

In [1]:

```
import numpy as np  
import pandas as pd
```

In [2]:

```
data_file = pd.read_csv('Test.csv')
```

In [3]:

```
print(data_file[0:5682])
```

	Item_Identifier	Item_Weight	Item_Fat_Content	Item_Visibility	\
0	FDW58	20.750	Low Fat	0.007565	
1	FDW14	8.300	reg	0.038428	
2	NCN55	14.600	Low Fat	0.099575	
3	FDQ58	7.315	Low Fat	0.015388	
4	FDY38	NaN	Regular	0.118599	
...	...	...	...	...	
5676	FDB58	10.500	Regular	0.013496	
5677	FDD47	7.600	Regular	0.142991	
5678	NC017	10.000	Low Fat	0.073529	
5679	FDJ26	15.300	Regular	0.000000	
5680	FDU37	9.500	Regular	0.104720	

	Item_Type	Item_MRP	Outlet_Identifier	\
0	Snack Foods	107.8622	OUT049	
1	Dairy	87.3198	OUT017	
2	Others	241.7538	OUT010	
3	Snack Foods	155.0340	OUT017	
4	Dairy	234.2300	OUT027	
...	...	...	...	
5676	Snack Foods	141.3154	OUT046	
5677	Starchy Foods	169.1448	OUT018	
5678	Health and Hygiene	118.7440	OUT045	
5679	Canned	214.6218	OUT017	
5680	Canned	79.7960	OUT045	

	Outlet_Establishment_Year	Outlet_Size	Outlet_Location_Type	\
0	1999	Medium	Tier 1	
1	2007	NaN	Tier 2	
2	1998	NaN	Tier 3	
3	2007	NaN	Tier 2	
4	1985	Medium	Tier 3	
...	...	...	...	
5676	1997	Small	Tier 1	
5677	2009	Medium	Tier 3	
5678	2002	NaN	Tier 2	
5679	2007	NaN	Tier 2	
5680	2002	NaN	Tier 2	

	Outlet_Type
0	Supermarket Type1
1	Supermarket Type1
2	Grocery Store
3	Supermarket Type1
4	Supermarket Type3
...	...
5676	Supermarket Type1
5677	Supermarket Type2
5678	Supermarket Type1
5679	Supermarket Type1
5680	Supermarket Type1

[5681 rows x 11 columns]

In [6]:

```
MRP = data_file['Item_MRP']
name = data_file['Item_Identifier']
weight = data_file['Item_Weight']
fat = data_file['Item_Fat_Content']
visibility = data_file['Item_Visibility']
Type = data_file['Item_Type']
identifier = data_file['Outlet_Identifier']
year = data_file['Outlet_Establishment_Year']
size = data_file['Outlet_Size']
L_Type = data_file['Outlet_Location_Type']
Outlet_Type = data_file['Outlet_Type']
```

In [5]:

```
avg_col = np.mean(MRP,0)
print(avg_col)
```

141.02327340256994

In [14]:

```
my_data = [name, Type, MRP, print('average of the MRP =',avg_col)]
result = pd.concat(my_data,axis=1)
print(result[0:5682])
```

average of the MRP = 141.02327340256994

	Item_Identifier	Item_Type	Item_MRP
0	FDW58	Snack Foods	107.8622
1	FDW14	Dairy	87.3198
2	NCN55	Others	241.7538
3	FDQ58	Snack Foods	155.0340
4	FDY38	Dairy	234.2300
...	...	...	...
5676	FDB58	Snack Foods	141.3154
5677	FDD47	Starchy Foods	169.1448
5678	NC017	Health and Hygiene	118.7440
5679	FDJ26	Canned	214.6218
5680	FDU37	Canned	79.7960

[5681 rows x 3 columns]

In [16]:

```
result.to_csv('result.csv')
```

In []:

In [1]:

```
import pandas as pd
```

In [2]:

```
df_train = pd.read_csv('Train.csv')
```

In [3]:

```
df_train
```

Out[3]:

id	Item_Fat_Content	Item_Visibility	Item_Type	Item_MRP	Outlet_Identifier	Outlet_Establishment_Year
100	Low Fat	0.016047	Dairy	249.8092	OUT049	2009
120	Regular	0.019278	Soft Drinks	48.2692	OUT018	2009
100	Low Fat	0.016760	Meat	141.6180	OUT049	2009
100	Regular	0.000000	Fruits and Vegetables	182.0950	OUT010	2009
130	Low Fat	0.000000	Household	53.8614	OUT013	2009
...	...	...	...	...	...	...
165	Low Fat	0.056783	Snack Foods	214.5218	OUT013	2009
180	Regular	0.046982	Baking Goods	108.1570	OUT045	2009
100	Low Fat	0.035186	Health and Hygiene	85.1224	OUT035	2009
110	Regular	0.145221	Snack Foods	103.1332	OUT018	2009
100	Low Fat	0.044878	Soft Drinks	75.4670	OUT046	2009



In [4]:

```
len(df_train['Item_Type'])
df_train['Item_Type'].count()
```

Out[4]:

8523

In [5]:



```
#Value count of Item_Type  
df_train['Item_Type'].value_counts()
```

Out[5]:

Fruits and Vegetables	1232
Snack Foods	1200
Household	910
Frozen Foods	856
Dairy	682
Canned	649
Baking Goods	648
Health and Hygiene	520
Soft Drinks	445
Meat	425
Breads	251
Hard Drinks	214
Others	169
Starchy Foods	148
Breakfast	110
Seafood	64

Name: Item_Type, dtype: int64

In [7]:



```
#percentage of Item_Type  
df_train['Item_Type'].value_counts(normalize=True)
```

Out[7]:

Fruits and Vegetables	0.144550
Snack Foods	0.140795
Household	0.106770
Frozen Foods	0.100434
Dairy	0.080019
Canned	0.076147
Baking Goods	0.076030
Health and Hygiene	0.061011
Soft Drinks	0.052212
Meat	0.049865
Breads	0.029450
Hard Drinks	0.025109
Others	0.019829
Starchy Foods	0.017365
Breakfast	0.012906
Seafood	0.007509

Name: Item_Type, dtype: float64

In [16]:

```
df_train[df_train.index.isin([723])]
```

Out[16]:

	Item_Identifier	Item_Weight	Item_Fat_Content	Item_Visibility	Item_Type	Item_MRP	Outle
723	FDO52	NaN	Regular	0.076791	Frozen Foods	172.1106	

In [11]:

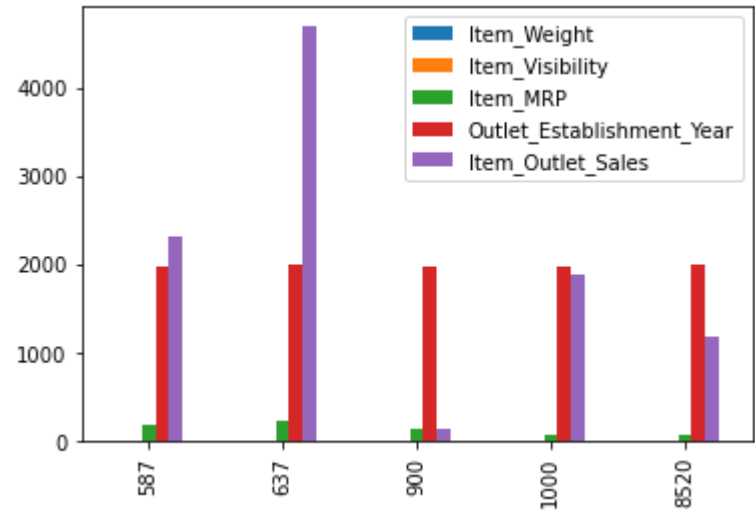
```
df_train_sample = df_train[df_train.index.isin([637, 587, 8520, 1000, 900])]
```

In [12]:

```
df_train_sample.plot(kind='bar')
```

Out[12]:

<AxesSubplot:>



In [1]:

```
# Python program to find the average of odd numbers in the given list
data=[68,71,21,68,32,80,64,52,92,97,93,37,61,62,23,81,67,84,33,41,92,51,71,52,92,43,80,74,6
# inp=input().split(' ')
# data=[int(x) for x in inp]
sm,cnt=0,0

for x in data:
    if x%2!=0:
        sm+=x
        cnt+=1

print(" avg of all odd numbers is ",str(sm/cnt))
```

avg of all odd numbers is 61.97391304347826

In [16]:

```

# python program to find the third largest number in the given list
import sys
def thirdLargest(arr, arr_size):

    # There should be
    # atleast three elements
    if (arr_size < 3):

        print(" Invalid Input ")
        return

    # Find first
    # largest element
    first = arr[0]
    for i in range(1, arr_size):
        if (arr[i] > first):
            first = arr[i]

    # Find second
    # largest element
    second = -sys.maxsize
    for i in range(0, arr_size):
        if (arr[i] > second and
            arr[i] < first):
            second = arr[i]

    # Find third
    # largest element
    third = -sys.maxsize
    for i in range(0, arr_size):
        if (arr[i] > third and
            arr[i] < second):
            third = arr[i]

    print("The Third Largest",
          "element is", third)

# Driver Code
arr = [68, 71, 21, 68, 32, 80, 64, 52, 92, 97, 93, 37, 61, 62, 23, 81, 67, 84, 33, 41, 92,
43, 80, 74, 69, 92, 79, 78, 70, 31, 34, 79, 73, 77, 96, 87, 53, 47, 85, 58, 58, 87, 44, 29,
71, 57, 96, 74, 30, 65, 62, 85, 34, 21, 81, 79, 84, 87, 22, 45, 59, 91, 26, 22, 46, 71, 60,
38, 82, 95, 77, 55, 37, 61, 59, 89, 61, 75, 97, 72, 58, 23, 78, 20, 95, 67, 59, 56, 93, 82, 3
95, 90, 80, 78, 73, 91, 52, 59, 96, 46, 94, 99, 78, 53, 40, 67, 76, 73, 46, 77, 40, 34, 38,
61, 76, 21, 90, 52, 65, 36, 93, 55, 86, 95, 65, 30, 89, 86, 20, 92, 46, 58, 75, 43, 57, 53, 4
58, 31, 73, 71, 63, 70, 72, 78, 26, 39, 35, 58, 84, 56, 24, 73, 36, 75, 32, 35, 56, 87, 87,
38, 94, 36, 90, 70, 65, 34, 61, 44, 85, 41, 50, 23, 34, 20, 87, 84, 30, 21, 84, 65, 85, 88,
41, 35, 71]
n = len(arr)
thirdLargest(arr, n)

```

The Third Largest element is 96

In [17]:

```
# python program to find the count of even and odd numbers in the given list
list1 = [68,71,21,68,32,80,64,52,92,97,93,37,61,62,23,81,67,84,33,41,92,51,71,52,92,43,80,74]

even_count, odd_count = 0, 0

for num in list1:

    if num % 2==0:
        even_count +=1

    else:
        odd_count +=1

print("Even numbers in the list: ", even_count)
print("Odd numbers in the list: ",odd_count)
```

Even numbers in the list: 105
Odd numbers in the list: 115

In []: