Task: Price Range vs. Online Delivery and Table Booking

3.3.1 Analyze if there is a relationship between the price range and the availability of online delivery and table booking.

In [1]:
```
import pandas as pd

data = pd.read_csv(r"C:\Users\HP\OneDrive\Documents\Cognifyz Internship Program\Dataset.

data
```

Out[1]:

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longi
0	6317637	Le Petit Souffle	162	Makati City	Third Floor, Century City Mall, Kalayaan Avenu...	Century City Mall, Poblacion, Makati City	Century City Mall, Poblacion, Makati City, Mak...	121.02
1	6304287	Izakaya Kikufuji	162	Makati City	Little Tokyo, 2277 Chino Roces Avenue, Legaspi...	Little Tokyo, Legaspi Village, Makati City	Little Tokyo, Legaspi Village, Makati City, Ma...	121.01
2	6300002	Heat - Edsa Shangri-La	162	Mandaluyong City	Edsa Shangri-La, 1 Garden Way, Ortigas, Mandal...	Edsa Shangri-La, Ortigas, Mandaluyong City	Edsa Shangri-La, Ortigas, Mandaluyong City, Ma...	121.05
3	6318506	Ooma	162	Mandaluyong City	Third Floor, Mega Fashion Hall, SM Megamall, O...	SM Megamall, Ortigas, Mandaluyong City	SM Megamall, Ortigas, Mandaluyong City, Mandal...	121.05
4	6314302	Sambo Kojin	162	Mandaluyong City	Third Floor, Mega Atrium, SM Megamall, Ortigas...	SM Megamall, Ortigas, Mandaluyong City	SM Megamall, Ortigas, Mandaluyong City, Mandal...	121.05
...	...	...	...	...	...	...	...	
9546	5915730	Naml\ Gurme	208	��stanbul	Kemanke�� Karamustafa Pa��a Mahallesi, R\ht\m ...	Karak�_y	Karak�_y, ��stanbul	28.97
9547	5908749	Ceviz A��ac\	208	��stanbul	Ko��uyolu Mahallesi, Muhittin ��st�_nda�� Cadd...	Ko��uyolu	Ko��uyolu, ��stanbul	29.04
9548	5915807	Huqqa	208	��stanbul	Kuru�_e��me Mahallesi, Muallim Naci Caddesi, N...	Kuru�_e��me	Kuru�_e��me, ��stanbul	29.03
9549	5916112	A���k Kahve	208	��stanbul	Kuru�_e��me Mahallesi, Muallim Naci Caddesi, N...	Kuru�_e��me	Kuru�_e��me, ��stanbul	29.03
9550	5927402	Walter's Coffee Roastery	208	��stanbul	Cafea��a Mahallesi, Bademalt\ Sokak, No 21/B, ...	Moda	Moda, ��stanbul	29.02

9551 rows × 21 columns

In [18]:
```
# Create columns for price range, delivery availability, and booking availability
```

In [2]:
```
data['Price range'] = pd.to_numeric(data['Price range'], errors='coerce')
```

In [3]:
```
data['Price range']
```

Out[3]:
```
0       3
1       3
2       4
3       4
4       4
       ..
9546    3
9547    3
9548    4
9549    4
9550    2
Name: Price range, Length: 9551, dtype: int64
```

In [9]:
```
data['PriceRange'] = pd.cut(data['Price range'], bins=[0, 1, 2, 3, 4, 5], labels=['$','$
```

In [10]:
```
data['PriceRange']
```

Out[10]:
```
0        $$$
1        $$$
2       $$$$
3       $$$$
4       $$$$
        ...
9546     $$$
9547     $$$
9548    $$$$
9549    $$$$
9550      $$
Name: PriceRange, Length: 9551, dtype: category
Categories (5, object): ['$' < '$$' < '$$$' < '$$$$' < '$$$$$']
```

In [11]:
```
data['Delivery'] = data['Has Online delivery'].map({True:1, False:0})
```

In [12]:
```
data['Delivery']
```

Out[12]:
```
0      NaN
1      NaN
2      NaN
3      NaN
4      NaN
        ..
9546   NaN
9547   NaN
9548   NaN
9549   NaN
9550   NaN
Name: Delivery, Length: 9551, dtype: float64
```

In [13]:
```
data['Booking'] = data['Has Table booking'].map({True:1, False:0})
```

In [14]:
```
data['Booking']
```

Out[14]:
```
0      NaN
1      NaN
2      NaN
3      NaN
4      NaN
        ..
9546   NaN
9547   NaN
9548   NaN
9549   NaN
9550   NaN
Name: Booking, Length: 9551, dtype: float64
```

In [ ]:
```
# Analyze correlations
```

In [15]:
```
print(data[['PriceRange', 'Delivery', 'Booking']].corr())
```

```
           Delivery  Booking
Delivery        NaN      NaN
Booking         NaN      NaN
```

In [ ]:
```
# ANOVA test for difference in means of delivery availability by price range and booking
```

In [16]:
```
import scipy.stats as stats
stats.f_oneway(data[data['PriceRange']=='$']['Delivery'],
               data[data['PriceRange']=='$$']['Delivery'],
               data[data['PriceRange']=='$$$']['Delivery'],
               data[data['PriceRange']=='$$$$']['Delivery'])
```

Out[16]:
```
F_onewayResult(statistic=nan, pvalue=nan)
```

In [17]:
```
import scipy.stats as stats
stats.f_oneway(data[data['PriceRange']=='$']['Booking'],
               data[data['PriceRange']=='$$']['Booking'],
               data[data['PriceRange']=='$$$']['Booking'],
               data[data['PriceRange']=='$$$$']['Booking'])
```

Out[17]:
```
F_onewayResult(statistic=nan, pvalue=nan)
```

3.3.2 Determine if higher-priced restaurants are more likely to offer these services.

In [21]:
```
# Create price buckets
price_buckets = [0,20,40,60,80,100]
data['PriceRange'] = pd.cut(data['Price range'], bins=price_buckets)
```

In [22]:
```
data['PriceRange']
```

Out[22]:
```
0       (0, 20]
1       (0, 20]
2       (0, 20]
3       (0, 20]
4       (0, 20]
         ...
9546    (0, 20]
9547    (0, 20]
9548    (0, 20]
9549    (0, 20]
9550    (0, 20]
Name: PriceRange, Length: 9551, dtype: category
Categories (5, interval[int64, right]): [(0, 20] < (20, 40] < (40, 60] < (60, 80] < (80,
100]]
```

In [23]:
```
# Create columns for delivery and booking
data['Delivery'] = data['Has Online delivery'].map({True:1, False:0})
data['Booking'] = data['Has Table booking'].map({True:1, False:0})
```

In [24]:
```
data['Delivery']
```

Out[24]:
```
0      NaN
1      NaN
2      NaN
3      NaN
4      NaN
       ..
9546   NaN
9547   NaN
9548   NaN
9549   NaN
9550   NaN
Name: Delivery, Length: 9551, dtype: float64
```

In [25]:
```
data['Booking']
```

Out[25]:
```
0      NaN
1      NaN
2      NaN
3      NaN
4      NaN
       ..
9546   NaN
9547   NaN
9548   NaN
9549   NaN
9550   NaN
Name: Booking, Length: 9551, dtype: float64
```

In [26]:
```
# Groupby price range and get mean of delivery and booking
means = data.groupby('PriceRange')[['Delivery','Booking']].mean()
```

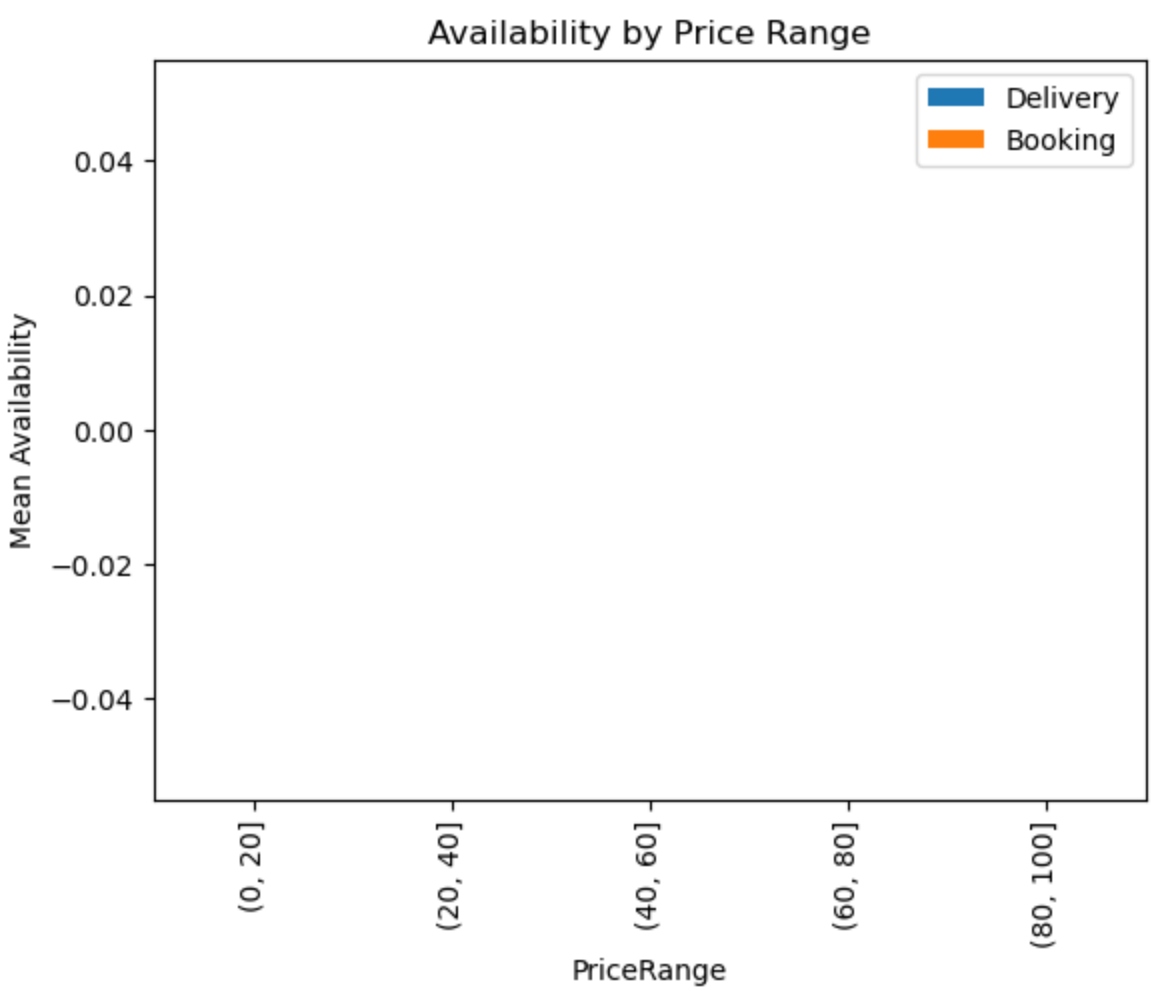
In [27]:
```
means
```

Out[27]:

PriceRange	Delivery	Booking
(0, 20]	NaN	NaN
(20, 40]	NaN	NaN
(40, 60]	NaN	NaN
(60, 80]	NaN	NaN
(80, 100]	NaN	NaN

In [28]:
```
# Plot the means
import matplotlib.pyplot as plt
means.plot(kind='bar')
plt.title('Availability by Price Range')
plt.ylabel('Mean Availability')
```

Out[28]:
```
Text(0, 0.5, 'Mean Availability')
```



In [29]:
```
from scipy.stats import f_oneway
```

In [31]:
```
f_oneway(data[data['PriceRange']==(0, 20)]['Delivery'],
         data[data['PriceRange']==(20, 40)]['Delivery'],
         data[data['PriceRange']==(40, 60)]['Delivery'])
```

```
C:\Users\HP\anaconda3\lib\site-packages\scipy\stats\_stats_py.py:3869: DegenerateDataWar
ning: at least one input has length 0
  warnings.warn(stats.DegenerateDataWarning('at least one input '
```

Out[31]:
```
F_onewayResult(statistic=nan, pvalue=nan)
```

In [32]:
```
f_oneway(data[data['PriceRange']==(0, 20)]['Booking'],
         data[data['PriceRange']==(20, 40)]['Booking'],
         data[data['PriceRange']==(40, 60)]['Booking'])
```

Out[32]:
```
F_onewayResult(statistic=nan, pvalue=nan)
```

In [ ]: