Supply Chain Shipment Pricing Data - Data Analysis and Modeling

Dataset Description: This data set provides supply chain health commodity shipment and pricing data.

File Descriptions

File Name :	**SCMS_Delivery_History_Dataset.csv**
File Size :	**Approx. 570kb**
Total Records :	**10,324**
File Updated :	**February 24, 2016**

Following are steps for Data Analysis and Modeling

- Import Packages
- Import CSV file
- Check Total Records in CSV file
- Check DataType of CSV file
- Print first 10 and last 10 recods from DataSet
- Total 10 Country wise count with graph
- Total Pack Price for Top 15 Countries with graph
- First Line Designation Wise Count with graph
- Shipment Mode percentage wise Pie Chart
- Unquie Manufacturing Site Names
- Shipment Mode, Min and Mean value for Air
- Top 10 Manufacturing Site for all Shipment Mode with Graph
- Top 10 Manufacturing Site for Air Shipment Mode with Graph
- Pack Price analysis using Distributions and Plot Graph
- Shipment Mode and Pack Price in Bar Plot Graph

Now, lets begin with Data!!!

Import Packages

In [1]:
```
import pandas as pd
import numpy as np
from decimal import Decimal
import matplotlib.pyplot as plt
%matplotlib inline
```

```
import seaborn as sn
import os
import plotly.graph_objs as go
import plotly.offline as py
py.init_notebook_mode(connected=True)
pd.options.mode.chained_assignment = None
pd.options.display.max_columns = 9999
pd.options.display.float_format = '{:20,.2f}'.format
```

Import CSV file

In [2]:
```
DataSet = pd.read_csv('SCMS_Delivery_History_Dataset.csv').fillna(0)
```

Check Total Records in CSV file

In [3]:
```
TotalRowCount = len(DataSet)
print("Total Number of Data Count :", TotalRowCount)
```

```
Total Number of Data Count : 10324
```

Check DataType of CSV file

In [4]:
```
DataSet.dtypes
```

Out[4]:
```
ID                                int64
Project Code                     object
PQ #                             object
PO / SO #                        object
ASN/DN #                         object
Country                          object
Managed By                       object
Fulfill Via                      object
Vendor INCO Term                 object
Shipment Mode                    object
PQ First Sent to Client Date     object
PO Sent to Vendor Date           object
Scheduled Delivery Date          object
Delivered to Client Date         object
Delivery Recorded Date           object
Product Group                    object
Sub Classification               object
Vendor                           object
Item Description                 object
Molecule/Test Type               object
Brand                            object
Dosage                           object
Dosage Form                      object
Unit of Measure (Per Pack)        int64
Line Item Quantity                int64
Line Item Value                 float64
Pack Price                      float64
Unit Price                      float64
Manufacturing Site               object
First Line Designation           object
Weight (Kilograms)               object
Freight Cost (USD)               object
Line Item Insurance (USD)       float64
dtype: object
```

Print first 10 and last 10 recods from DataSet

In [5]:
```
DataSet.head(10)
```

Out[5]:

	ID	Project Code	PQ #	PO / SO #	ASN/DN #	Country	Managed By	Fulfill Via	Vendor INCO Term	Shipment Mode	
0	1	100-CI-T01	Pre-PQ Process	SCMS-4	ASN-8	Côte d'Ivoire	PMO - US	Direct Drop	EXW	Air	F P
1	3	108-VN-T01	Pre-PQ Process	SCMS-13	ASN-85	Vietnam	PMO - US	Direct Drop	EXW	Air	F P
2	4	100-CI-T01	Pre-PQ Process	SCMS-20	ASN-14	Côte d'Ivoire	PMO - US	Direct Drop	FCA	Air	F P
3	15	108-VN-T01	Pre-PQ Process	SCMS-78	ASN-50	Vietnam	PMO - US	Direct Drop	EXW	Air	F P
4	16	108-VN-T01	Pre-PQ Process	SCMS-81	ASN-55	Vietnam	PMO - US	Direct Drop	EXW	Air	F P
5	23	112-NG-T01	Pre-PQ Process	SCMS-87	ASN-57	Nigeria	PMO - US	Direct Drop	EXW	Air	F P
6	44	110-ZM-T01	Pre-PQ Process	SCMS-139	ASN-130	Zambia	PMO - US	Direct Drop	DDU	Air	F P
7	45	109-TZ-T01	Pre-PQ Process	SCMS-140	ASN-94	Tanzania	PMO - US	Direct Drop	EXW	Air	F P
8	46	112-NG-T01	Pre-PQ Process	SCMS-156	ASN-93	Nigeria	PMO - US	Direct Drop	EXW	Air	F P
9	47	110-ZM-T01	Pre-PQ Process	SCMS-165	ASN-199	Zambia	PMO - US	Direct Drop	CIP	Air	F P

