Taco Sales Analysis

In [1]:
```python
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
import seaborn as sns
```

In [2]:
```python
Rdata=pd.read_csv("taco_sales_(2024-2025).csv")
```

In [3]:
```python
Rdata.isnull().sum()
```

Out[3]:
```
Order ID                  0
Restaurant Name           0
Location                  0
Order Time                0
Delivery Time             0
Delivery Duration (min)   0
Taco Size                 0
Taco Type                 0
Toppings Count            0
Distance (km)             0
Price ($)                 0
Tip ($)                   0
Weekend Order             0
dtype: int64
```

In [4]:
```python
data=Rdata.dropna
```

In [5]:
```python
print(data)
```

```
<bound method DataFrame.dropna of      Order ID   Restaurant Name     Location      Order Time  \
0      770487      El Taco Loco      New York  01-08-2024 14:55
1      671858      El Taco Loco   San Antonio  23-11-2024 17:11
2      688508        Taco Haven        Austin  21-11-2024 20:24
3      944962   Spicy Taco House        Dallas  21-09-2024 06:43
4      476417     Casa del Taco   San Antonio  24-07-2024 11:01
..        ...               ...          ...               ...
995    164891      La Vida Taco        Austin  27-05-2024 11:12
996    232442    The Taco Stand   San Antonio  03-04-2025 20:51
997    251729       Urban Tacos       Houston  23-01-2025 05:33
998    940215       Taco Fiesta   San Antonio  05-09-2024 15:49
999    326183      Grande Tacos       Phoenix  19-10-2024 14:21

         Delivery Time  Delivery Duration (min) Taco Size    Taco Type  \
0     01-08-2024 15:36                       41   Regular  Chicken Taco
1     23-11-2024 17:25                       14   Regular     Beef Taco
2     21-11-2024 21:02                       38     Large     Pork Taco
3     21-09-2024 07:28                       45   Regular  Chicken Taco
4     24-07-2024 11:16                       15     Large     Pork Taco
..                 ...                      ...       ...           ...
995   27-05-2024 12:11                       59     Large   Veggie Taco
996   03-04-2025 21:29                       38   Regular     Fish Taco
997   23-01-2025 06:41                       68   Regular     Pork Taco
998   05-09-2024 16:38                       49   Regular   Veggie Taco
999   19-10-2024 14:38                       17     Large  Chicken Taco

     Toppings Count  Distance (km)  Price ($)  Tip ($)  Weekend Order
0                 5           3.01       9.25     2.22          False
1                 1           6.20       4.25     3.01           True
2                 2          20.33       7.00     0.02          False
3                 2           3.00       5.50     1.90           True
4                 0          24.34       4.50     1.14          False
..              ...            ...        ...      ...            ...
995               2          20.04       7.00     2.80          False
996               0          13.69       3.00     1.38          False
997               2           4.07       5.50     2.00          False
998               5          14.56       9.25     1.89          False
999               5           5.66      10.75     3.66           True

[1000 rows x 13 columns]>
```

In [6]:
```python
numerical_data=Rdata.select_dtypes(include=['int64','float64'])
print(numerical_data)
```

```
       Order ID  Delivery Duration (min)  Toppings Count  Distance (km)  \
0        770487                       41               5           3.01
1        671858                       14               1           6.20
2        688508                       38               2          20.33
3        944962                       45               2           3.00
4        476417                       15               0          24.34
..          ...                      ...             ...            ...
995      164891                       59               2          20.04
996      232442                       38               0          13.69
997      251729                       68               2           4.07
998      940215                       49               5          14.56
999      326183                       17               5           5.66

     Price ($)  Tip ($)
0         9.25     2.22
1         4.25     3.01
2         7.00     0.02
3         5.50     1.90
4         4.50     1.14
..         ...      ...
995       7.00     2.80
996       3.00     1.38
997       5.50     2.00
998       9.25     1.89
999      10.75     3.66

[1000 rows x 6 columns]
```

