Data Solution 360

Battle of Insights

Transaction Data Analysis

Presentation By

Md Zillur Rahman

Objective and Dataset Overview

01

Objective:

This analysis aims to explore various aspects of transaction data to derive meaningful insights. The primary objectives are to:

1. Determine the average transaction amount across different store types and seasons.
2. Identify the most commonly used payment methods in high-value transactions and analyze how they vary across cities.
3. Compare sales amounts in transactions with and without discounts, and identify trends over the month.
4. Identify the top three cities with the highest average number of items per transaction and analyze how their sales amounts vary across seasons.
5. Evaluate the effectiveness of different promotions in driving higher transaction amounts, and determine which promotion type performs best in each season."

Objective and Dataset Overview Cont.

02

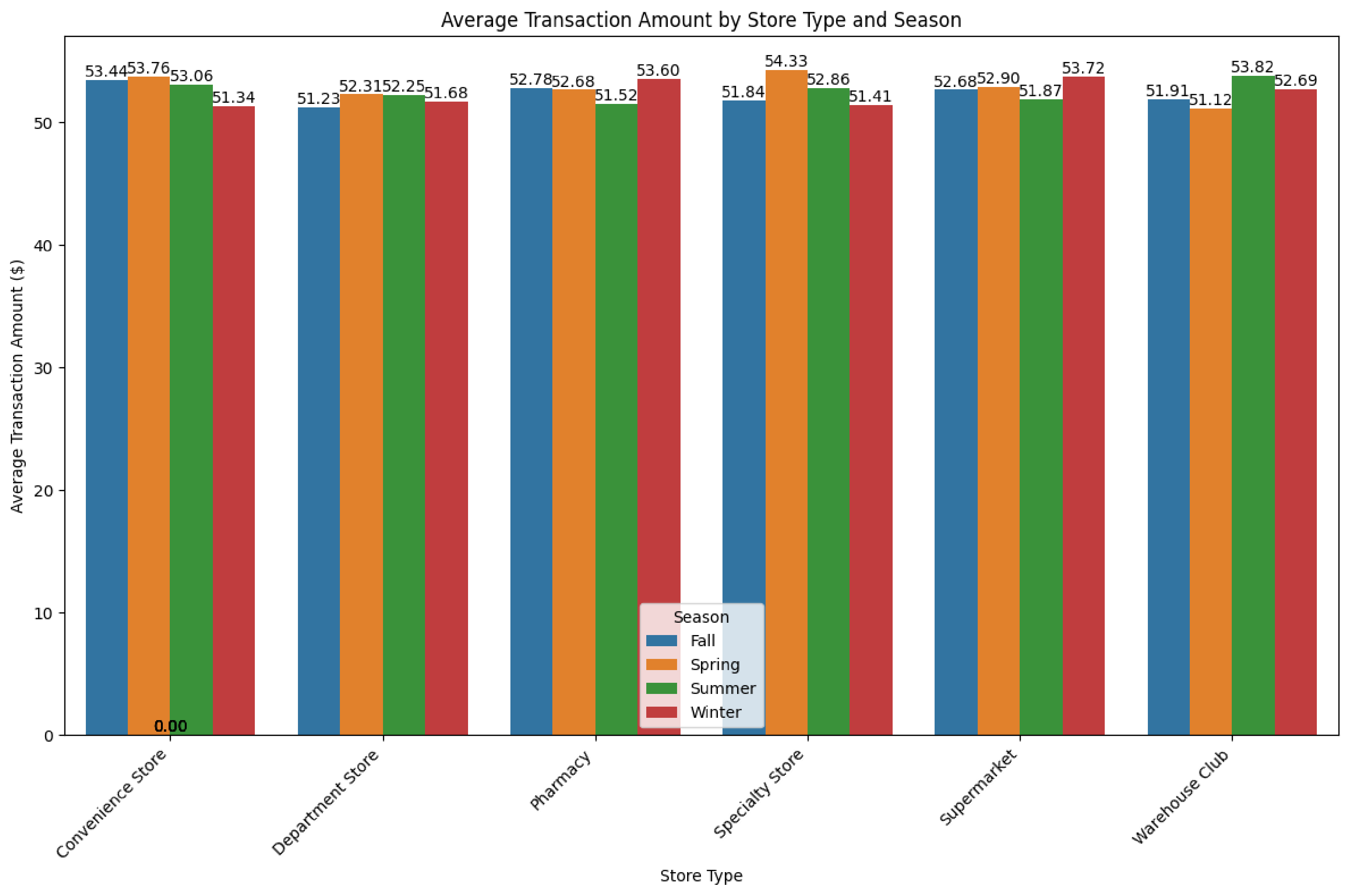
Dataset Overview:

•**Dataset Source:** Provided by Data Solution 360
•**Key Columns:**

- **Transaction_ID:** Unique identifier for each transaction
- **Date:** Date and time of the transaction
- **Customer_Name:** Name of the customer
- **Store_Type:** Type of store where the transaction occurred (e.g., Convenience Store, Supermarket, etc.)
- **Amount($):** Total cost of the transaction
- **Season:** Season in which the transaction occurred (e.g., Spring, Summer, Fall, Winter)
- **Payment_Method, City, Discount_Applied, Customer_Category, Promotion:** Other columns that provide insights into the nature of the transactions

Question – 01

What is the average transaction amount across different store types, and how does it vary by season?



Question – 01

What is the average transaction amount across different store types, and how does it vary by season?

Insights:

- **Summary:** The average transaction amount varies across both store types and seasons. For instance, in the Supermarket, the winter season has the highest average transaction amount ($53.72), while in the Warehouse Club, the summer season shows the highest ($53.81).

- **Seasonal Trends:** Winter tends to have slightly higher transaction amounts for stores like Supermarkets and Pharmacies, whereas Warehouse Clubs see higher spending in the summer.

- **Store Insights:** The Specialty Store shows a higher average transaction in the spring ($54.33), which could suggest an uptick in customer spending during this season.

Question - 01

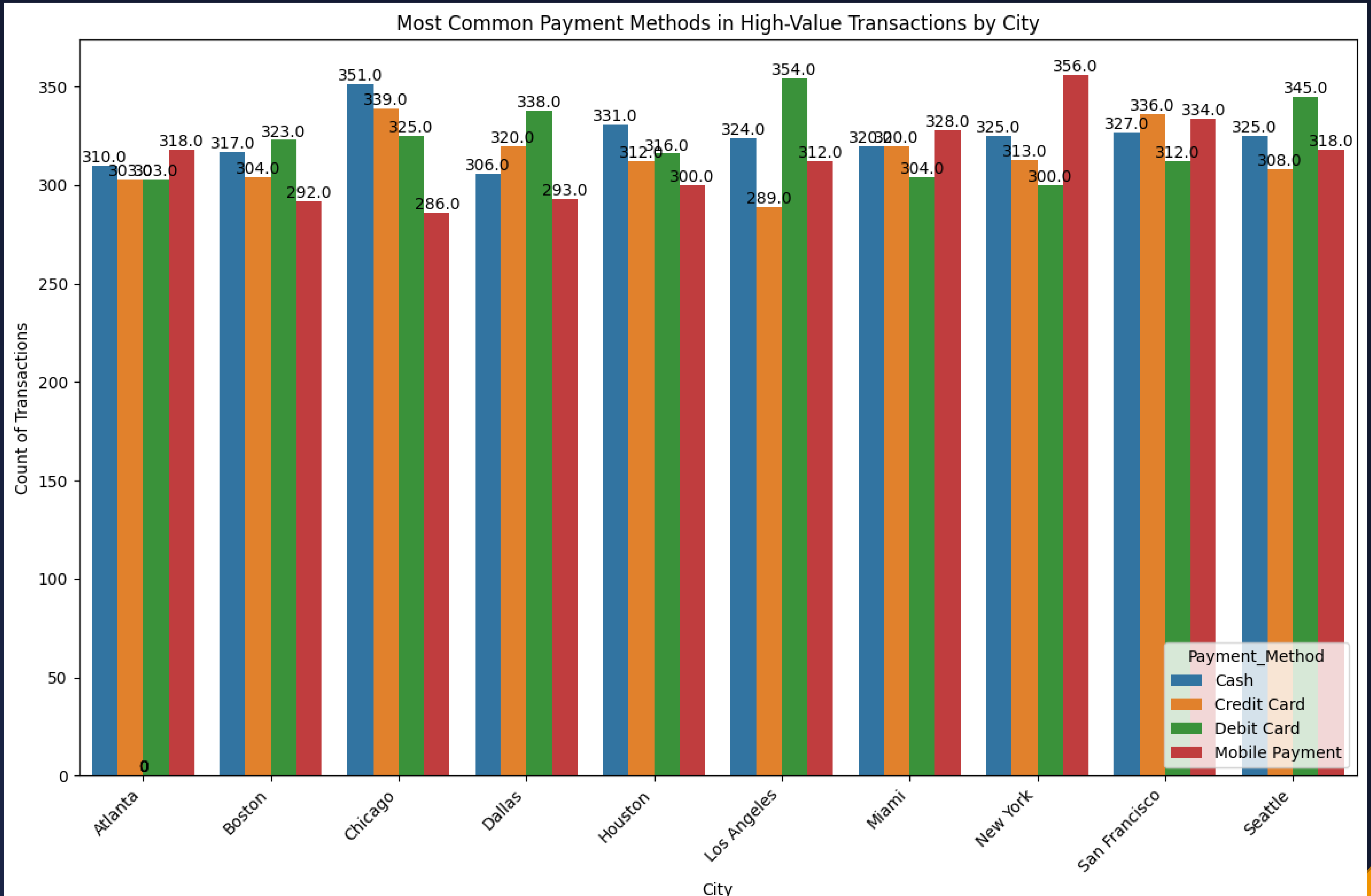
What is the average transaction amount across different store types, and how does it vary by season?

Python Code:

```python
# Visualizing the average transaction amount by Store Type and Season
average_transaction = df.groupby(['Store_Type', 'Season'])['Amount($)'].mean().reset_index()

# Create the plot
plt.figure(figsize=(12, 8))  # Increase figure size
ax = sns.barplot(x='Store_Type', y='Amount($)', hue='Season', data=average_transaction, dodge=True)

# Adding labels and title
plt.title('Average Transaction Amount by Store Type and Season')
plt.xlabel('Store Type')
plt.ylabel('Average Transaction Amount ($)')
plt.xticks(rotation=45, ha='right')  # Rotate x-axis labels for better readability

# Annotate each bar with the value
for p in ax.patches:
    ax.annotate(f'{p.get_height():.2f}',  # Display the height value (amount) rounded to 2 decimal places
                (p.get_x() + p.get_width() / 2., p.get_height()),  # Position of the label
                ha='center', va='center',  # Center the label
                fontsize=10, color='black',  # Font size and color
                xytext=(0, 5), textcoords='offset points')  # Adjust label position slightly above the bar

# Show the plot
plt.tight_layout()
plt.show()
```

Question - 02

Which payment method is most commonly used in high-value transactions (above the average transaction amount), and how does it differ across cities?



Question - 02

Which payment method is most commonly used in high-value transactions (above the average transaction amount), and how does it differ across cities?

Insights:

- **Key Findings:** In cities like New York, Miami, and Atlanta, Mobile Payment is the most commonly used method for high-value transactions, with New York leading the count with 356 transactions. However, cities like Chicago and Houston show a preference for Cash as the payment method.

- **City Comparison:** The payment method varies greatly across cities. While Debit Card is the most common in cities like Boston and Dallas, Mobile Payment is more prevalent in cities such as Atlanta and New York. Cash is also significant in Chicago and Houston.

- **Conclusion:** This indicates that payment preferences for high-value transactions are geographically diverse, potentially influenced by local payment infrastructure and customer preferences.

Question - 02

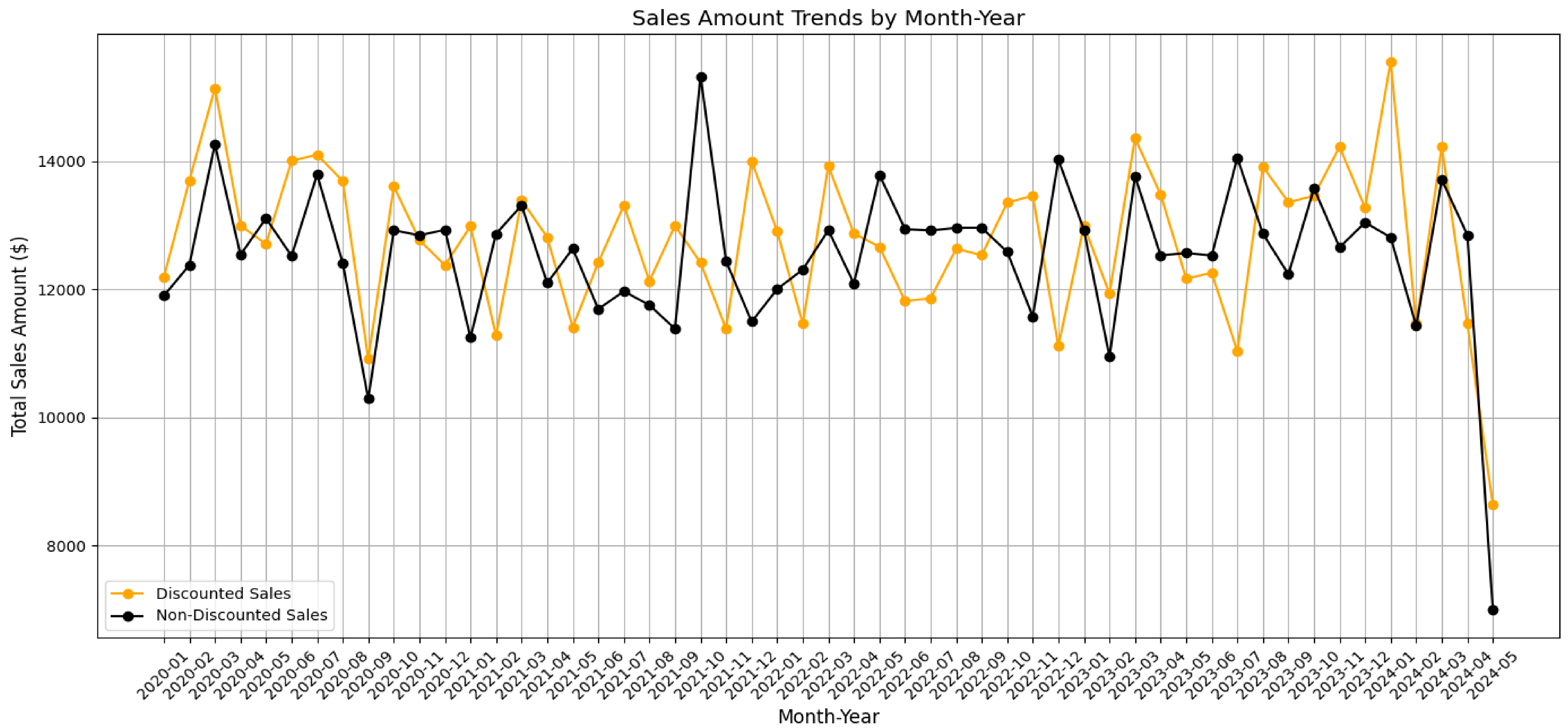
Which payment method is most commonly used in high-value transactions (above the average transaction amount), and how does it differ across cities?

Python Code:

```
# visualization of the most common payment methods across cities
plt.figure(figsize=(12, 8))
ax = sns.barplot(x='City', y='Count', hue='Payment_Method', data=payment_method_counts)

# Adding labels and title
plt.title('Most Common Payment Methods in High-Value Transactions by City')
plt.xlabel('City')
plt.ylabel('Count of Transactions')
plt.xticks(rotation=45, ha='right')

# Annotate each bar with the value
for p in ax.patches:
    ax.annotate(f'{p.get_height()}',  # Display the height value (count of transactions)
                (p.get_x() + p.get_width() / 2., p.get_height()),  # Position of the label
                ha='center', va='center',  # Center the label
                fontsize=10, color='black',  # Font size and color
                xytext=(0, 5), textcoords='offset points')  # Adjust label position slightly above the bar

# Show the plot
plt.tight_layout()
plt.show()
```

Question - 03

How do the sales amounts in transactions with discounts compare to those without discounts, and what trends can be observed over the month?



Question - 03

How do the sales amounts in transactions with discounts compare to those without discounts, and what trends can be observed over the month?

Insights:

1. Comparison of Discounted vs Non-Discounted Sales

- Across the dataset, discounted sales tend to contribute slightly more or are comparable to non-discounted sales. This suggests that discounts play a significant role in driving total sales.
- For most months, the difference between the two categories is not drastic, indicating a balanced strategy between regular and discounted pricing.

2. Monthly Trends

- Peaks in Sales: Discounted sales peaked in January 2024 ($15,548.30), suggesting effective seasonal promotions, possibly during New Year sales. Non-discounted sales peaked in December 2020 ($14,020.84), likely due to year- end or holiday shopping.
- Lows in Sales: Both discounted and non-discounted sales dropped significantly in May 2024 ($8,634.17 and $6,995.35, respectively). This could indicate a seasonal dip or less promotional activity.

Question - 03

How do the sales amounts in transactions with discounts compare to those without discounts, and what trends can be observed over the month?

Insights:

3. Seasonal Influence:

- Higher sales volumes are observed in November-December, aligning with global holiday and shopping seasons.
- A consistent increase in discounted sales during the first quarter (January-March) is noticeable, suggesting effective promotional campaigns after the holiday season.

4. Year-on-Year Insights:

- Sales for both categories have generally increased year-on-year, reflecting potential business growth or expansion.
- Some months (e.g., December 2022 vs December 2023) show stability or slight declines, hinting at saturation or changing customer behavior.

5. Effectiveness of Discounts:

- In some months (e.g., December 2020, October 2021), non-discounted sales exceed discounted sales, indicating strong product demand without reliance on discounts.
- In months with heavy discounts (e.g., January 2024), sales are significantly boosted, highlighting the importance of strategic discounting.

Question - 03

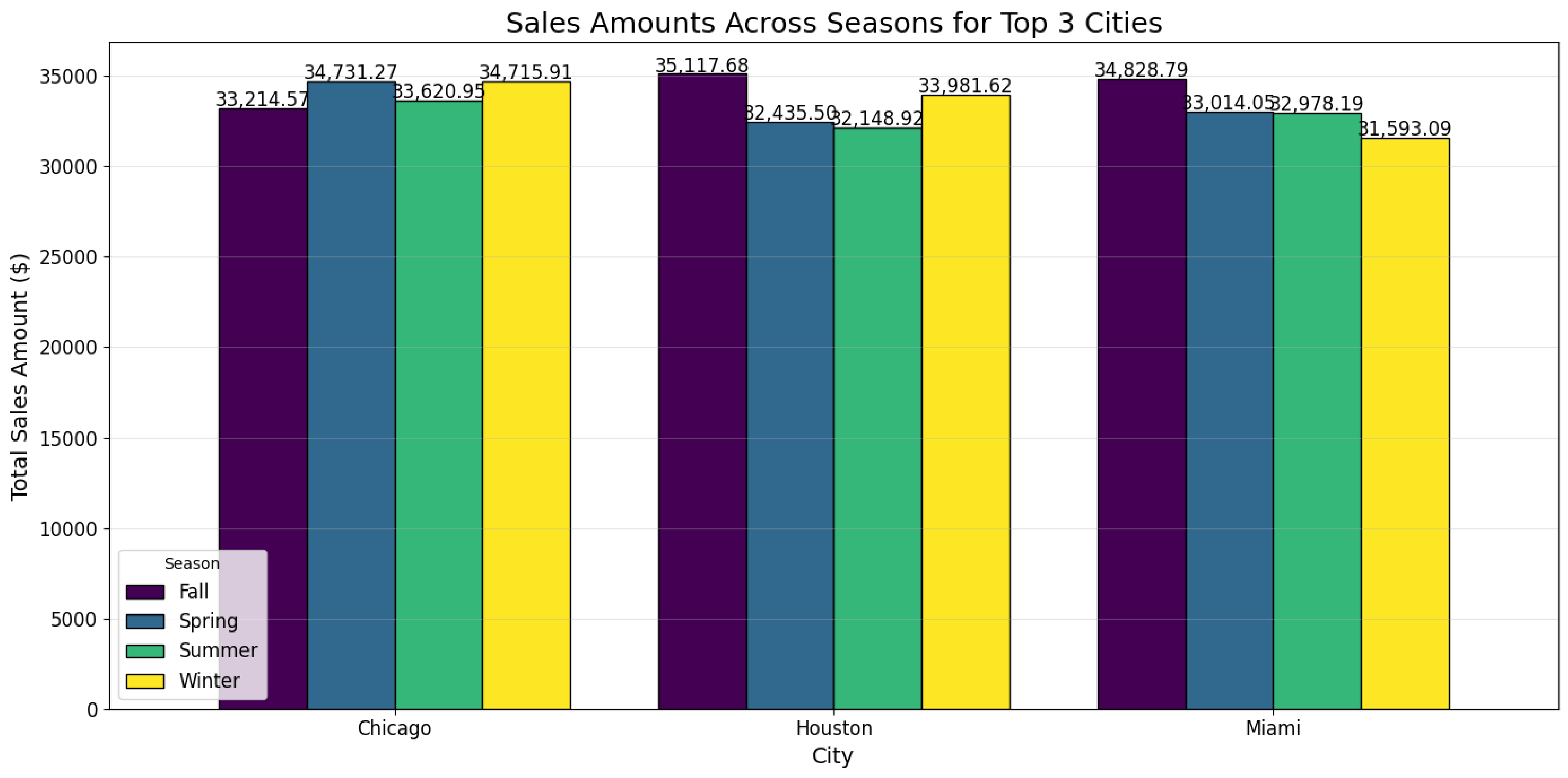
How do the sales amounts in transactions with discounts compare to those without discounts, and what trends can be observed over the month?

Python Code:

```
# Convert the Date column to datetime
df['Date'] = pd.to_datetime(df['Date'])

# Extract Month-Year for trend analysis
df['Month_Year'] = df['Date'].dt.to_period('M').astype(str)

# Filter data for transactions with and without discounts
discounted = df[df['Discount_Applied'] == True]
non_discounted = df[df['Discount_Applied'] == False]

# Calculate total sales amounts for each category by month-year
discounted_sales = discounted.groupby('Month_Year')['Amount($)'].sum()
non_discounted_sales = non_discounted.groupby('Month_Year')['Amount($)'].sum()

# Plotting the trends over the month-years
plt.figure(figsize=(14, 7))

plt.plot(discounted_sales.index, discounted_sales.values, label='Discounted Sales', marker='o', color='orange')
plt.plot(non_discounted_sales.index, non_discounted_sales.values, label='Non-Discounted Sales', marker='o', color='black')

plt.title('Sales Amount Trends by Month-Year', fontsize=14)
plt.xlabel('Month-Year', fontsize=12)
plt.ylabel('Total Sales Amount ($)', fontsize=12)
plt.legend()
plt.grid()
plt.xticks(rotation=45)
plt.tight_layout()

plt.show()
```

Question - 04

What are the top three cities with the highest average number of items per transaction, and how do their sales amounts vary across seasons?



Question - 04

What are the top three cities with the highest average number of items per transaction, and how do their sales amounts vary across seasons?

Insights:

Top 3 Cities with the Highest Average Number of Items per Transaction:

1. Houston: 5.55 items per transaction
2. Chicago: 5.54 items per transaction
3. Miami: 5.53 items per transaction

These cities indicate a higher tendency for customers to purchase multiple items in a single transaction, suggesting potentially more family-sized purchases or bulk-buying behavior.

Sales Amounts Across Seasons for the Top 3 Cities:

Chicago:

- Fall: $33,214.57
- Spring: $34,731.27
- Summer: $33,620.95
- Winter: $34,715.91
- Trend: Chicago shows fairly consistent sales throughout the year, with Winter slightly outperforming other seasons. Winter likely has higher demand, possibly due to holiday shopping.

Question - 04

What are the top three cities with the highest average number of items per transaction, and how do their sales amounts vary across seasons?

Insights:

Sales Amounts Across Seasons for the Top 3 Cities:

Houston:

- Fall: $35,117.68
- Spring: $32,435.50
- Summer: $32,148.92
- Winter: $33,981.62
- Trend: Houston's sales peak in Fall, but Spring sees a notable dip. The drop in Spring could indicate lower demand post-holiday season.

Miami:

- Fall: $34,828.79
- Spring: $33,014.05
- Summer: $32,978.19
- Winter: $31,593.09
- Trend: Fall sees the highest sales, followed by Spring and Summer. Winter has the lowest sales, potentially due to lower tourism or seasonal buying behavior.

Question - 04

What are the top three cities with the highest average number of items per transaction, and how do their sales amounts vary across seasons?

Insights:

Seasonal Sales Difference (Peak - Off-Peak) for Each City:

- Chicago: $1,516.70 (Peak: Winter, Off-Peak: Summer)
- Houston: $2,968.76 (Peak: Fall, Off-Peak: Spring)
- Miami: $3,235.70 (Peak: Fall, Off-Peak: Winter)
- Trend: Miami shows the largest seasonal sales difference, indicating a strong peak during Fall and a significant dip in Winter. Houston also sees a drop in Spring sales compared to Fall. Chicago’s difference is smaller, suggesting less fluctuation in demand across seasons.

Conclusion:

- Chicago shows stability in sales, with Winter slightly outperforming other seasons.
- Houston peaks in Fall, but experiences a significant dip in Spring.
- Miami shows the strongest Fall sales but a notable decline in Winter, likely due to seasonal purchasing behavior.

Question - 04

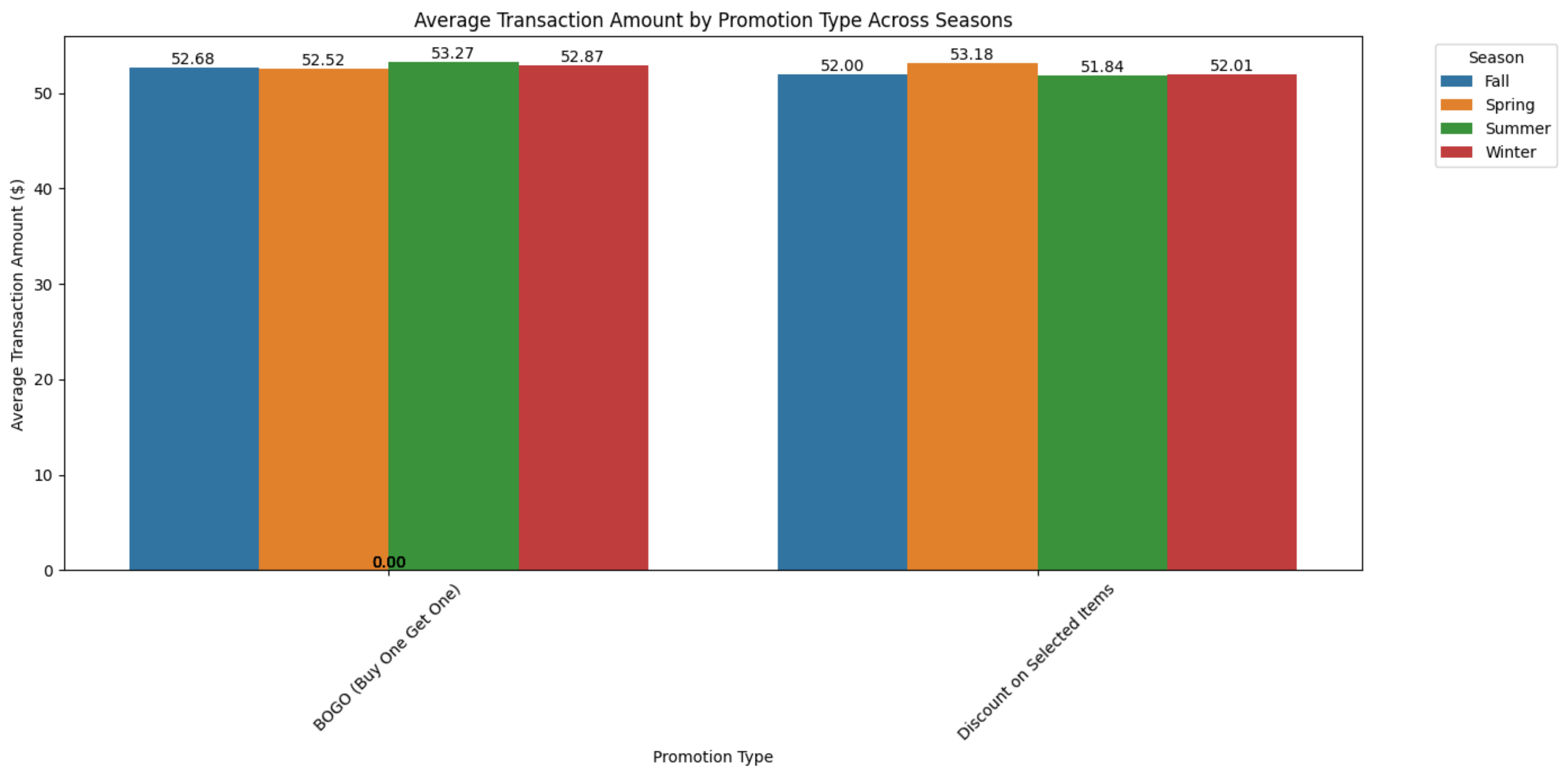
What are the top three cities with the highest average number of items per transaction, and how do their sales amounts vary across seasons?

Python Code:

```
# Calculate average items per transaction for each city

# Ensure 'Date' is in datetime format
df['Date'] = pd.to_datetime(df['Date'])

# Calculate the average number of items per transaction for each city
city_avg_items = df.groupby('City')['Total_Items'].mean().sort_values(ascending=False)

# Get the top three cities
top_cities = city_avg_items.head(3).index
print("Top 3 Cities with Highest Average Items per Transaction:")
print(city_avg_items.head(3))

# Filter the data for the top three cities
top_cities_data = df[df['City'].isin(top_cities)]

# Group the data by City and Season to calculate total sales amounts
city_season_sales = top_cities_data.groupby(['City', 'Season'])['Amount($)'].sum().unstack()

# Create a bar chart with more space and larger figure size
fig, ax = plt.subplots(figsize=(14, 7))  # Increase figure size
city_season_sales.plot(kind='bar', ax=ax, colormap='viridis', edgecolor='black', width=0.8)

# Add data labels with adjusted position
for p in ax.patches:
    height = p.get_height()
    ax.annotate(f'{height:,.2f}',
                (p.get_x() + p.get_width() / 2., height),
                ha='center', va='center',
                fontsize=12, color='black',
                xytext=(0, 5), textcoords='offset points')

# Customize the chart for better readability
plt.title('Sales Amounts Across Seasons for Top 3 Cities', fontsize=18)
plt.xlabel('City', fontsize=14)
plt.ylabel('Total Sales Amount ($)', fontsize=14)
plt.legend(title='Season', fontsize=12)
plt.xticks(rotation=0, fontsize=12)
plt.yticks(fontsize=12)
plt.grid(axis='y', alpha=0.3)
plt.tight_layout()

# Show the chart
plt.show()
```

Question - 05

How effective are different promotions in driving higher transaction amounts, and which promotion type performs best in each season?



Question - 05

How effective are different promotions in driving higher transaction amounts, and which promotion type performs best in each season?

	Promotion	Season	mean	sum	count
0	BOGO (Buy One Get One)	Fall	52.678438	167306.72	3176
1	BOGO (Buy One Get One)	Spring	52.519019	167483.15	3189
2	BOGO (Buy One Get One)	Summer	53.271761	170949.08	3209
3	BOGO (Buy One Get One)	Winter	52.872204	166230.21	3144
4	Discount on Selected Items	Fall	51.997053	169042.42	3251
5	Discount on Selected Items	Spring	53.177148	169475.57	3187
6	Discount on Selected Items	Summer	51.842832	166933.92	3220
7	Discount on Selected Items	Winter	52.012388	163995.06	3153

Question - 05

How effective are different promotions in driving higher transaction amounts, and which promotion type performs best in each season?

	Promotion	Season	mean	sum	count
0	BOGO (Buy One Get One)	Fall	52.678438	167306.72	3176
1	BOGO (Buy One Get One)	Spring	52.519019	167483.15	3189
2	BOGO (Buy One Get One)	Summer	53.271761	170949.08	3209
3	BOGO (Buy One Get One)	Winter	52.872204	166230.21	3144
4	Discount on Selected Items	Fall	51.997053	169042.42	3251
5	Discount on Selected Items	Spring	53.177148	169475.57	3187
6	Discount on Selected Items	Summer	51.842832	166933.92	3220
7	Discount on Selected Items	Winter	52.012388	163995.06	3153

Question - 05

How effective are different promotions in driving higher transaction amounts, and which promotion type performs best in each season?

Insights:

1. Effectiveness of Promotions in Driving Higher Transaction Amounts:

BOGO (Buy One Get One) promotion has slightly higher average transaction amounts across all seasons compared to Discount on Selected Items.

2. Best Performing Promotion in Each Season:

- Fall: BOGO (Buy One Get One) promotion leads with an average transaction amount of $52.68, while Discount on Selected Items trails slightly at $51.99.
- Spring: Discount on Selected Items slightly outperforms BOGO, with an average of $53.18, compared to BOGO's $52.52.
- Summer: BOGO (Buy One Get One) takes the lead in summer with an average of $53.27, whereas Discount on Selected Items lags behind at $51.84.
- Winter: Both promotions perform similarly, but BOGO (Buy One Get One) slightly edges out Discount on Selected Items with $52.87 versus $52.01.

Question - 05

How effective are different promotions in driving higher transaction amounts, and which promotion type performs best in each season?

Insights:

3. General Observations:

•BOGO (Buy One Get One) consistently performs well across all seasons, maintaining relatively high average transaction values.

•Discount on Selected Items appears to perform better in Spring, where it slightly surpasses BOGO in average transaction amounts, while in Summer, BOGO is the top performer.

•The overall average transaction amounts are very similar for both promotions, with BOGO having a marginal edge in some seasons, except for Spring where Discount on Selected Items shows a slight advantage.

Conclusion:

•BOGO (Buy One Get One) promotion is generally more effective in driving higher transaction amounts, particularly in Summer and Winter.

•Discount on Selected Items seems to perform best in Spring, suggesting a seasonal preference for this promotion type.

•Both promotions are quite close in performance across all seasons, but specific seasonal preferences may influence the choice of promotion for maximizing sales.

Question - 05

How effective are different promotions in driving higher transaction amounts, and which promotion type performs best in each season?

Python Code:

```
df['Date'] = pd.to_datetime(df['Date'])

# Extract year and month from Date column
df['Year_Month'] = df['Date'].dt.to_period('M')

# Filter data based on promotion type
promotion_data = df.groupby(['Promotion', 'Season'])['Amount($)'].agg(['mean', 'sum', 'count']).reset_index()

# Plotting the data
plt.figure(figsize=(14, 7))
ax = sns.barplot(x='Promotion', y='mean', hue='Season', data=promotion_data)

# Annotate each bar with its corresponding value
for p in ax.patches:
    ax.annotate(f'{p.get_height():.2f}',
                (p.get_x() + p.get_width() / 2., p.get_height()),
                ha='center', va='center',
                fontsize=10, color='black',
                xytext=(0, 5), textcoords='offset points')

# Add title and labels
plt.title('Average Transaction Amount by Promotion Type Across Seasons')
plt.xlabel('Promotion Type')
plt.ylabel('Average Transaction Amount ($)')
plt.xticks(rotation=45)
plt.legend(title='Season', bbox_to_anchor=(1.05, 1), loc='upper left')
plt.tight_layout()
plt.show()

# Insights by promotion type and season
promotion_data
```

Conclusion

Key Insights from Data Analysis

Store Type & Seasonality Impact

- Specialty Stores and Supermarkets show consistent transaction amounts, with Spring and Winter driving higher sales.
- Seasonal Trends: Winter and Spring experience higher sales, particularly in Supermarkets, with a slight dip in Summer.

Payment Method Preferences

- Debit Cards dominate in cities like Dallas and Los Angeles, while Mobile Payments are gaining popularity in New York and Miami.
- Payment Method Insights: Debit cards continue to lead in larger transactions, but mobile payments show strong growth, particularly in tech-savvy regions.

Discounted vs Non-Discounted Sales

- Discounted Sales generally drive higher revenue, especially in January and November-December periods, with a peak in January 2024.
- Trends: Both discounted and non-discounted sales have shown growth over the years, but discounts significantly boost sales, especially in the first quarter.

Conclusion

Key Insights from Data Analysis

Top Cities with High Transaction Item Volume

Houston, Chicago, and Miami show the highest number of items per transaction, with Chicago leading in overall sales.

Seasonal Trends: Chicago remains consistent year-round, while Houston and Miami show peak sales in Fall and Spring.

Promotion Effectiveness

BOGO (Buy One Get One) promotion consistently outperforms Discount on Selected Items in terms

Seasonal Insights: Discount on Selected Items performs slightly better in Spring, while BOGO dominates Summer and Winter.

Key Takeaways:

- Strategic Promotions like BOGO are highly effective year-round, especially during peak seasons like Summer and Winter.
- Seasonal fluctuations are evident, and understanding seasonal buying behavior can optimize marketing strategies and promotions.
- Cities like Chicago, Houston, and Miami offer opportunities for tailored promotions based on seasonal and transaction trends.

Thanks