In [6]:
```
DataSet.tail(10)
```

Out[6]:

	ID	Project Code	PQ #	PO / SO #	ASN/DN #	Country	Managed By	Fulfill Via	Vendor INCO Term	Ship M
10314	86813	151-NG-T30	FPQ-14989	SO-51422	DN-4274	Nigeria	PMO - US	From RDC	N/A - From RDC	Ch
10315	86814	151-NG-T30	FPQ-14989	SO-51424	DN-4276	Nigeria	PMO - US	From RDC	N/A - From RDC	Ch
10316	86815	151-NG-T30	FPQ-16313	SO-51420	DN-4279	Nigeria	PMO - US	From RDC	N/A - From RDC	Ch
10317	86816	151-NG-T30	FPQ-16313	SO-51440	DN-4282	Nigeria	PMO - US	From RDC	N/A - From RDC	
10318	86817	103-ZW-T30	FPQ-15197	SO-50020	DN-4307	Zimbabwe	PMO - US	From RDC	N/A - From RDC	
10319	86818	103-ZW-T30	FPQ-15197	SO-50020	DN-4307	Zimbabwe	PMO - US	From RDC	N/A - From RDC	
10320	86819	104-CI-T30	FPQ-15259	SO-50102	DN-4313	Côte d'Ivoire	PMO - US	From RDC	N/A - From RDC	
10321	86821	110-ZM-T30	FPQ-14784	SO-49600	DN-4316	Zambia	PMO - US	From RDC	N/A - From RDC	
10322	86822	200-ZW-T30	FPQ-16523	SO-51680	DN-4334	Zimbabwe	PMO - US	From RDC	N/A - From RDC	
10323	86823	103-ZW-T30	FPQ-15197	SO-50022	DN-4336	Zimbabwe	PMO - US	From RDC	N/A - From RDC	

Total 10 Country wise count with graph

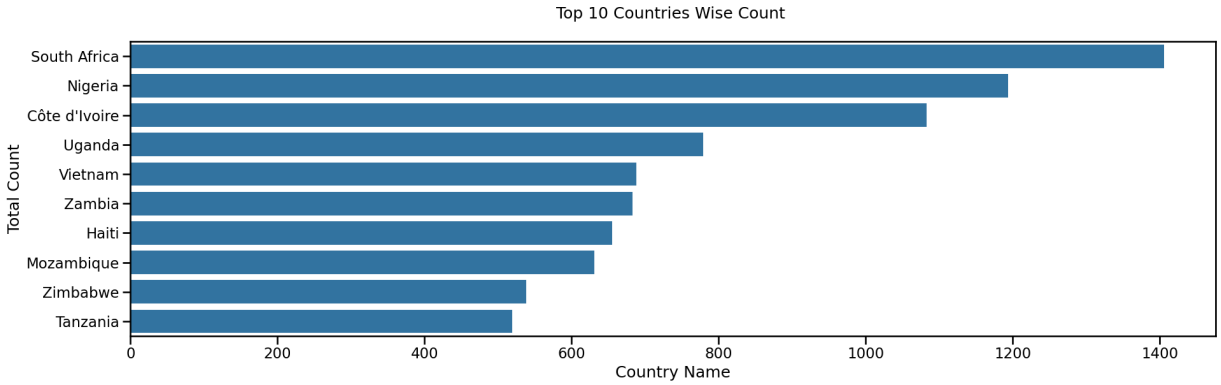
In [7]:
```
DataSet = DataSet.dropna()
ItemCount = DataSet["Country"].value_counts().nlargest(10)
print("Top 10 Countries Wise Count \n")
print(ItemCount)
sn.set_context("talk",font_scale=1)
```

```
plt.figure(figsize=(22,6))
sn.countplot(DataSet['Country'],order = DataSet['Country'].value_counts().nlargest(
plt.title('Top 10 Countries Wise Count \n')
plt.ylabel('Total Count')
plt.xlabel('Country Name')
```

```
Top 10 Countries Wise Count

Country
South Africa     1406
Nigeria          1194
Côte d'Ivoire    1083
Uganda            779
Vietnam           688
Zambia            683
Haiti             655
Mozambique        631
Zimbabwe          538
Tanzania          519
Name: count, dtype: int64
```

Out[7]: Text(0.5, 0, 'Country Name')



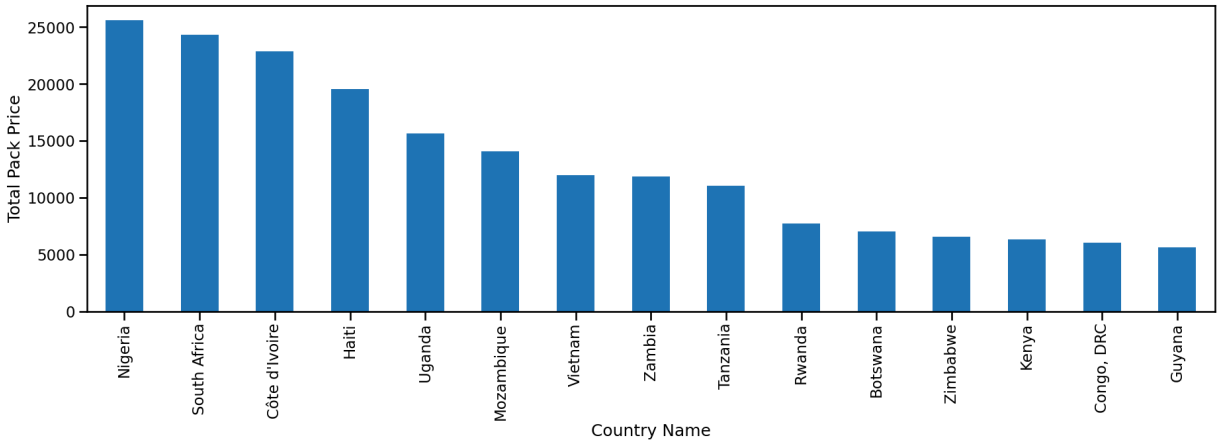
Total Pack Price for Top 15 Countries with graph

In [8]:
```
TotalPrice = DataSet.groupby(['Country'])['Pack Price'].sum().nlargest(15)
print("Total Pack Price for Top 15 Countries\n")
print(TotalPrice)
plt.figure(figsize=(22,6))
GraphData=DataSet.groupby(['Country'])['Pack Price'].sum().nlargest(15)
GraphData.plot(kind='bar')
plt.ylabel('Total Pack Price')
plt.xlabel('Country Name')
```

```
Total Pack Price for Top 15 Countries

Country
Nigeria                    25,620.72
South Africa               24,318.90
Côte d'Ivoire              22,882.35
Haiti                      19,521.30
Uganda                     15,664.08
Mozambique                 14,092.40
Vietnam                    11,963.78
Zambia                     11,889.67
Tanzania                   11,047.72
Rwanda                      7,724.29
Botswana                    7,052.54
Zimbabwe                    6,563.56
Kenya                       6,349.14
Congo, DRC                  6,032.41
Guyana                      5,659.35
Name: Pack Price, dtype: float64
```

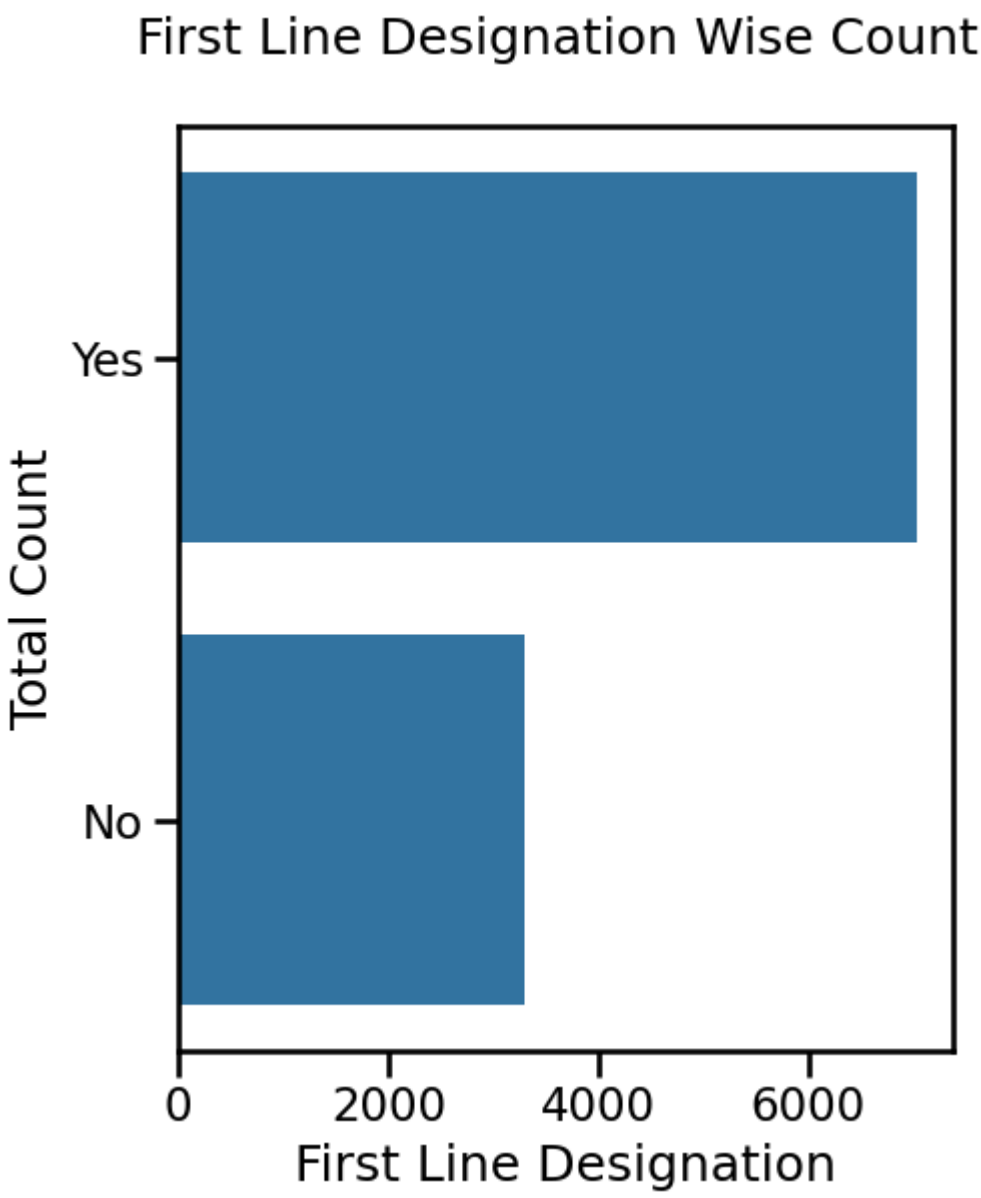
Out[8]: Text(0.5, 0, 'Country Name')



First Line Designation Wise Count

In [9]:
```
sn.set_context("talk",font_scale=1)
plt.figure(figsize=(5,6))
sn.countplot(DataSet['First Line Designation'],order = DataSet['First Line Designat
plt.title('First Line Designation Wise Count \n')
plt.ylabel('Total Count')
plt.xlabel('First Line Designation')
```

Out[9]: Text(0.5, 0, 'First Line Designation')

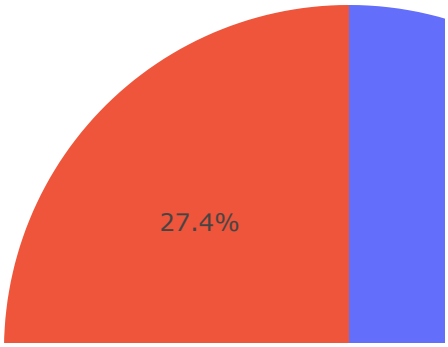


Shipment Mode percentage wise Pie Chart

In [10]:
```
ShippingMode = DataSet["Shipment Mode"].value_counts()
labels = (np.array(ShippingMode.index))
sizes = (np.array((ShippingMode / ShippingMode.sum())*100))

trace = go.Pie(labels=labels, values=sizes)
layout = go.Layout(title="Shipment Mode")
dat = [trace]
fig = go.Figure(data=dat, layout=layout)
py.iplot(fig, filename="Shipment Mode")
```

