In [1]:
```
pip install pandas seaborn matplotlib statsmodels patsy
```

```
Requirement already satisfied: pandas in c:\users\nihar\anaconda3\lib\site-packages (2.1.4)
Requirement already satisfied: seaborn in c:\users\nihar\anaconda3\lib\site-packages (0.12.2)
Requirement already satisfied: matplotlib in c:\users\nihar\anaconda3\lib\site-packages (3.8.0)
Requirement already satisfied: statsmodels in c:\users\nihar\anaconda3\lib\site-packages (0.14.0)
Requirement already satisfied: patsy in c:\users\nihar\anaconda3\lib\site-packages (0.5.3)
Requirement already satisfied: numpy<2,>=1.23.2 in c:\users\nihar\anaconda3\lib\site-packages (from pandas) (1.26.4)
Requirement already satisfied: python-dateutil>=2.8.2 in c:\users\nihar\anaconda3\lib\site-packages (from pandas) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in c:\users\nihar\anaconda3\lib\site-packages (from pandas) (2023.3.post1)
Requirement already satisfied: tzdata>=2022.1 in c:\users\nihar\anaconda3\lib\site-packages (from pandas) (2023.3)
Requirement already satisfied: contourpy>=1.0.1 in c:\users\nihar\anaconda3\lib\site-packages (from matplotlib) (1.2.0)
Requirement already satisfied: cycler>=0.10 in c:\users\nihar\anaconda3\lib\site-packages (from matplotlib) (0.11.0)
Requirement already satisfied: fonttools>=4.22.0 in c:\users\nihar\anaconda3\lib\site-packages (from matplotlib) (4.25.0)
Requirement already satisfied: kiwisolver>=1.0.1 in c:\users\nihar\anaconda3\lib\site-packages (from matplotlib) (1.4.4)
Requirement already satisfied: packaging>=20.0 in c:\users\nihar\anaconda3\lib\site-packages (from matplotlib) (23.1)
Requirement already satisfied: pillow>=6.2.0 in c:\users\nihar\anaconda3\lib\site-packages (from matplotlib) (10.2.0)
Requirement already satisfied: pyparsing>=2.3.1 in c:\users\nihar\anaconda3\lib\site-packages (from matplotlib) (3.0.9)
Requirement already satisfied: scipy!=1.9.2,>=1.4 in c:\users\nihar\anaconda3\lib\site-packages (from statsmodels) (1.11.4)
Requirement already satisfied: six in c:\users\nihar\anaconda3\lib\site-packages (from patsy) (1.16.0)
Note: you may need to restart the kernel to use updated packages.
```

In [7]:
```python
# Load necessary packages
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
import statsmodels.api as sm
from patsy import dmatrices
```

In [9]:
```python
# Load the data
file_path = "C:/Users/nihar/OneDrive/Desktop/Bootcamp/SCMA 632/DataSet/pizza_data.csv"
pizza_data = pd.read_csv(file_path)
```

In [11]:
```python
# Convert columns to category (equivalent to factors in R)
factor_columns = ['brand', 'price', 'weight', 'crust', 'cheese', 'size', 'toppings', 'spicy']
for column in factor_columns:
    pizza_data[column] = pizza_data[column].astype('category')
```

In [13]:
```python
# Create profiles
profiles = pizza_data.drop(columns=['ranking'])
```

In [15]:
```python
# Create preferences dataset
preferences = pd.DataFrame({
    'Respondent': pizza_data.index.repeat(len(profiles)),
    'Profile': pd.Series(range(len(profiles))).repeat(len(pizza_data)),
    'Rating': pizza_data['ranking'].repeat(len(profiles))
})
```

In [17]:
```python
# Merge profiles with preferences
conjoint_data = preferences.merge(profiles, left_on='Profile', right_index=True)
```

In [19]:
```python
# Perform conjoint analysis using OLS (equivalent to lm in R)
formula = 'Rating ~ ' + ' + '.join(factor_columns)
y, X = dmatrices(formula, data=conjoint_data, return_type='dataframe')
conjoint_model = sm.OLS(y, X).fit()
```

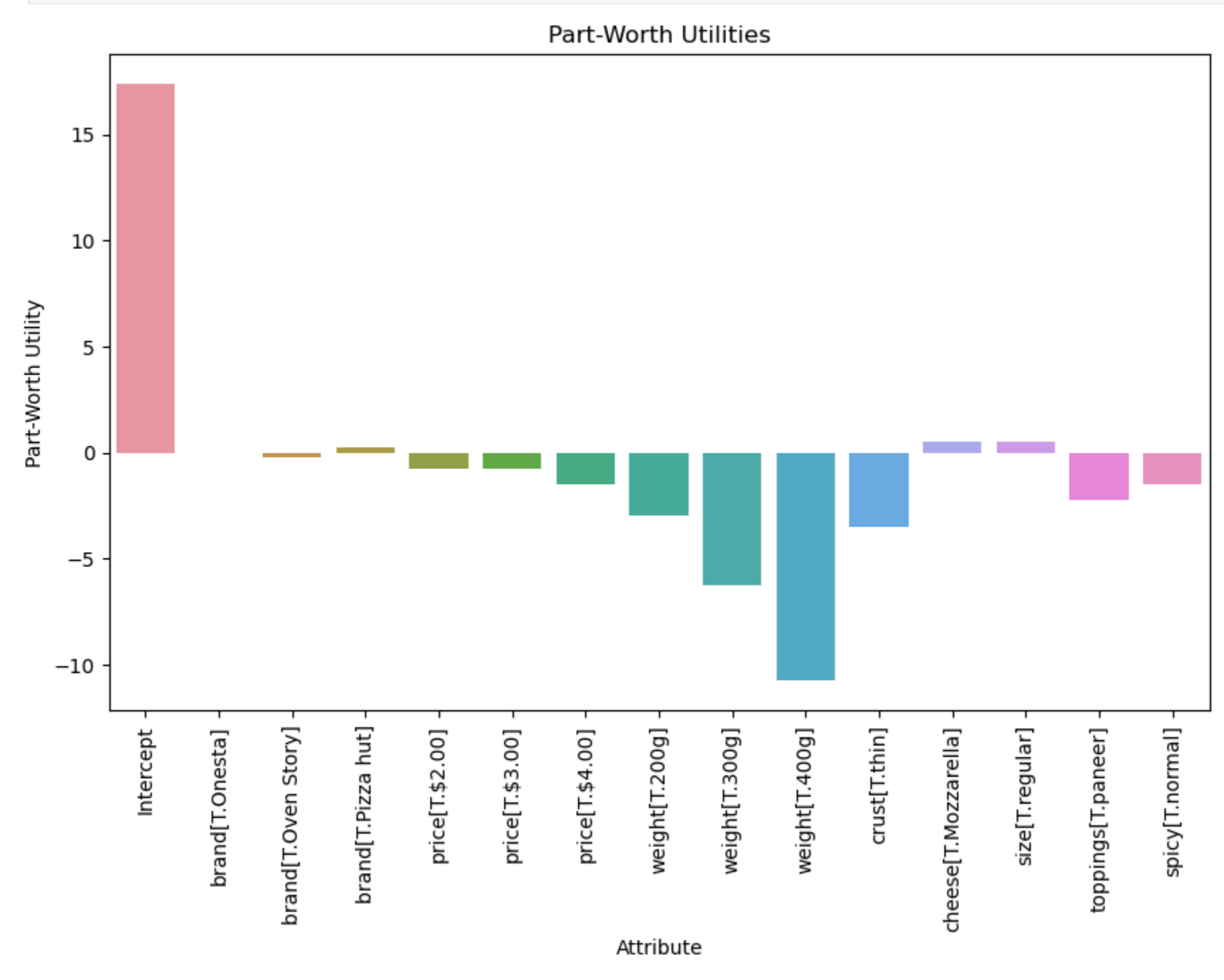
In [21]:
```python
# Summary of results
print(conjoint_model.summary())
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                 Rating   R-squared:                       0.999
Model:                            OLS   Adj. R-squared:                  0.999
Method:                 Least Squares   F-statistic:                 2.339e+04
Date:                Tue, 09 Jul 2024   Prob (F-statistic):               0.00
Time:                        07:13:16   Log-Likelihood:                 169.09
No. Observations:                 256   AIC:                            -308.2
Df Residuals:                     241   BIC:                            -255.0
Df Model:                          14
Covariance Type:            nonrobust
=========================================================================================
                            coef    std err          t      P>|t|      [0.025      0.975]
-----------------------------------------------------------------------------------------
Intercept              17.3750      0.031    557.157      0.000      17.314      17.436
brand[T.Onesta]      9.437e-15      0.023   4.14e-13      1.000      -0.045       0.045
brand[T.Oven Story]    -0.2500      0.023    -10.977      0.000      -0.295      -0.205
brand[T.Pizza hut]      0.2500      0.023     10.977      0.000       0.205       0.295
price[T.$2.00]         -0.7500      0.023    -32.932      0.000      -0.795      -0.705
price[T.$3.00]         -0.7500      0.023    -32.932      0.000      -0.795      -0.705
price[T.$4.00]         -1.5000      0.023    -65.863      0.000      -1.545      -1.455
weight[T.200g]         -3.0000      0.023   -131.727      0.000      -3.045      -2.955
weight[T.300g]         -6.2500      0.023   -274.431      0.000      -6.295      -6.205
weight[T.400g]        -10.7500      0.023   -472.022      0.000     -10.795     -10.705
crust[T.thin]          -3.5000      0.016   -217.338      0.000      -3.532      -3.468
cheese[T.Mozzarella]    0.5000      0.016     31.048      0.000       0.468       0.532
size[T.regular]         0.5000      0.016     31.048      0.000       0.468       0.532
toppings[T.paneer]     -2.2500      0.016   -139.718      0.000      -2.282      -2.218
spicy[T.normal]        -1.5000      0.016    -93.145      0.000      -1.532      -1.468
==============================================================================
Omnibus:                     1257.426   Durbin-Watson:                   0.125
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               42.667
Skew:                           0.000   Prob(JB):                     5.43e-10
Kurtosis:                       1.000   Cond. No.                         9.06
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

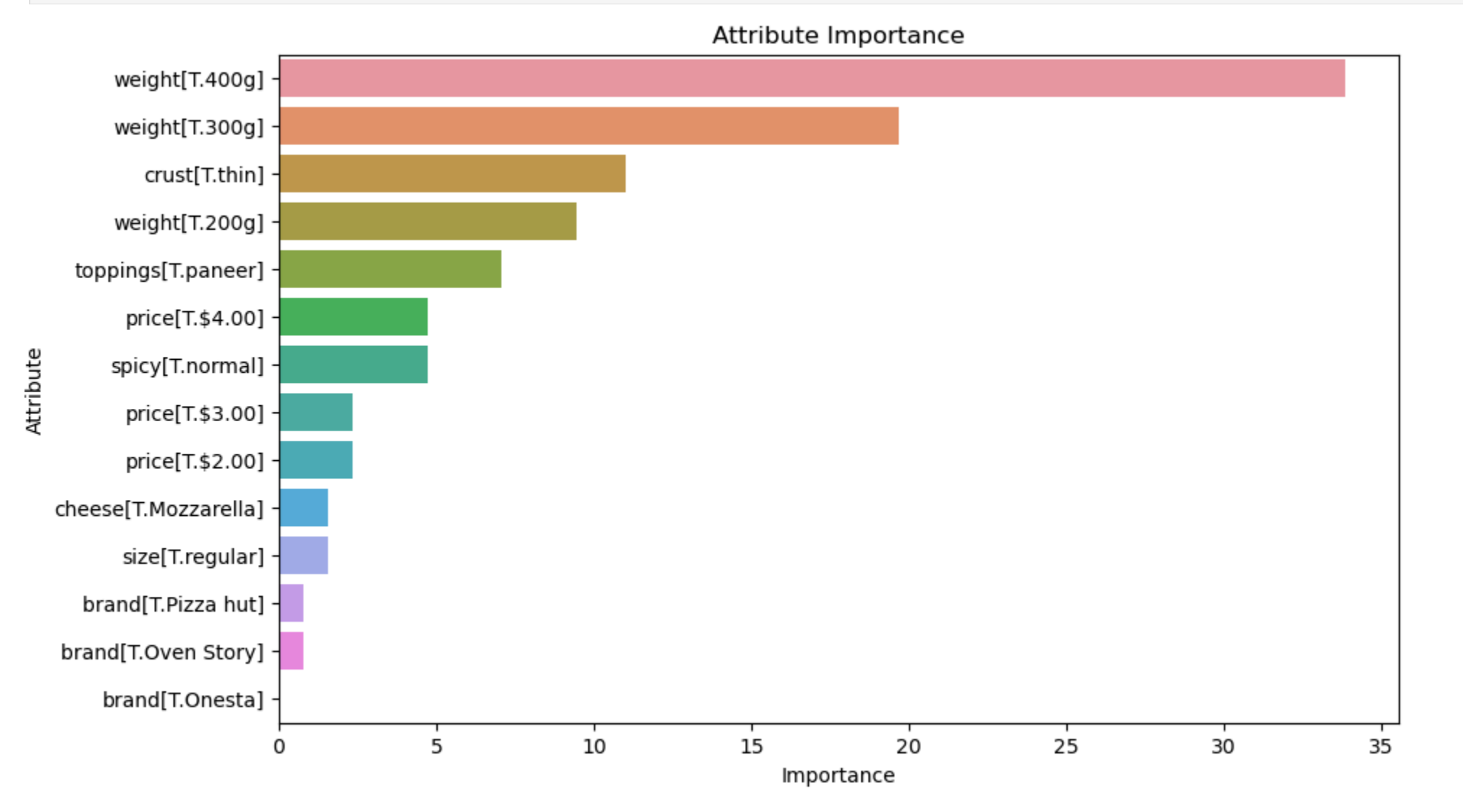
In [23]:
```python
# Extract part-worth utilities
part_worths = conjoint_model.params
part_worths_df = part_worths.reset_index()
part_worths_df.columns = ['Attribute', 'PartWorth']
```

In [25]:
```python
# Plot part-worth utilities
plt.figure(figsize=(10, 6))
sns.barplot(x='Attribute', y='PartWorth', data=part_worths_df)
plt.title("Part-Worth Utilities")
plt.xlabel("Attribute")
plt.ylabel("Part-Worth Utility")
plt.xticks(rotation=90)
plt.show()
```



In [27]:
```python
# Compute importance
importance = part_worths.drop('Intercept').abs() / part_worths.drop('Intercept').abs().sum() * 100
importance_df = importance.reset_index()
importance_df.columns = ['Attribute', 'Importance']
```

In [29]:
```python
# Plot attribute importance
plt.figure(figsize=(10, 6))
sns.barplot(x='Importance', y='Attribute', data=importance_df.sort_values('Importance', ascending=False))
plt.title("Attribute Importance")
plt.xlabel("Importance")
plt.ylabel("Attribute")
plt.show()
```

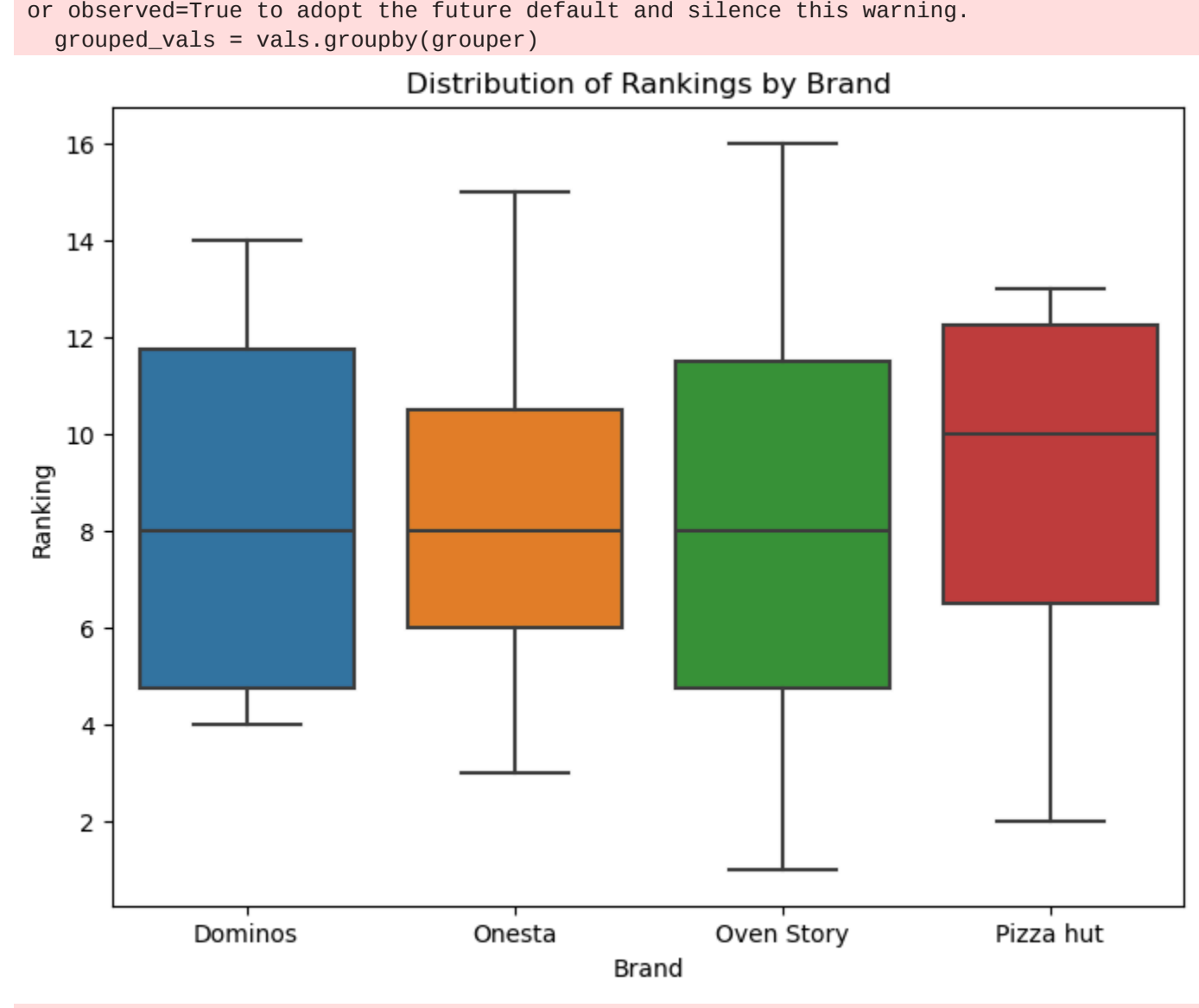


In [31]:
```python
# Create boxplots for each attribute
attributes = ['brand', 'price', 'weight', 'crust', 'cheese', 'size', 'toppings', 'spicy']
```

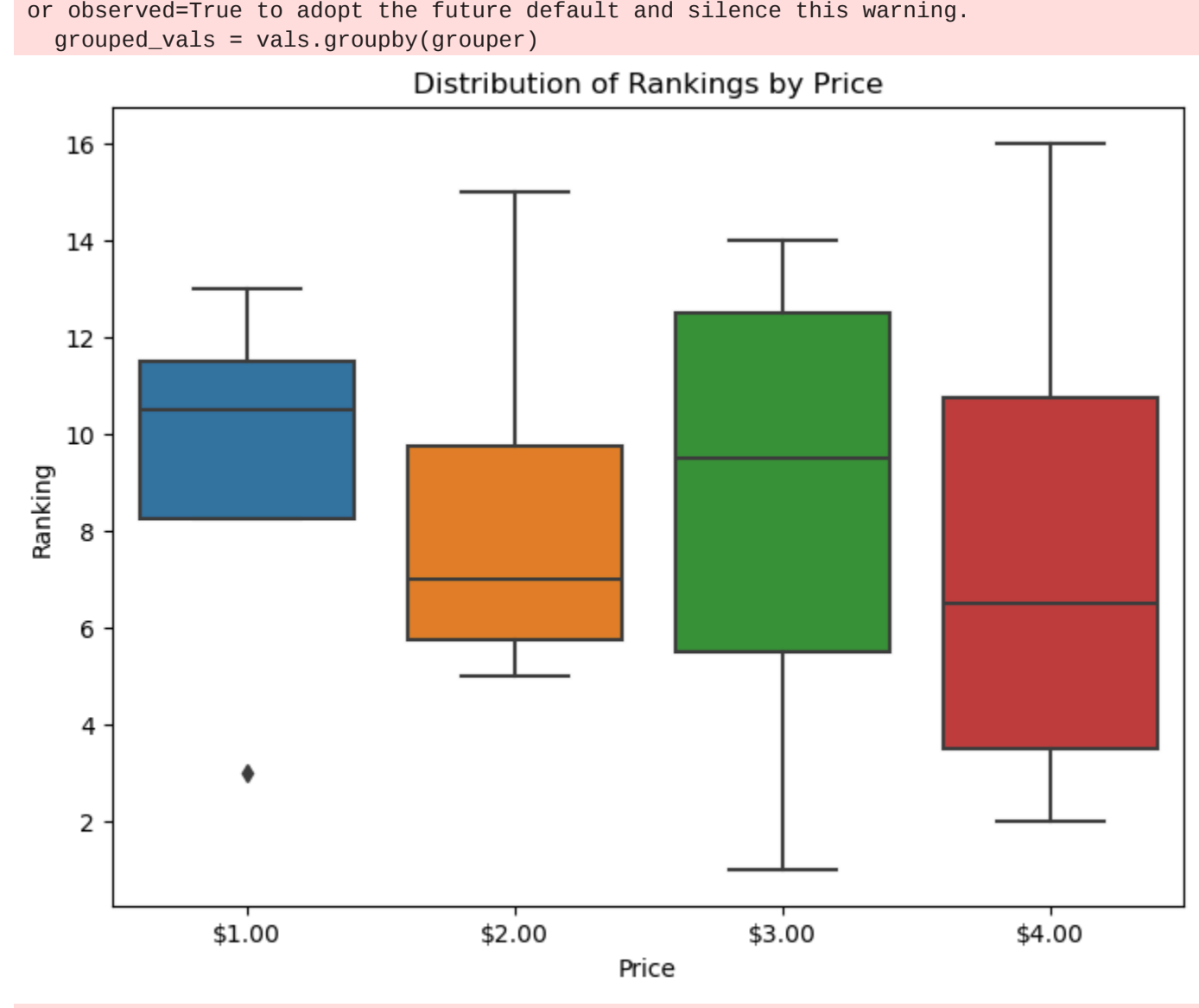
In [33]:
```python
# Function to create a boxplot for an attribute
def create_boxplot(attribute):
    plt.figure(figsize=(8, 6))
    sns.boxplot(x=attribute, y='ranking', data=pizza_data)
    plt.title(f"Distribution of Rankings by {attribute.capitalize()}")
    plt.xlabel(attribute.capitalize())
    plt.ylabel("Ranking")
    plt.show()
```

In [35]:
```python
# Generate boxplots for all attributes
for attribute in attributes:
    create_boxplot(attribute)
```

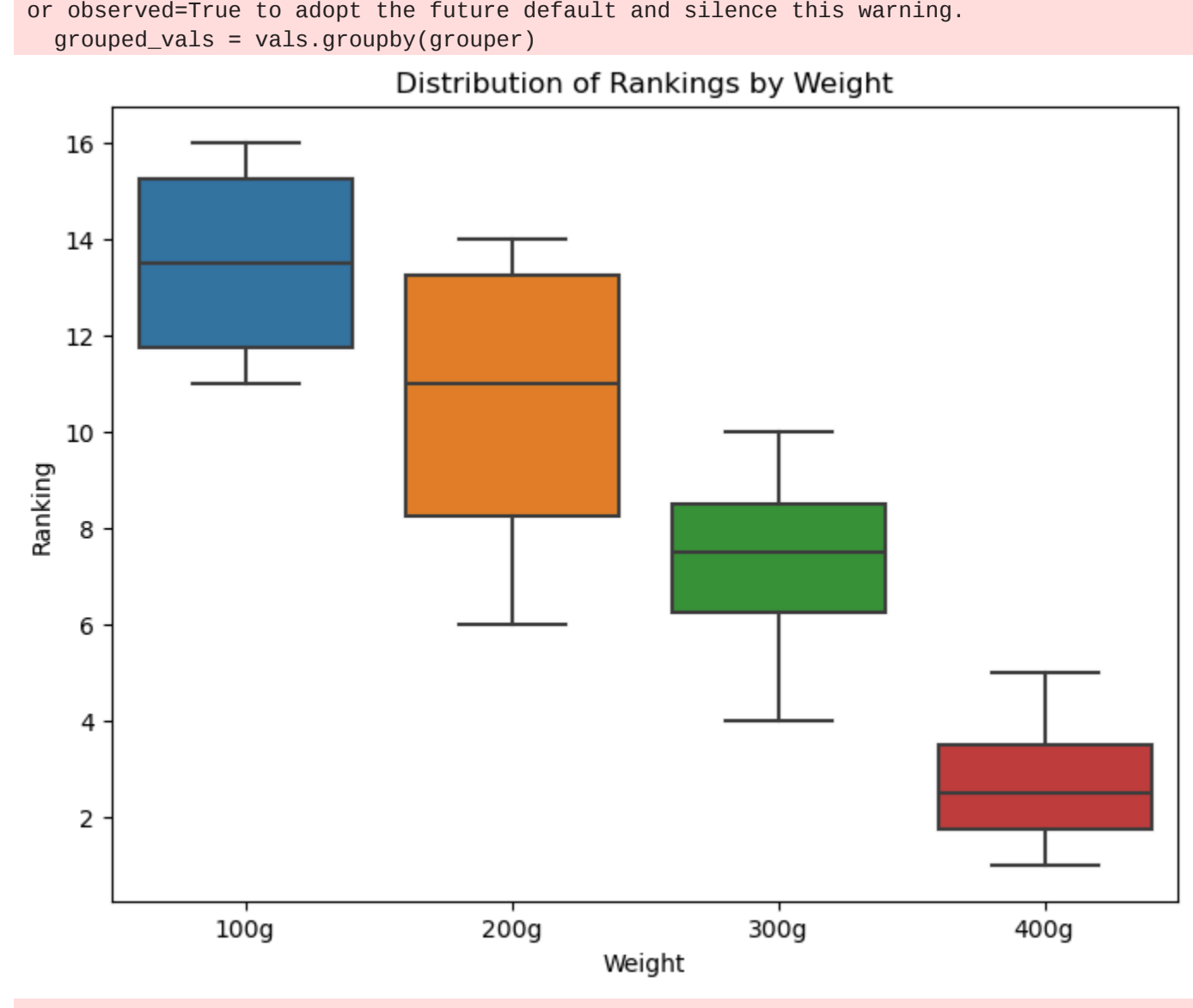
```
C:\Users\nihar\anaconda3\Lib\site-packages\seaborn\categorical.py:641: FutureWarning: The default of observed=False is deprecated and will be changed to True in a future version of pandas. Pass observed=False to retain current behavior or observed=True to adopt the future default and silence this warning.
  grouped_vals = vals.groupby(grouper)
```



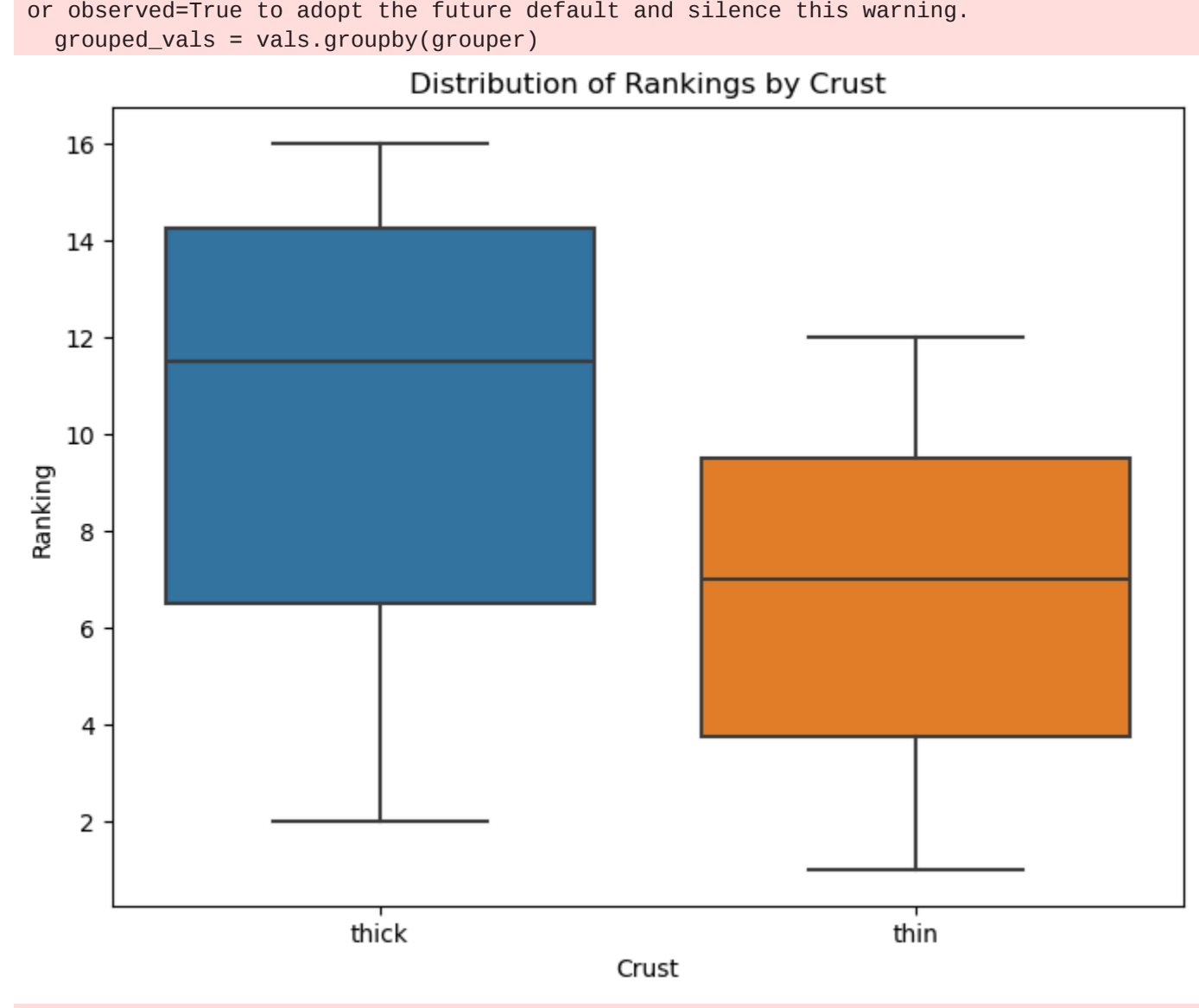

```
C:\Users\nihar\anaconda3\Lib\site-packages\seaborn\categorical.py:641: FutureWarning: The default of observed=False is deprecated and will be changed to True in a future version of pandas. Pass observed=False to retain current behavior or observed=True to adopt the future default and silence this warning.
  grouped_vals = vals.groupby(grouper)
```



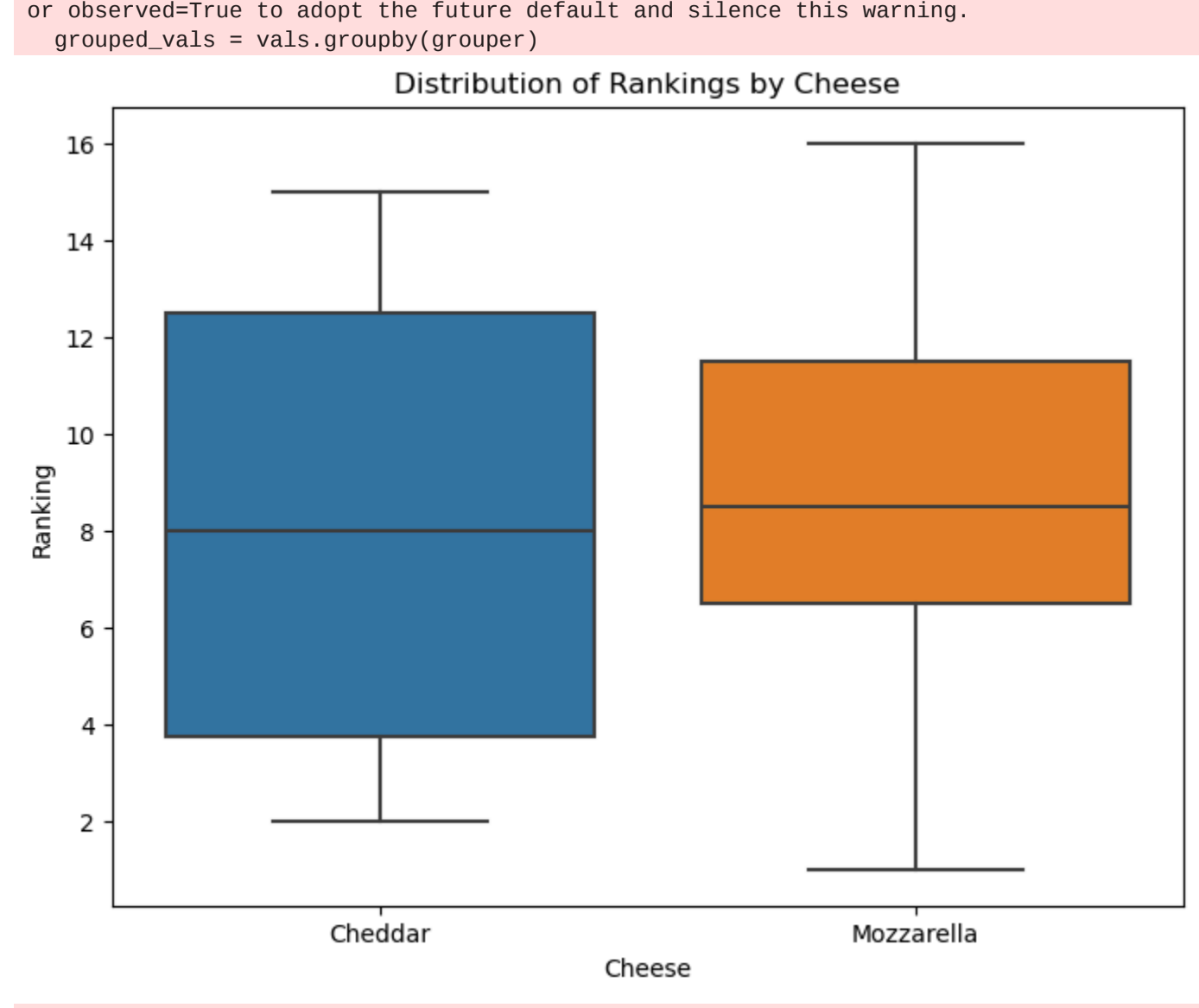

```
C:\Users\nihar\anaconda3\Lib\site-packages\seaborn\categorical.py:641: FutureWarning: The default of observed=False is deprecated and will be changed to True in a future version of pandas. Pass observed=False to retain current behavior or observed=True to adopt the future default and silence this warning.
  grouped_vals = vals.groupby(grouper)
```



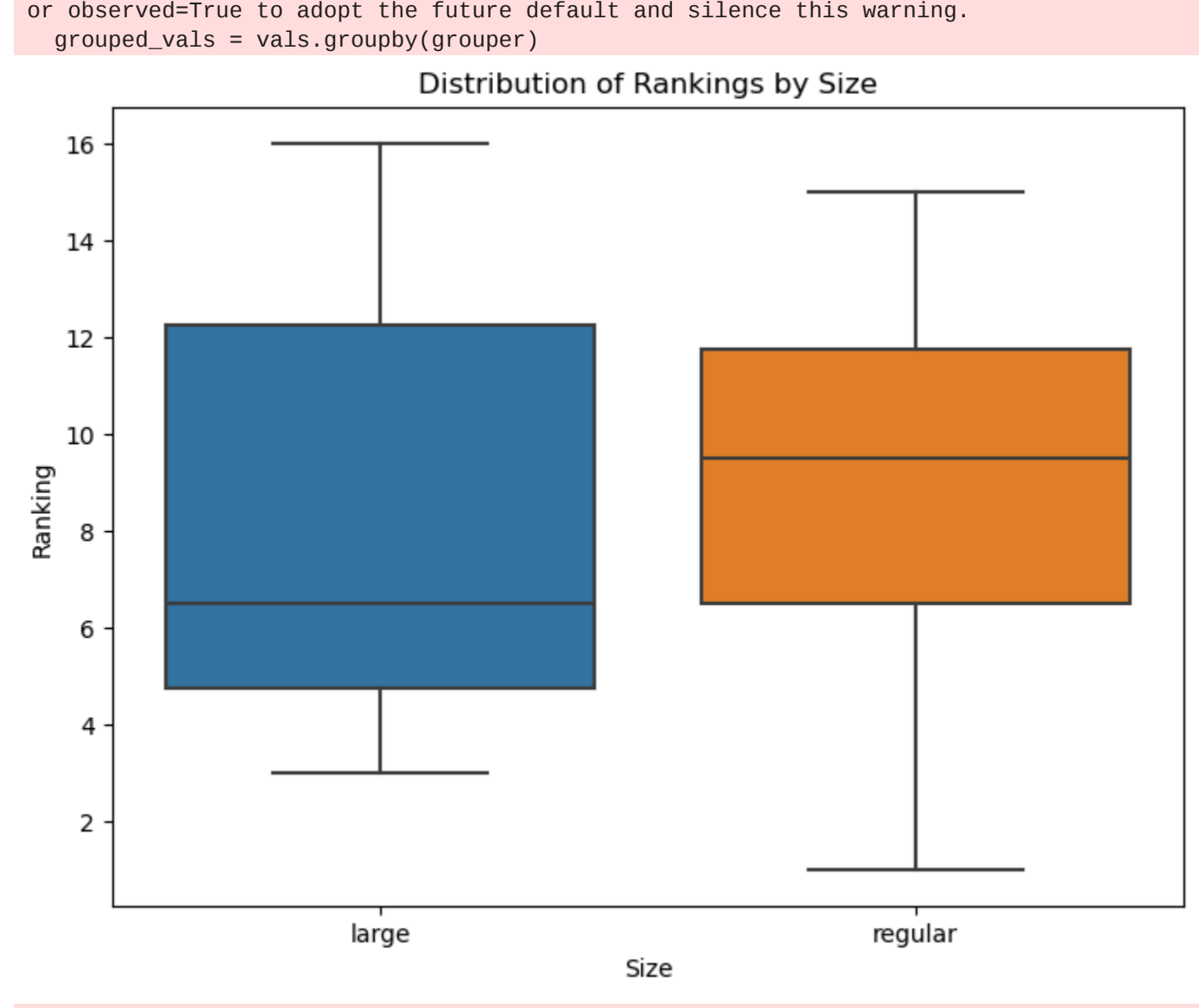

```
C:\Users\nihar\anaconda3\Lib\site-packages\seaborn\categorical.py:641: FutureWarning: The default of observed=False is deprecated and will be changed to True in a future version of pandas. Pass observed=False to retain current behavior or observed=True to adopt the future default and silence this warning.
  grouped_vals = vals.groupby(grouper)
```



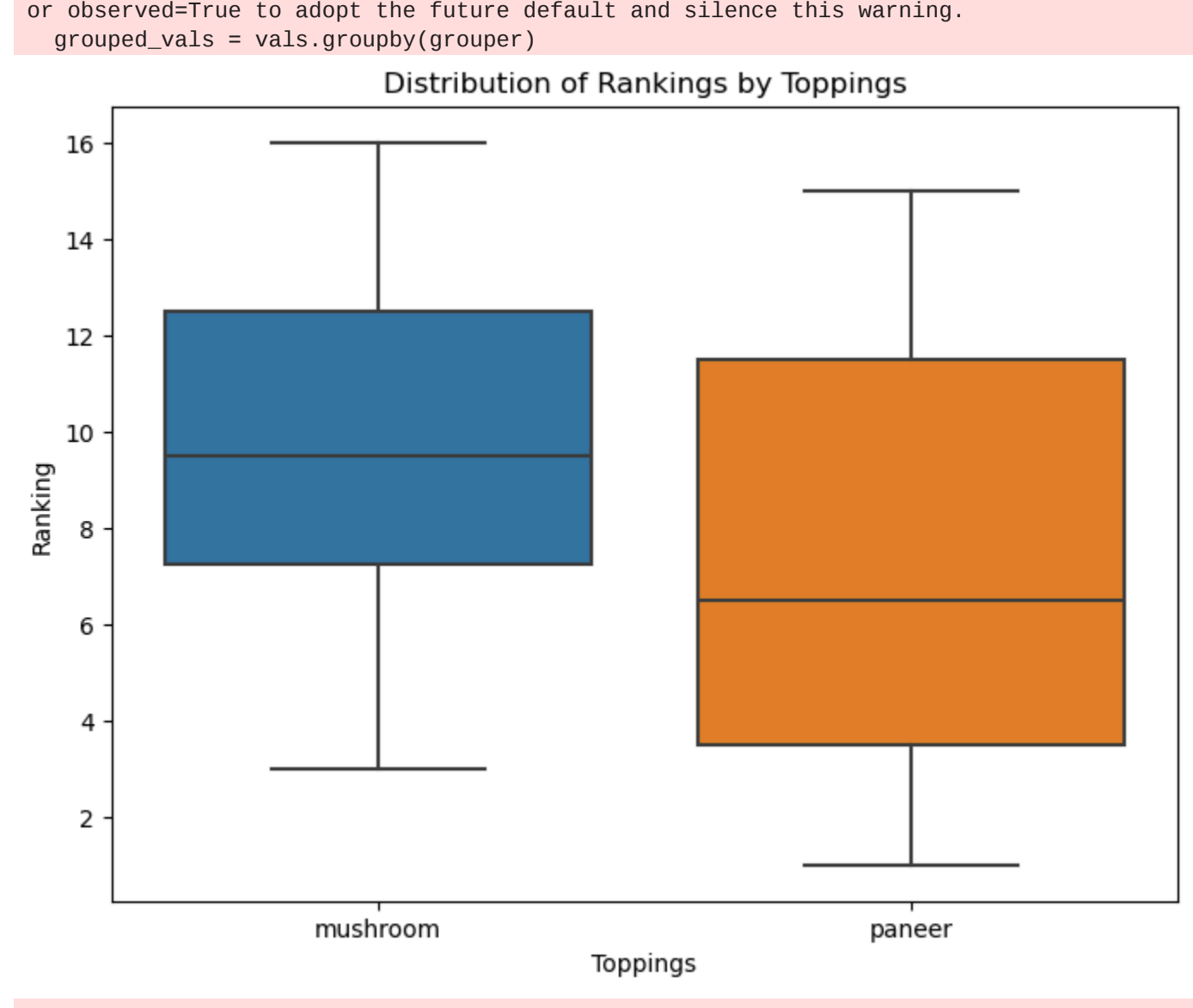

