Question 1: Exploratory Data Analysis

In [2]:
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from collections import Counter

#Load the dataset
data = pd.read_csv("Restaurant_Menu.csv")
```

In [7]:
```
#Print the first 5 rows of data

print ("Head of the dataset:")
print (data.head())
```

```
Head of the dataset:
  RestaurantID MenuCategory               MenuItem  \
0         R003    Beverages                   Soda
1         R001   Appetizers  Spinach Artichoke Dip
2         R003     Desserts    New York Cheesecake
3         R003  Main Course        Chicken Alfredo
4         R002  Main Course          Grilled Steak

                                        Ingredients  Price Profitability
0                                  ['confidential']   2.55           Low
1       ['Tomatoes', 'Basil', 'Garlic', 'Olive Oil']  11.12        Medium
2           ['Chocolate', 'Butter', 'Sugar', 'Eggs']  18.66          High
3  ['Chicken', 'Fettuccine', 'Alfredo Sauce', 'Pa...  29.55          High
4  ['Chicken', 'Fettuccine', 'Alfredo Sauce', 'Pa...  17.73        Medium
```

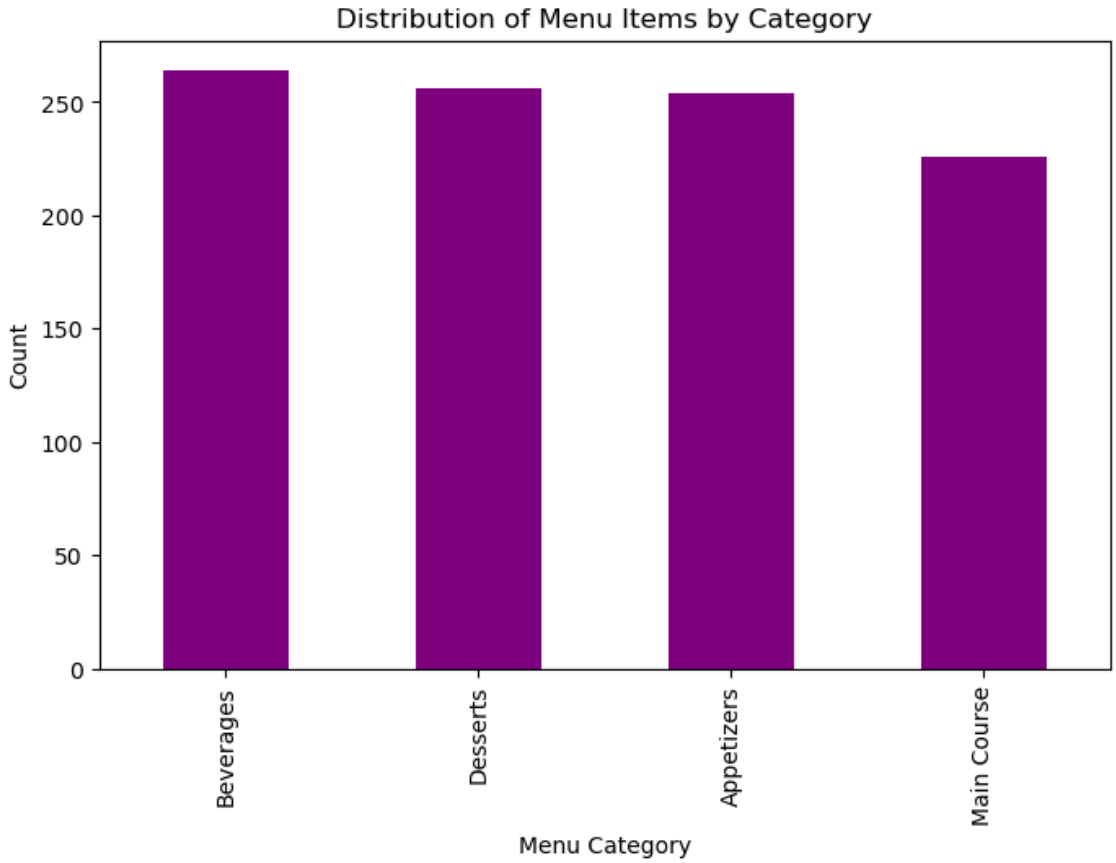
In [12]:
```
# Print the lat 5 rows  of data

print ("Tail of the dataset:")
print (data.tail())
```

```
Tail of the dataset:
    RestaurantID MenuCategory            MenuItem  \
995         R003    Beverages                Soda
996         R001   Appetizers        Caprese Salad
997         R003  Main Course  Vegetable Stir-Fry
998         R002     Desserts            Tiramisu
999         R001    Beverages                Soda

