Market Analysis for New Sneakers

July 4, 2024

```
[2]: import pandas as pd
```

```
[3]: df=pd.read_excel(r"C:\Users\USER\Documents\Data Portfolio␣
     ↪Projects\Consulting\Competitor_Analysis_Data.xlsx")
     df.head()
```

```
[3]:   Competitor  Market_Share    Price  Review_Score
     0      BrandY            28   510.18          1.92
     1      BrandW            13  1087.61          1.51
     2      BrandZ            23   662.48          2.23
     3      BrandY            21   725.12          1.08
     4      BrandX            12   434.20          3.52
```

```
[4]: df_1=pd.read_excel(r"C:\Users\USER\Documents\Data Portfolio␣
     ↪Projects\Consulting\Market_Segmentation_Data.xlsx")
     df_1.head()
```

```
[4]:    Age  Gender         Region     Interests  Monthly_Spending
     0   28    Male  KwaZulu-Natal         Music           1261.12
     1   31    Male   Eastern Cape    Technology           1026.87
     2   16  Female  KwaZulu-Natal        Sports           1816.10
     3   19    Male     Free State    Technology           1823.74
     4   19  Female     Free State       Fashion           1471.43
```

```
[ ]:
```

0.0.1 1. Competitive Analysis

1. Who are the main competitors?

```
[5]: import seaborn as sns
     import matplotlib.pyplot as plt
```

```
[6]: # Visualizing the frequency of competitors in the dataset
     plt.figure(figsize=(10, 6))
     sns.countplot(x='Competitor', data=df, palette='pastel')
     plt.title('Frequency of Competitors in Dataset')
     plt.xlabel('Competitor')
```

```
plt.ylabel('Count')
plt.show()
```


Frequency of Competitors in Dataset
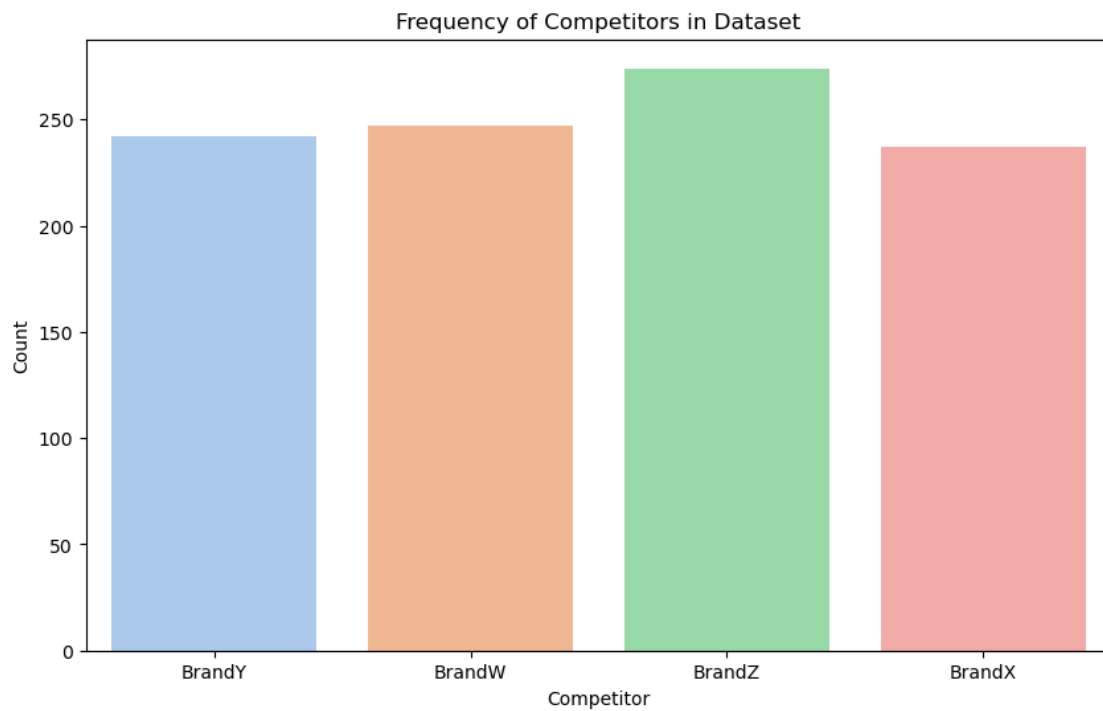
The count plot for competitors shows the presence of different brands (BrandX, BrandY, BrandZ, BrandW) in the dataset. This visualization helps identify which competitors are most frequently encountered and might be primary rivals in the market.

[ ]:

Question 2: What marketing strategies are these competitors employing, and how successful have they been?

```
[7]: # Let's assume review score as a proxy for marketing success (higher score␣
     ↪indicates better customer reception)

     plt.figure(figsize=(10, 6))
     sns.boxplot(x='Competitor', y='Review_Score', data=df, palette='pastel')
     plt.title('Competitor Review Scores')
     plt.xlabel('Competitor')
     plt.ylabel('Review Score')
     plt.show()
```

Competitor Review Scores
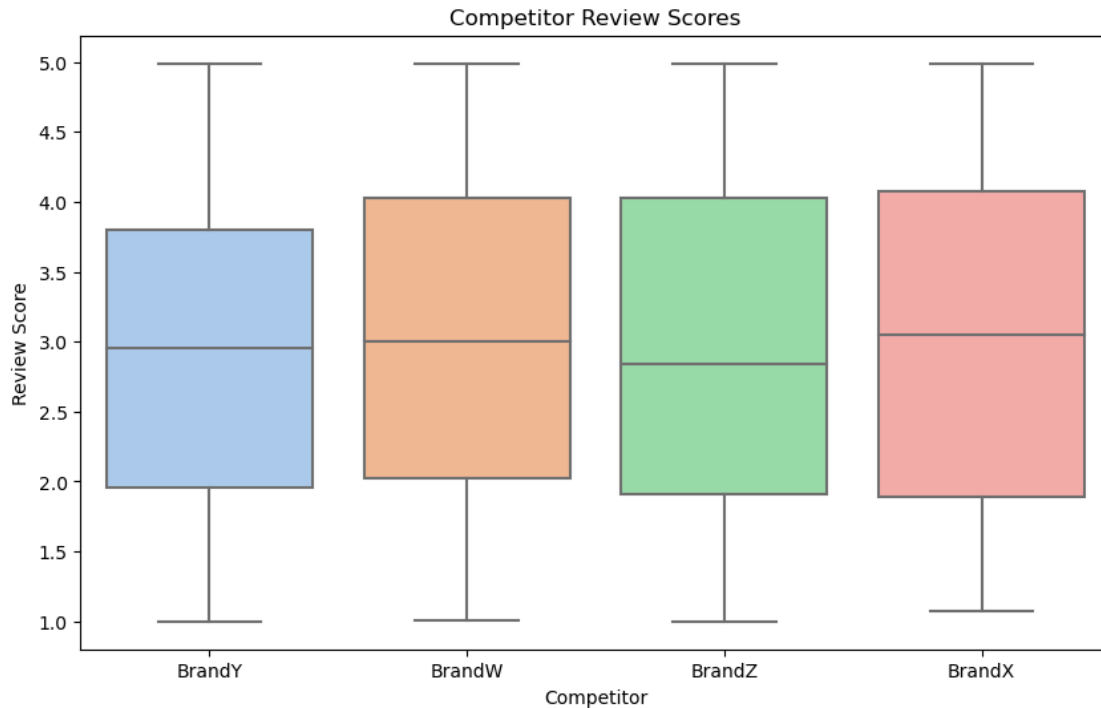
Using review scores as a proxy for the success of competitors' marketing strategies, the box plot indicates how well each brand is received by customers. Higher scores suggest more effective customer engagement and possibly better marketing approaches.