```
C:\Users\nihar\anaconda3\Lib\site-packages\seaborn\categorical.py:641: FutureWarning: The default of observed=False is deprecated and will be changed to True in a future version of pandas. Pass observed=False to retain current behavior or observed=True to adopt the future default and silence this warning.
  grouped_vals = vals.groupby(grouper)
```



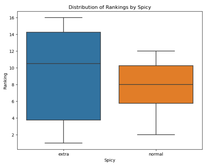

```
C:\Users\nihar\anaconda3\Lib\site-packages\seaborn\categorical.py:641: FutureWarning: The default of observed=False is deprecated and will be changed to True in a future version of pandas. Pass observed=False to retain current behavior or observed=True to adopt the future default and silence this warning.
  grouped_vals = vals.groupby(grouper)
```




```
C:\Users\nihar\anaconda3\Lib\site-packages\seaborn\categorical.py:641: FutureWarning: The default of observed=False is deprecated and will be changed to True in a future version of pandas. Pass observed=False to retain current behavior or observed=True to adopt the future default and silence this warning.
  grouped_vals = vals.groupby(grouper)
```




```
C:\Users\nihar\anaconda3\Lib\site-packages\seaborn\categorical.py:641: FutureWarning: The default of observed=False is deprecated and will be changed to True in a future version of pandas. Pass observed=False to retain current behavior or observed=True to adopt the future default and silence this warning.
  grouped_vals = vals.groupby(grouper)
```


Distribution of Rankings by Spicy
Ranking
16
14
12
10
8
6
4
2
extra
normal
Spicy

[ ]