Sparta Python Skills Sprint

Ausura Eccleston

My Process

Data cleaning and validation

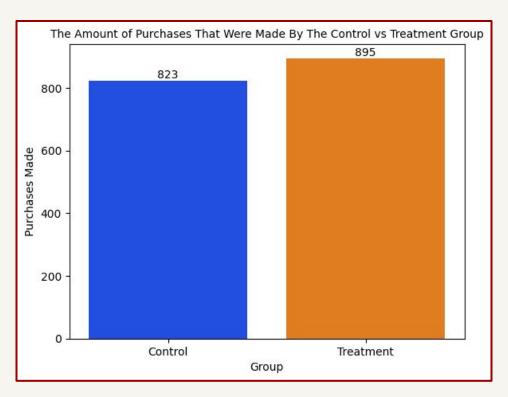
Exploratory analysis using pivot tables and summary functions

Create visualisations

Analysis & write up

Visualisations

The Group of Customers Exposed to Treatment Made More Purchases.

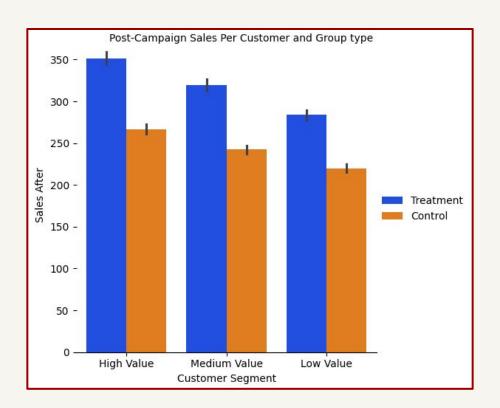


made by the control and treatment group #Filtering for only customers that made purchases df true = df d[df d['Purchase Made'] == 'Yes'l #grouping by the customer group and aggregating by the no. of purchases made q=sns.barplot(data=df true.groupby('Group', as index=False)['Purchase Made'].count(x='Group'. y='Purchase Made', palette='bright' #adding data labels to each bar for container in g.containers: # loop through each BarContainer g.bar label(container, fontsize=10)

#a vis to see how many purchases were

#styling elements g.set_ylabel("Purchases Made") #g.legend.set_title("") g.set_title("The Amount of Purchases That Were Made By The Control vs Treatment Group", fontsize=10)

Across all three different customer segments, customers exposed to the treatment/campaign spent more in sales.

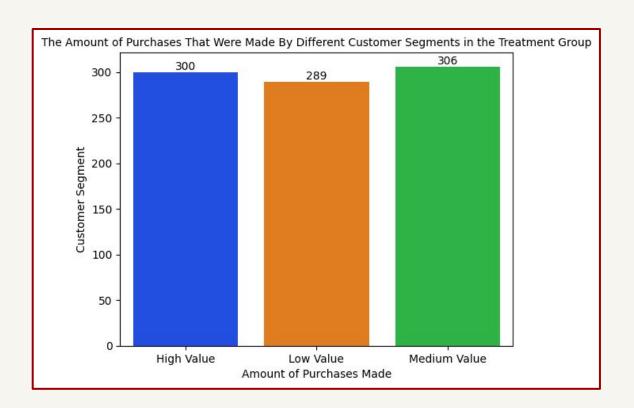


#a vis to see the affect of the campaign on sales in different customer groups

#creating a catplot g=sns.catplot(data=df_d, kind='bar', x='Customer_Segment', y='Sales_After', hue='Group', palette='bright'

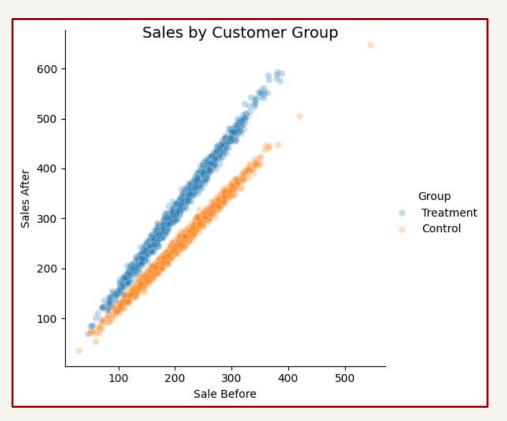
#styling elements g.despine(left=True) g.set_axis_labels("Customer Segment", "Sales After") g.legend.set_title("") g.fig.suptitle("Post-Campaign Sales Per Customer and Group type", fontsize=10)

Medium Value Customers Made The Most Purchases After Being Exposed To The Treatment. This was closely followed by high value customers.



```
#a vis to see how many purchases were made by
the different types of customers
#filtering for customers that mad a purchase and
were exposed to the treatment/campaign
df_true = df_d[(df_d['Purchase_Made'] == 'Yes') &
(df d['Group'] == 'Treatment')]
#Creating barplot, goruping by customer segment
and aggregating by the no. of purchases
g1=sns.barplot(
  data=df true.groupby('Customer Segment',
as index=False)['Purchase Made'].count(),
  x='Customer Segment',
  v='Purchase Made'.
  palette='bright',
  hue='Customer Segment'
#creating data labels for each bar
for container in g.containers: # loop through
each BarContainer
  q1.bar label(container, fontsize=10)
#styling elements
q1.set xlabel("Customer Segment")
g1.set ylabel("Amount of Purchases Made")
#a.leaend.set title("")
q1.set title("The Amount of Purchases That Were
Made By Different Customer Segments in the
Treatment Group", fontsize=10)
```

There is a strong positive correlation between the amount or sales made before the campaign and sales made after the campaign across both customer groups.



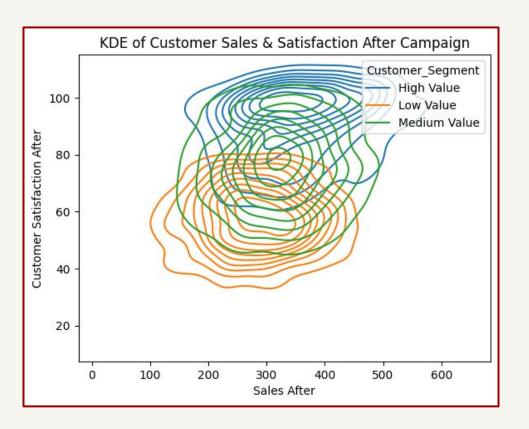
#a vis to see the correlation between sales made before the campaign and sales after the campaign by customer group

#creating scatter plot
sp = sns.relplot(data=df_d,
x='Sales_Before', y='Sales_After',
hue='Group', alpha=.25)

#create axis labes and title sp.set_axis_labels("Sale Before", "Sales After")

sp.fig.suptitle("Sales by Customer Group",
fontsize=14)
#sp.despine(trim=True)

Low value customers typically had lower sales and lower satisfaction compared to other customer segments. It was the opposite for high value customers.



```
#Filtering by customers who were
exposed to the treatment
df true = df d[(df d['Group'] ==
'Treatment')]
#creating the kde plot
kde graph = sns.kdeplot(
  data=df true,
  x='Sales After',
y='Customer Satisfaction After',
  hue='Customer Segment',
  thresh=.1.
#styling elements
kde graph.set xlabel("Sales After")
kde graph.set ylabel("Customer
Satisfaction After")
kde graph.set title("KDE of Customer
Sales & Satisfaction After Campaign")
```

Thank You