

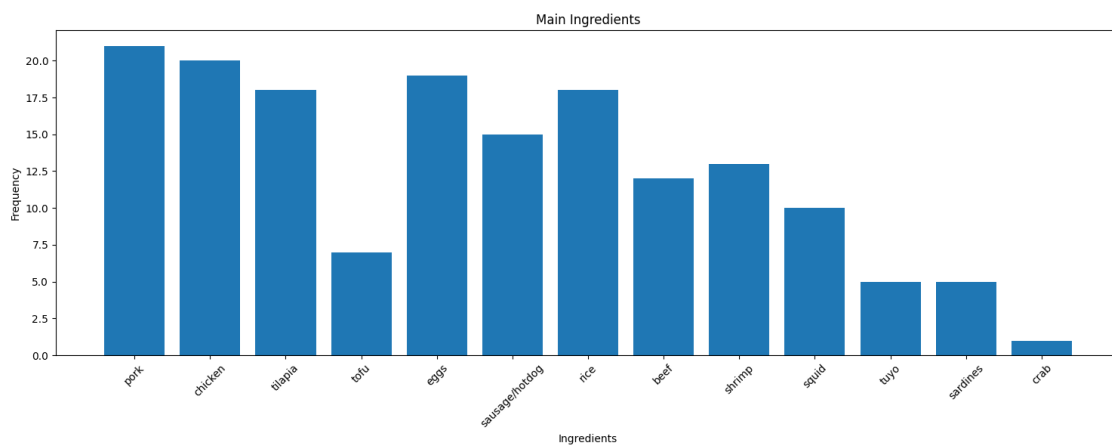
# clustering

March 6, 2025

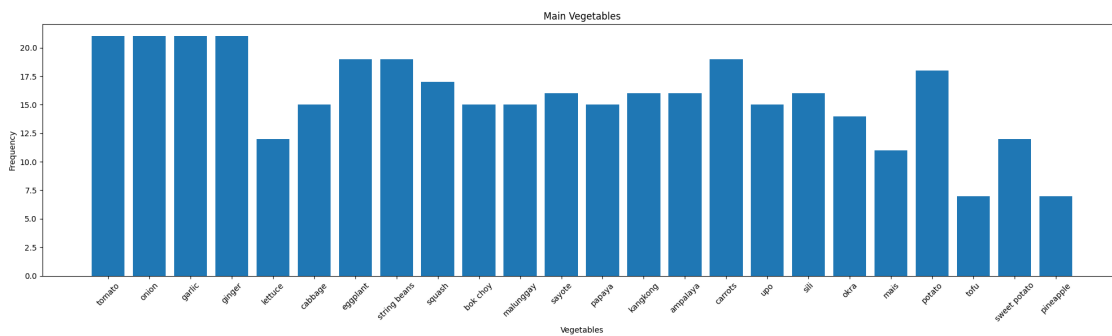
Import libraries

Read the raw dataset

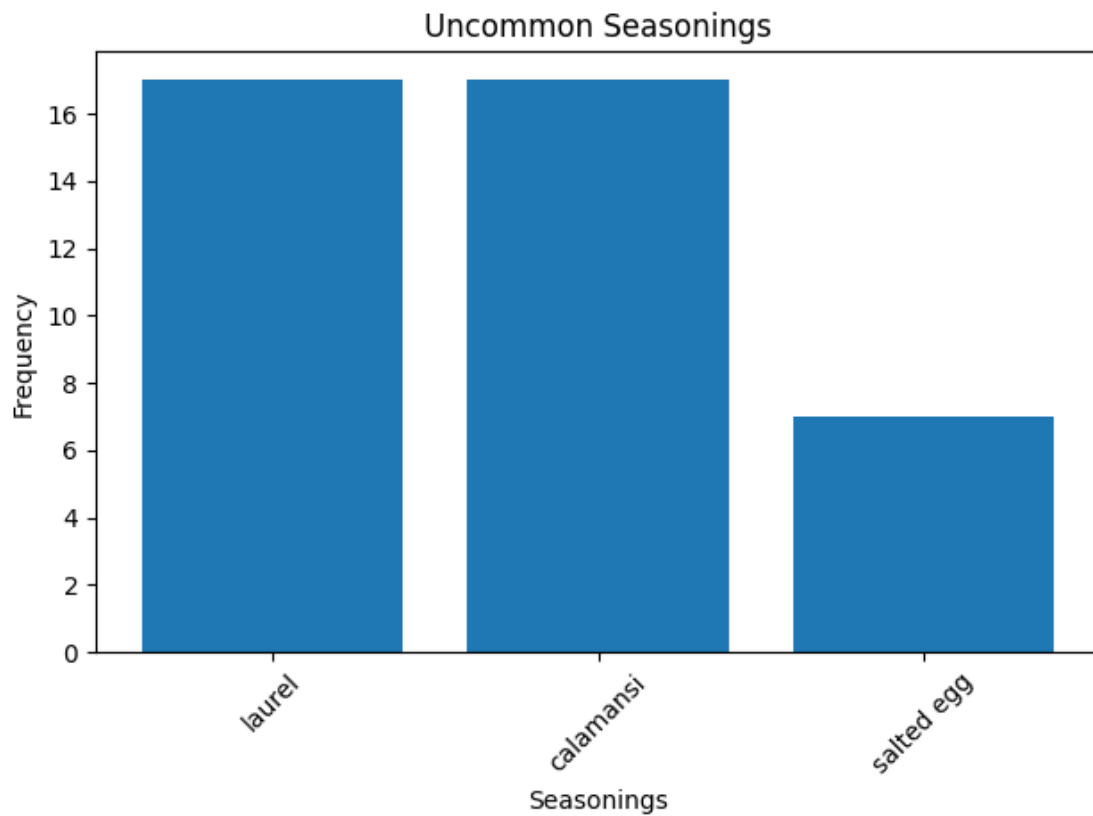
Process the feature 'main\_ingredients' - check the frequency



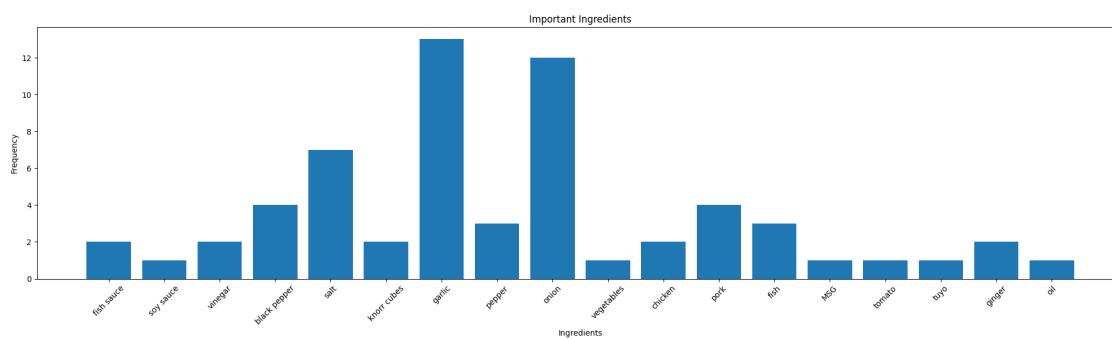
Process the feature 'vegetables'



Process the feature 'seasonings'

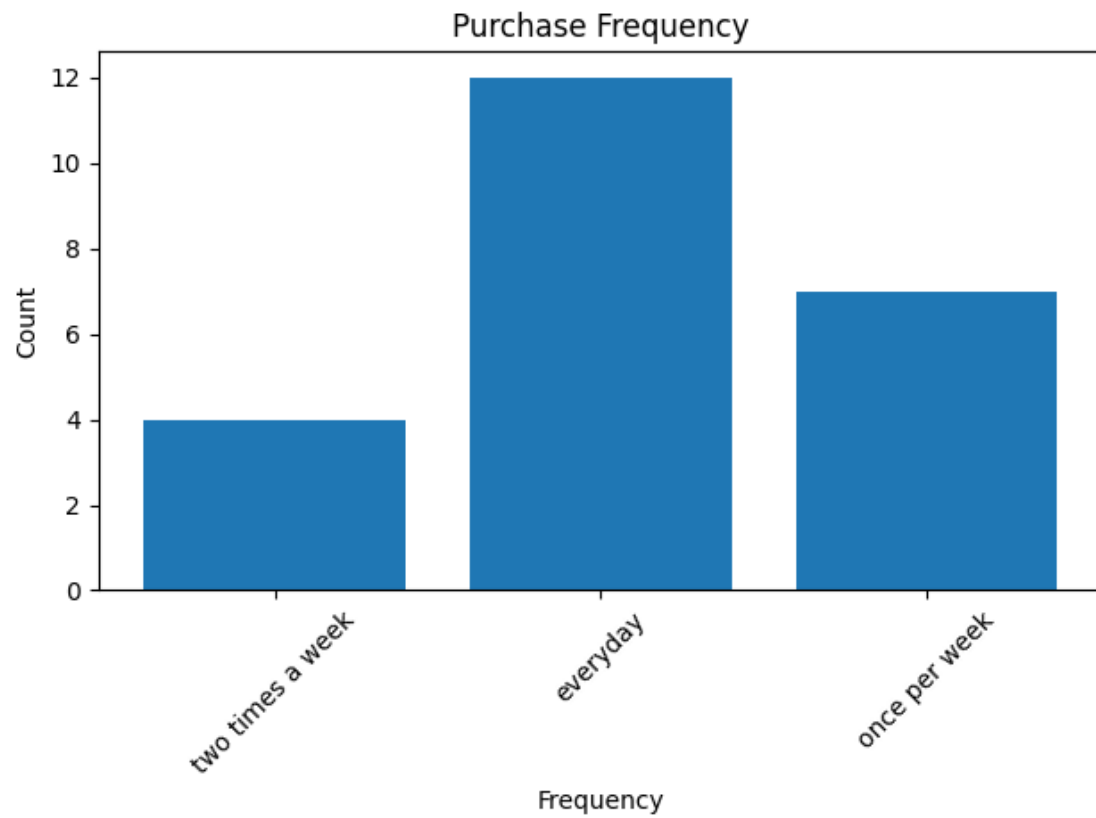


Process 'important\_ingredients'



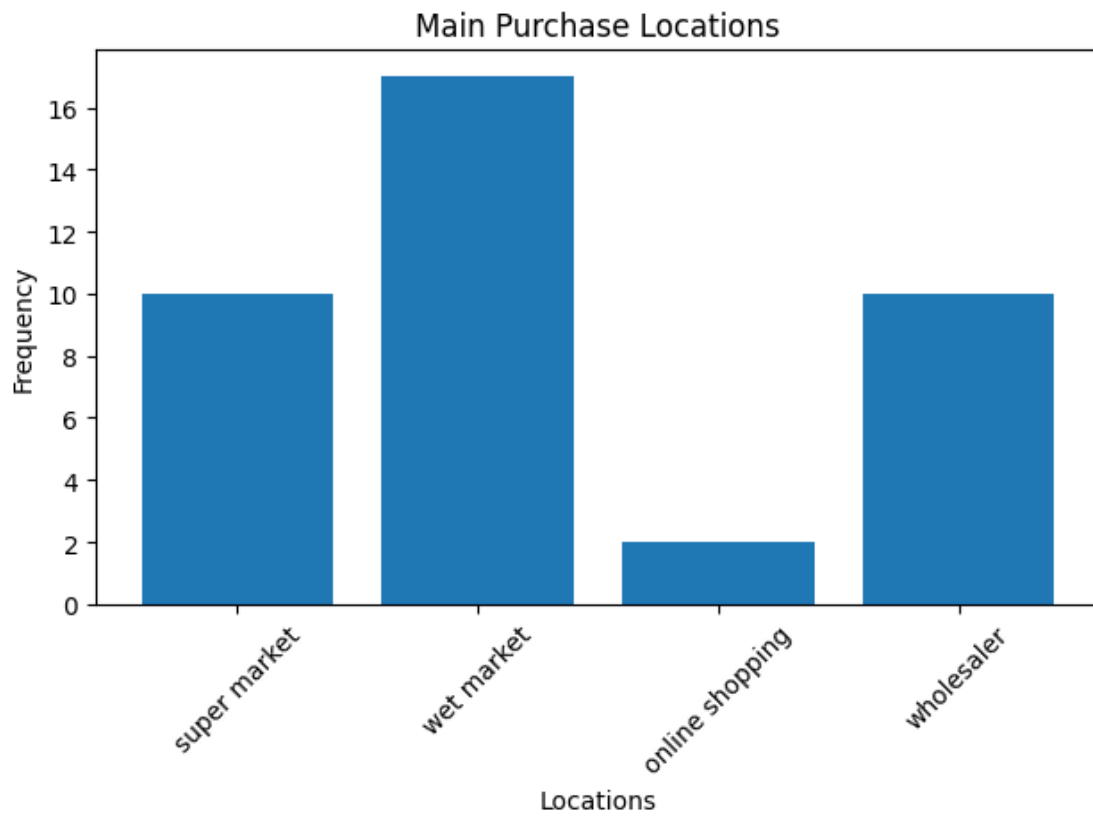
Process 'purchase\_frequency'

```
Counter({'everyday': 12, 'once per week': 7, 'two times a week': 4})
```



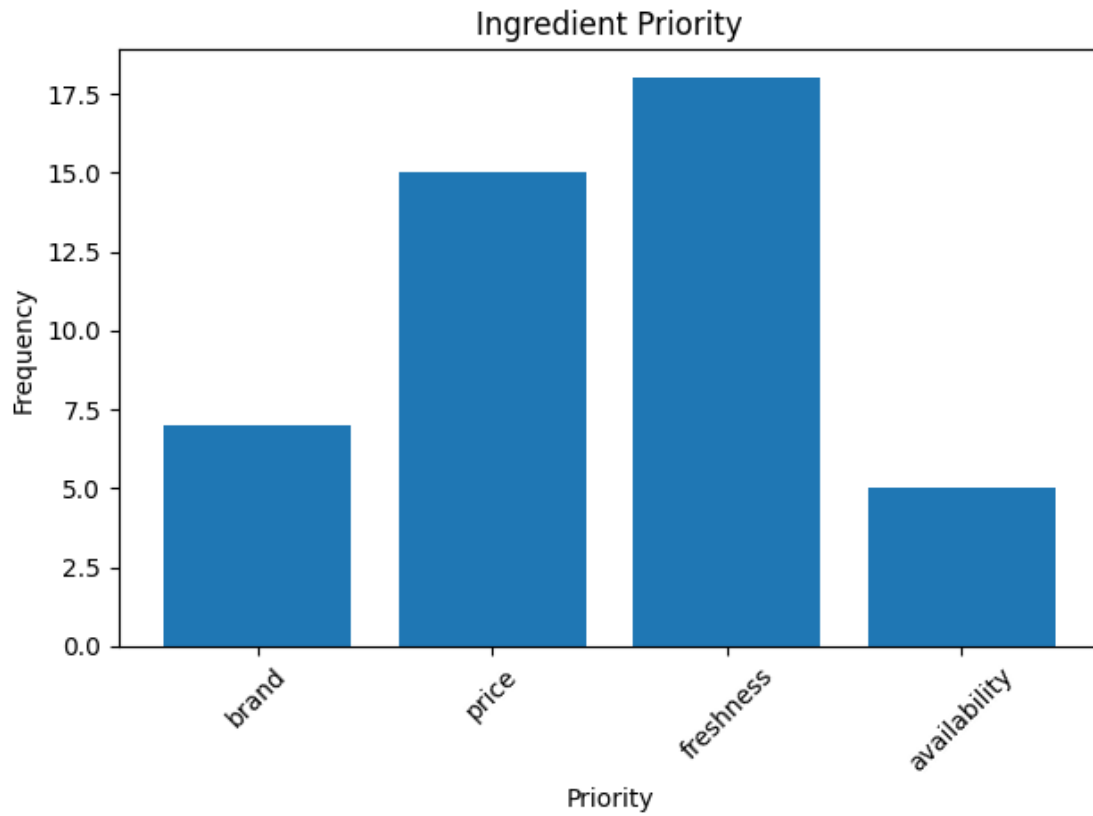
Process 'purchase location'

```
Counter({'wet market': 17, 'super market': 10, 'wholesaler': 10, 'online shopping': 2})
```



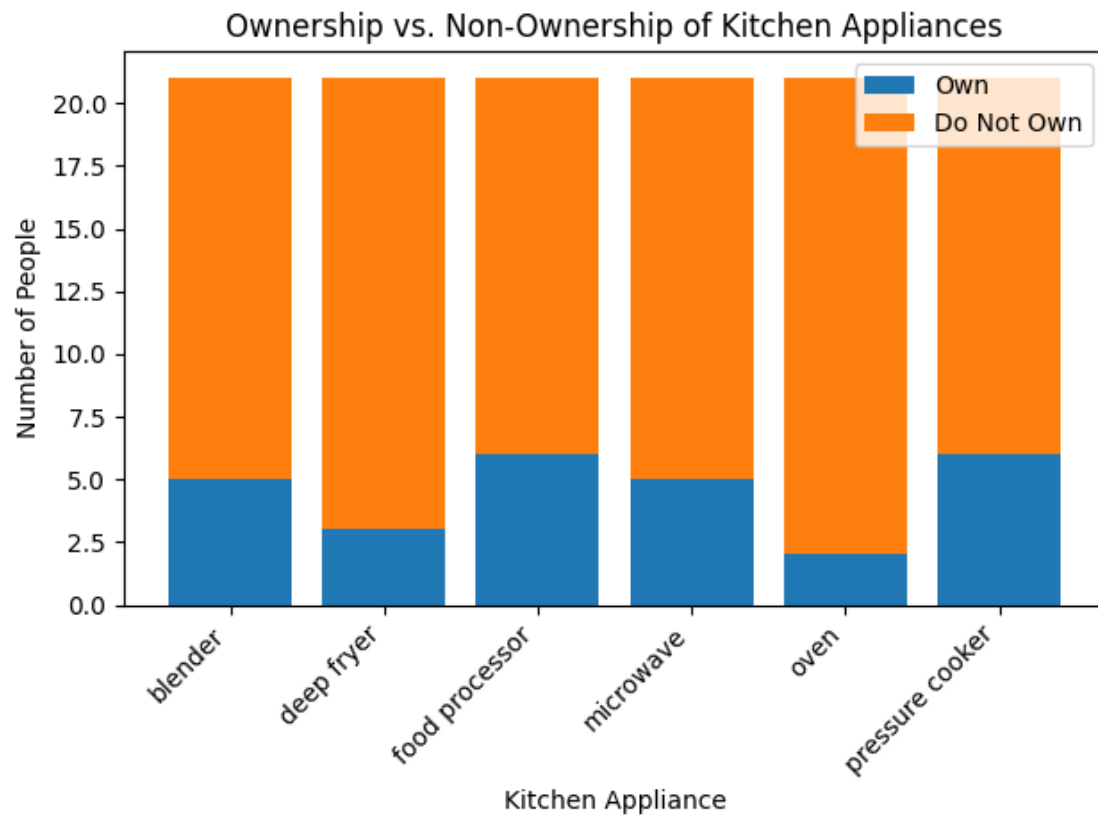
Process 'ingredient\_priority'

```
Counter({'freshness': 18, 'price': 15, 'brand': 7, 'availability': 5})
```



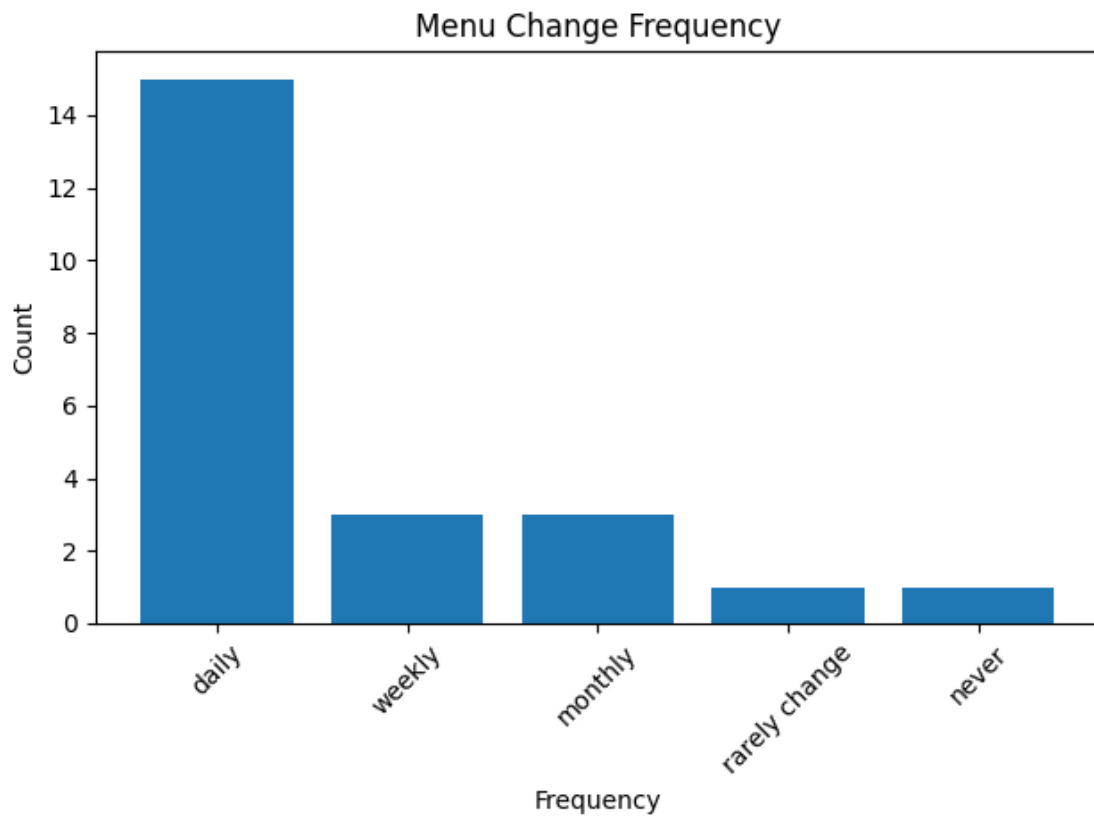
Process 'kitchen\_appliances' - used to assess whether they use uncommon appliances - no appliance = total\_respondents - owned

```
Counter({'food processor': 6, 'pressure cooker': 6, 'microwave': 5, 'blender': 5, 'deep fryer': 3, 'oven': 2})
```



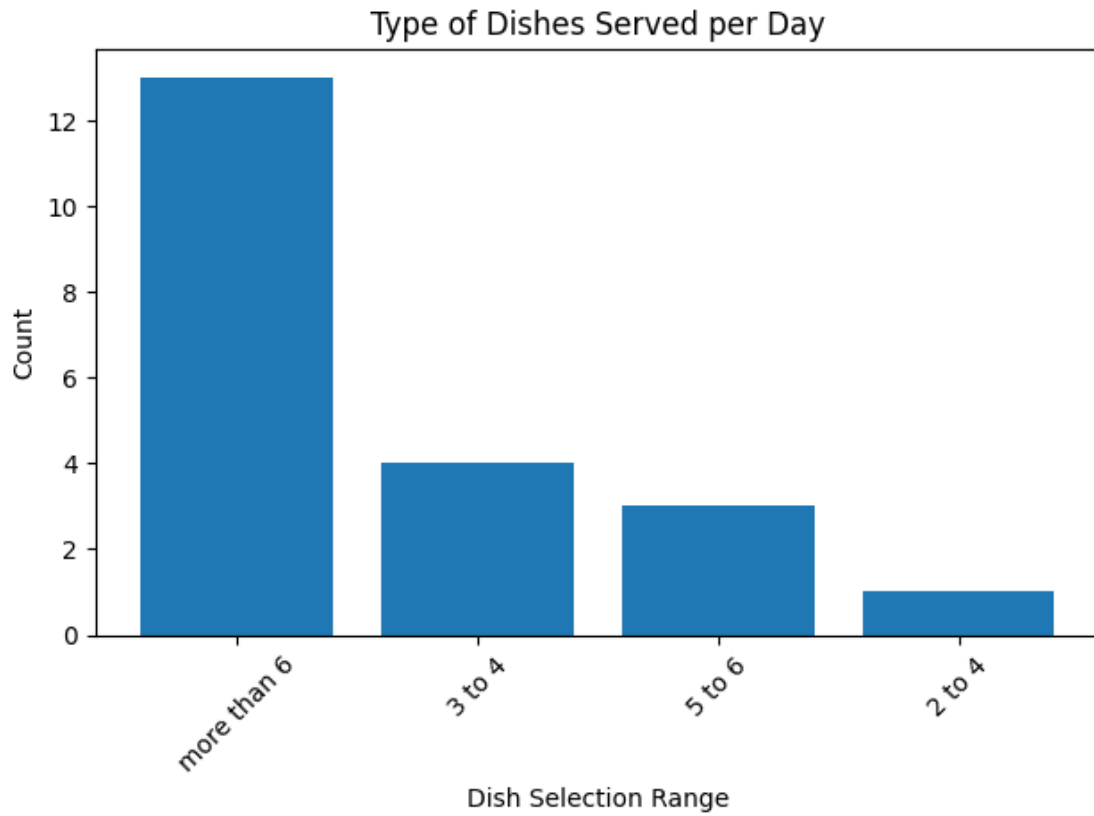
Process 'menu change\_frequency'

```
Counter({'daily': 15, 'weekly': 3, 'monthly': 3, 'rarely change': 1, 'never': 1})
```



Process 'daily\_dish\_selection'

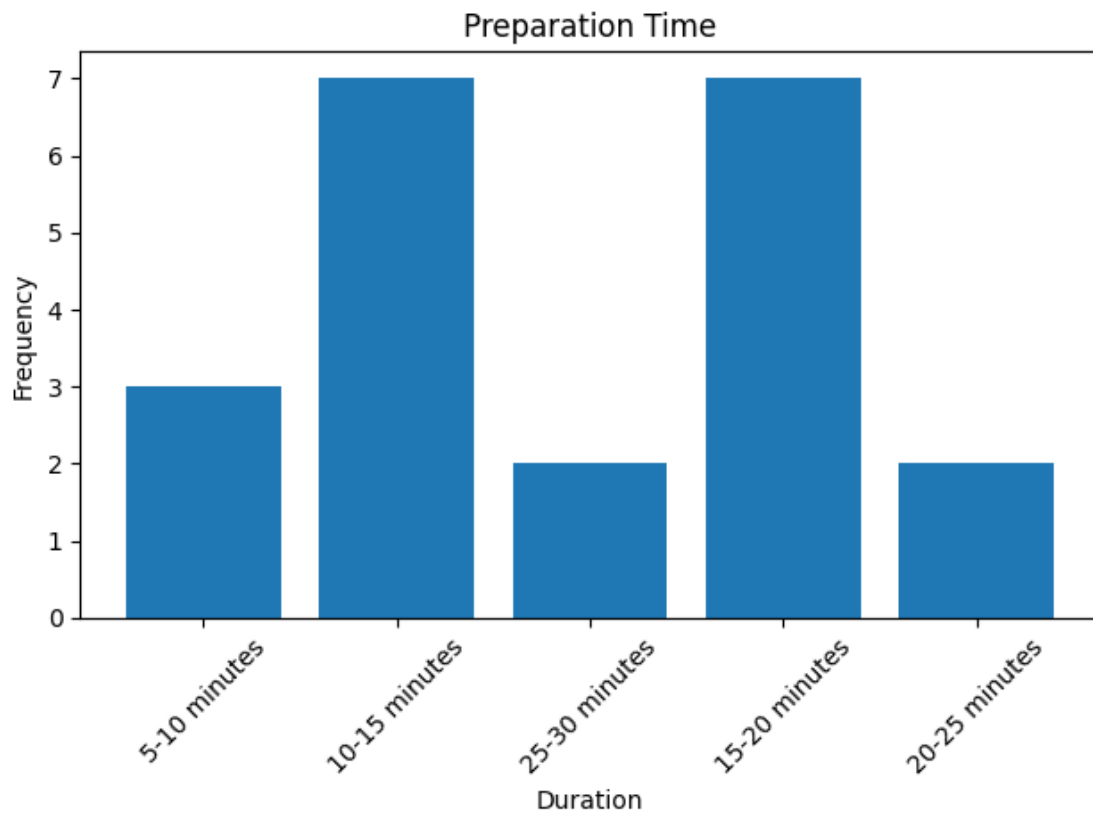
```
Counter({'more than 6': 13, '3 to 4': 4, '5 to 6': 3, '2 to 4': 1})
```



Process 'prep\_time'

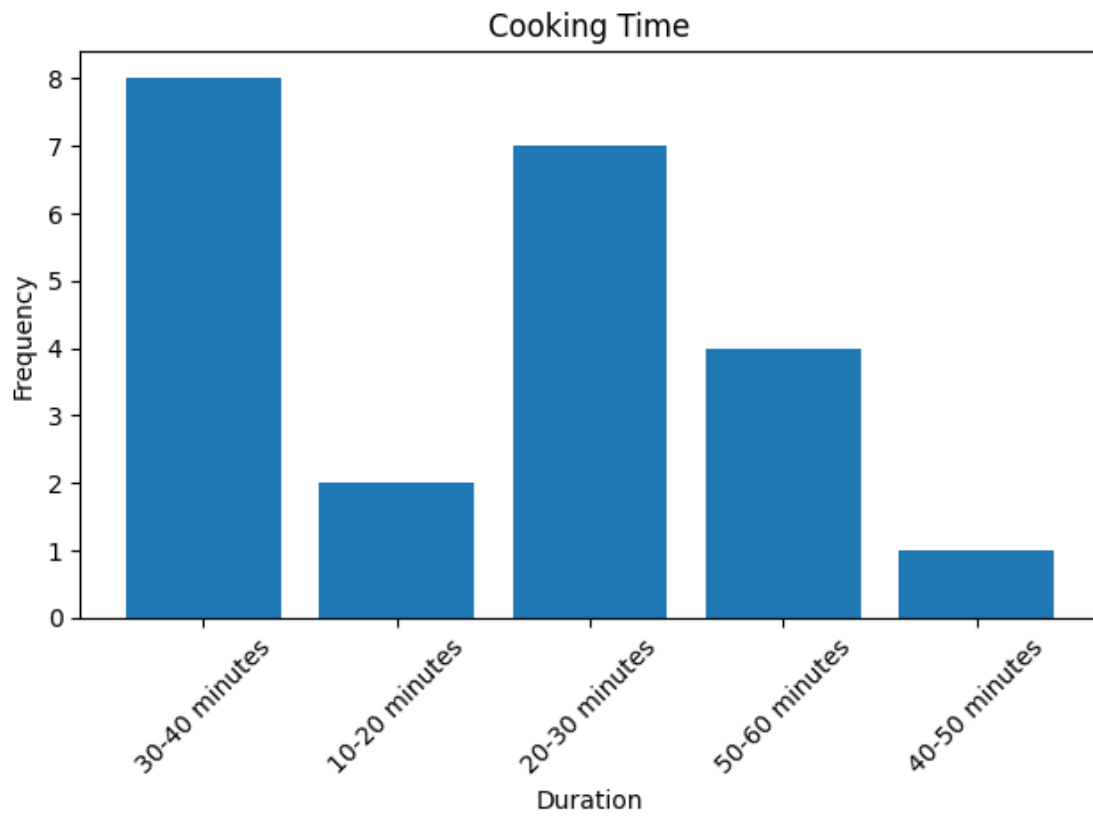
```
Counter({'10-15 minutes': 7, '15-20 minutes': 7, '5-10 minutes': 3, '25-30 minutes': 2, '20-25 minutes': 2})
```