                                        Ingredients  Price Profitabili
ty
995                                ['confidential']   2.16           L
ow
996       ['Tomatoes', 'Basil', 'Garlic', 'Olive Oil']  11.86        Med
um
997  ['Chicken', 'Fettuccine', 'Alfredo Sauce', 'Pa...  20.57          Hi
gh
998           ['Chocolate', 'Butter', 'Sugar', 'Eggs']  18.80          Hi
gh
999                                ['confidential']   4.26        Medi
um
```

In [18]:
```
category_counts = data['MenuCategory'].value_counts()
plt.figure(figsize=(8, 5))
category_counts.plot(kind='bar', color='purple')
plt.title('Distribution of Menu Items by Category')
plt.xlabel('Menu Category')
plt.ylabel('Count')
plt.show()
```



Question 2: Average Price by Category

In [24]:
```
avg_price_by_category = data.groupby('MenuCategory')['Price'].mean().round(
print("Average Price by Category:")
print(avg_price_by_category)
```

```
Average Price by Category:
MenuCategory
Appetizers     11.42
Beverages       3.51
Desserts       14.97
Main Course    22.83
Name: Price, dtype: float64
```

Question 3: Price Range by Profitability

In [26]:
```
price_range_by_profitability = data.groupby('Profitability')['Price'].agg(
print("Price range by profitability:")
print(price_range_by_profitability)
```

```
Price range by profitability:
                min    max
Profitability
High           2.01  29.84
Low            2.01  28.32
Medium         2.25  29.54
```

Question 4: Most Frequent Ingredient

In [27]:
```
ingredient_lists = data['Ingredients'].apply(lambda x: eval(x))   #Convert
ingredients_flat = [ingredient for sublist in ingredient_lists for ingredie
ingredient_counts = Counter(ingredients_flat)
most_common_ingredient = ingredient_counts.most_common(1)
print ("Most common ingredient:", most_common_ingredient)
```

```
Most common ingredient: [('confidential', 264)]
```

Question 5: Restaurant with the Highest-Priced Item

In [31]:
```
highest_priced_item = data.loc[data['Price'].idxmax()]
print("Restaurant with the highesr-priced menu item:")
print(highest_priced_item)
```

```
Restaurant with the highesr-priced menu item:
RestaurantID                                                  R003
MenuCategory                                           Main Course
MenuItem                                          Chicken Alfredo
Ingredients      ['Chicken', 'Fettuccine', 'Alfredo Sauce', 'Pa...
Price                                                        29.84
Profitability                                                 High
Name: 406, dtype: object
```

Question 6: High Profitability, Below-Average Price

In [32]:
```
average_price_by_category = data.groupby('MenuCategory')['Price'].transform
below_avg_high_profit = data[(data['Price'] < average_price_by_category) &
print("High profitability items with below-average prices:")
print(below_avg_high_profit)
```

```
High profitability items with below-average prices:
    RestaurantID MenuCategory               MenuItem  \
30          R001    Beverages                 Coffee
33          R003   Appetizers  Spinach Artichoke Dip
41          R003     Desserts    New York Cheesecake
114         R001  Main Course          Shrimp Scampi
137         R002  Main Course          Grilled Steak
192         R001  Main Course          Shrimp Scampi
198         R001  Main Course          Grilled Steak
215         R001  Main Course          Grilled Steak
224         R002  Main Course        Chicken Alfredo
243         R003  Main Course        Chicken Alfredo
268         R001  Main Course     Vegetable Stir-Fry
280         R003     Desserts    New York Cheesecake
284         R003     Desserts    Chocolate Lava Cake
292         R003  Main Course          Grilled Steak
300         R001  Main Course          Shrimp Scampi
321         R002  Main Course     Vegetable Stir-Fry
333         R001  Main Course        Chicken Alfredo
```

Question 7: Average Number of Ingrediants Per Menu Item

In [36]:
```
data['IngredientCount'] = ingredient_lists.apply(len)
avg_ingredients = data['IngredientCount'].mean()
print(f"Average number of ingredient per menu item: {avg_ingredients:.2f}")
```

```
Average number of ingredient per menu item: 3.21
```

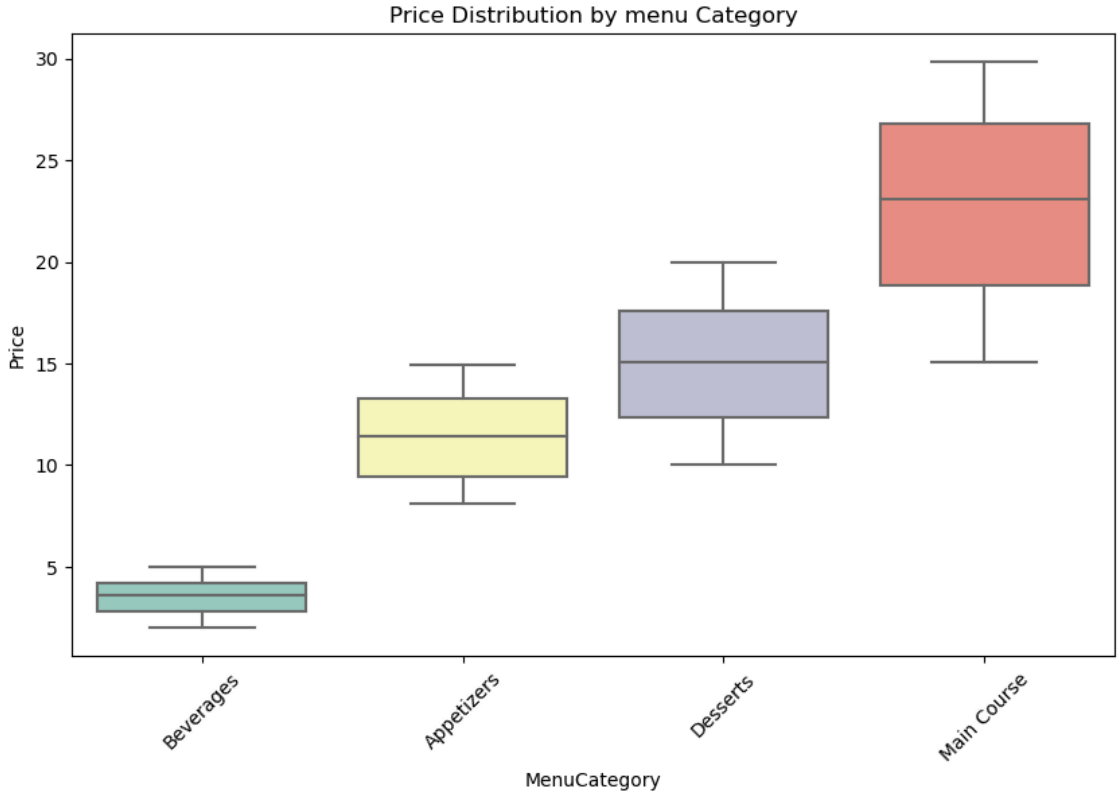
Question 8: Correlation Between Price and Profitability

In [38]:
```
profitability_map = {'Low': 1, 'Medium': 2, 'High': 3}
data['ProfitabilityScore'] = data['Profitability'].map(profitability_map)
correlation = data['Price'].corr(data['ProfitabilityScore'])
print(f"Correlation between price and profitability: {correlation:.2f}")
```

```
Correlation between price and profitability: 0.62
```

Question 9 : Box Plot for Price by Category

In [39]:
```
plt.figure(figsize=(10, 6))
sns.boxplot(x='MenuCategory', y ='Price', data=data, palette='Set3')
plt.title('Price Distribution by menu Category')
plt.xticks(rotation=45)
plt.show()
```



Question 10: Filter Items with Chicken and High Profitability

In [40]:
```
chicken_high_profit = data[data['Ingredients'].str.contains('Chicken') & (
print("Menu items with 'Chicken' as an ingredient and High profitability:")
print(chicken_high_profit)
```

```
Menu items with 'Chicken' as an ingredient and High profitability:
    RestaurantID MenuCategory            MenuItem  \
3           R003  Main Course     Chicken Alfredo
8           R003  Main Course       Grilled Steak
17          R002  Main Course       Grilled Steak
19          R002  Main Course       Grilled Steak
43          R001  Main Course  Vegetable Stir-Fry
..           ...          ...                 ...
978         R002  Main Course  Vegetable Stir-Fry
980         R001  Main Course     Chicken Alfredo
984         R001  Main Course  Vegetable Stir-Fry
992         R001  Main Course     Chicken Alfredo
997         R003  Main Course  Vegetable Stir-Fry

                                           Ingredients  Price Profitabili
ty  \
3    ['Chicken', 'Fettuccine', 'Alfredo Sauce', 'Pa...  29.55          Hi
gh
8    ['Chicken', 'Fettuccine', 'Alfredo Sauce', 'Pa...  26.78          Hi
gh
17   ['Chicken', 'Fettuccine', 'Alfredo Sauce', 'Pa...  23.52          Hi
gh
19   ['Chicken', 'Fettuccine', 'Alfredo Sauce', 'Pa...  28.90          Hi
gh
43   ['Chicken', 'Fettuccine', 'Alfredo Sauce', 'Pa...  27.06          Hi
gh
..                                                 ...    ...
...
978  ['Chicken', 'Fettuccine', 'Alfredo Sauce', 'Pa...  23.90          Hi
gh
980  ['Chicken', 'Fettuccine', 'Alfredo Sauce', 'Pa...  29.11          Hi
gh
984  ['Chicken', 'Fettuccine', 'Alfredo Sauce', 'Pa...  29.04          Hi
gh
992  ['Chicken', 'Fettuccine', 'Alfredo Sauce', 'Pa...  20.40          Hi
gh
997  ['Chicken', 'Fettuccine', 'Alfredo Sauce', 'Pa...  20.57          Hi
gh

     IngredientCount  ProfitabilityScore
3                  4                   3
8                  4                   3
17                 4                   3
19                 4                   3
43                 4                   3
..               ...                 ...
978                4                   3
980                4                   3
984                4                   3
992                4                   3
997                4                   3

[143 rows x 8 columns]
```