Shipment Mode



Unquie Manufacturing Site Names

In [11]:
```
UniqueItem = DataSet['Manufacturing Site'].unique()
print("All Unique Manufacturing Site \n")
print(UniqueItem)
```

```
All Unique Manufacturing Site

['Ranbaxy Fine Chemicals LTD' 'Aurobindo Unit III, India'
 'ABBVIE GmbH & Co.KG Wiesbaden' 'Ranbaxy, Paonta Shahib, India'
 'MSD South Granville Australia' "ABBVIE (Abbott) St. P'burg USA"
 'ABBVIE Ludwigshafen Germany' 'Trinity Biotech, Plc'
 'EY Laboratories, USA' 'Cipla, Goa, India' 'BMS Meymac, France'
 'Premier Med. Corp Ltd. India' 'Chembio Diagnostics Sys. Inc.'
 'Inverness Japan' 'Pacific Biotech, Thailand'
 'Standard Diagnostics, Korea' 'GSK Mississauga (Canada)'
 'Gilead(Nycomed) Oranienburg DE' 'Bio-Rad Laboratories'
 'Mylan (formerly Matrix) Nashik' 'Roche Basel' 'GSK Ware (UK)'
 'Novartis Pharma AG, Switzerland' 'BI, Ingelheim, Germany'
 'Not Applicable' 'Ipca Dadra/Nagar Haveli IN' 'MSD, Haarlem, NL'
 'Aspen-OSD, Port Elizabeth, SA' 'ABBVIE (Abbott) Logis. UK'
 'Gland Pharma Ltd Pally Factory' 'GSK Aranda'
 'Hetero Unit III Hyderabad IN' 'ABBVIE (Abbott) France'
 'Strides, Bangalore, India.' 'ABBSP' 'Cipla, Patalganga, India'
 'GSK Cape Town Factory (South Africa)' "MSD Midrand, J'burg, SA"
 'KHB Test Kit Facility, Shanghai China' 'bioLytical Laboratories'
 'Inverness USA' 'Boehringer Ing., Koropi, GR' 'GSK Crawley'
 'OMEGA Diagnostics, UK' 'Roche Madrid'
 'BUNDI INTERNATIONAL DIAGNOSTICS LTD' 'INVERNESS ORGENICS LINE'
 'Novartis Pharma Suffern, USA' 'Micro Labs Ltd. (Brown & Burk), India'
 'Meditab (for Cipla) Daman IN' 'Medopharm Malur Factory, INDIA'
 'ABBVIE (Abbott) Japan Co. Ltd.' 'MSD Elkton USA'
 'Orasure Technologies, Inc USA'
 'Weifa A.S., Hausmanngt. 6, P.O. Box 9113 GrÃ¸nland, 0133, Oslo, Norway'
 'MSD Manati, Puerto Rico, (USA)' 'MSD Patheon, Canada'
 'Emcure Plot No.P-2, I.T-B.T. Park, Phase II, MIDC, Hinjwadi, Pune, India'
 'Alere Medical Co., Ltd.' 'Premier Medical Corporation'
 'ABBVIE Labs North Chicago US' 'Janssen-Cilag, Latina, IT'
 'Aurobindo Unit VII, IN' 'Micro labs, Verna, Goa, India'
 'Orasure Technologies, Inc' 'Mylan,  H-12 & H-13, India'
 'Hetero, Jadcherla, unit 5, IN' 'Bristol-Myers Squibb Anagni IT'
 'Cipla Ltd A-42 MIDC Mahar. IN' 'Medochemie Factory A, CY'
 'Boehringer Ingelheim Roxane US' 'BMS Evansville, US'
 'GSK, U1, Poznan, Poland' 'Janssen Ortho LLC, Puerto Rico'
 'Micro Labs, Hosur, India' 'Remedica, Limassol, Cyprus' 'INVERNESS ANY'
 'Guilin OSD site, No 17, China' 'Cipla, Kurkumbh, India'
 'GSK Barnard Castle UK' 'Gland Pharma, Hyderabad, IN' 'Access BIO, L.C.'
 'Human Diagnostic' 'Mepro Pharm Wadhwan Unit II'
 'Ranbaxy per Shasun Pharma Ltd' 'MedMira Inc.'
 'Ranbaxy per Shasun Pharma' 'Macleods Daman Plant INDIA']
```

Shipment Mode, Min and Mean value for Air

In [12]:
```
ItemData=DataSet[DataSet['Shipment Mode']=='Air']
print ("The Max Air Shipment Mode is :",ItemData['Unit of Measure (Per Pack)'].max(
print ("The Min Air Shipment is :",ItemData['Unit of Measure (Per Pack)'].min())
ItemTypeMean = ItemData['Unit of Measure (Per Pack)'].mean()
print ("The Mean Air Shipment is :", round(ItemTypeMean,2))
```

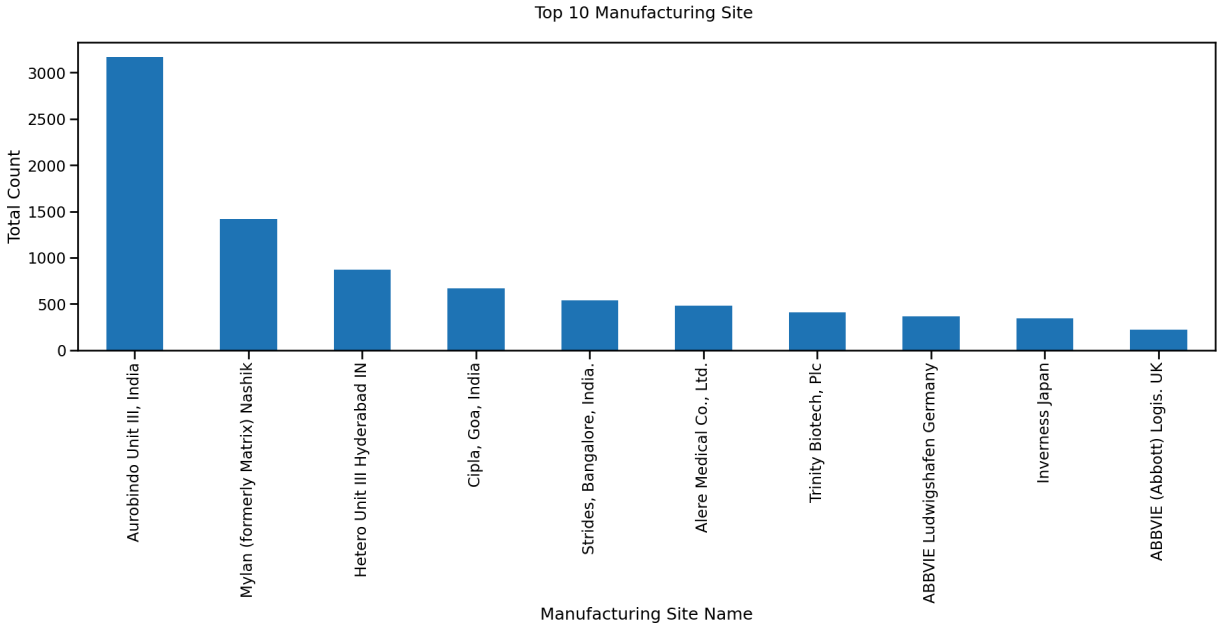
```
The Max Air Shipment Mode is : 1000
The Min Air Shipment is : 1
The Mean Air Shipment is : 82.34
```

Top 10 Manufacturing Site for all Shipment Mode with Graph

In [13]:
```
plt.figure(figsize=(22,6))
TopFiveManufacturingSite=DataSet.groupby('Manufacturing Site').size().nlargest(10)
print(TopFiveManufacturingSite)
TopFiveManufacturingSite.plot(kind='bar')
plt.title('Top 10 Manufacturing Site \n')
plt.ylabel('Total Count')
plt.xlabel('Manufacturing Site Name')
```

```
Manufacturing Site
Aurobindo Unit III, India         3172
Mylan (formerly Matrix) Nashik    1415
Hetero Unit III Hyderabad IN       869
Cipla, Goa, India                  665
Strides, Bangalore, India.         540
Alere Medical Co., Ltd.            481
Trinity Biotech, Plc               405
ABBVIE Ludwigshafen Germany        366
Inverness Japan                    345
ABBVIE (Abbott) Logis. UK          219
dtype: int64
```

Out[13]: Text(0.5, 0, 'Manufacturing Site Name')



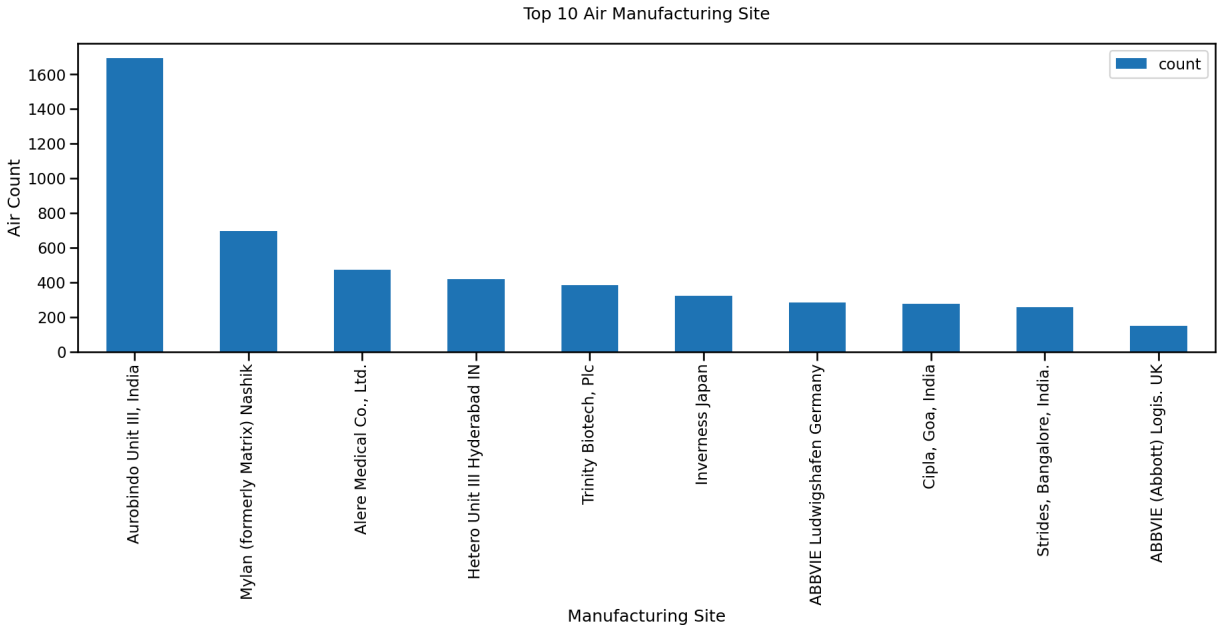
Top 10 Manufacturing Site for Air Shipment Mode with Graph

In [14]:
```
# Top 10 Air Shipment Mode in Bar Chart
ItemData=DataSet[DataSet['Shipment Mode']=='Air']
DataSet[DataSet["Shipment Mode"]=='Air']['Manufacturing Site'].value_counts()[0:10]
ItemSupplier = DataSet[DataSet["Shipment Mode"]=='Air']['Manufacturing Site'].value
print("Top 10 Air Manufacturing Site \n")
print(ItemSupplier)
plt.title('Top 10 Air Manufacturing Site\n')
plt.ylabel('Air Count')
plt.xlabel('Manufacturing Site')
```

```
Top 10 Air Manufacturing Site

Manufacturing Site
Aurobindo Unit III, India         1694
Mylan (formerly Matrix) Nashik     695
Alere Medical Co., Ltd.            473
Hetero Unit III Hyderabad IN       417
Trinity Biotech, Plc               383
Inverness Japan                    320
ABBVIE Ludwigshafen Germany        285
Cipla, Goa, India                  274
Strides, Bangalore, India.         256
ABBVIE (Abbott) Logis. UK          150
Name: count, dtype: int64
```

Out[14]:
```
Text(0.5, 0, 'Manufacturing Site')
```



Conclusion

- Top Country for Pack Price : Nigeria - 25,620.72
- Top Shipping Mode : Air
- The Max Air Shipment Mode is : 1000
- The Min Air Shipment is : 1
- The Mean Air Shipment is : 82.35

- Top Manufacturing Site : Aurobindo Unit III, India - 3172
- Top Air Manufacturing Site : Aurobindo Unit III, India - 1694