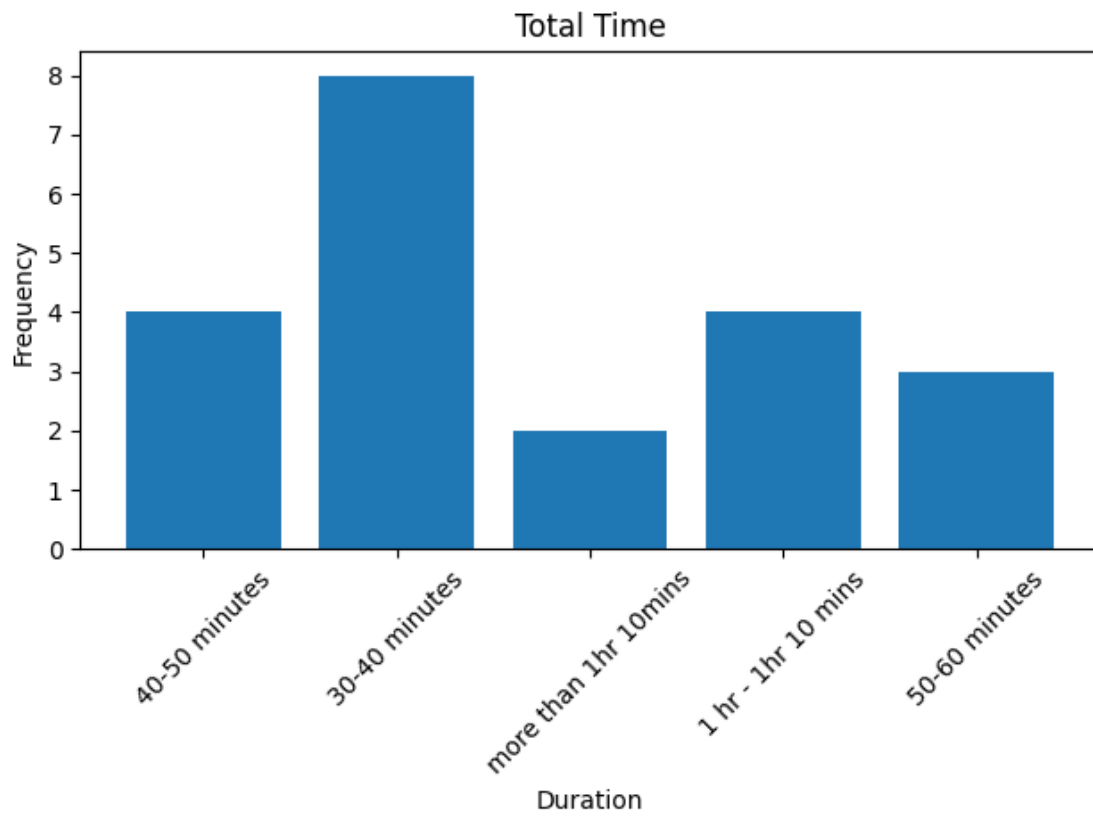
Process 'cook\_time'

```
Counter({'30-40 minutes': 8, '20-30 minutes': 7, '50-60 minutes': 4, '10-20 minutes': 2, '40-50 minutes': 1})
```



Process 'total\_time'

```
Counter({'30-40 minutes': 8, '40-50 minutes': 4, '1 hr - 1hr 10 mins': 4, '50-60 minutes': 3, 'more than 1hr 10mins': 2})
```



Process 'ingredients\_per\_dish'

Counter({'2 to 4': 8, '4 to 6': 8, '6 to 8': 4, '10 to 12': 1})