In [7]:
```python
object_data=Rdata.select_dtypes(include=['object'])
print(object_data)
```

```
       Restaurant Name     Location        Order Time      Delivery Time  \
0         El Taco Loco     New York  01-08-2024 14:55  01-08-2024 15:36
1         El Taco Loco  San Antonio  23-11-2024 17:11  23-11-2024 17:25
2          Taco Haven       Austin  21-11-2024 20:24  21-11-2024 21:02
3     Spicy Taco House       Dallas  21-09-2024 06:43  21-09-2024 07:28
4        Casa del Taco  San Antonio  24-07-2024 11:01  24-07-2024 11:16
..                 ...          ...               ...               ...
995        La Vida Taco       Austin  27-05-2024 11:12  27-05-2024 12:11
996     The Taco Stand  San Antonio  03-04-2025 20:51  03-04-2025 21:29
997         Urban Tacos      Houston  23-01-2025 05:33  23-01-2025 06:41
998         Taco Fiesta  San Antonio  05-09-2024 15:49  05-09-2024 16:38
999        Grande Tacos      Phoenix  19-10-2024 14:21  19-10-2024 14:38

     Taco Size     Taco Type
0      Regular  Chicken Taco
1      Regular     Beef Taco
2        Large     Pork Taco
3      Regular  Chicken Taco
4        Large     Pork Taco
..         ...          ...
995      Large   Veggie Taco
996    Regular     Fish Taco
997    Regular     Pork Taco
998    Regular   Veggie Taco
999      Large  Chicken Taco

[1000 rows x 6 columns]
```

In [8]:
```python
Rdata.describe().T
```

Out[8]:

	count	mean	std	min	25%	50%	75%	max
Order ID	1000.0	552504.86500	255948.779709	101139.00	331796.7500	559740.00	771781.7500	999138.00
Delivery Duration (min)	1000.0	50.93000	23.227540	10.00	30.0000	53.00	71.0000	90.00
Toppings Count	1000.0	2.52900	1.717005	0.00	1.0000	3.00	4.0000	5.00
Distance (km)	1000.0	13.07342	7.142268	0.51	6.9725	13.20	19.2425	24.98
Price ($)	1000.0	6.90825	2.310137	3.00	4.5000	6.75	9.2500	10.75
Tip ($)	1000.0	1.80611	1.132035	0.01	0.9075	1.76	2.5200	4.98

In [9]:
```python
#Total Sales By Restuarant
sales_restaurant=Rdata.groupby('Restaurant Name')['Price ($)'].sum()
```

In [10]:
```python
sales_restaurant
```

```
Out[10]:  Restaurant Name
          Casa del Taco          708.75
          El Taco Loco           574.00
          Grande Tacos           727.00
          La Vida Taco           666.25
          Spicy Taco House       732.75
          Taco Fiesta            712.00
          Taco Haven             654.00
          Taco Time Express      616.75
          The Taco Stand         722.25
          Urban Tacos            794.50
          Name: Price ($), dtype: float64
```

```
In [11]:  plt.figure(figsize=(5,3))
          sales_restaurant.plot(kind='barh',color='Red',edgecolor='blue')
          plt.title("Total Sales by Restaurant")
          plt.xlabel("Sales")
          plt.ylabel("Restaurant")
          plt.show()
```


Total Sales by Restaurant
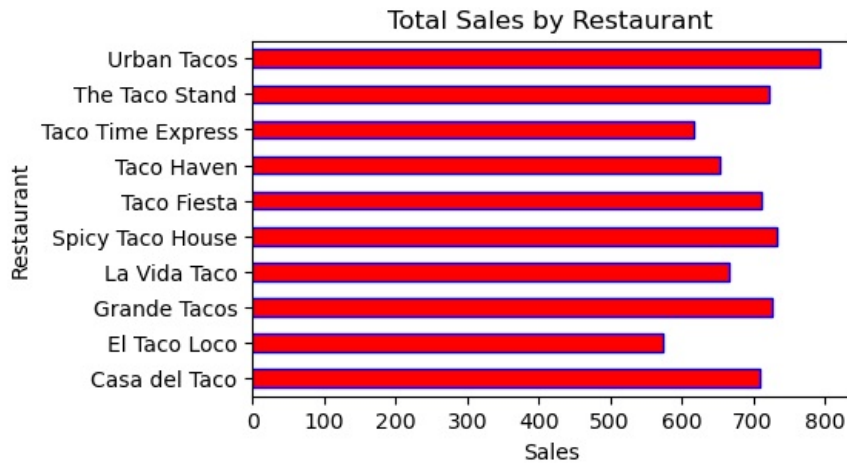
INSIGHTS: Urban tacos restaurant is the top selling restaurant and el taco loco is the lowest selling restaurant.

```
In [12]:  #Top selling taco type
          top_tacotype=Rdata.groupby('Taco Type')['Order ID'].value_counts().reset_index()
          top_tacotypes=top_tacotype.groupby('Taco Type')['count'].sum()
```

```
In [13]:  top_tacotypes
```

```
Out[13]:  Taco Type
          Beef Taco        182
          Chicken Taco     218
          Fish Taco        211
          Pork Taco        192
          Veggie Taco      197
          Name: count, dtype: int64
```

```
In [14]:  plt.figure(figsize=(5,3))
          top_tacotypes.plot(kind='bar', color='skyblue',edgecolor='blue')
          plt.title("Sales by Taco Type")
          plt.xlabel("Taco Type")
          plt.ylabel("Total Sales")
          plt.show()
```
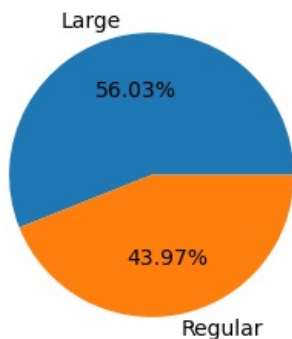
Sales by Taco Type
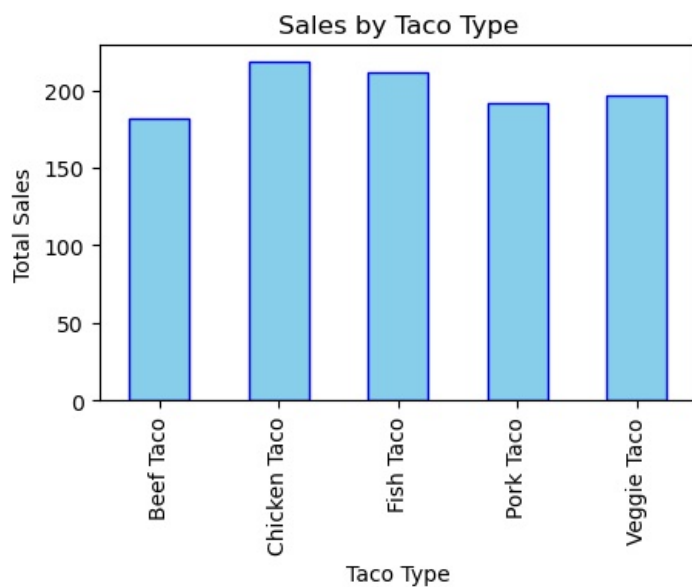
INSIGHTS: Chicken taco is highest selling taco type followed by fish taco, veggie taco and pork taco. Also, beef taco is the lowest selling taco type.

```
In [15]: #revenue by taco size
         revenue_tacosize=Rdata.groupby('Taco Size')['Price ($)'].sum()
         revenue_tacosize
```

```
Out[15]: Taco Size
         Large      3871.00
         Regular    3037.25
         Name: Price ($), dtype: float64
```

```
In [16]: plt.figure(figsize=(5,3))
         plt.pie(revenue_tacosize,labels=revenue_tacosize.index,autopct='%1.2f%%')
         plt.title("Proportion of Sales by Taco Size")
         plt.show()
```

Proportion of Sales by Taco Size

INSIGHTS:When analysing the highest sales by taco size, large size tacos was sold higher than the regular size tacos.

```
In [17]: #average delivery time by location
         avg_deliverytime=round(Rdata.groupby('Location')['Delivery Duration (min)'].mean(),2)
         avg_deliverytime
```

```
Out[17]: Location
         Austin         48.96
         Chicago        50.78
         Dallas         55.04
         Houston        50.92
         Los Angeles    49.93
         New York       49.79
         Phoenix        51.11
         San Antonio    49.79
         San Diego      52.67
         San Jose       51.15
         Name: Delivery Duration (min), dtype: float64
```

```
In [18]: plt.figure(figsize=(5,5))
         avg_deliverytime.plot(kind='barh',color='yellow',edgecolor='black')
         plt.title("Average Delivery Time Taken by Location")
         plt.xlabel("Average Delivery Time (min)")
         plt.ylabel("Location")
         plt.show()
```

Average Delivery Time Taken by Location
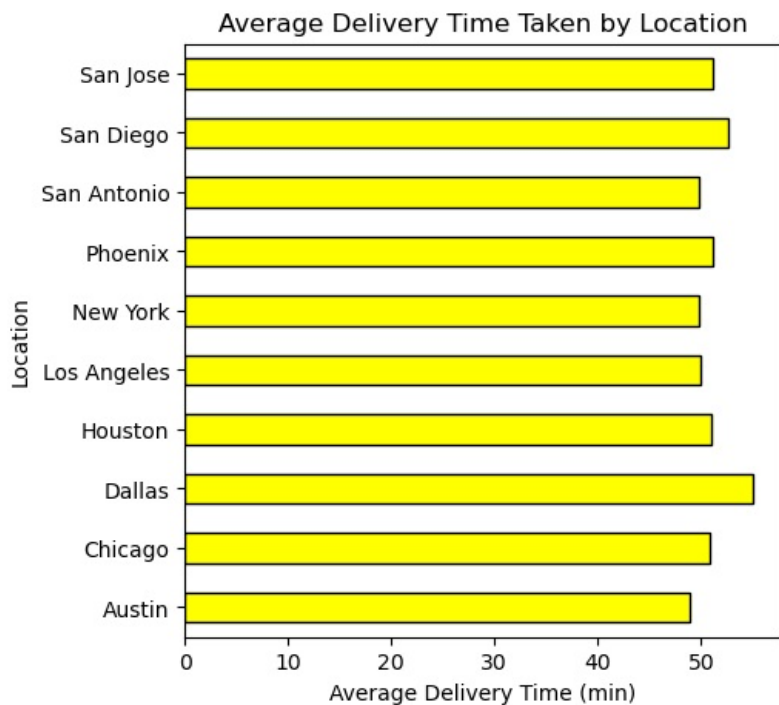
INSIGHTS: When analysing the average time taken to deliver the order in each location, shows Dallas location takes more time to delivery when compared to other locations.

```
In [19]: Rdata['tip_percent']=(Rdata['Tip ($)']/Rdata['Price ($)'])*100
         Rdata['tip_percent']
```

```
Out[19]: 0        24.000000
         1        70.823529
         2         0.285714
         3        34.545455
         4        25.333333
                    ...
         995      40.000000
         996      46.000000
         997      36.363636
         998      20.432432
         999      34.046512
         Name: tip_percent, Length: 1000, dtype: float64
```

```
In [20]: location_tip=round(Rdata.groupby('Location')['tip_percent'].mean(),2)
         location_tip
```

```
Out[20]: Location
         Austin         27.86
         Chicago        30.88
         Dallas         28.91
         Houston        31.00
         Los Angeles    25.84
         New York       27.27
         Phoenix        29.93
         San Antonio    29.98
         San Diego      32.20
         San Jose       29.94
         Name: tip_percent, dtype: float64
```

```
In [21]: #Average tipping percent based on location
         plt.figure(figsize=(5,3))
         location_tip.plot(kind='bar',color='orange',edgecolor='red')
         plt.title("Average Tipping Percent Based on Location")
         plt.xlabel("Location")
         plt.ylabel("Avg Tip Percentage")
         plt.show()
```

