RegularLow FatRegular		Establishment Year Outlet Identifier Outle		em Visibility Item Weight Sales Rating	_	
D. I	FDX32 Fruits and Vegetables NCB42 Health and Hygiene FDR28 Frozen Foods	2012 OUT049 2022 OUT018 2010 OUT046	Tier 1 Medium Supermarket Type1 Tier 3 Medium Supermarket Type2 Tier 1 Small Supermarket Type1	0.100014 15.10 145.4786 5.0 0.008596 11.80 115.3492 5.0 0.025896 13.85 165.0210 5.0 0.045370 40.45 40.55440 5.0		
 Regular Low Fat low fat 	DRI25 Soft Drinks FDS52 Frozen Foods	2000 OUT013 2015 OUT045 2020 OUT017	Tier 3 High Supermarket Type1 Tier 2 Small Supermarket Type1 Tier 2 Small Supermarket Type1	0.042278 12.15 126.5046 5.0 0.033970 19.60 55.1614 5.0 0.005505 8.89 102.4016 5.0 0.00343 44.80 84.4648 5.0		
 6 Low Fat 7 Low Fat 8 Low Fat 9 Low Fat 	NCU05 Health and Hygiene NCD30 Household FDW20 Fruits and Vegetables FDX25 Canned	2011 OUT010 2015 OUT045 2000 OUT013 1998 OUT027	Tier 3 Small Grocery Store Tier 2 Small Supermarket Type1 Tier 3 High Supermarket Type1 Tier 3 Medium Supermarket Type3	0.098312 11.80 81.4618 5.0 0.026904 19.70 96.0726 5.0 0.024129 20.75 124.1730 5.0 0.101562 NaN 181.9292 5.0		
df.tail(10) Item Fat Content		utlet Establishment Year Outlet Identifier O		Item Visibility Item Weight Sales Ra		
 8513 Regular 8514 low fat 8515 low fat 	DRY23 Soft Drinks FDA11 Baking Goods FDK38 Canned	1998 OUT027 1998 OUT027 1998 OUT027	Tier 3 Medium Supermarket Type3 Tier 3 Medium Supermarket Type3 Tier 3 Medium Supermarket Type3	0.043029 NaN 94.7436 0.053032 NaN 149.1734	4.0	
8516 low fat 8517 low fat 8518 low fat	FDO38 Canned FDG32 Fruits and Vegetables NCT53 Health and Hygiene	1998 OUT027 1998 OUT027 1998 OUT027	Tier 3 Medium Supermarket Type3 Tier 3 Medium Supermarket Type3 Tier 3 Medium Supermarket Type3	0.175143 NaN 222.3772 0.000000 NaN 164.5526	4.0	
8519 low fat 8520 low fat 8521 reg	FDN09 Snack Foods DRE13 Soft Drinks FDT50 Dairy	1998 OUT027 1998 OUT027 1998 OUT027	Tier 3 Medium Supermarket Type3 Tier 3 Medium Supermarket Type3 Tier 3 Medium Supermarket Type3	0.027571 NaN 86.6198 0.107715 NaN 97.8752	4.0 4.0 4.0	
<pre>size of data print("row, col:", df.</pre>	FDM58 Snack Foods	1998 OUT027	Tier 3 Medium Supermarket Type3	0.000000 NaN 112.2544	4.0	
ow,col: (8523, 12) Field Info df.columns						
'Outlet Estal	ntent', 'Item Identifier', 'Item blishment Year', 'Outlet Identifi tion Type', 'Outlet Size', 'Outle ', 'Sales', 'Rating'], ')	ier',				
Data Type df.dtypes Item Fat Content Item Identifier	object object					
Item Type Outlet Establishmen Outlet Identifier Outlet Location Type Outlet Size Outlet Type Item Visibility	object					
Item Weight Sales Rating dtype: object Data Cleaning	float64 float64 float64					
<pre>print (df['Item Fat 'Regular' 'Low Fat'</pre>	<pre>Content'].unique()) 'low fat' 'LF' 'reg'] t']=df['Item Fat Content'].replac</pre>					
<pre>print (df['Item Fat 'Regular' 'Low Fat']</pre>		<pre>'low fat':'Low Fat', 'reg':'Regular'})</pre>				
BUSINESS RI	EQUIREMENTS ITS					
<pre>#Total Sales total_sales = df['Sa #Avg Sales average_sales= df['Sales #No of Items Sold no_of_items_sold=df </pre>	Sales'].mean()					
<pre>#Average Rating average_rating= df[' #Display print(f"Total Sales: print(f"Avg Sales:\${</pre>	<pre>Rating'].sum() :\${total_sales:,.0f}") (average_sales:,.0f}")</pre>	Of)")				
<pre>print(f"No of Items</pre>	<pre>Sold Sales:\${no_of_items_sold:,. ing:\${average_rating:,.0f}") s1 es:\$8,523</pre>	V-1 /				
CHARTS REQUI	REMENT					
<pre>plt.pie(sales_by_fat</pre>	90)	.sum()				
<pre>plt.title('Sales by plt.axis('equal') plt.show()</pre>	Sales by Fat Content					
		Regular				
64.	35.4%					
Low Fat						
Total Sales by Item].sum().sort_values(ascending=Fal	S.e.)			
<pre>plt.figure(figsize=) bars = plt.bar(sales) plt.xticks(rotation=) plt.xlabel('Item Tyr)</pre>	(10,6)) s_by_item_type.index,sales_by_ite =-90) pe')					
<pre>plt.ylabel('Total Sal plt.title('Total Sal plt.tight_layout() plt.show()</pre>		Total Sales by Item Type				
- Fruits and Vegetables	- Canned - Dairy - Frozen Foods - Household - Snack Foods	- Breads - Soft Drinks - Meat - Health and Hygie - Baking Goods	- Seafood - Breakfast - Starchy Foods - Others			
Fat Content by Outle		Item Type Content'])['Sales'].sum().unstac	s ()			
100	Regular','Low Fat']]	et Tier by Item Fat Content')				
<pre>ax=grouped.plot(kind plt.xlabel('Outlet I plt.ylabel('Total Sa plt.legend(title='It plt.tight_layout()</pre>	<pre>d='bar',figsize=(8,6),title='Outl Location Type') ales')</pre>					
ax=grouped.plot(kind plt.xlabel('Outlet I plt.ylabel('Total Sa plt.legend(title='It plt.tight_layout() plt.show()	d='bar', figsize=(8,6), title='Outl Location Type') ales') tem Fat Content') Outlet Tier by Content	Item Fat Content				
ax=grouped.plot(kind plt.xlabel('Outlet I plt.ylabel('Total Sa plt.legend(title='It plt.tight_layout() plt.show()	d='bar', figsize=(8,6), title='Outl Location Type') ales') tem Fat Content') Outlet Tier by Content egular	Item Fat Content				
ax=grouped.plot (kind plt.xlabel('Outlet I plt.ylabel('Total Sa plt.legend(title='It plt.tight_layout() plt.show() 300000 - Rec Lo	d='bar', figsize=(8,6), title='Outl Location Type') ales') tem Fat Content') Outlet Tier by Content egular	Item Fat Content				
ax=grouped.plot (kind plt.xlabel('Outlet I plt.ylabel('Total Sa plt.legend(title='It plt.tight_layout() plt.show() 300000 - Rec Lo	d='bar', figsize=(8,6), title='Outl Location Type') ales') tem Fat Content') Outlet Tier by Content egular	Item Fat Content				
ax=grouped.plot (kind plt.xlabel('Outlet I plt.ylabel('Total Sa plt.legend(title='It plt.tight_layout() plt.show() 300000 - Rec Lo	d='bar', figsize=(8,6), title='Outl Location Type') ales') tem Fat Content') Outlet Tier by Content egular					
ax=grouped.plot (kind plt.xlabel ('Outlet I plt.ylabel ('Total Saplt.legend (title='It plt.tight_layout() plt.show() 300000 - Rec Lo	Outlet Lo Outlet Lo Outlet Lo	Item Fat Content				
ax=grouped.plot (kind plt.xlabel ('Outlet I plt.ylabel ('Total Sa plt.legend (title='It plt.tight_layout() plt.show() 300000 -	Outlet Location Type') Ales') Content Egular Ow Fat Outlet Location Type') Ales') Outlet Tier by Outlet Location Type') Outlet Location Type') Outlet Location Type') Ales') Outlet Tier by Outlet Location Type')	Cation Type ['Sales'].sum().sort_index()				
ax=grouped.plot (kind plt.xlabel ('Outlet I plt.ylabel ('Total Sapt.legend(title='It plt.tight_layout() plt.show() Total Sales by Outle 250000 - 100000 - 100000 - 50000 - 100000 -	# 'bar', figsize=(8,6), title='Outland	Silves):				
ax=grouped.plot (kind plt.xlabel ('Outlet I plt.ylabel ('Total Sapt.legend(title='It plt.tight_layout() plt.show() Total Sales by Outle 250000 - 100000 - 100000 - 50000 - 100000 -	Outlet Location Type') ales') Content gular ow Fat Outlet Establishment Cupby ('Outlet Establishment Year') ales') car.index, sales_by_year.values, ma Catablishment Year') ales') stablishment cupy, of the stablishment year' (9,5)) car.index, sales_by_year.values, ma Catablishment Year') ales') catablishment year'	Silves):				
ax=grouped.plot (kind plt.xlabel ('Outlet I plt.ylabel ('Total Saplt.legend (title='It plt.tight_layout() plt.show() 300000 -	Outlet Location Type') ales') Content gular ow Fat Outlet Establishment Cupby ('Outlet Establishment Year') ales') car.index, sales_by_year.values, ma Catablishment Year') ales') stablishment cupy, of the stablishment year' (9,5)) car.index, sales_by_year.values, ma Catablishment Year') ales') catablishment year'	ocation Type (Sales'].sum().sort_index() (rker='o',linestyle='-') (lues): (tom',fontsize=8)				
ax=grouped.plot (kind plt.xlabel ('Outlet I plt.ylabel ('Total Sa plt.legend (title='It plt.tight_layout() plt.show() 300000 - Record	Outlet Location Type') ales') Content gular ow Fat Outlet Establishment Cupby ('Outlet Establishment Year') ales') car.index, sales_by_year.values, ma Catablishment Year') ales') stablishment cupy, of the stablishment year' (9,5)) car.index, sales_by_year.values, ma Catablishment Year') ales') catablishment year'	ocation Type ('Sales'].sum().sort_index() (arker='o',linestyle='-') (alues): (attom', fontsize=8) (attom', fontsize=8) (attom', fontsize=8)	3.104 129.104 131,478			
ax=grouped.plot (kind plt.xlabel ('Outlet I plt.ylabel ('Total Saplt.legend (title='It plt.tight_layout() plt.show() 300000	Outlet Location Type') ales') item Fat Content') Outlet Tier by Content agular bw Fat Outlet Establishment Outlet Establishment Year' (9,5)) bar.index,sales_by_year.values,ma Catablishment Year') ales') stablishment Year') ales') stablishment year') ales') stablishment year') ales') content Year') ales') content Year' (9,5)) Outlet Location Type') ales') content Year' (9,5)) content Year' (9,5) c	ocation Type) ['Sales'].sum().sort_index() arker='o',linestyle='-') alues): attom',fontsize=8) ttlet Establishment	3.104 129.104 131.478			
ax=grouped.plot (kind plt.xlabel ('Outlet I plt.ylabel ('Total Sapt.legend (title='Itplt.tight_layout() plt.show() 300000 -	de'bar', figsize=(8,6), title='Outl	ocation Type ('Sales'].sum().sort_index() (arker='o',linestyle='-') (alues): (attom', fontsize=8) (attom', fontsize=8) (attom', fontsize=8)	3104 129,104 131,478			
ax=grouped.plot (kind plt.xlabel ('Outlet I plt.ylabel ('Total Saplt.legend (title='It plt.tight_layout() plt.show() Total Sales by Outle 300000 - Register of the plt.graph of the plt.graph of the plt.graph of the plt.ylabel ('Outlet Englt.ylabel ('Outlet Englt.ylabel ('Outlet Englt.tight_layout() plt.show() 200000 - Register of the plt.ylabel ('Outlet Englt.ylabel ('Outlet Englt.ylabel ('Outlet Englt.tight_layout() plt.show() 200000 - Register of the plt.ylabel ('Outlet Englt.ylabel ('Outlet Englt.ylabel ('Outlet Englt.tight_layout() plt.show() 200000 - Register of the plt.ylabel ('Outlet Englt.ylabel (Dutlet Location Type') ales') ales') cem Fat Content') Outlet Tier by Content cyular ow Fat Outlet Location Coupby ('Outlet Establishment Year' (9,5)) car.index, sales_by_year.values, ma Cstablishment Year') class') cstablishment') cs_by_year.index, sales_by_year.va [y:,.of}', ha = 'center', va='bot Outlet Location Outlet	Sales'].sum().sort_index() Sales'].sum().sum().sort_index() Sales'].sum().sum().sum().sum().sum().sum().sum().sum().sum().sum(
ax=grouped.plot (kind plt.xlabel('Outlet I plt.ylabel('Outlet I plt.ylabel('Total Saplt.legend(title='Itplt.tight_layout() plt.show() Total Sales by Outle 300000 - Re 250000 - Re 150000 - Re 1500000 - Re 150000 - Re 15	Dutlet Location Type') siles') sem Fat Content') Outlet Tier by Content egular w Fat Outlet Establishment Dupby ('Outlet Establishment Year' (9,5)) sar.index, sales_by_year.values, ma (stablishment Year') stablishment Stablishment') ses_by_year.index, sales_by_year.va (y:,.0f)', ha = 'center', va='bot Out Out Out Out Out Out Out O	Sales'].sum().sort_index() Sales'].sum().sum().sort_index() Sales'].sum().sum().sum().sum().sum().sum().sum().sum().sum().sum(
ax=grouped.plot (kind plt.xlabel('Outlet I plt.ylabel('Total Sa plt.legend(title='It plt.tight_layout() plt.show() Total Sales by Outle sales_by_year=df.gro plt.figure(figsize='plt.plot(sales_by_ye plt.xlabel('Total Sa plt.title('Outlet Es for x, y in zip (sale plt.tight_layout()) plt.show() Sales by Outlet Size for x, y in zip (sale plt.tight_layout()) plt.show() 200000 - 100000 -	Dutlet Location Type') siles') sem Fat Content') Outlet Tier by Content egular w Fat Outlet Establishment Dupby ('Outlet Establishment Year' (9,5)) sar.index, sales_by_year.values, ma (stablishment Year') stablishment Stablishment') ses_by_year.index, sales_by_year.va (y:,.0f)', ha = 'center', va='bot Out Out Out Out Out Out Out O	Sales'].sum().sort_index() Sales'].sum().sum().sort_index() Sales'].sum().sum().sum().sum().sum().sum().sum().sum().sum().sum(
ax=grouped.plot (kind plt.xlabel('Outlet I plt.ylabel('Total Sa plt.legend(title='It plt.ylabel('Total Sapt.legend(title='It plt.show()) 300000	Dutlet Location Type') iles'n fat Content') Outlet Tier by Content gular w Fat Outlet Establishment Dupby ('Outlet Establishment Year' (9,5)) Par.index, sales_by_year.values, ma Establishment') Par.index, sales_by_year.values, ma Establishment') Par.index, sales_by_year.values, ma Establishment') Par.index, sales_by_year.values, ma Coutlet Location Outlet Locat	Sales'].sum().sort_index() Sales'].sum().sum().sort_index() Sales'].sum().sum().sum().sum().sum().sum().sum().sum().sum().sum(
ax=grouped.plot (kind plt.xlabel('Outlet I plt.xlabel('Outlet I plt.ylabel('Total Saplt.legend(title='It plt.tight_layout() plt.show() Total Sales by Outle 250000	Outlet Tier by Outlet Tier by Content gular gw Fat Outlet Establishment Dupby ('Outlet Establishment Year' (9,5)) 9ar.index,sales_by_year.values,ma Costablishment Distablishment Outlet Establishment Year' (19,5) Par.index,sales_by_year.values,ma Costablishment Year') Soloyyear.index,sales_by_year.values,ma Costablishment') Outlet Establishment Year' (19,5) Outlet Establishment Year' (19,5) Outlet Establishment Year' (19,5) Soloyyear.index,sales_by_year.values,ma (19,1) Outlet Establishment Year' (19,5) Soloyyear.index,sales_by_size.index,sales_by_year.values,ma (19,5) Outlet Establishment Year' (19,5) Soloyyear.index,sales_by_size.index,sales_by_year.values,ma (19,5) Soloyyear.index,sales_by_size.	Sales'].sum().sort_index() Sales'].sum().sum().sort_index() Sales'].sum().sum().sum().sum().sum().sum().sum().sum().sum().sum(
ax=grouped.plot (kine plt.xlabel ('Outlet Signature) Sales by Outlet Size 250000 - South 250000	Outlet Tier by Content Guglar W Fat Outlet Tier by Content Guglar W Fat Outlet Establishment Supply ('Outlet Establishment Year' (9,5)) Stablishment Year') Stablishment Year' (19,5)) Stablishment Year' (19,5) Stablishment') Sa by year index, sales by year values (11,01)', ha = 'center', va='bot Outlet Size') Outlet Long Outlet Tier by Outlet Long Outlet Tier by Outlet Long Outlet Tier by Outlet Long Outlet Long Outlet Size') Outlet Long Out	Sales'].sum().sort_index() ('Sales'].sum().sort_index() ('Sales').sum().sort_index() ('Sales').sum().sum().sort_index() ('Sales').sum().sum().sort_index() ('Sales').sum().sum				
ax=grouped.plot (kine plt.xlabel ('Outlet I plt.ylabel ('Total Sales by Outlet Size I plt.tight_layout() plt.show() Total Sales by Outlet Sales_by_year=df.grouple.gr	Outlet Idea Outlet Tier by Content Egular We Fat Outlet Establishment Outlet Content') Outlet Idea Outl	Sales'].sum().sort_index() ('Sales'].sum().sort_index() ('Sales').sum().sort_index() ('Sales').sum().sum().sort_index() ('Sales').sum().sum().sort_index() ('Sales').sum().sum				
ax=grouped.plot (king plt.xlabel ('Outlet I plt.ylabel ('Total Sales by Outle Sales_by_year=df.group lt.figure (figsize=plt.plot (sales_by_year=df.group lt.xlabel ('Outlet E plt.ylabel ('Total Sales_by_year=df.group lt.xlabel ('Outlet E plt.xlabel ('Outlet E plt.ylabel ('Total Sales_by_year=df.group lt.xlabel ('Outlet E plt.xlabel ('Outlet E plt.xlabel ('Outlet E plt.xlabel ('Outlet E plt.xlabel ('Italia Sales_by_year=df.group lt.xlabel ('Outlet E plt.xlabel ('Ou	Outlet Lost t Establishment Supply ('Outlet Establishment Year') (3,5)) Sar, index, sales_by_year.values, manuple ('outlet Establishment Year') (3,5)) Sar, index, sales_by_year.values, manuple ('outlet Establishment Year') (2000	Sales'].sum().sort_index() ('Sales'].sum().sort_index() ('Sales').sum().sort_index() ('Sales').sum().sum().sort_index() ('Sales').sum().sum().sort_index() ('Sales').sum().sum				
ax=grouped.plot (kine plt.xlabel('Outlet I plt.ylabel('Total Sales by Outlet Size I plt.tight_layout() plt.show() Total Sales by Outlet Sales_by_year=df.grouple.show() Total Sales by Outlet Fight.grouple.show() Total Sales by Outlet Sales_by_year=df.grouple.show() Total Sales by Outlet Fight.grouple.show() Total Sales by Outlet Fight.grouple.show() Total Sales by Outlet Sales_by_year=df.grouple.show() Total Sales by Outlet Fight.grouple.show() Total Sales by Outlet Sales_by_year=df.grouple.show() Total Sales_by_year=df.groupl	Dutlet Tier by Outlet Tier by Content gular w Fat Outlet Establishment Dupby('Outlet Establishment Year') (9,5)) ser.index, sales_by_year.values, ma Establishment') Dupby('Outlet Size')['Sales'].sum (y;,0f)', ha = 'center', va='hot (yi,.0f)', ha = 'center', va='hot (yi,.0f	Sales'].sum().sort_index() ('Sales'].sum().sort_index() ('Sales').sum().sort_index() ('Sales').sum().sum().sort_index() ('Sales').sum().sum().sort_index() ('Sales').sum().sum				
Ax=grouped.plot (kine plt.xlabel ('Outlet I plt.ylabel ('Outlet I plt.ylabed (tittle='If plt.tight_layout()) plt.show() Total Sales by Outle 300000 -	Outlet Identify (1) (1) (1) (2) (2) (2) (2) (3) (3) (4) (4) (4) (4) (4) (4) (4) (4) (4) (4	Sales'].sum().sort_index() ('Sales'].sum().sort_index() ('Sales').sum().sort_index() ('Sales').sum().sum().sort_index() ('Sales').sum().sum().sort_index() ('Sales').sum().sum				
Axegrouped.plot (kind plt.xlabel (10tal Sales (10tal Sales plt.legend (12tal Splt.legend	Outlet Lie Stablishment Year' Sugular by Fat Content') Outlet Lie Stablishment Year' Sugular by Fat Content') Outlet Lie Stablishment Year' (9,5)) Sar_index, sales_by_year.values, mast content's content', various, mast content', various, ma	132-113 130-477 130-443 132-113 130-477 130-443 131-113 130-477 130-443 131-113 130-477 130-443 131-113 130-477 130-443 131-113 130-477 130-443 131-113 130-477 130-443 131-113 130-477 130-443 131-113 130-477 130-443 131-113 130-477 130-443 131-113 130-477 130-443 131-113 130-477 130-443 131-113 130-477 130-443 131-113 130-477 130-443 131-113 130-477 130-443 131-113 130-477 130-443 131-113 130-477 130-443 130-477 130-				

DATA ANALYSIS PYTHON PROJECT - BLINKIT ANALYSIS

IMPORT LIBRARIES