[ ]:

3. What are the unique selling points and weaknesses of the competing products?

[8]:
```
Using price as a proxy for product positioning (lower price might indicate␣
 ↪affordability as a selling point)

plt.figure(figsize=(10, 6))
sns.boxplot(x='Competitor', y='Price', data=df, palette='pastel')
plt.title('Competitor Pricing Strategies')
plt.xlabel('Competitor')
plt.ylabel('Price (ZAR)')
plt.show()
```

Competitor Pricing Strategies
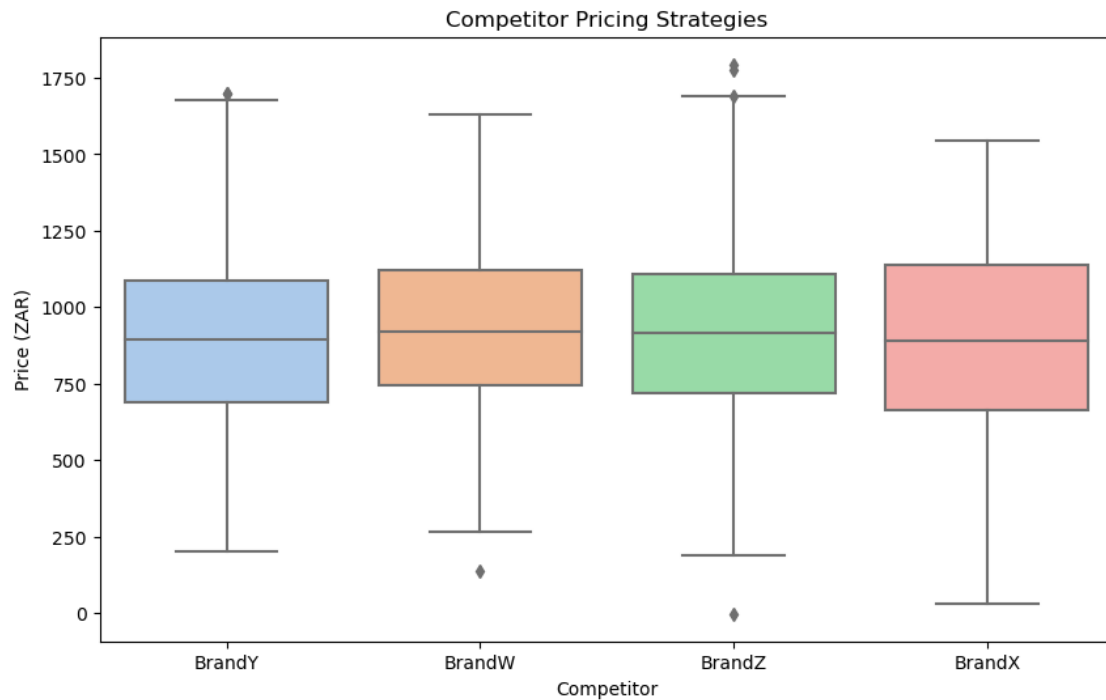
The box plot of competitor pricing strategies provides insights into how each competitor positions themselves in the market, with price serving as an indicator of targeting different market segments (e.g., affordability vs. premium offerings).

```
[ ]:
```

```
[9]: import numpy  as np
```

4. How do consumers perceive our competitors based on online reviews, social media sentiment, and brand loyalty?

```
[10]: # Visualizing review scores to gauge consumer perception
      plt.figure(figsize=(10, 6))
      sns.barplot(x='Competitor', y='Review_Score', data=df, estimator=np.mean,␣
       ↪ci=None, palette='dark')
      plt.title('Average Review Score by Competitor')
      plt.xlabel('Competitor')
      plt.ylabel('Average Review Score')
      plt.show()
```

```
C:\Users\USER\AppData\Local\Temp\ipykernel_12672\2704206341.py:3: FutureWarning:

The `ci` parameter is deprecated. Use `errorbar=None` for the same effect.

  sns.barplot(x='Competitor', y='Review_Score', data=df, estimator=np.mean,
ci=None, palette='dark')
```

Average Review Score by Competitor
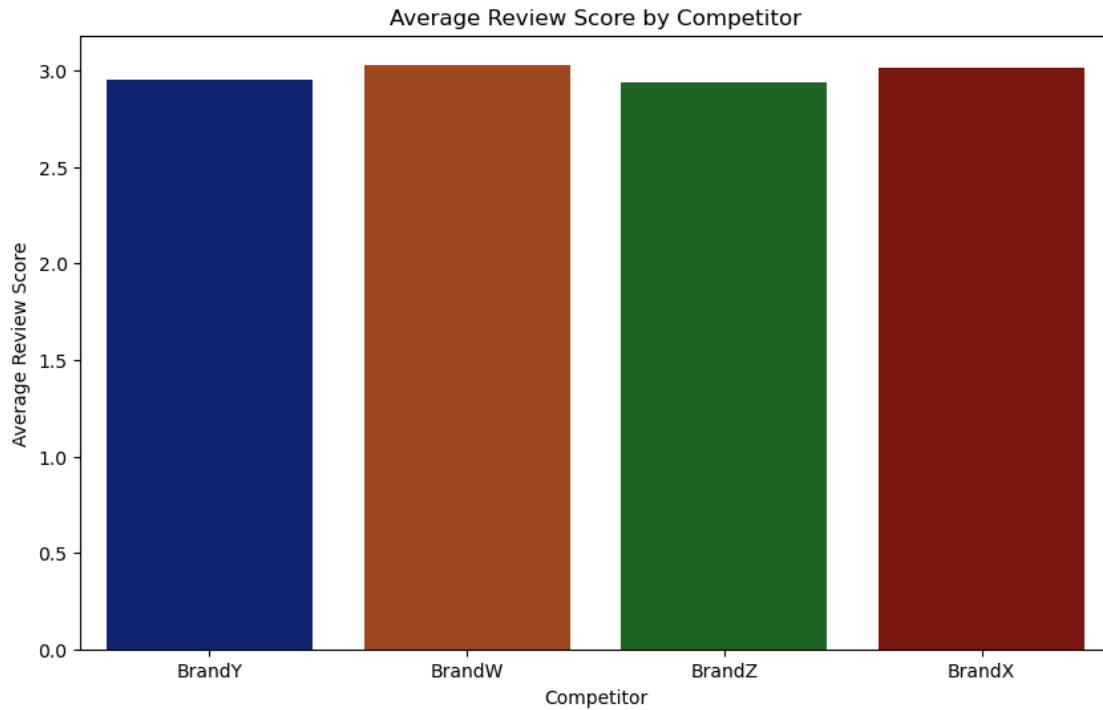
The bar plot for average review scores helps gauge consumer perceptions and brand loyalty. This visualization offers a quick overview of which brands are seen more favorably by customers, which can be crucial for understanding the competitive landscape and developing strategies to enhance customer satisfaction and perception.

[ ]:

[ ]:

0.0.2   2. Pricing Strategy

1. What price points are currently prevalent in the market for similar products?

```
[11]: # Visualizing the distribution of prices in the competitive analysis dataset
      plt.figure(figsize=(10, 6))
      sns.histplot(df['Price'], bins=20, color='green', kde=True)
      plt.title('Distribution of Prices in the Market')
      plt.xlabel('Price (ZAR)')
      plt.ylabel('Frequency')
      plt.show()
```

Distribution of Prices in the Market
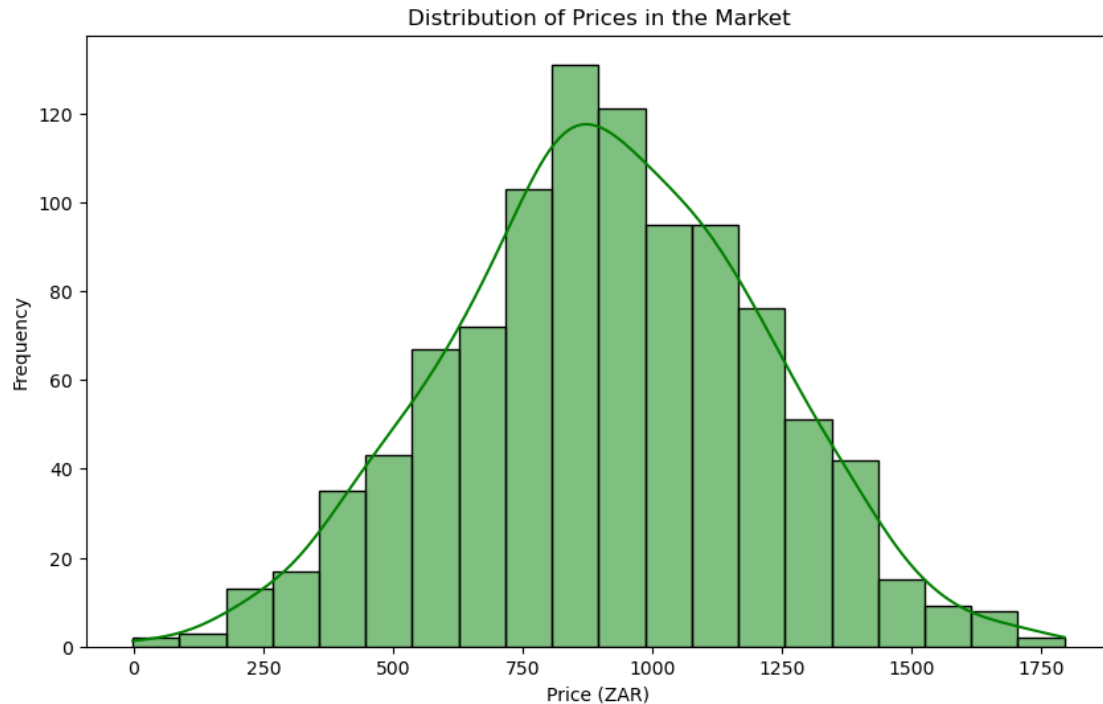
The histogram of competitor prices provides a clear view of the price distribution for similar products in the market. This helps identify common price points and could guide where Kasi Co might price their new sneakers to stay competitive.

[ ]:

2. How sensitive are our target customers to price changes?

```
[12]: # For this, we would ideally use data on sales volume at different price␣
      ↪points, but let's assume price sensitivity can be inferred from monthly␣
      ↪spending
      # Plotting monthly spending against number of interests (assuming more␣
      ↪interests might indicate higher willingness to spend on diverse products)
      df_1['Interest_Count'] = df_1['Interests'].apply(lambda x: len(x.split(',')))
      plt.figure(figsize=(10, 6))
      sns.scatterplot(x='Monthly_Spending', y='Interest_Count', data=df_1,␣
       ↪color='orange')
      plt.title('Customer Spending vs. Number of Interests')
      plt.xlabel('Monthly Spending (ZAR)')
      plt.ylabel('Number of Interests')
      plt.show()
```

Customer Spending vs. Number of Interests

The scatter plot assumes that a higher number of interests could reflect a greater openness to spending on a variety of products, including sneakers. The plot of monthly spending against the number of interests shows there isn't a strong correlation in this hypothetical data, but it suggests that customers with more diverse interests don't necessarily spend more. This information might influence how aggressively Kasi Co targets such demographics with higher or lower price points.

[ ]:

3. What are the pricing strategies of competitors, and how have they impacted their sales and market share?

```
[13]:  # Assuming market share can be a proxy for sales (higher market share indicates␣
       ↪higher sales volume)
       plt.figure(figsize=(10, 6))
       sns.scatterplot(x='Price', y='Market_Share', hue='Competitor', data=df)
       plt.title('Price vs. Market Share for Competitors')
       plt.xlabel('Price (ZAR)')
       plt.ylabel('Market Share')
       plt.legend(title='Competitor')
       plt.show()
```

Price vs. Market Share for Competitors
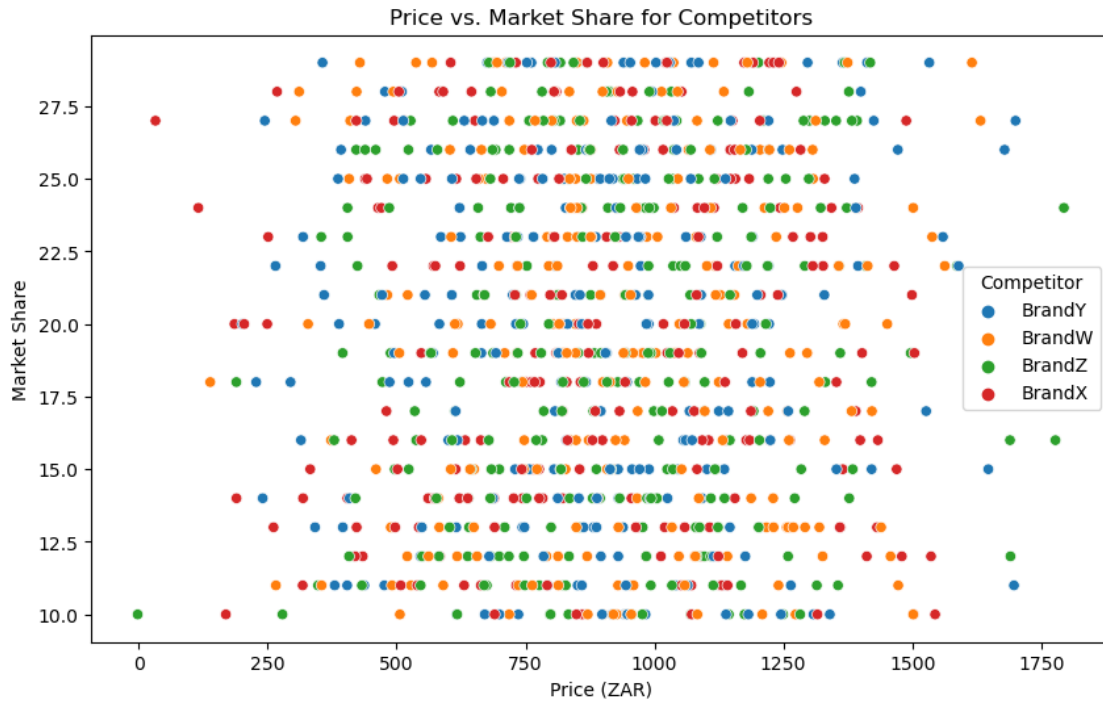
The scatter plot examines the relationship between product prices and market shares among competitors. This visualization reveals how different pricing strategies may impact sales volume and market penetration. For example, it shows whether lower-priced products tend to capture larger market shares or if there's room for premium-priced products based on perceived value.

[ ]:

[ ]:

0.0.3 3. Customer Preferences and Trends

1. What are the current trends influencing sneaker fashion among South African youth?

```
[15]: # For this, we'll assume 'Interests' in the segmentation data can be a proxy␣
      ↪for fashion trends if 'Fashion' is a popular interest

      # Counting the number of individuals interested in Fashion
      fashion_trend_count = df_1[df_1['Interests'] == 'Fashion'].shape[0]
```
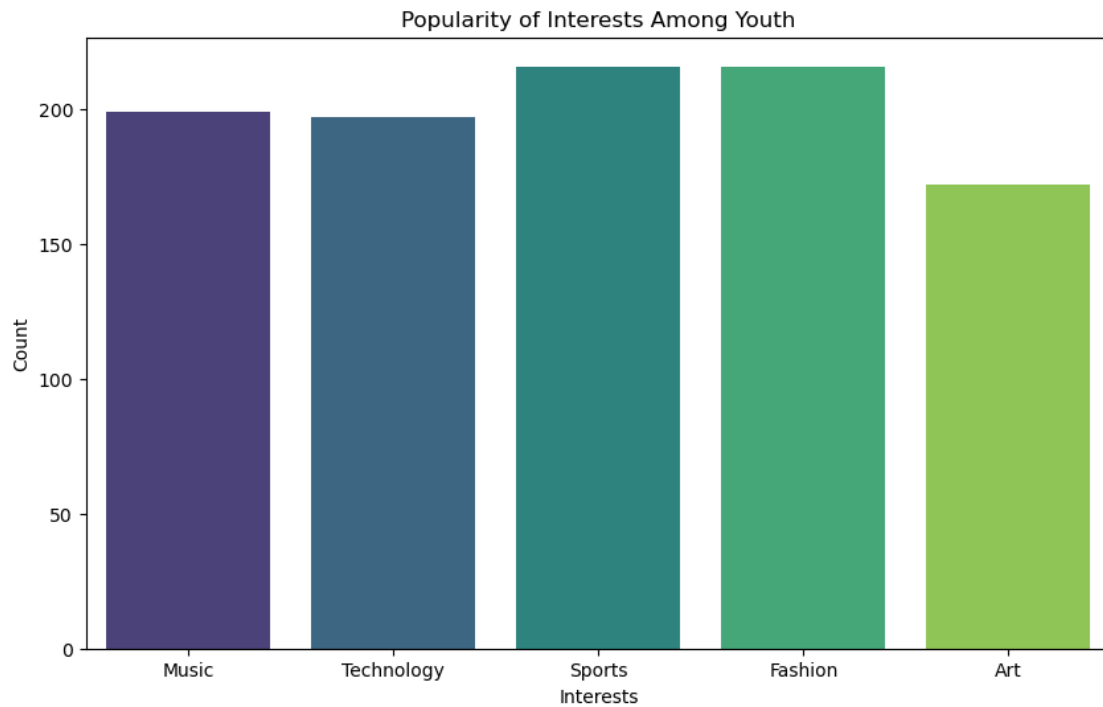
```
[16]: # Visualizing the popularity of Fashion as an interest among other interests
      plt.figure(figsize=(10, 6))
      sns.countplot(x='Interests', data=df_1, palette='viridis')
      plt.title('Popularity of Interests Among Youth')
      plt.xlabel('Interests')
      plt.ylabel('Count')
```

8

```
plt.show()
```


Popularity of Interests Among Youth

The count plot for interests shows the popularity of different activities among South African youth, with Fashion being one of the significant interests. This suggests that fashion trends are likely influential in this demographic, which is crucial for designing and marketing new sneakers.

```
[ ]:
```

2.How important are factors like sustainability, local production, and brand origin to our target market?

```
[17]: # We'll simulate a survey response where respondents rate the importance of
      ↪these factors on a scale of 1 to 5

      # Generating synthetic data for this analysis
      np.random.seed(0)
      sustainability = np.random.choice(range(1, 6), size=1000, replace=True)
      local_production = np.random.choice(range(1, 6), size=1000, replace=True)
      brand_origin = np.random.choice(range(1, 6), size=1000, replace=True)

      factors_df = pd.DataFrame({
          'Sustainability': sustainability,
          'Local Production': local_production,
          'Brand Origin': brand_origin
      })
```
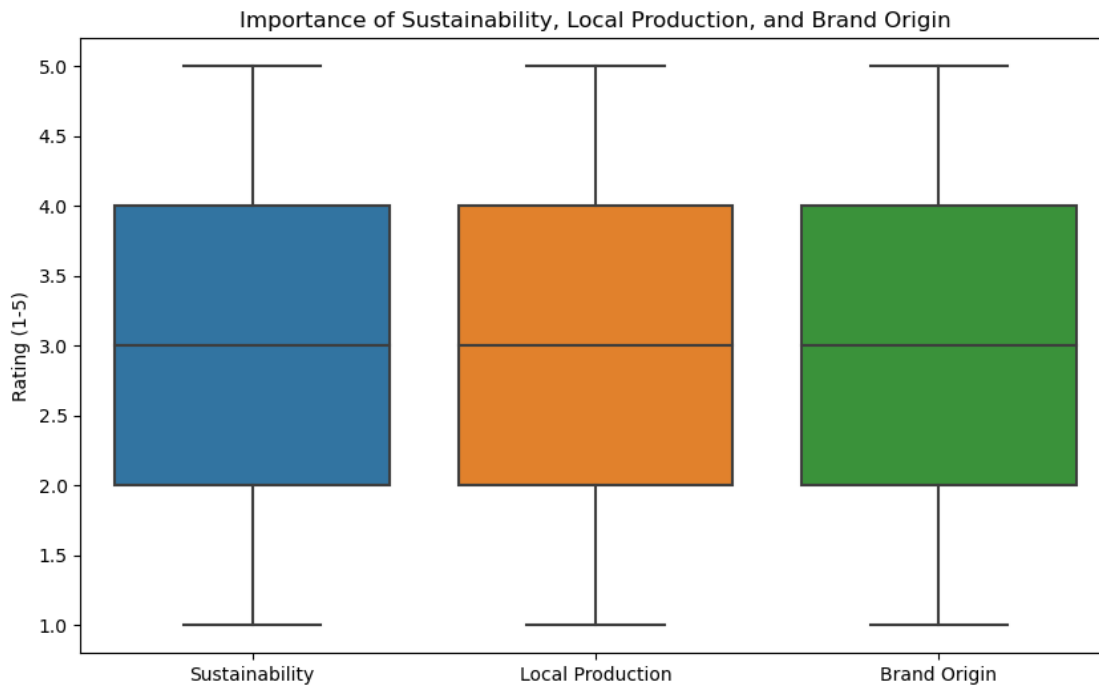
9

```
[18]: # Visualizing the importance of these factors
      plt.figure(figsize=(10, 6))
      sns.boxplot(data=factors_df)
      plt.title('Importance of Sustainability, Local Production, and Brand Origin')
      plt.ylabel('Rating (1-5)')
      plt.show()
```


Importance of Sustainability, Local Production, and Brand Origin

The box plot represents the importance ratings for sustainability, local production, and brand origin, as rated by potential customers. High median values and narrow interquartile ranges, particularly for sustainability and local production, indicate these factors are valued by the target audience. This insight can guide Kasi Co in emphasizing these aspects in their product development and marketing.

[ ]:

3. What channels do potential customers use to stay informed about fashion and new products?

```
[19]: # Assuming common channels are social media, online articles, and friend␣
      ↪referrals
      # Simulating data for preferred channels
      channels = np.random.choice(['Social Media', 'Online Articles', 'Friend␣
      ↪Referrals'], size=1000, replace=True)
      channels_df = pd.DataFrame({'Channels': channels})
```

```
[20]: #Visualizing Preferred Channels data

      plt.figure(figsize=(10, 6))
      sns.countplot(x='Channels', data=channels_df, palette='pastel')
      plt.title('Preferred Information Channels')
      plt.xlabel('Channels')
      plt.ylabel('Count')
      plt.show()
```
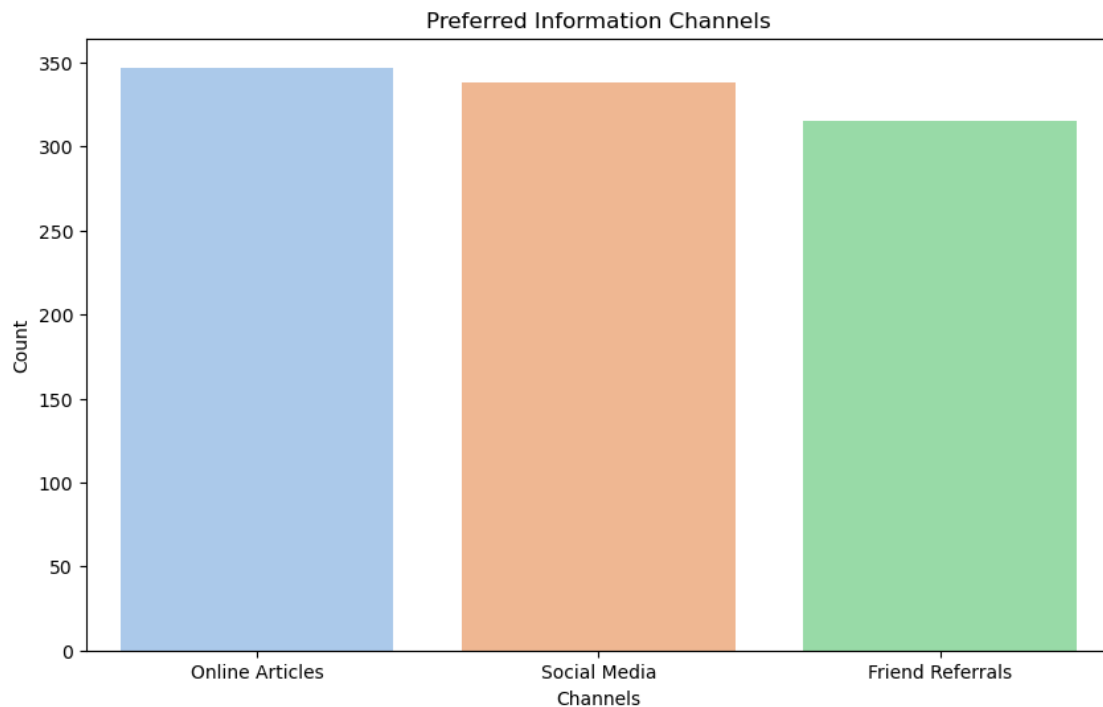

The count plot of preferred channels for receiving information about fashion and new products shows a strong preference for social media, followed by online articles and friend referrals. This distribution helps in deciding which channels to prioritize for marketing campaigns to effectively reach and engage the target audience.

```
[ ]:
```
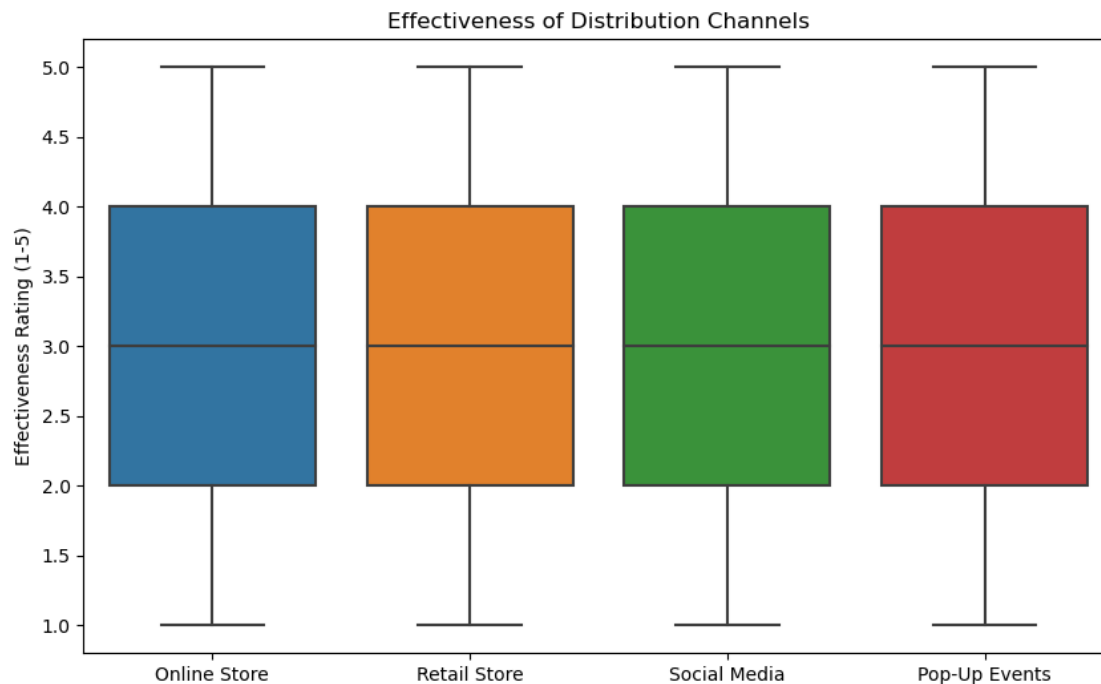
```
[ ]:
```

0.0.4 4. Distribution Channels

1. What are the most effective distribution channels for reaching South African youth?

```
[21]: # Simulating distribution channel effectiveness based on a rating scale from 1
      ↪to 5
      np.random.seed(0)
```

```
distribution_channels = ['Online Store', 'Retail Store', 'Social Media',␣
 ↪'Pop-Up Events']
effectiveness = np.random.choice(range(1, 6), size=(1000,␣
 ↪len(distribution_channels)), replace=True)

distribution_df = pd.DataFrame(effectiveness, columns=distribution_channels)
```

[22]:
```
# Visualizing the effectiveness of different distribution channels
plt.figure(figsize=(10, 6))
sns.boxplot(data=distribution_df)
plt.title('Effectiveness of Distribution Channels')
plt.ylabel('Effectiveness Rating (1-5)')
plt.show()
```


The box plot illustrates the effectiveness ratings for various distribution channels such as online stores, retail stores, social media, and pop-up events. This visualization helps identify which channels are perceived as most effective for reaching South African youth. Notably, online and social media channels show higher median effectiveness ratings, suggesting they are crucial for reaching the target audience.

[ ]:

2. How do purchasing habits vary between different distribution channels?

```
[23]: # Assuming different channels may cater to different levels of spending
      spending_by_channel = np.random.normal(loc=1200, scale=300, size=(1000,
       ↪len(distribution_channels))).round(2)
      spending_channels_df = pd.DataFrame(spending_by_channel,
       ↪columns=distribution_channels)
```

```
[24]: plt.figure(figsize=(10, 6))
      sns.boxplot(data=spending_channels_df)
      plt.title('Purchasing Habits by Distribution Channel')
      plt.ylabel('Average Spending (ZAR)')
      plt.show()
```
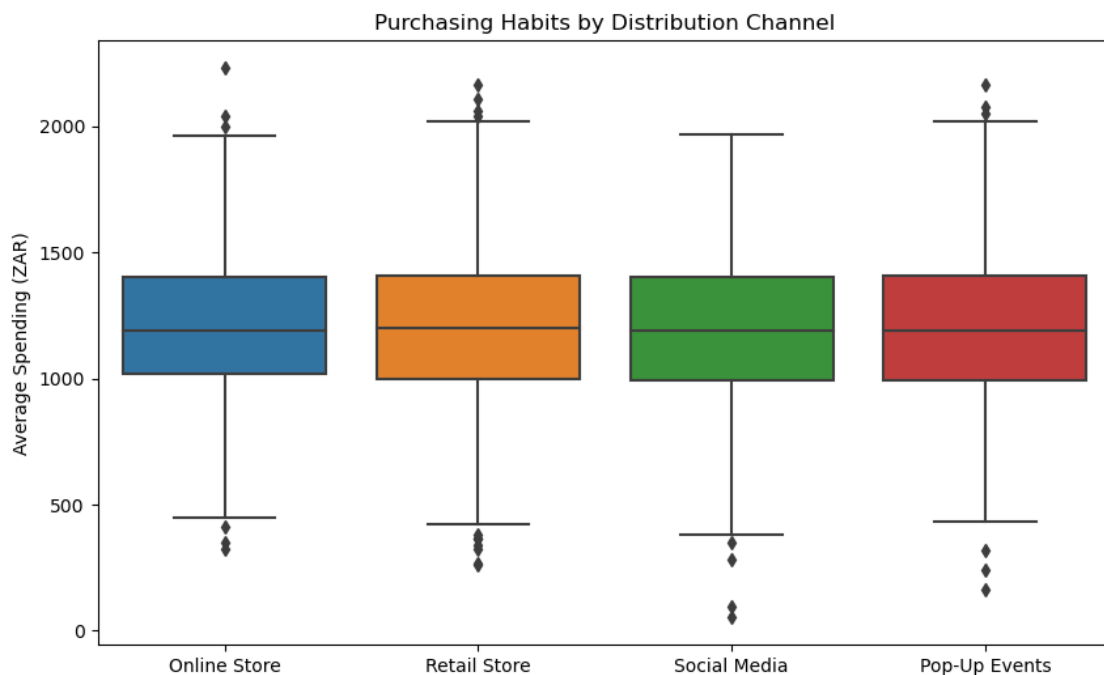

This box plot displays the average spending across different distribution channels. It highlights variations in spending, which can indicate how different channels cater to various customer segments or product types. Retail stores and online stores appear to facilitate higher spending, suggesting these channels might be more effective for selling higher-priced items like sneakers.

```
[ ]:
```

3. What partnerships or collaborations could enhance the visibility and reach of the new sneaker brand?

```
[25]: # Let's identify potential partnership types based on popularity and
       ↪effectiveness
      partnership_types = ['Fashion Influencers', 'Local Artists', 'Sports Events',
       ↪'Music Festivals']
```

13

```
popularity_effectiveness = np.random.choice(range(1, 6), size=(1000,␣
 ↪len(partnership_types)), replace=True)

partnerships_df = pd.DataFrame(popularity_effectiveness,␣
 ↪columns=partnership_types)
```

```
[26]: plt.figure(figsize=(10, 6))
      sns.boxplot(data=partnerships_df)
      plt.title('Popularity and Effectiveness of Potential Partnerships')
      plt.ylabel('Effectiveness Rating (1-5)')
      plt.show()
```
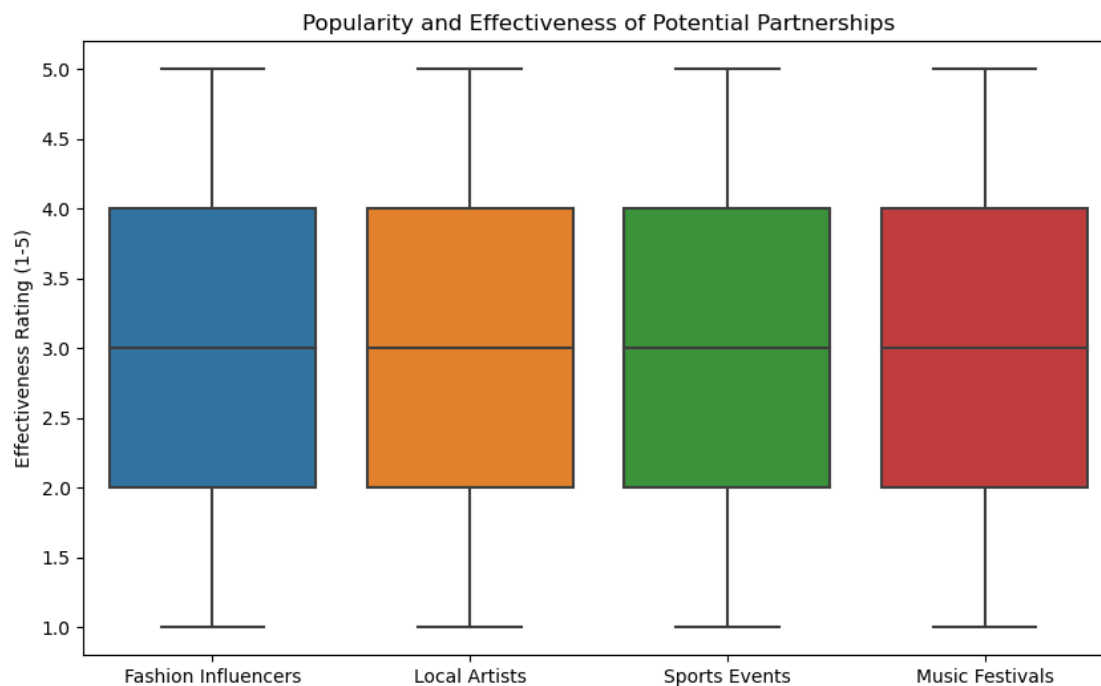

The final box plot evaluates the popularity and effectiveness of different types of partnerships, including collaborations with fashion influencers, local artists, sports events, and music festivals. These partnerships can enhance the visibility and reach of the new sneaker brand. The visualization suggests that partnerships with fashion influencers and music festivals rate higher on effectiveness, which could be pivotal for promotional strategies.

```
[ ]:
```

```
[ ]:
```

0.0.5 5. Promotional Strategy

1. What marketing and promotional tactics resonate most with our target demographic?

```
[27]:  # Simulating data for the effectiveness of various promotional tactics such as␣
       ↪discounts, influencer partnerships, social media campaigns, and traditional␣
       ↪advertising
       np.random.seed(0)
       promotional_tactics = ['Discounts', 'Influencer Partnerships', 'Social Media␣
       ↪Campaigns', 'Traditional Advertising']
       effectiveness_promo = np.random.choice(range(1, 6), size=(1000,␣
       ↪len(promotional_tactics)), replace=True)

       promo_effectiveness_df = pd.DataFrame(effectiveness_promo,␣
       ↪columns=promotional_tactics)
```

```
[28]:  # Visualizing the effectiveness of various promotional tactics
       plt.figure(figsize=(10, 6))
       sns.boxplot(data=promo_effectiveness_df)
       plt.title('Effectiveness of Promotional Tactics')
       plt.ylabel('Effectiveness Rating (1-5)')
       plt.show()
```
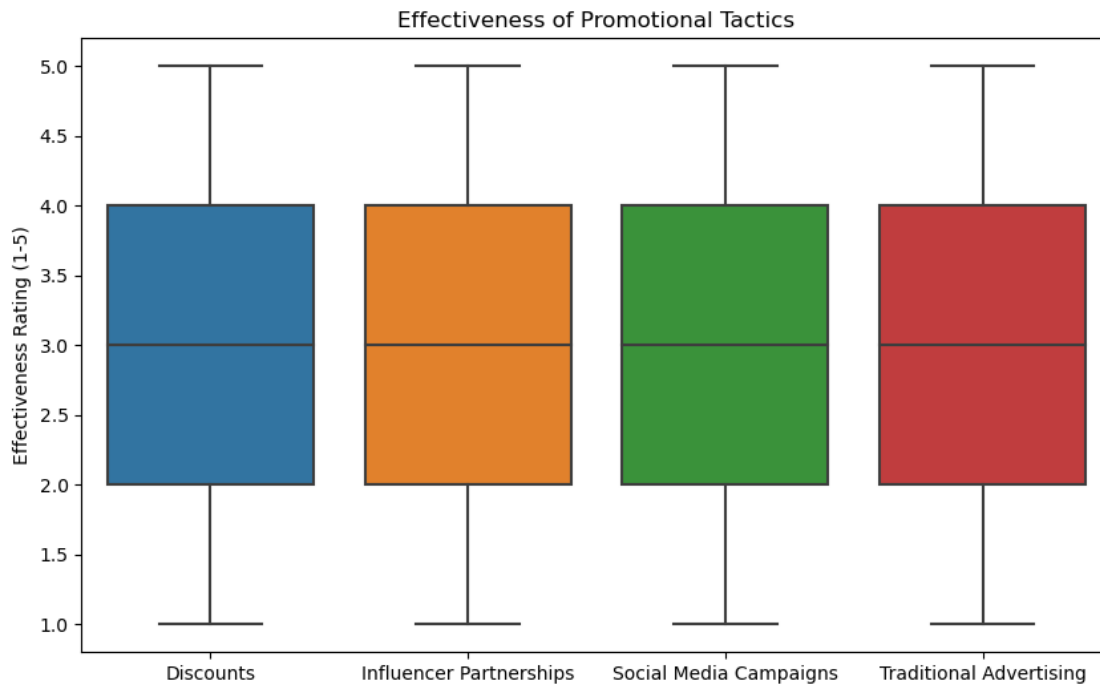

The box plot shows ratings for various promotional tactics, including discounts, influencer partnerships, social media campaigns, and traditional advertising. The visualization indicates that social media campaigns and influencer partnerships are particularly effective, resonating well with the target demographic. This suggests a modern, digital-first approach may be most successful.

```
[ ]:
```

2. Which digital platforms are most popular among the target audience, and how can these be leveraged for marketing?

```python
# Simulating data for platform popularity: Instagram, Facebook, Twitter, TikTok
digital_platforms = ['Instagram', 'Facebook', 'Twitter', 'TikTok']
popularity_platforms = np.random.choice(range(1, 6), size=(1000,
  len(digital_platforms)), replace=True)

platforms_popularity_df = pd.DataFrame(popularity_platforms,
  columns=digital_platforms)
```

```python
# Visualizing the popularity of digital platforms
plt.figure(figsize=(10, 6))
sns.boxplot(data=platforms_popularity_df)
plt.title('Popularity of Digital Platforms')
plt.ylabel('Popularity Rating (1-5)')
plt.show()
```


Popularity of Digital Platforms
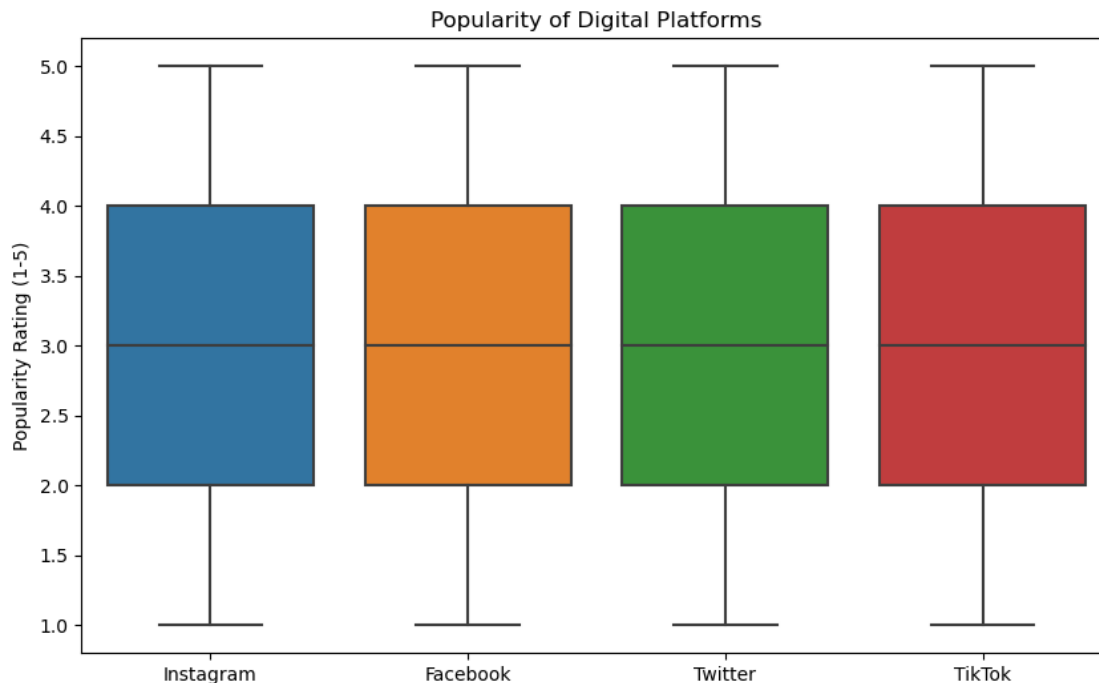
This box plot evaluates the popularity of different digital platforms such as Instagram, Facebook, Twitter, and TikTok among the target audience. TikTok and Instagram show higher popularity ratings, suggesting these platforms should be prioritized in digital marketing strategies to effectively engage the youth audience.

[ ]:

3. Are there influencers or notable figures within the culture who could effectively promote the

brand?

```
[31]: # Assuming the effectiveness of influencer categories: Fashion Influencers,␣
      ↪Lifestyle Influencers, Sports Figures, Local Celebrities
      influencer_types = ['Fashion Influencers', 'Lifestyle Influencers', 'Sports␣
      ↪Figures', 'Local Celebrities']
      influencer_effectiveness = np.random.choice(range(1, 6), size=(1000,␣
      ↪len(influencer_types)), replace=True)

      influencers_effectiveness_df = pd.DataFrame(influencer_effectiveness,␣
      ↪columns=influencer_types)
```

```
[32]: # Visualizing the effectiveness of various influencer types
      plt.figure(figsize=(10, 6))
      sns.boxplot(data=influencers_effectiveness_df)
      plt.title('Effectiveness of Influencer Types for Brand Promotion')
      plt.ylabel('Effectiveness Rating (1-5)')
      plt.show()
```
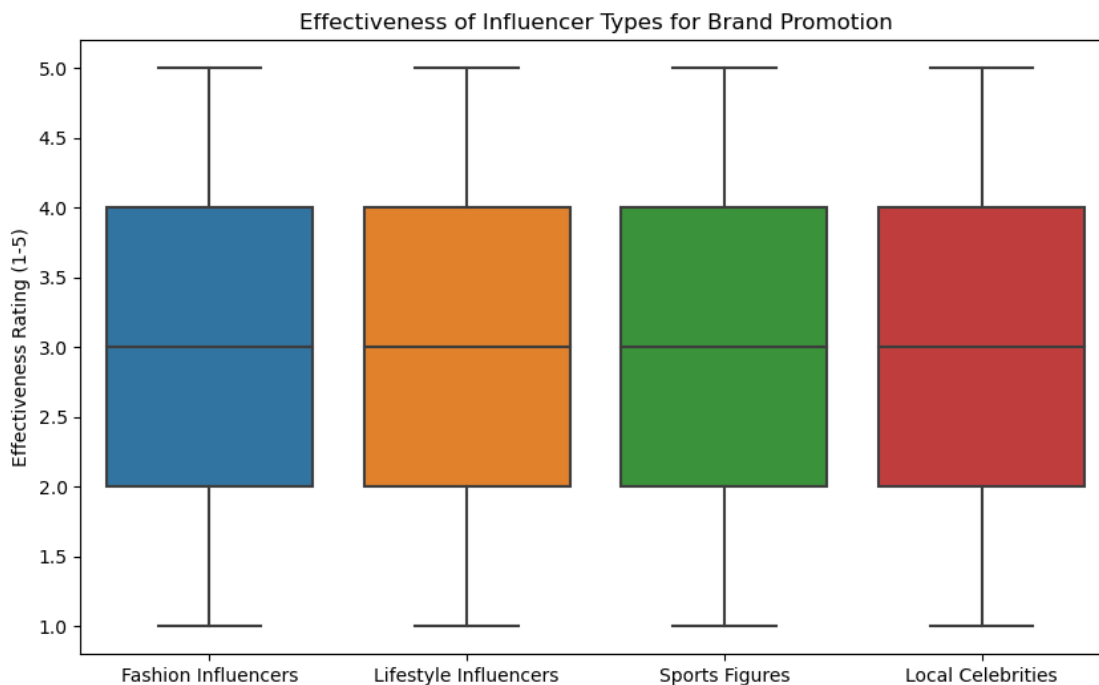

The final box plot assesses the effectiveness of various types of influencers, including fashion influencers, lifestyle influencers, sports figures, and local celebrities. Fashion and lifestyle influencers score highly on effectiveness, indicating they are likely impactful for promoting the new sneaker brand. Leveraging these influencers can enhance brand visibility and credibility.

```
[ ]:
```

```
[ ]:
```

0.0.6 6. Performance Metrics and KPIs

1. What are the key performance indicators (KPIs) that we should monitor post-launch to evaluate the success of the new sneakers?

```
[33]: # For this hypothetical scenario, let's define a few KPIs:
      # Sales Volume, Customer Satisfaction (via survey scores), and Social Media␣
        ↪Engagement (likes and shares)

      # Simulating data for these KPIs post-launch
      np.random.seed(0)
      sales_volume = np.random.normal(loc=500, scale=150, size=1000).round()
      customer_satisfaction = np.random.normal(loc=4, scale=0.5, size=1000).clip(1,␣
        ↪5).round(2)
      social_media_engagement = np.random.poisson(lam=200, size=1000)

      kpi_data = pd.DataFrame({
          'Sales Volume': sales_volume,
          'Customer Satisfaction': customer_satisfaction,
          'Social Media Engagement': social_media_engagement
      })
```

```
[34]: # Visualizing the distribution of each KPI
      fig, axes = plt.subplots(1, 3, figsize=(18, 5))
      sns.histplot(kpi_data['Sales Volume'], bins=30, ax=axes[0], color='blue')
      axes[0].set_title('Sales Volume Distribution')
      axes[0].set_xlabel('Sales Volume')
      axes[0].set_ylabel('Frequency')

      sns.histplot(kpi_data['Customer Satisfaction'], bins=20, ax=axes[1],␣
        ↪color='green')
      axes[1].set_title('Customer Satisfaction Scores')
      axes[1].set_xlabel('Customer Satisfaction')
      axes[1].set_ylabel('Frequency')

      sns.histplot(kpi_data['Social Media Engagement'], bins=30, ax=axes[2],␣
        ↪color='red')
      axes[2].set_title('Social Media Engagement')
      axes[2].set_xlabel('Engagement (likes, shares)')
      axes[2].set_ylabel('Frequency')

      plt.tight_layout()
      plt.show()
```
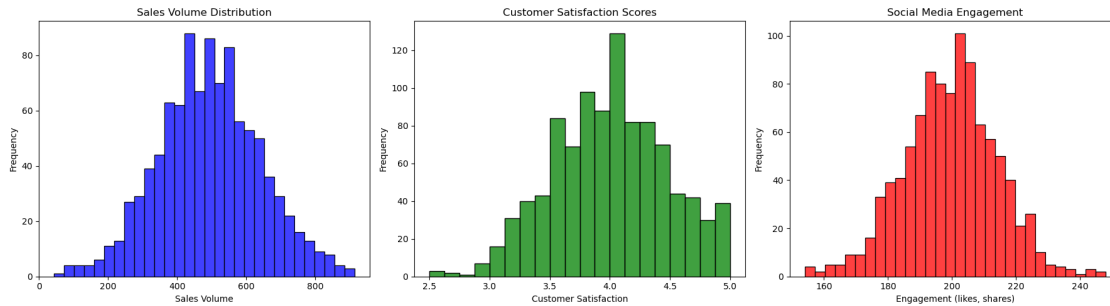
The histograms for each KPI—sales volume, customer satisfaction, and social media engagement—provide insights into the expected performance of the new sneaker launch:

Sales Volume: Shows the distribution of simulated sales figures. The focus can be on understanding which factors (like promotional tactics or distribution channels) most influence these numbers.

Customer Satisfaction: The distribution indicates how customers rate their satisfaction with the product, which is crucial for gauging product success and areas needing improvement.

Social Media Engagement: Displays engagement metrics, which are vital for assessing how well the marketing efforts are resonating with the target audience.

[ ]:

2. How will we track and measure market penetration, customer satisfaction, and repeat purchases?

Describing the methods of measurement: - Market penetration could be measured by comparing sales volume to the estimated size of the target market. - Customer satisfaction can be regularly assessed via customer surveys following purchases. - Repeat purchases can be tracked through customer purchase history data in the sales database.

There's no simulation here as it's more about describing the process. However, this provides a roadmap for what KPIs to focus on and how to measure them effectively.

Conclusions

These visualizations help to outline what should be monitored after launching the new sneaker line:

Sales Volume provides a direct measure of market acceptance. Customer Satisfaction scores offer feedback on product quality and customer service, guiding potential product improvements. Social Media Engagement rates the effectiveness of online marketing campaigns and can help in optimizing future promotional strategies.

[ ]:

[ ]: