Analytical Insights in Retail and Finance

Problem Statement

Pelican Stores, a division of National Clothing, is a chain of women's apparel stores operating throughout the country. The chain recently ran a promotion in which discount coupons were sent to customers of other National Clothing stores. Data collected for a sample of 100 in-store credit card transactions at Pelican Stores during one day while the promotion was running are contained in the data file available at the following link: https://docs.google.com/spreadsheets/d/1Ltc8MpadeyMfz0YExUPsRZztyf8iHoag/edit?usp=sharing&ouid=103941265042628485322&rtpof=true&sd=true (https://docs.google.com/spreadsheets/d/1Ltc8MpadeyMfz0YExUPsRZztyf8iHoag/edit?usp=sharing&ouid=103941265042628485322&rtpof=true&sd=true)

The Proprietary Card method of payment refers to charges made using a National Clothing charge card. Customers who made a purchase using a discount coupon are referred to as promotional customers and customers who made a purchase but did not use a discount coupon are referred to as regular customers. Because the promotional coupons were not sent to regular Pelican Stores customers, management considers the sales made to people presenting the promotional coupons as sales it would not otherwise make. Of course, Pelican also hopes that the promotional customers will continue to shop at its stores.

Most of the variables shown in data file are self-explanatory, but two of the variables require some clarification. • Items: The total number of items purchased • Net Sales: The total amount ($) charged to the credit card

Pelican's management would like to use this sample data to learn about its customer base and to evaluate the promotion involving discount coupons.

Required: Use the tabular and graphical methods of descriptive statistics to help management develop a customer profile and to evaluate the promotional campaign. Your report should include the following:

A. Percent frequency distributions for each of the key variables: number of items purchased, net sales, method of payment, gender, marital status, and age.

B. A sorted bar chart showing the number of customer purchases attributable to the method of payment.

C. A cross tabulation of type of customer (regular or promotional) versus net sales. Comment on any similarities of differences observed.

D. A scatter chart to explore the relationship between net sales and customer age.

E. A chart to examine whether the relationship between net sales and age depends on the marital status of the customer.

F. A side-by-side bar chart to examine the method of payment by customer type (regular or promotional). Comment on any differences you observe between the methods of payments used by the different types of customers.

Question 2 [4 Marks]

Mean Annual Growth Rate of Asset. If a financial asset declines in value from $5,000 to$ 3,500 over nine years, what is the mean annual growth rate in the asset's value over those nine years?

Question 1-A

Percent frequency distributions for each of the key variables: number of items purchased, net sales, method of payment, gender, marital status, and age.

Answer 1-A

In [1]:
```
#  libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

# reading from csv
data = pd.read_csv("pelicanstores.csv")
data.head()
```

Out[1]:

	Customer	Type of Customer	Items	Net Sales	Method of Payment	Gender	Marital Status	Age
0	1	Regular	1	39.5	Discover	Male	Married	32
1	2	Promotional	1	102.4	Proprietary Card	Female	Married	36
2	3	Regular	1	22.5	Proprietary Card	Female	Married	32
3	4	Promotional	5	100.4	Proprietary Card	Female	Married	28
4	5	Regular	2	54.0	MasterCard	Female	Married	34

In [2]:
```
# Calculate percent frequency distributions for key variables
key_variables = ["Items", "Net Sales", "Method of Payment", "Gender

for var in key_variables:
    percent_freq = data[var].value_counts(normalize=True) * 100
    print(f"Percent Frequency Distribution for {var}:\n{percent_fre
```

```
Percent Frequency Distribution for Items:
Items
1     29.0
2     27.0
3     10.0
4     10.0
5      9.0
6      7.0
9      3.0
7      1.0
13     1.0
8      1.0
17     1.0
10     1.0
Name: proportion, dtype: float64

Percent Frequency Distribution for Net Sales:
Net Sales
31.60     4.0
39.50     3.0
49.50     3.0
44.50     3.0
39.60     2.0
          ...
52.50     1.0
13.23     1.0
117.50    1.0
102.50    1.0
28.44     1.0
Name: proportion, Length: 84, dtype: float64

Percent Frequency Distribution for Method of Payment:
Method of Payment
Proprietary Card    70.0
MasterCard          14.0
Visa                10.0
Discover             4.0
American Express     2.0
Name: proportion, dtype: float64

Percent Frequency Distribution for Gender:
Gender
Female    93.0
Male       7.0
Name: proportion, dtype: float64

Percent Frequency Distribution for Marital Status:
Marital Status
Married    84.0
Single     16.0
Name: proportion, dtype: float64

Percent Frequency Distribution for Age:
Age
46    9.0
32    8.0
30    8.0
54    8.0
44    7.0
42    7.0
48    6.0
```

```
36    6.0
38    5.0
28    5.0
40    4.0
62    3.0
50    3.0
34    3.0
56    2.0
22    2.0
60    2.0
52    2.0
68    2.0
20    2.0
24    1.0
70    1.0
58    1.0
78    1.0
74    1.0
72    1.0
Name: proportion, dtype: float64
```

Question 1 B

A sorted bar chart showing the number of customer purchases attributable to the method of payment?

Answer 1 B

In [3]:
```
data.head()
```

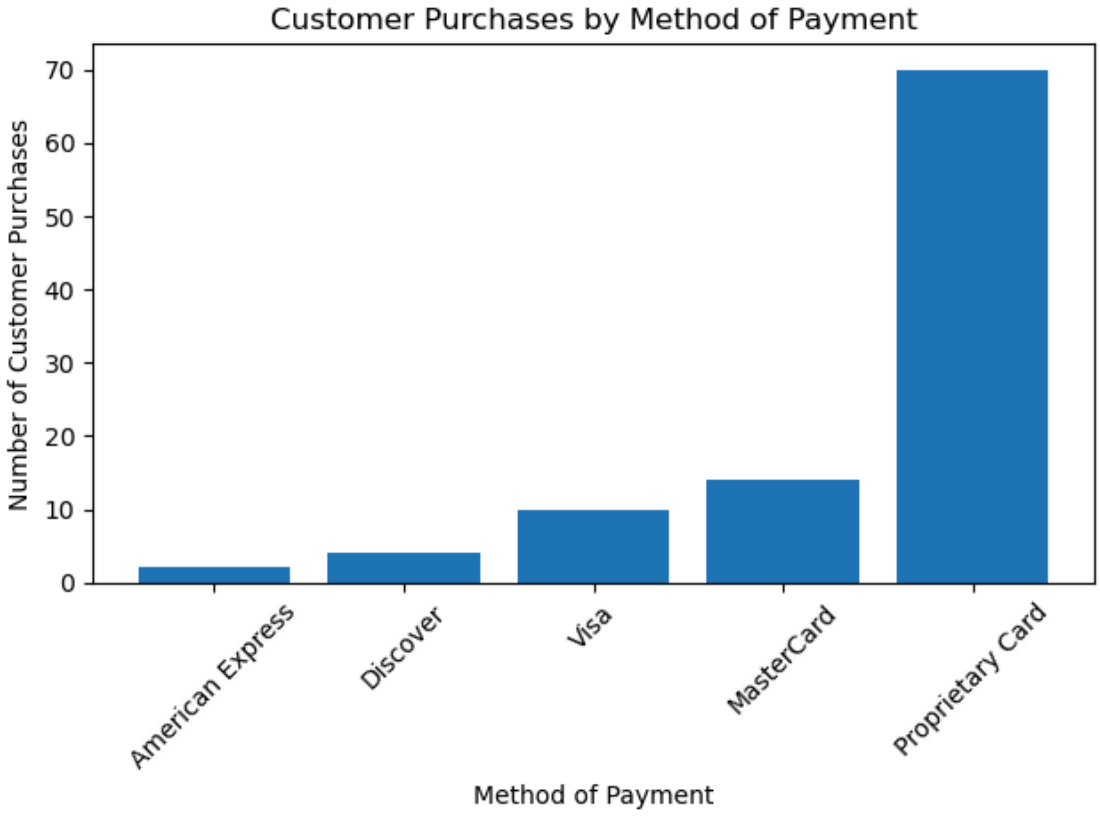
Out[3]:

	Customer	Type of Customer	Items	Net Sales	Method of Payment	Gender	Marital Status	Age
0	1	Regular	1	39.5	Discover	Male	Married	32
1	2	Promotional	1	102.4	Proprietary Card	Female	Married	36
2	3	Regular	1	22.5	Proprietary Card	Female	Married	32
3	4	Promotional	5	100.4	Proprietary Card	Female	Married	28
4	5	Regular	2	54.0	MasterCard	Female	Married	34

In [4]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

# Count the occurrences of each method of payment and sort the valu
sorted_bar_data = data["Method of Payment"].value_counts().sort_val

# Plot the bar chart
plt.bar(sorted_bar_data.index, sorted_bar_data.values)
plt.xlabel("Method of Payment")
plt.ylabel("Number of Customer Purchases")
plt.title("Customer Purchases by Method of Payment")
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```



Question 1 C

A cross tabulation of type of customer (regular or promotional) versus net sales. Comment on any similarities of differences observed.

Answer 1 C

In [5]:
```
data.head()
```

Out[5]:

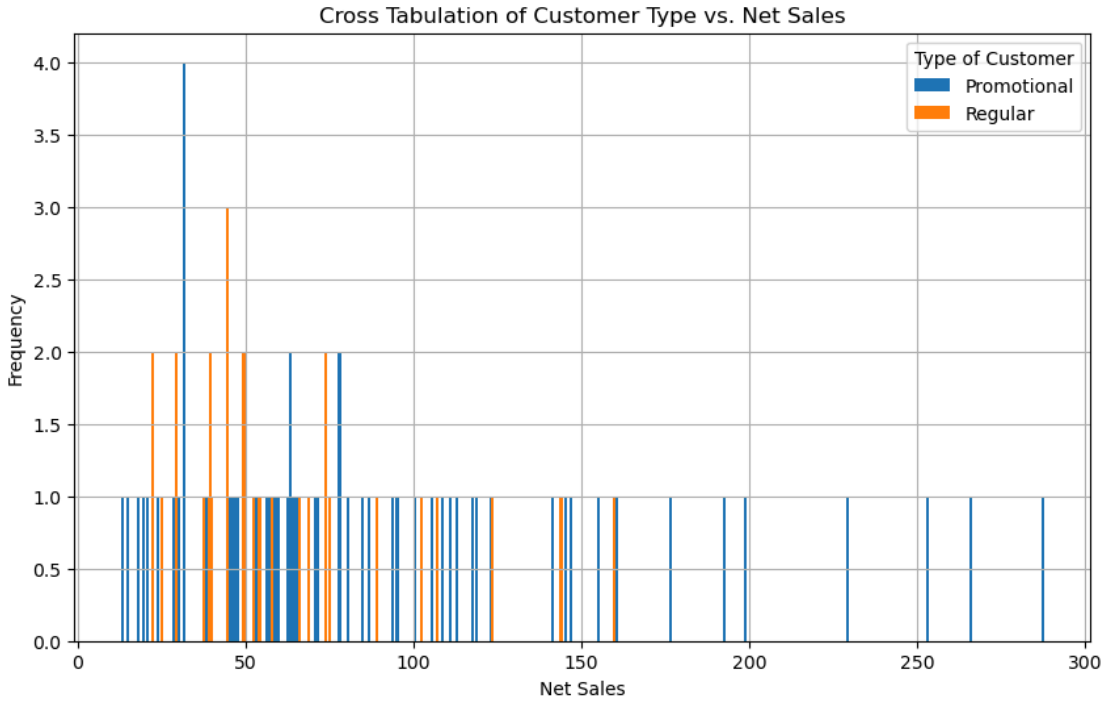
	Customer	Type of Customer	Items	Net Sales	Method of Payment	Gender	Marital Status	Age
0	1	Regular	1	39.5	Discover	Male	Married	32
1	2	Promotional	1	102.4	Proprietary Card	Female	Married	36
2	3	Regular	1	22.5	Proprietary Card	Female	Married	32
3	4	Promotional	5	100.4	Proprietary Card	Female	Married	28
4	5	Regular	2	54.0	MasterCard	Female	Married	34

In [6]:
```
cross_tab = pd.crosstab(data["Type of Customer"],data["Net Sales"])

fig, ax = plt.subplots(figsize=(10, 6))

for idx, row in cross_tab.iterrows():
    ax.bar(row.index, row.values, label=idx)

ax.set_xlabel("Net Sales")
ax.set_ylabel("Frequency")
ax.set_title("Cross Tabulation of Customer Type vs. Net Sales")
ax.legend(title="Type of Customer")
ax.grid(True)
plt.show()
```



Both customer types have made purchases for several net sales amounts such as 13.23, 19.50, 20.80, etc. Overall, the cross-tabulation provides insights into the purchasing behavior of different customer types based on net sales amounts. Further analysis could involve exploring the distribution of net sales amounts for each customer type or investigating any patterns or trends in purchasing behavior across different price points.

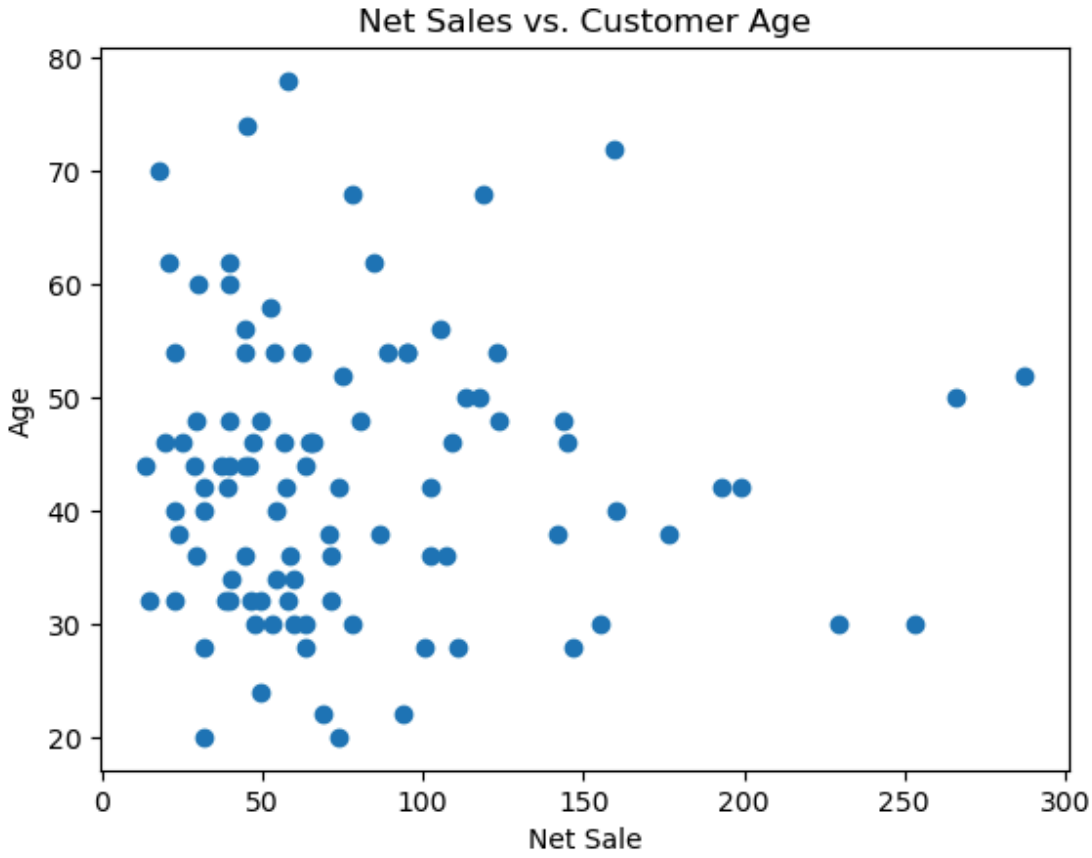
Question 1 D

A scatter chart to explore the relationship between net sales and customer age.

Answer 1 D

In [7]:
```
import matplotlib.pyplot as plt

plt.scatter(data["Net Sales"],data["Age"])
plt.ylabel("Age")
plt.xlabel("Net Sale")
plt.title("Net Sales vs. Customer Age")
plt.show()
```

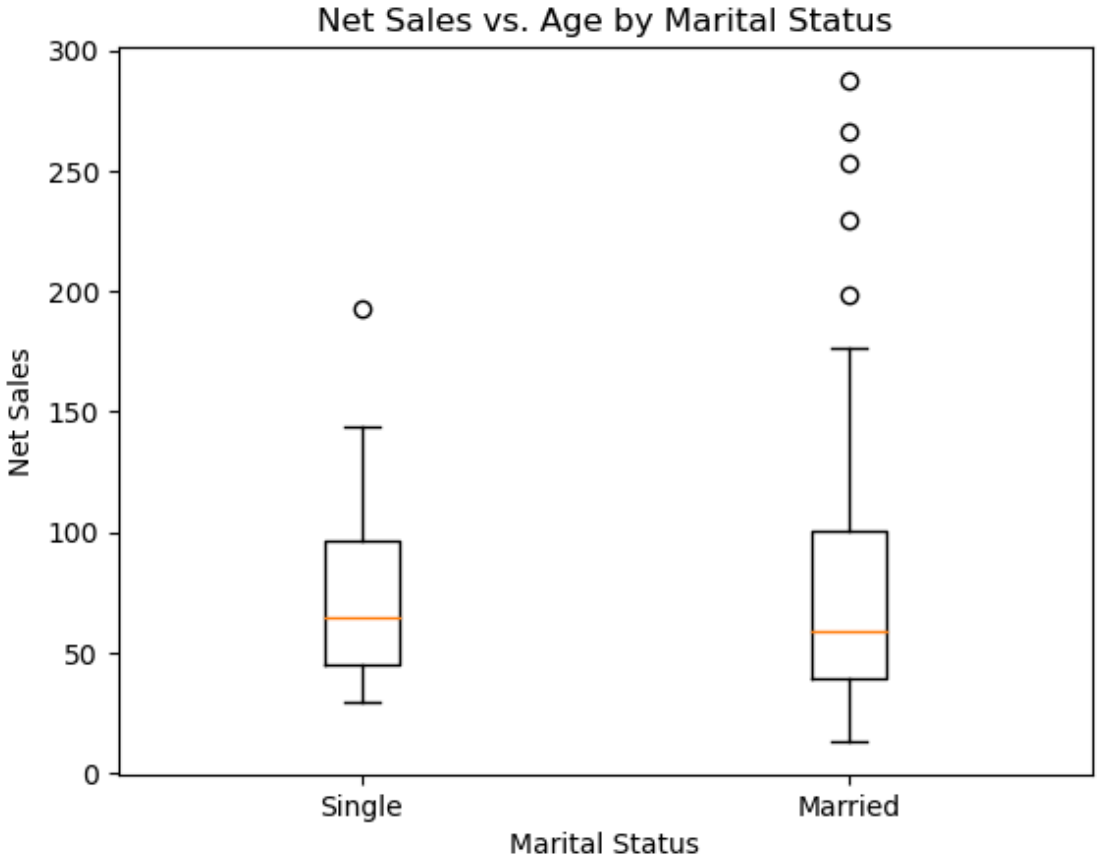


Question 1 E

A chart to examine whether the relationship between net sales and age depends on the marital status of the customer.

Answer 1 E

In [8]:
```
# Create a box plot to examine the relationship between net sales,
plt.boxplot([data[data["Marital Status"] == "Single"]["Net Sales"],
             data[data["Marital Status"] == "Married"]["Net Sales"]
            labels=["Single", "Married"])
plt.xlabel("Marital Status")
plt.ylabel("Net Sales")
plt.title("Net Sales vs. Age by Marital Status")
plt.show()
```

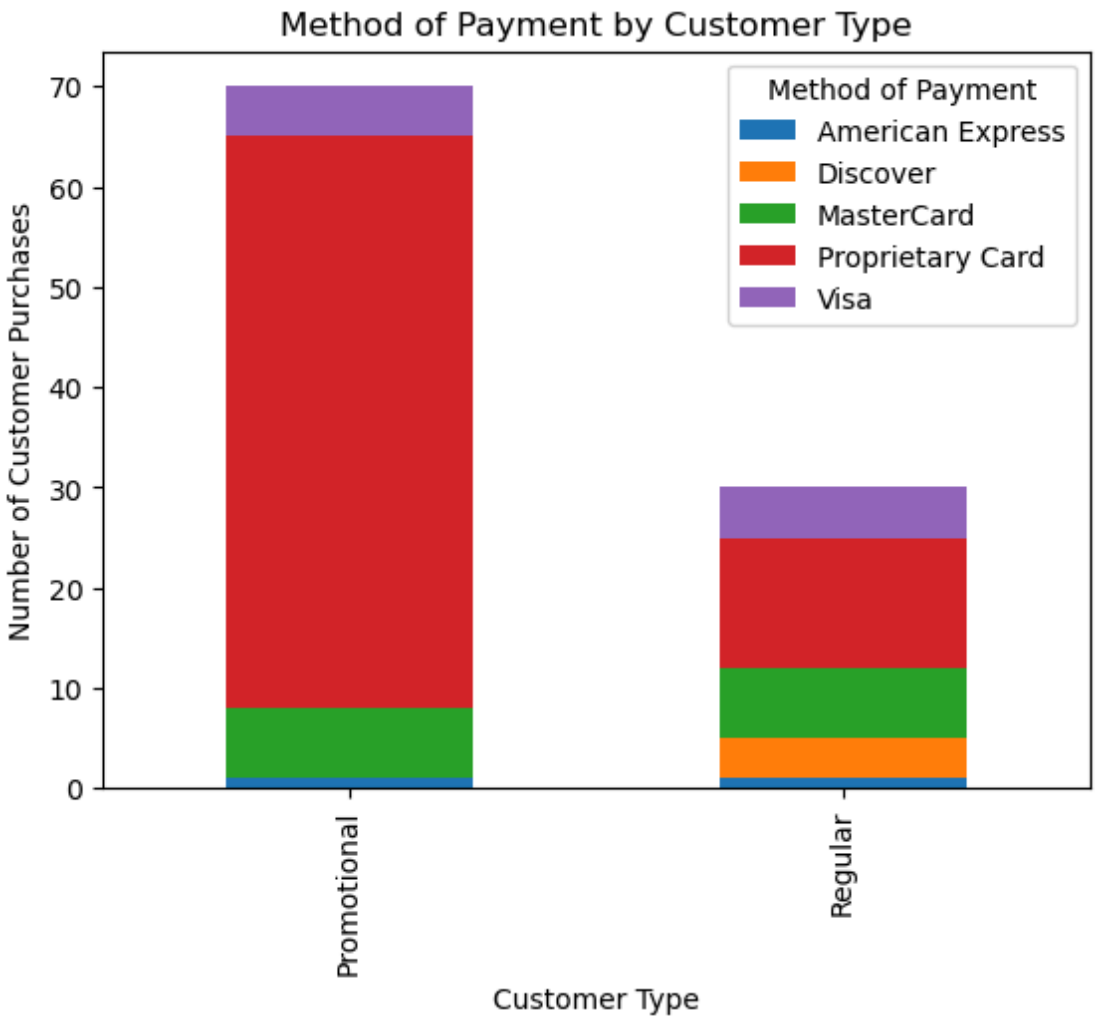


Question 1 F

side-by-side bar chart to examine the method of payment by customer type (regular or promotional). Comment on any differences you observe between the methods of payments used by the different types of customers.

Answer 1 F

In [9]:
```
# Create a side-by-side bar chart
payment_by_type = data.groupby(["Type of Customer", "Method of Paym
payment_by_type.plot(kind="bar", stacked=True)
plt.xlabel("Customer Type")
plt.ylabel("Number of Customer Purchases")
plt.title("Method of Payment by Customer Type")
plt.show()
```



Credit Card Usage: Both Regular and Promotional customers seem to prefer using credit cards (such as MasterCard, Discover, and Visa) as their primary method of payment. This indicates that credit cards are commonly used across both customer types for making purchases. Proprietary Card Usage: There is a noticeable difference in the usage of proprietary cards between Regular and Promotional customers. Regular customers show a higher tendency to use proprietary cards compared to Promotional customers. This suggests that proprietary cards might be more popular or incentivized among Regular customers. Other Payment Methods: While credit cards and proprietary cards dominate the payment methods for both customer types, there are also occurrences of other payment methods such as cash, check, and gift card. However, these alternative methods seem to be less frequently used compared to credit cards and proprietary cards.

Question 2 [4 Marks]

Mean Annual Growth Rate of Asset. If a financial asset declines in value from $5,000 to$ 3,500 over nine years, what is the mean annual growth rate in the asset's value over these nine years?

Answer

In [10]:
```
initial_value = 5000
final_value = 3500
years = 9

annual_decline = (final_value - initial_value) / years
mean_growth_rate = (annual_decline / initial_value) * 100
print(f"Mean Annual Growth Rate: {mean_growth_rate:.2f}%")
```

```
Mean Annual Growth Rate: -3.33%
```

Question 3

The results of a national survey showed that on average adults sleep 6.9 hours per night. Suppose that the standard deviation is 1.2 hours and that the number of hours of sleep follows a bell-shaped distribution.

A. Use the empirical rule to calculate the percentage of in dividuals who sleep between 4.5 and 9.3 hours per day.

B. What is the z-score for an adult who sleeps 8 hours per night?

C. What is the z-score for an adult who sleeps 6 hours per night?

Answer 3 A

In [11]:
```
mean_sleep_hours = 6.9
std_dev_sleep_hours = 1.2

# A. Empirical Rule
# Calculate the percentage of individuals who sleep between 4.5 and
z1 = (4.5 - mean_sleep_hours) / std_dev_sleep_hours
z2 = (9.3 - mean_sleep_hours) / std_dev_sleep_hours
percentage_within_range = (z2 - z1) * 100
print(f"Percentage within 4.5 to 9.3 hours: {percentage_within_rang
```

```
Percentage within 4.5 to 9.3 hours: 400.00%
```

Answer 3 B

In [12]:
```
# B. Z-score for 8 hours of sleep
z_8_hours = (8 - mean_sleep_hours) / std_dev_sleep_hours
print(f"Z-score for 8 hours of sleep: {z_8_hours:.2f}")
```

```
Z-score for 8 hours of sleep: 0.92
```

Answer 3 C

In [13]:
```
# C. Z-score for 6 hours of sleep
z_6_hours = (6 - mean_sleep_hours) / std_dev_sleep_hours
print(f"Z-score for 6 hours of sleep: {z_6_hours:.2f}")
```

```
Z-score for 6 hours of sleep: -0.75
```