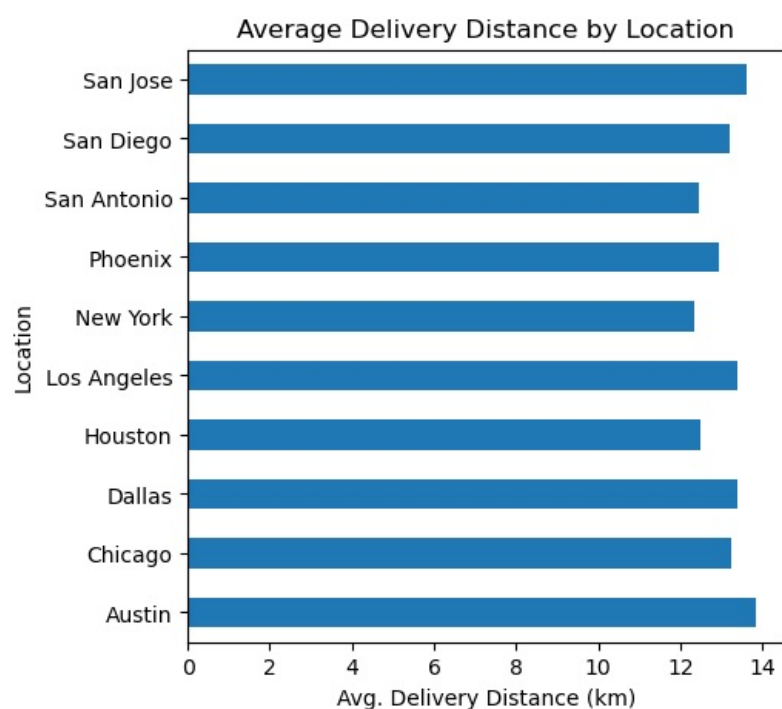## Average Tipping Percent Based on Location
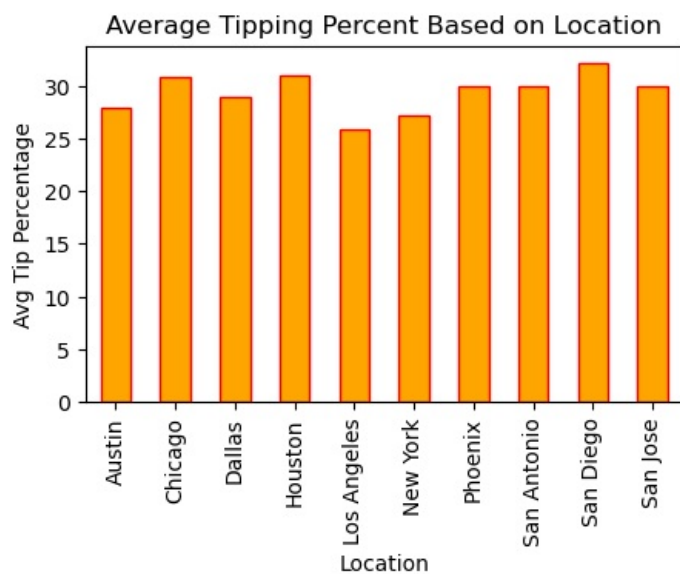
INSIGHTS: Analysing the average tipping by location shows, San Diego, Houston and Chicago are the top three location providing more tip to delivery person.

```
In [22]: avgdist_location=round(Rdata.groupby('Location')['Distance (km)'].mean(),2)
         avgdist_location
```

```
Out[22]: Location
         Austin          13.84
         Chicago         13.22
         Dallas          13.40
         Houston         12.50
         Los Angeles     13.38
         New York        12.34
         Phoenix         12.95
         San Antonio     12.45
         San Diego       13.19
         San Jose        13.61
         Name: Distance (km), dtype: float64
```

```
In [23]: #Average Delivery Distance by Location
         plt.figure(figsize=(5,5))
         avgdist_location.plot(kind='barh')
         plt.title("Average Delivery Distance by Location")
         plt.xlabel("Avg. Delivery Distance (km)")
         plt.ylabel("Location")
         plt.show()
```

Average Delivery Distance by Location

INSIGHTS:Anlaysing the average delivery distance by location shows, restaurants located in Austin delivery to longest distance when compared to other locations.

```
In [24]:  Rdata['Order Time'] = pd.to_datetime(Rdata['Order Time'], format='%d-%m-%Y %H:%M')
          Rdata['year'] = Rdata['Order Time'].dt.year
          Rdata['year']

Out[24]:  0        2024
          1        2024
          2        2024
          3        2024
          4        2024
                   ...
          995      2024
          996      2025
          997      2025
          998      2024
          999      2024
          Name: year, Length: 1000, dtype: int32

In [25]:  Rdata['month'] = Rdata['Order Time'].dt.month
          Rdata['month']

Out[25]:  0         8
          1        11
          2        11
          3         9
          4         7
                   ..
          995       5
          996       4
          997       1
          998       9
          999      10
          Name: month, Length: 1000, dtype: int32

In [26]:  sales_month=round(Rdata.groupby('month')['Price ($)'].sum(),2)
          sales_month

Out[26]:  month
          1      961.50
          2      700.50
          3      876.50
          4      731.75
          5      810.25
          6      429.25
          7      429.00
          8      404.25
          9      352.25
          10     420.75
          11     431.00
          12     361.25
          Name: Price ($), dtype: float64

In [27]:  #Average Sale by Month
          plt.figure(figsize=(5,3))
          sales_month.plot(kind='line',x='month',y='Price ($)',marker='o')
          plt.title("Total Sales by Month")
          plt.xlabel("Month")
          plt.ylabel("Total Sales")
          plt.show()
```
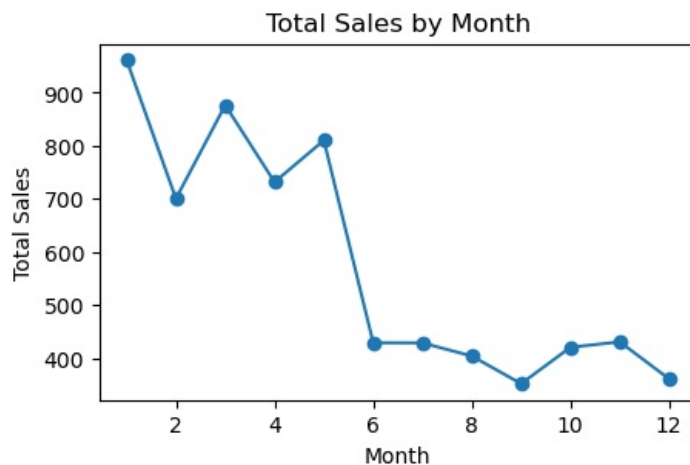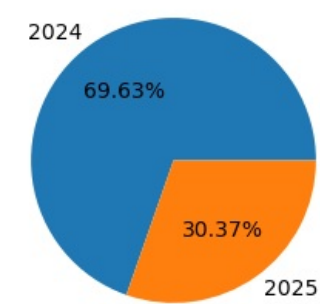

Total Sales by Month

```
In [28]:  sales_year=round(Rdata.groupby('year')['Price ($)'].sum(),2)
          sales_year
```

year
        2024   4810.50
        2025   2097.75
        Name: Price ($), dtype: float64

In [29]:
```python
plt.figure(figsize=(5,3))
plt.pie(sales_year,labels=sales_year.index,autopct='%1.2f%%')
plt.title("Proportion of Sales by Year")
plt.show()
```
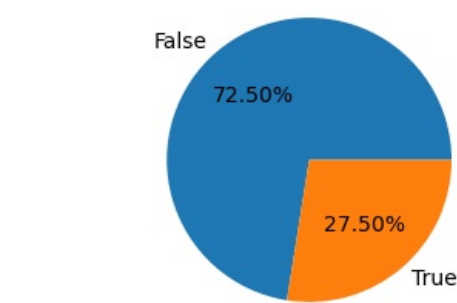
Proportion of Sales by Year

In [31]:
```python
#Weekend Order Analysis:
weekend_order=Rdata['Weekend Order'].value_counts()
weekend_order
```

Out[31]:   Weekend Order
        False   725
        True   275
        Name: count, dtype: int64

In [33]:
```python
plt.figure(figsize=(5,3))
plt.pie(weekend_order,labels=weekend_order.index, autopct='%1.2f%%')
plt.title("Pie Chart Showing Proportion of Weekend Orders")
plt.show()
```

Pie Chart Showing Proportion of Weekend Orders

INSIGHTS: From the data, most of the Orders are weekday orders.

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js