

```

In [7]: import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from scipy import stats
import os

# Шаг 0: Загрузка данных
current_dir = os.getcwd()
file_path = os.path.join(current_dir, 'marketing_campaign.csv') # Убедит
df = pd.read_csv(file_path, sep='\t')

# Убедимся, что лишние пробелы в названиях колонок устранены
df.columns = df.columns.str.strip()

# Предварительная обработка данных (пример удаления лишних колонок и пров
# df.drop(['Z_CostContact', 'Z_Revenue'], axis=1, inplace=True) # Пример

# Шаг 1: Описательный анализ данных
print(df.head())
print(df.describe())

# Шаг 2: Построение графиков для анализа
columns_to_plot = ['MntWines', 'MntFruits', 'MntMeatProducts', 'MntFishPr
for column in columns_to_plot:
    plt.figure(figsize=(10, 6)) # Установить размер графика
    sns.boxplot(x=df[column])
    plt.title(f'Распределение покупок по {column}')
    plt.xlabel('Сумма')
    plt.show()

# Распределение ответов на маркетинговые кампании и по семейному положени
campaign_columns = ['AcceptedCmp1', 'AcceptedCmp2', 'AcceptedCmp3', 'Acce
for column in campaign_columns + ['Marital_Status']:
    plt.figure(figsize=(10, 6))
    sns.countplot(x=column, data=df)
    plt.title(f'Распределение по {column}')
    plt.ylabel('Количество')
    plt.show()

# Шаг 3: Проверка гипотез
# Проверка различий в доходах между семьями с детьми и без
t_stat, p_value = stats.ttest_ind(df[df['Kidhome'] > 0]['Income'], df[df[
print(f'T-статистика: {t_stat}, p-значение: {p_value}')

# Анализ влияния образования на принятие предложений первой кампании
for education in df['Education'].unique():
    plt.figure(figsize=(10, 6))
    sns.countplot(x='AcceptedCmp1', data=df[df['Education'] == education])
    plt.title(f'Ответы на 1-ю кампанию среди {education}')
    plt.xlabel('Принято предложение')
    plt.ylabel('Количество')
    plt.show()

# Шаг 4: Выводы из исследования
print("""

```

На основе проведенного анализа можно сделать следующие выводы:

- Средние траты на различные категории товаров варьируются. Например, покупатели с детьми и без отличаются по уровню дохода, что может влиять на различные сегменты по образованию и семейному положению показывают разный анализ ответов на маркетинговые кампании может помочь в планировании бюджета

Эти выводы могут помочь компании лучше понять своих клиентов и оптимизировать маркетинговую стратегию.)

	ID	Year_Birth	Education	Marital_Status	Income	Kidhome	Teenhome
0	5524	1957	Graduation	Single	58138.0	0	0
1	2174	1954	Graduation	Single	46344.0	1	1
2	4141	1965	Graduation	Together	71613.0	0	0
3	6182	1984	Graduation	Together	26646.0	1	0
4	5324	1981	PhD	Married	58293.0	1	0

	Dt_Customer	Recency	MntWines	...	NumWebVisitsMonth	AcceptedCmp3	\
0	04-09-2012	58	635	...	7	0	
1	08-03-2014	38	11	...	5	0	
2	21-08-2013	26	426	...	4	0	
3	10-02-2014	26	11	...	6	0	
4	19-01-2014	94	173	...	5	0	

	AcceptedCmp4	AcceptedCmp5	AcceptedCmp1	AcceptedCmp2	Complain	\
0	0	0	0	0	0	
1	0	0	0	0	0	
2	0	0	0	0	0	
3	0	0	0	0	0	
4	0	0	0	0	0	

	Z_CostContact	Z_Revenue	Response
0	3	11	1
1	3	11	0
2	3	11	0
3	3	11	0
4	3	11	0

[5 rows x 29 columns]

	ID	Year_Birth	Income	Kidhome	Teenhome
count	2240.000000	2240.000000	2216.000000	2240.000000	2240.000000
mean	5592.159821	1968.805804	52247.251354	0.444196	0.506250
std	3246.662198	11.984069	25173.076661	0.538398	0.544538
min	0.000000	1893.000000	1730.000000	0.000000	0.000000
25%	2828.250000	1959.000000	35303.000000	0.000000	0.000000
50%	5458.500000	1970.000000	51381.500000	0.000000	0.000000
75%	8427.750000	1977.000000	68522.000000	1.000000	1.000000
max	11191.000000	1996.000000	66666.000000	2.000000	2.000000

	Recency	MntWines	MntFruits	MntMeatProducts	\
count	2240.000000	2240.000000	2240.000000	2240.000000	
mean	49.109375	303.935714	26.302232	166.950000	
std	28.962453	336.597393	39.773434	225.715373	
min	0.000000	0.000000	0.000000	0.000000	
25%	24.000000	23.750000	1.000000	16.000000	
50%	49.000000	173.500000	8.000000	67.000000	

75%	74.000000	504.250000	33.000000	232.000000
max	99.000000	1493.000000	199.000000	1725.000000

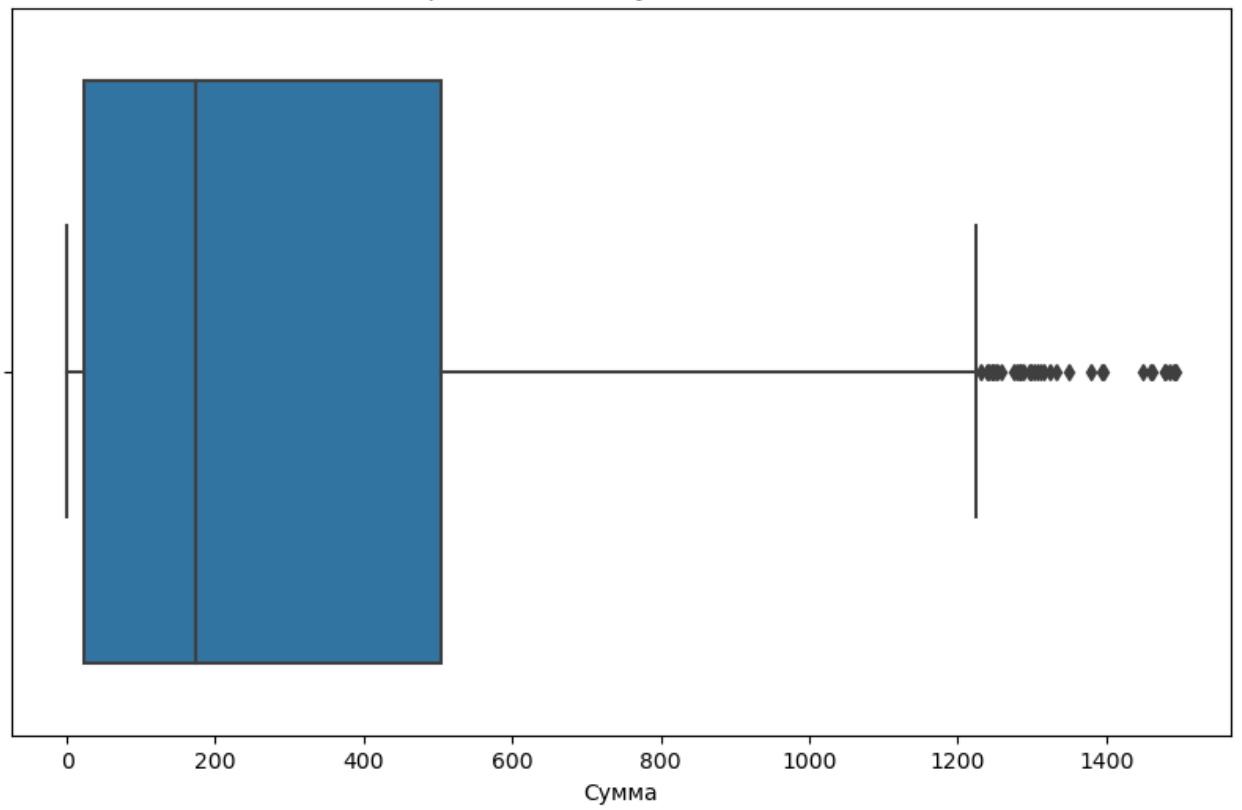
	MntFishProducts	...	NumWebVisitsMonth	AcceptedCmp3	AcceptedCmp4
\					
count	2240.000000	...	2240.000000	2240.000000	2240.000000
mean	37.525446	...	5.316518	0.072768	0.074554
std	54.628979	...	2.426645	0.259813	0.262728
min	0.000000	...	0.000000	0.000000	0.000000
25%	3.000000	...	3.000000	0.000000	0.000000
50%	12.000000	...	6.000000	0.000000	0.000000
75%	50.000000	...	7.000000	0.000000	0.000000
max	259.000000	...	20.000000	1.000000	1.000000

