```
import pandas as pd
```

In [2]:
```
dataprod.read_excel("C:\Users\KULKY\Downloads\Financials Sample Data.xlsx")
data
```

Out[2]:

	Segment	Country	Product	Units Sold	Manufacturing Price	Sale Price	Gross Sales	Discounts	Sales	COGS	Profit	Date	Month Number	Month Name	Year
0	Government	France	Carretera	1618.5	3	350	5670000	NaN	5670000	5292000	378000	2014-01-01 00:00:00	1	January	2014
1	Government	Germany	Carretera	1321	3	350	4626500	NaN	4626500	4318200	308300	2014-01-01 00:00:00	1	January	2014
2	Midmarket	Mexico	Carretera	2178	3	350	7623000	NaN	7623000	7114800	508200	2014-06-01 00:00:00	6	June	2014
3	Midmarket	Germany	Carretera	888	3	350	3108000	NaN	3108000	2901600	206400	2014-06-01 00:00:00	6	June	2014
4	Midmarket	France	Carretera	2470	3	350	8645000	NaN	8645000	8070000	575000	2014-06-01 00:00:00	6	June	2014
5	Enterprise	Canada	Carretera	1513	3	350	5295500	NaN	5295500	4943400	352100	2014-12-01 00:00:00	12	December	2014
6	Small Business	Mexico	Carretera	921	3	350	3223500	NaN	3223500	3009060	214440	2015-03-01 00:00:00	3	March	2015
7	Government	France	Montana	2518	5	300	7554000	NaN	7554000	6797400	756600	2013-07-01 00:00:00	7	July	2013
8	Enterprise	Canada	Montana	1899	5	300	5697000	NaN	5697000	5127300	569700	2013-08-01 00:00:00	8	August	2013
9	Small Business	Germany	Montana	1545	5	300	4635000	NaN	4635000	4171500	463500	2013-09-01 00:00:00	9	September	2013
10	Midmarket	France	Montana	2470	5	300	7410000	NaN	7410000	6669000	741000	2013-12-01 00:00:00	12	December	2013
11	Government	Germany	Montana	1893	5	300	5679000	NaN	5679000	5111100	567900	2014-03-01 00:00:00	3	March	2014
12	Enterprise	Canada	Montana	1375	5	300	4125000	NaN	4125000	3712500	412500	2014-04-01 00:00:00	4	April	2014
13	Small Business	Mexico	Montana	1492	5	300	4476000	NaN	4476000	4028400	447600	2014-05-01 00:00:00	5	May	2014
14	Government	France	Paseo	1804	10	350	6314000	NaN	6314000	5682600	631400	2014-07-01 00:00:00	7	July	2014
15	Enterprise	Canada	Paseo	1731	10	350	6058500	NaN	6058500	5452650	605850	2014-08-01 00:00:00	8	August	2014
16	Small Business	Mexico	Paseo	1799	10	350	6296500	NaN	6296500	5666850	629650	2014-09-01 00:00:00	9	September	2014
17	Midmarket	Germany	Paseo	1246	10	350	4361000	NaN	4361000	3924900	436100	2014-12-01 00:00:00	12	December	2014
18	Government	France	VTT	1098	250	300	329400	NaN	329400	296460	32940	2013-02-01 00:00:00	2	February	2013
19	Enterprise	Canada	VTT	1267	250	300	380100	NaN	380100	342090	38010	2013-03-01 00:00:00	3	March	2013
20	Small Business	Mexico	VTT	1294	250	300	388200	NaN	388200	349380	38820	2013-04-01 00:00:00	4	April	2013
21	Midmarket	Germany	VTT	1012	250	300	303600	NaN	303600	273240	30360	2013-07-01 00:00:00	7	July	2013
22	Enterprise	France	VTT	1805	250	300	541500	NaN	541500	487350	54150	2013-08-01 00:00:00	8	August	2013
23	Small Business	Canada	VTT	1314	250	300	394200	NaN	394200	354780	39420	2013-09-01 00:00:00	9	September	2013
24	Midmarket	Mexico	VTT	1773	250	300	531900	NaN	531900	478710	53190	2013-12-01 00:00:00	12	December	2013
25	Government	Germany	Velo	1620	350	300	486000	NaN	486000	437400	48600	2014-03-01 00:00:00	3	March	2014
26	Enterprise	France	Velo	1376	350	300	412800	NaN	412800	371520	41280	2014-04-01 00:00:00	4	April	2014
27	Small Business	Canada	Velo	1494	350	300	448200	NaN	448200	403380	44820	2014-05-01 00:00:00	5	May	2014
28	Midmarket	Mexico	Velo	1843	350	300	552900	NaN	552900	497610	55290	2014-08-01 00:00:00	8	August	2014
29	Government	France	Amarilla	1773	260	300	531900	NaN	531900	478710	53190	2014-09-01 00:00:00	9	September	2014
30	Enterprise	Canada	Amarilla	1377	260	300	413100	NaN	413100	371790	41310	2014-12-01 00:00:00	12	December	2014
31	Small Business	Mexico	Amarilla	1444	260	300	433200	NaN	433200	389880	43320	2015-03-01 00:00:00	3	March	2015
32	Midmarket	Germany	Amarilla	1004	260	300	301200	NaN	301200	271080	30120	2015-06-01 00:00:00	6	June	2015
33	Government	France	Montana	1636	5	300	4908000	NaN	4908000	4417200	490800	2013-07-01 00:00:00	7	July	2013
34	Enterprise	Canada	Montana	1884	5	300	5652000	NaN	5652000	5086800	565200	2013-08-01 00:00:00	8	August	2013
35	Small Business	Germany	Montana	1610	5	300	4830000	NaN	4830000	4347000	483000	2013-09-01 00:00:00	9	September	2013
36	Midmarket	France	Montana	2705	5	300	8115000	NaN	8115000	7303500	811500	2013-12-01 00:00:00	12	December	2013
37	Government	Germany	Montana	1943	5	300	5829000	NaN	5829000	5246100	582900	2014-03-01 00:00:00	3	March	2014
38	Enterprise	Canada	Montana	1387	5	300	4161000	NaN	4161000	3744900	416100	2014-04-01 00:00:00	4	April	2014
39	Small Business	Mexico	Montana	1505	5	300	4515000	NaN	4515000	4063500	451500	2014-05-01 00:00:00	5	May	2014
40	Midmarket	France	Montana	1828	5	300	5484000	NaN	5484000	4935600	548400	2014-08-01 00:00:00	8	August	2014
41	Government	Germany	Paseo	1799	10	350	6296500	NaN	6296500	5666850	629650	2014-09-01 00:00:00	9	September	2014
42	Enterprise	Canada	Paseo	1377	10	350	4819500	NaN	4819500	4337550	481950	2014-12-01 00:00:00	12	December	2014
43	Small Business	Mexico	Paseo	1444	10	350	5054000	NaN	5054000	4548600	505400	2015-03-01 00:00:00	3	March	2015
44	Midmarket	Germany	Paseo	1004	10	350	3514000	NaN	3514000	3162600	351400	2015-06-01 00:00:00	6	June	2015
45	Government	France	VTT	1098	250	300	329400	NaN	329400	296460	32940	2013-02-01 00:00:00	2	February	2013
46	Enterprise	Canada	VTT	1267	250	300	380100	NaN	380100	342090	38010	2013-03-01 00:00:00	3	March	2013
47	Small Business	Mexico	VTT	1294	250	300	388200	NaN	388200	349380	38820	2013-04-01 00:00:00	4	April	2013
48	Midmarket	Germany	VTT	1012	250	300	303600	NaN	303600	273240	30360	2013-07-01 00:00:00	7	July	2013
49	Enterprise	France	VTT	1805	250	300	541500	NaN	541500	487350	54150	2013-08-01 00:00:00	8	August	2013
50	Small Business	Canada	VTT	1314	250	300	394200	NaN	394200	354780	39420	2013-09-01 00:00:00	9	September	2013
51	Midmarket	Mexico	VTT	1773	250	300	531900	NaN	531900	478710	53190	2013-12-01 00:00:00	12	December	2013
52	Government	Germany	Velo	1620	350	300	486000	NaN	486000	437400	48600	2014-03-01 00:00:00	3	March	2014
53	Enterprise	France	Velo	1376	350	300	412800	NaN	412800	371520	41280	2014-04-01 00:00:00	4	April	2014
54	Small Business	Canada	Velo	1494	350	300	448200	NaN	448200	403380	44820	2014-05-01 00:00:00	5	May	2014
55	Midmarket	Mexico	Velo	1843	350	300	552900	NaN	552900	497610	55290	2014-08-01 00:00:00	8	August	2014
56	Government	France	Amarilla	1773	260	300	531900	NaN	531900	478710	53190	2014-09-01 00:00:00	9	September	2014
57	Enterprise	Canada	Amarilla	1377	260	300	413100	NaN	413100	371790	41310	2014-12-01 00:00:00	12	December	2014
58	Small Business	Mexico	Amarilla	1444	260	300	433200	NaN	433200	389880	43320	2015-03-01 00:00:00	3	March	2015
59	Midmarket	Germany	Amarilla	1004	260	300	301200	NaN	301200	271080	30120	2015-06-01 00:00:00	6	June	2015

700 rows × 17 columns

In [3]:
```
data.describe()
```

Out[3]:

	Units Sold	Manufacturing Price	Sale Price	Gross Sales	Discounts	Sales	COGS	Profit	Month Number
count	700.000000	700.000000	700.000000	700.000000	700.000000	700.000000	700.000000	700.000000	700.000000
mean	1611.718571	3.395714	349.857143	5608470.000000	75494.285714	5532975.714286	4661275.714286	871700.000000	8.958571
std	731.062368	2.351327	107.261074	2345908.262828	44267.140966	2299382.857360	1914467.948283	384571.237961	3.267915
min	200.000000	1.000000	120.000000	300000.000000	1500.000000	293100.000000	242100.000000	39000.000000	1.000000
25%	1043.000000	3.000000	250.000000	3883000.000000	44865.000000	3804375.000000	3149400.000000	607890.000000	7.000000
50%	1507.000000	5.000000	300.000000	5320000.000000	74070.000000	5235000.000000	4366500.000000	842000.000000	9.000000
75%	2099.000000	5.000000	350.000000	7035000.000000	104805.000000	6932625.000000	5827500.000000	1099250.000000	11.000000
max	2816.000000	10.000000	700.000000	12096000.000000	162375.000000	12033375.000000	10290000.000000	1743375.000000	12.000000

In [4]:
```
data.head()
```

Out[4]:

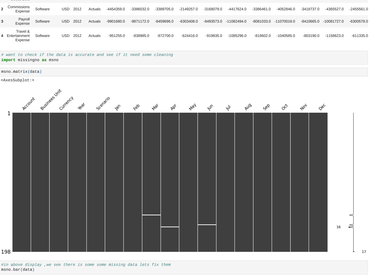
	Segment	Country	Product	Units Sold	Manufacturing Price	Sale Price	Gross Sales	Discounts	Sales	COGS	Profit	Date	Month Number	Month Name	Year
0	Government	France	Carretera	1618.5	3	350	5670000	NaN	5670000	5292000	378000	2014-01-01 00:00:00	1	January	2014
1	Government	Germany	Carretera	1321	3	350	4626500	NaN	4626500	4318200	308300	2014-01-01 00:00:00	1	January	2014
2	Midmarket	Mexico	Carretera	2178	3	350	7623000	NaN	7623000	7114800	508200	2014-06-01 00:00:00	6	June	2014
3	Midmarket	Germany	Carretera	888	3	350	3108000	NaN	3108000	2901600	206400	2014-06-01 00:00:00	6	June	2014
4	Midmarket	France	Carretera	2470	3	350	8645000	NaN	8645000	8070000	575000	2014-06-01 00:00:00	6	June	2014

```
# want to check if the data is accurate and see if it need some cleaning
import missingno as msno
```

In [6]:
```
msno.matrix(data)
```

•AxesSubplot:

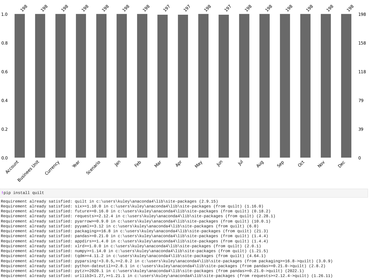
Out[6]:



In [7]:
```
#in above display ,we see there is some some missing data lets fix them
msno.bar(data)
```

•AxesSubplot:

Out[7]:



In [8]:
```
!pip install quilt
```

```
Requirement already satisfiel: quilt in c:\users\kulky\anaconda4\lib\site-packages (2.9.15)
Requirement already satisfiel: requests in c:\users\kulky\anaconda4\lib\site-packages (from quilt) (0.10.0)
Requirement already satisfiel: future>=0.16.0 in c:\users\kulky\anaconda4\lib\site-packages (from quilt) (0.16.2)
Requirement already satisfiel: requests>=2.12.4 in c:\users\kulky\anaconda4\lib\site-packages (from quilt) (2.20.1)
Requirement already satisfiel: pyyaml>=3.12 in c:\users\kulky\anaconda4\lib\site-packages (from quilt) (3.13)
Requirement already satisfiel: pyyaml>=3.12 in c:\users\kulky\anaconda4\lib\site-packages (from quilt) (3.13)
Requirement already satisfiel: packaging>=16.8 in c:\users\kulky\anaconda4\lib\site-packages (from quilt) (21.3)
Requirement already satisfiel: pandas>=0.21.0 in c:\users\kulky\anaconda4\lib\site-packages (from quilt) (1.4.4)
Requirement already satisfiel: appdirs>=1.4.0 in c:\users\kulky\anaconda4\lib\site-packages (from quilt) (1.4.4)
Requirement already satisfiel: requests>=2.12.4 in c:\users\kulky\anaconda4\lib\site-packages (from quilt) (2.28.1)
Requirement already satisfiel: numpy>=1.14.0 in c:\users\kulky\anaconda4\lib\site-packages (from quilt) (1.21.5)
Requirement already satisfiel: tqdm>=4.11.2 in c:\users\kulky\anaconda4\lib\site-packages (from quilt) (4.66.1)
Requirement already satisfiel: requests-kerberos>=0.12.0,<0.12.1 in c:\users\kulky\anaconda4\lib\site-packages (from quilt) (0.12.0)
Requirement already satisfiel: requests-aws4auth>=1.1.1,<1.1.2 in c:\users\kulky\anaconda4\lib\site-packages (from quilt) (1.1.2)
Requirement already satisfiel: requests-aws4auth>=1.1.1,<1.1.2 in c:\users\kulky\anaconda4\lib\site-packages (from quilt) (1.1.2)
Requirement already satisfiel: boto3>=1.9.184,<1.9.185 in c:\users\kulky\anaconda4\lib\site-packages (from quilt) (1.9.184)
Requirement already satisfiel: boto3>=1.9.184,<1.9.185 in c:\users\kulky\anaconda4\lib\site-packages (from quilt) (1.9.184)
Requirement already satisfiel: pyjwt>=2020.1 in c:\users\kulky\anaconda4\lib\site-packages (from pandas>=0.21.0->quilt) (2022.1)
Requirement already satisfiel: pyjwt>=2020.1 in c:\users\kulky\anaconda4\lib\site-packages (from pandas>=0.21.0->quilt) (2022.1)
Requirement already satisfiel: pytz>=2020.1 in c:\users\kulky\anaconda4\lib\site-packages (from pandas>=0.21.0->quilt) (2022.1)
Requirement already satisfiel: certifi>=2017.4.17 in c:\users\kulky\anaconda4\lib\site-packages (from requests>=2.12.4->quilt) (2022.9.14)
Requirement already satisfiel: idna>4,<5.3 in c:\users\kulky\anaconda4\lib\site-packages (from requests>=2.12.4->quilt) (3.3)
Requirement already satisfiel: idna>4,<5.3 in c:\users\kulky\anaconda4\lib\site-packages (from requests>=2.12.4->quilt) (2.8.4)
Requirement already satisfiel: colorama in c:\users\kulky\anaconda4\lib\site-packages (from tqdm>=4.11.2->quilt) (0.4.5)
```

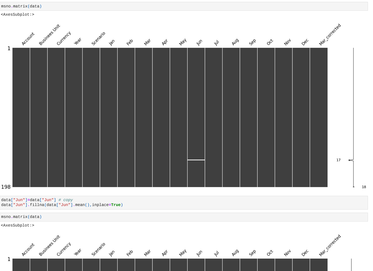
In [9]:
```
from quilt.data.ResidentMario import missingno,data
```

In [10]:
```
# I may just replace the missing data with the mean which make some sense
data["Mar"].fillna(data["Mar"].mean(),inplace=True)
```

In [11]:
```
msno.matrix(data)
```

•AxesSubplot:

Out[11]:

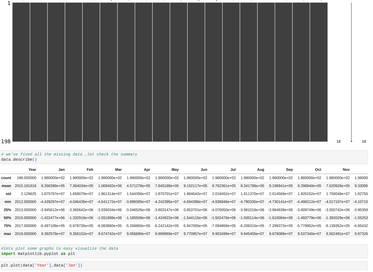


In [12]:
```
# lets replace all the missing data
data["Apr"].fillna(data["Apr"].mean(),inplace=True)
data["May"].fillna(data["May"].mean(),inplace=True)
```

In [13]:
```
msno.matrix(data)
```

•AxesSubplot:

Out[13]:



In [14]:
```
# we've fixed all the missing data ,let check the summary
data.describe()
```

Out[14]:

	Units Sold	Manufacturing Price	Sale Price	Gross Sales	Discounts	Sales	COGS	Profit	Month Number
count	199.000000	199.000000	199.000000	199.000000	199.000000	199.000000	199.000000	199.000000	199.000000
mean	2015.181818	3.262688e-05	7.364664e-05	1.094624e-04	4.571279e-05	7.645188e-05	9.152117e-05	6.762814e-05	5.341796e-05
std	2.129625	1.975767e-07	1.890875e-07	1.962314e-07	1.544258e-07	1.970701e-07	1.990804e-07	2.010462e-07	1.911370e-07
min	2012.000000	-4.426097e-07	-4.044054e-07	-4.641215e-07	3.000000e-07	-4.242599e-07	-4.004889e-07	-4.790200e-07	-4.738141e-07
25%	2013.000000	3.946513e-06	3.306043e-06	3.930164e-06	3.040526e-06	3.831147e-06	3.853703e-06	4.170503e-06	3.961518e-06
50%	2015.000000	5.423747e-06	5.325063e-06	5.559811e-06	1.089869e-06	5.434823e-06	5.546129e-06	5.552767e-06	5.595114e-06
75%	2017.000000	6.497108e-06	5.670736e-06	6.156430e-06	5.301500e-06	6.242142e-06	5.965316e-06	7.094866e-06	6.280156e-06
max	2018.000000	9.382578e-07	9.282152e-07	9.674742e-07	9.258886e-07	8.888989e-07	9.770857e-07	9.953496e-07	9.845456e-07

```
# store plot some graph to easy visualize the data
plot = matplotlib.pyplot.ms.plot
```

In [15]:
```
plt.plot(data["Year"],data["Jan"])
```

Out[15]:

```
[matplotlib.lines.Line2D at 0x24f9835f2f0b]
```

In [16]:
```
plt.hist(data["Jan"])
```

Out[16]:

```
(array([  9.,   0.,   0.,  198.,   0.,   0.,   0.,   0.,   0.,   1.],
       [  0
```