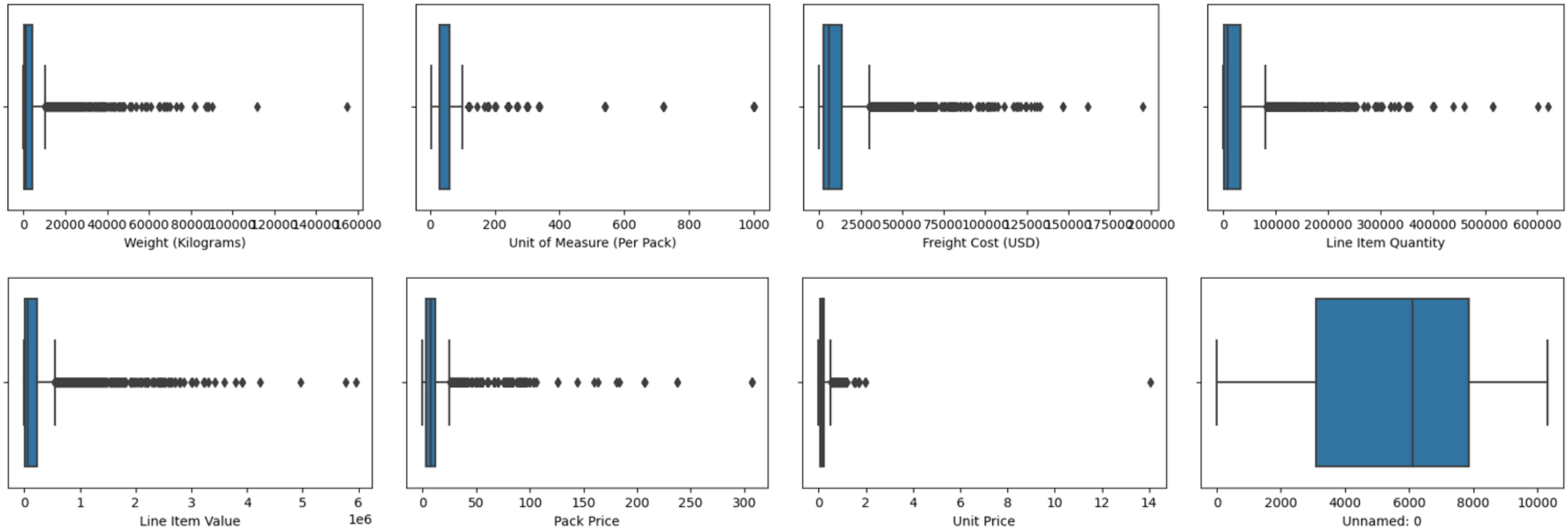
	count	mean	std	min	25%	50%	75%	max
Unnamed: 0	4638.0	5583.948254	2832.553762	13.00	3099.2500	6119.500	7870.750	10316.00
Unit of Measure (Per Pack)	4638.0	81.035144	86.592105	5.00	30.0000	60.000	60.000	1000.00
Line Item Quantity	4638.0	27065.223588	47838.896882	1.00	1500.0000	8341.000	33129.500	619999.00
Line Item Value	4638.0	206221.628648	409749.149714	0.00	10503.0000	58085.700	223776.885	5951990.40
Pack Price	4638.0	12.469987	18.945518	0.00	3.4000	7.400	12.110	306.88
Unit Price	4638.0	0.201419	0.291985	0.00	0.0600	0.140	0.240	14.04
Weight (Kilograms)	4638.0	3865.767141	7831.314758	1.00	263.0000	1308.000	4234.500	154780.00
Freight Cost (USD)	4638.0	10744.380509	15203.464191	14.36	2267.1350	5803.415	13398.060	194623.44
Line Item Insurance (USD)	4638.0	309.154254	591.001517	0.00	16.6625	79.385	329.555	7005.49

Total categories in each non-numeric column:

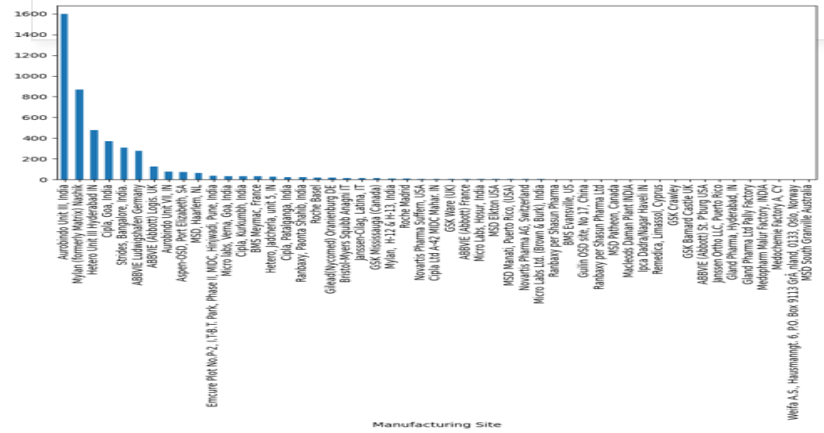
```
for cat in categorical_columns:
  print(cat, ": ", df[cat].nunique())
```

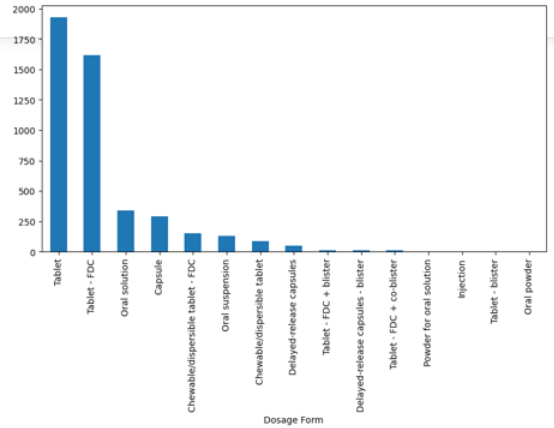
```
Project Code :  82
PQ # :  815
PO / SO # :  4291
Country :  31
Managed By :  1
Fulfill Via :  2
Vendor INCO Term :  7
Shipment Mode :  4
PQ First Sent to Client Date :  546
PO Sent to Vendor Date :  493
Scheduled Delivery Date :  1591
Delivered to Client Date :  1703
Delivery Recorded Date :  1638
Product Group :  3
Sub Classification :  4
Vendor :  34
Item Description :  114
Molecule/Test Type :  40
Brand :  21
Dosage :  51
Dosage Form :  15
Manufacturing Site :  51
First Line Designation :  1
```

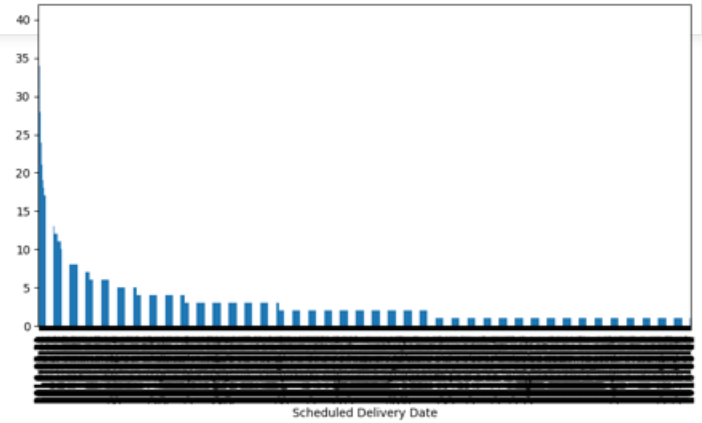
Outlier Analysis

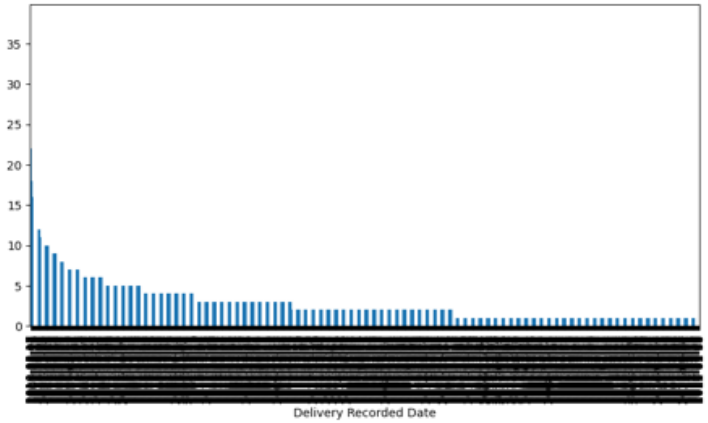


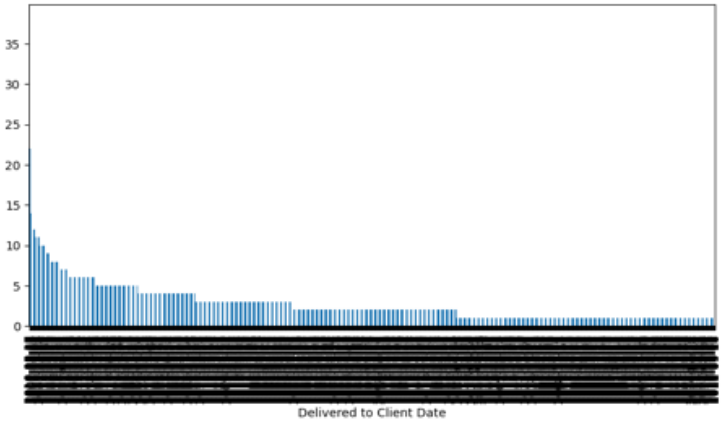
Histograms

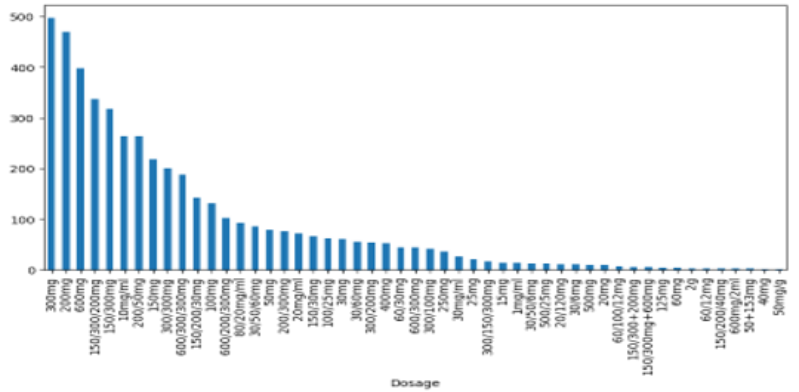

Manufacturing Site

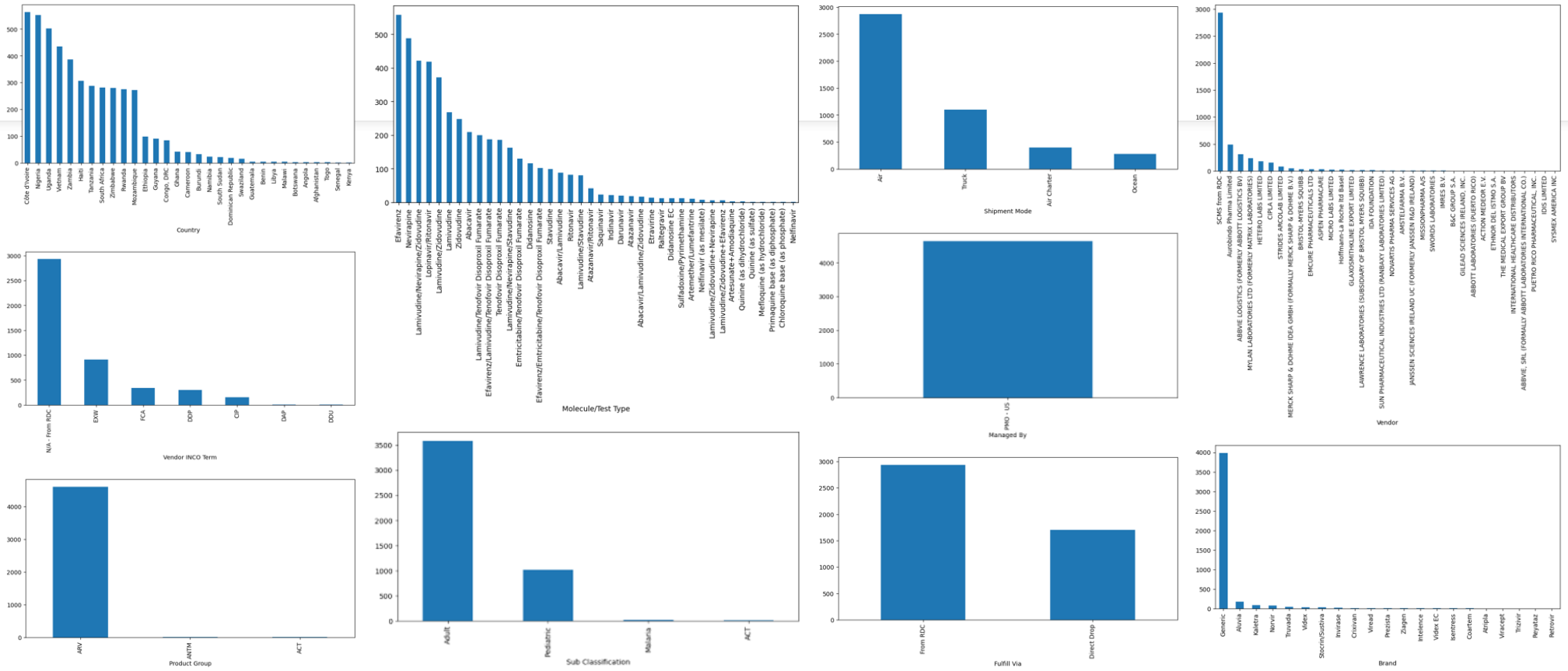

Tablet
Tablet - FDC
Oral solution
Capsule
Chewabledispersible tablet - FDC
Oral suspension
Chewabledispersible tablet
Delayed-release capsules
Tablet - FDC + blister
Delayed-release capsules - blister
Tablet - FDC + co-blister
Powder for oral solution
Injection
Tablet - blister
Oral powder
Dosage Form


Scheduled Delivery Date


Delivery Recorded Date


Delivered to Client Date


Dosage

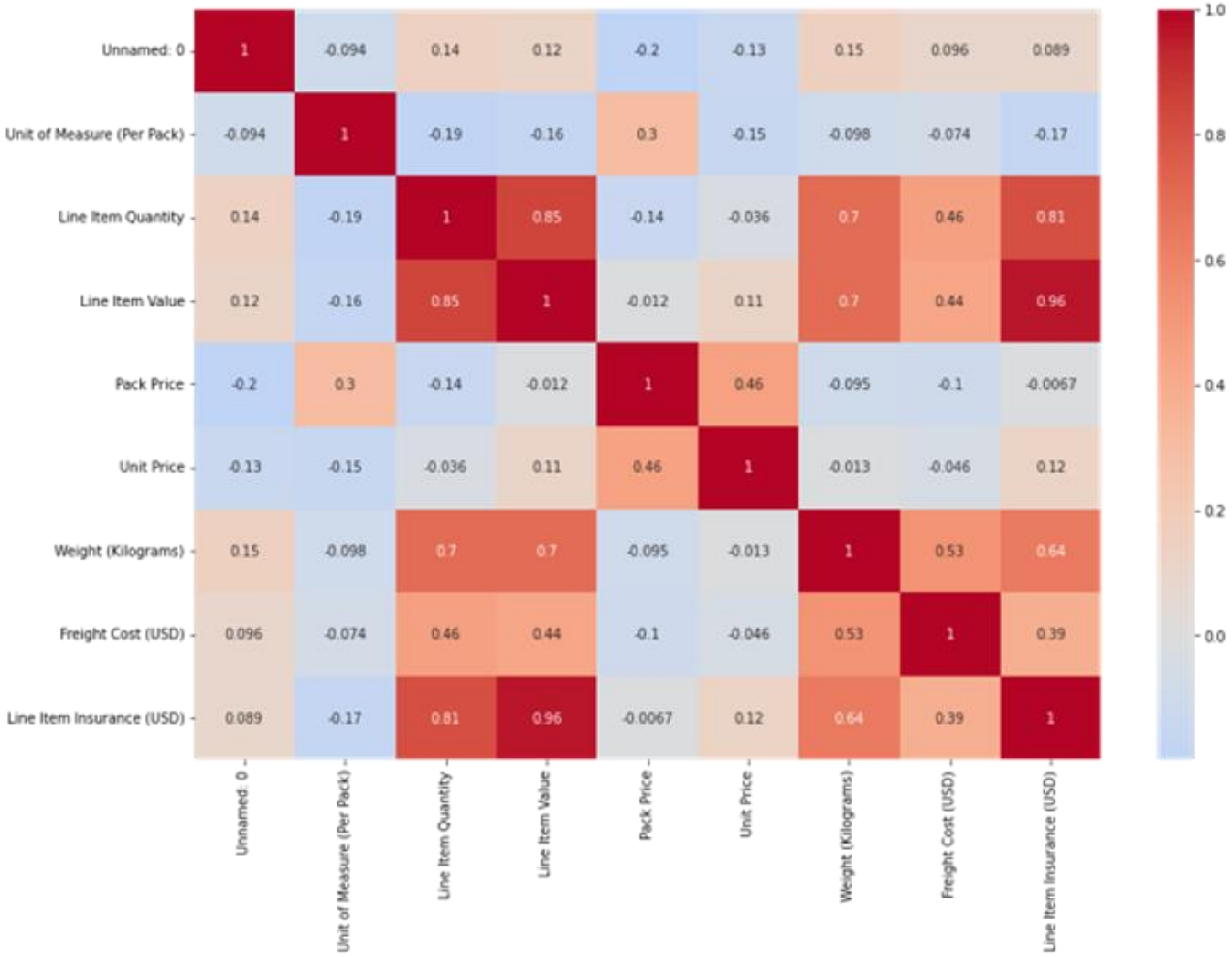


Notice that majority of the graphs are POSITIVELY/RIGHT SKEWED

Statistical Testing

Correlation

- *Strong positive correlation* between Line Item Value with Line Item Quantity and Line Item Insurance.
- *Strong positive correlation* between Weight and Line Item Quantity.
- Other relations exist but are relatively weaker.



Anova Test

- Delivery Delay (Fact) against Categorical Variables (Dimensional attributes)

Feature	P-Value
Vendor	4.136564e-31
Brand	0.084145
Product Group	0.334551
Sub-Classification	0.223597
Item Description	0.002792
Dosage Form	0.000043
Dosage	8.413971e-07
Molecule Test-Type	0.000528
Manufacturing Country	0.000003
Delivery Country	7.293645e-49

- P-value < 0.05
  Reject H0
- Null hypothesis:
  H0 : All Delivery delay means are equal across selected feature for e.g. Vendor

Ideate

Explore BI Solutions Creatively

How might we analyze and improve the supply chain efficiency?

- Come up with ideas of different charts and dashboard and pen them down
- Once you learn all the charts you will be in better shape to suggest different BI solutions.

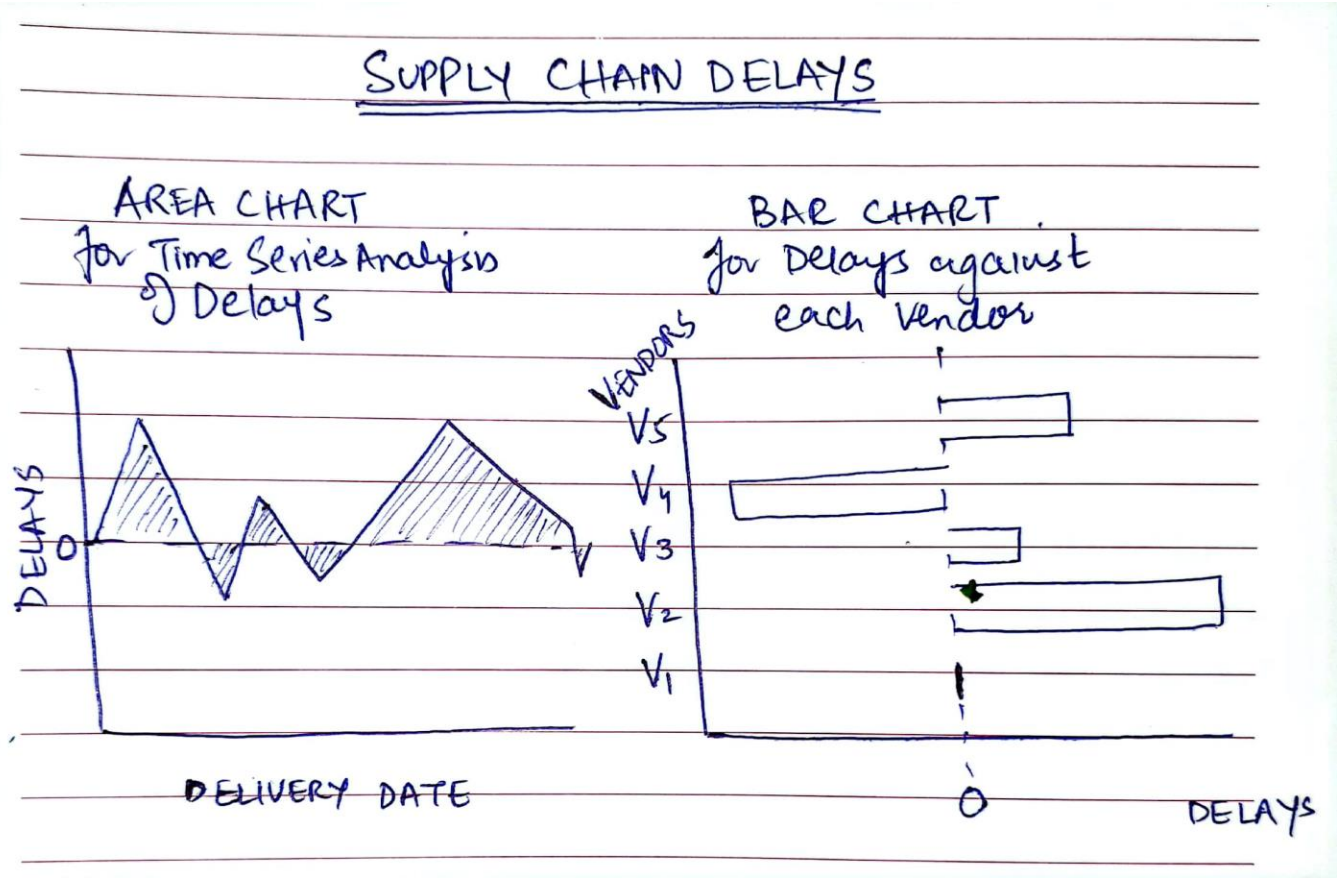
Prototype

Build BI Solutions Iteratively

BI Analysis


Storytelling on paper
Using the Tool for Analysis
Dashboards

Storytelling on Paper



Transform Data in Power BI

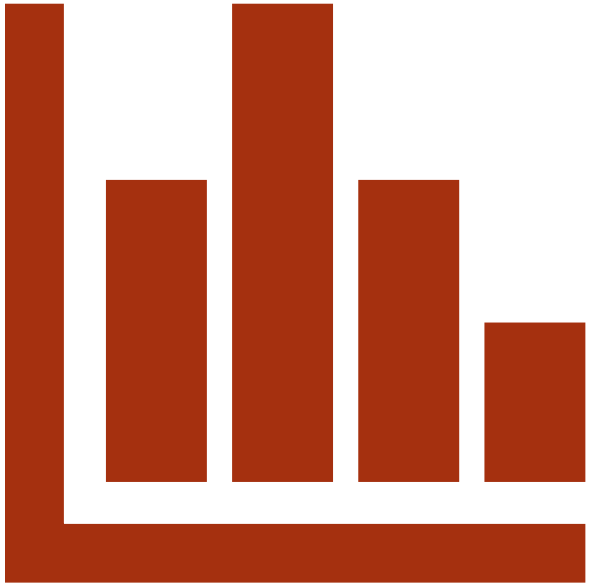
Several Options – Some examples

- Remove irrelevant columns
  - Column1,
  - Managed by, First Line Designation (same value, not useful for analysis)
- Change column names as needed – for better readability on charts and better understanding
- Create columns as required
  - E.g Create a manufacturing country column using the existing columns.

Identified Hierarchies

Create hierarchies in Power BI for easier accessibility for analysis.

H1	H2	H3	H4
Project code PQ# PO/SO#	Vendor Delivery Country Manuf. Country	Vendor Brand Product Group Sub classification Item Description	Dosage Form Dosage Molecule Test Type

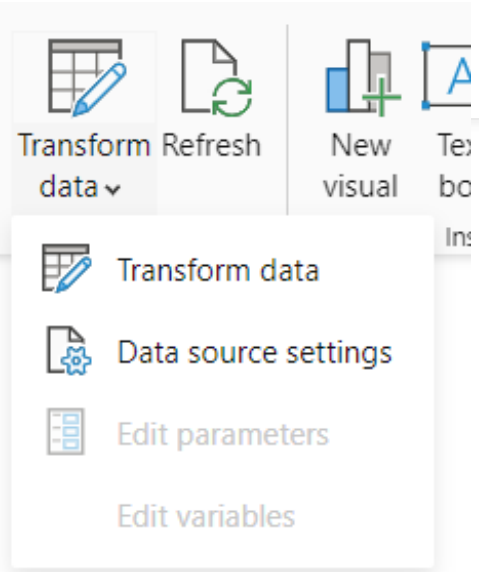


Begin Creating a Dashboard on Power BI

Test

Validate with Stakeholders

Transformations made



- Deleted useless columns 1,2,3
- Rename columns:
  - Delivered to Client Date → Delivery Date
  - Item Description → Item Desc
  - Others as required
- Full Via Replace “From RDC” to “RDC”
- Trimmed the text in Manuf. Country column