	AcceptedCmp5	AcceptedCmp1	AcceptedCmp2	Complain	Z_CostContac
t \					
count	2240.000000	2240.000000	2240.000000	2240.000000	2240.
0					
mean	0.072768	0.064286	0.013393	0.009375	3.
0					
std	0.259813	0.245316	0.114976	0.096391	0.
0					
min	0.000000	0.000000	0.000000	0.000000	3.
0					
25%	0.000000	0.000000	0.000000	0.000000	3.
0					
50%	0.000000	0.000000	0.000000	0.000000	3.
0					
75%	0.000000	0.000000	0.000000	0.000000	3.
0					
max	1.000000	1.000000	1.000000	1.000000	3.
0					

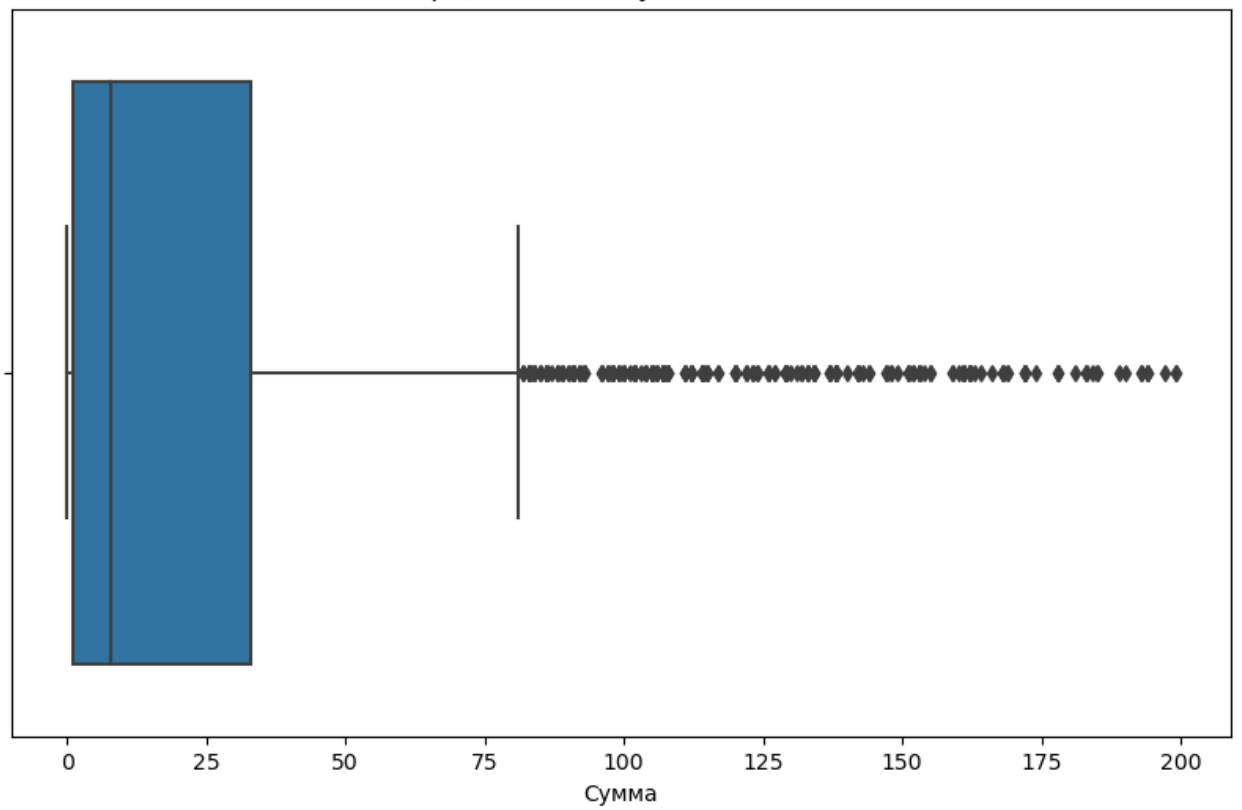
	Z_Revenue	Response
count	2240.0	2240.000000
mean	11.0	0.149107
std	0.0	0.356274
min	11.0	0.000000
25%	11.0	0.000000
50%	11.0	0.000000
75%	11.0	0.000000
max	11.0	1.000000

[8 rows x 26 columns]

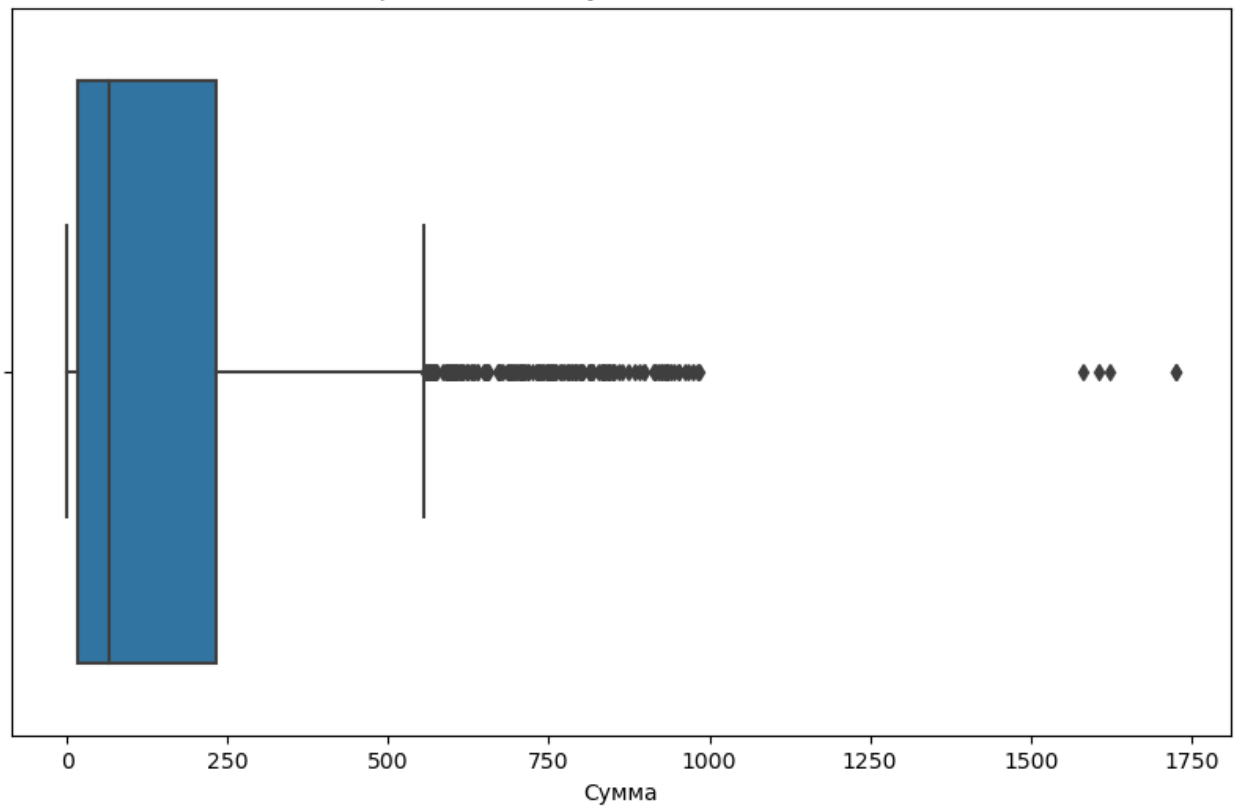
Распределение покупок по MntWines



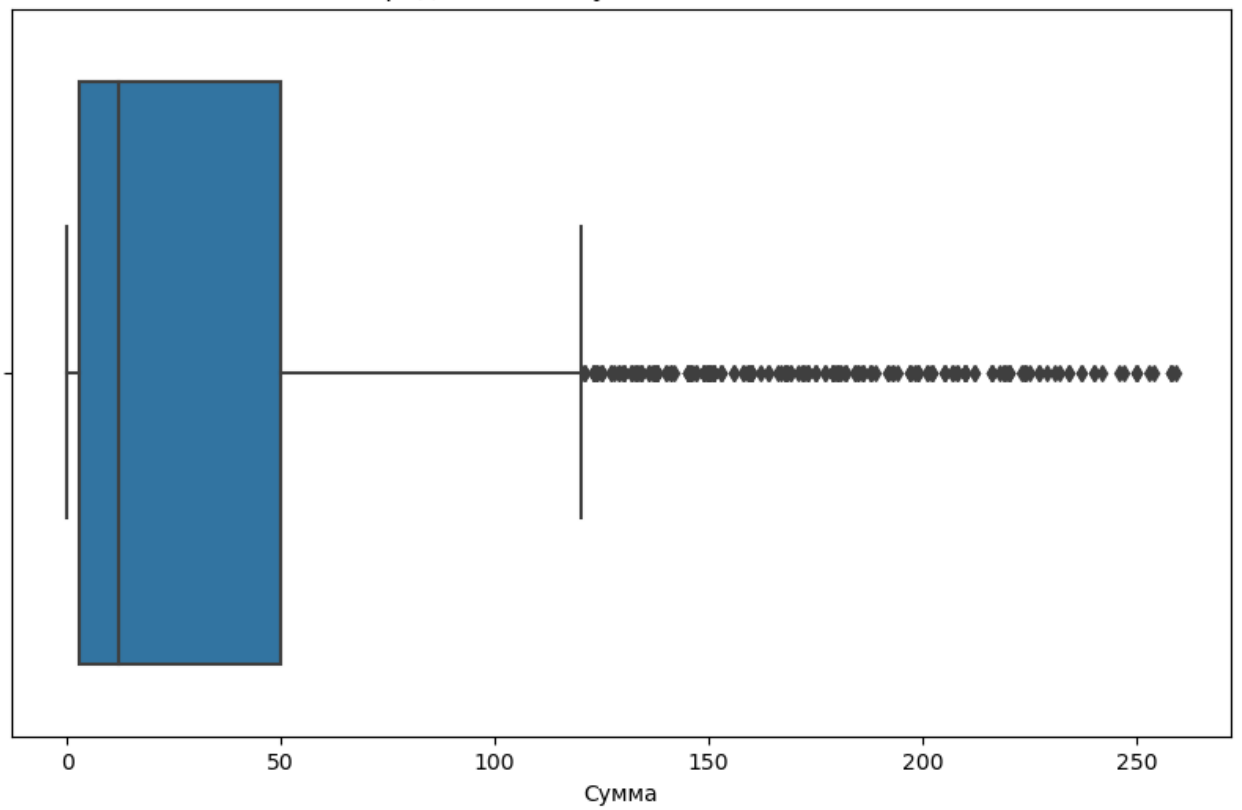
Распределение покупок по MntFruits



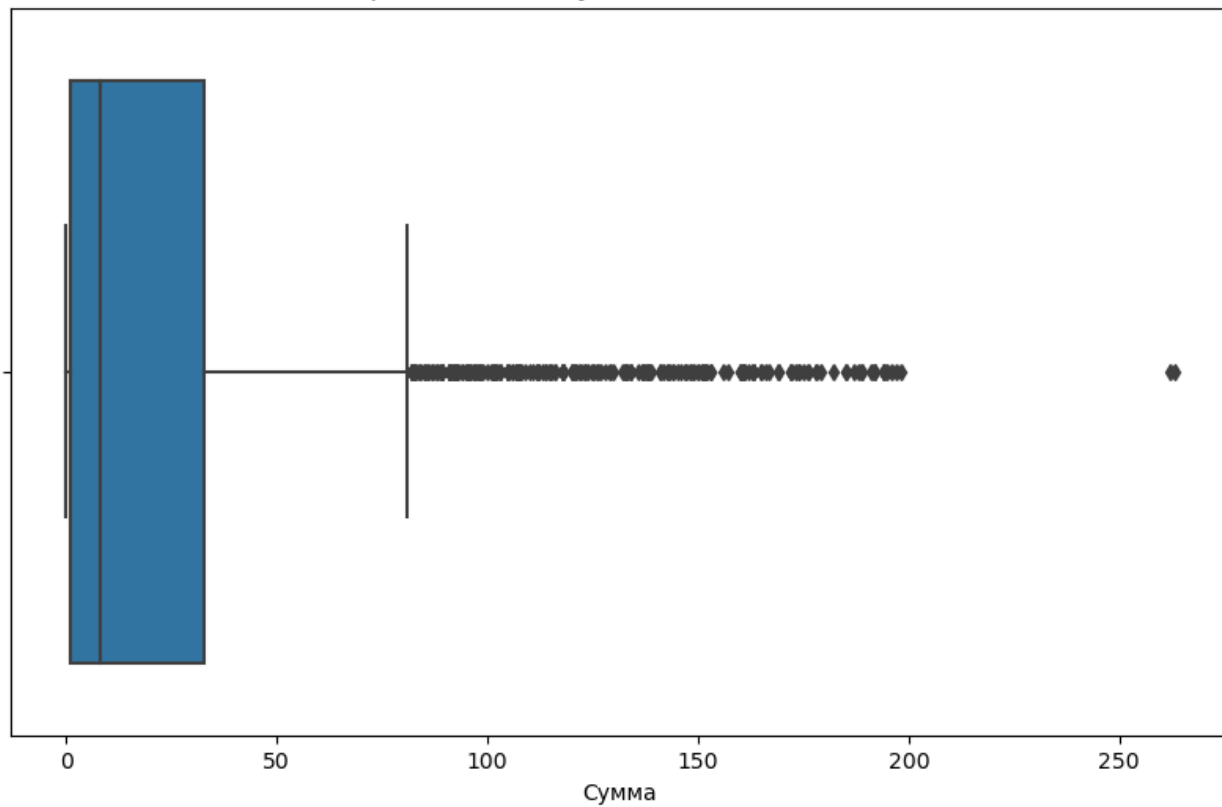
Распределение покупок по MntMeatProducts



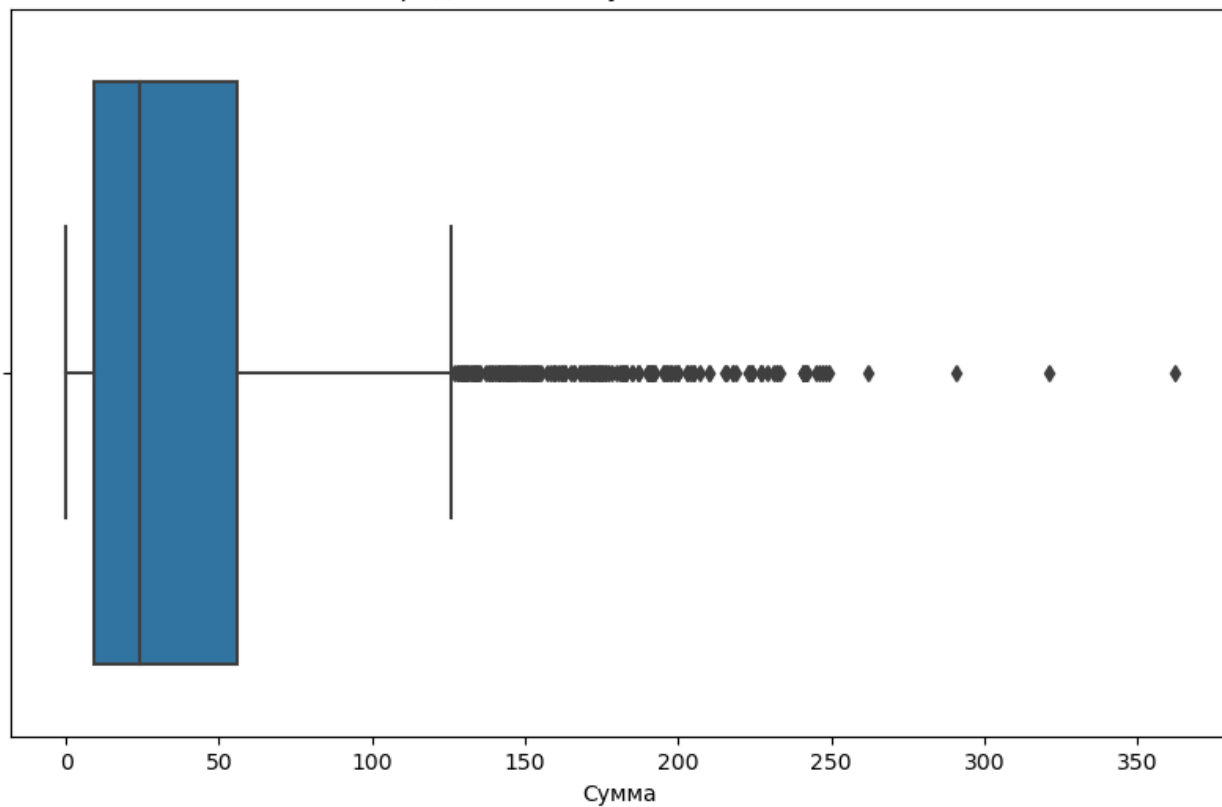
Распределение покупок по MntFishProducts



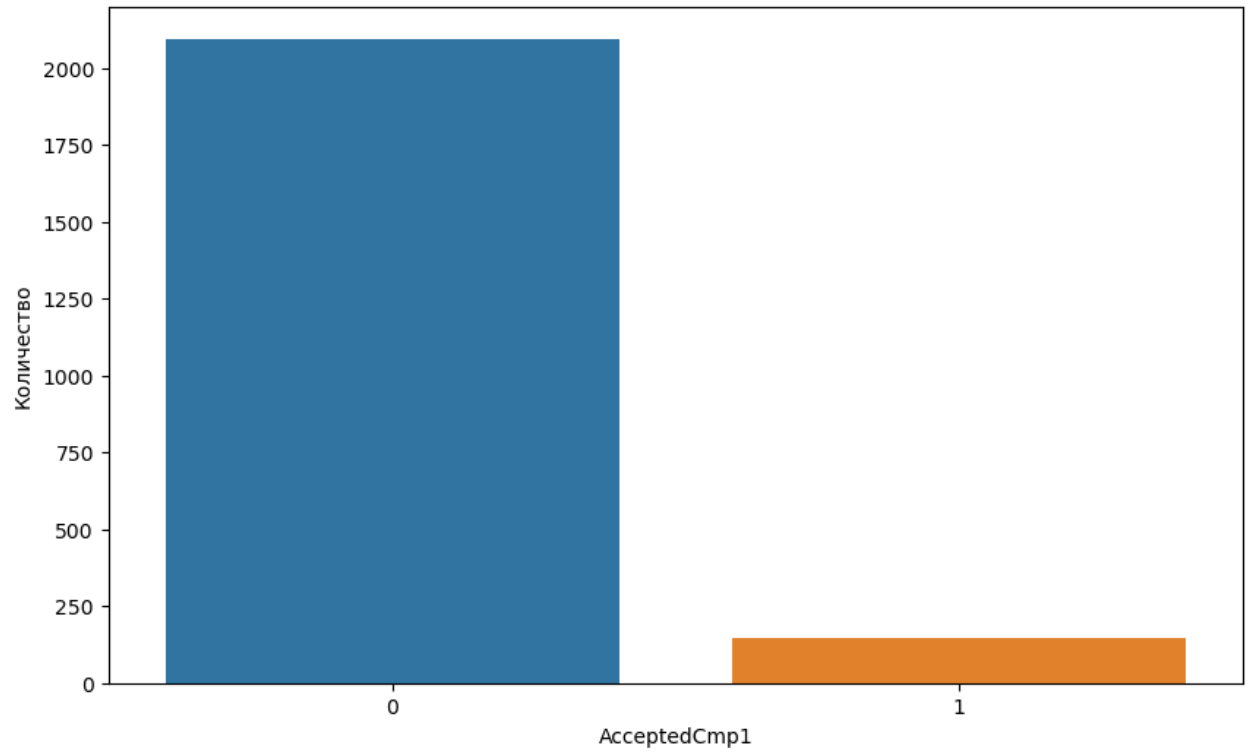
Распределение покупок по MntSweetProducts



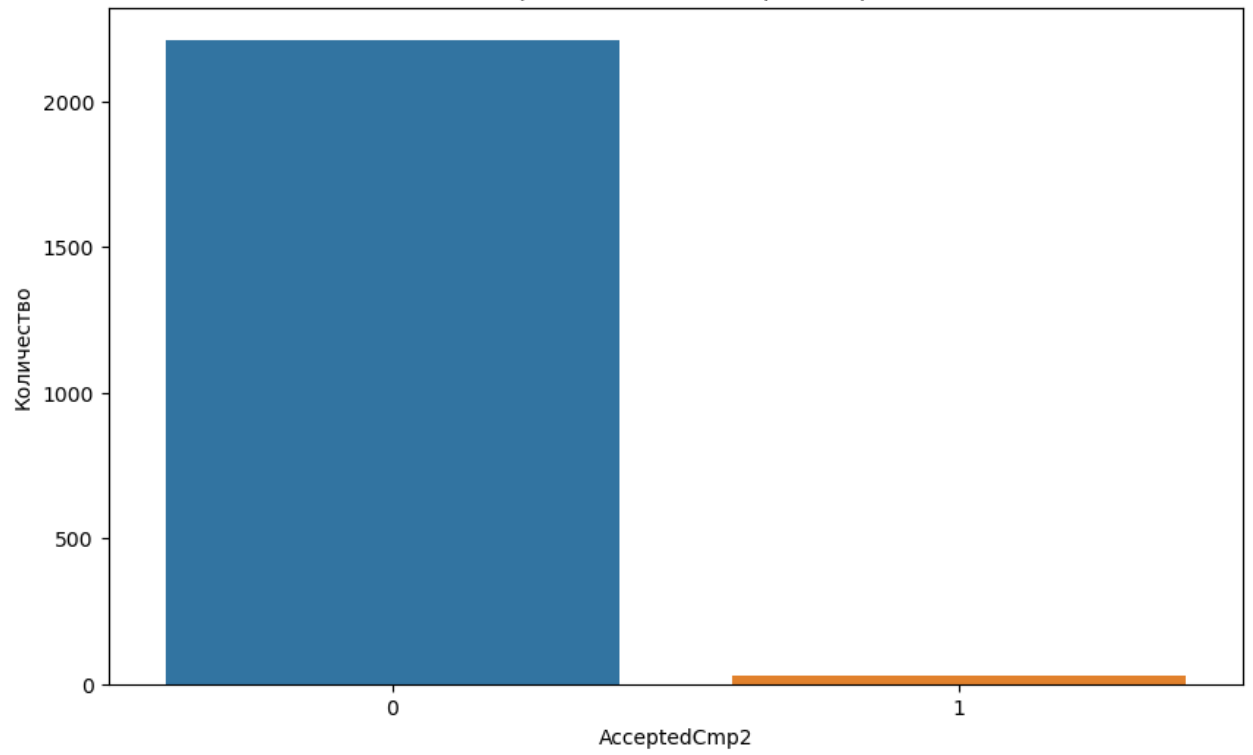
Распределение покупок по MntGoldProds



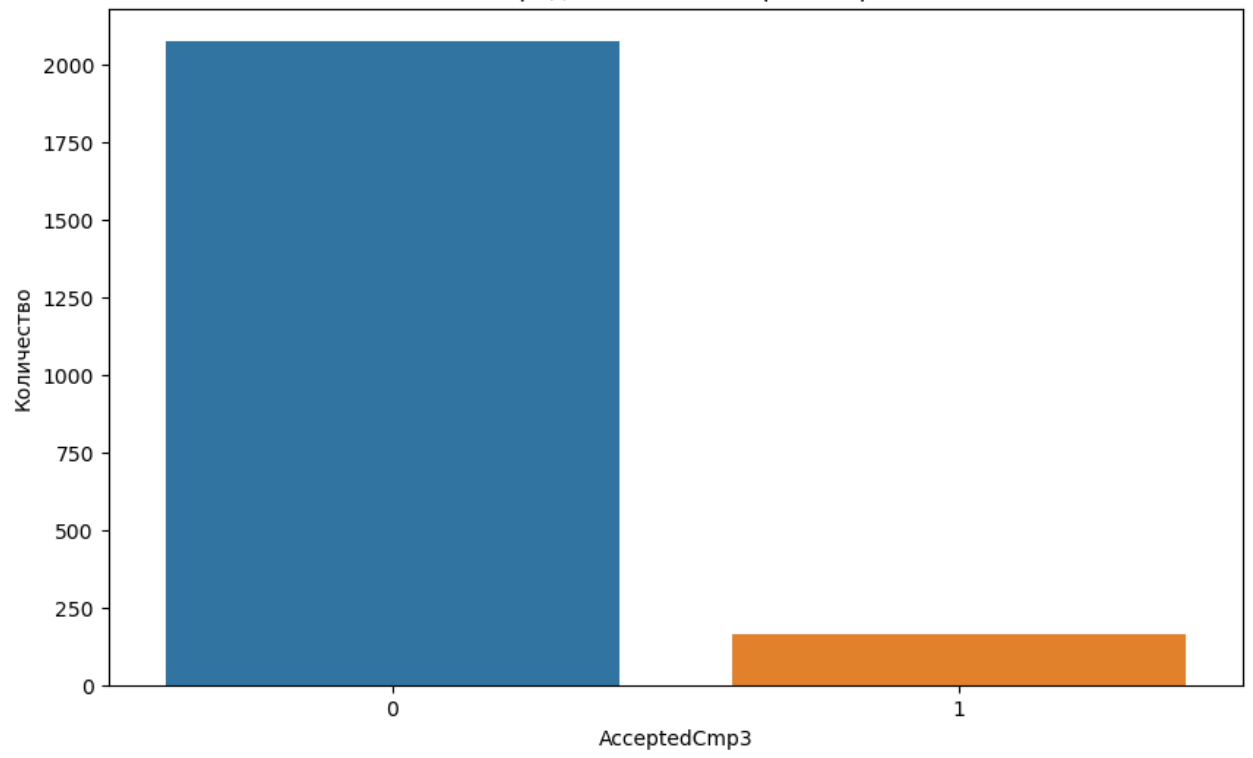
Распределение по AcceptedCmp1



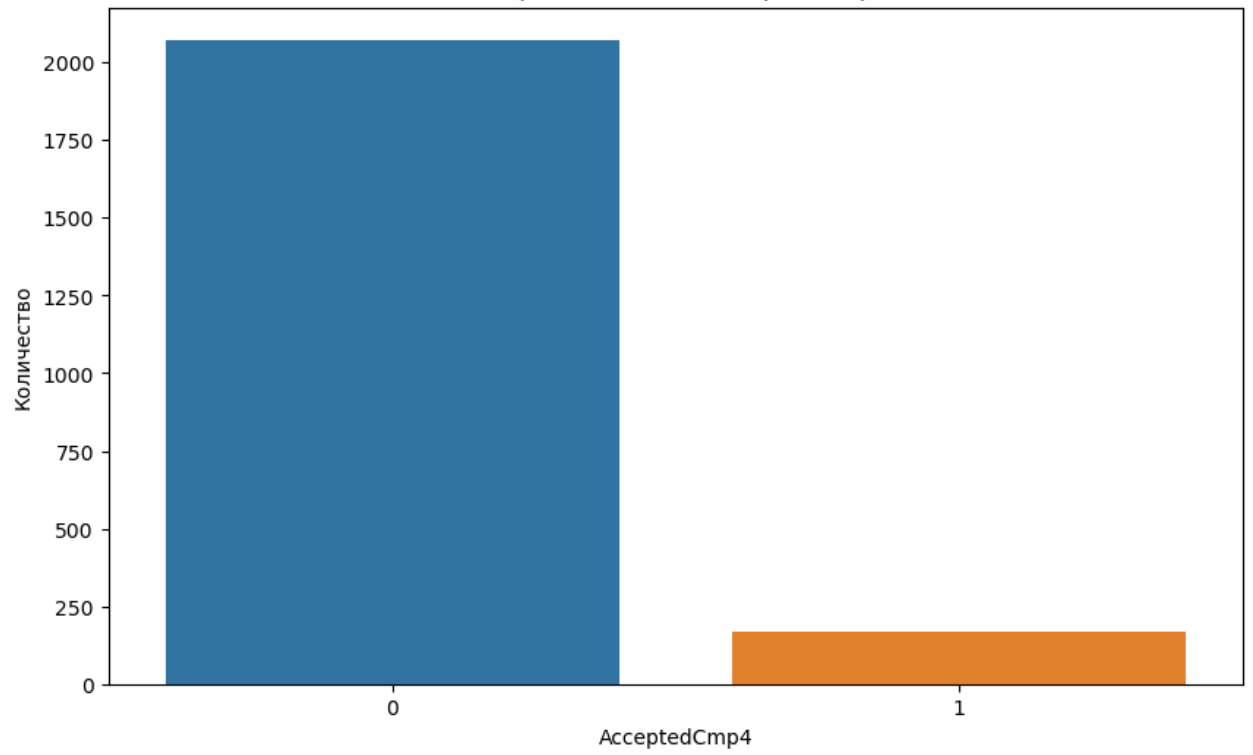
Распределение по AcceptedCmp2

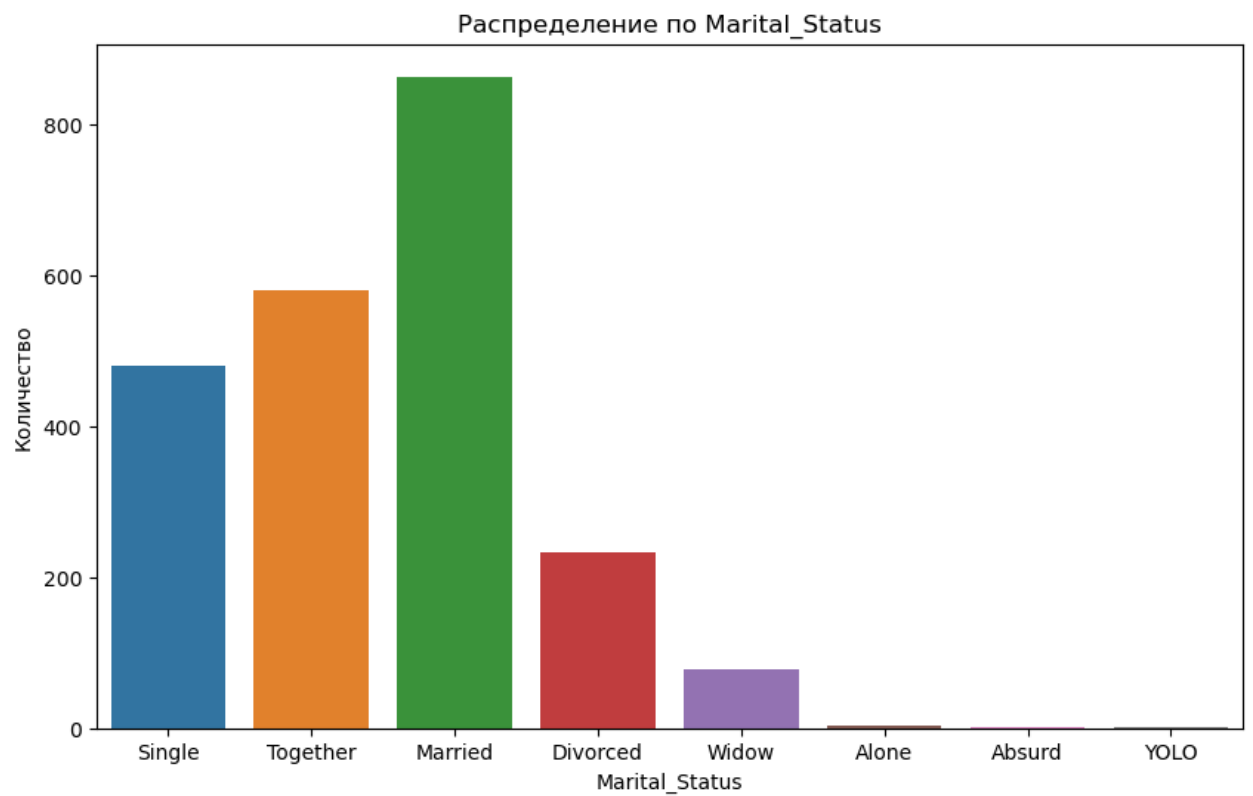
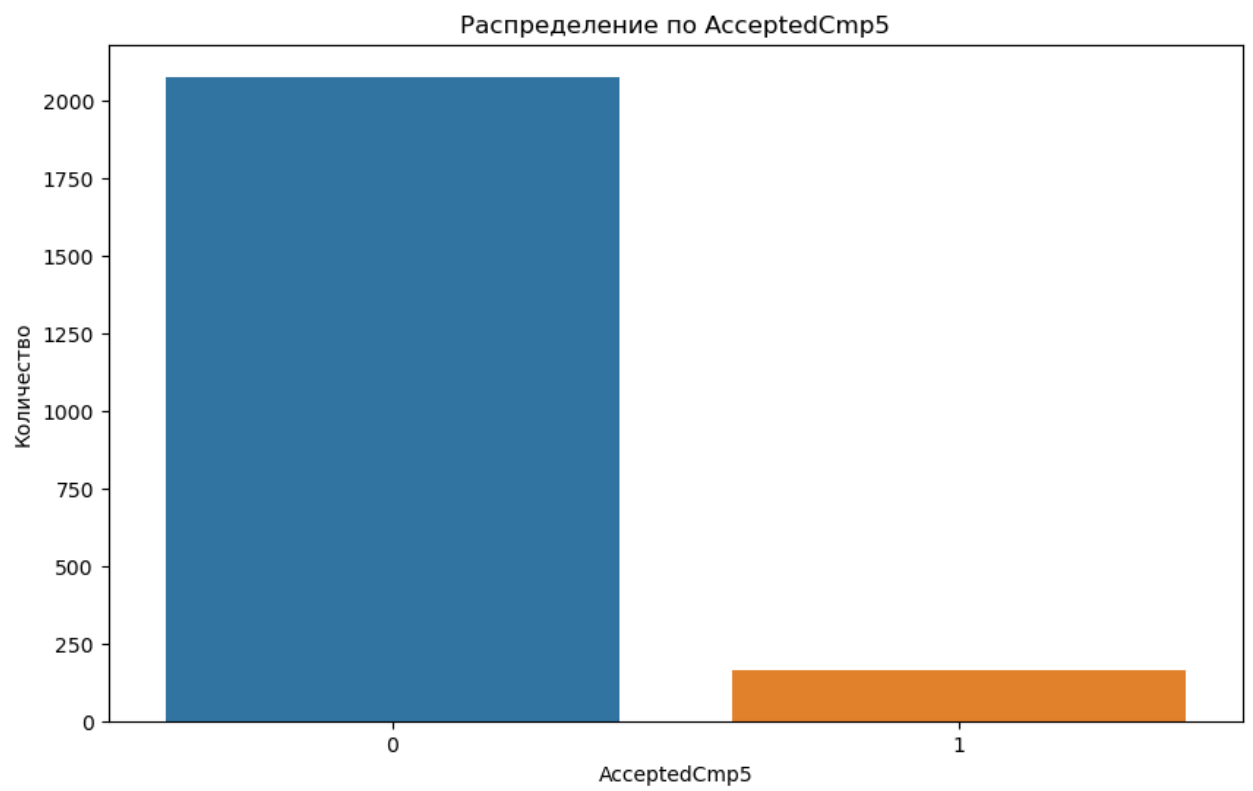


Распределение по AcceptedCmp3



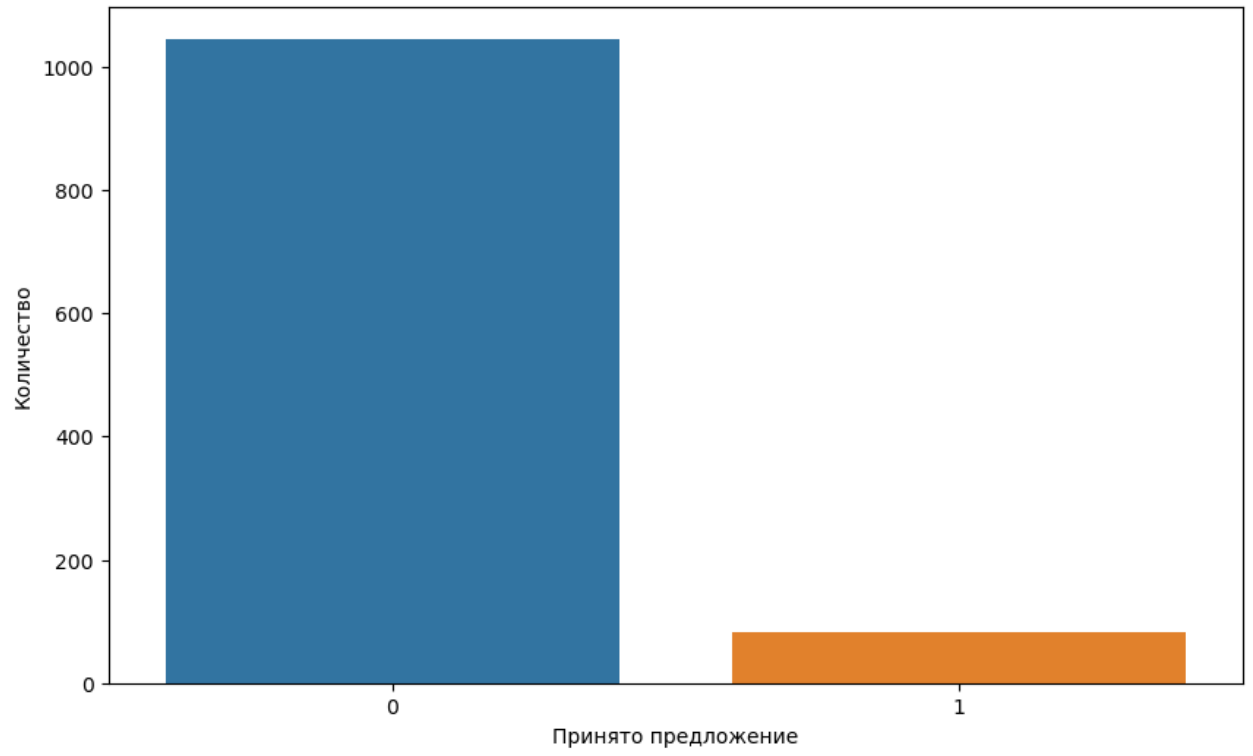
Распределение по AcceptedCmp4



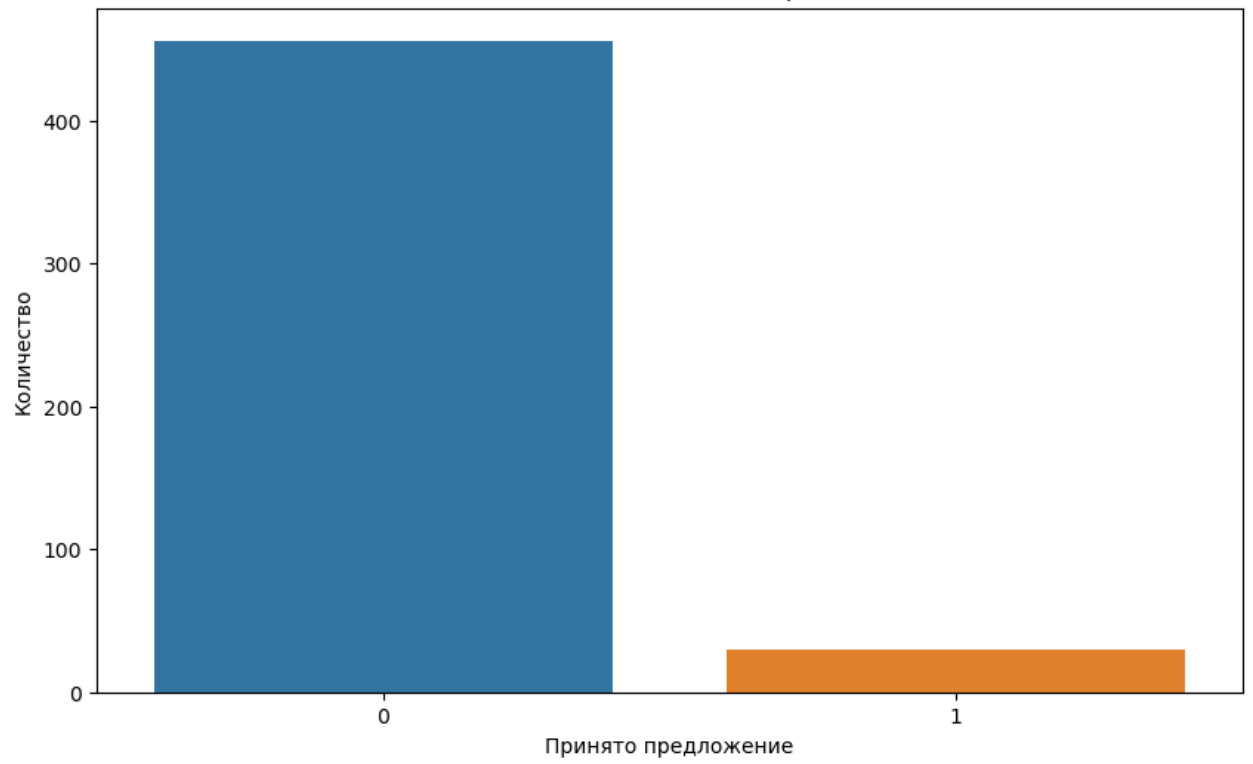


T-статистика: -23.32672344019383, p-значение: 8.531495230506484e-108

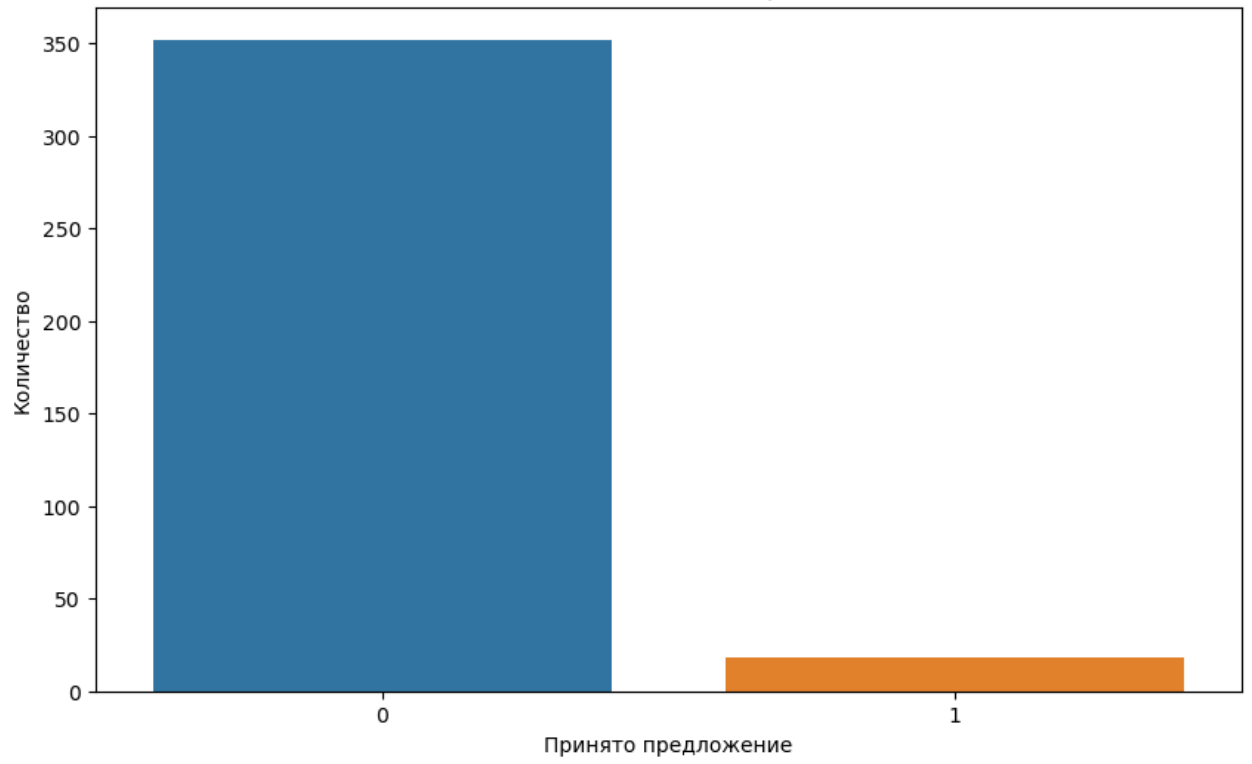
Ответы на 1-ю кампанию среди Graduation



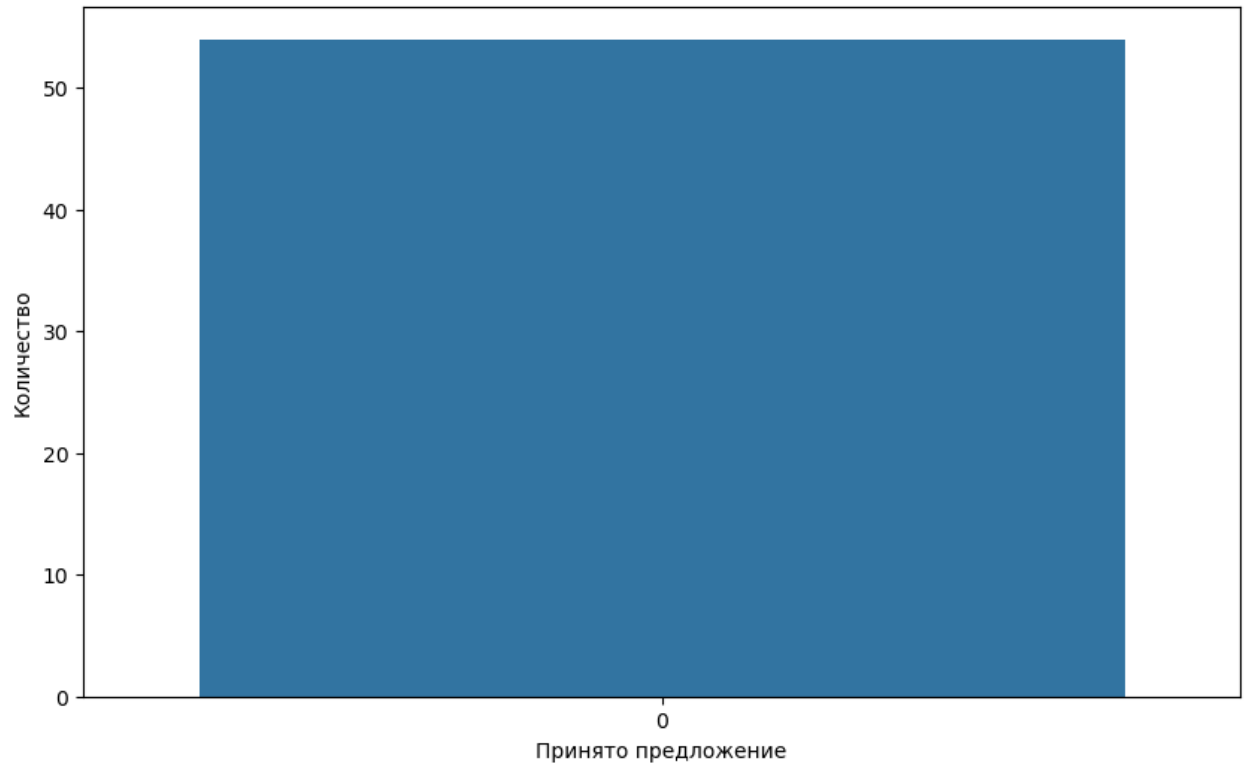
Ответы на 1-ю кампанию среди PhD

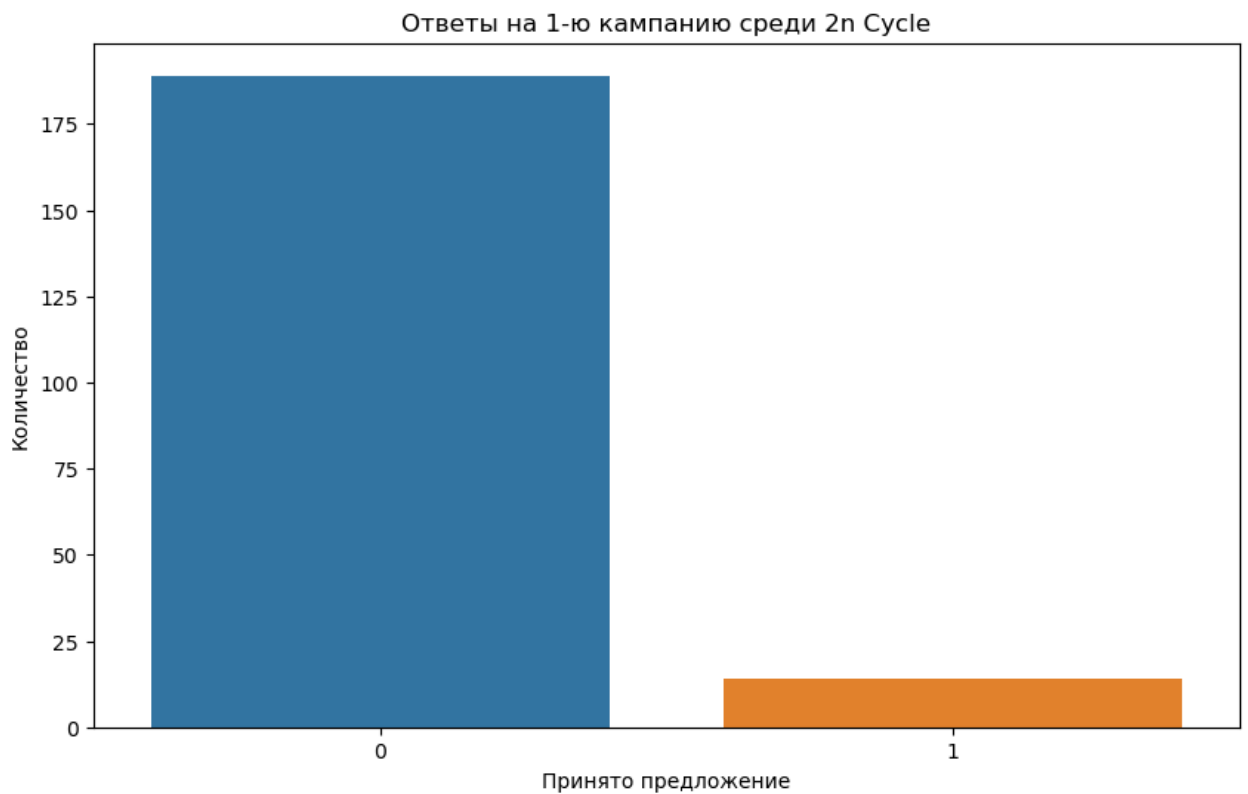


Ответы на 1-ю кампанию среди Master



Ответы на 1-ю кампанию среди Basic





На основе проведенного анализа можно сделать следующие выводы:

- Средние траты на различные категории товаров варьируются. Например, покупатели тратят больше на вино и мясо, чем на фрукты или сладости.
- Покупатели с детьми и без отличаются по уровню дохода, что может влиять на их потребительские предпочтения.
- Различные сегменты по образованию и семейному положению показывают разноеобразие в ответах на маркетинговые кампании.
- Анализ ответов на маркетинговые кампании может помочь в планировании будущих стратегий продвижения.

Эти выводы могут помочь компании лучше понять своих клиентов и оптимизировать маркетинговые кампании для различных сегментов.

In []: