

Temperatura X Vendas Walmart

```
In [ ]: df = read.csv("../Walmart_sales.csv")
```

```
In [ ]: #Conferir se já não foi alterado e converter coluna 'Date para formato Date
if ( class(df$Date) != "Date" ) {
  df$Date = as.Date(df$Date, format = "%d-%m-%Y")
}
```

```
In [ ]: summary(df)
```

	Store	Date	Weekly_Sales	Holiday_Flag
Min.	: 1	Min. :2010-02-05	Min. : 209986	Min. :0.00000
1st Qu.	:12	1st Qu.:2010-10-08	1st Qu.: 553350	1st Qu.:0.00000
Median	:23	Median :2011-06-17	Median : 960746	Median :0.00000
Mean	:23	Mean :2011-06-17	Mean :1046965	Mean :0.06993
3rd Qu.	:34	3rd Qu.:2012-02-24	3rd Qu.:1420159	3rd Qu.:0.00000
Max.	:45	Max. :2012-10-26	Max. :3818686	Max. :1.00000
	Temperature	Fuel_Price	CPI	Unemployment
Min.	: -2.06	Min. :2.472	Min. :126.1	Min. : 3.879
1st Qu.	: 47.46	1st Qu.:2.933	1st Qu.:131.7	1st Qu.: 6.891
Median	: 62.67	Median :3.445	Median :182.6	Median : 7.874
Mean	: 60.66	Mean :3.359	Mean :171.6	Mean : 7.999
3rd Qu.	: 74.94	3rd Qu.:3.735	3rd Qu.:212.7	3rd Qu.: 8.622
Max.	:100.14	Max. :4.468	Max. :227.2	Max. :14.313

```
In [ ]: Gráfico de dispersão de Vendas Semanais vs. Temperatura:
```

Os pontos em azul representam todas as vendas.

Os pontos em vermelho destacam vendas ocorrendo em temperaturas abaixo de 32°F (congelamento).

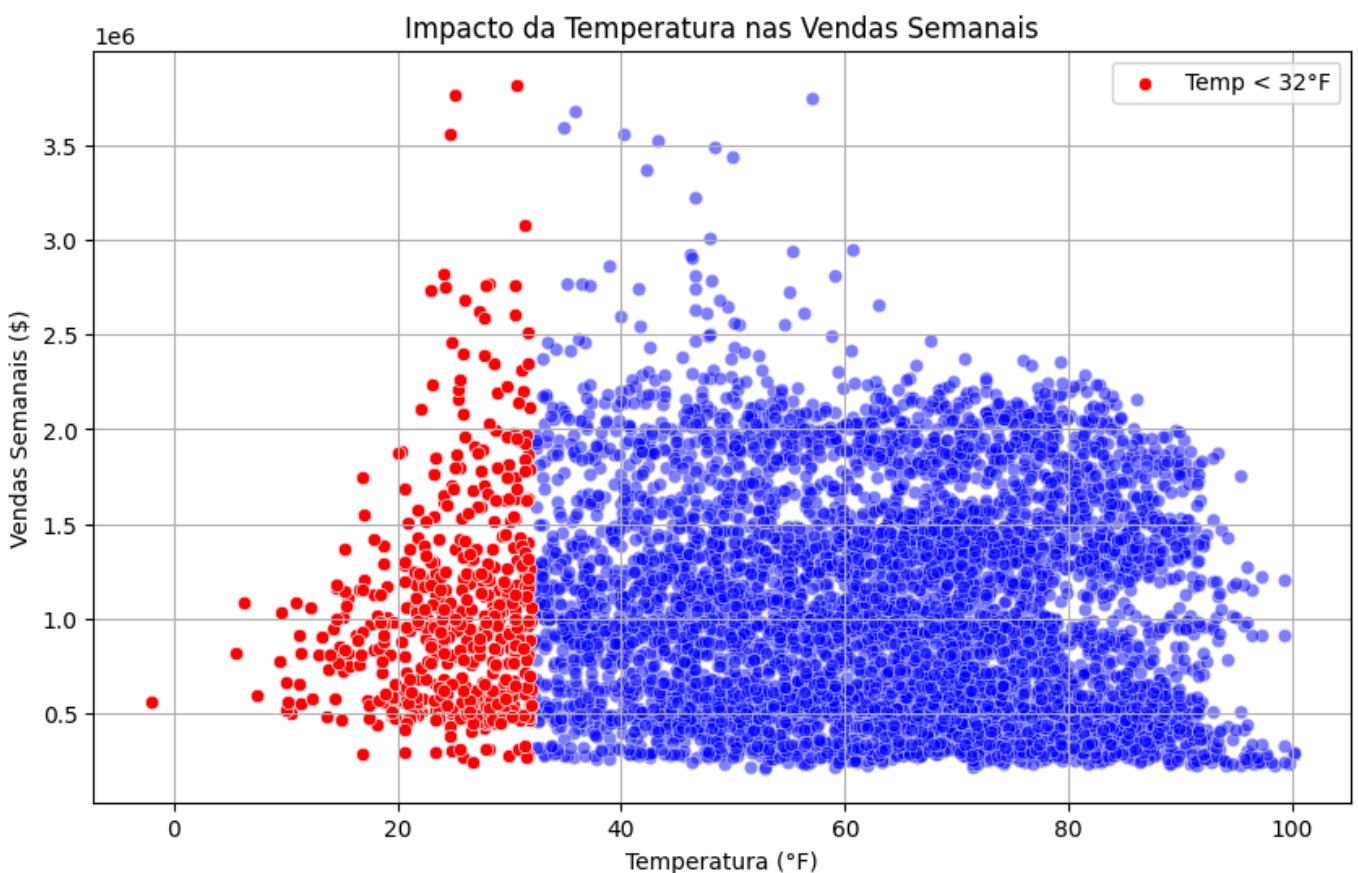
```
In [17]: import matplotlib.pyplot as plt
import seaborn as sns

# Converter a coluna de data para o formato correto
df['Date'] = pd.to_datetime(df['Date'], format='%d-%m-%Y')

# Criar o gráfico de dispersão Vendas x Temperatura
plt.figure(figsize=(10, 6))
sns.scatterplot(x=df['Temperature'], y=df['Weekly_Sales'], alpha=0.5, color='blue')

# Destacar temperaturas abaixo de 32°F
extreme_cold = df[df['Temperature'] < 32]
sns.scatterplot(x=extreme_cold['Temperature'], y=extreme_cold['Weekly_Sales'], color='red', label='Congelamento')

plt.xlabel('Temperatura (°F)')
plt.ylabel('Vendas Semanais ($)')
plt.title('Impacto da Temperatura nas Vendas Semanais')
plt.legend()
plt.grid(True)
plt.show()
```



Comentário: Nas lojas onde a temperatura cai muito abaixo do ponto de congelamento (32°F), o impacto nas vendas é mais evidente. Isso se reflete especialmente no primeiro quartil do boxplot (rabicho inferior), indicando que temperaturas extremamente baixas estão associadas a quedas significativas nas vendas.

```
In [45]: import seaborn as sns
import matplotlib.pyplot as plt
import warnings

# Ignorar avisos futuros (opcional)
warnings.simplefilter(action='ignore', category=FutureWarning)

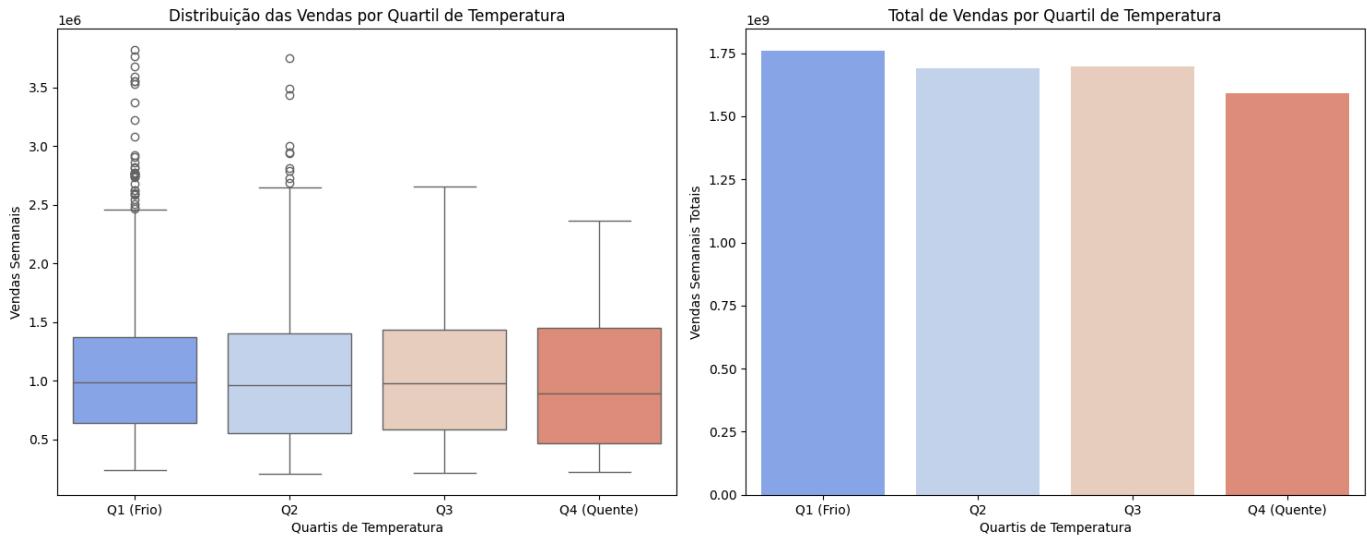
# Definir quartis da temperatura para categorização
df["Temp_Quartile"] = pd.qcut(df["Temperature"], q=4, labels=["Q1 (Frio)", "Q2", "Q3", "Q4 (Quente)"])

# Criar figuras para os gráficos
fig, axes = plt.subplots(1, 2, figsize=(15, 6))

# Boxplot das vendas por quartil de temperatura
sns.boxplot(x="Temp_Quartile", y="Weekly_Sales", data=df, ax=axes[0], hue="Temp_Quartile", palette="coolwarm")
axes[0].set_title("Distribuição das Vendas por Quartil de Temperatura")
axes[0].set_xlabel("Quartis de Temperatura")
axes[0].set_ylabel("Vendas Semanais")

# Barplot somando as vendas por quartil de temperatura
sales_by_temp = df.groupby("Temp_Quartile", observed=False)[["Weekly_Sales"]].sum().reset_index()
sns.barplot(x="Temp_Quartile", y="Weekly_Sales", data=sales_by_temp, ax=axes[1], palette="coolwarm")
axes[1].set_title("Total de Vendas por Quartil de Temperatura")
axes[1].set_xlabel("Quartis de Temperatura")
axes[1].set_ylabel("Vendas Semanais Totais")

# Melhorar o Layout
plt.tight_layout()
plt.show()
```



In [46]:

```

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

# Ler os dados
data_path = 'Walmart_Sales.csv'
df = pd.read_csv(data_path)

# Certificar-se de que a coluna 'Date' está em formato de data
df['Date'] = pd.to_datetime(df['Date'], format='%d-%m-%Y')

# Função para filtrar outliers
def filter_outliers(df, nstore):
    Q1 = df[df['Store'] == nstore]['Weekly_Sales'].quantile(0.25)
    Q3 = df[df['Store'] == nstore]['Weekly_Sales'].quantile(0.75)
    IQR = Q3 - Q1
    lower_bound = Q1 - 1.5 * IQR
    upper_bound = Q3 + 1.5 * IQR
    return df[(df['Store'] == nstore) &
               (df['Weekly_Sales'] > lower_bound) &
               (df['Weekly_Sales'] < upper_bound)]

# Função para criar os gráficos com a paleta 'coolwarm'
def create_plots(nstore):
    # Filtrar dados sem outliers e criar uma cópia
    dfplot = filter_outliers(df, nstore).copy()

    # Quartis da Temperatura
    Q0_TEMP = dfplot['Temperature'].min()
    Q1_TEMP = dfplot['Temperature'].quantile(0.25)
    Q2_TEMP = dfplot['Temperature'].quantile(0.50)
    Q3_TEMP = dfplot['Temperature'].quantile(0.75)
    Q4_TEMP = dfplot['Temperature'].max()

    # Labels para os quartis de temperatura
    labels = [f"A: {round(Q0_TEMP, 2)} - {round(Q1_TEMP, 2)}",
              f"B: {round(Q1_TEMP, 2)} - {round(Q2_TEMP, 2)}",
              f"C: {round(Q2_TEMP, 2)} - {round(Q3_TEMP, 2)}",
              f"D: {round(Q3_TEMP, 2)} - {round(Q4_TEMP, 2)}"]

    # Atribuir quartis de temperatura
    dfplot.loc[:, 'Temp_Quantile'] = pd.cut(dfplot['Temperature'],
                                              bins=[Q0_TEMP, Q1_TEMP, Q2_TEMP, Q3_TEMP, Q4_TEMP],
                                              labels=labels,
                                              include_lowest=True)

    # Agregar vendas semanais por quartil de temperatura
    sums_temp = dfplot.groupby('Temp_Quantile', observed=False)[['Weekly_Sales']].sum()

```

```

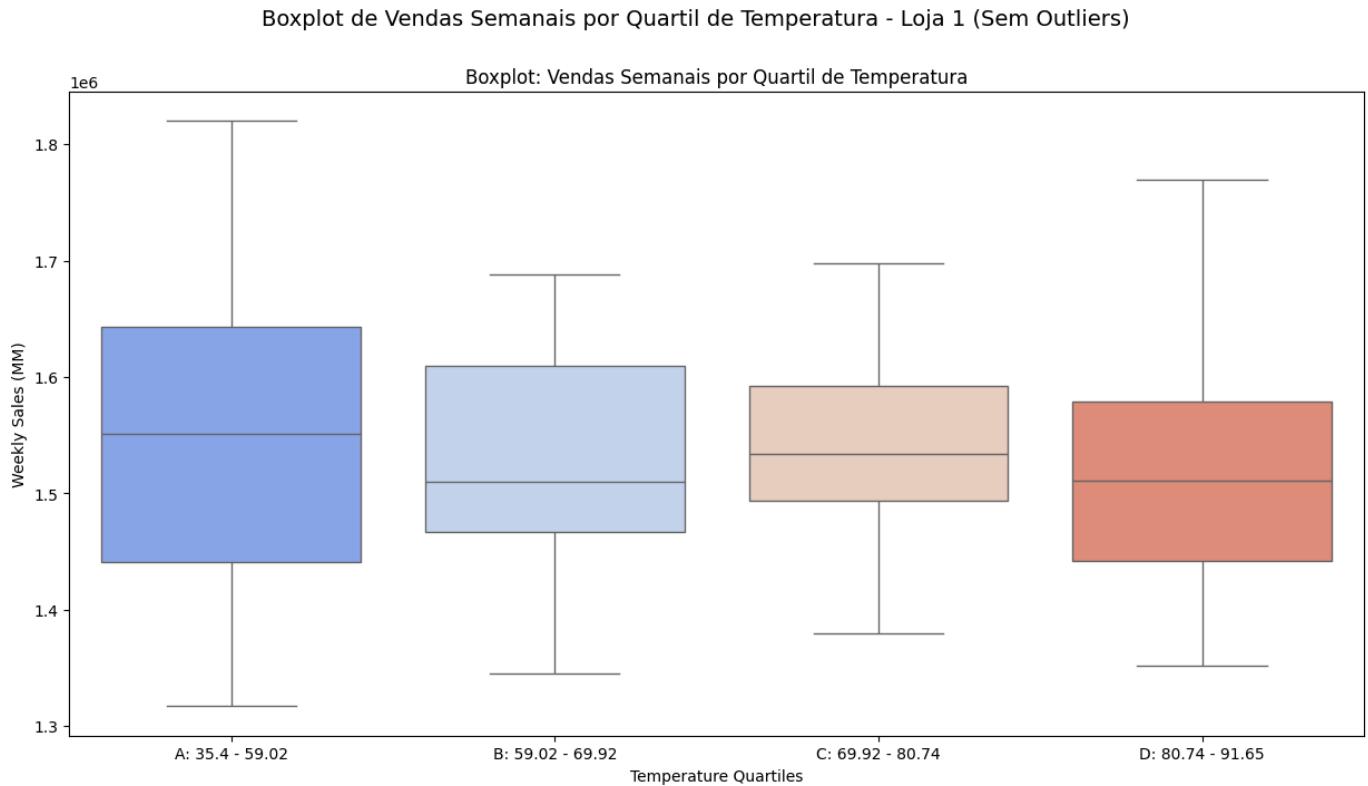
# Criar o gráfico 1 (Boxplot) com paleta de cores
plt.figure(figsize=(15, 7.5))
sns.boxplot(x='Temp_Quantile', y='Weekly_Sales', data=dfplot, palette='coolwarm', legend=False)
plt.title('Boxplot: Vendas Semanais por Quartil de Temperatura')
plt.ylabel('Weekly Sales (MM)')
plt.xlabel('Temperature Quartiles')
plt.suptitle(f'Boxplot de Vendas Semanais por Quartil de Temperatura - Loja {nstore} (Sem Outliers)')
plt.show()

# Criar o gráfico 2 (Barplot) com paleta de cores
plt.figure(figsize=(15, 7.5))
sns.barplot(x=sums_temp.index.astype(str), y=sums_temp / 1e6, palette='coolwarm', legend=False)
plt.title('Barplot: Vendas Totais por Quartil de Temperatura')
plt.ylabel('Weekly Sales (MM)')
plt.xlabel('Temperature Quartiles')
plt.suptitle(f'Vendas Totais por Quartil de Temperatura - Loja {nstore} (Sem Outliers)', fontweight='bold')
plt.show()

# Criar o gráfico de Vendas vs Temperatura (Dispersão) com cores
plt.figure(figsize=(15, 7.5))
plt.scatter(dfplot['Temperature'], dfplot['Weekly_Sales'],
            c=dfplot['Temperature'], cmap='coolwarm', alpha=0.7)
plt.colorbar(label="Temperature (°F)")
plt.title(f'Vendas Semanais vs Temperatura - Loja {nstore}')
plt.xlabel('Temperature (°F)')
plt.ylabel('Weekly Sales (MM)')
plt.suptitle(f'Gráfico de Vendas Semanais vs Temperatura - Loja {nstore} (Sem Outliers)', fontweight='bold')
plt.show()

# Exemplo de uso para todas as lojas de 1 a 45
for nstore in range(1, 46):
    create_plots(nstore)

```



Vendas Totais por Quartil de Temperatura - Loja 1 (Sem Outliers)

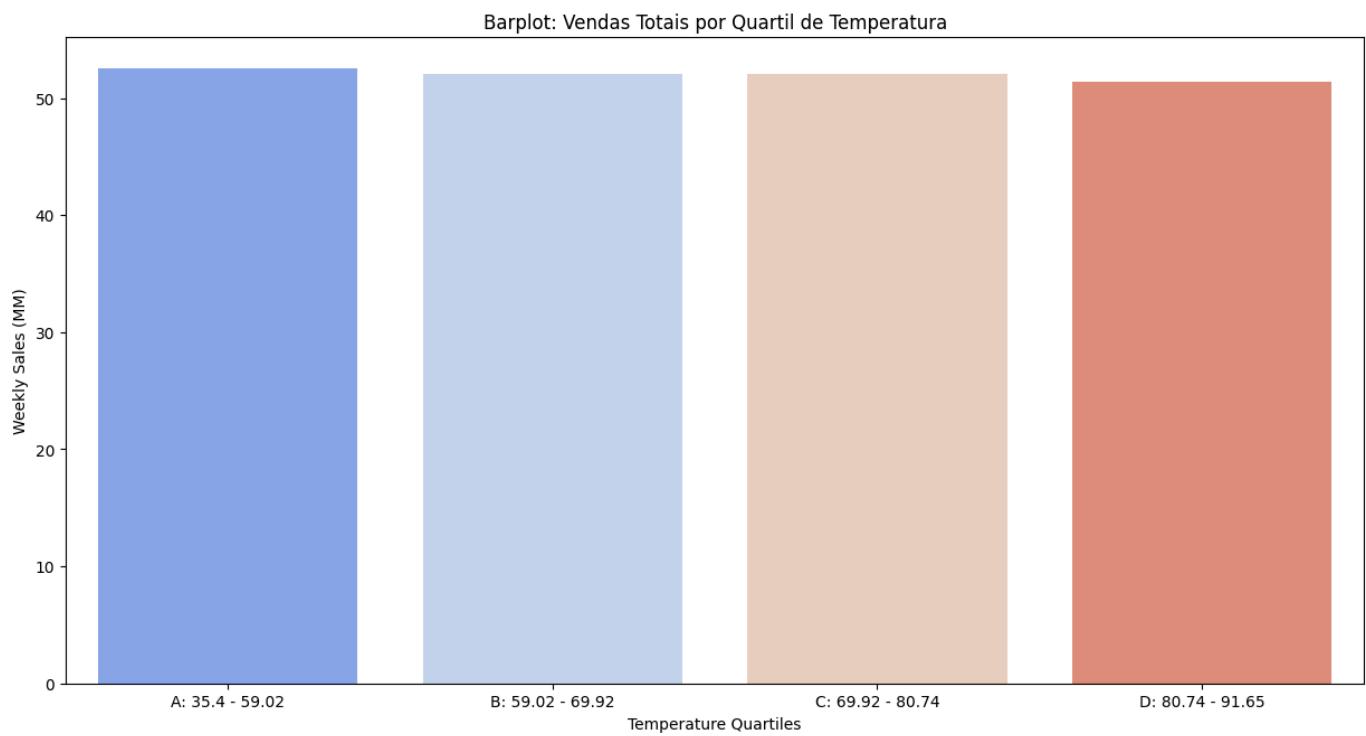
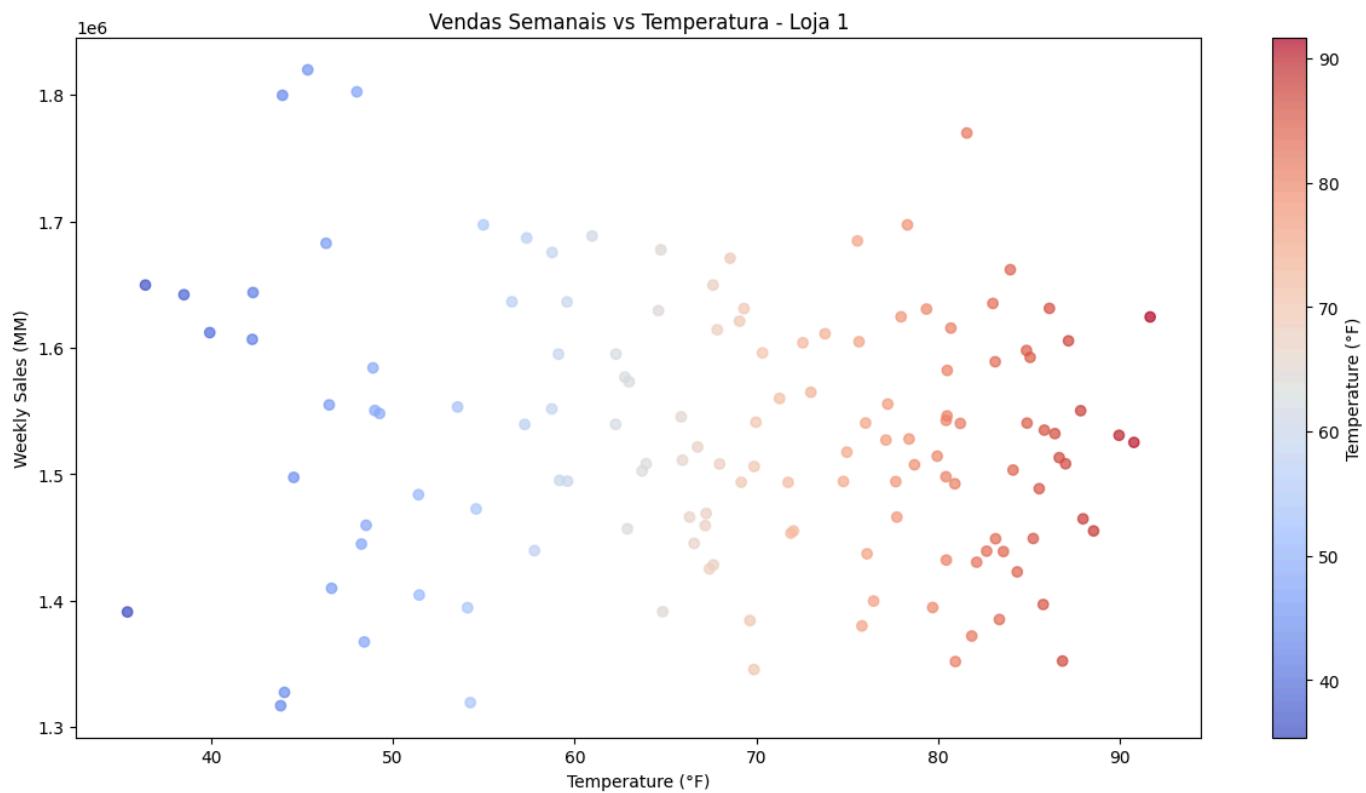
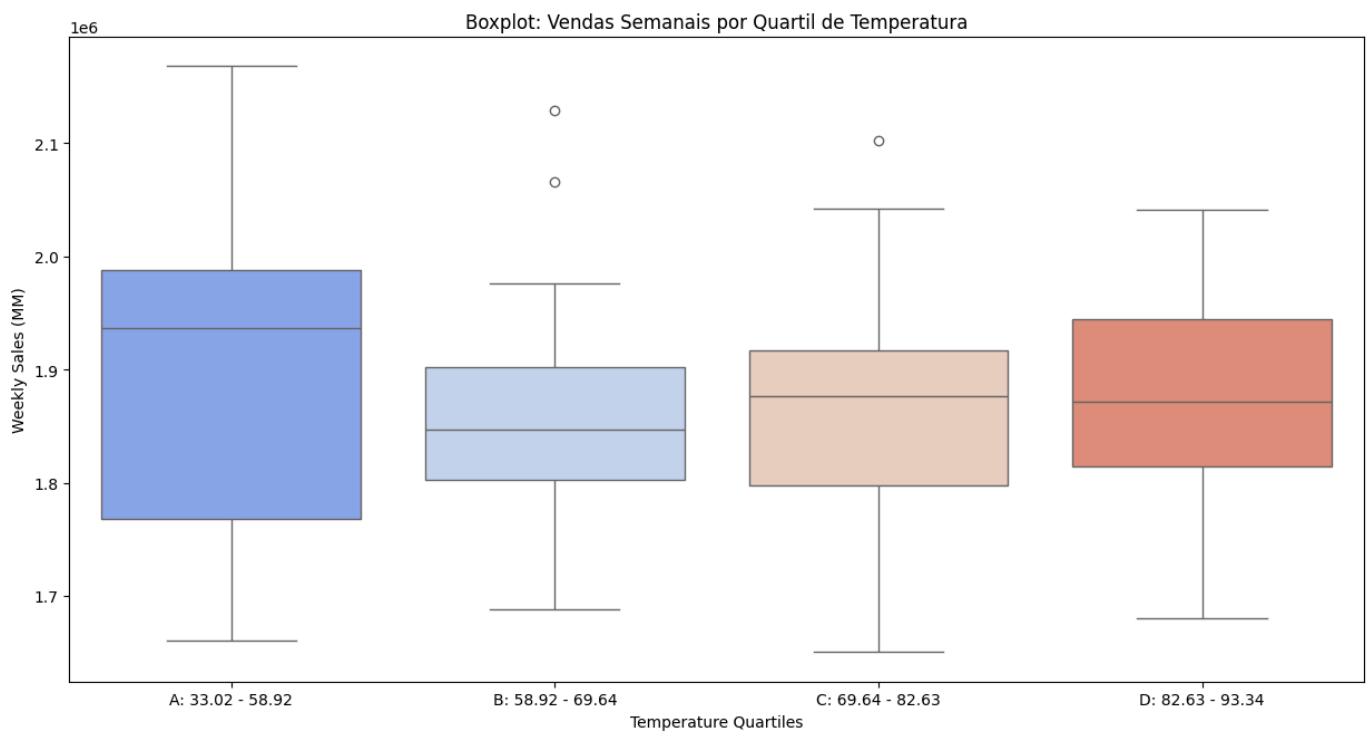


Gráfico de Vendas Semanais vs Temperatura - Loja 1 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 2 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 2 (Sem Outliers)

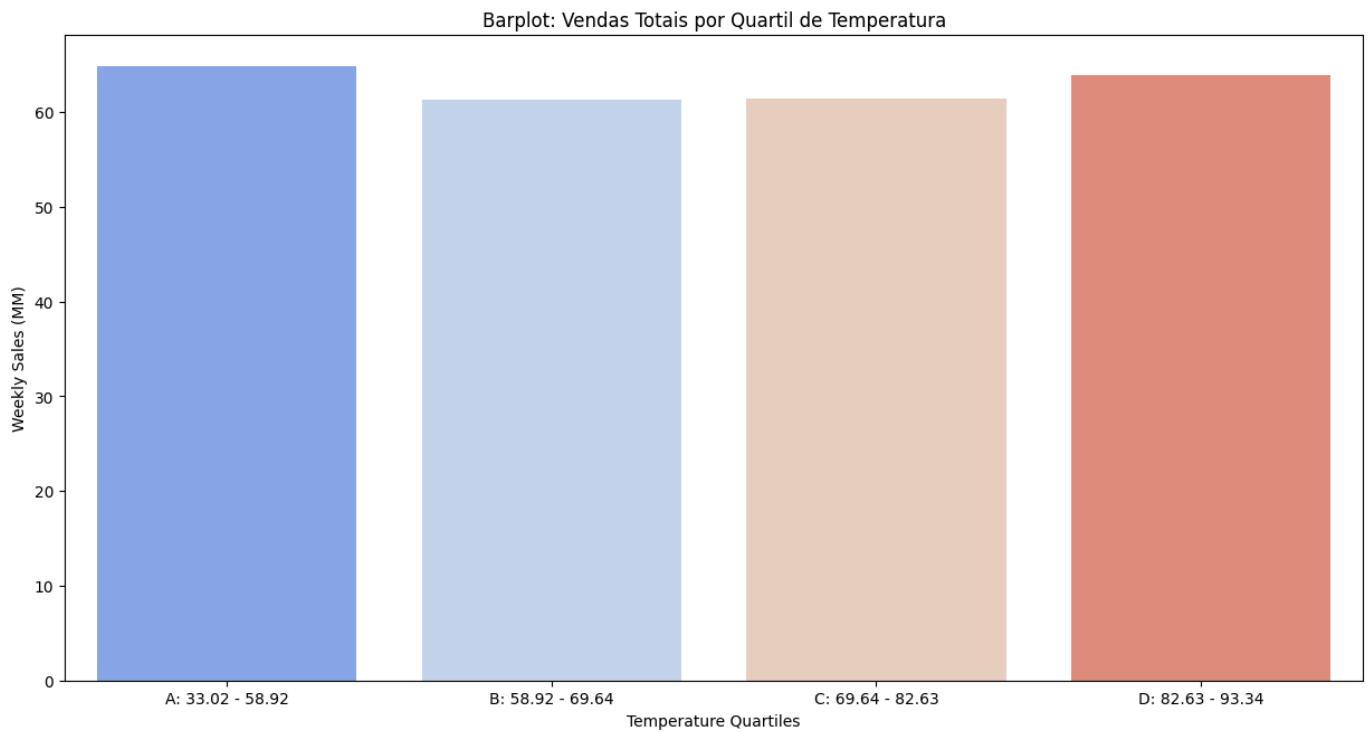
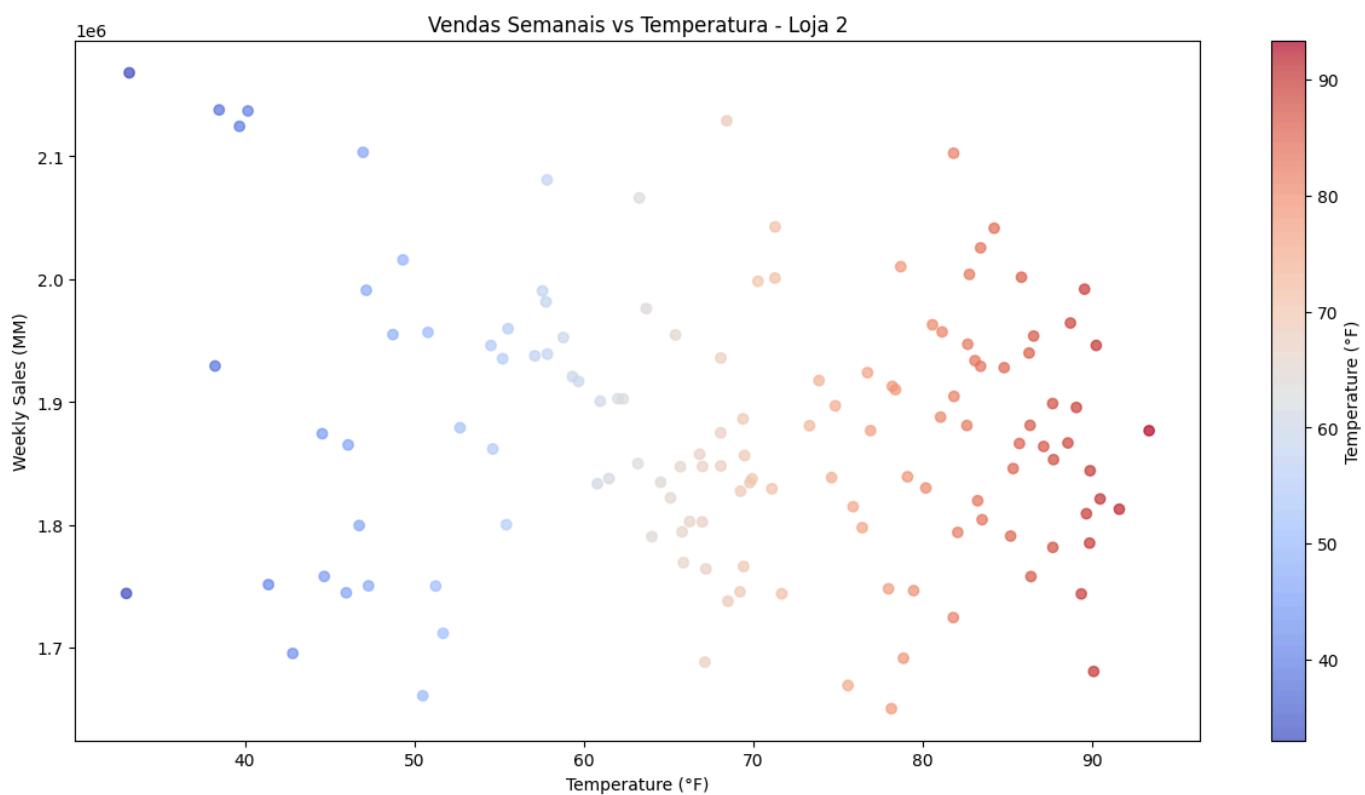
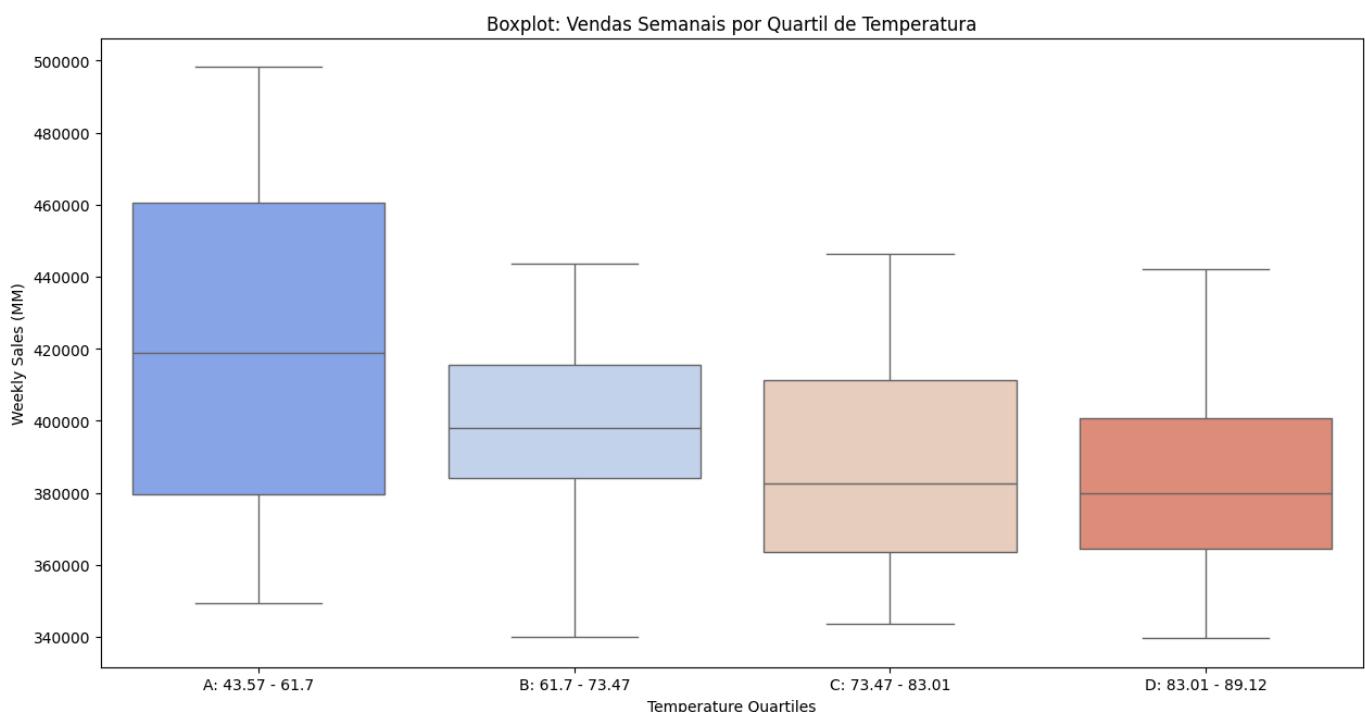


Gráfico de Vendas Semanais vs Temperatura - Loja 2 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 3 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 3 (Sem Outliers)

Barplot: Vendas Totais por Quartil de Temperatura

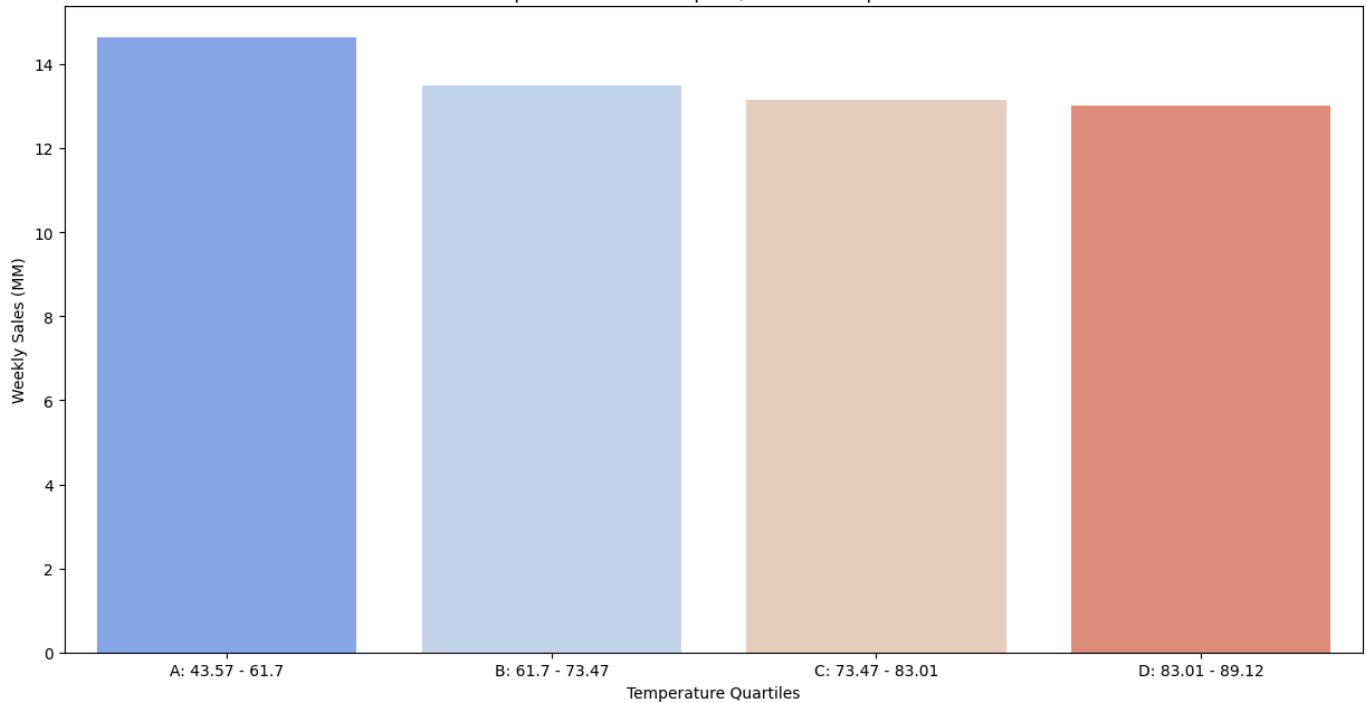
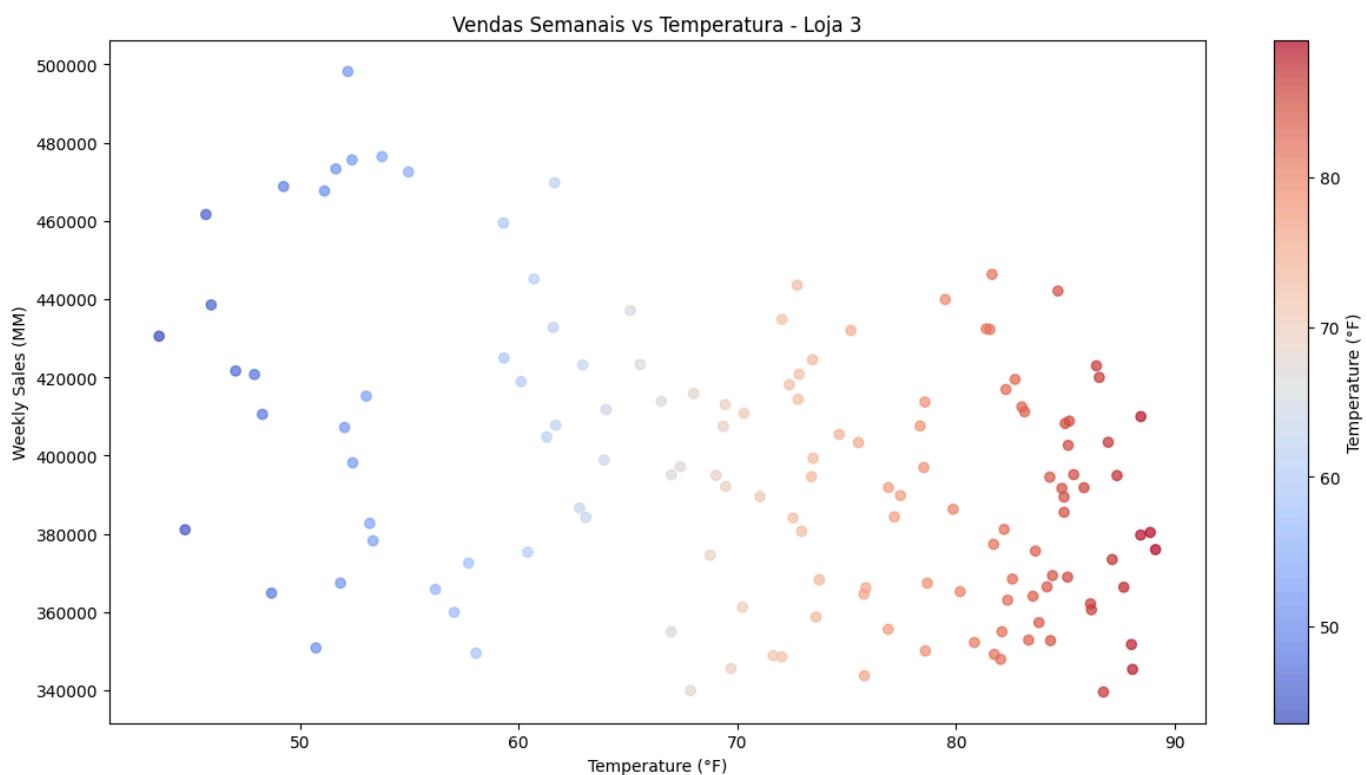
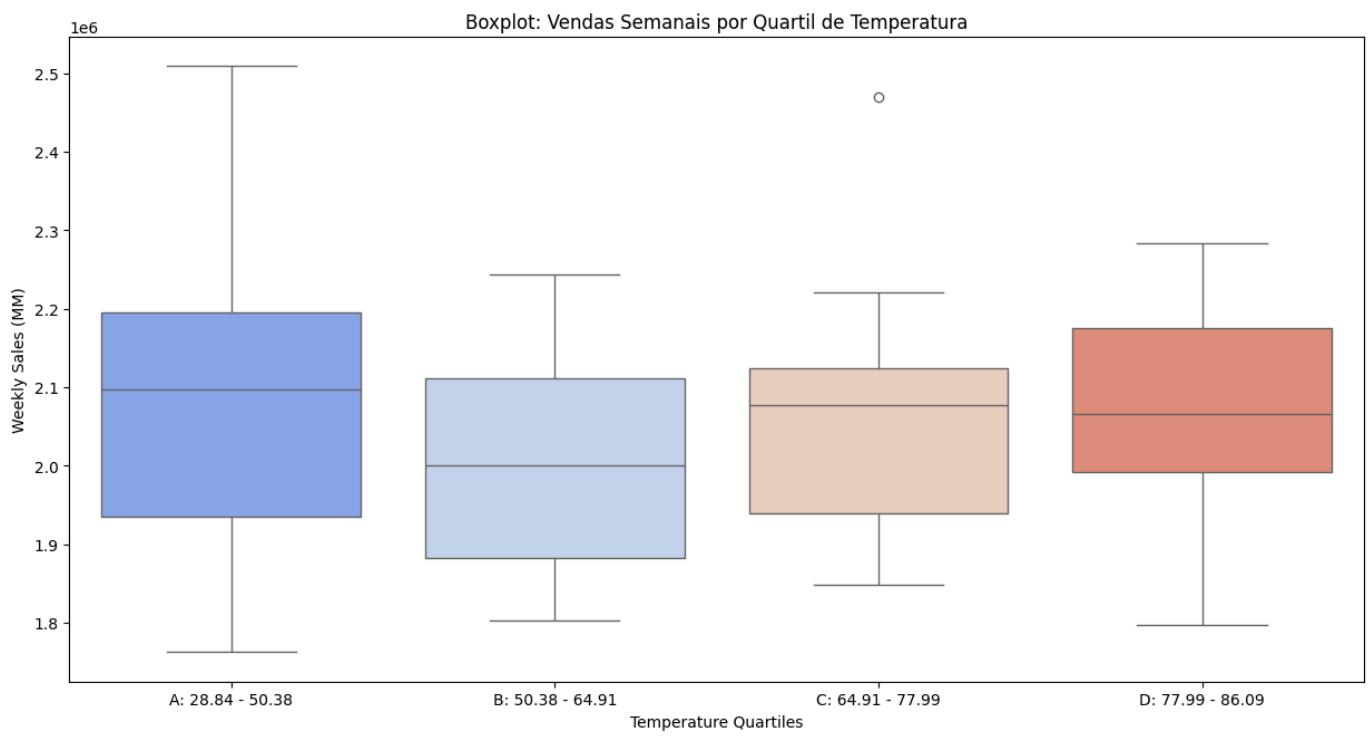


Gráfico de Vendas Semanais vs Temperatura - Loja 3 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 4 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 4 (Sem Outliers)

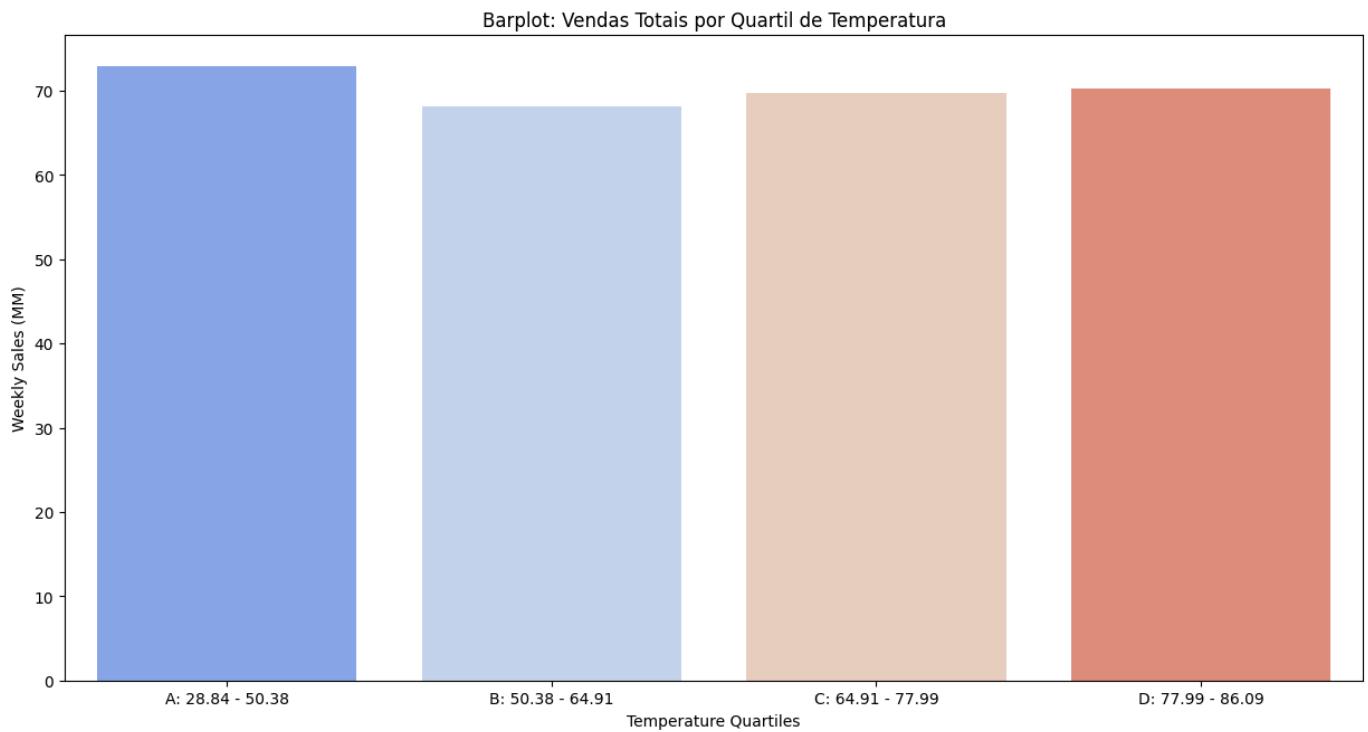
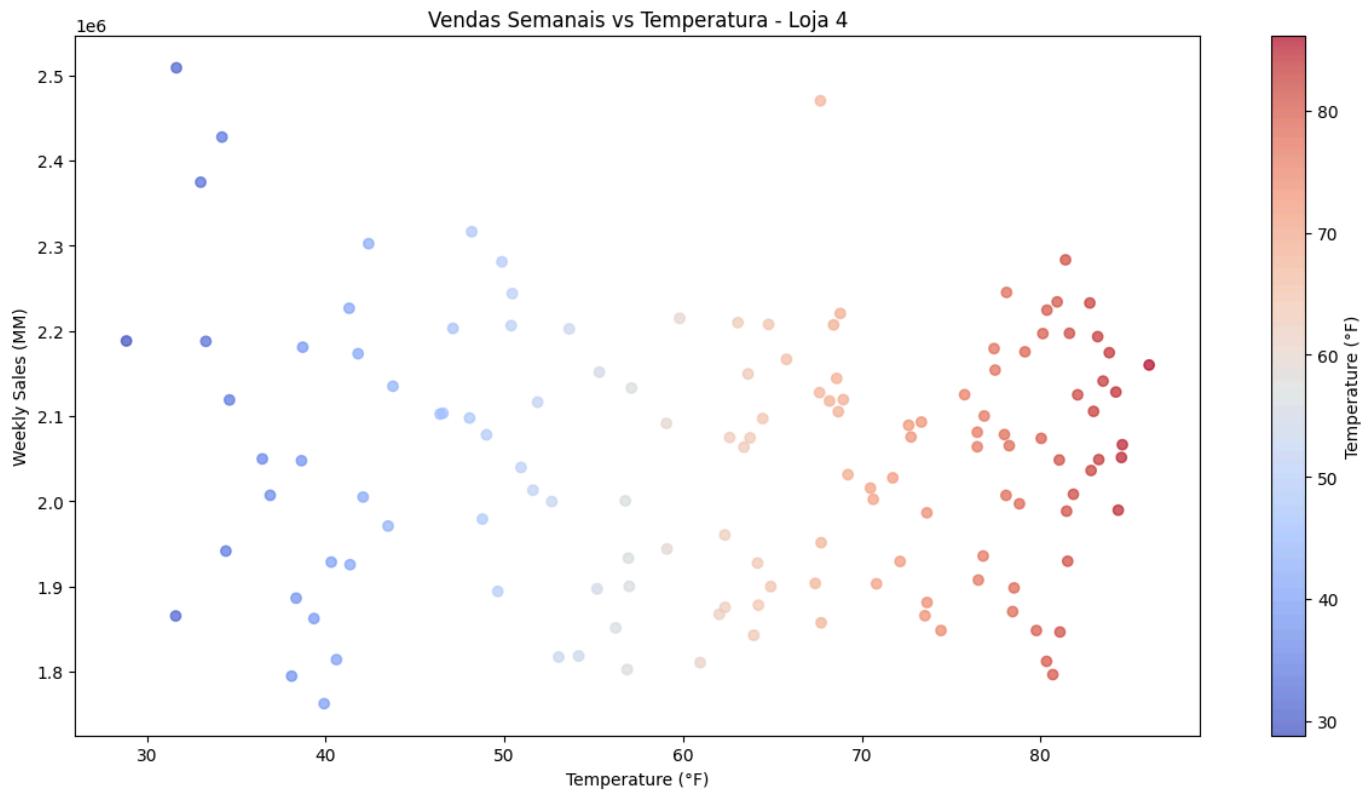
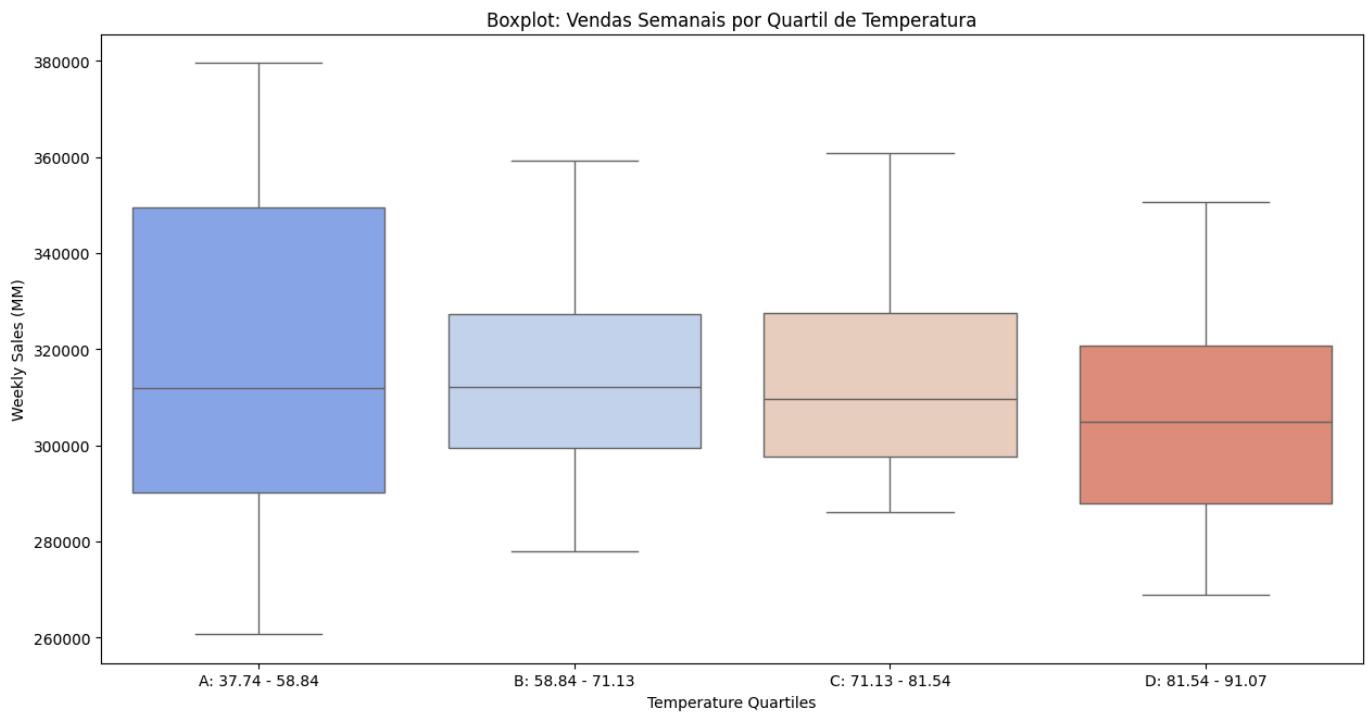


Gráfico de Vendas Semanais vs Temperatura - Loja 4 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 5 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 5 (Sem Outliers)

Barplot: Vendas Totais por Quartil de Temperatura

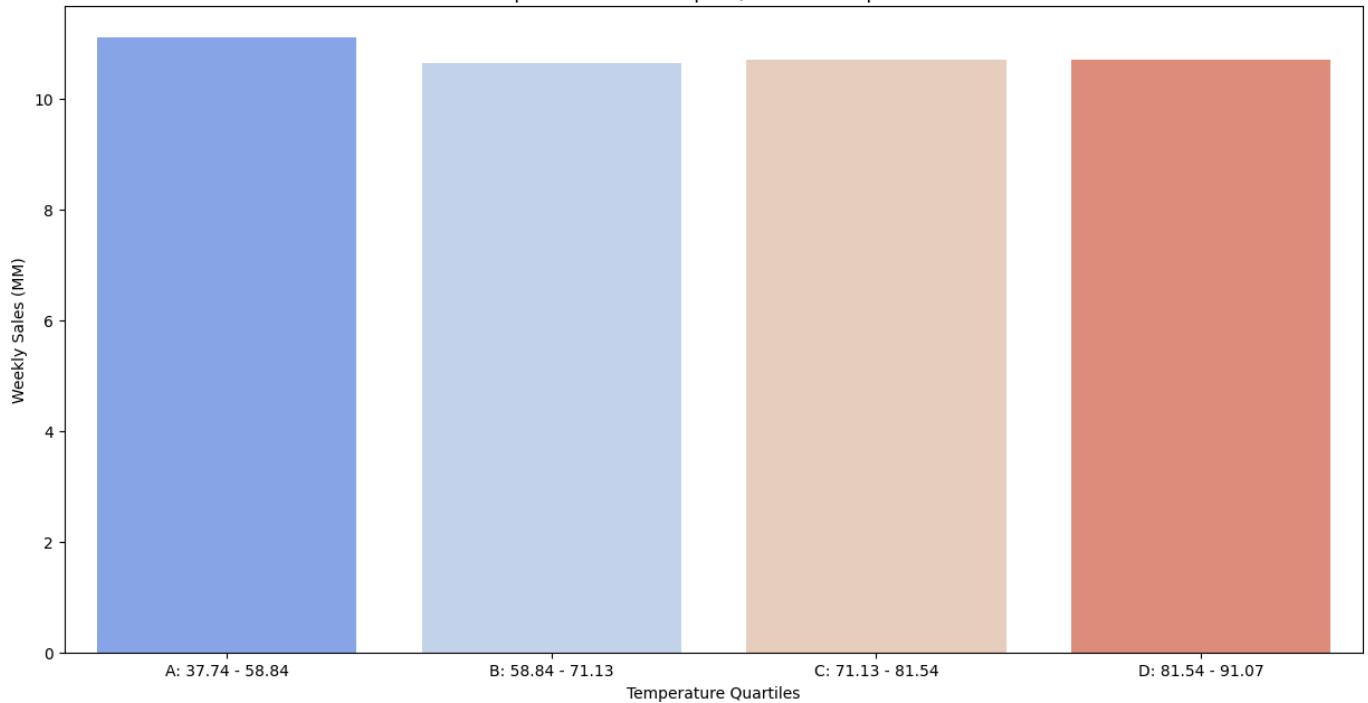
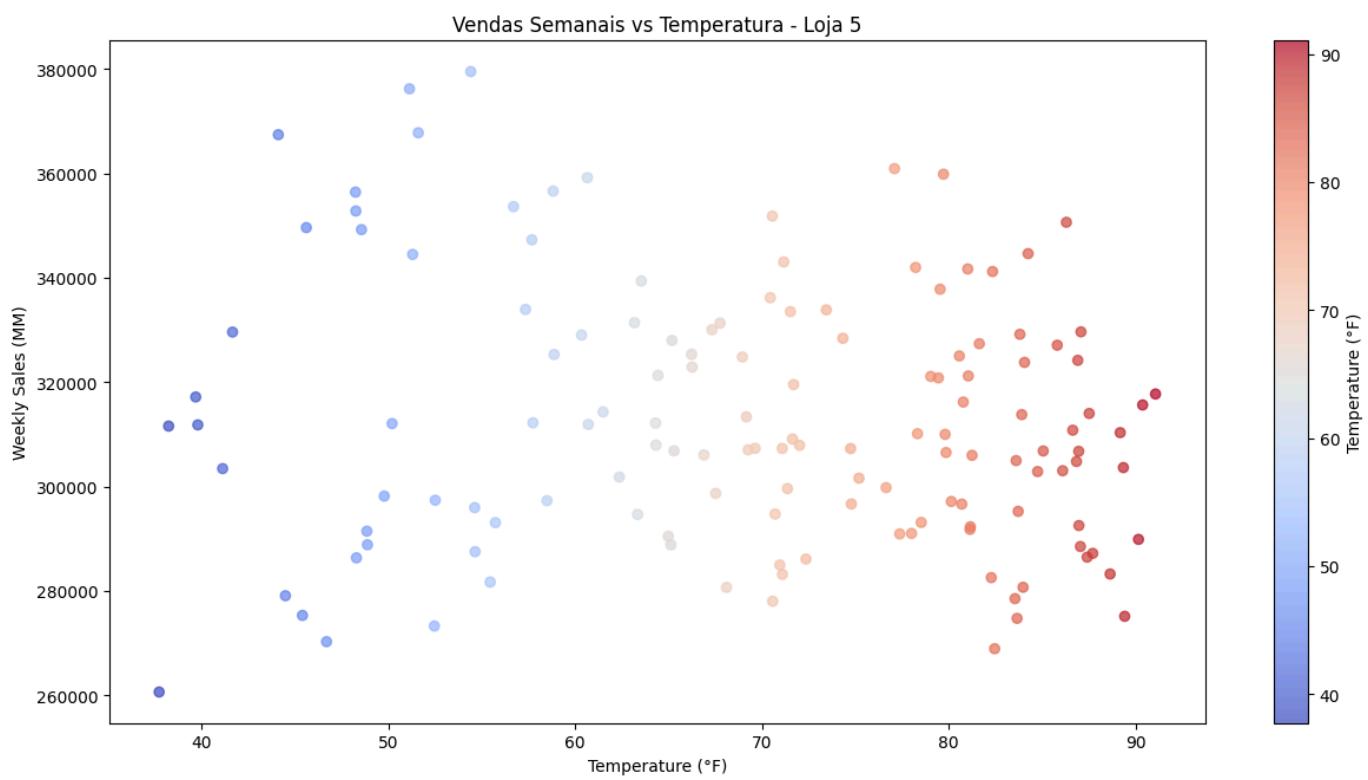
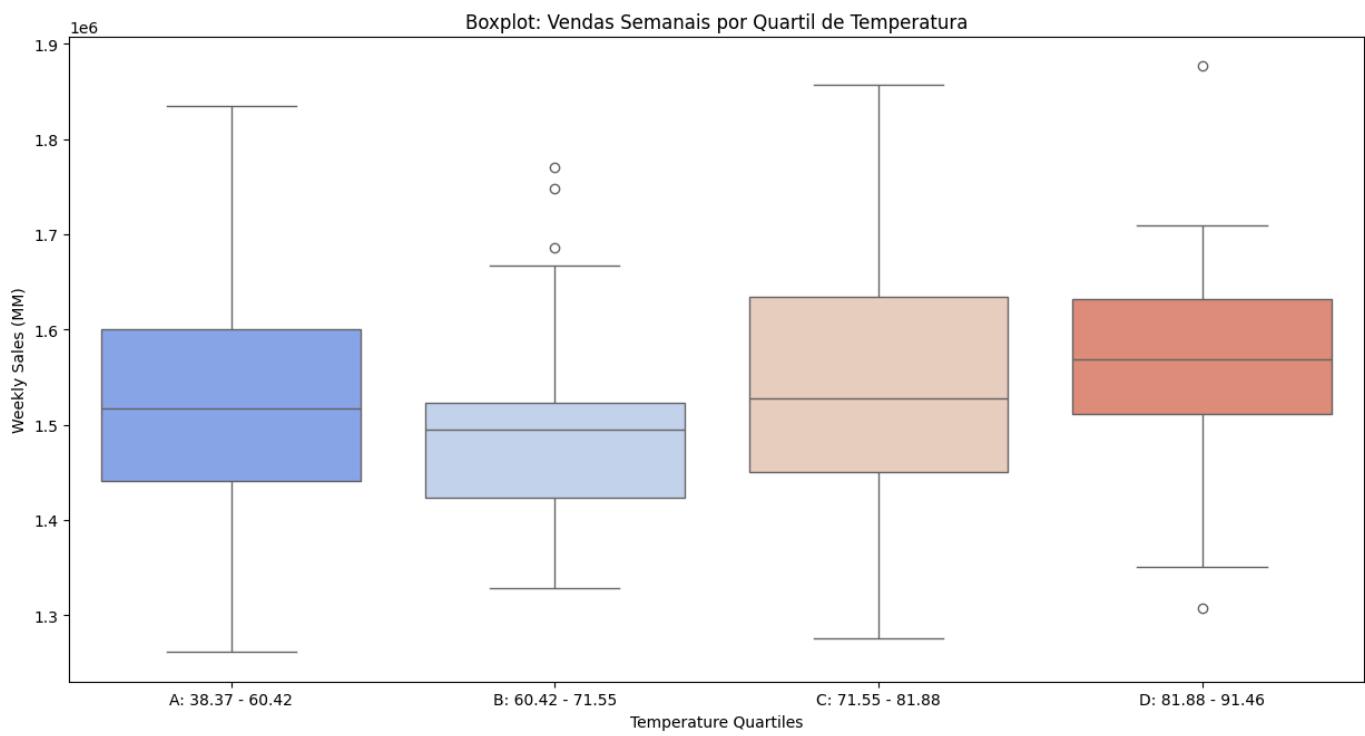


Gráfico de Vendas Semanais vs Temperatura - Loja 5 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 6 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 6 (Sem Outliers)

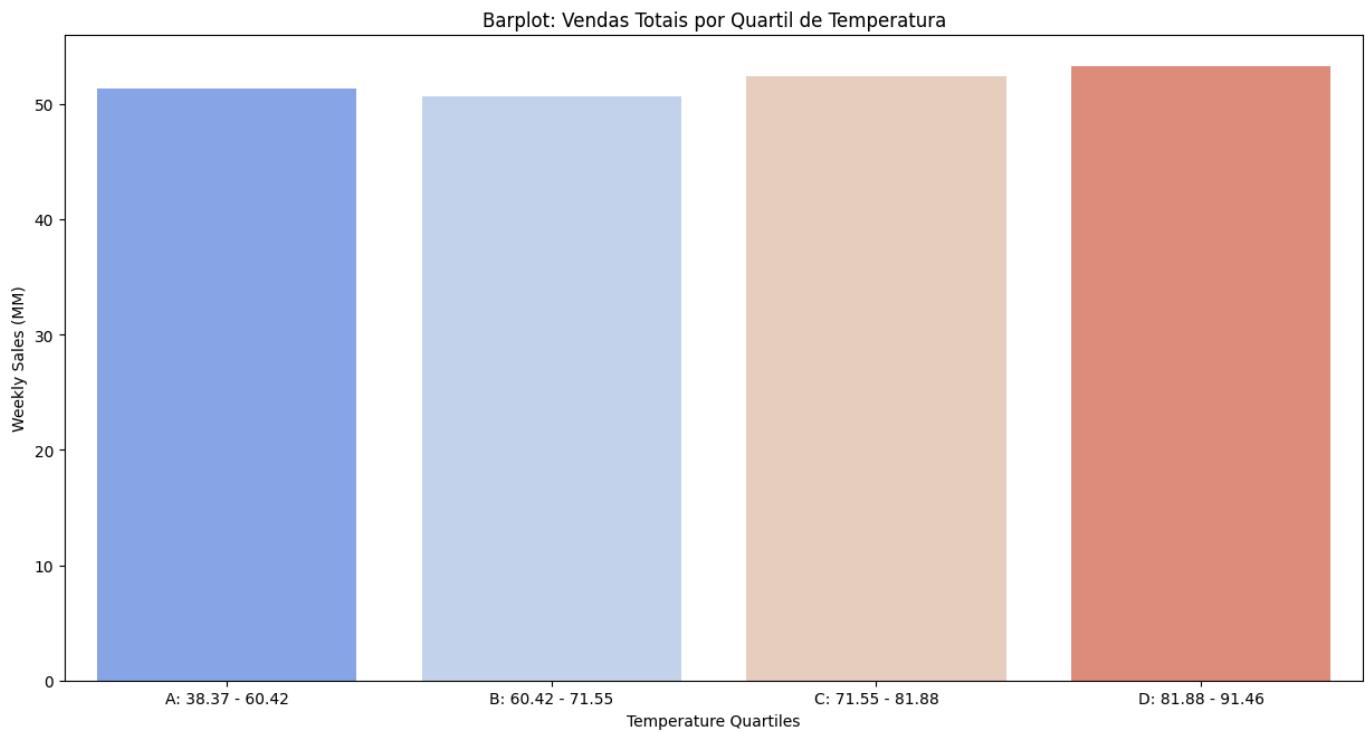
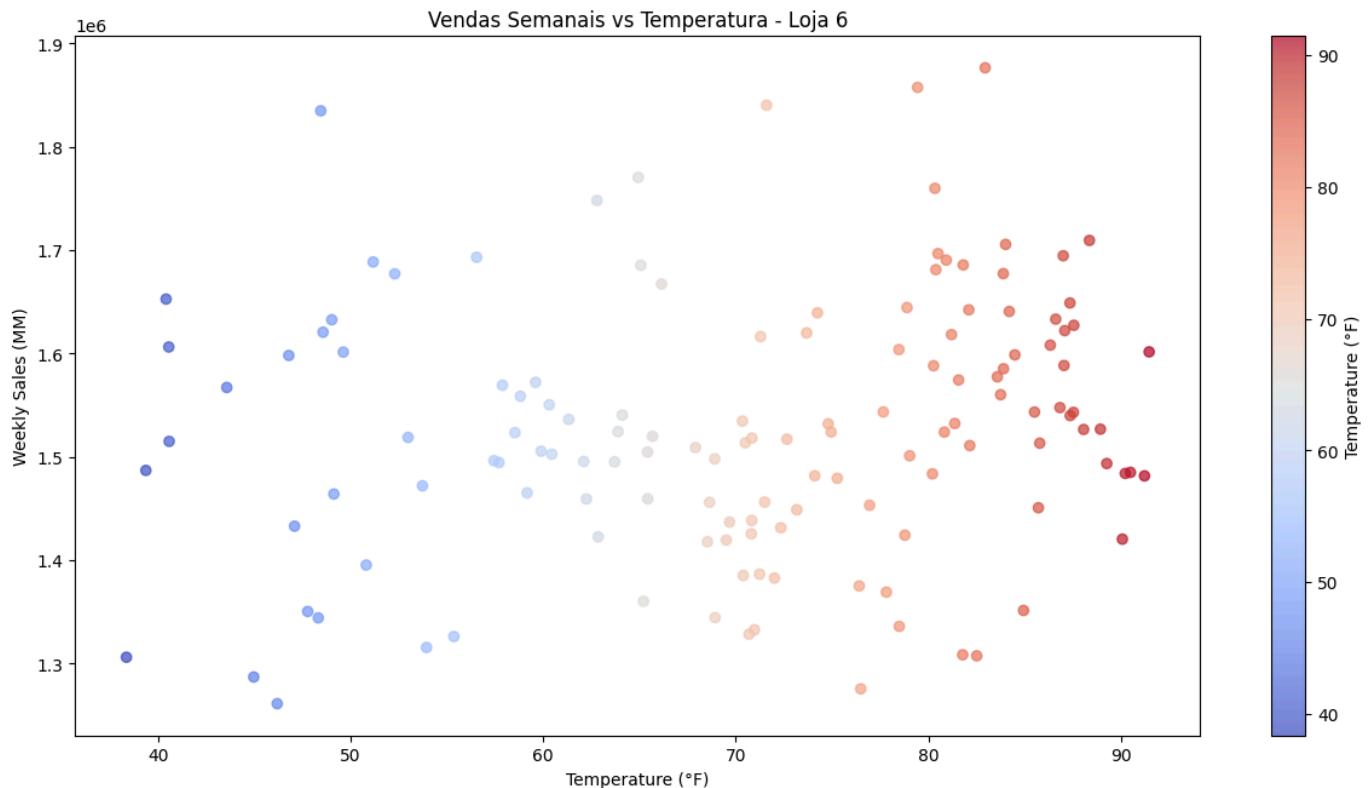
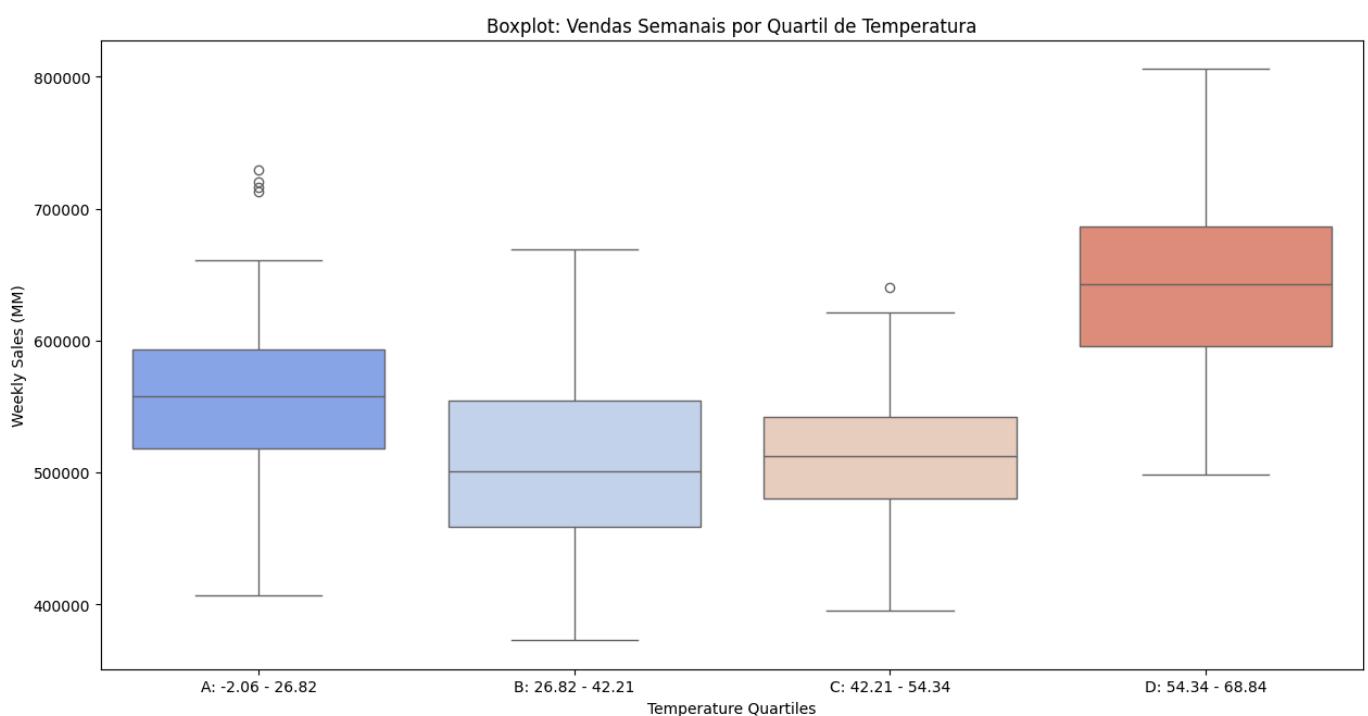


Gráfico de Vendas Semanais vs Temperatura - Loja 6 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 7 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 7 (Sem Outliers)

Barplot: Vendas Totais por Quartil de Temperatura

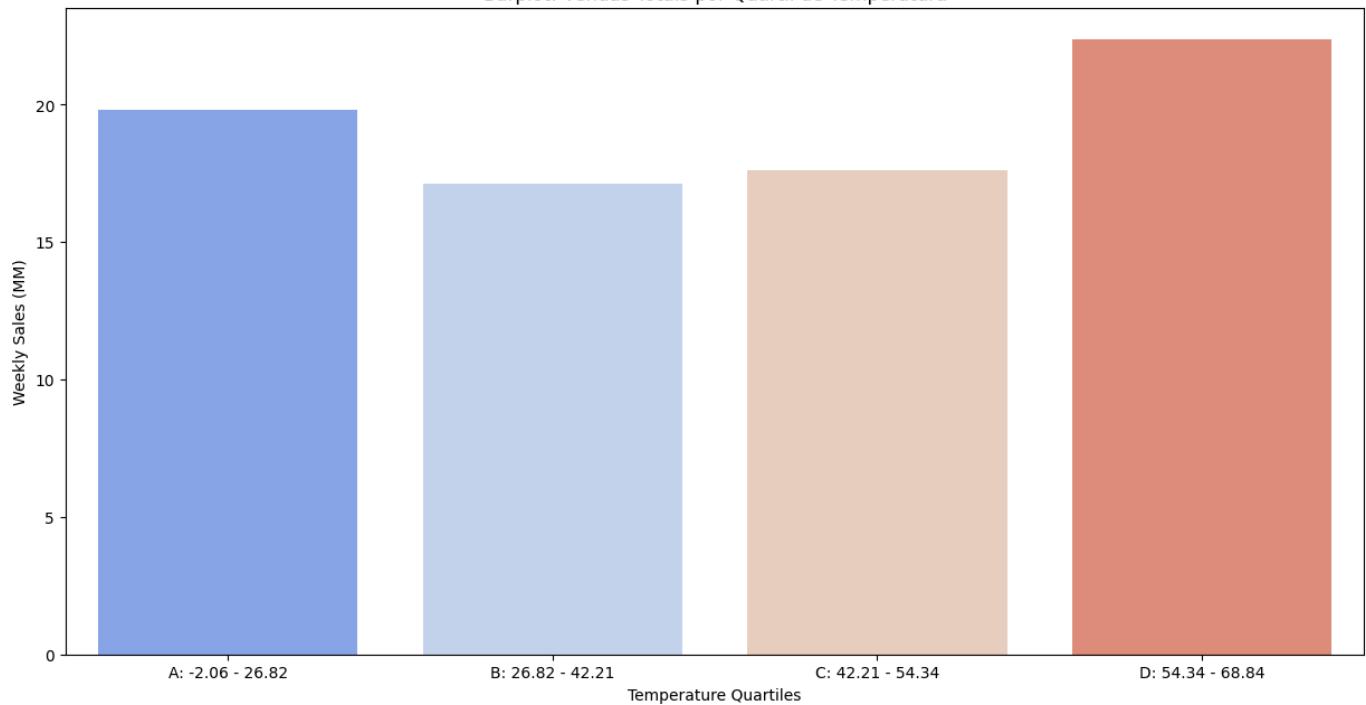
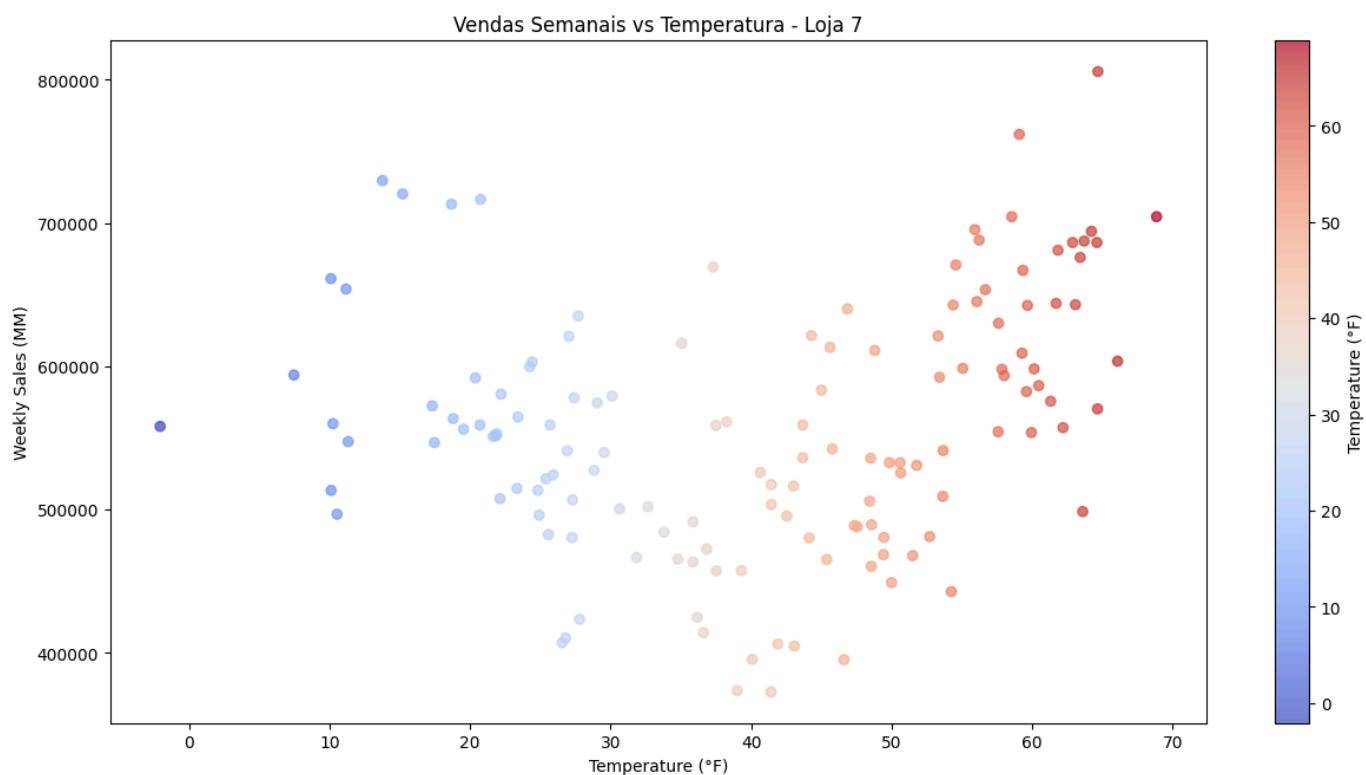
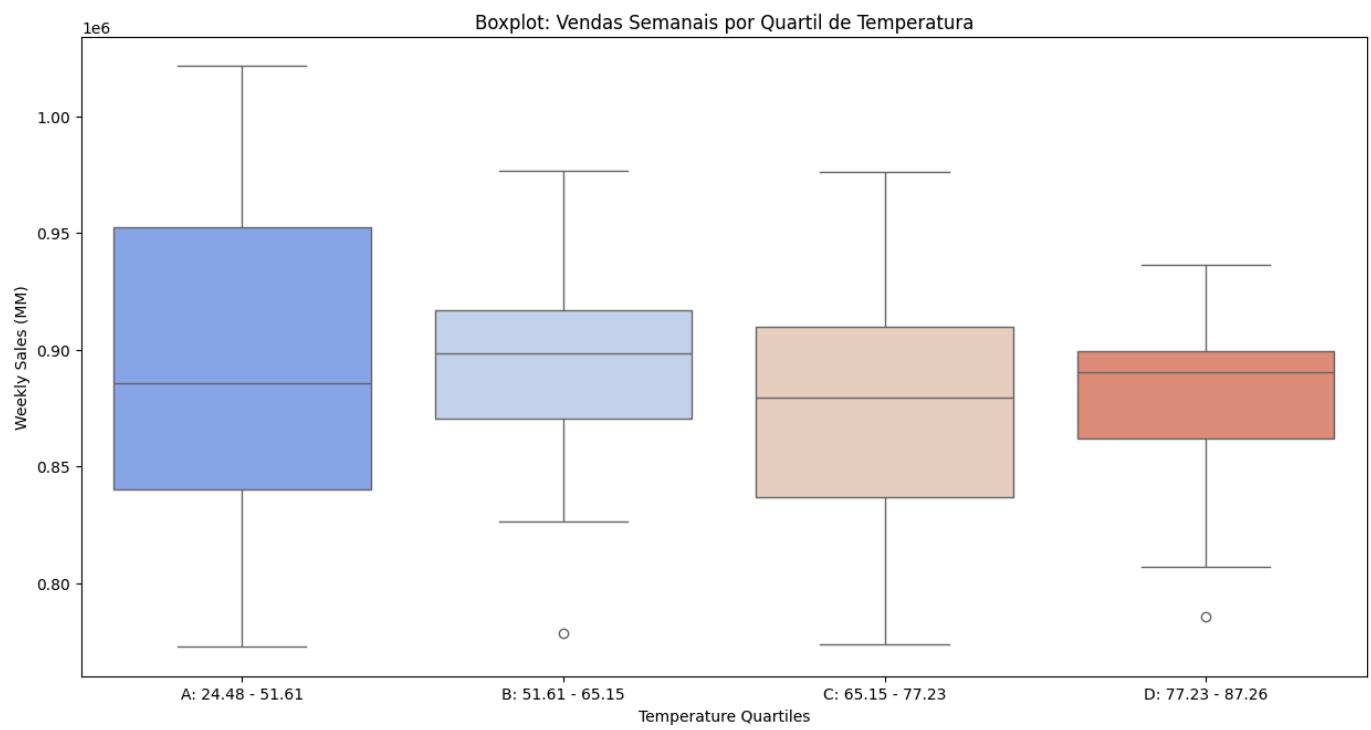


Gráfico de Vendas Semanais vs Temperatura - Loja 7 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 8 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 8 (Sem Outliers)

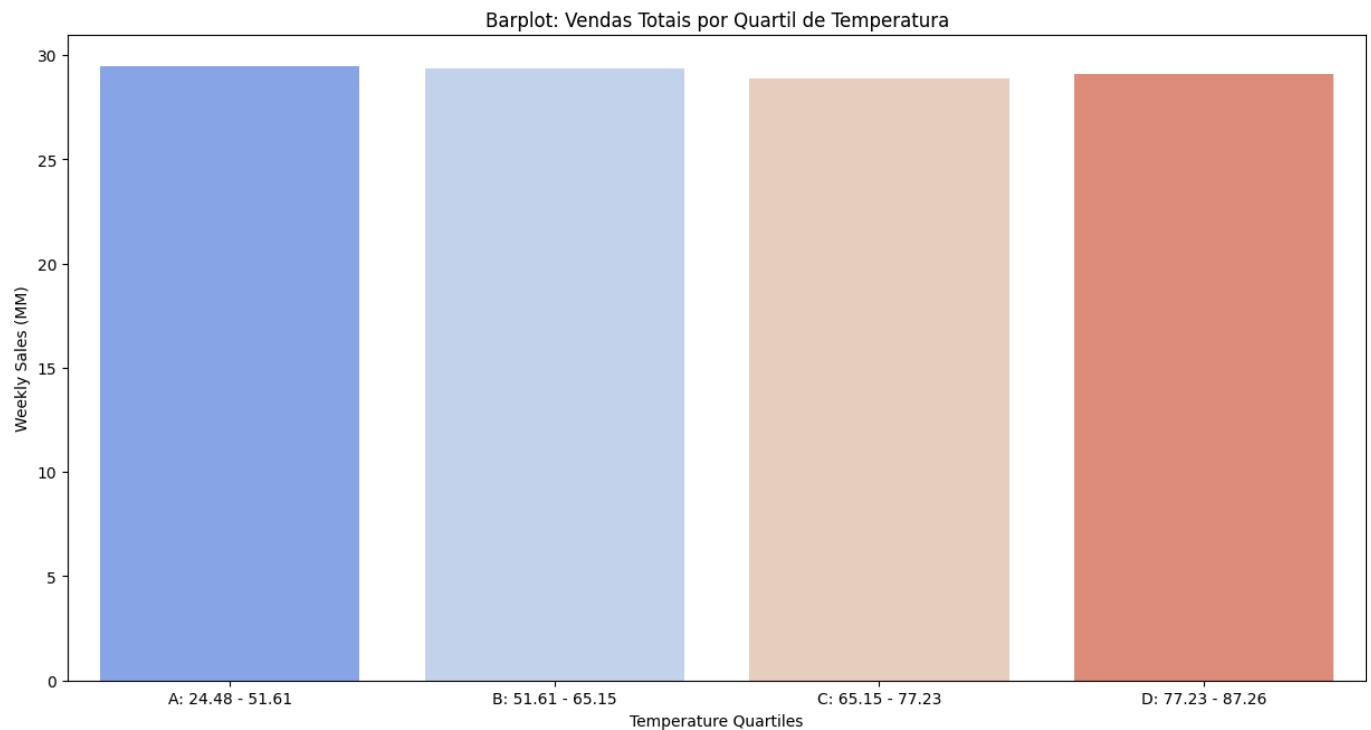
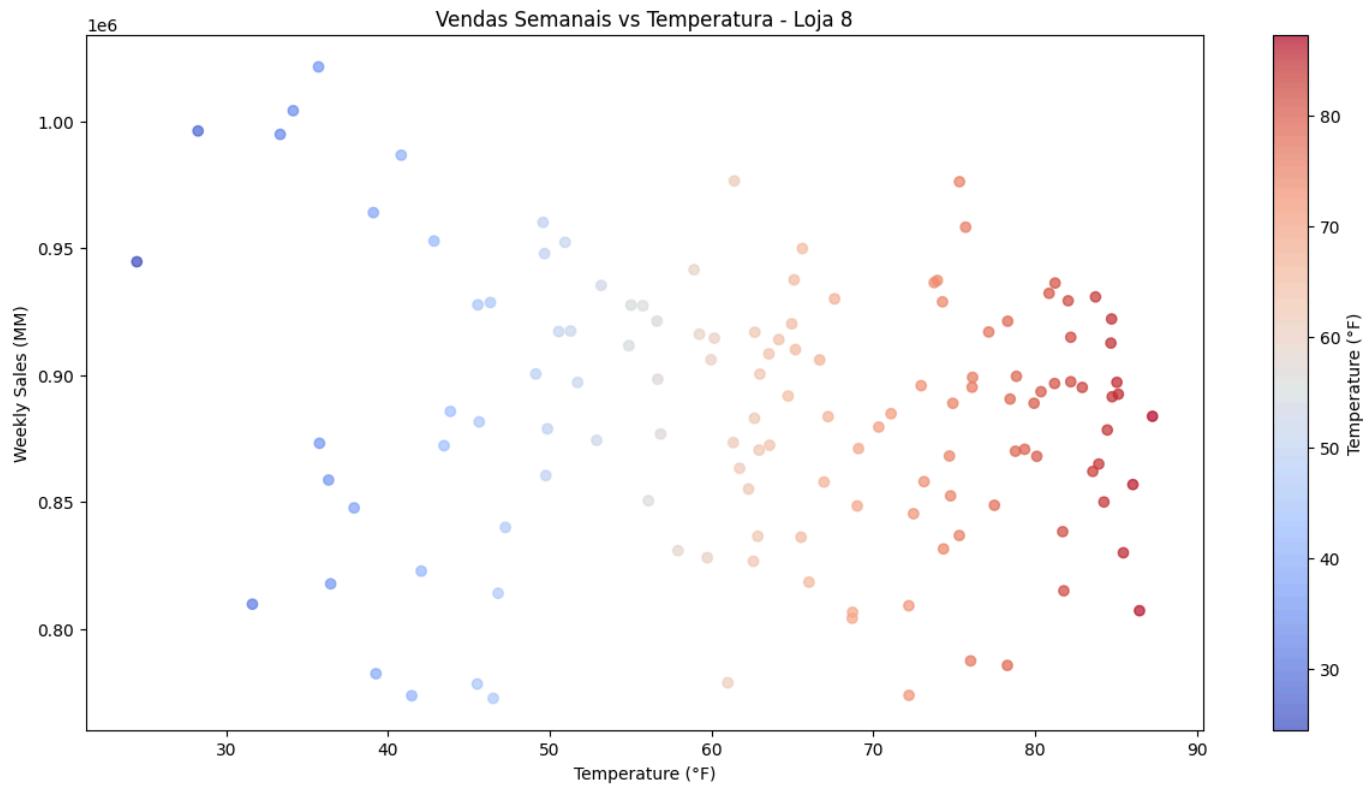
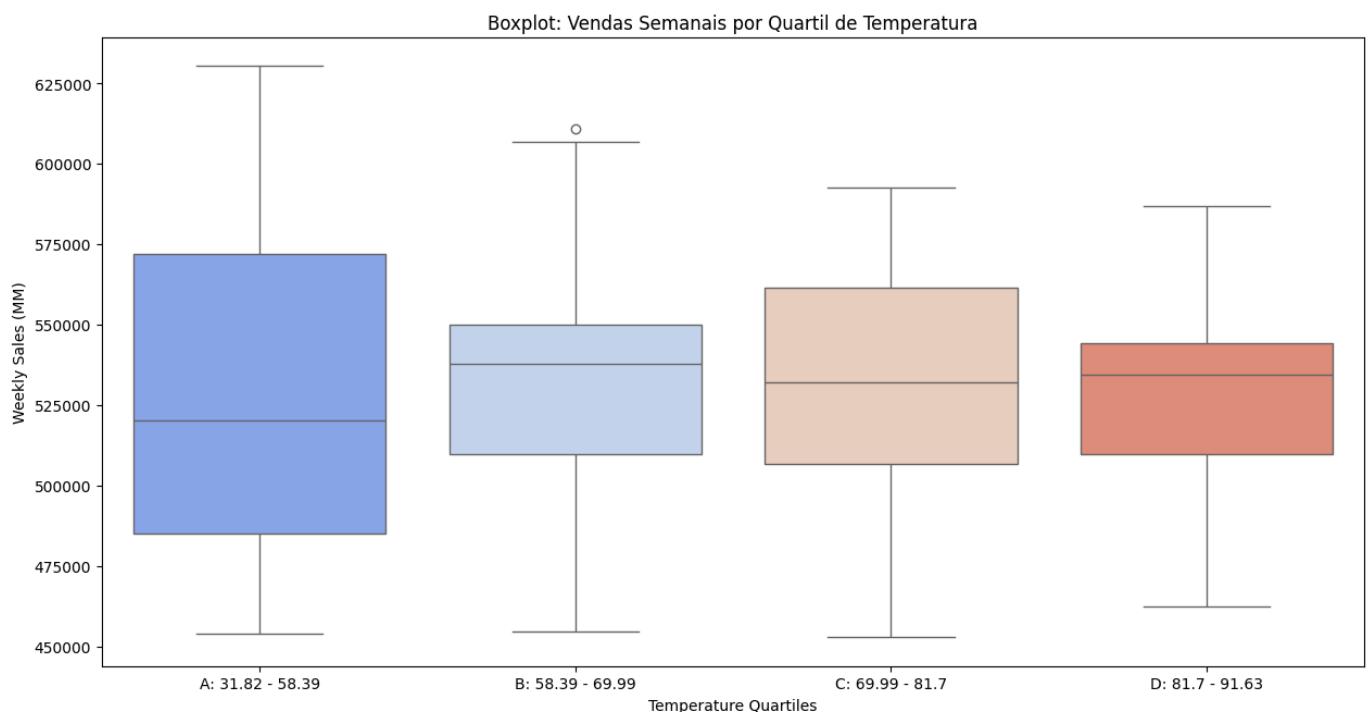


Gráfico de Vendas Semanais vs Temperatura - Loja 8 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 9 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 9 (Sem Outliers)

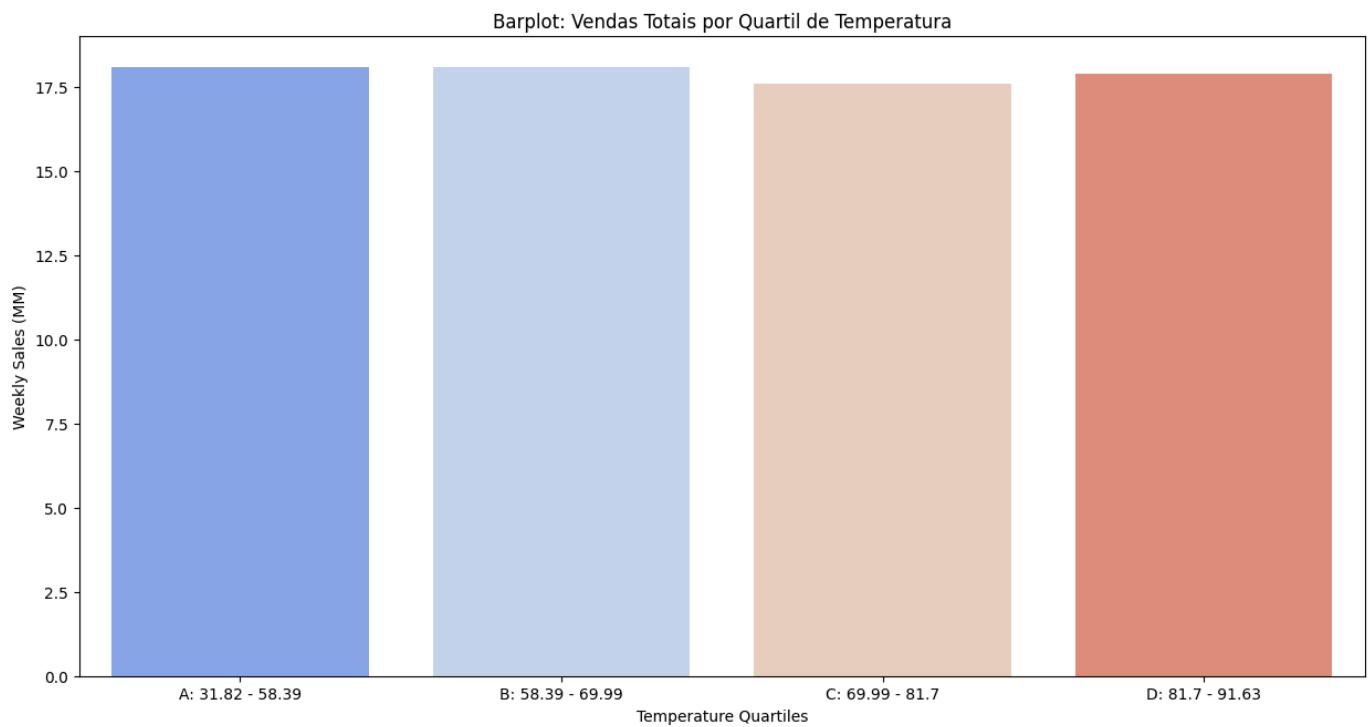
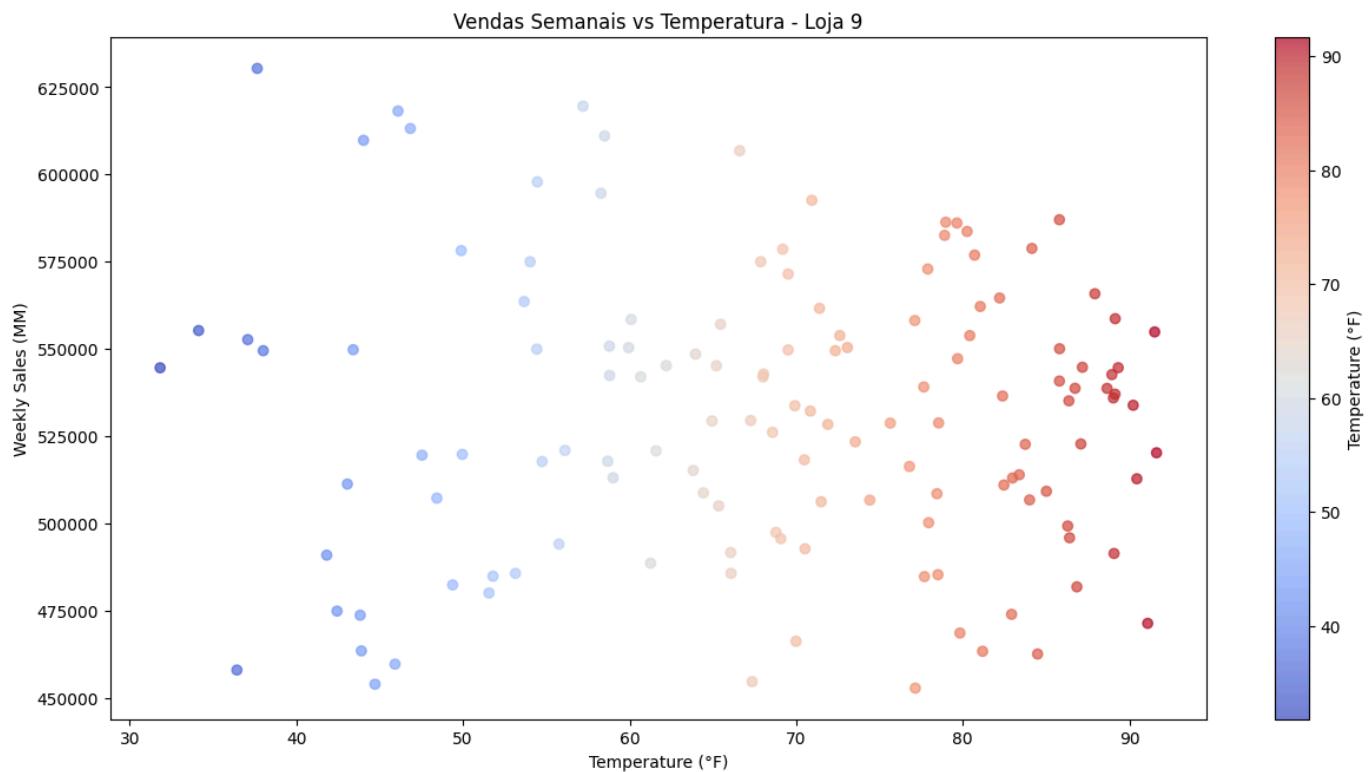
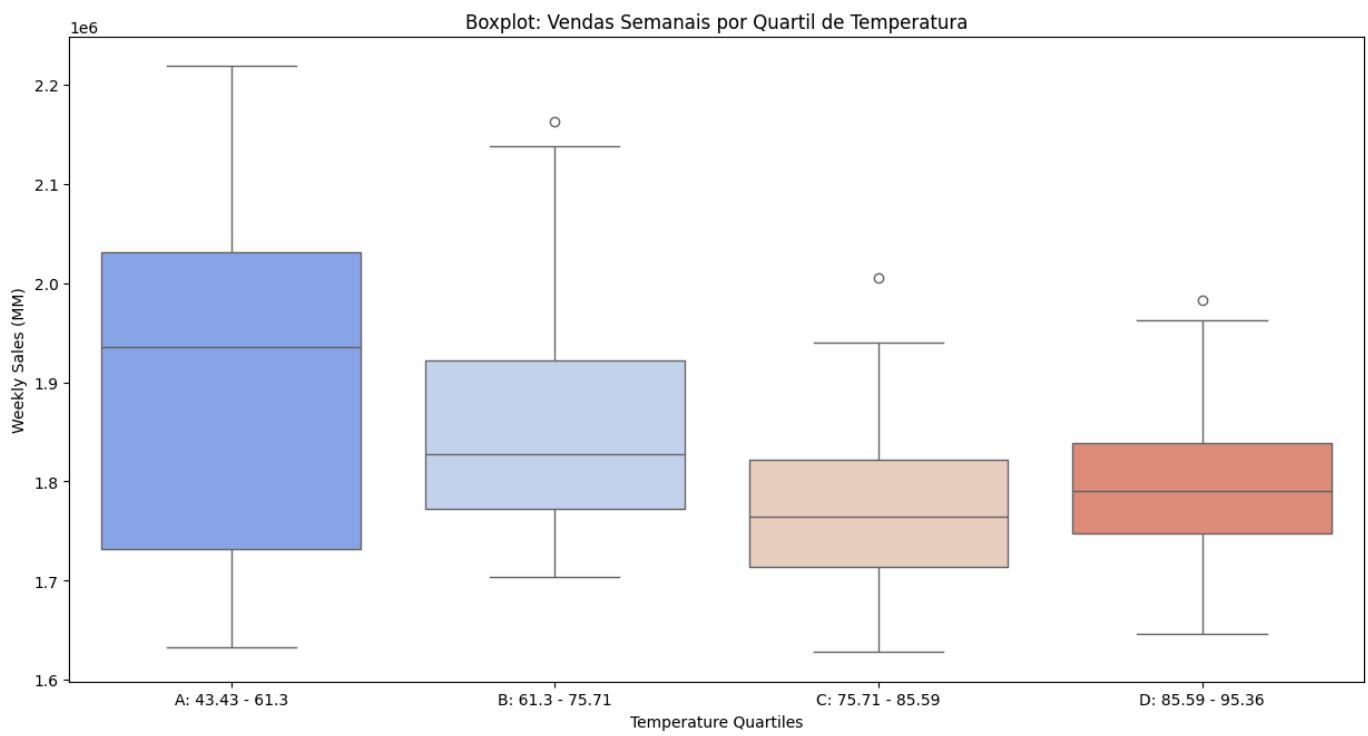


Gráfico de Vendas Semanais vs Temperatura - Loja 9 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 10 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 10 (Sem Outliers)

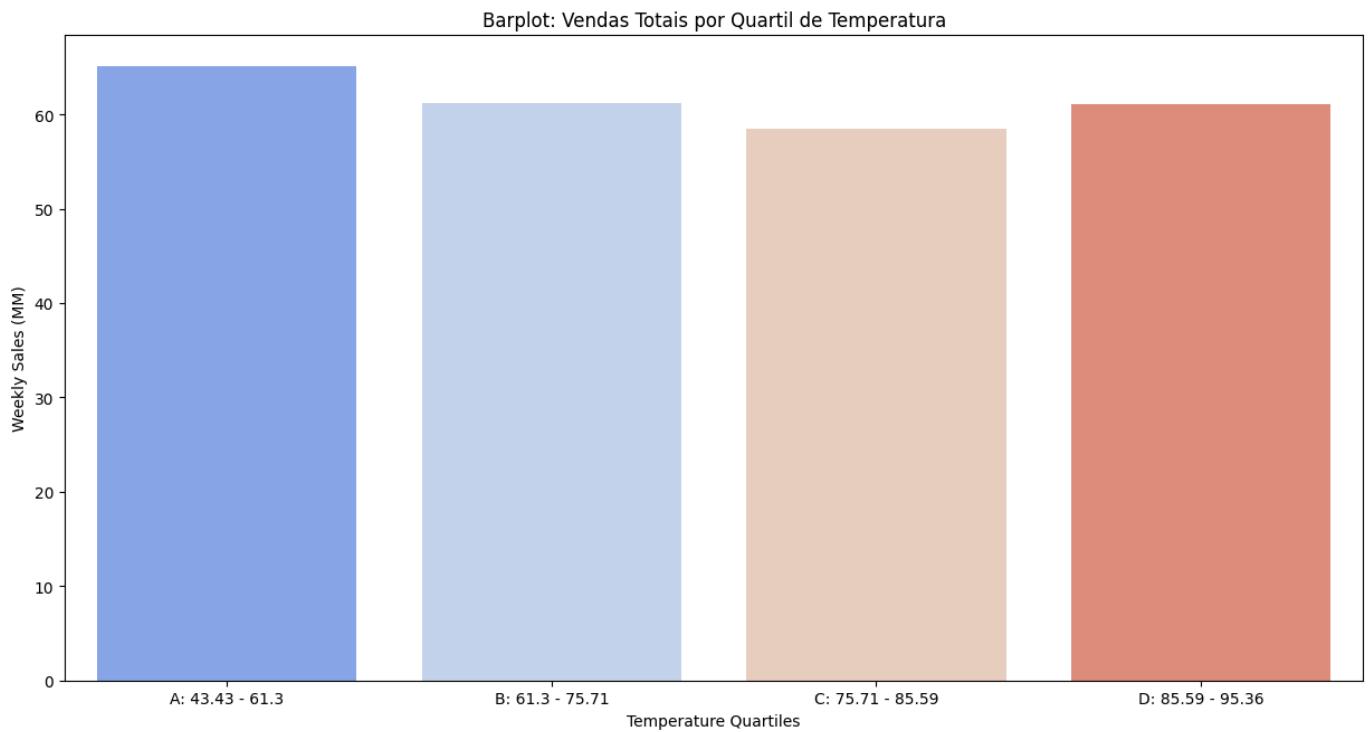
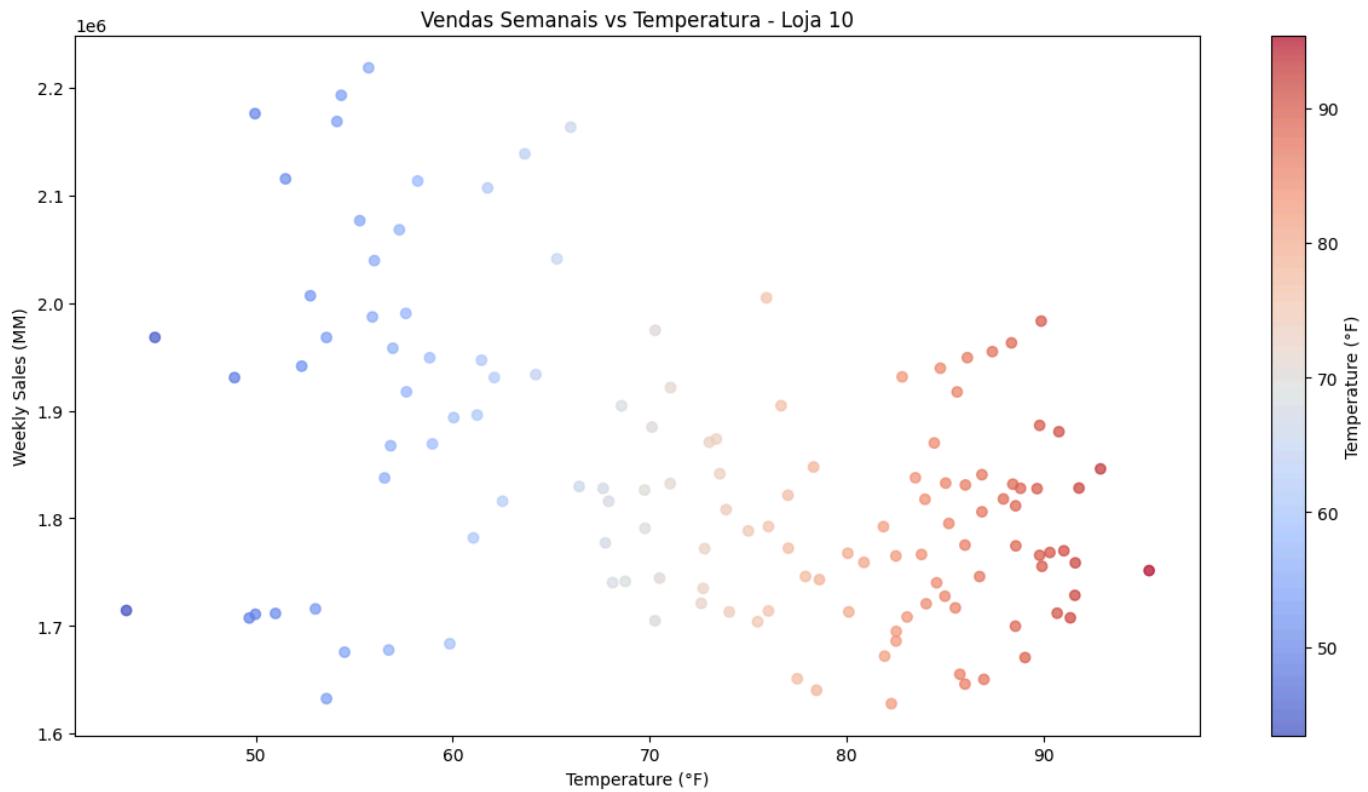
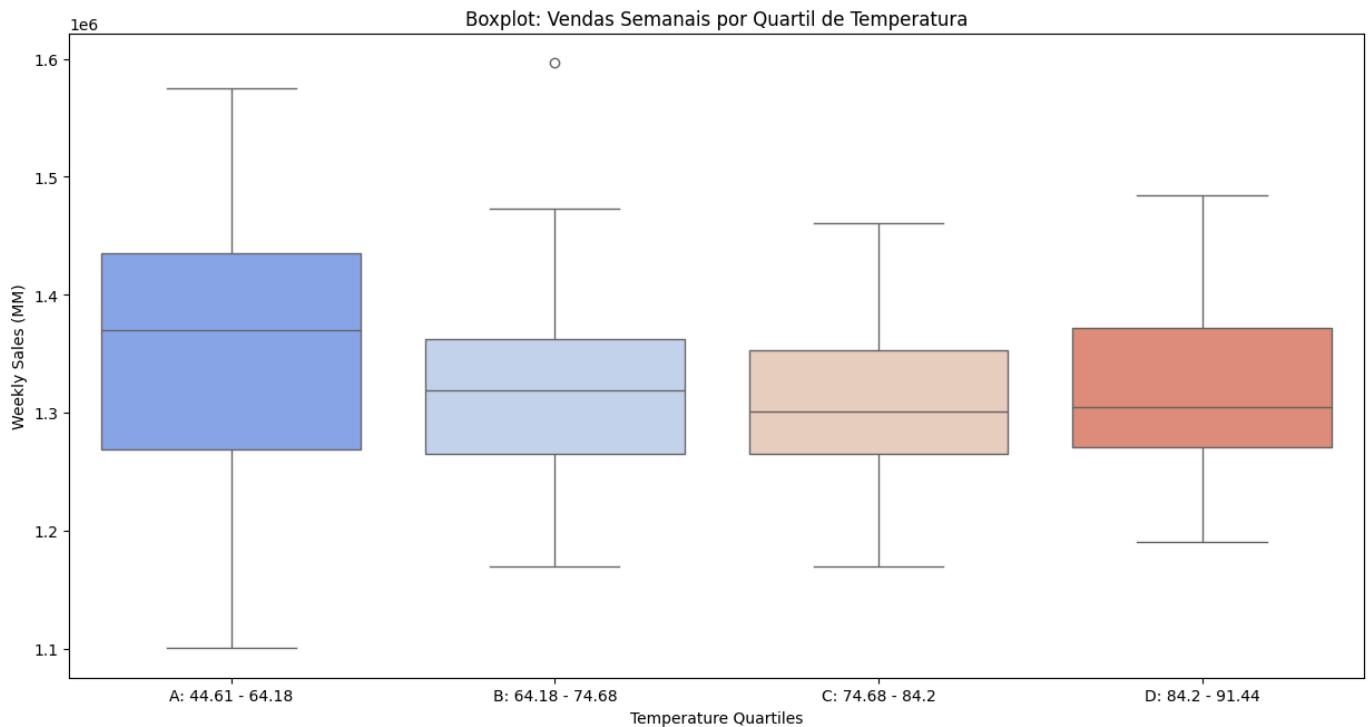


Gráfico de Vendas Semanais vs Temperatura - Loja 10 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 11 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 11 (Sem Outliers)

Barplot: Vendas Totais por Quartil de Temperatura

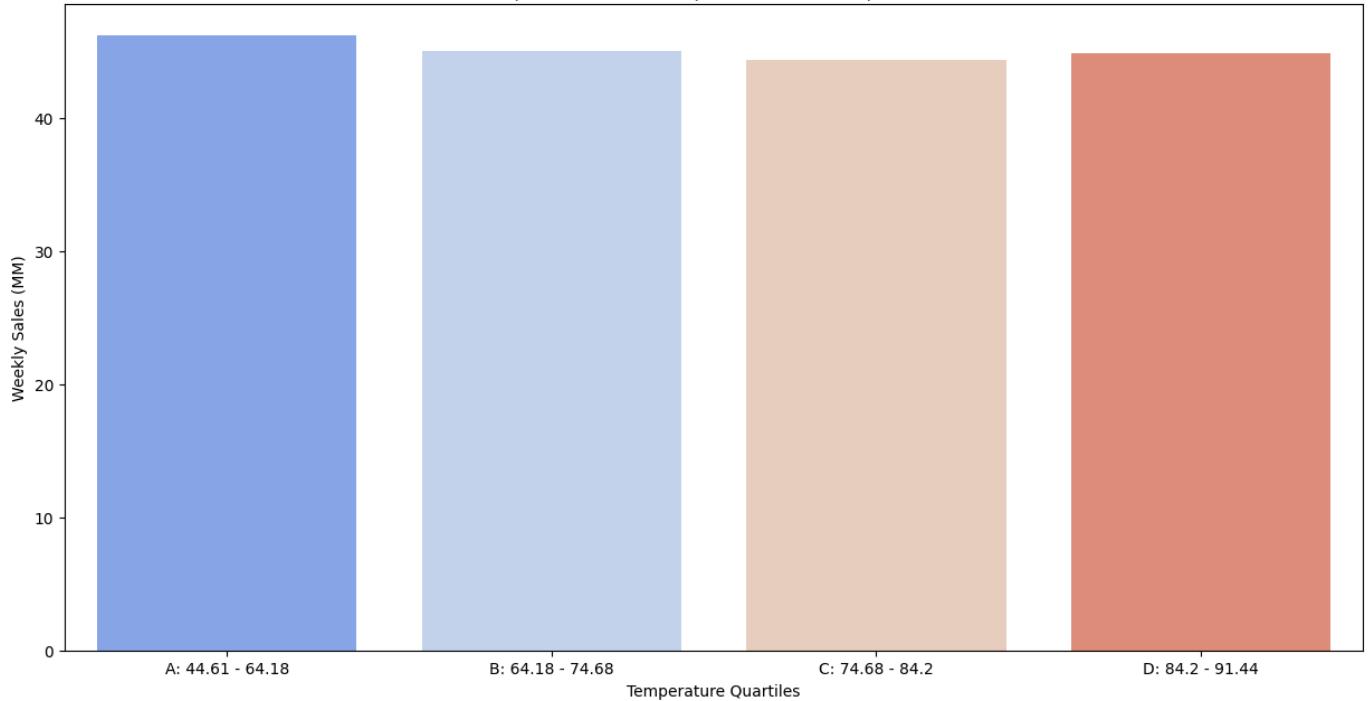
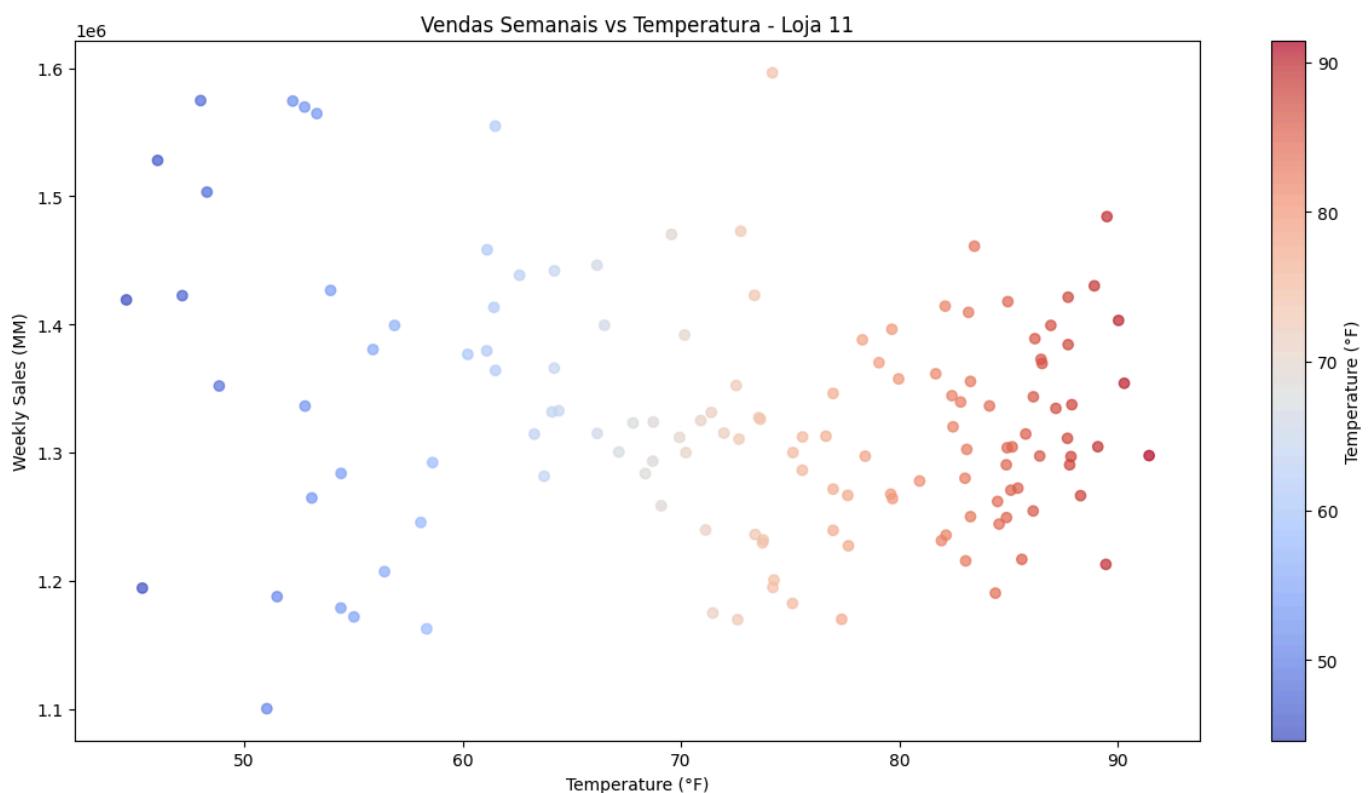
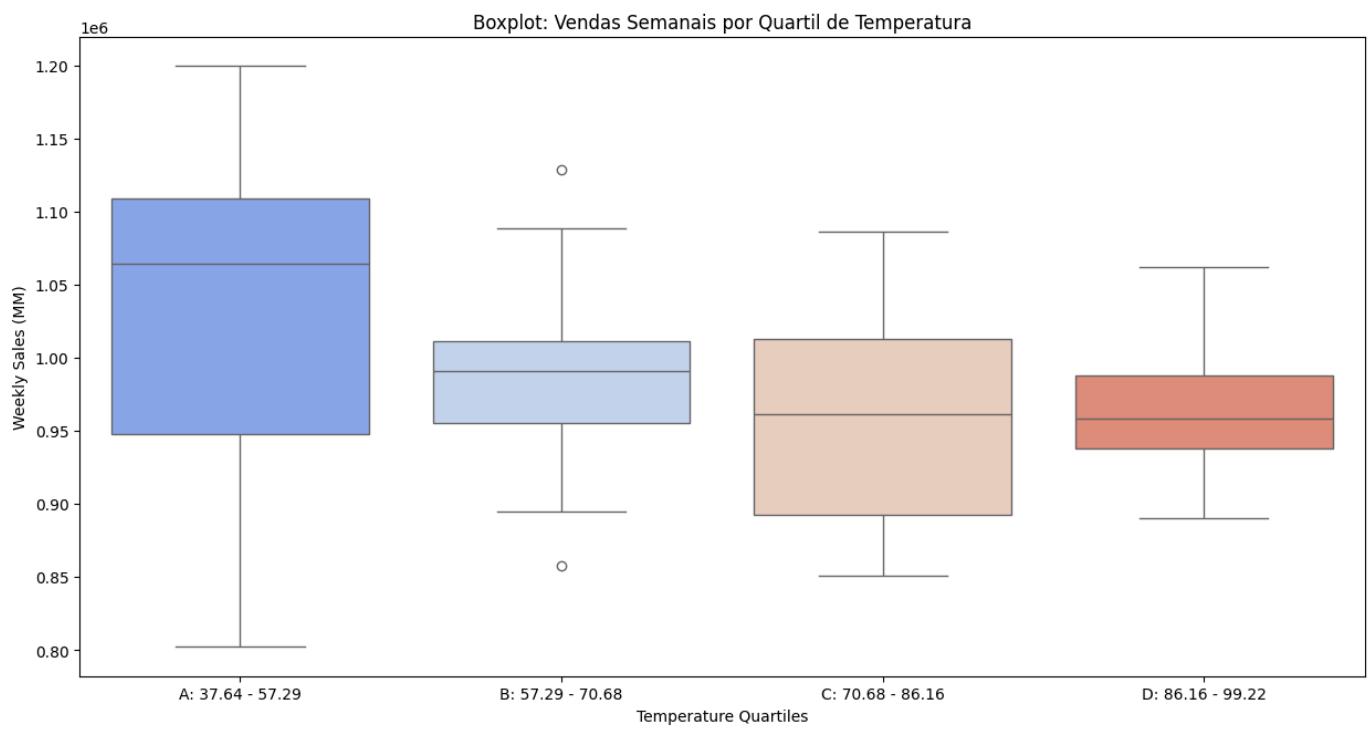


Gráfico de Vendas Semanais vs Temperatura - Loja 11 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 12 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 12 (Sem Outliers)

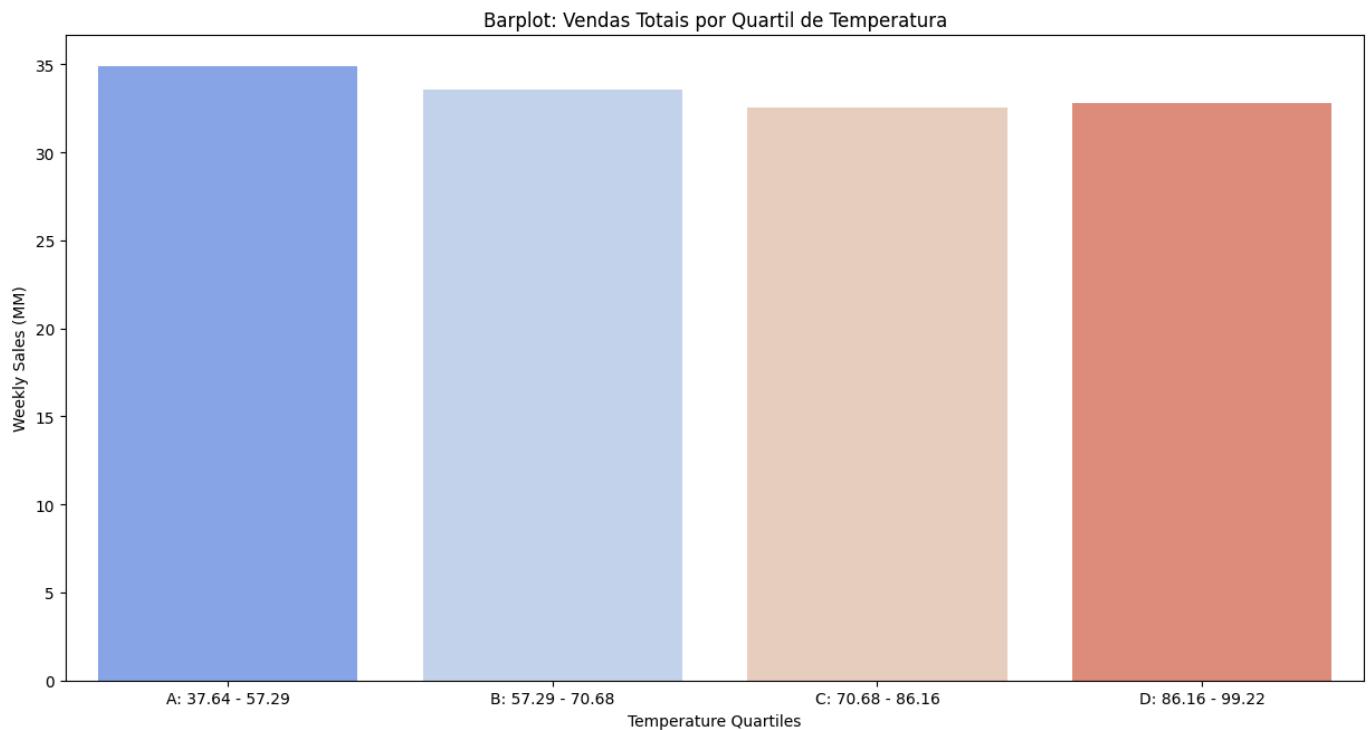
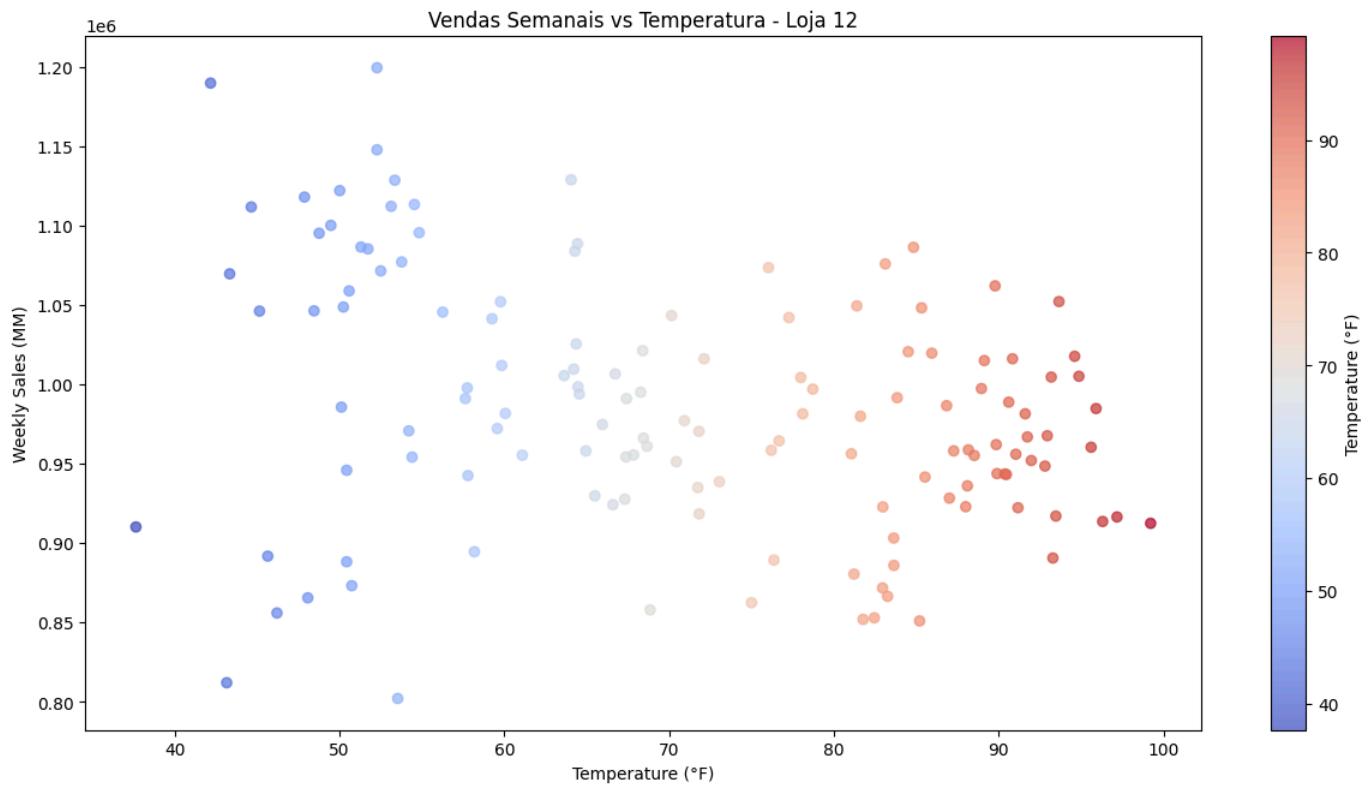
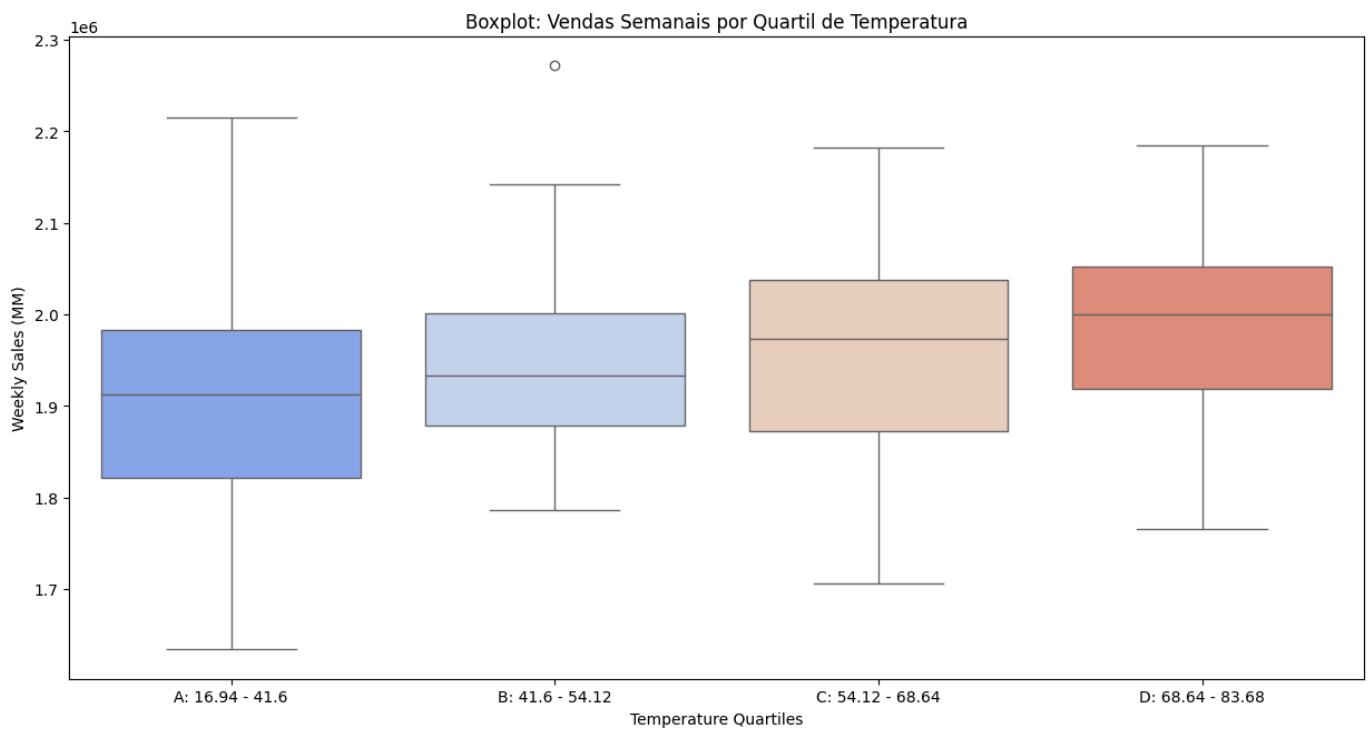


Gráfico de Vendas Semanais vs Temperatura - Loja 12 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 13 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 13 (Sem Outliers)

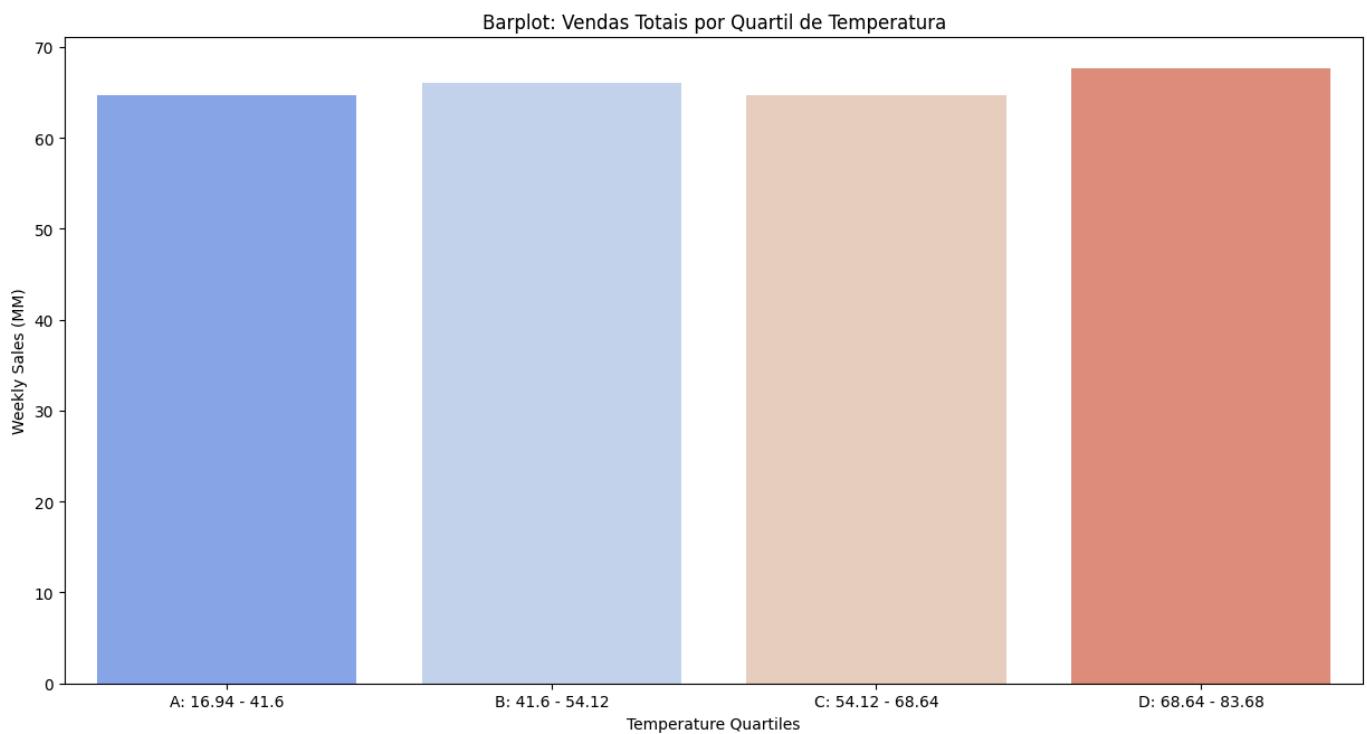
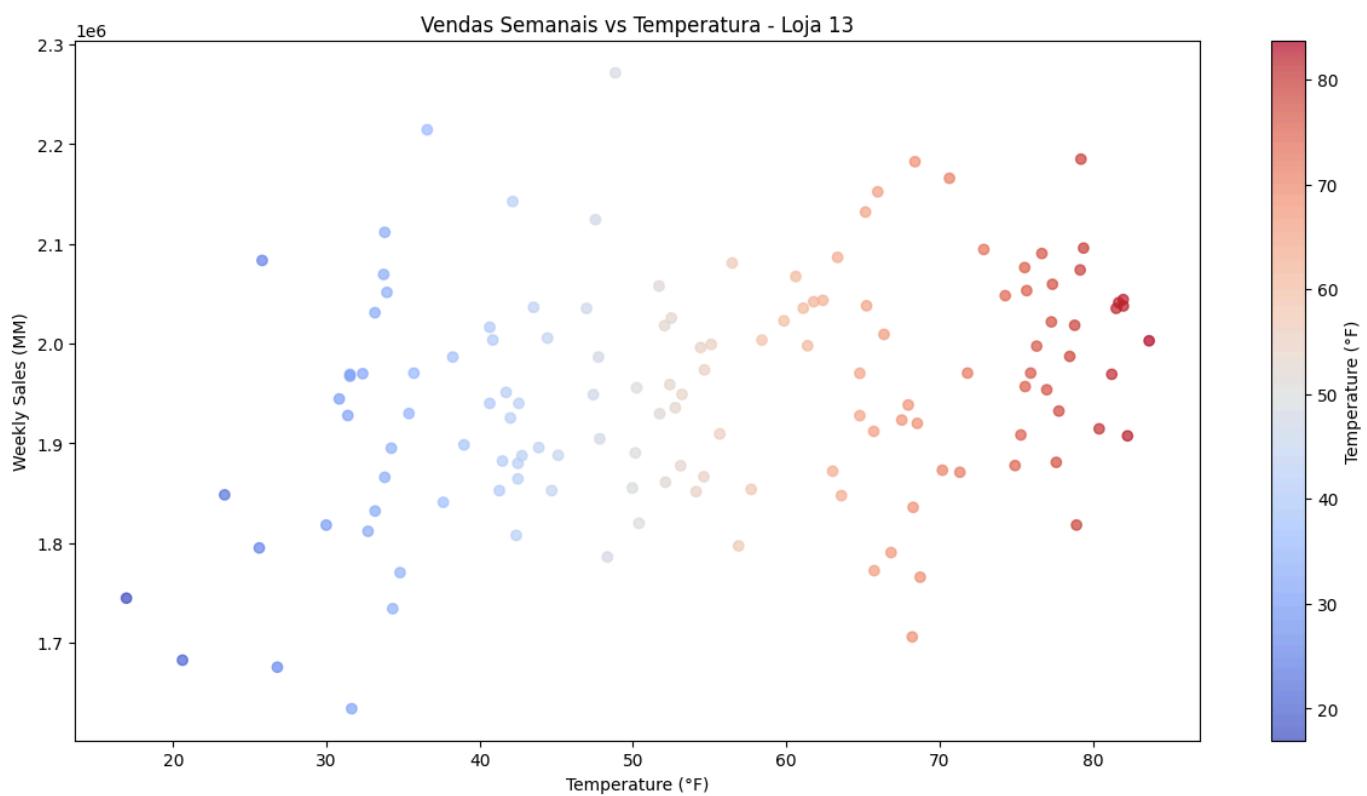
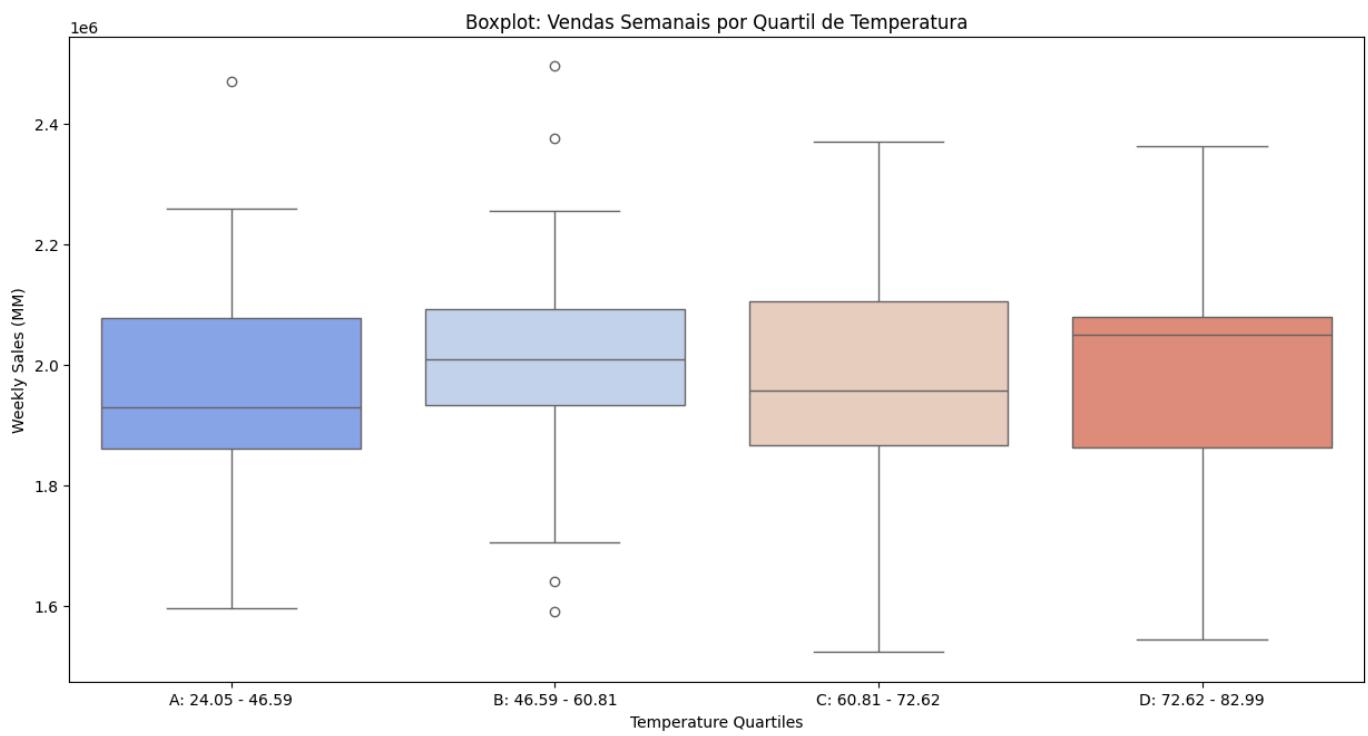


Gráfico de Vendas Semanais vs Temperatura - Loja 13 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 14 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 14 (Sem Outliers)

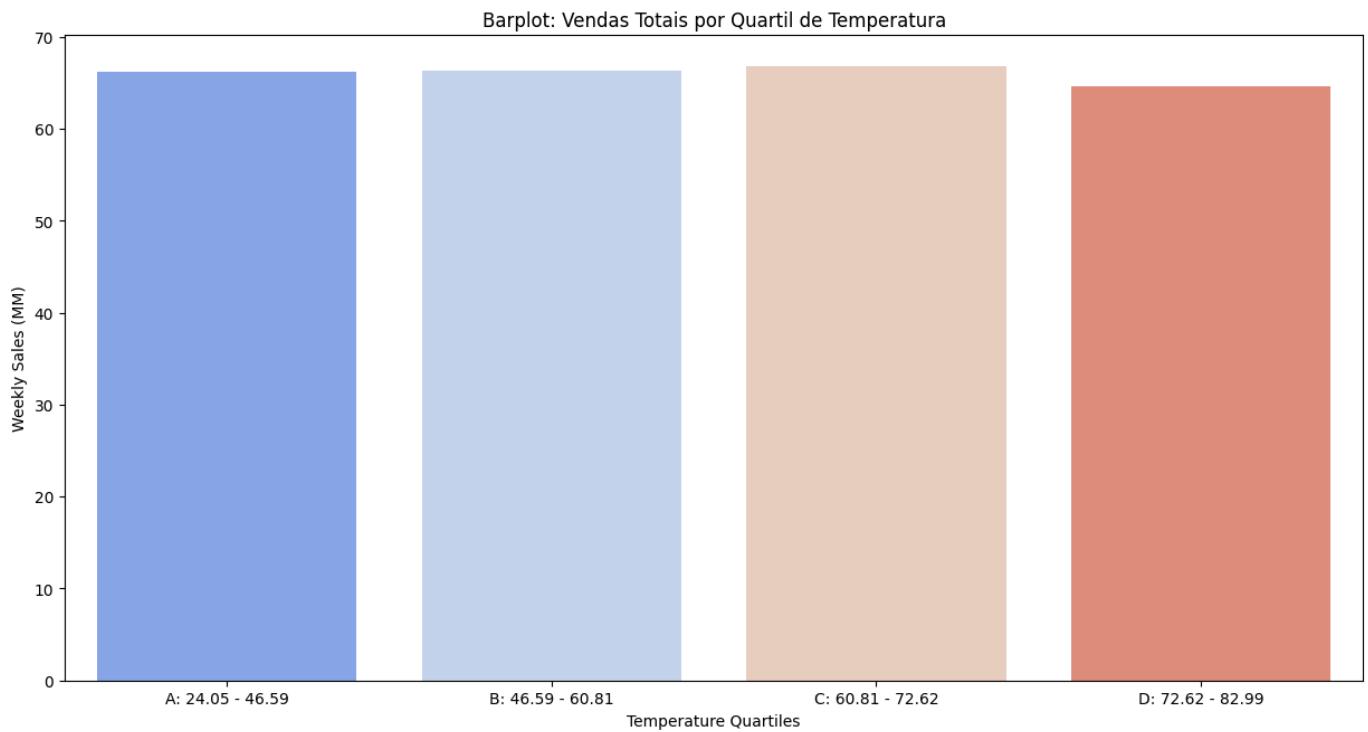
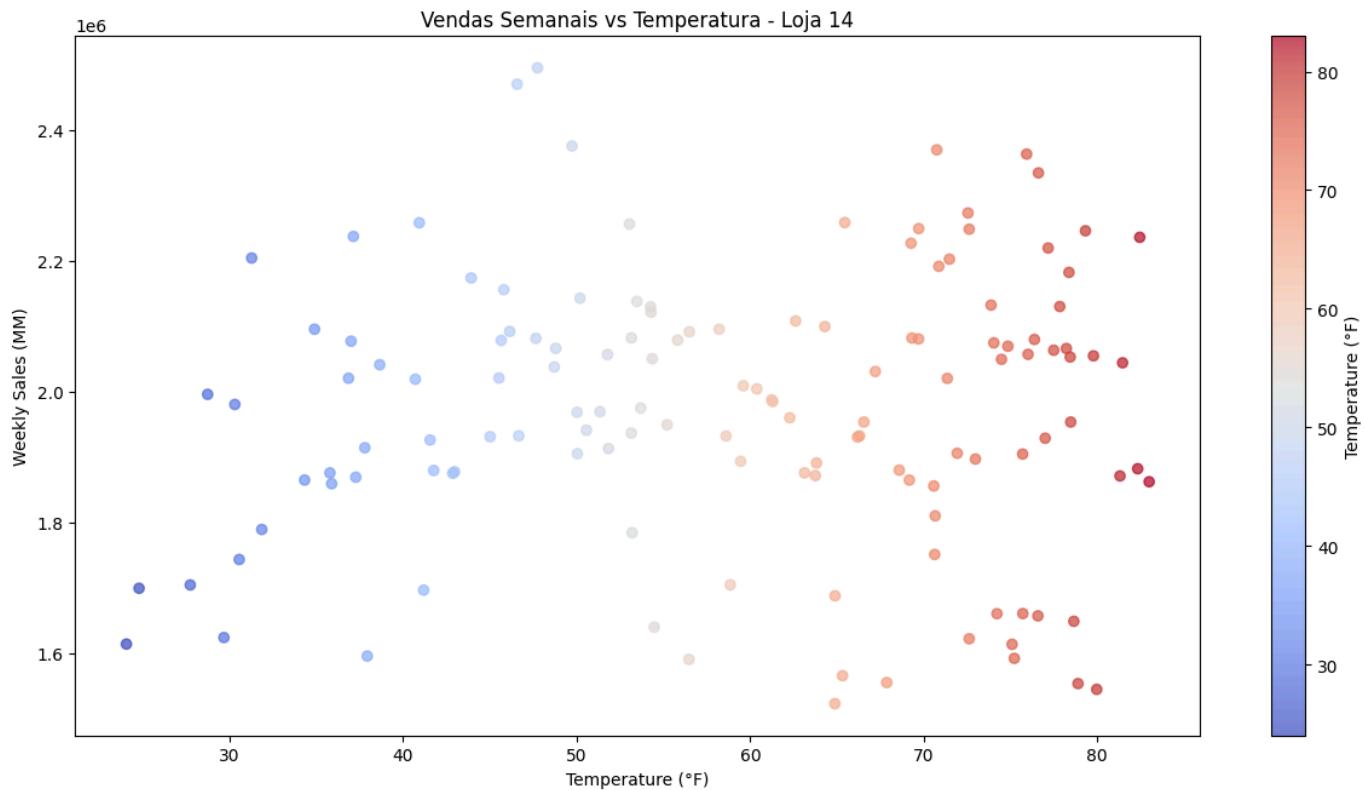
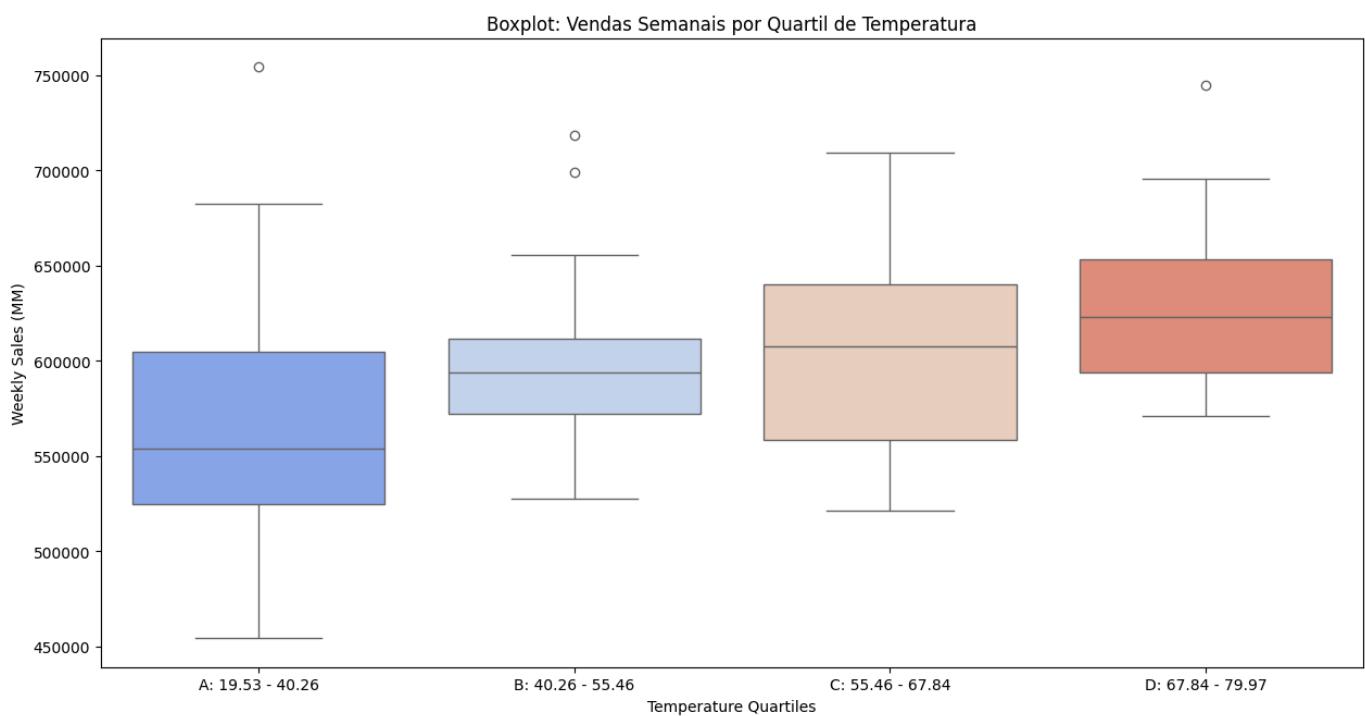


Gráfico de Vendas Semanais vs Temperatura - Loja 14 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 15 (Sem Outliers)



Barplot: Vendas Totais por Quartil de Temperatura

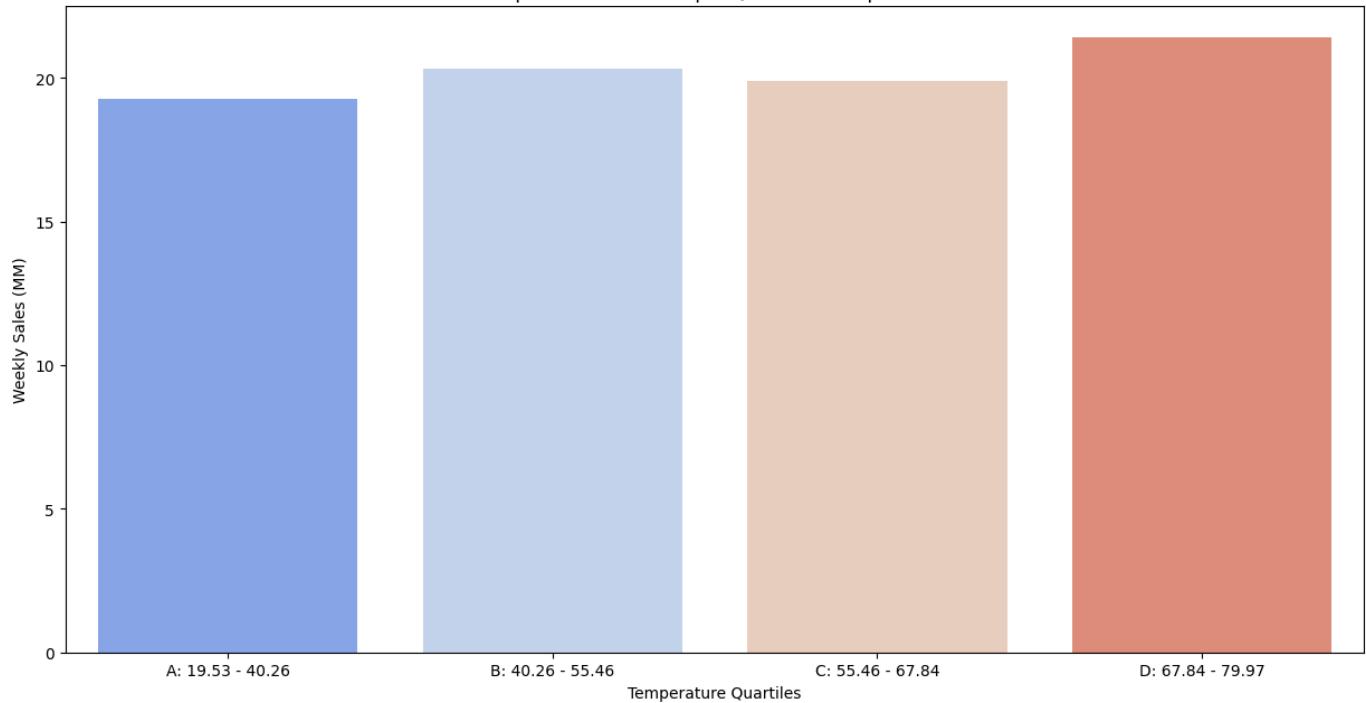
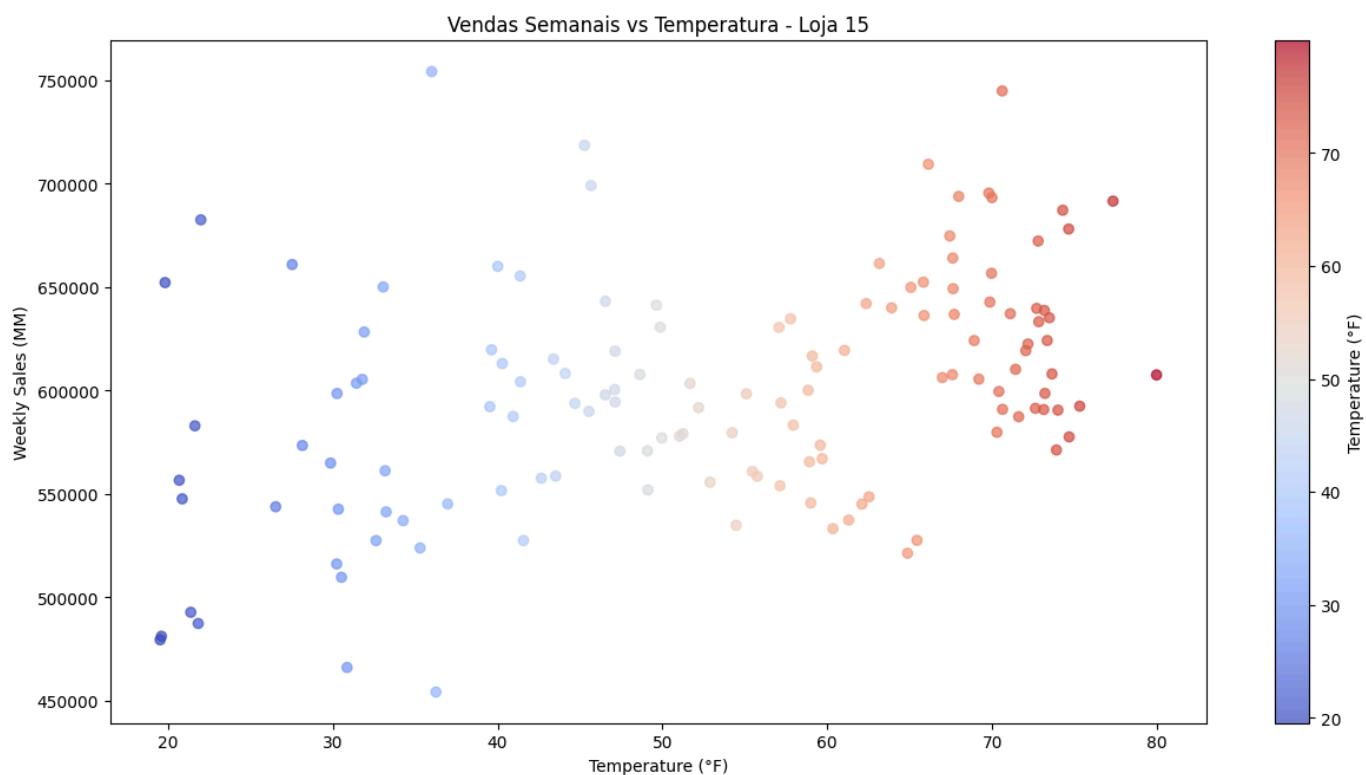
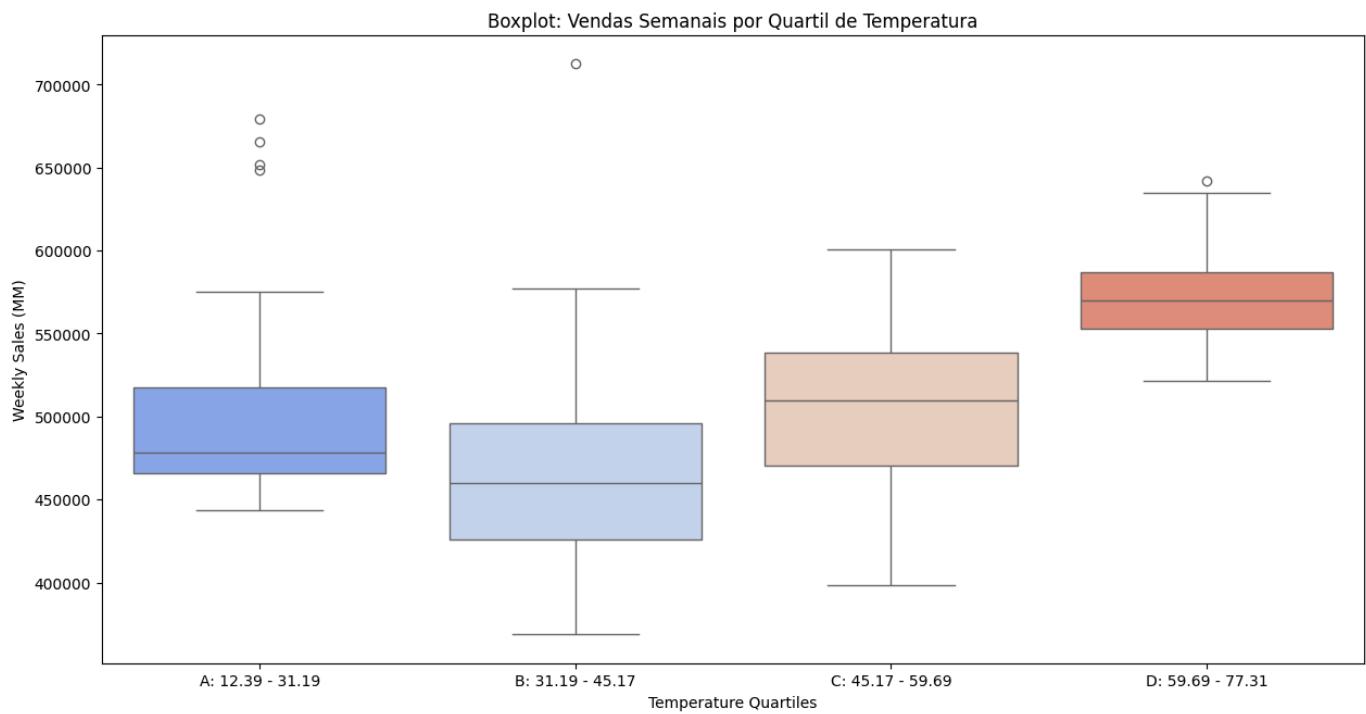


Gráfico de Vendas Semanais vs Temperatura - Loja 15 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 16 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 16 (Sem Outliers)

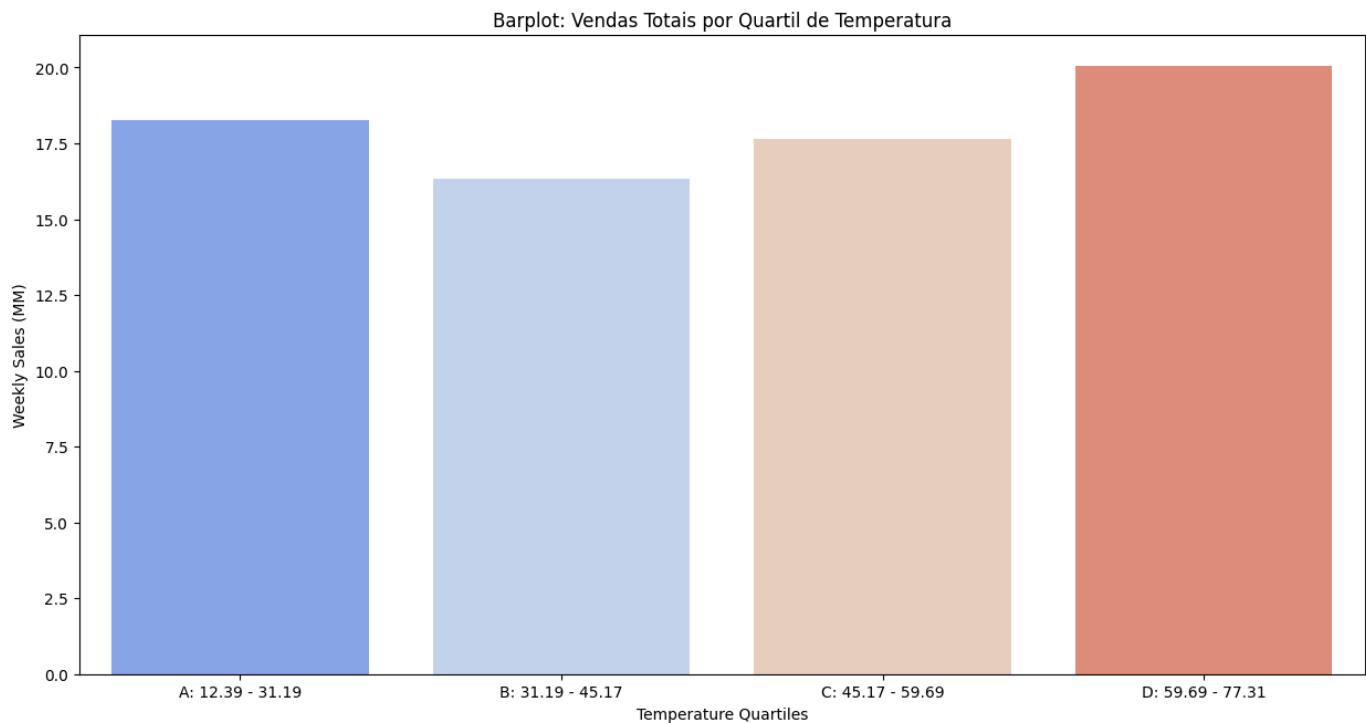
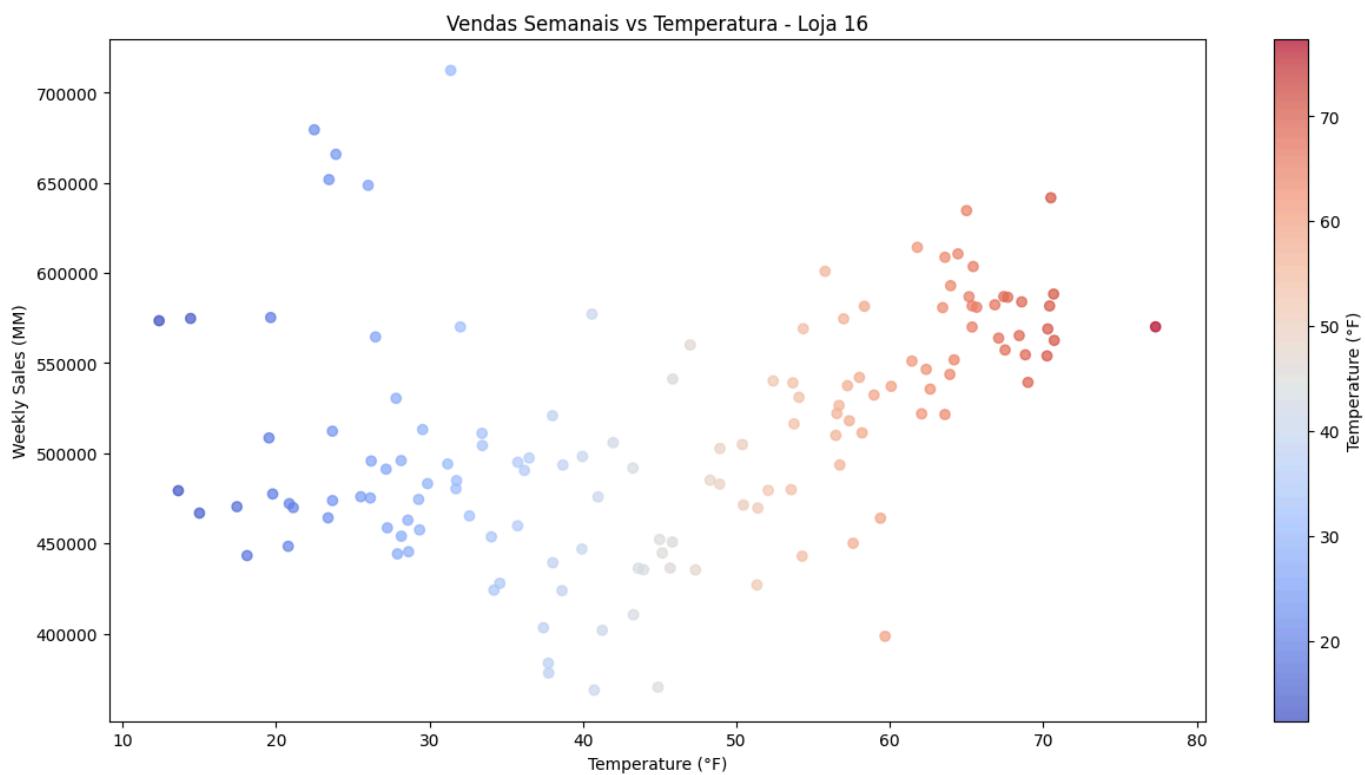
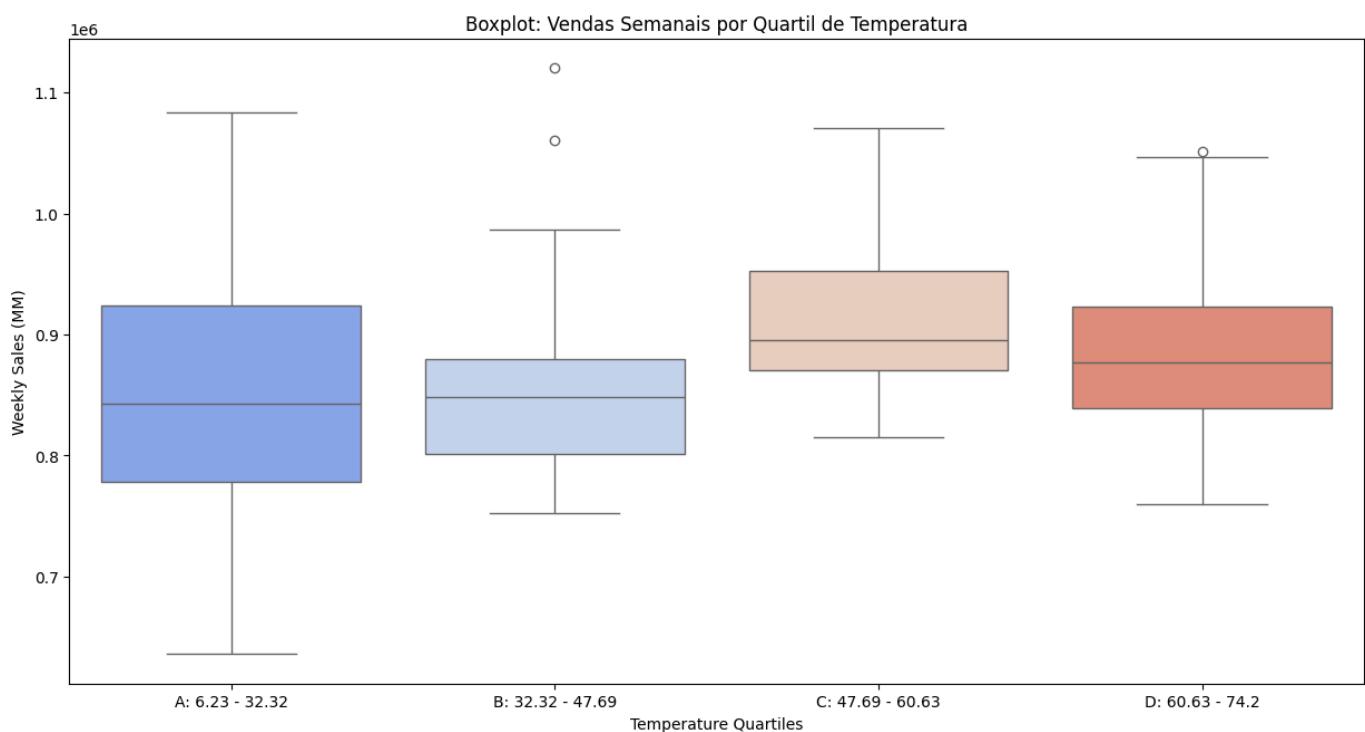


Gráfico de Vendas Semanais vs Temperatura - Loja 16 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 17 (Sem Outliers)



Barplot: Vendas Totais por Quartil de Temperatura

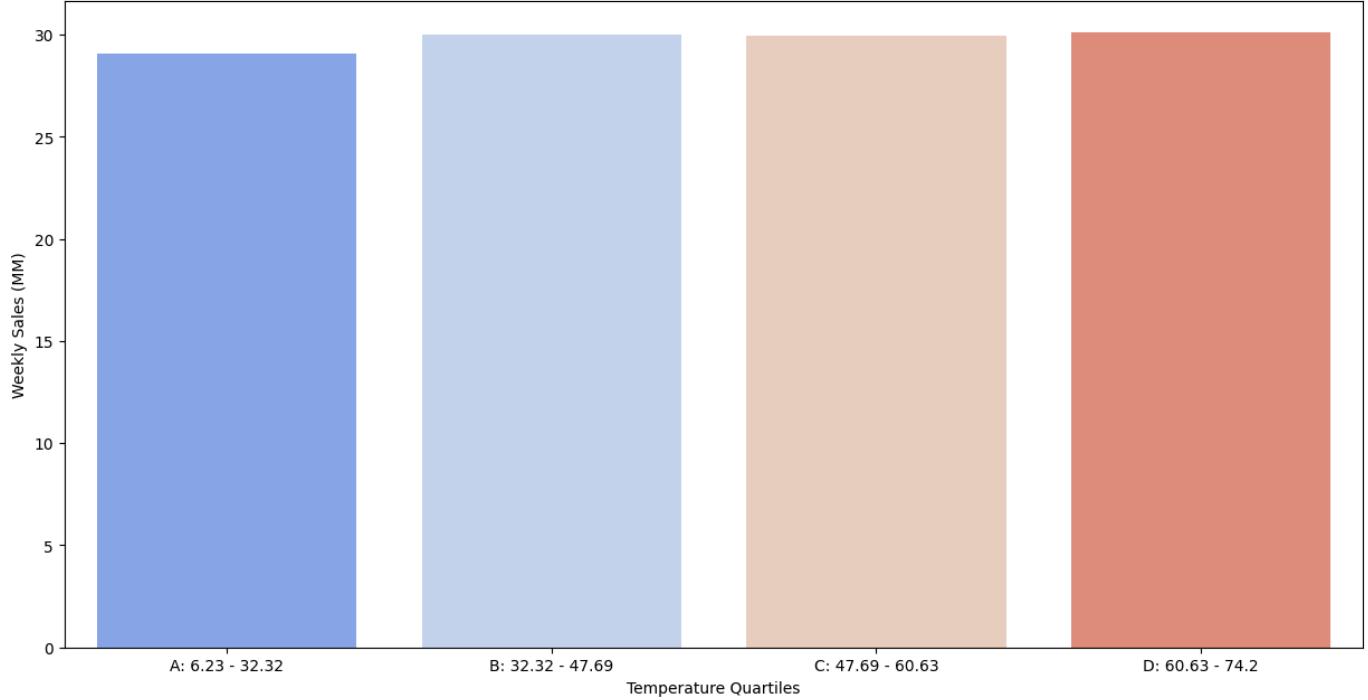
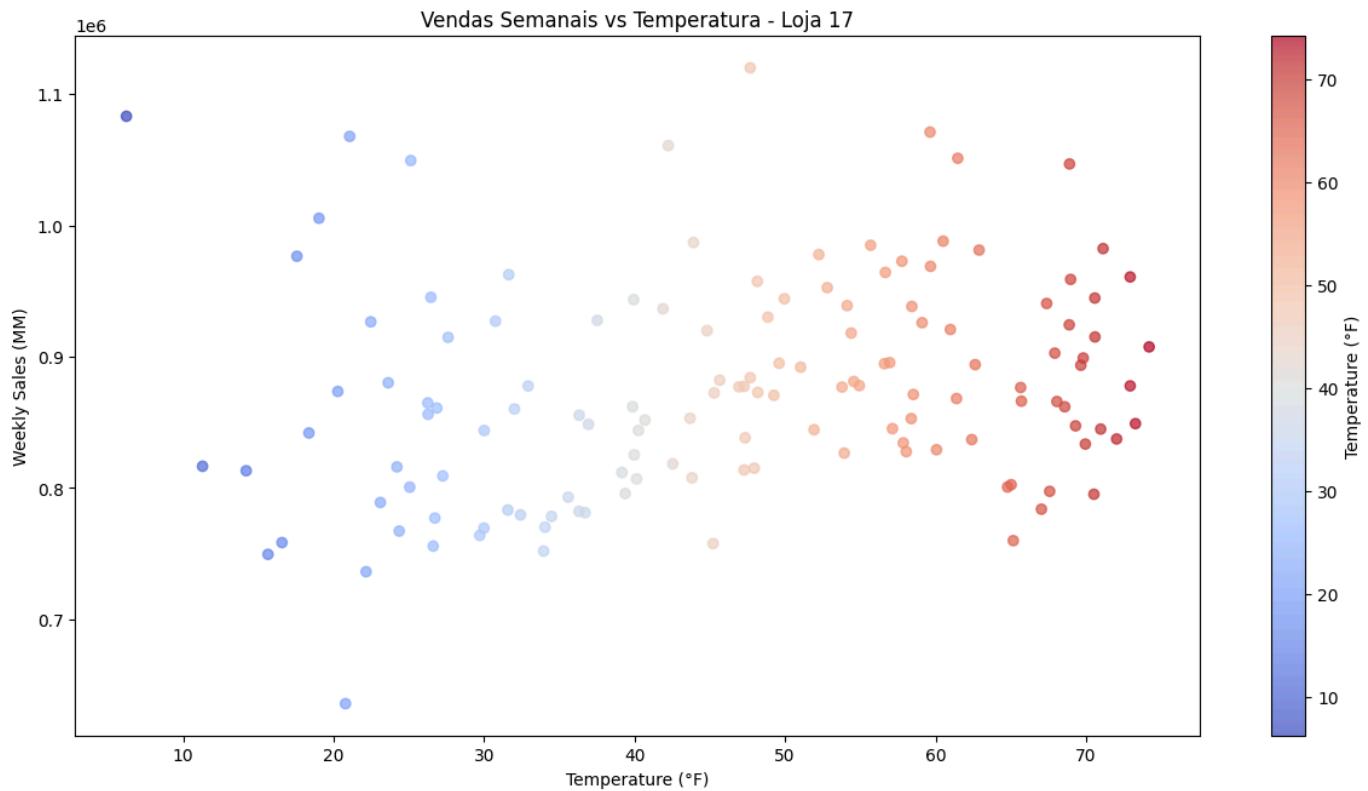
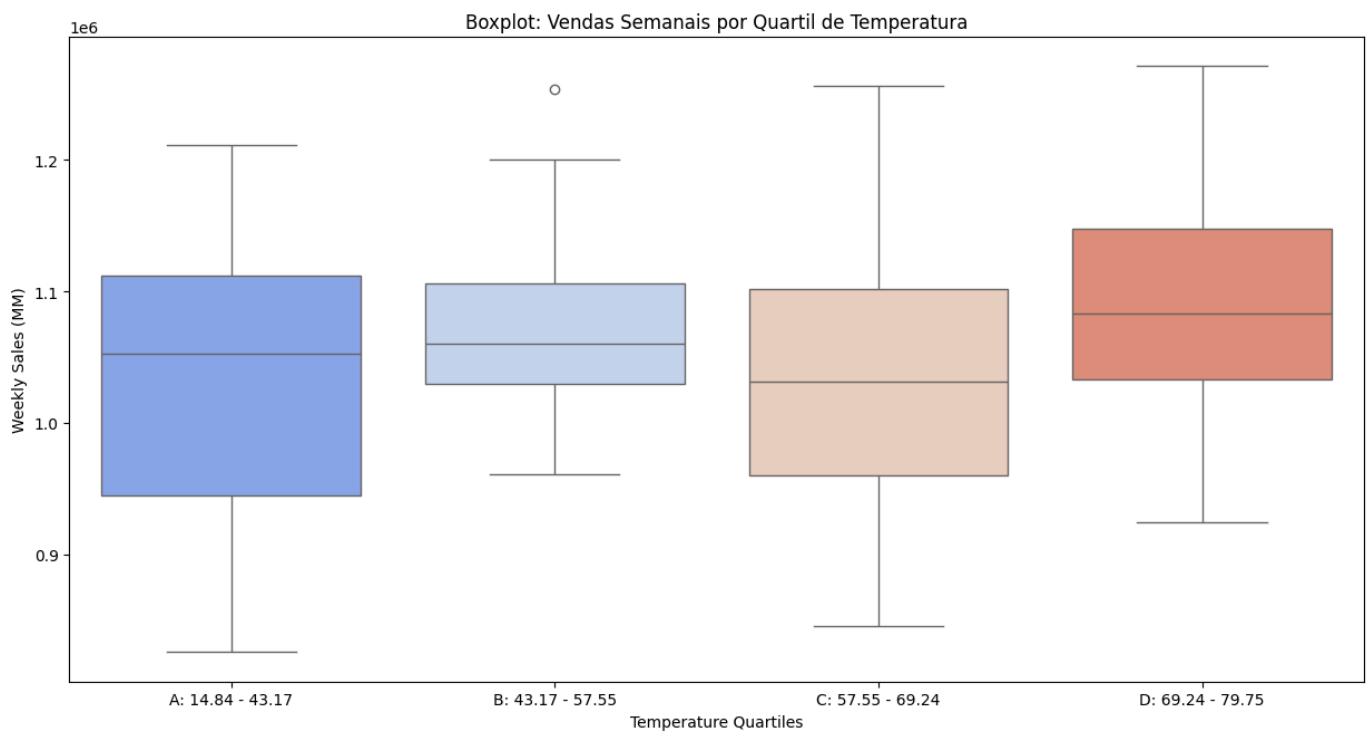


Gráfico de Vendas Semanais vs Temperatura - Loja 17 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 18 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 18 (Sem Outliers)

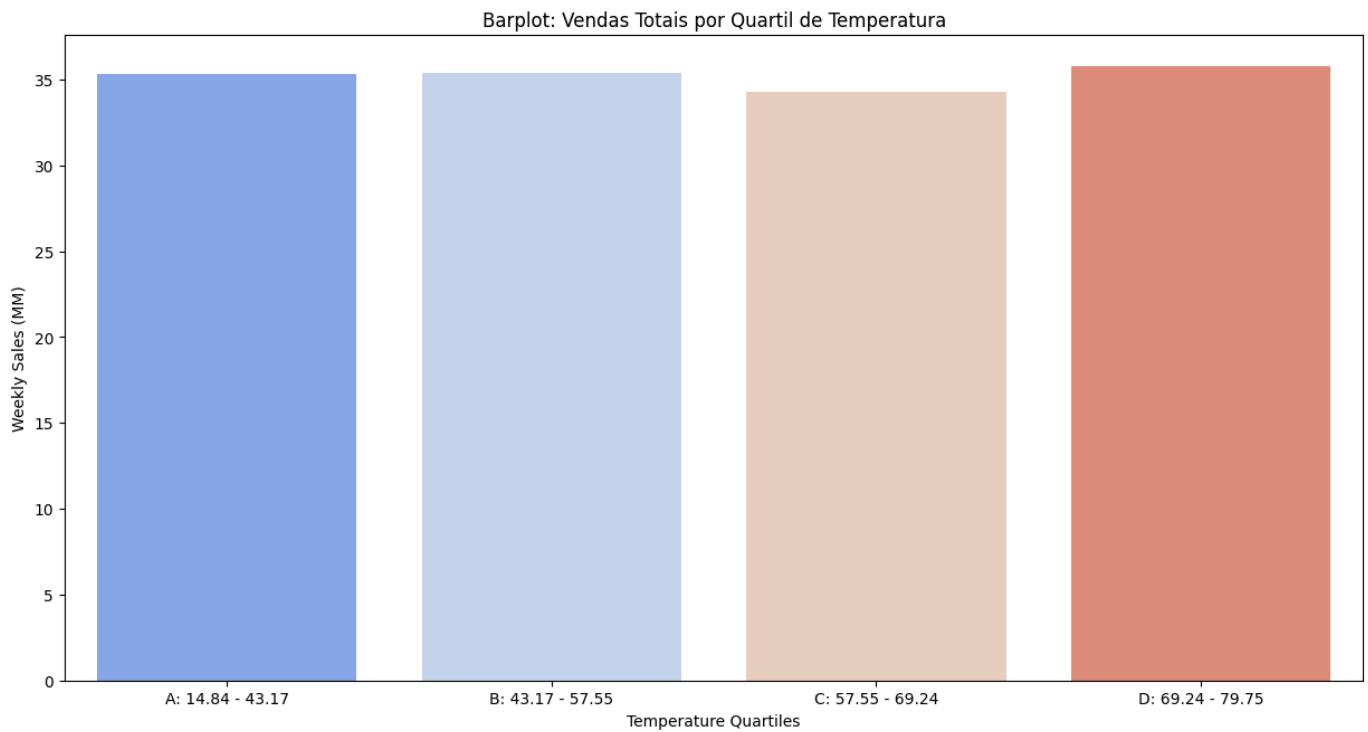
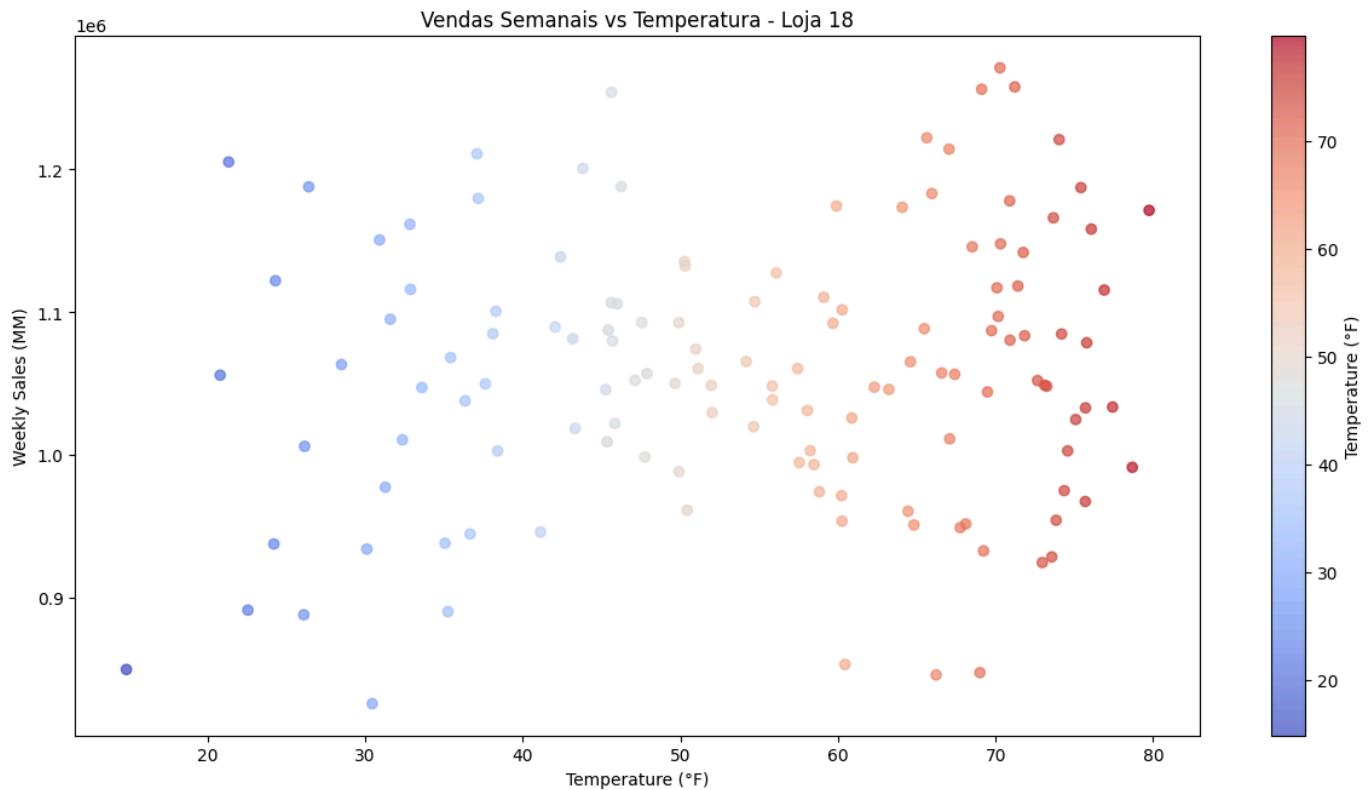
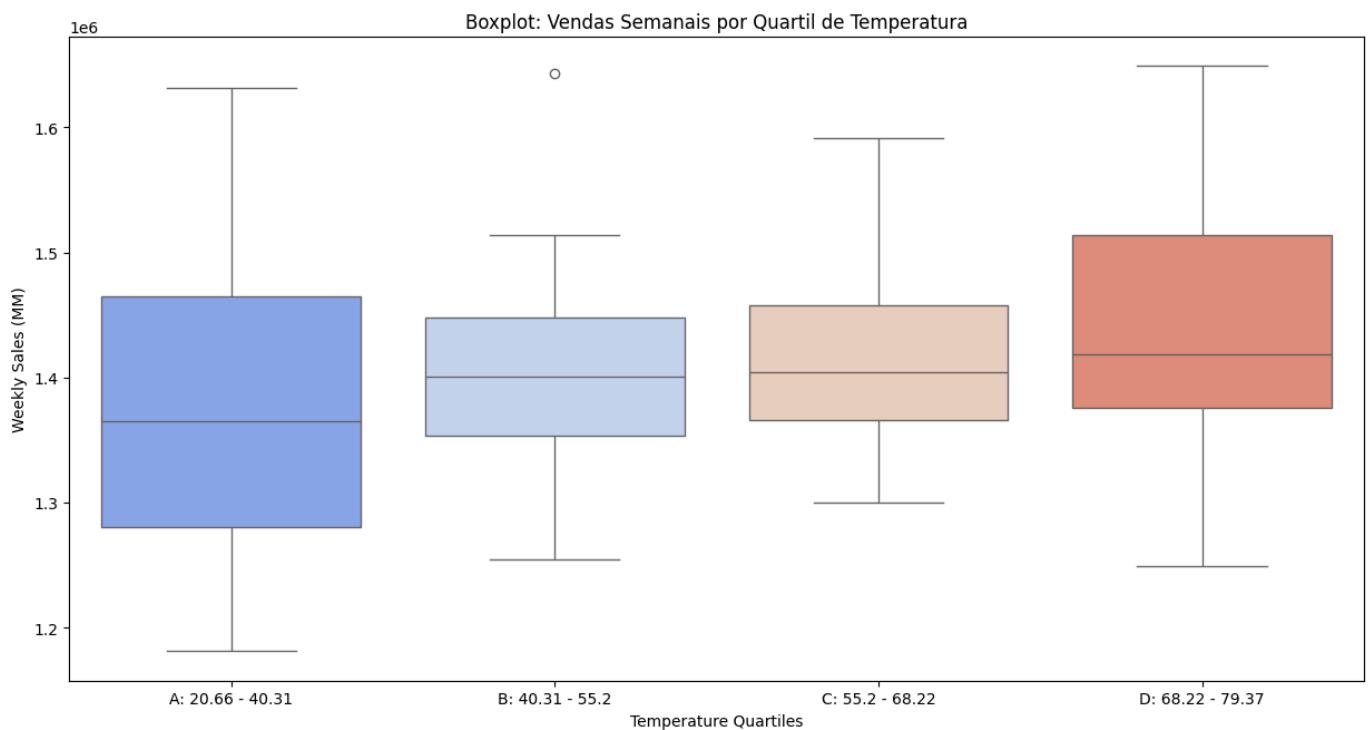


Gráfico de Vendas Semanais vs Temperatura - Loja 18 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 19 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 19 (Sem Outliers)

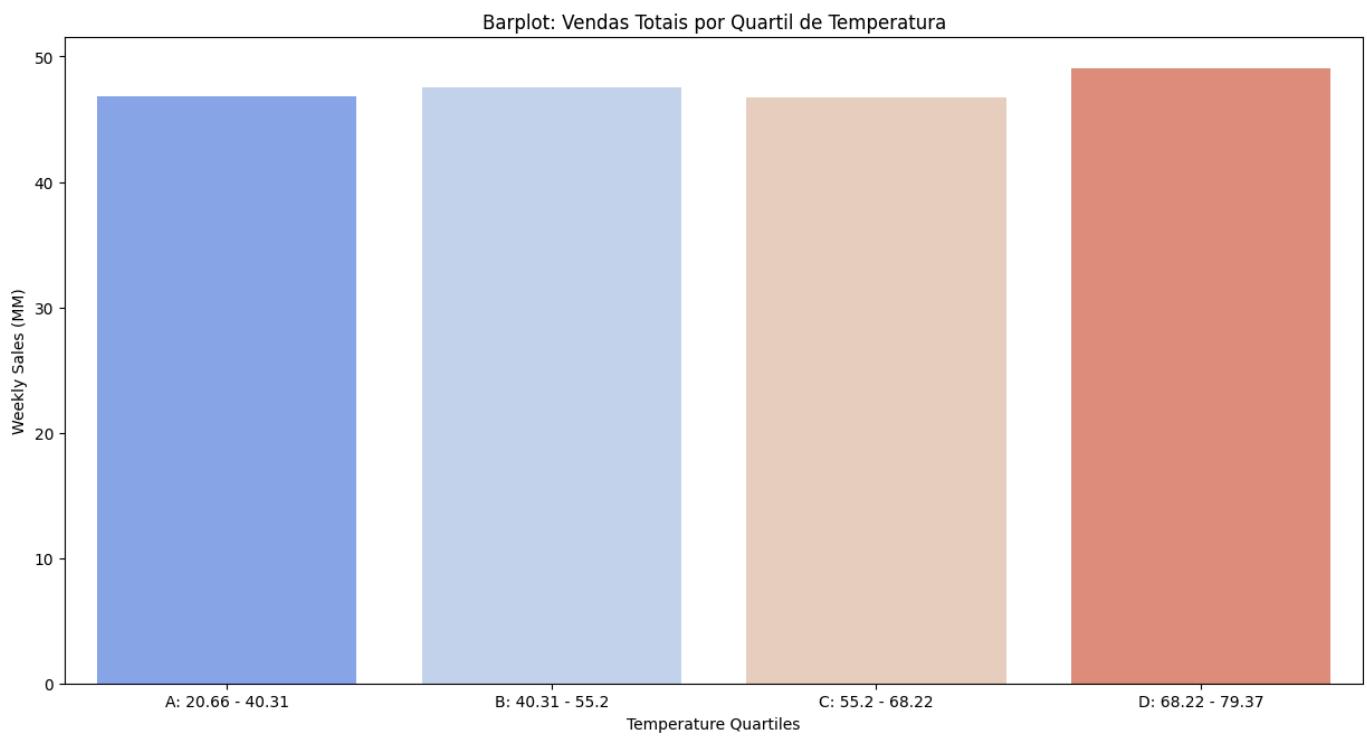
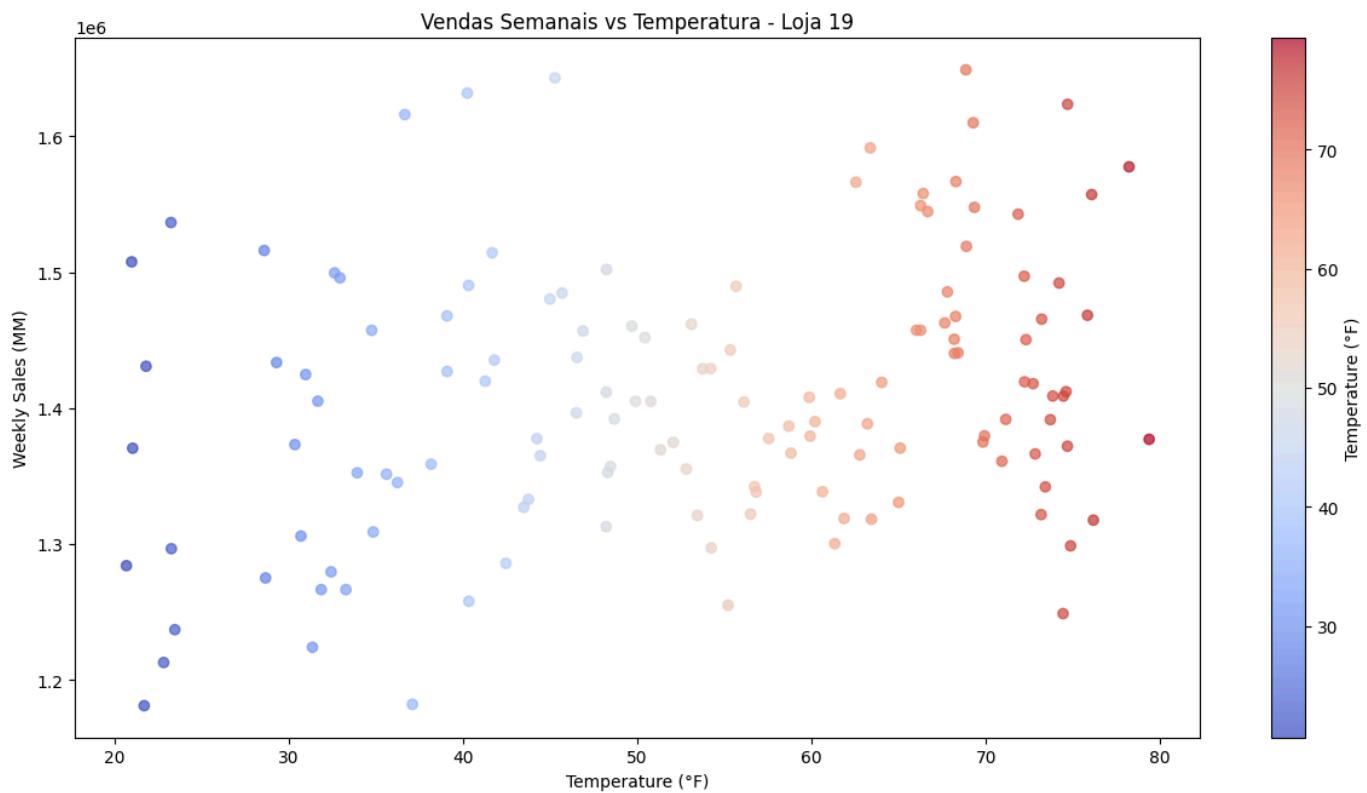
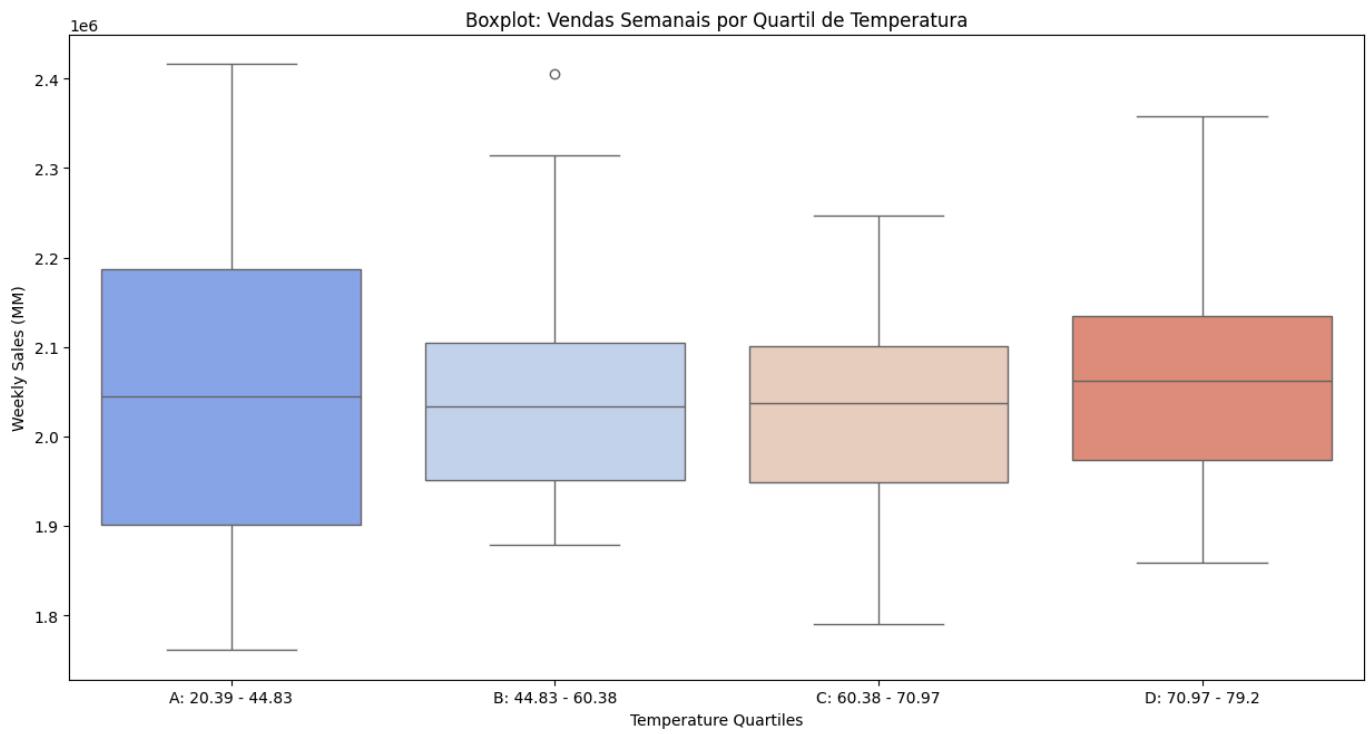


Gráfico de Vendas Semanais vs Temperatura - Loja 19 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 20 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 20 (Sem Outliers)

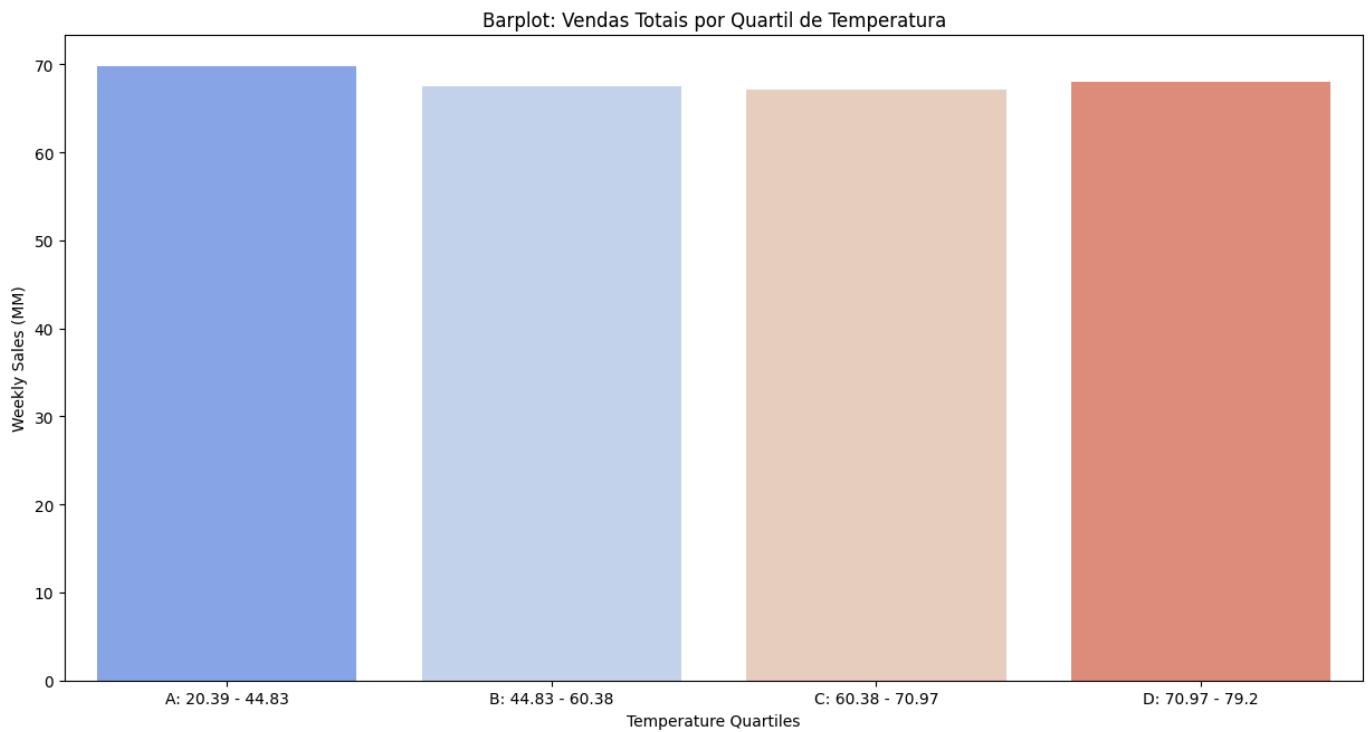
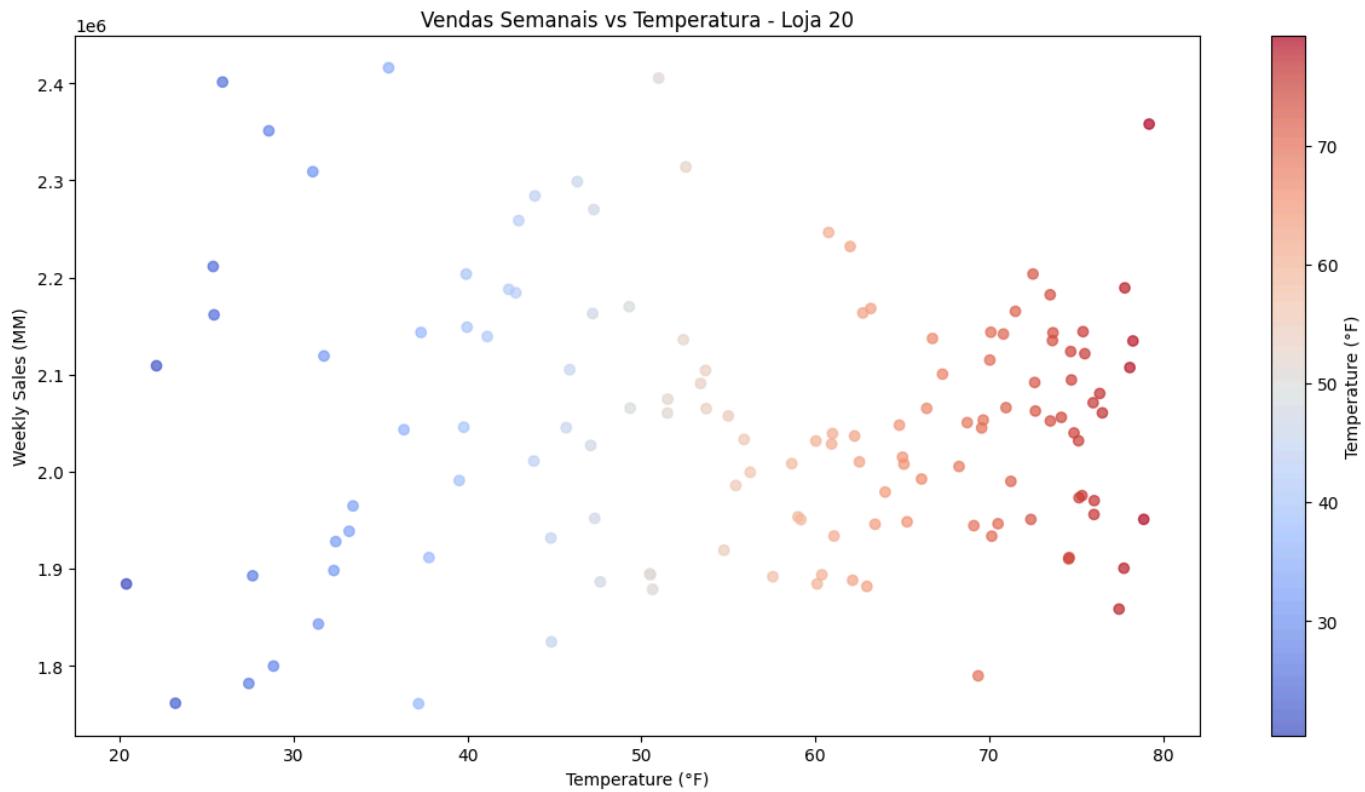
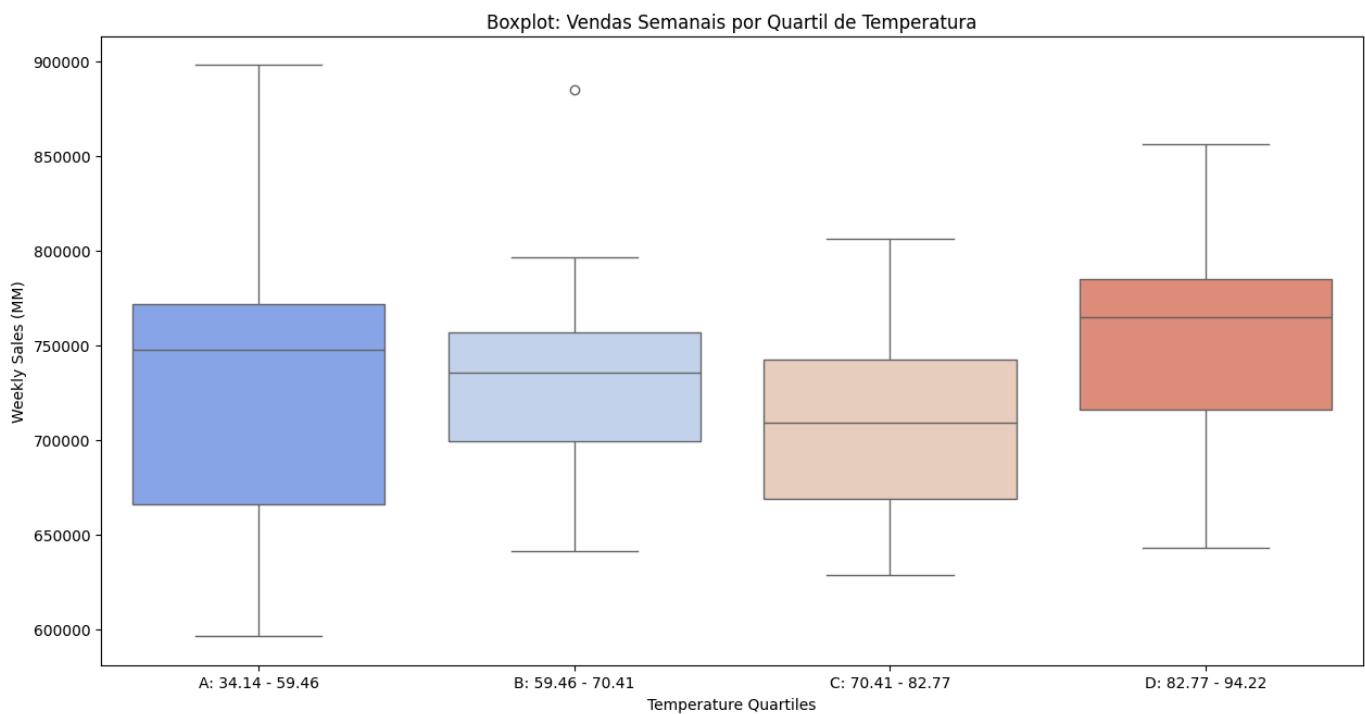


Gráfico de Vendas Semanais vs Temperatura - Loja 20 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 21 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 21 (Sem Outliers)

Barplot: Vendas Totais por Quartil de Temperatura

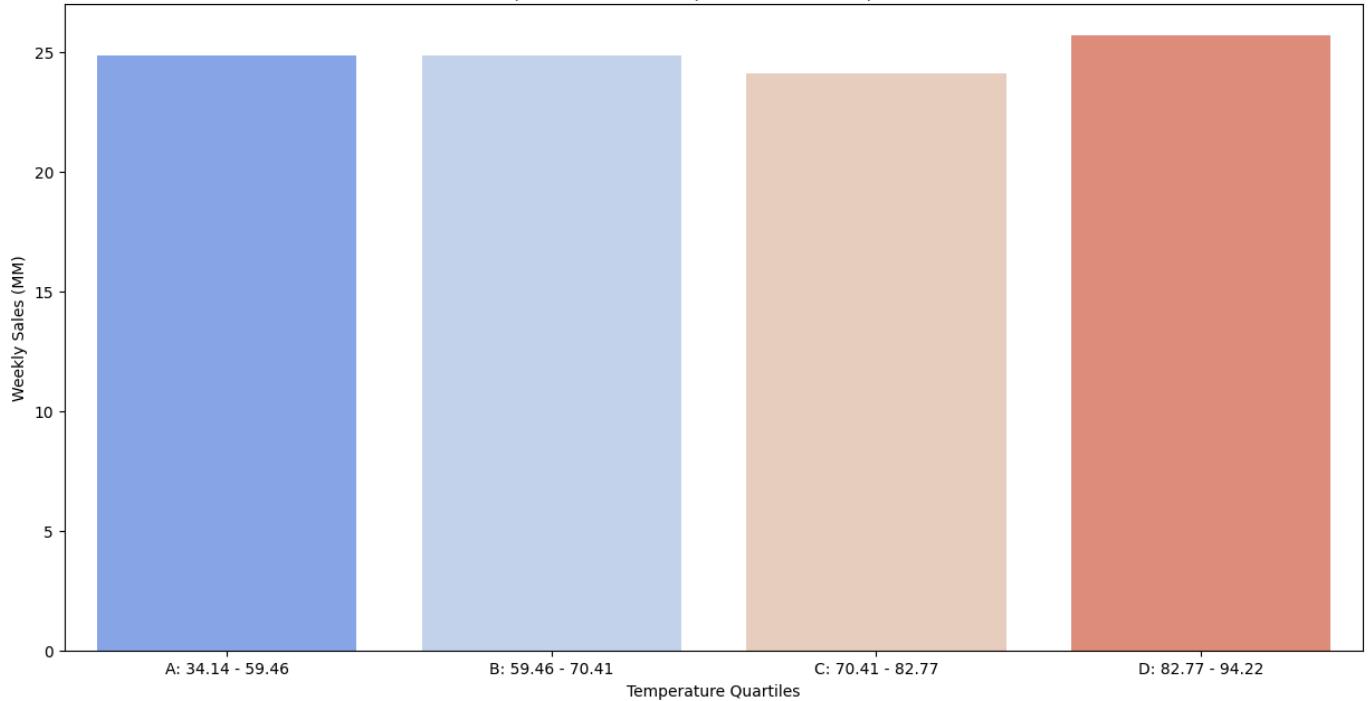
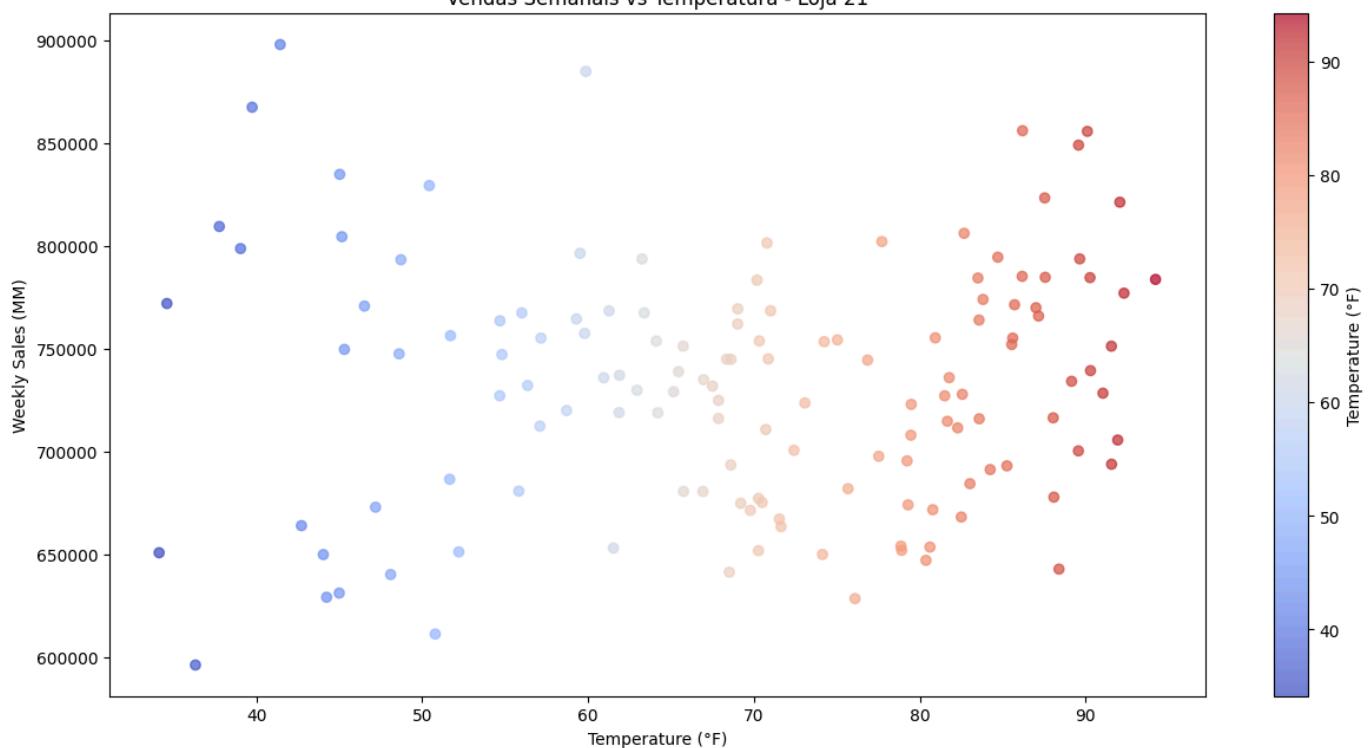
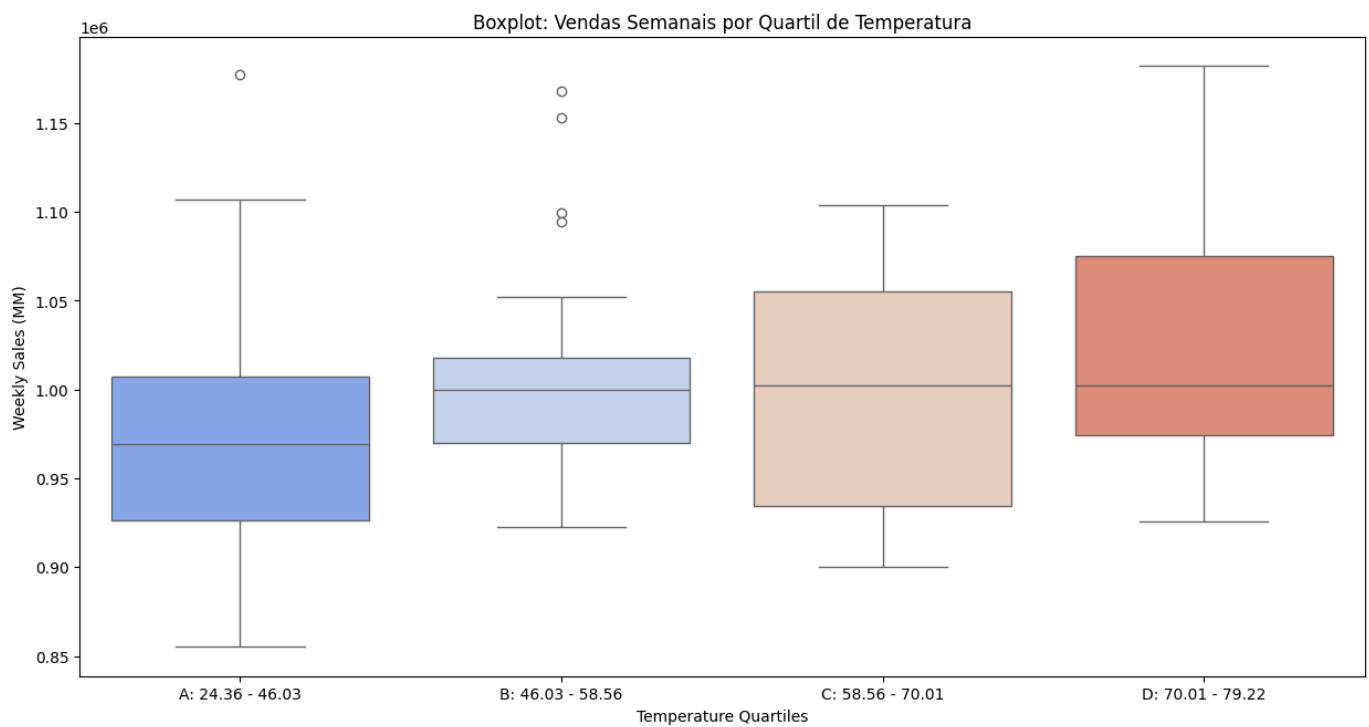


Gráfico de Vendas Semanais vs Temperatura - Loja 21 (Sem Outliers)

Vendas Semanais vs Temperatura - Loja 21



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 22 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 22 (Sem Outliers)

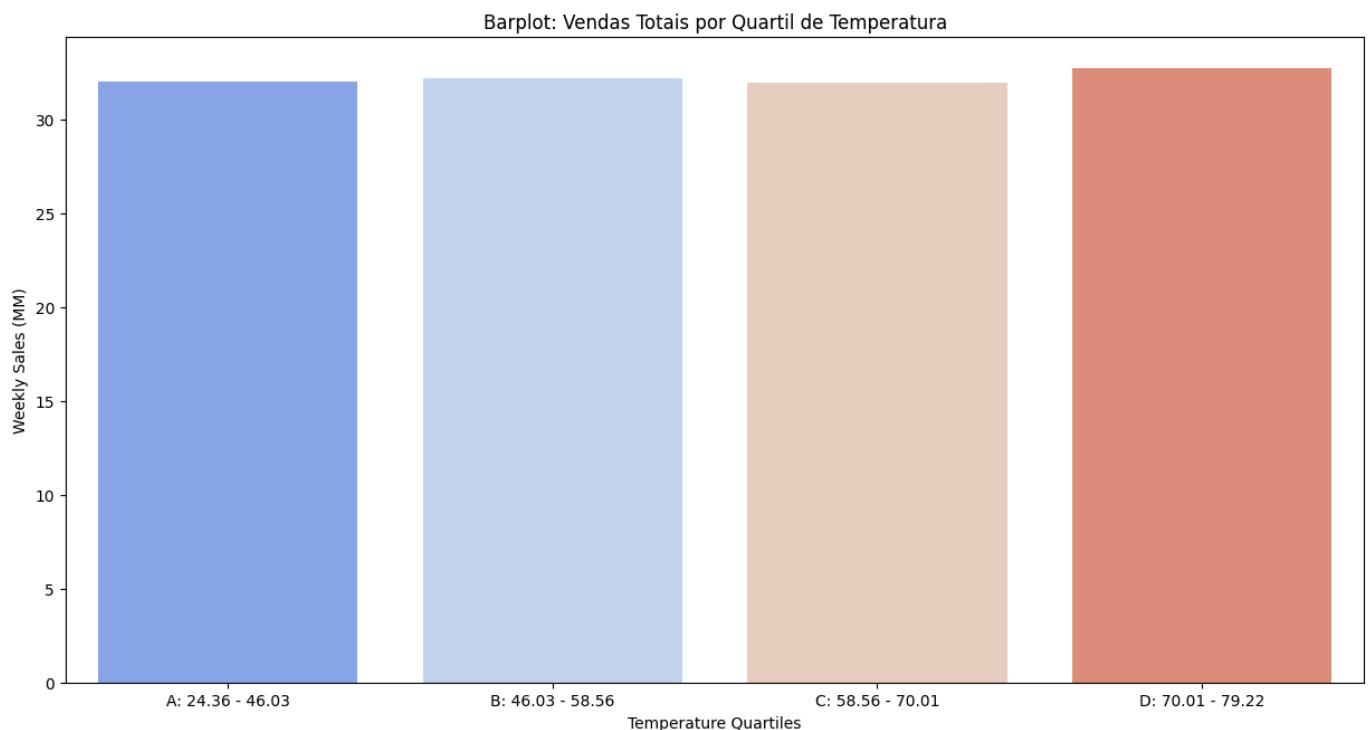
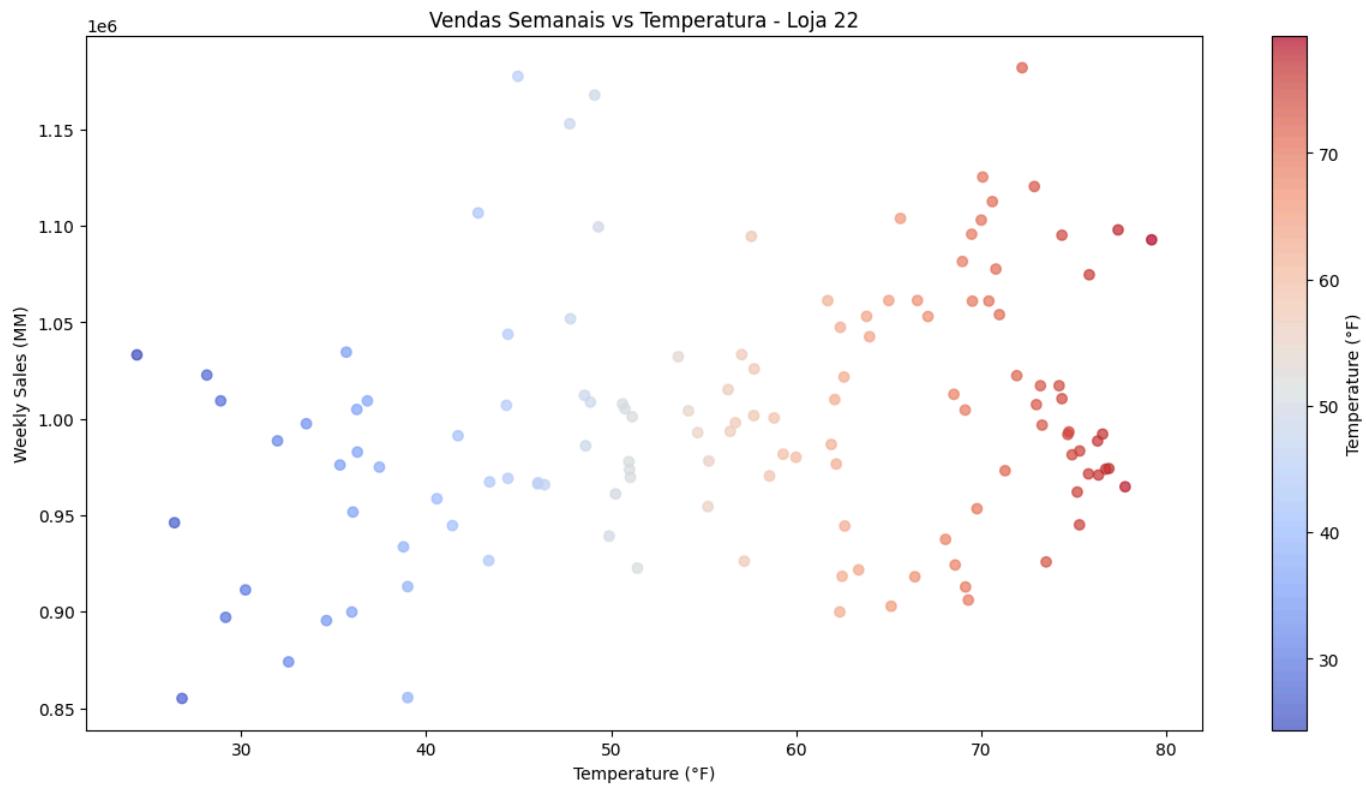
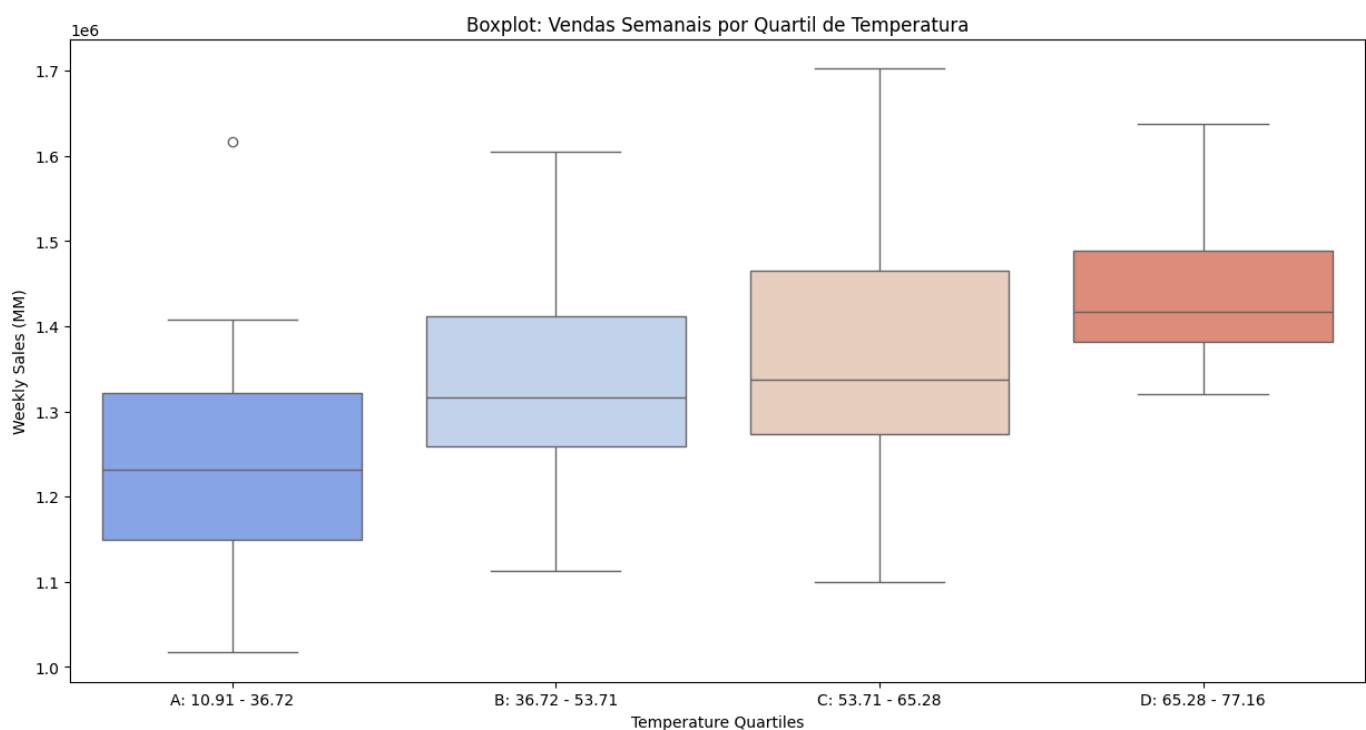


Gráfico de Vendas Semanais vs Temperatura - Loja 22 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 23 (Sem Outliers)



Barplot: Vendas Totais por Quartil de Temperatura

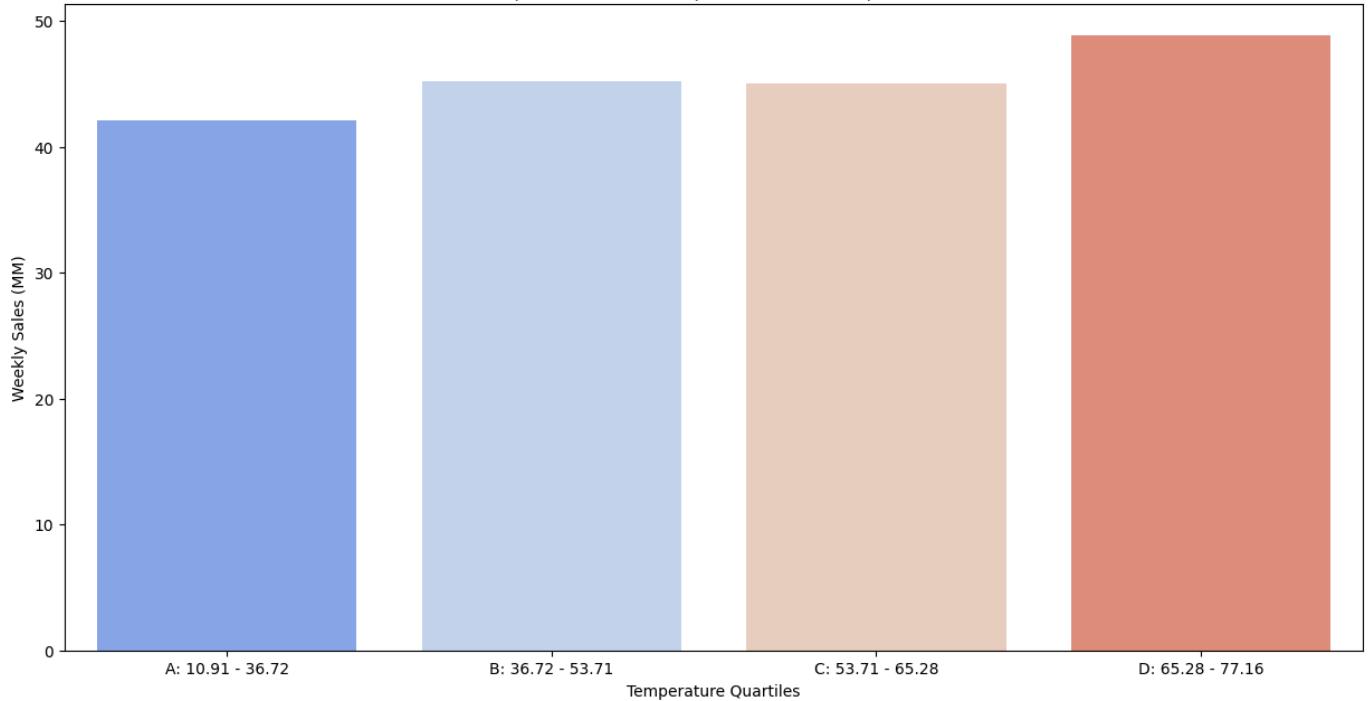
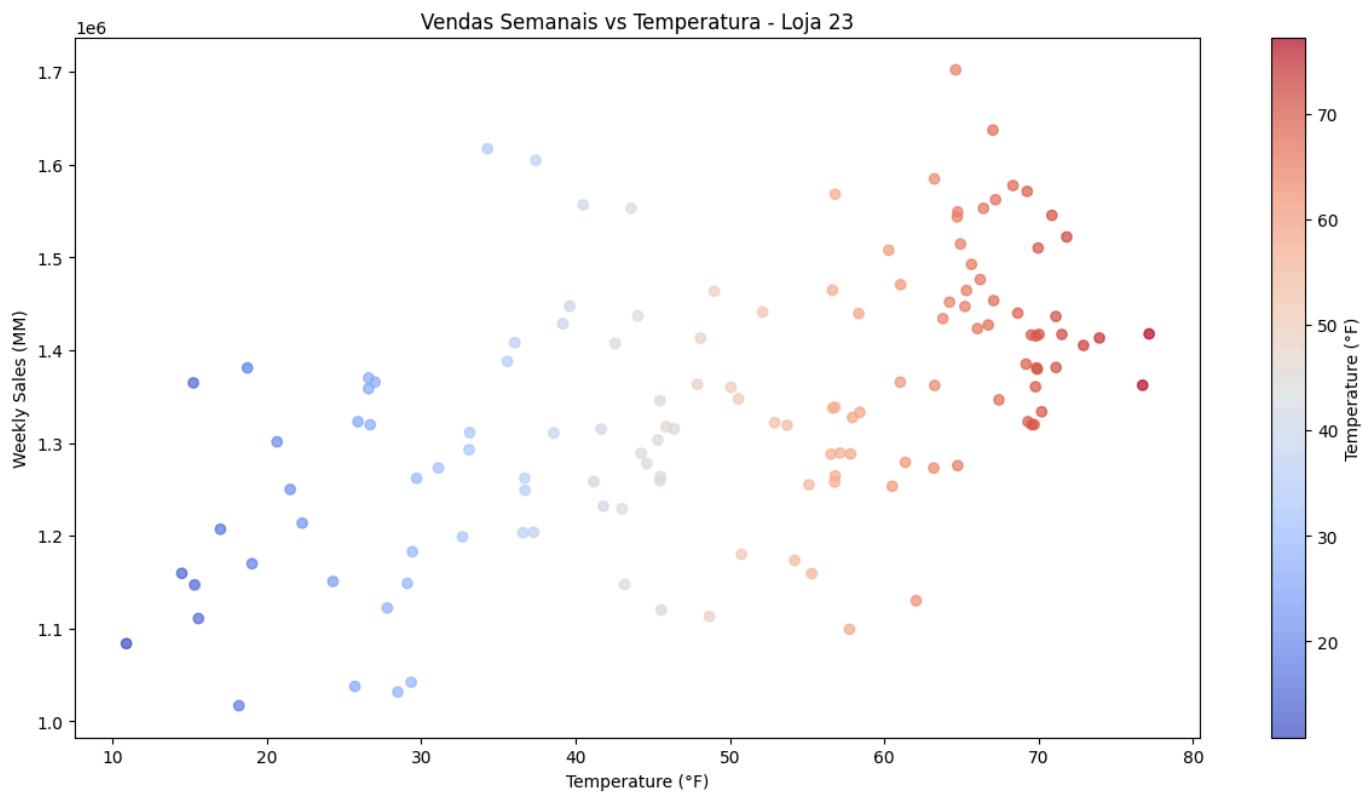
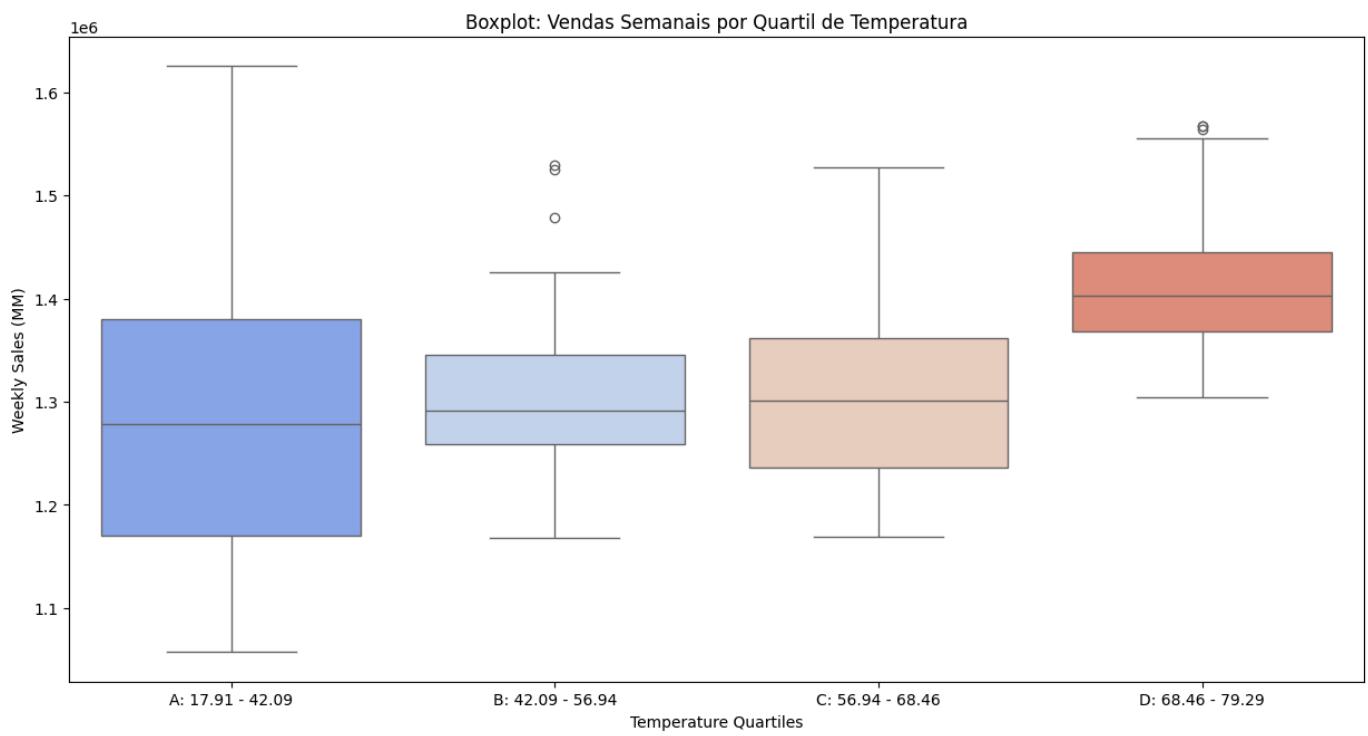


Gráfico de Vendas Semanais vs Temperatura - Loja 23 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 24 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 24 (Sem Outliers)

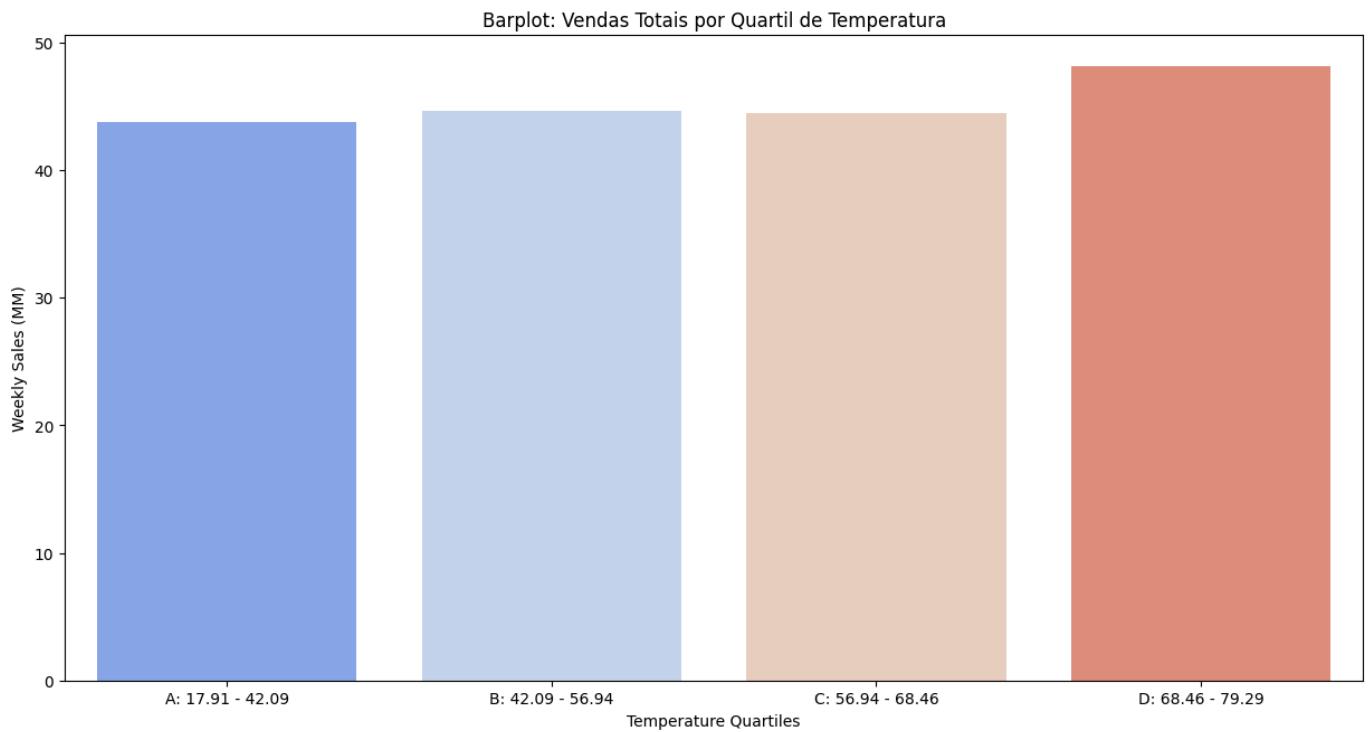
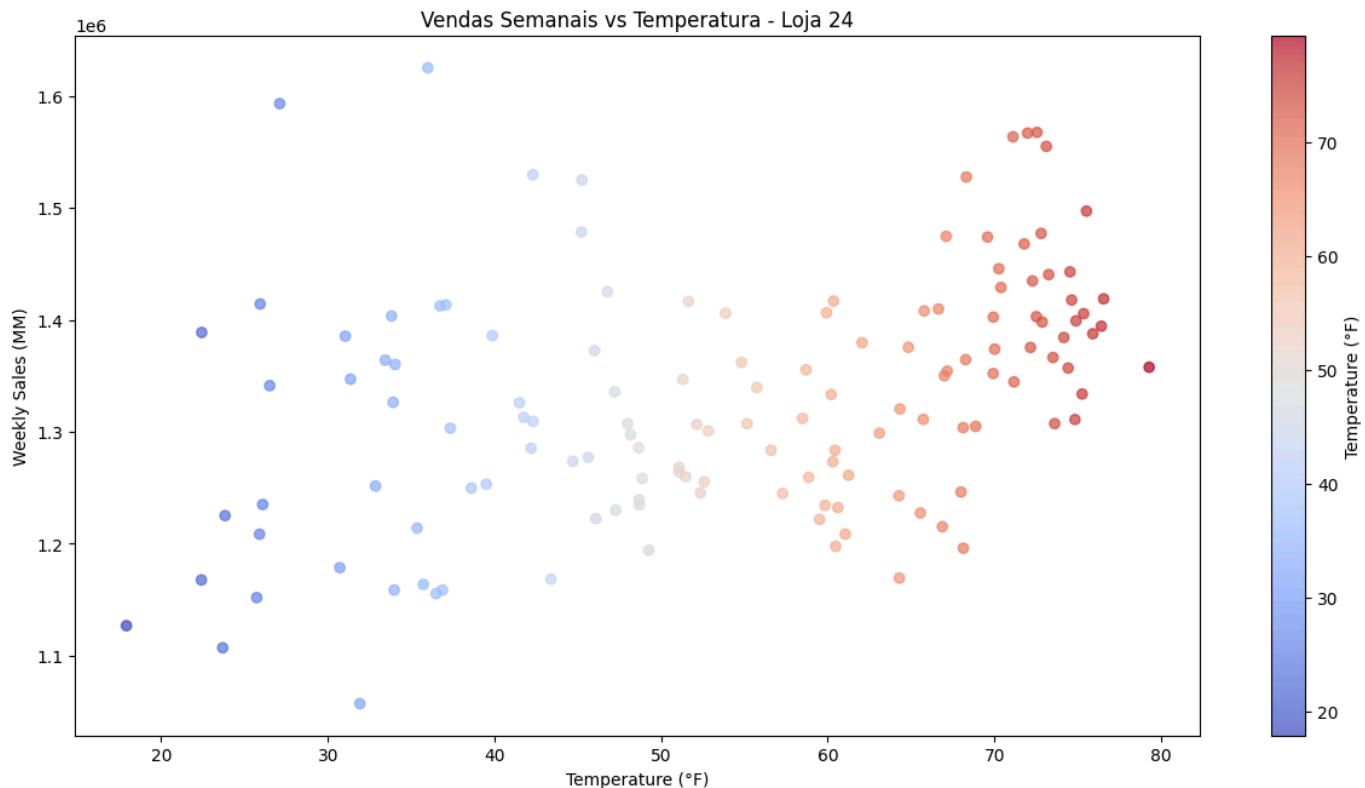
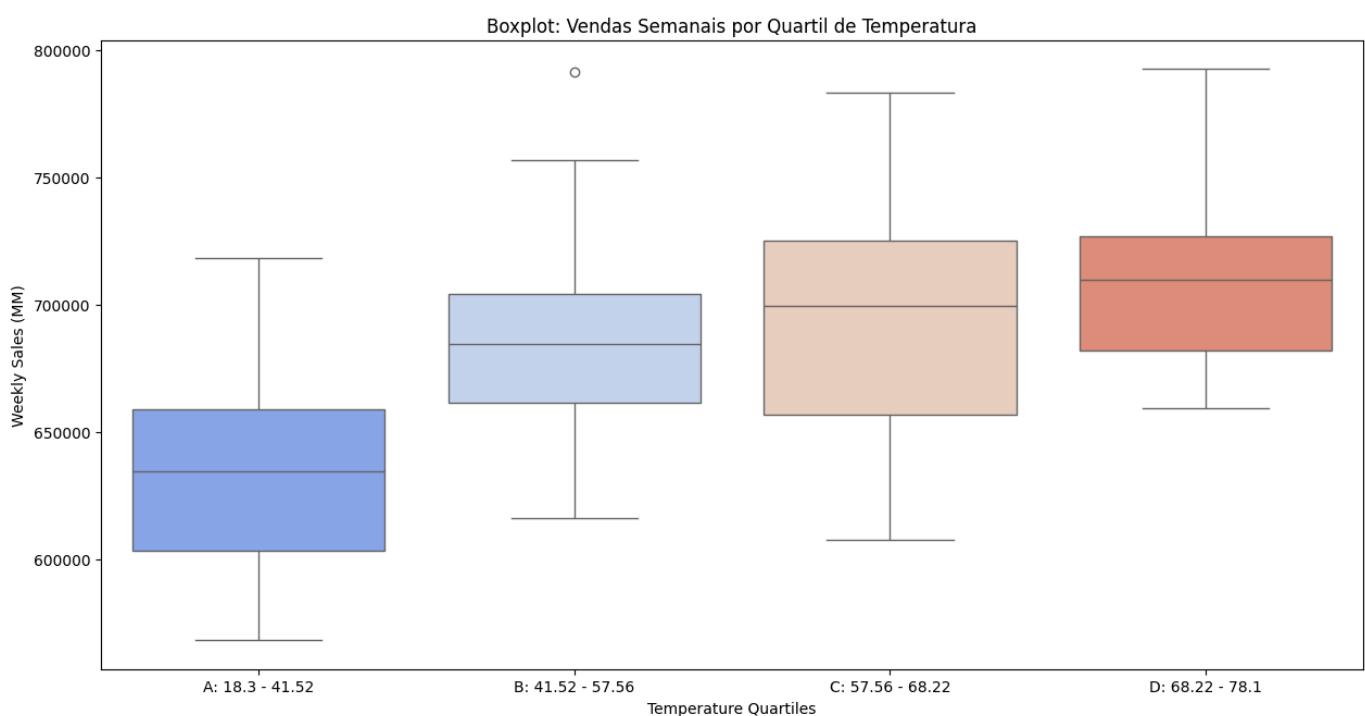


Gráfico de Vendas Semanais vs Temperatura - Loja 24 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 25 (Sem Outliers)



Barplot: Vendas Totais por Quartil de Temperatura

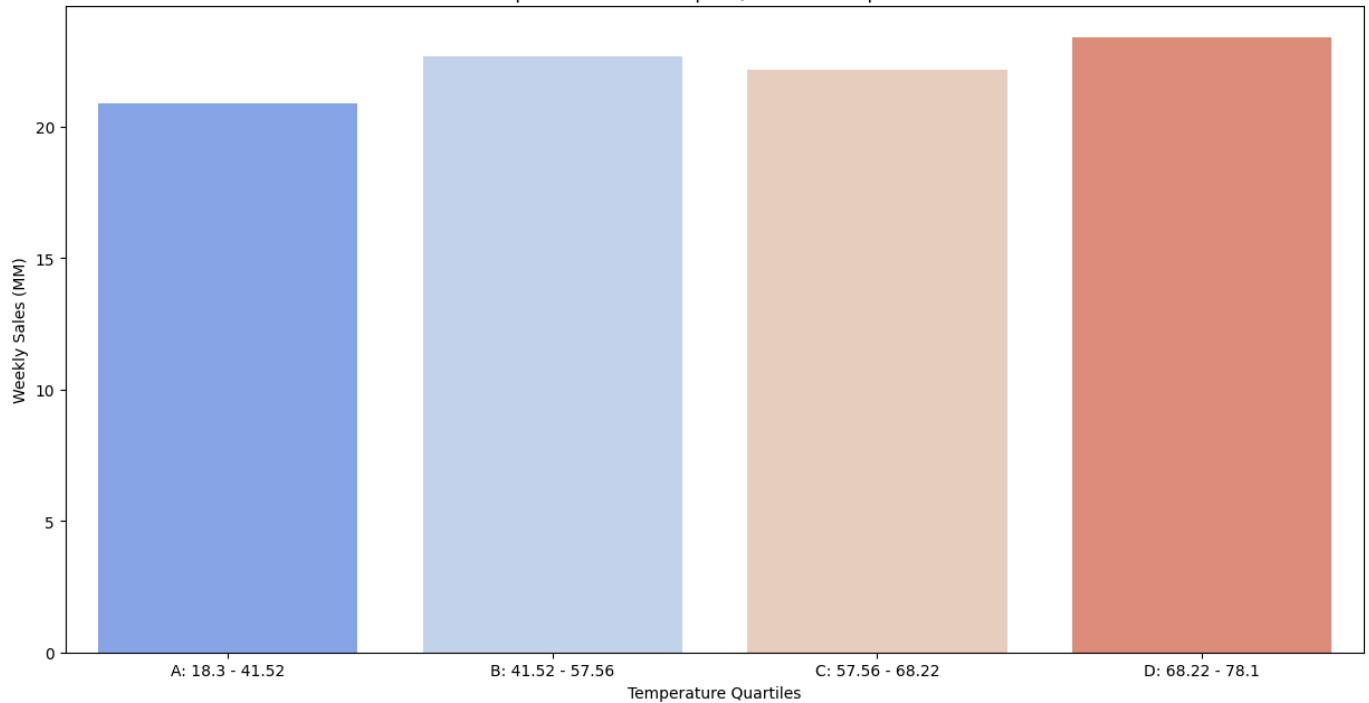
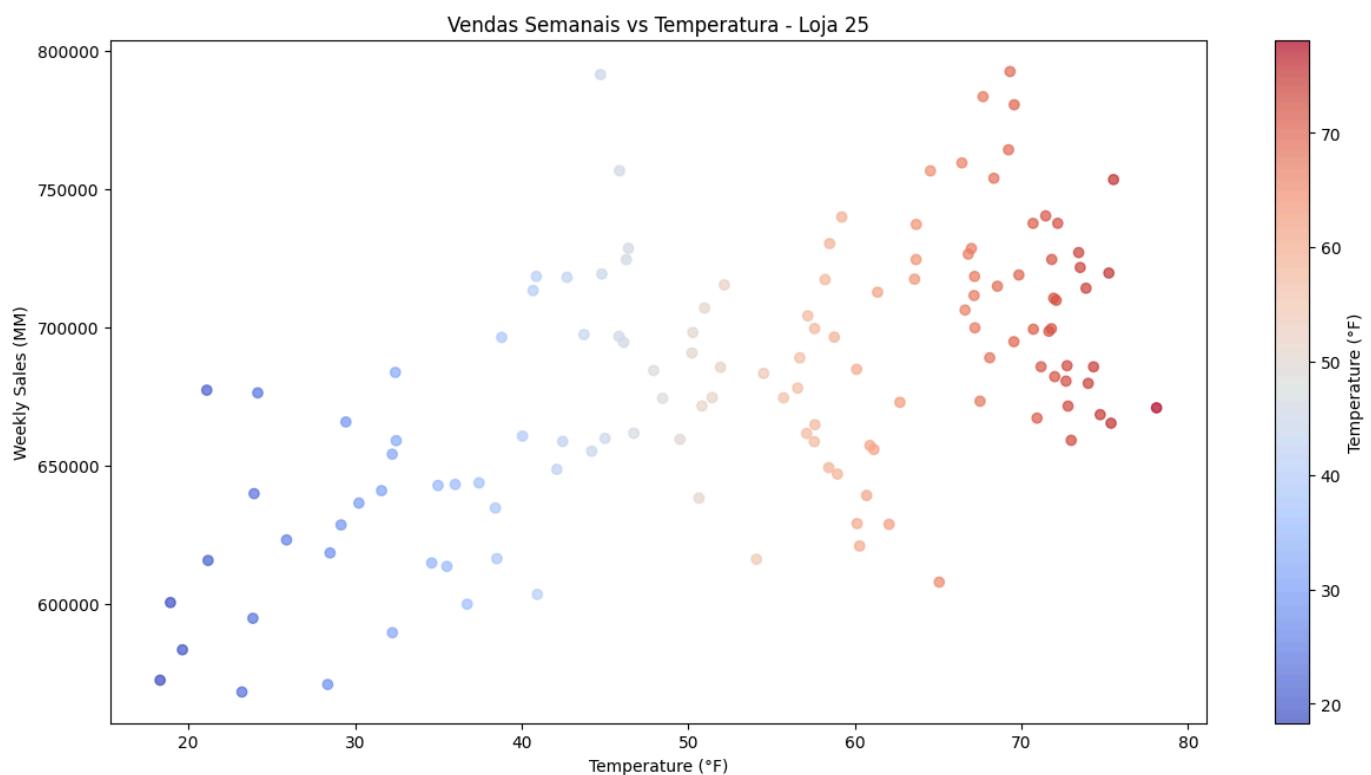
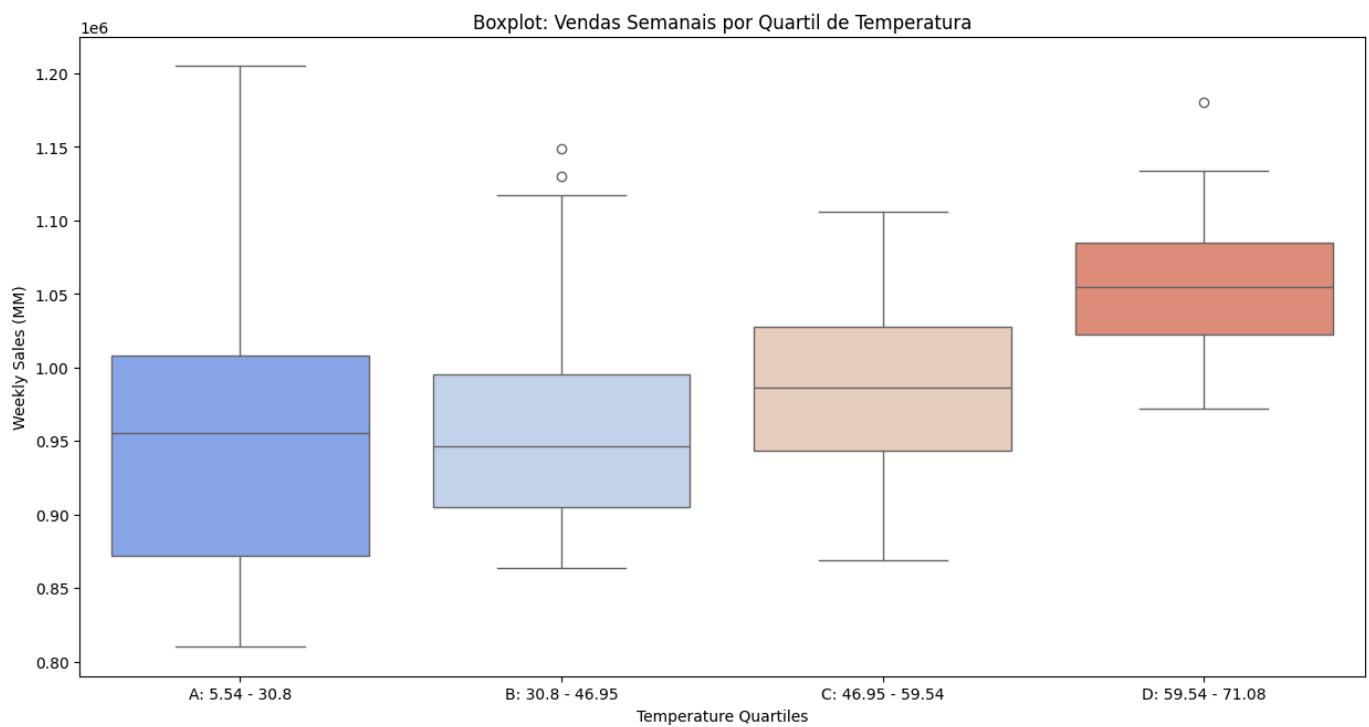


Gráfico de Vendas Semanais vs Temperatura - Loja 25 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 26 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 26 (Sem Outliers)

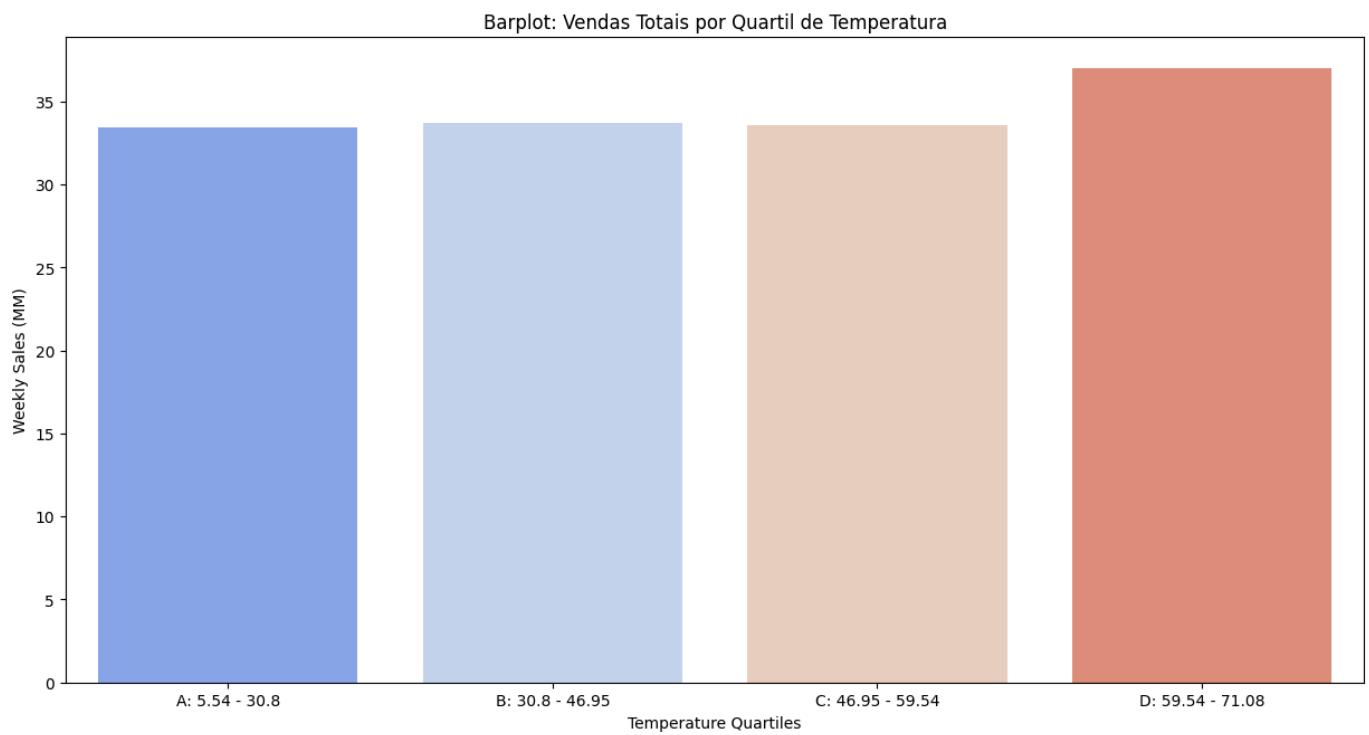
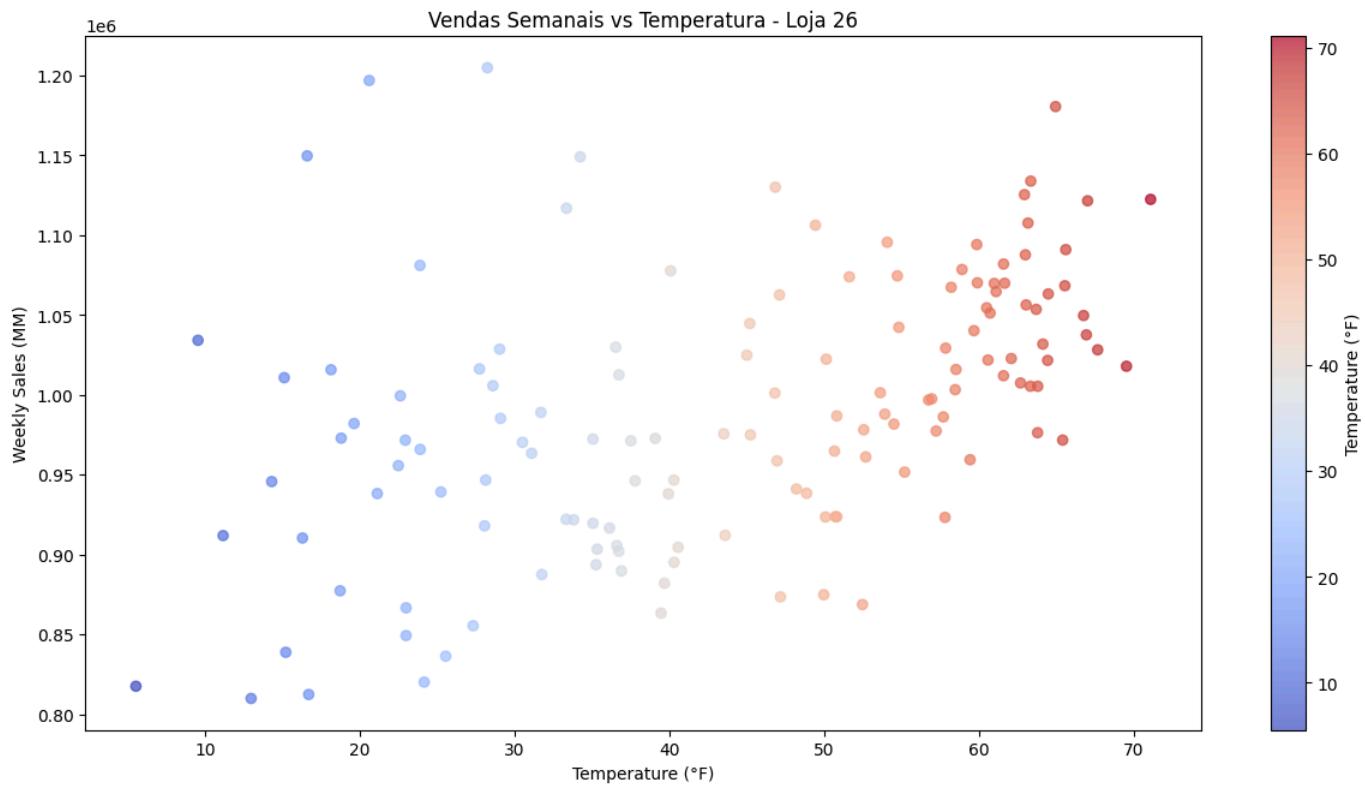
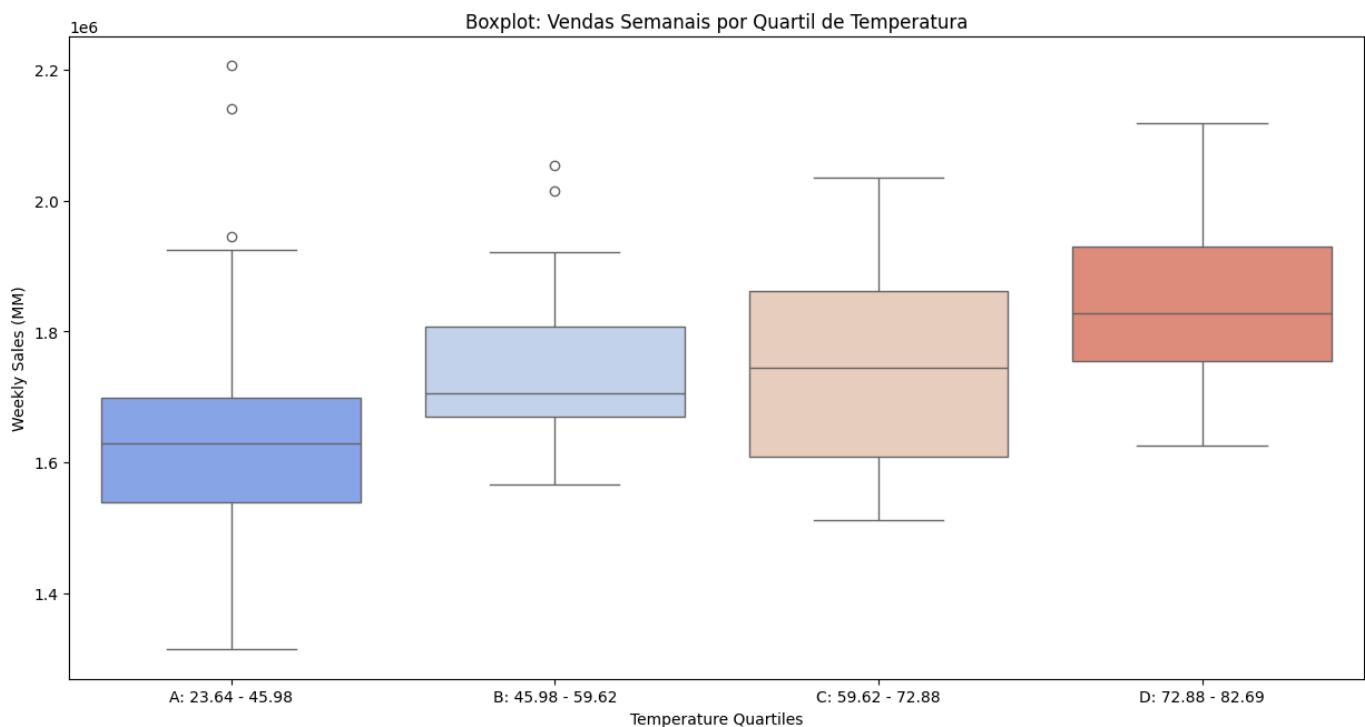


Gráfico de Vendas Semanais vs Temperatura - Loja 26 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 27 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 27 (Sem Outliers)

Barplot: Vendas Totais por Quartil de Temperatura

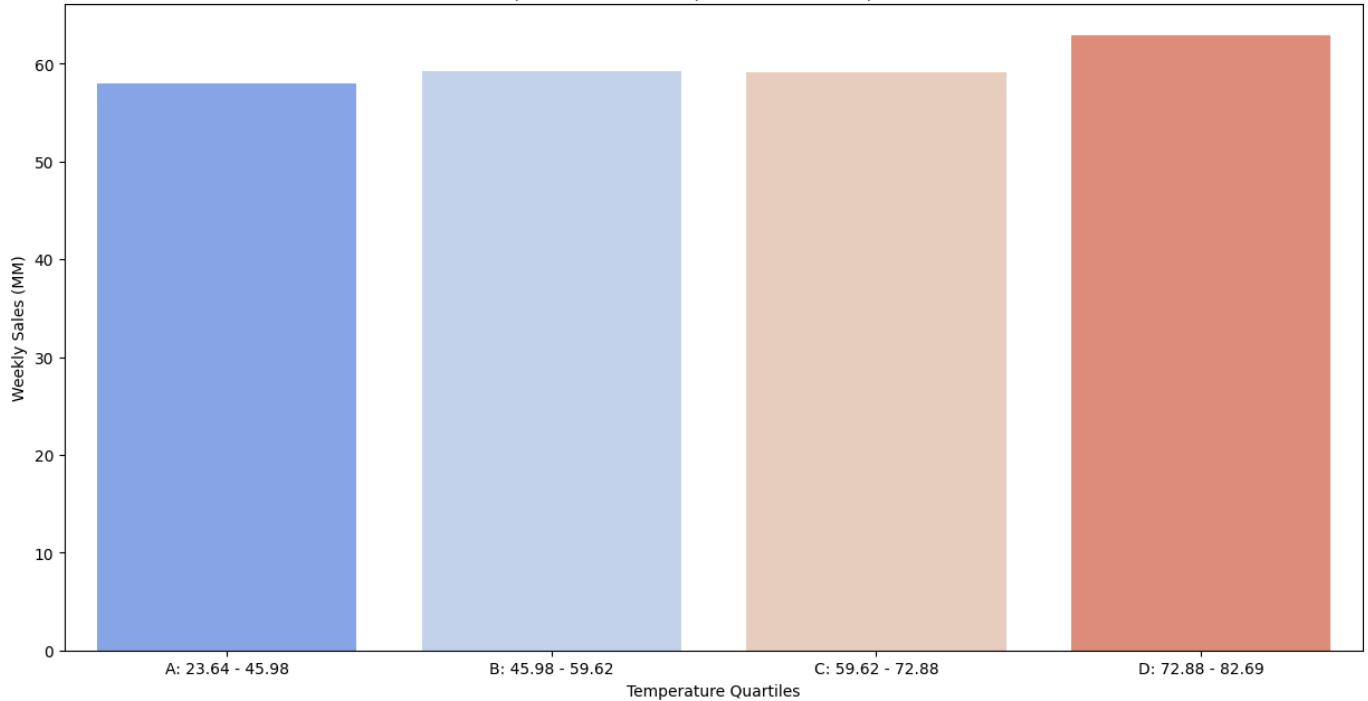
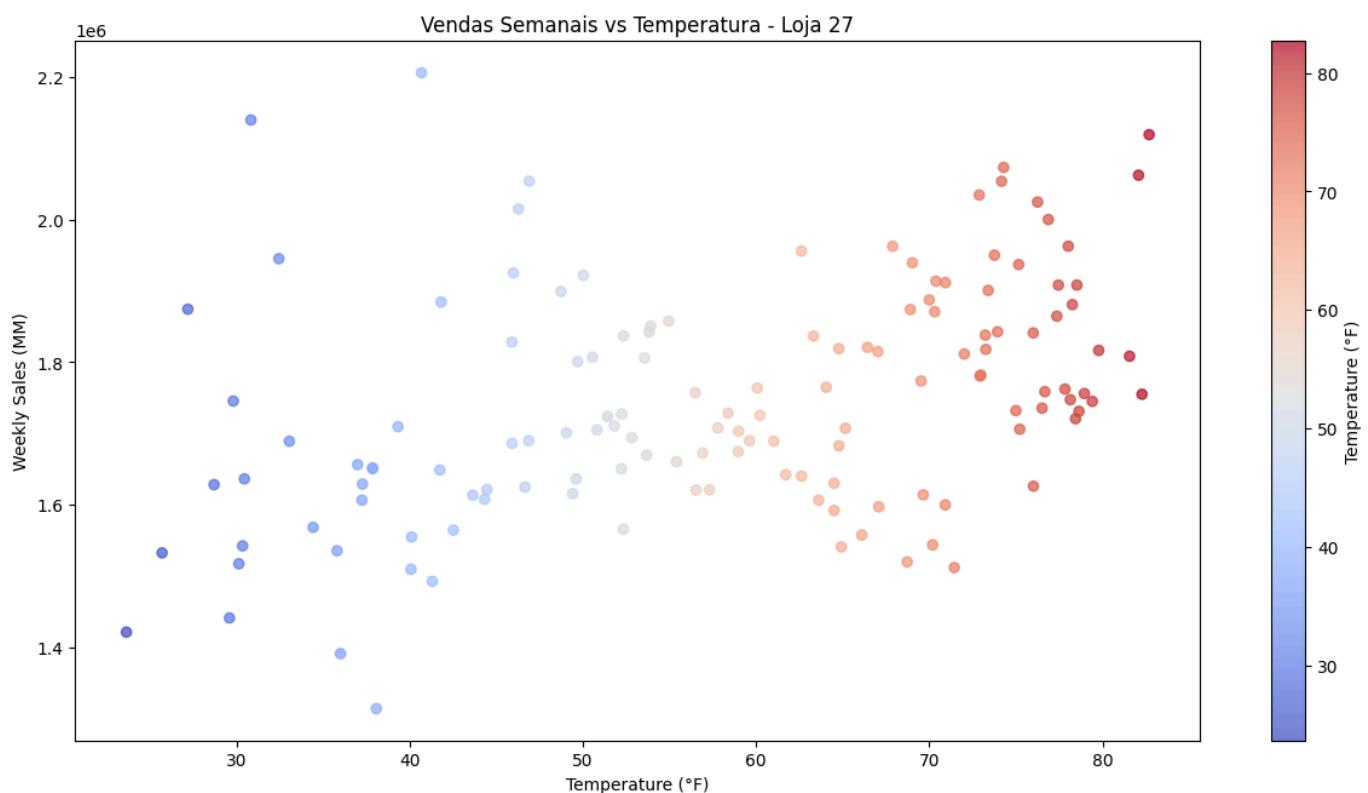
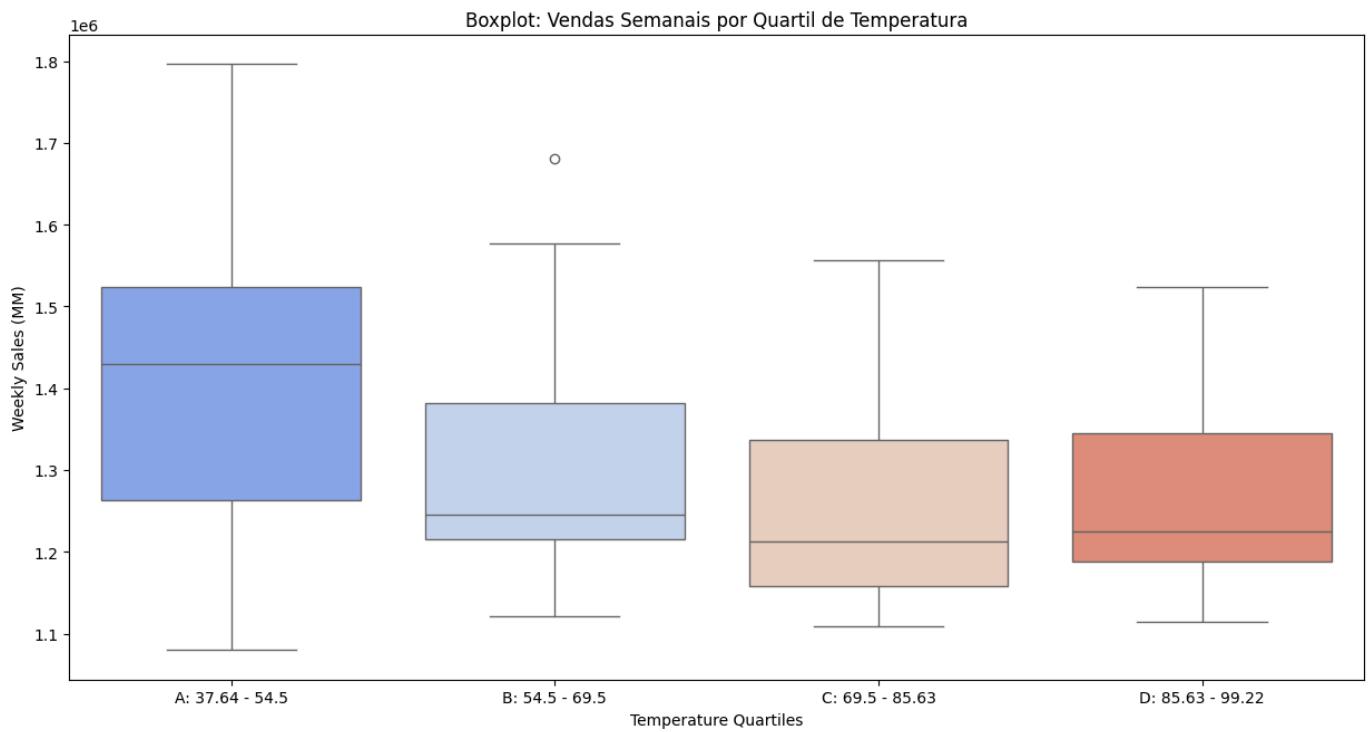


Gráfico de Vendas Semanais vs Temperatura - Loja 27 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 28 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 28 (Sem Outliers)

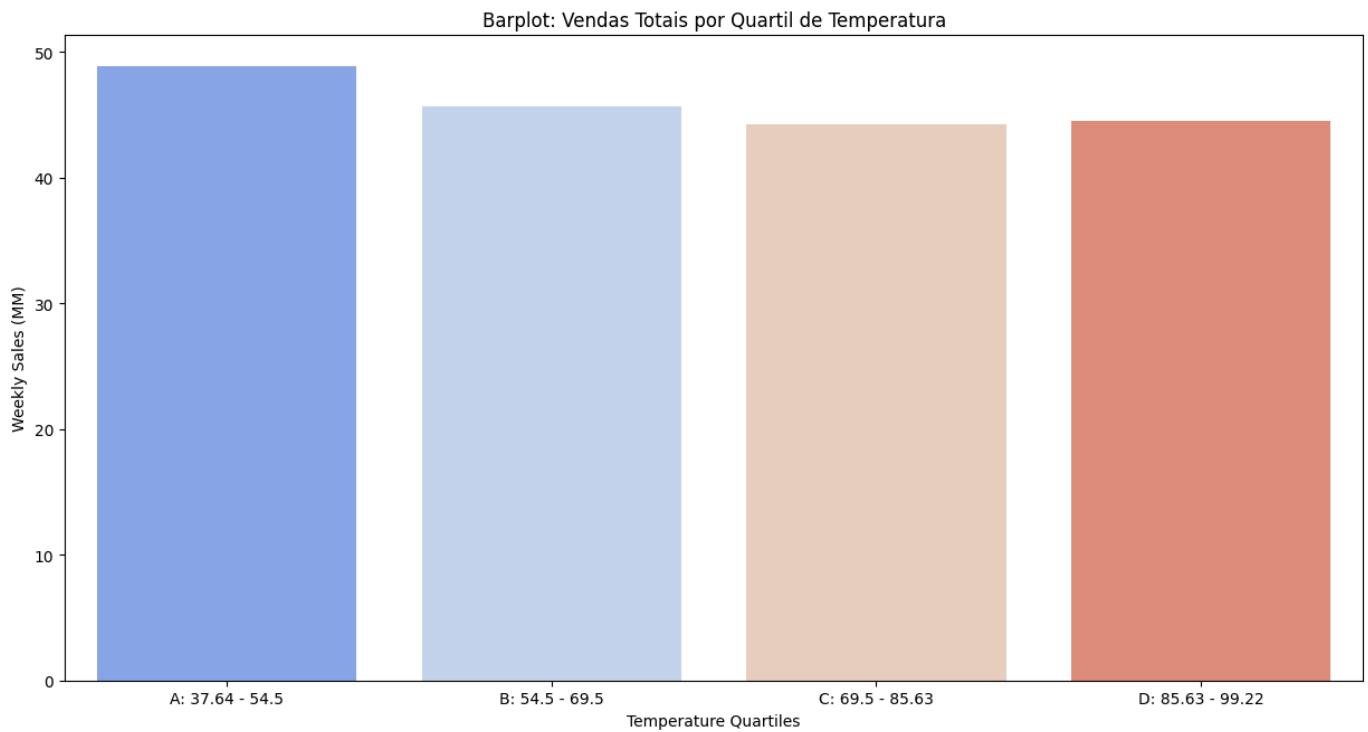
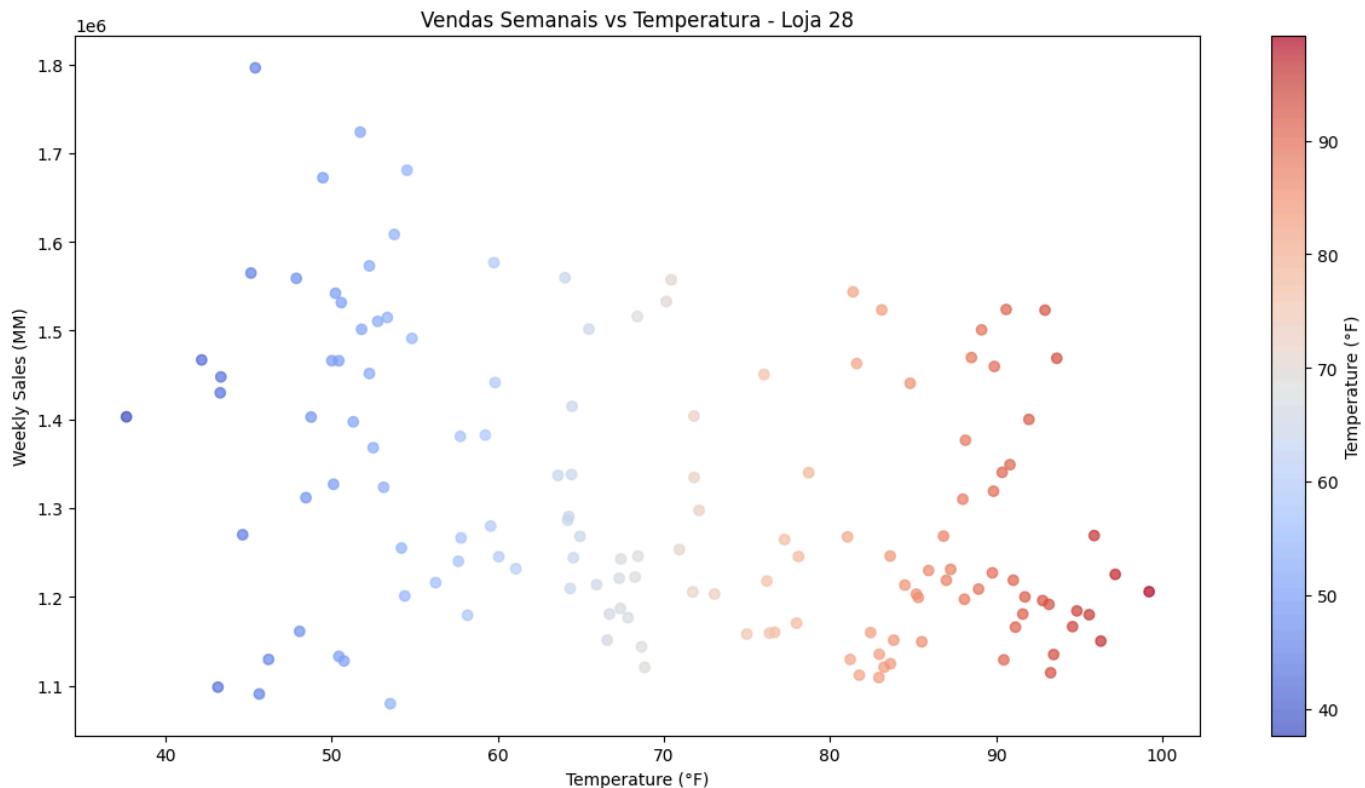
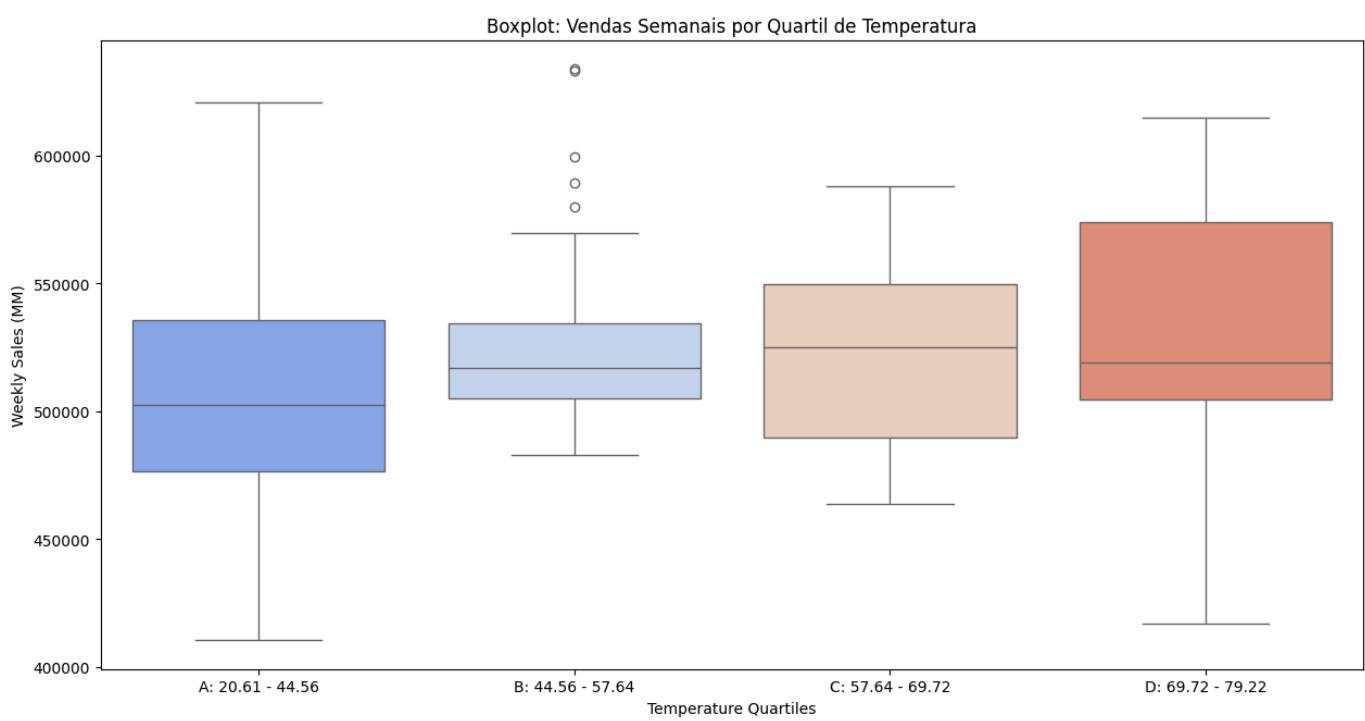


Gráfico de Vendas Semanais vs Temperatura - Loja 28 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 29 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 29 (Sem Outliers)

Barplot: Vendas Totais por Quartil de Temperatura

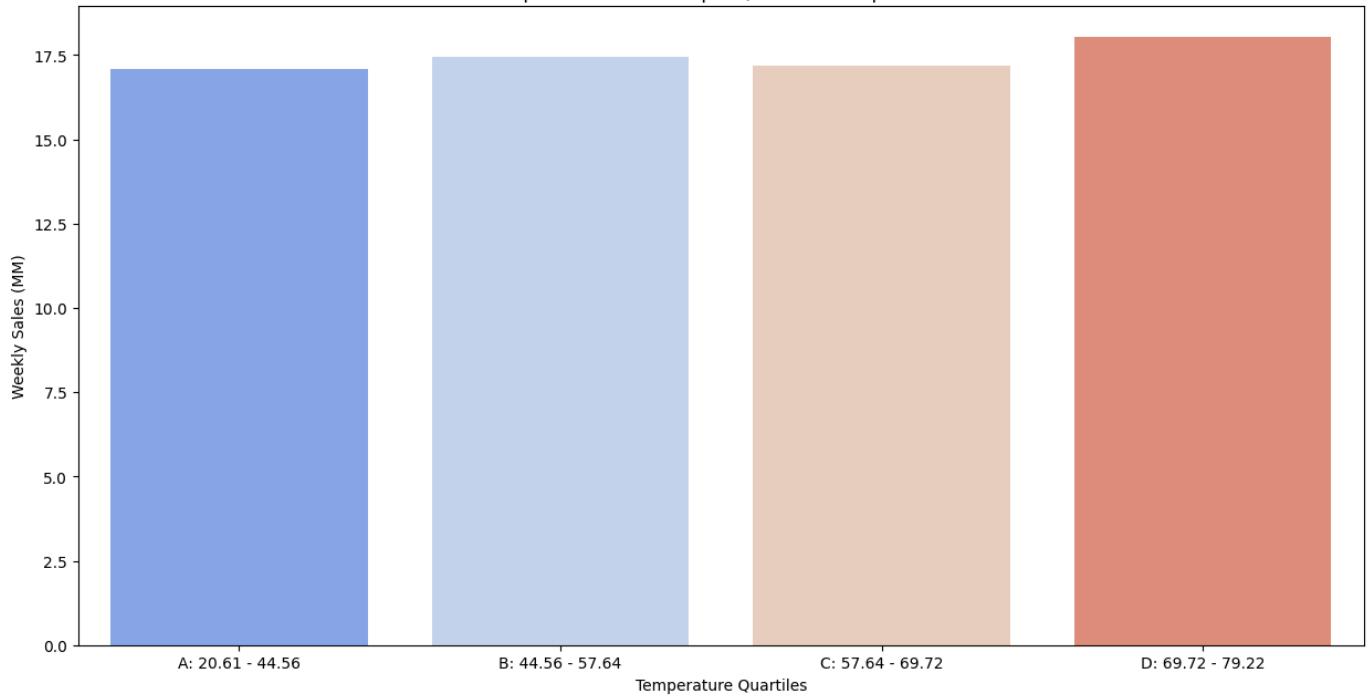
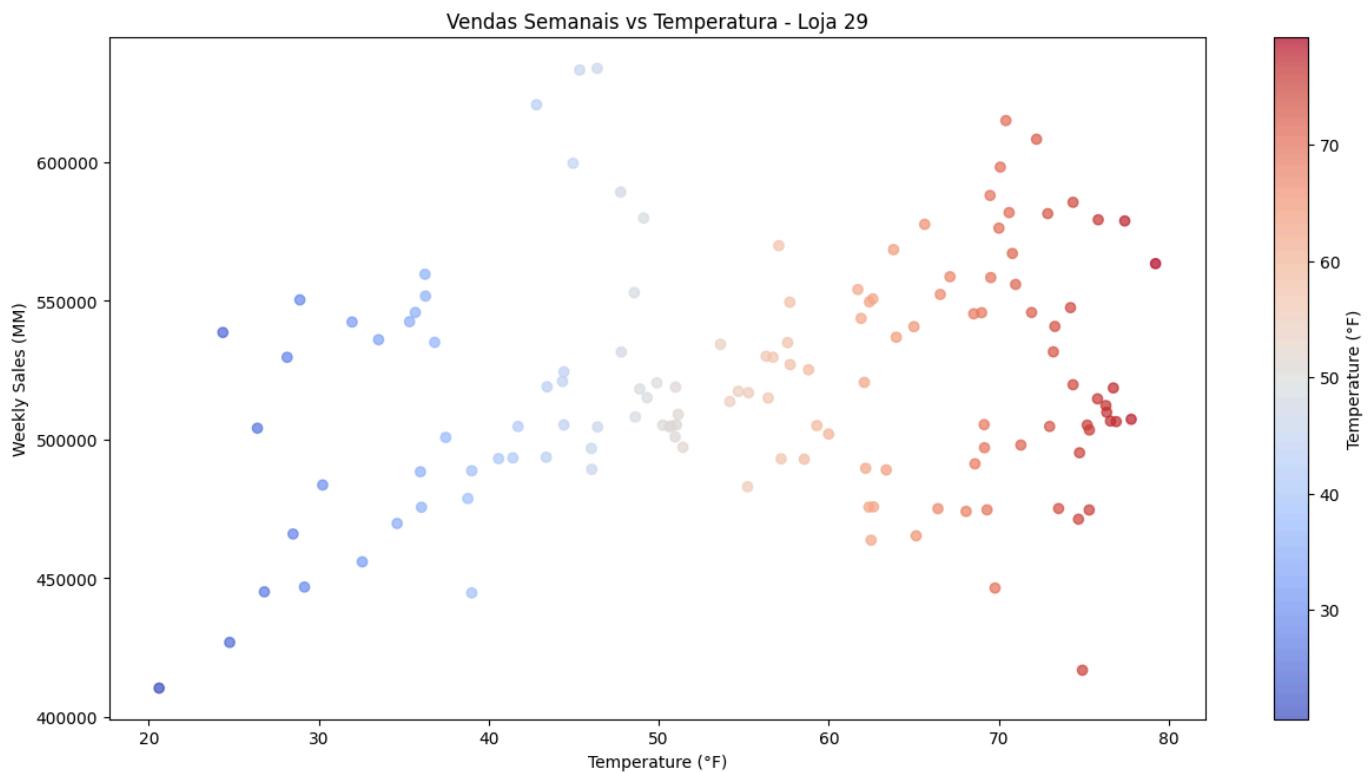
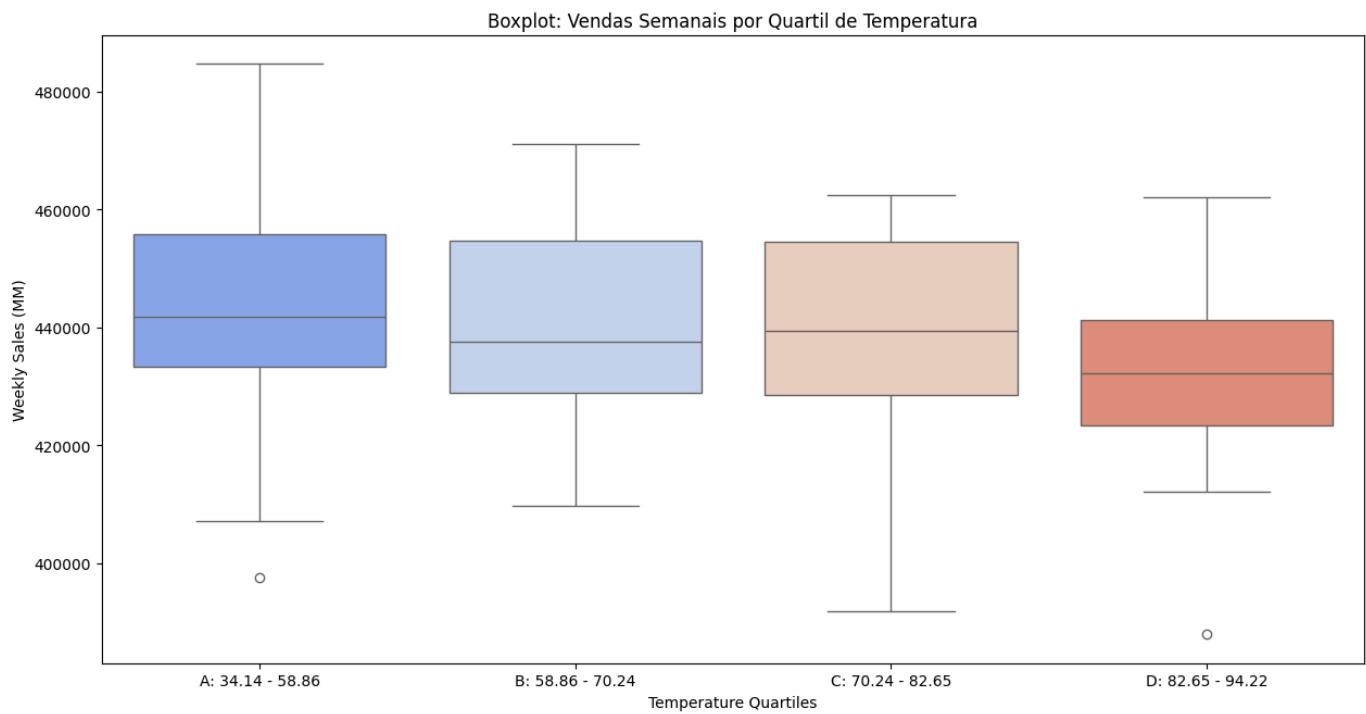


Gráfico de Vendas Semanais vs Temperatura - Loja 29 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 30 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 30 (Sem Outliers)

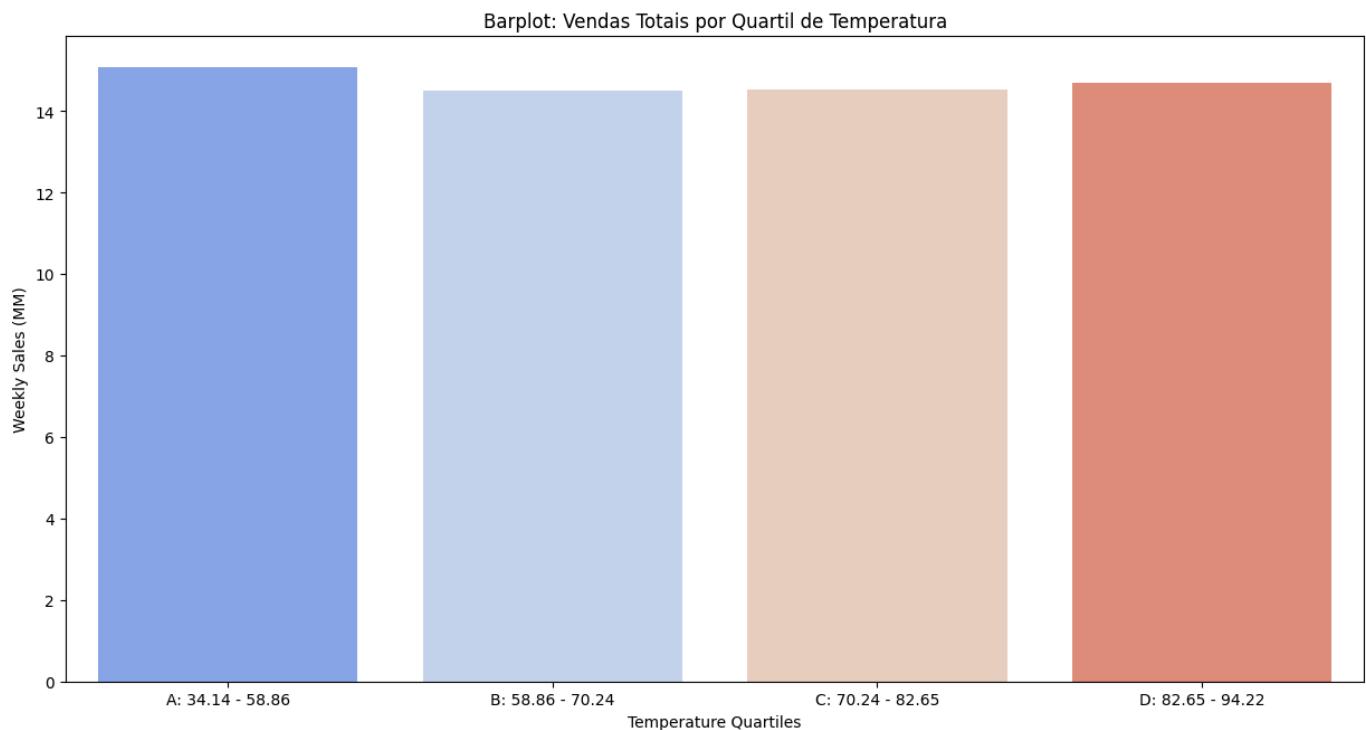
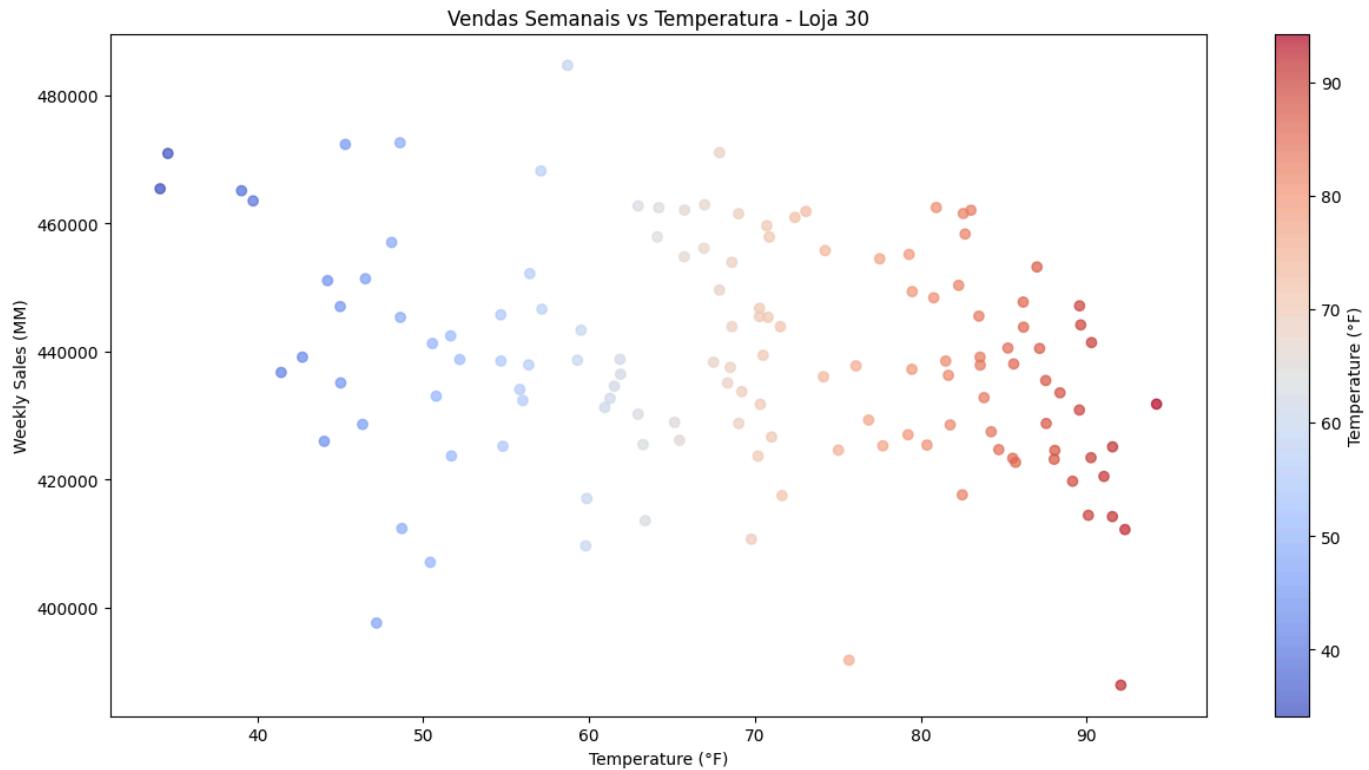
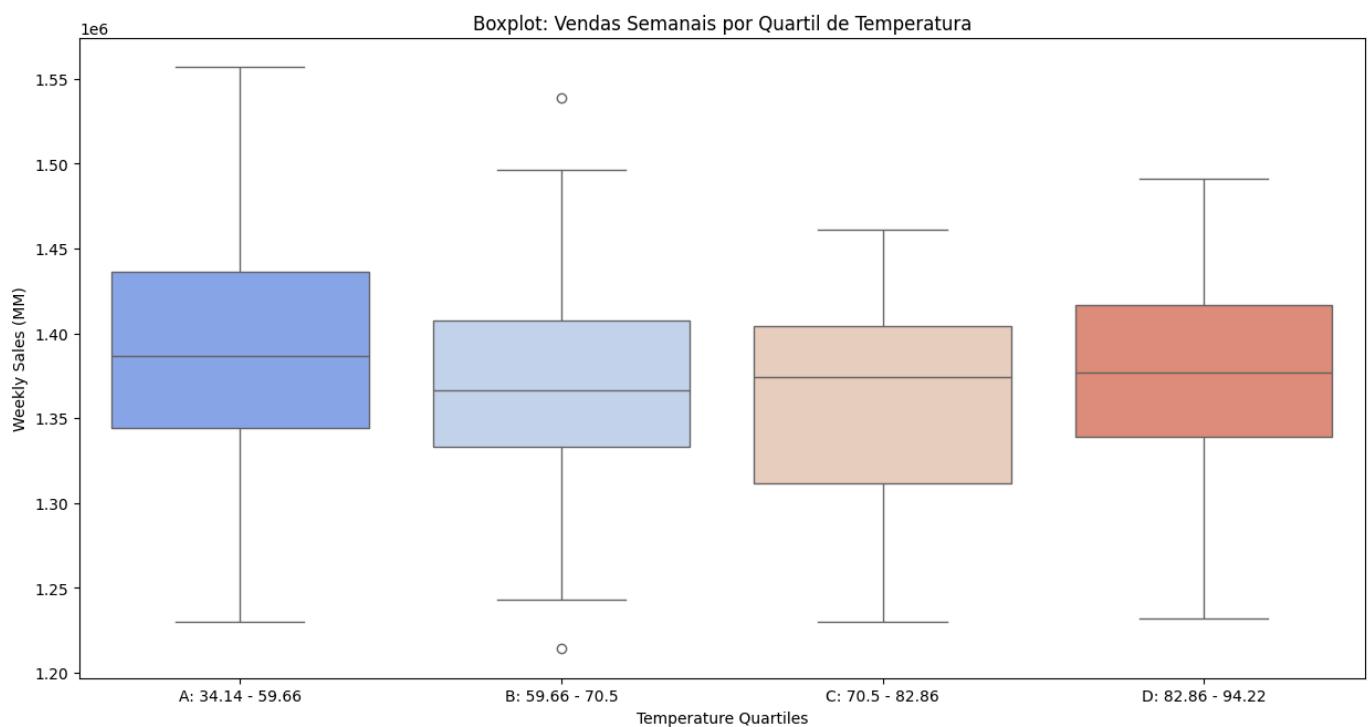


Gráfico de Vendas Semanais vs Temperatura - Loja 30 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 31 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 31 (Sem Outliers)

Barplot: Vendas Totais por Quartil de Temperatura

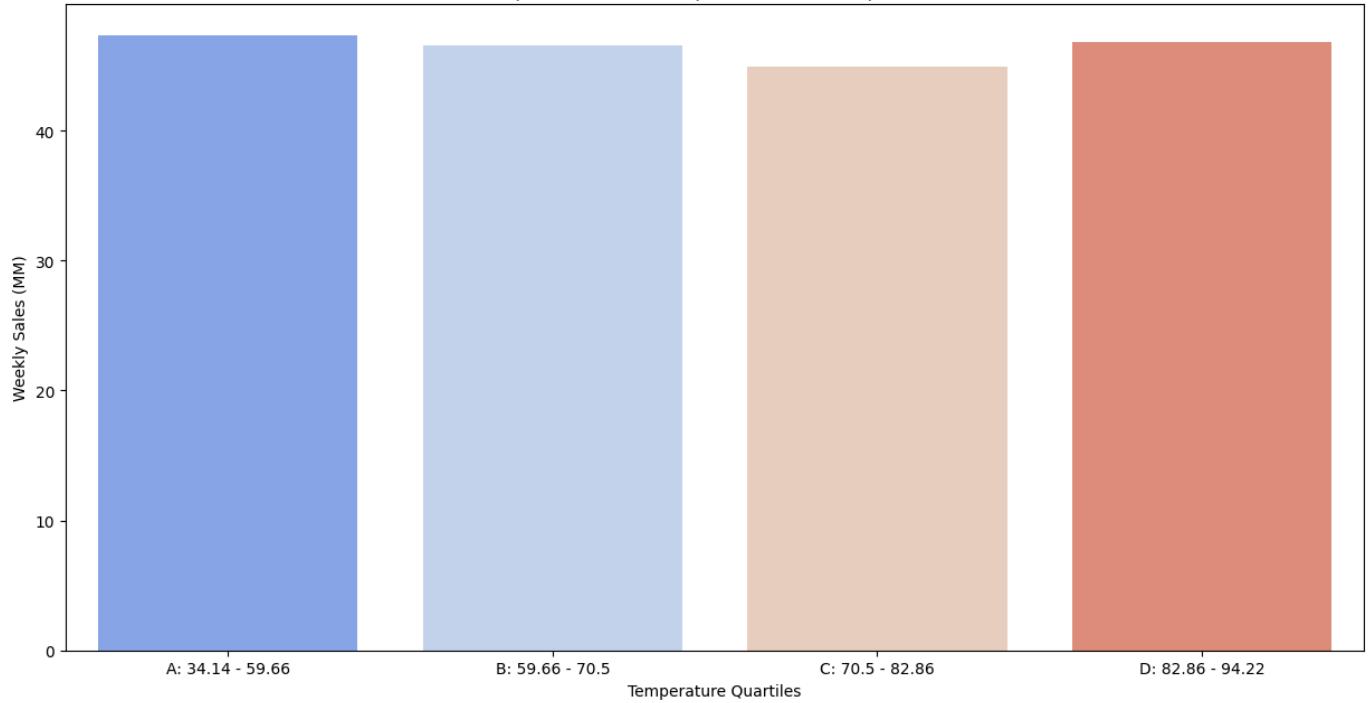
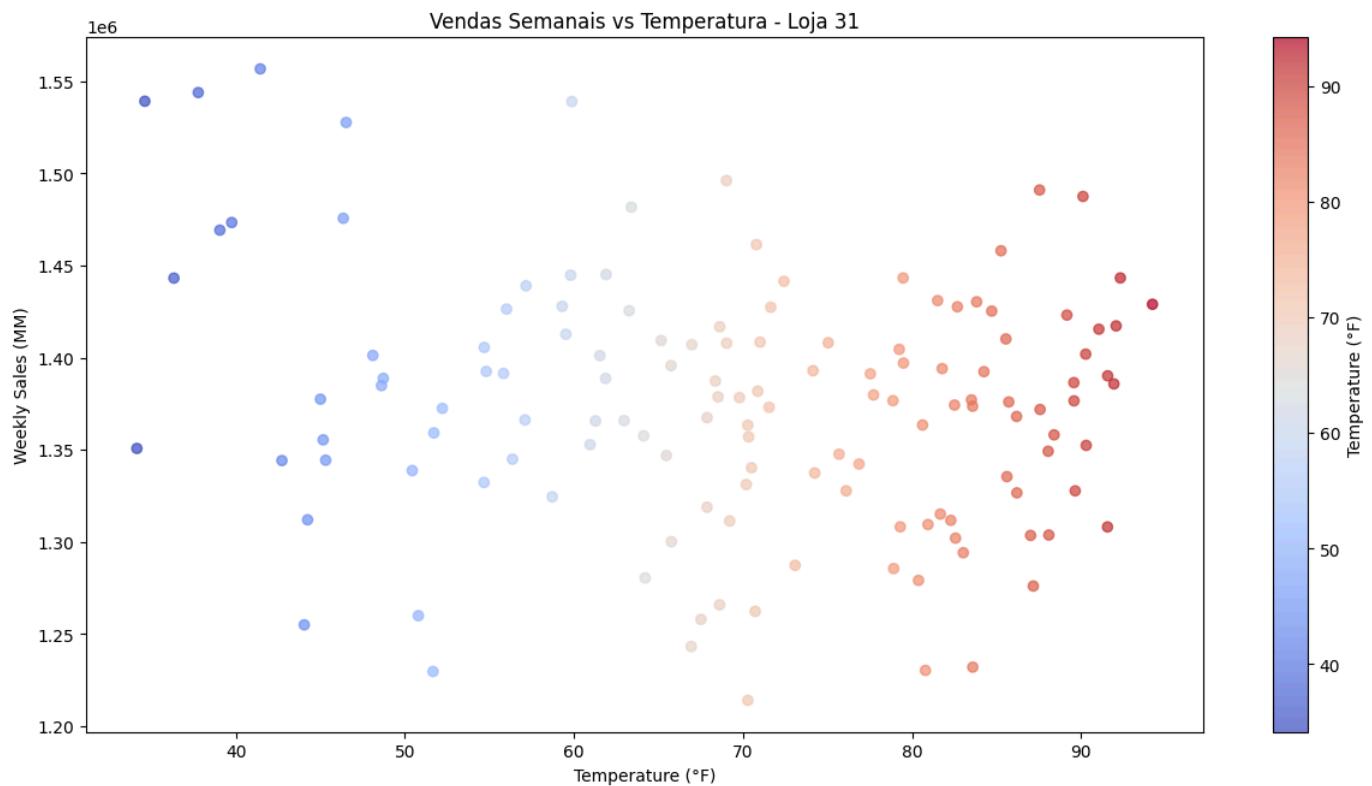
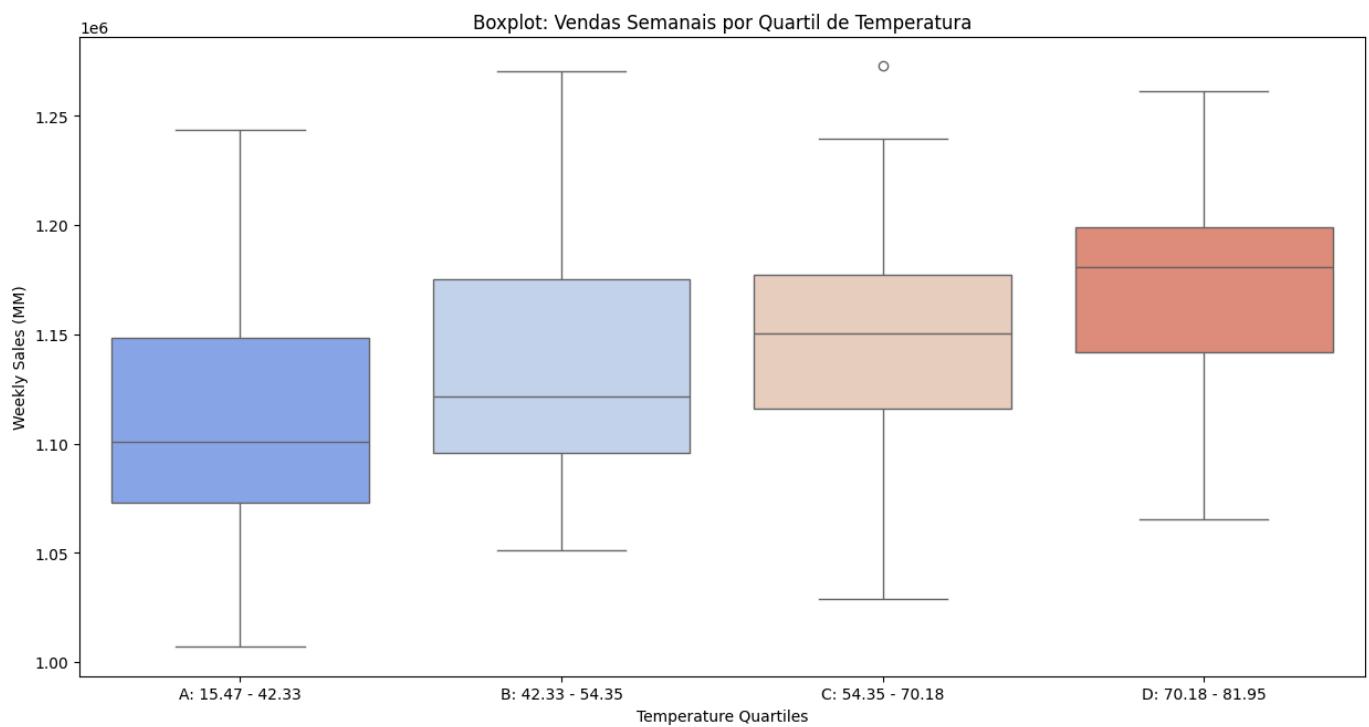


Gráfico de Vendas Semanais vs Temperatura - Loja 31 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 32 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 32 (Sem Outliers)

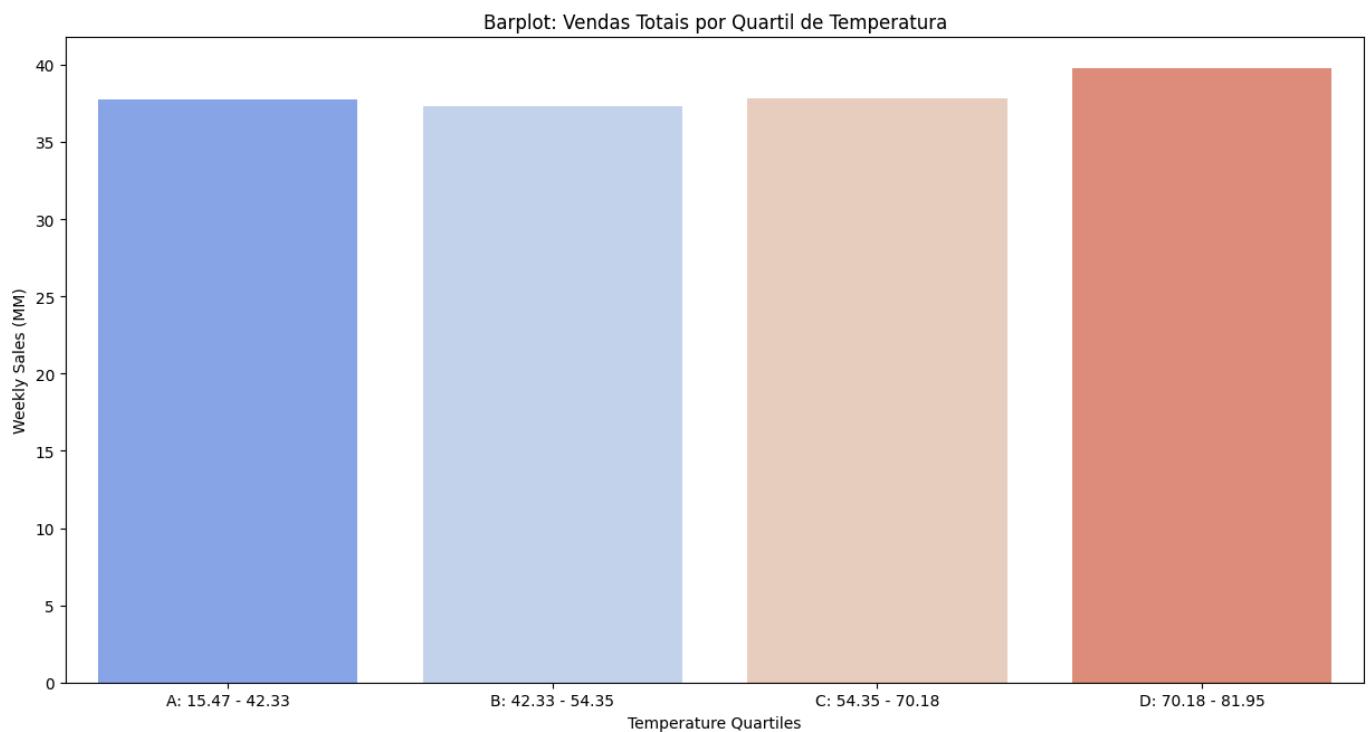
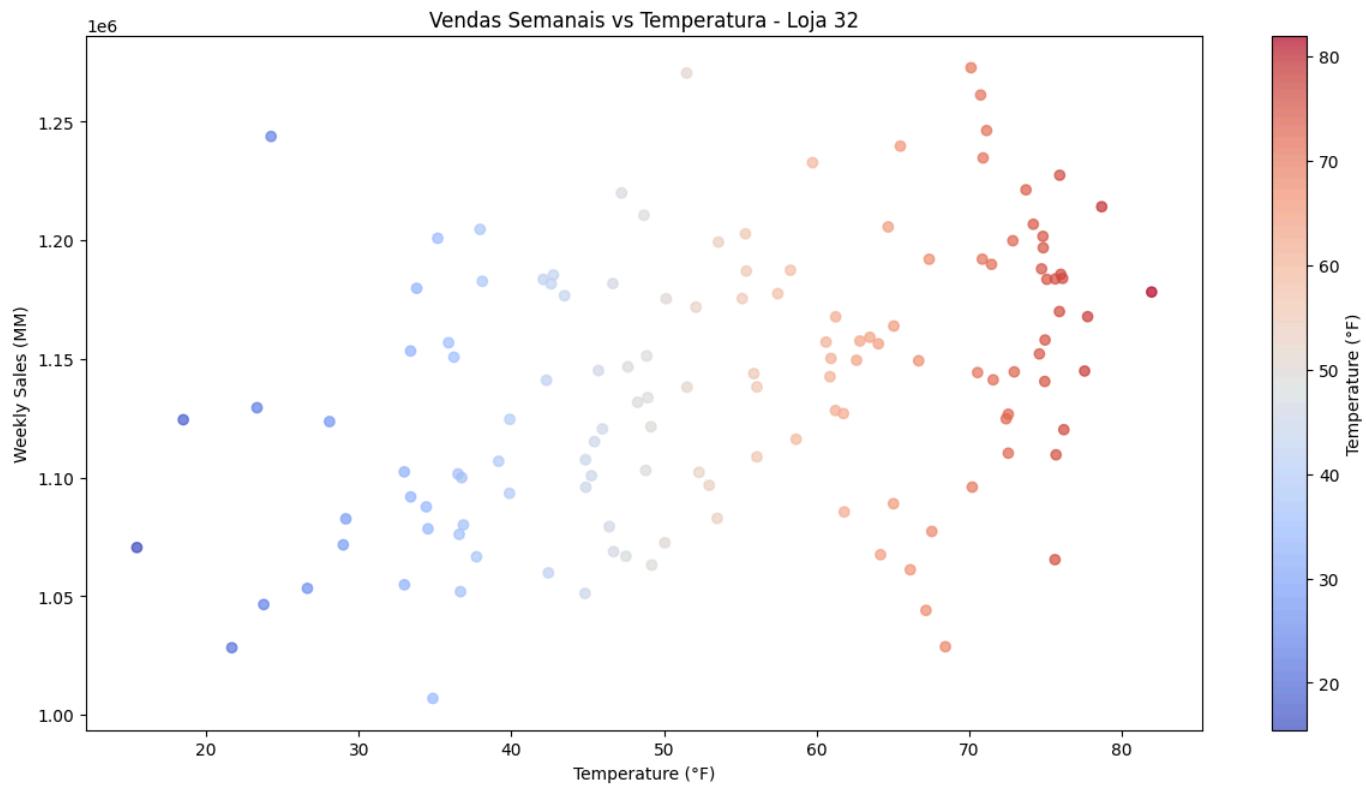
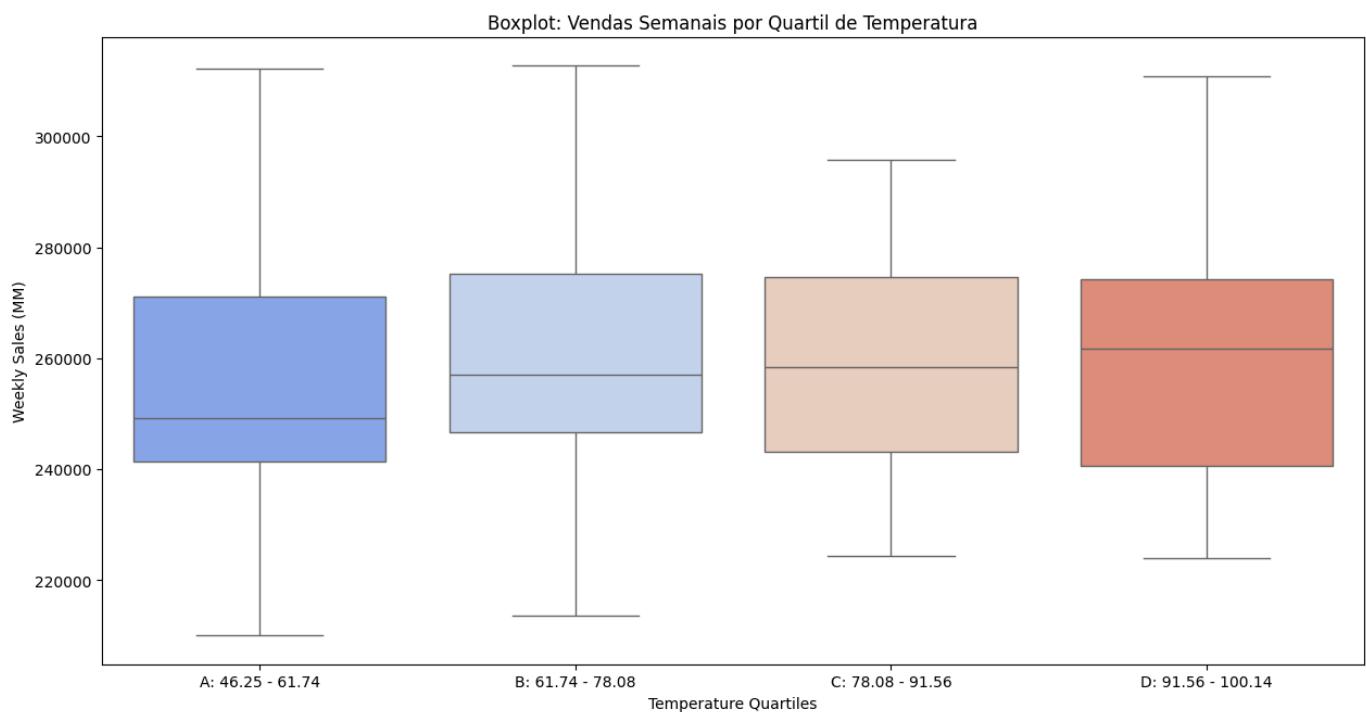


Gráfico de Vendas Semanais vs Temperatura - Loja 32 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 33 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 33 (Sem Outliers)

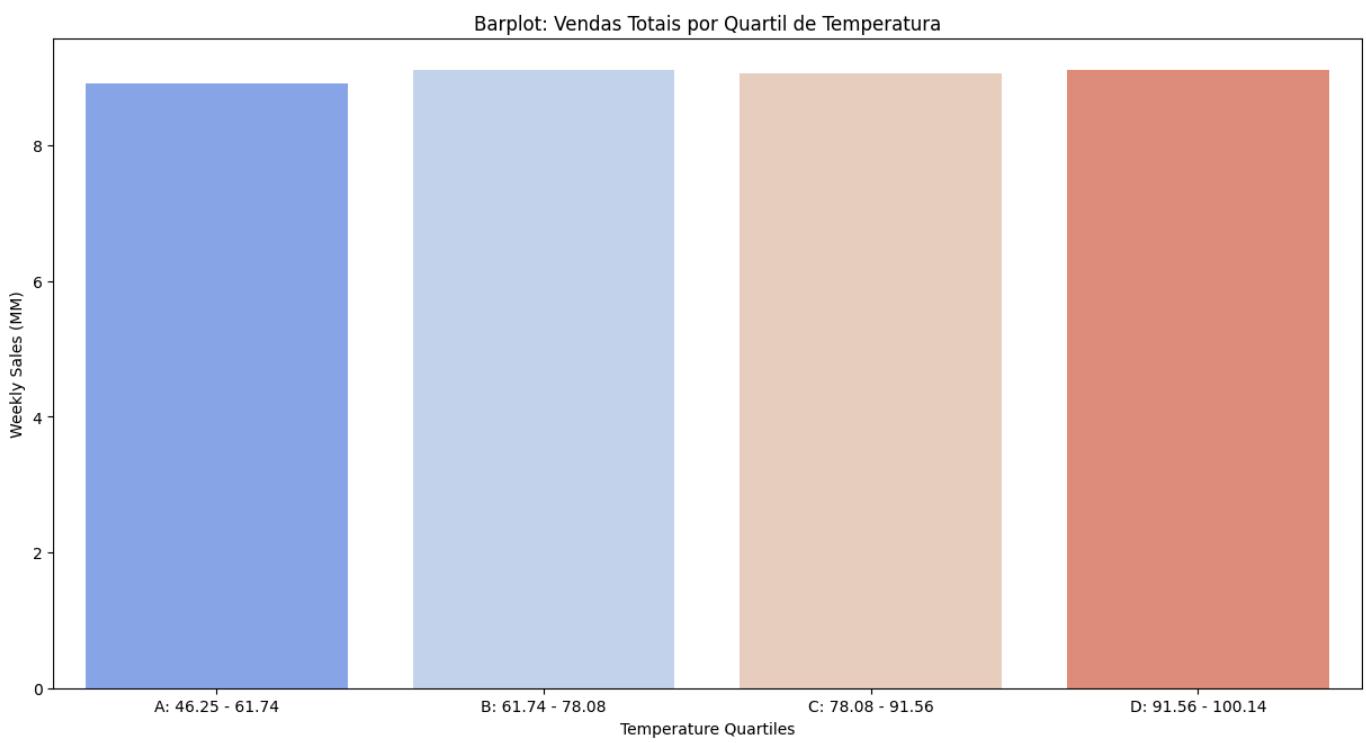
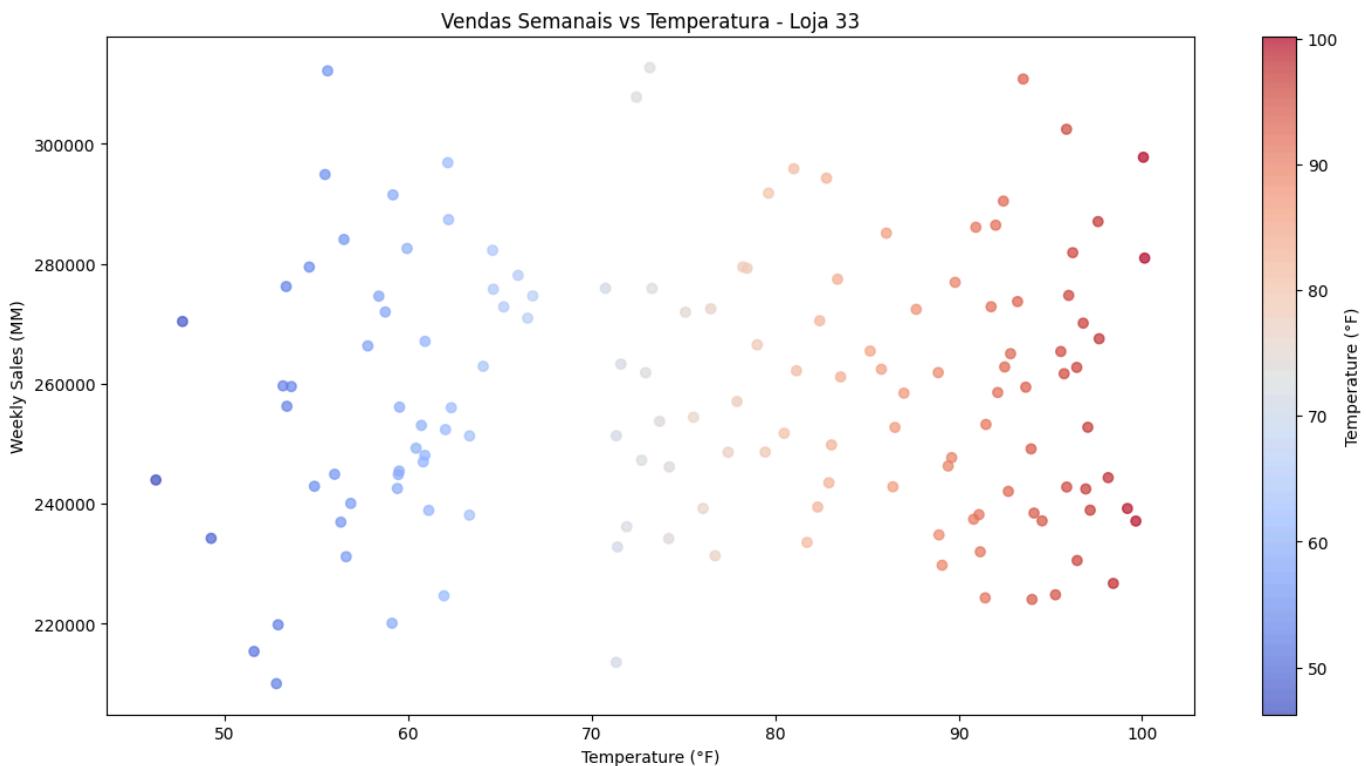
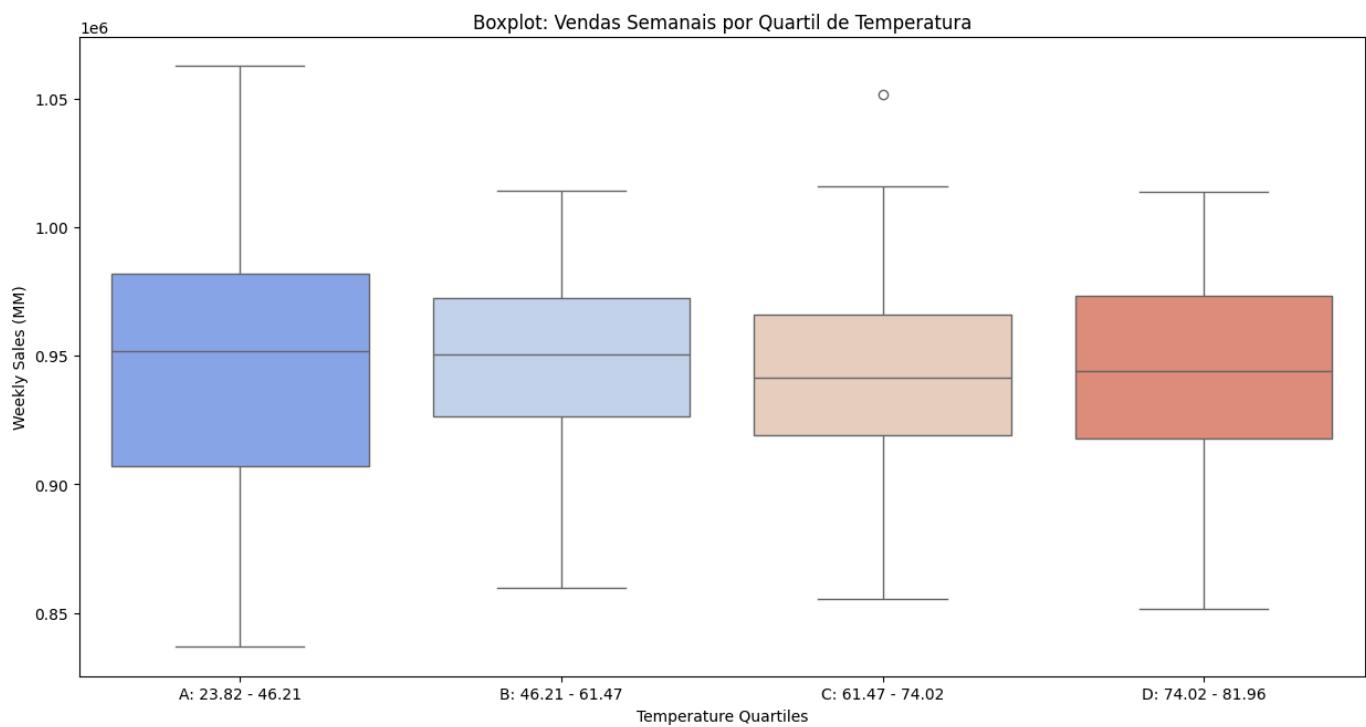


Gráfico de Vendas Semanais vs Temperatura - Loja 33 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 34 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 34 (Sem Outliers)

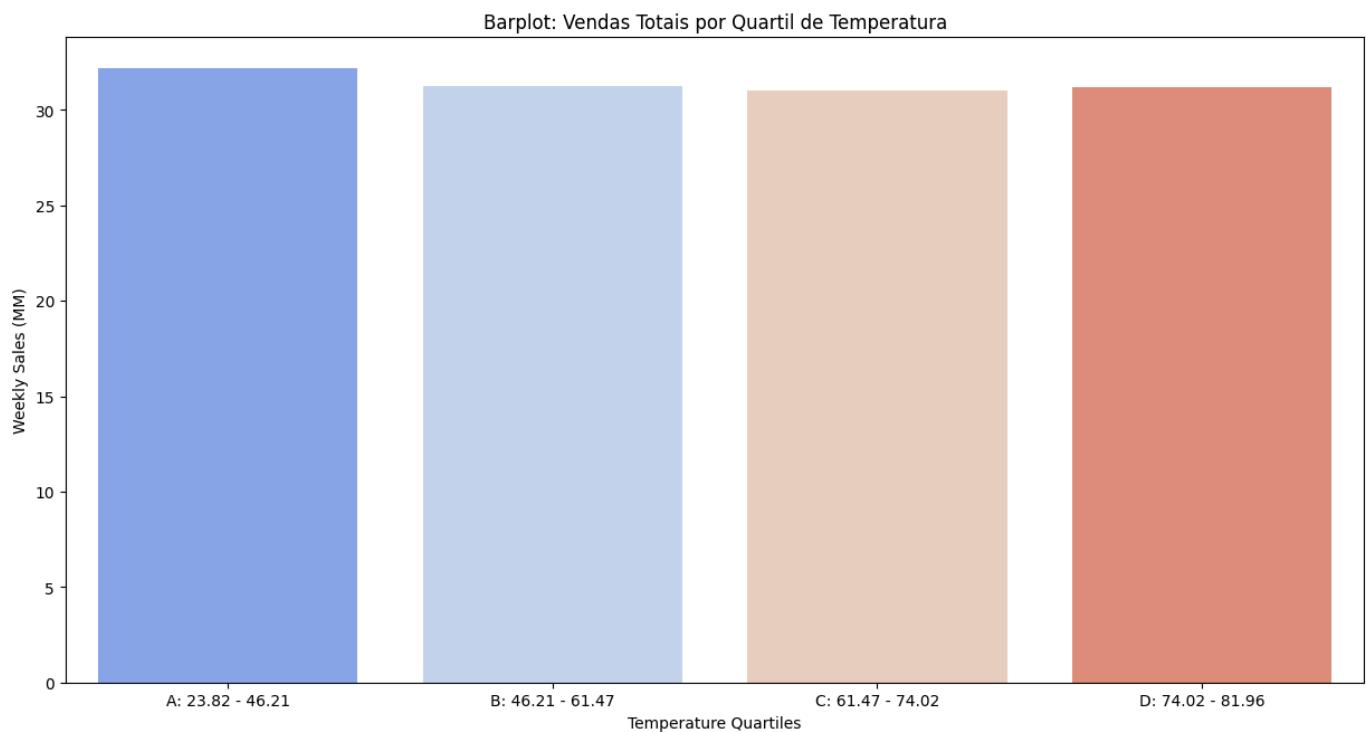
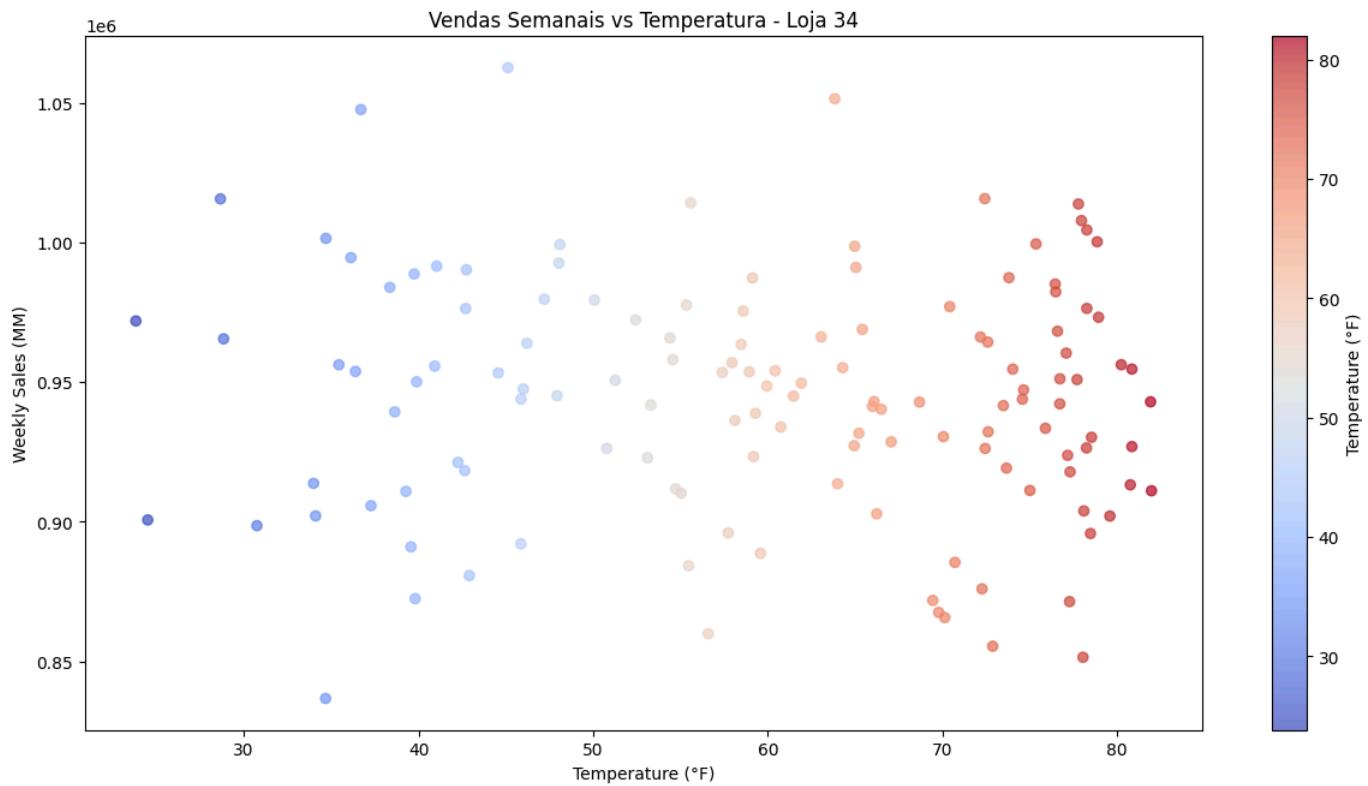
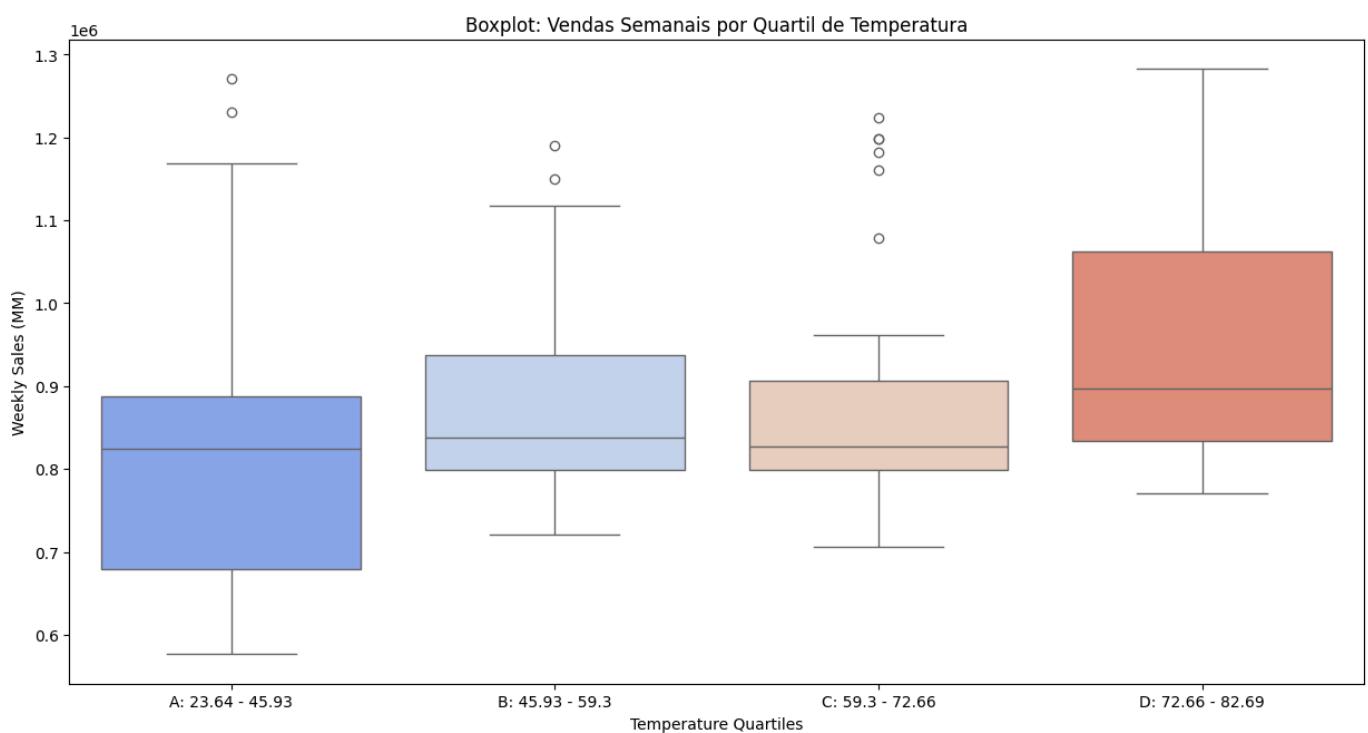


Gráfico de Vendas Semanais vs Temperatura - Loja 34 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 35 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 35 (Sem Outliers)

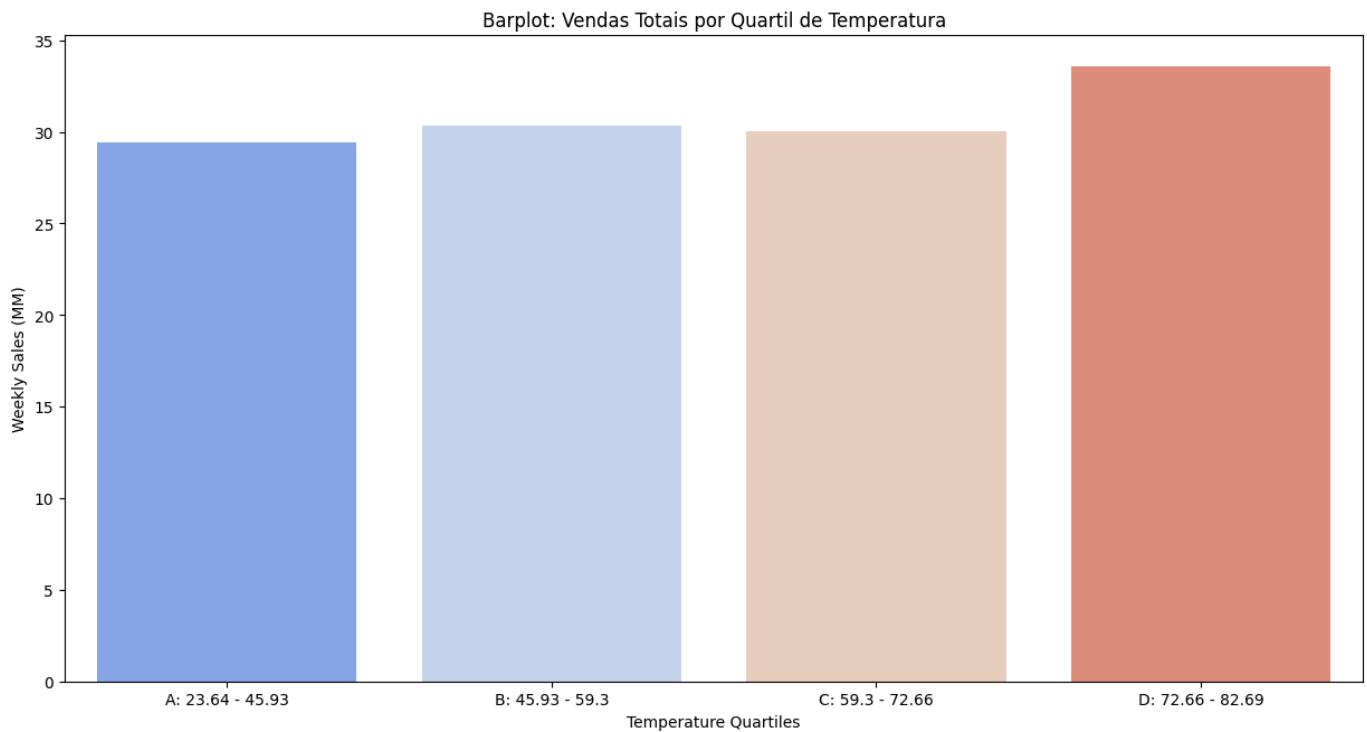
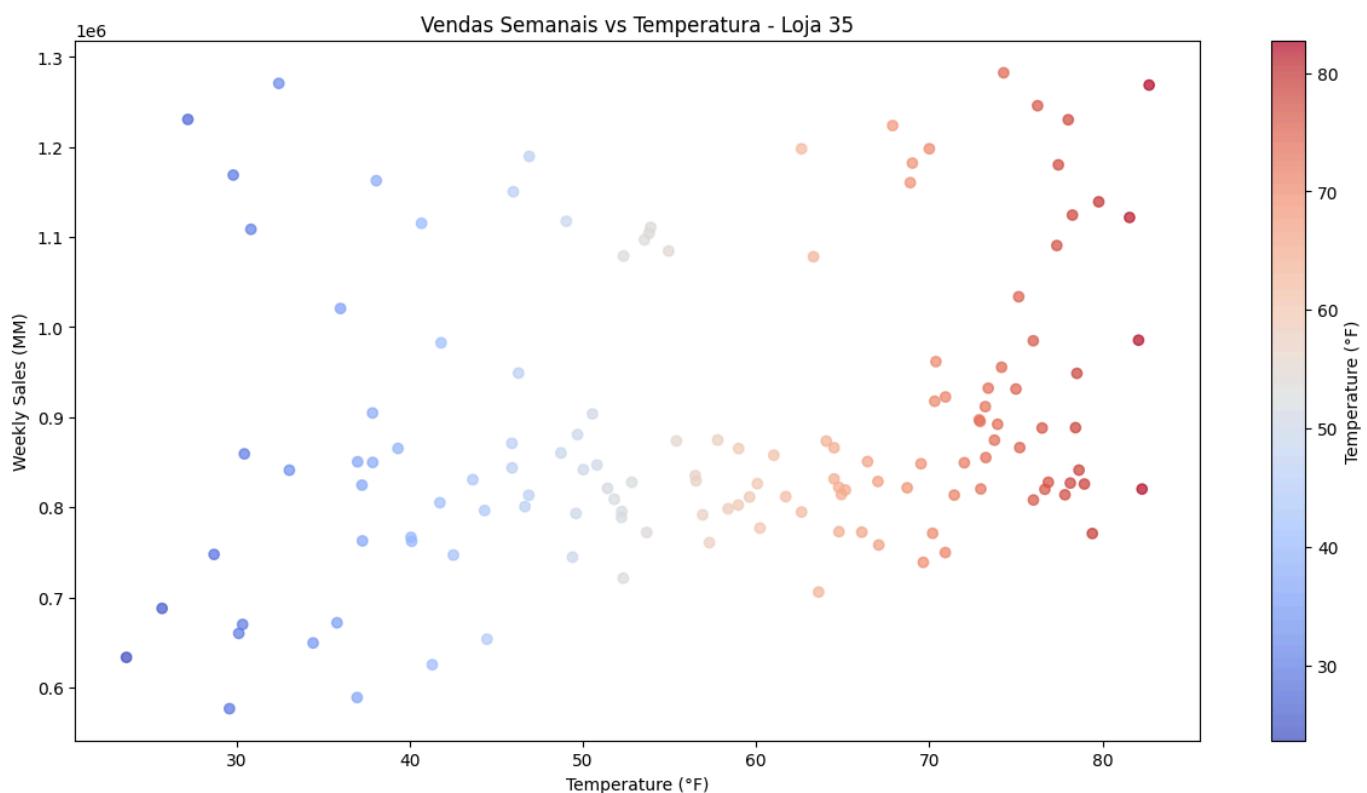
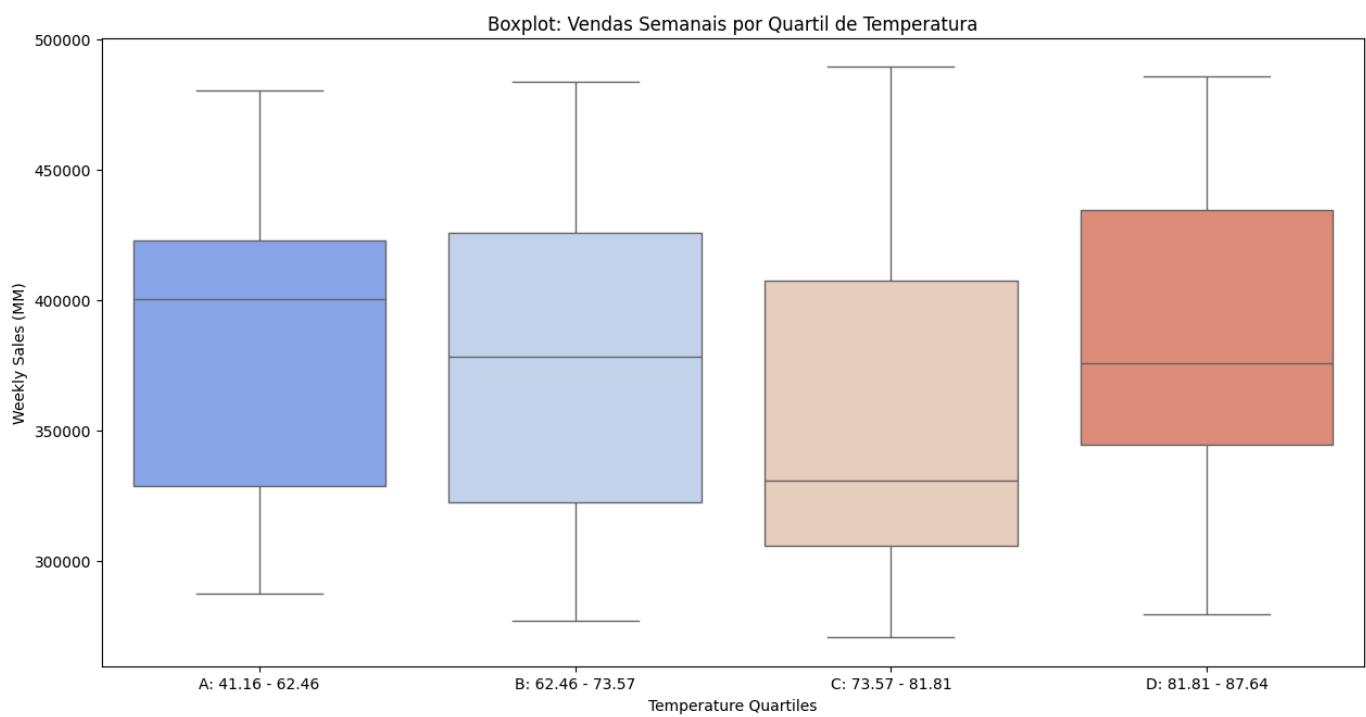


Gráfico de Vendas Semanais vs Temperatura - Loja 35 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 36 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 36 (Sem Outliers)

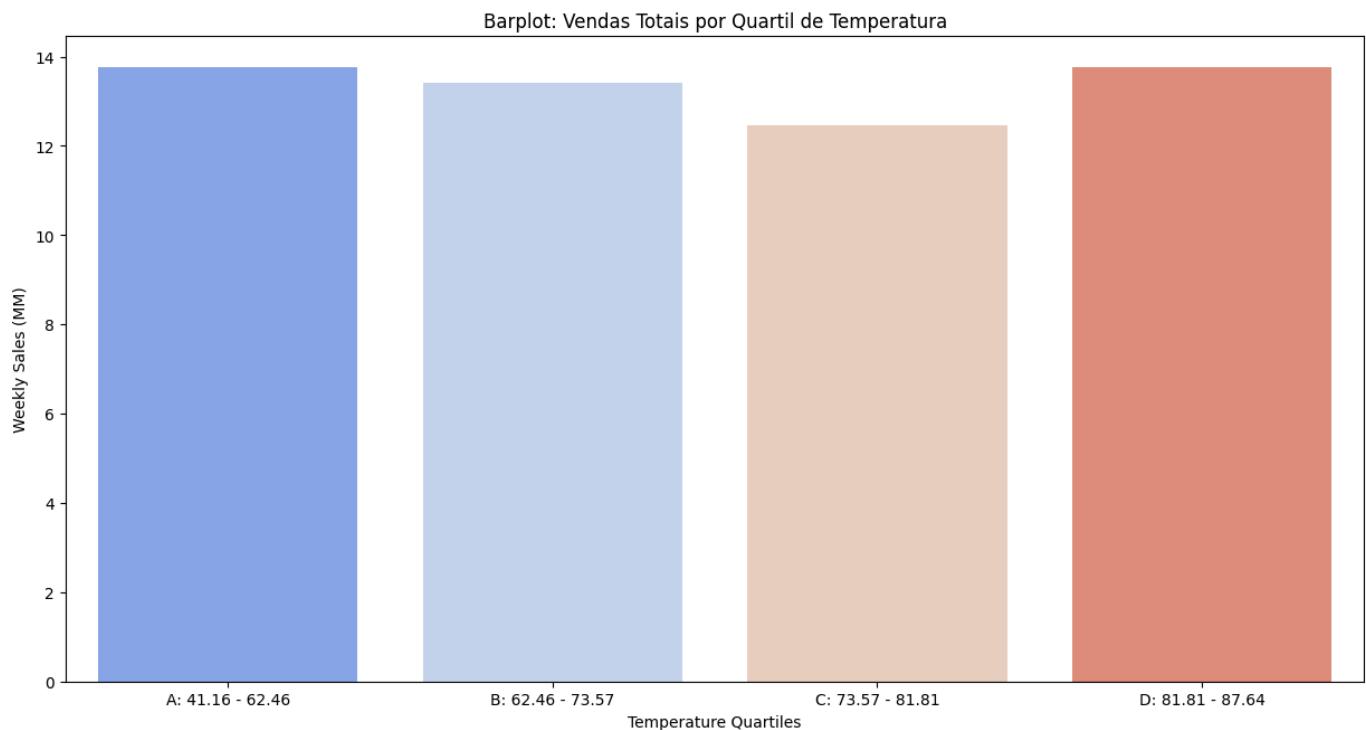
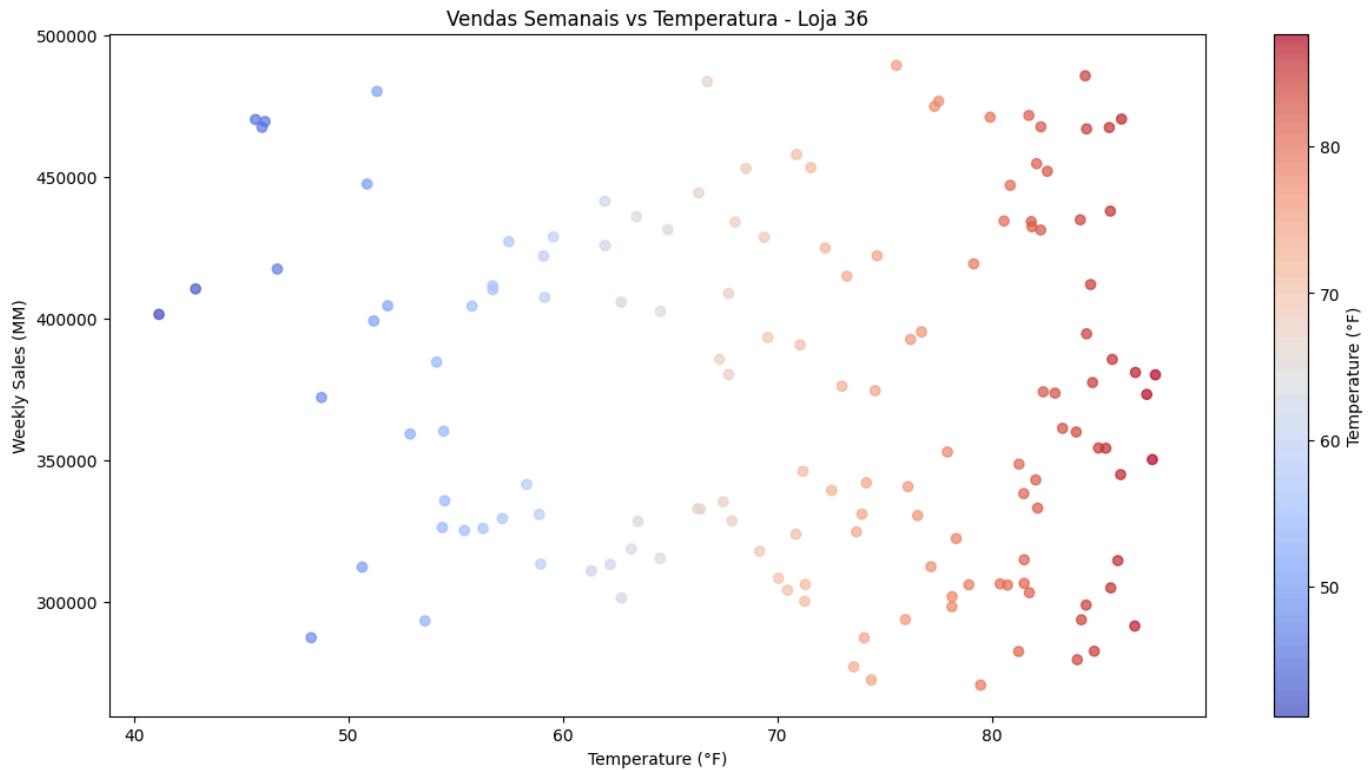
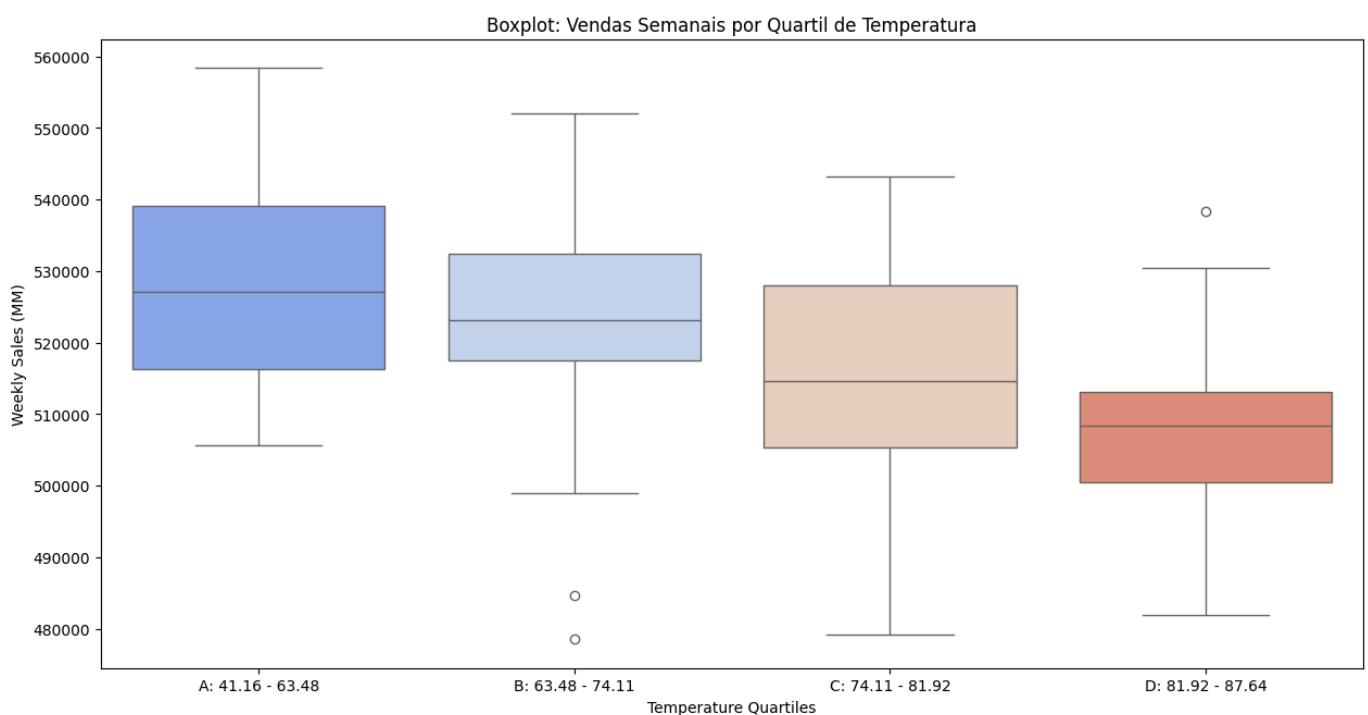


Gráfico de Vendas Semanais vs Temperatura - Loja 36 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 37 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 37 (Sem Outliers)

Barplot: Vendas Totais por Quartil de Temperatura

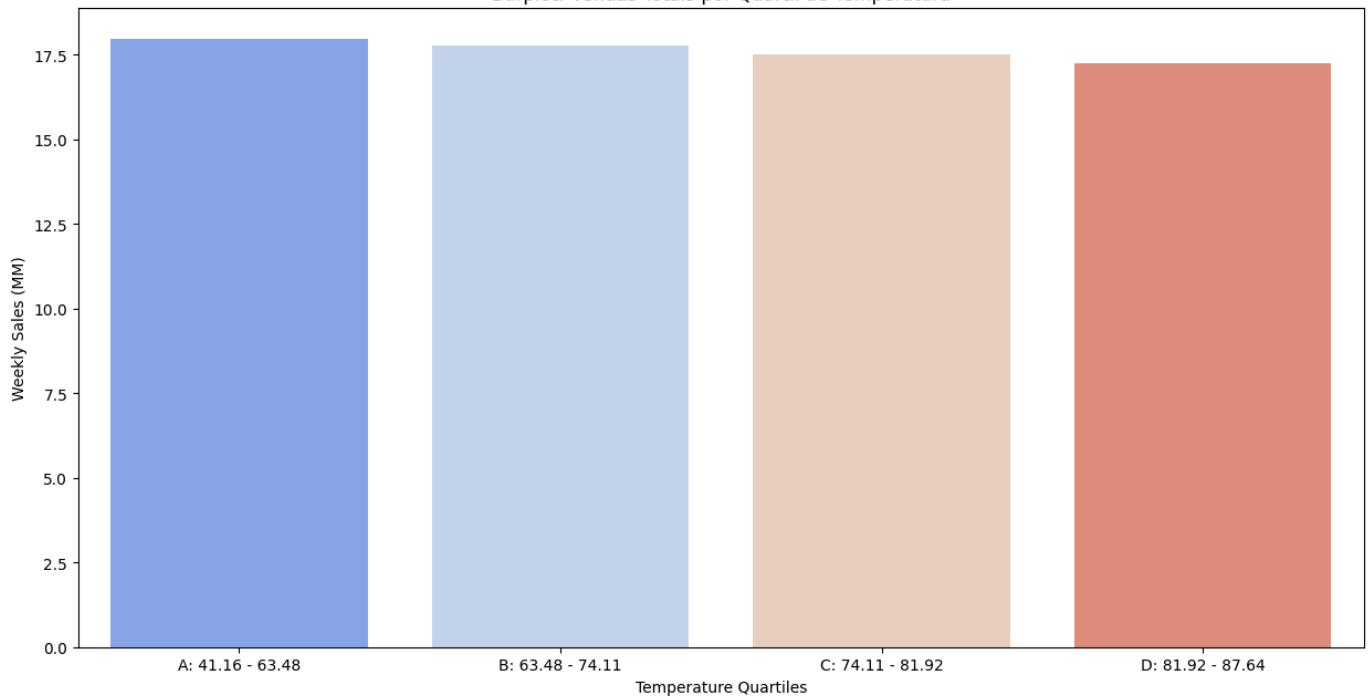
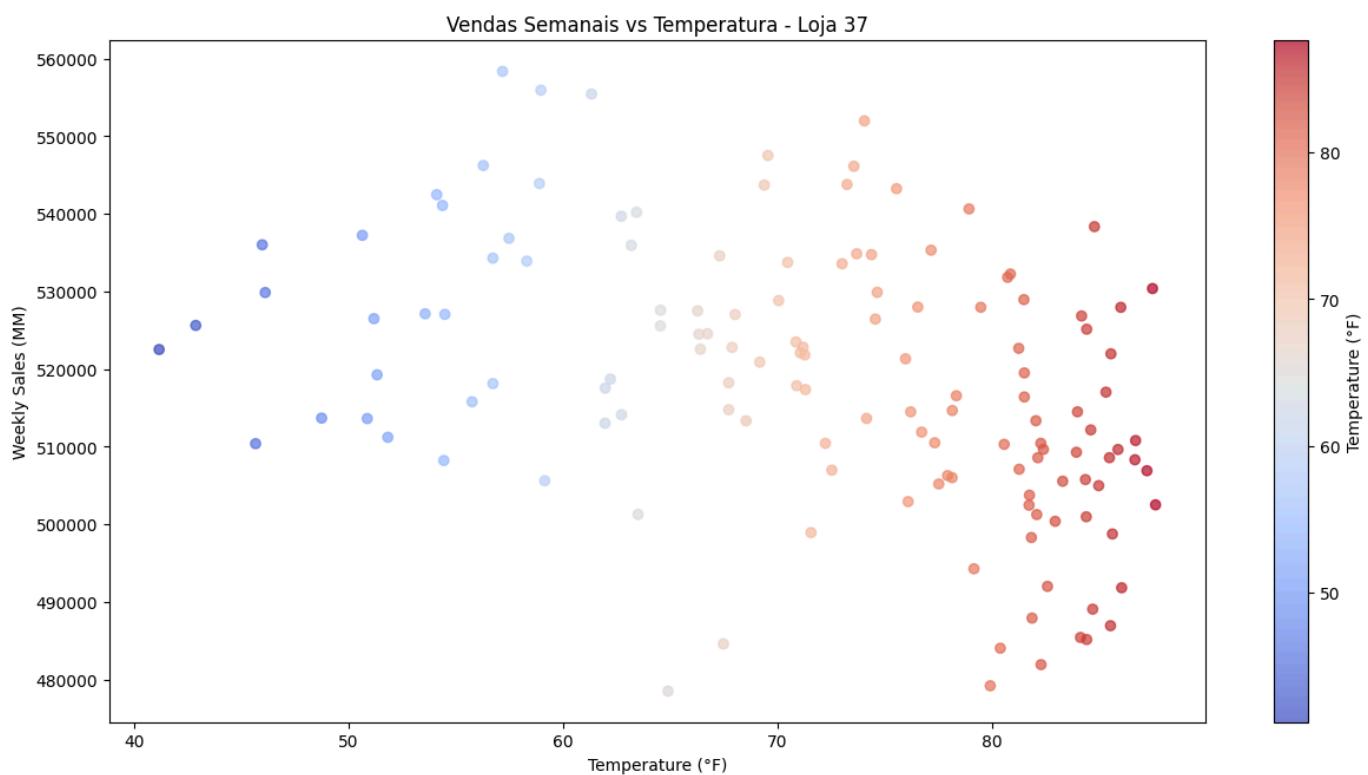
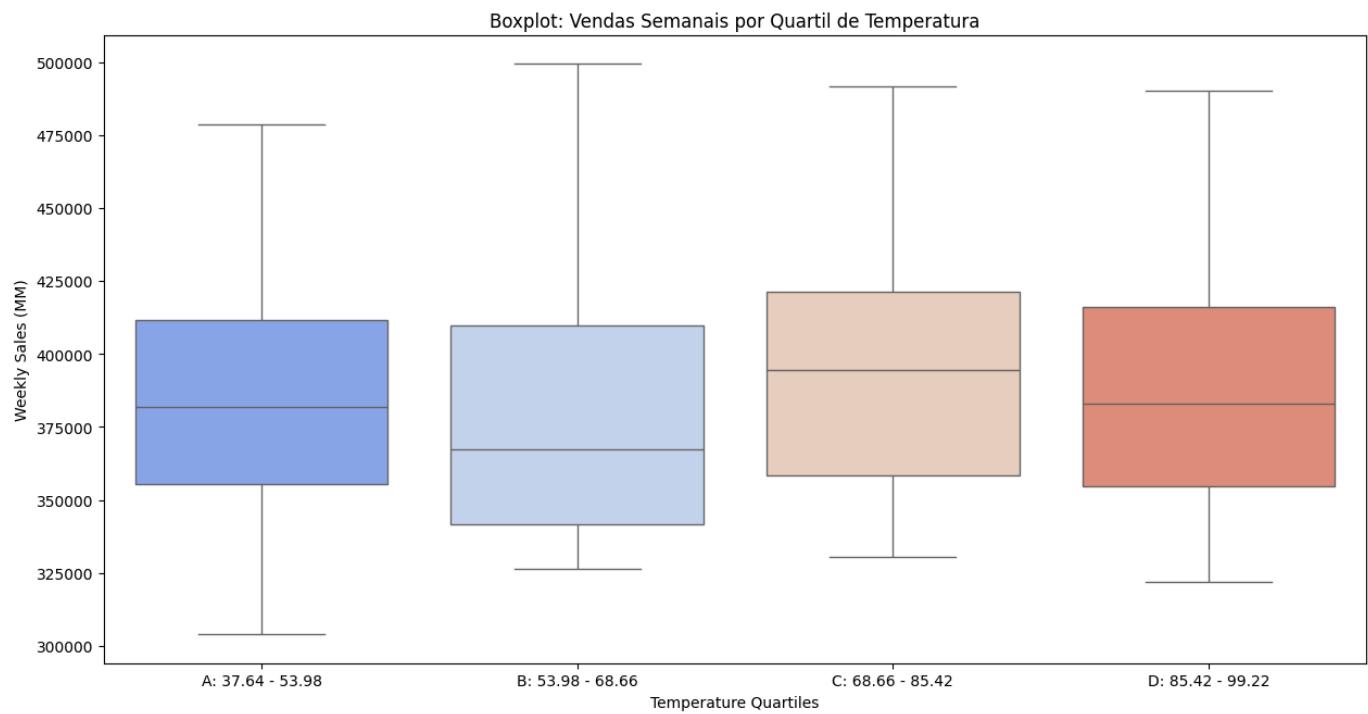


Gráfico de Vendas Semanais vs Temperatura - Loja 37 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 38 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 38 (Sem Outliers)

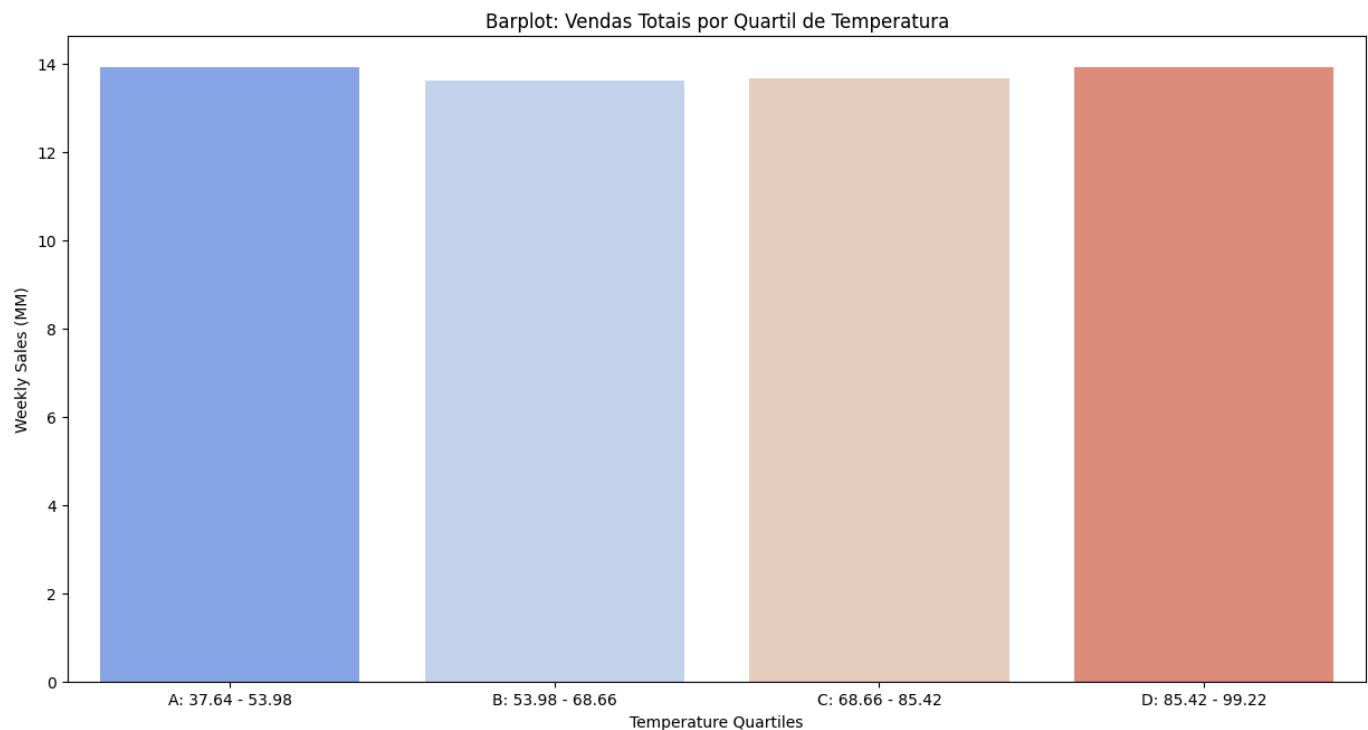
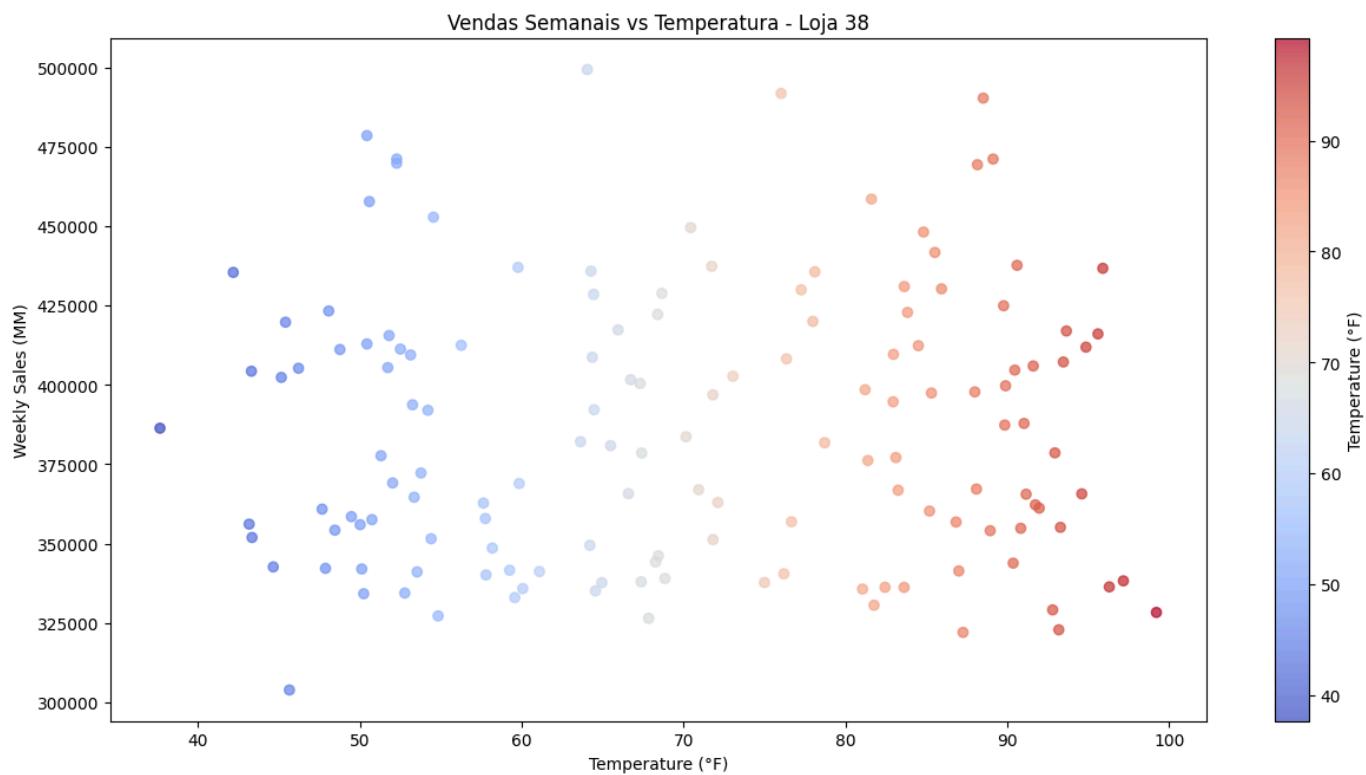
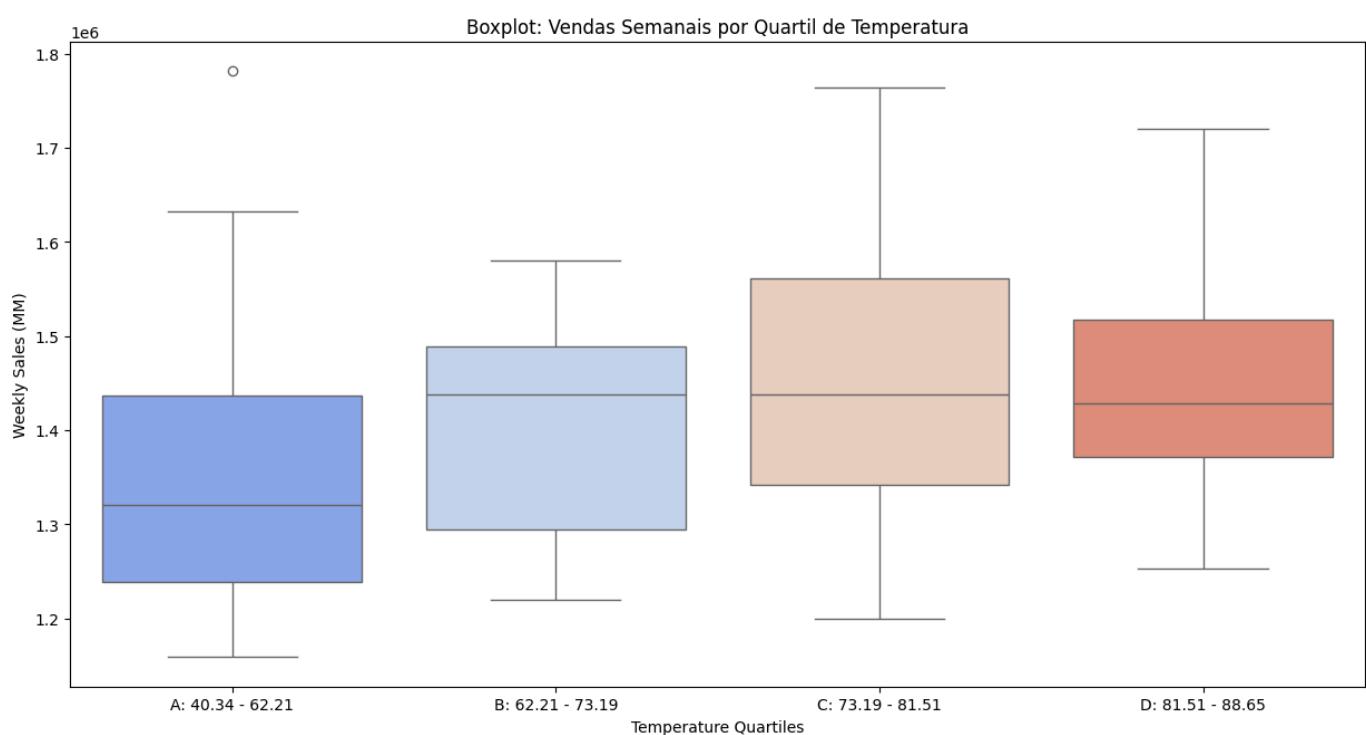


Gráfico de Vendas Semanais vs Temperatura - Loja 38 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 39 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 39 (Sem Outliers)

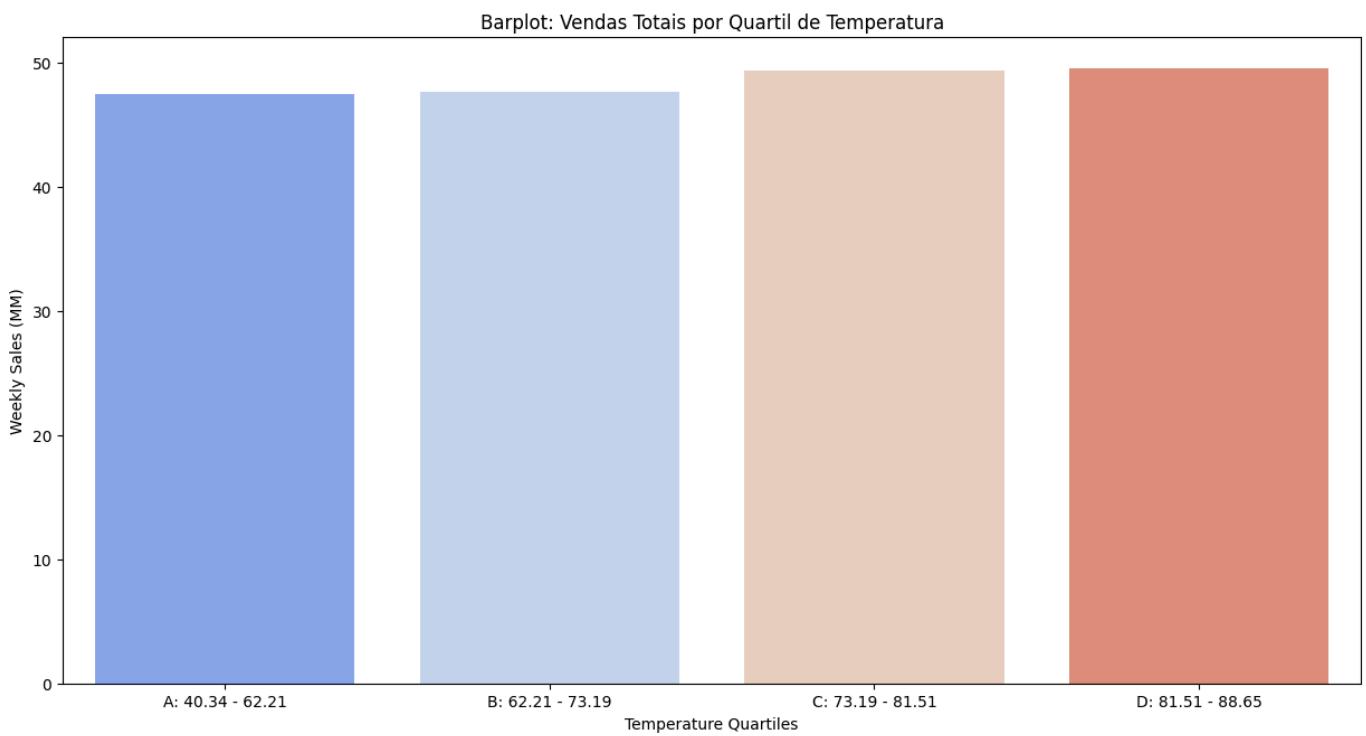
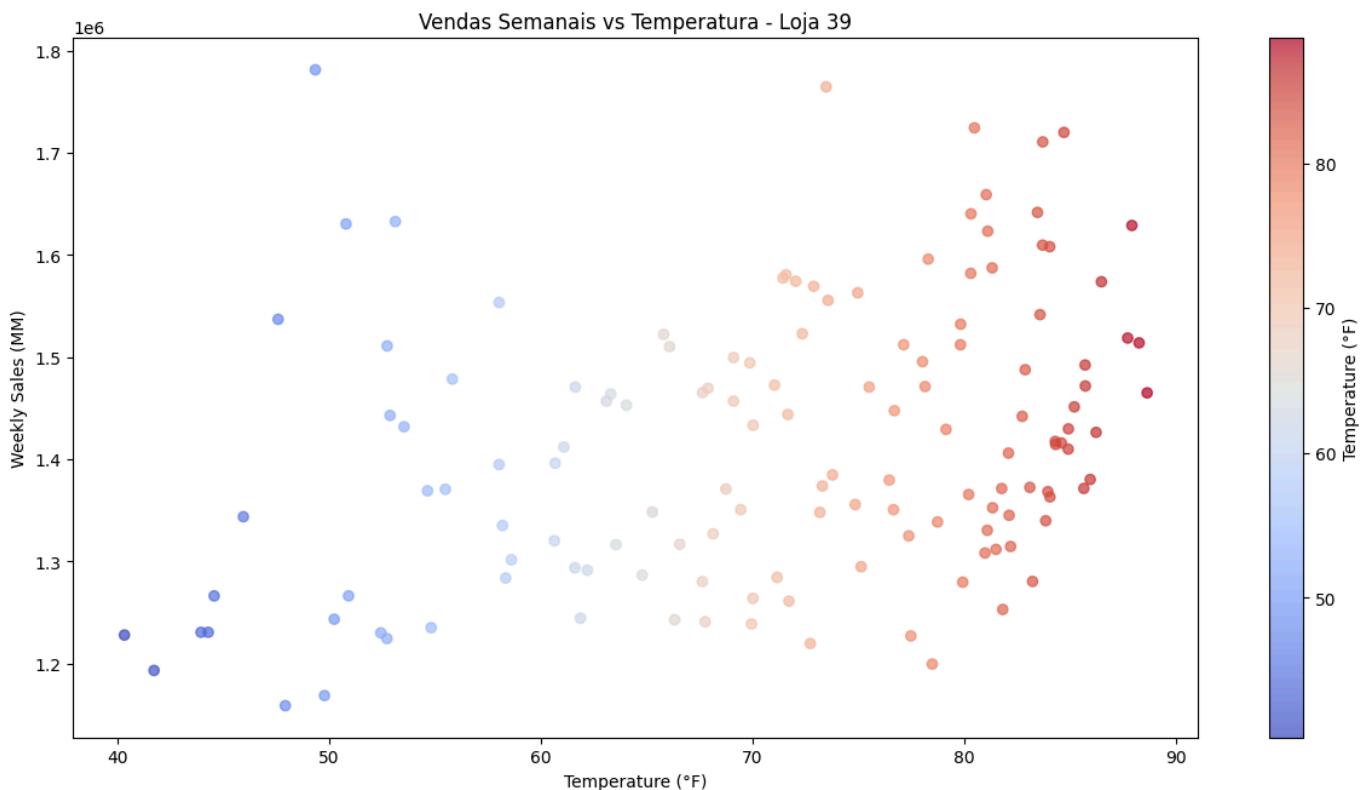
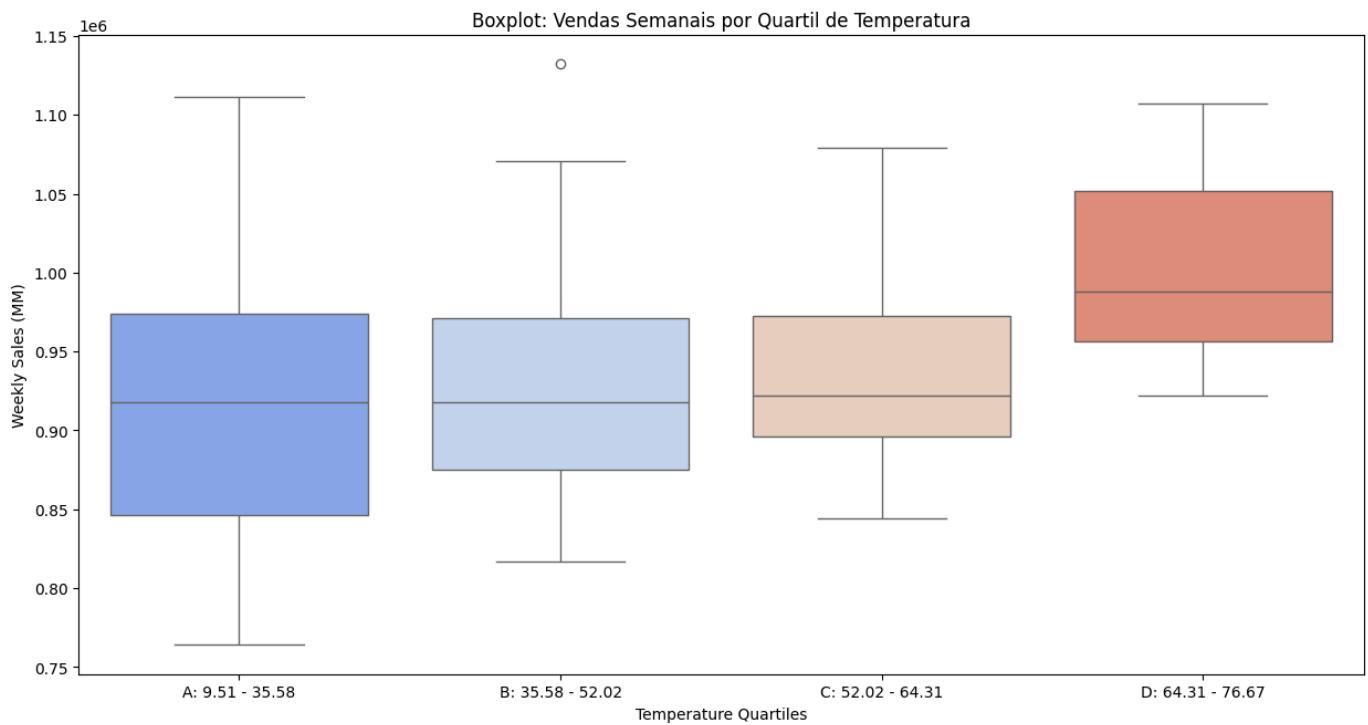


Gráfico de Vendas Semanais vs Temperatura - Loja 39 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 40 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 40 (Sem Outliers)

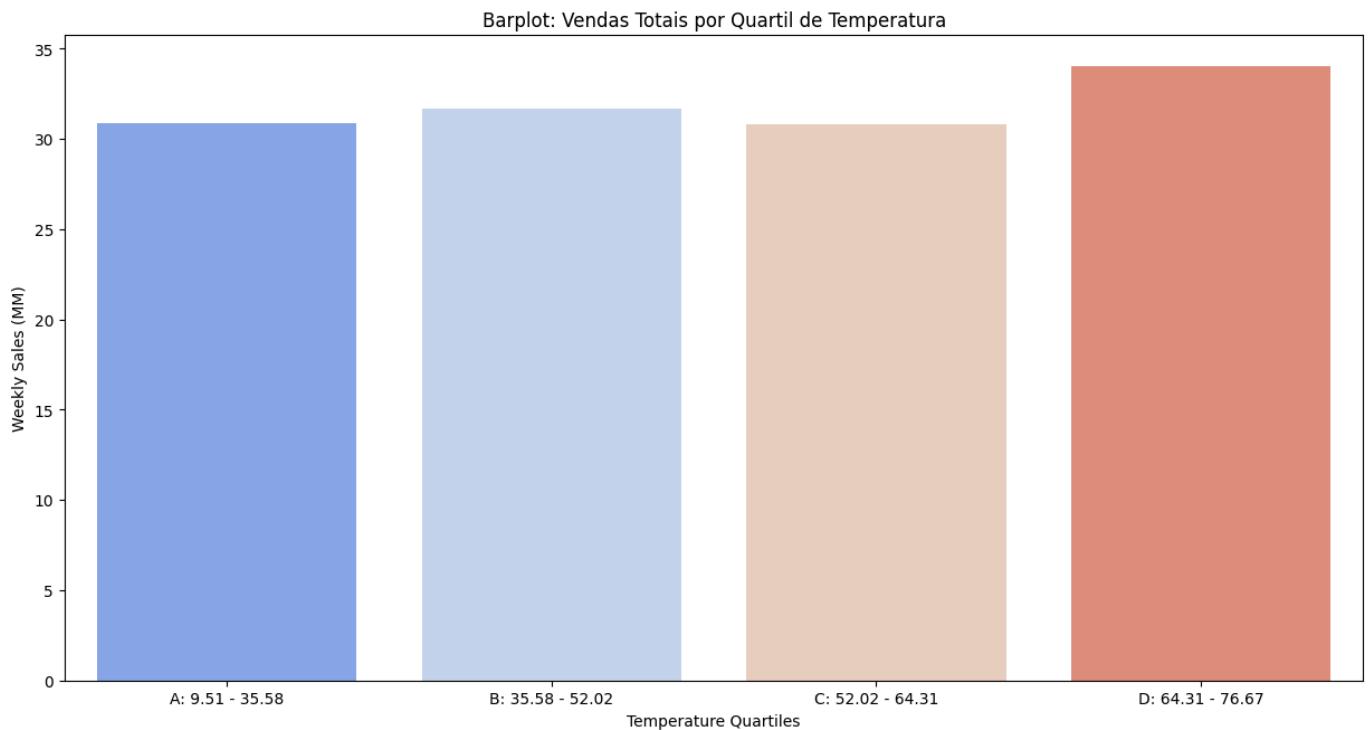
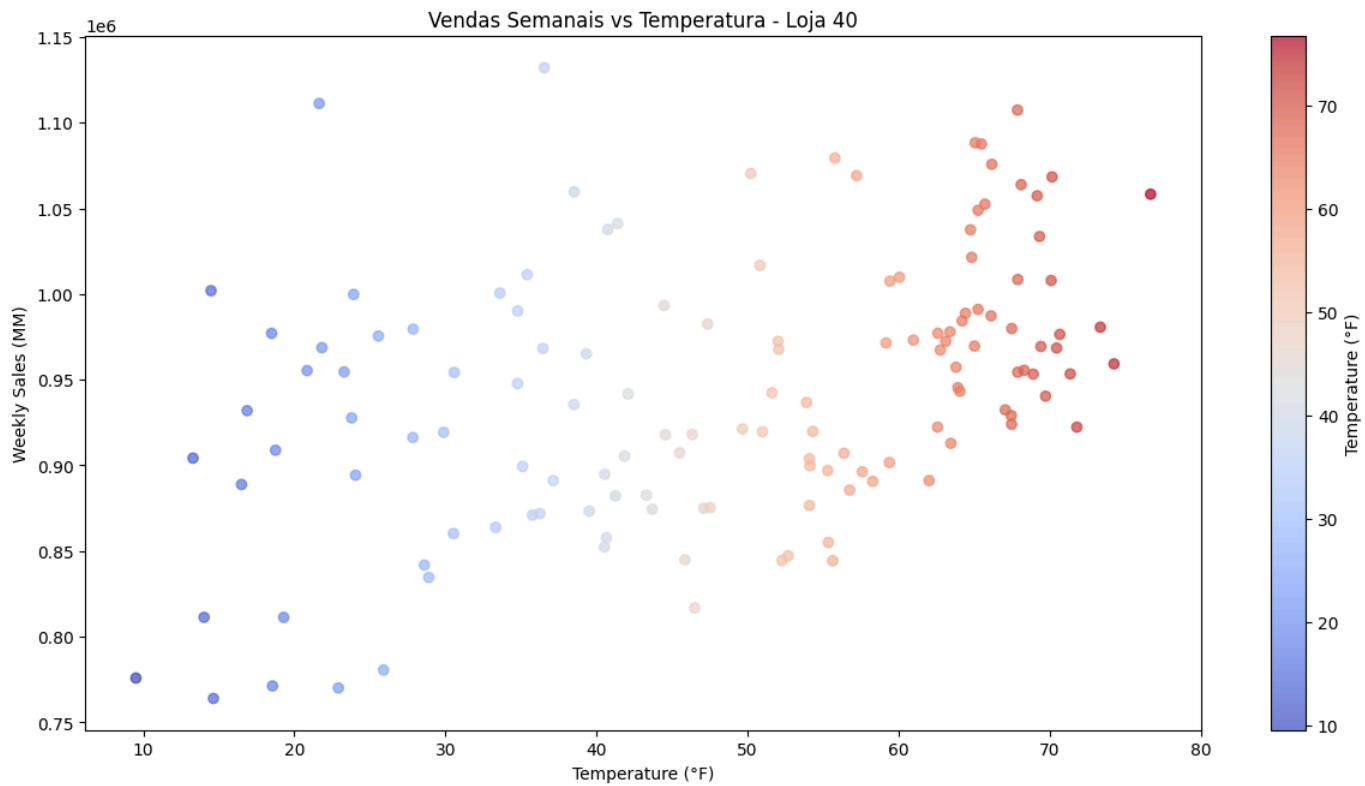
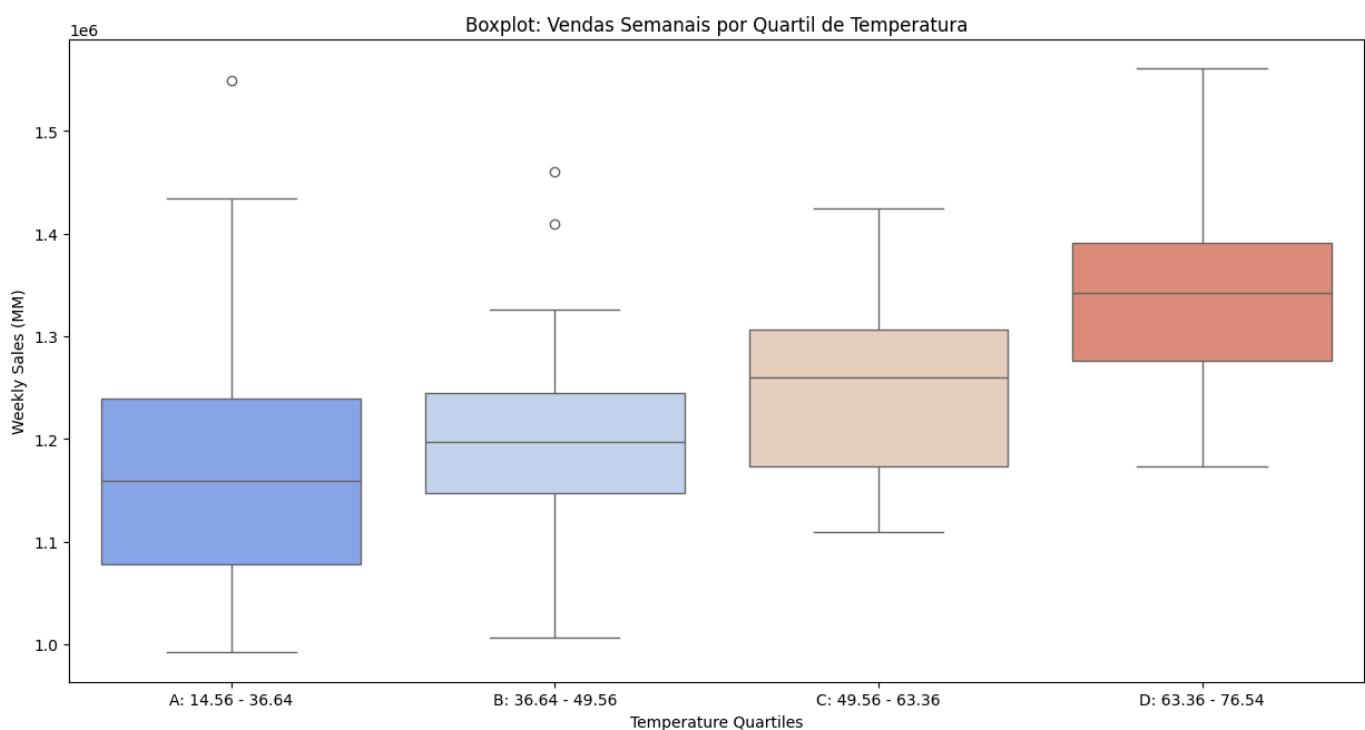


Gráfico de Vendas Semanais vs Temperatura - Loja 40 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 41 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 41 (Sem Outliers)

Barplot: Vendas Totais por Quartil de Temperatura

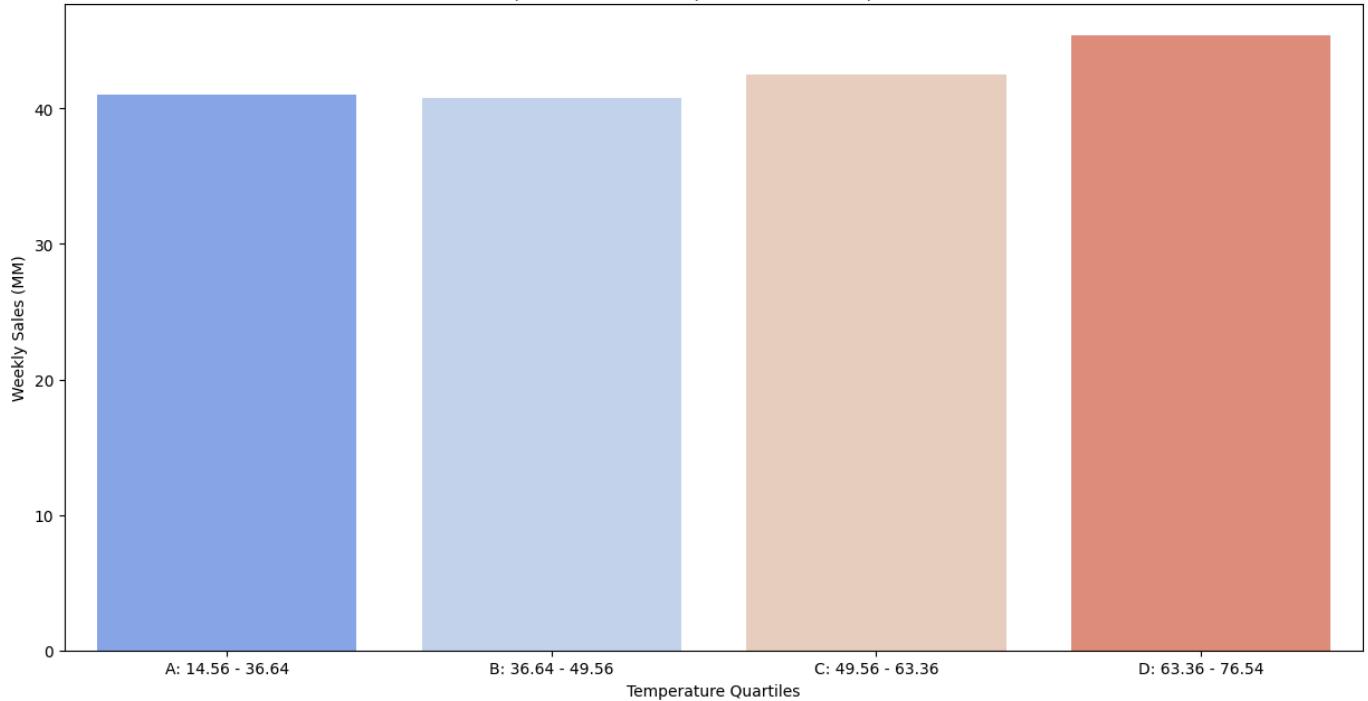
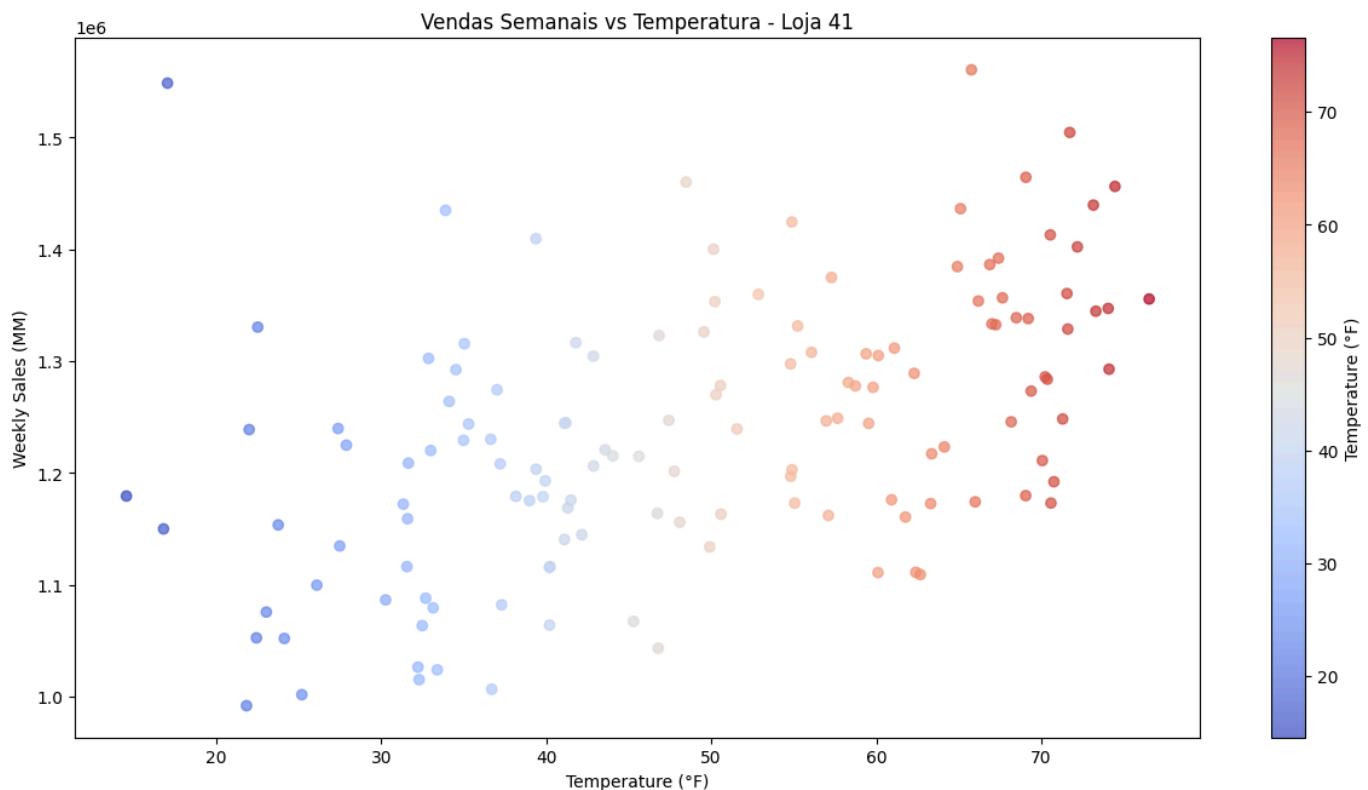
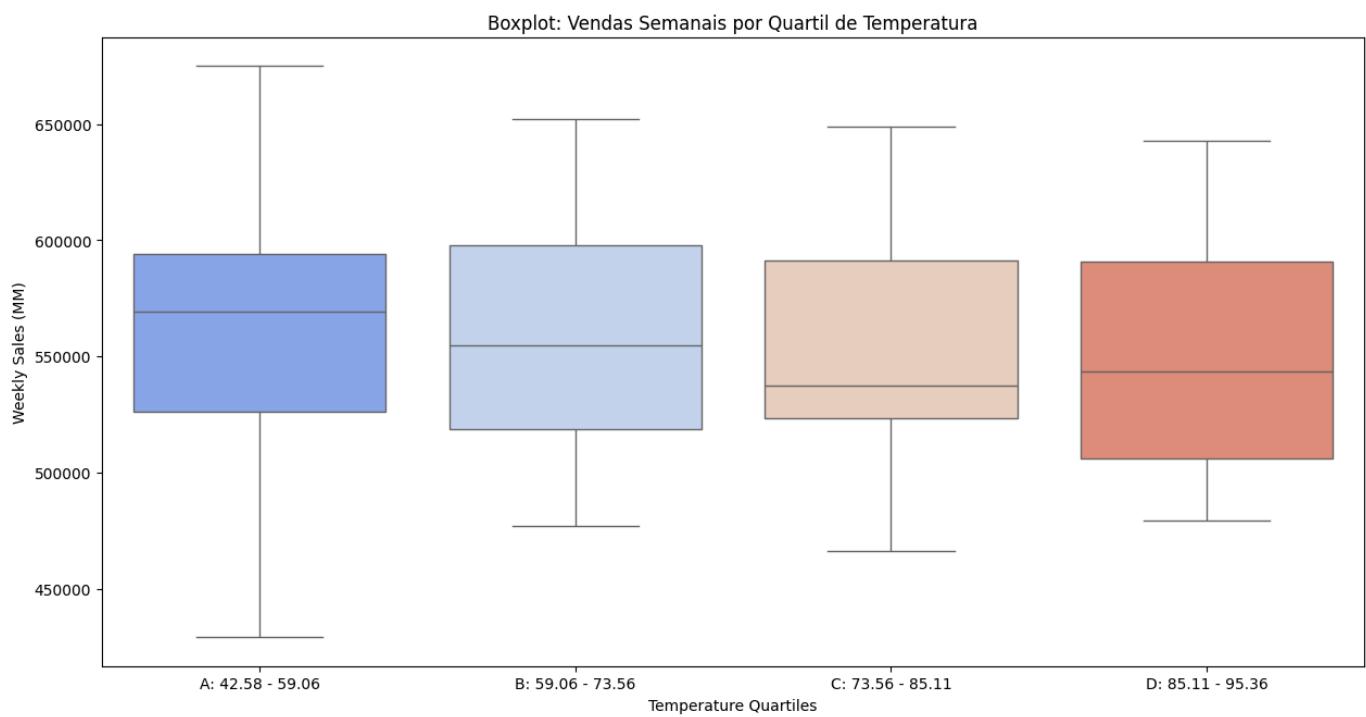


Gráfico de Vendas Semanais vs Temperatura - Loja 41 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 42 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 42 (Sem Outliers)

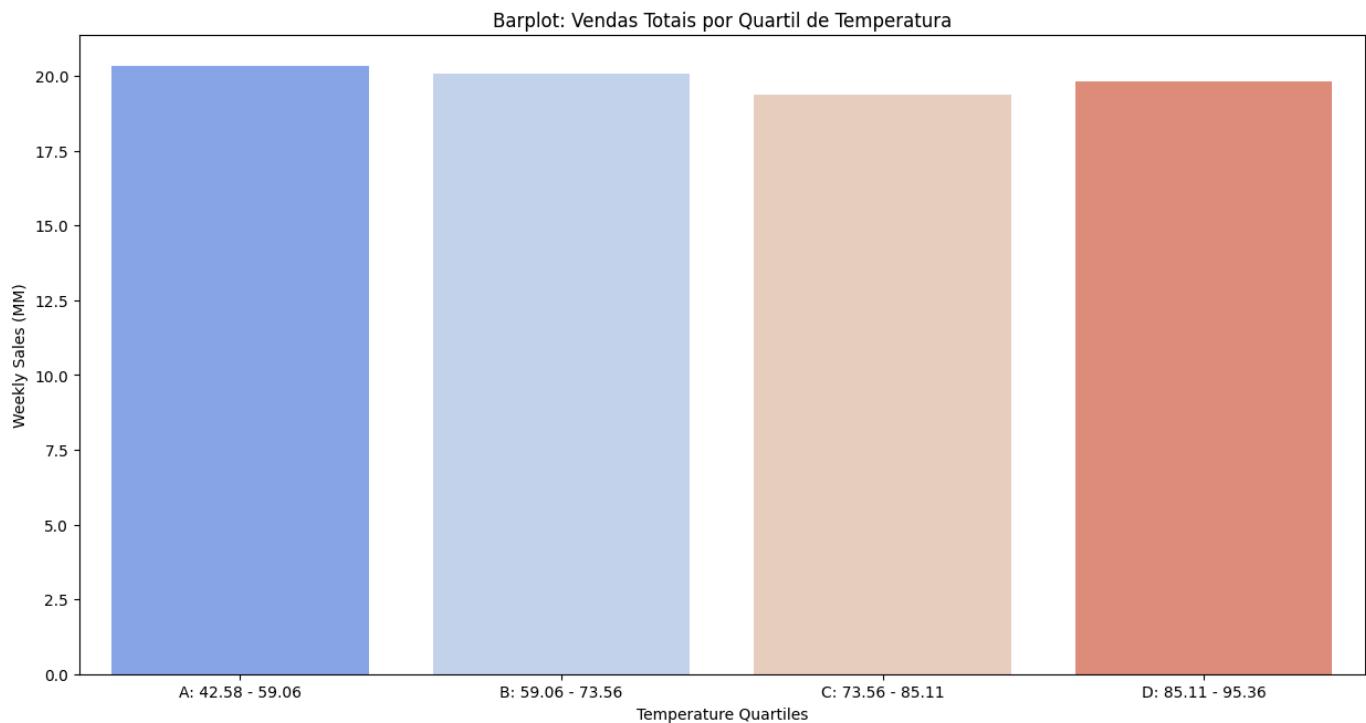
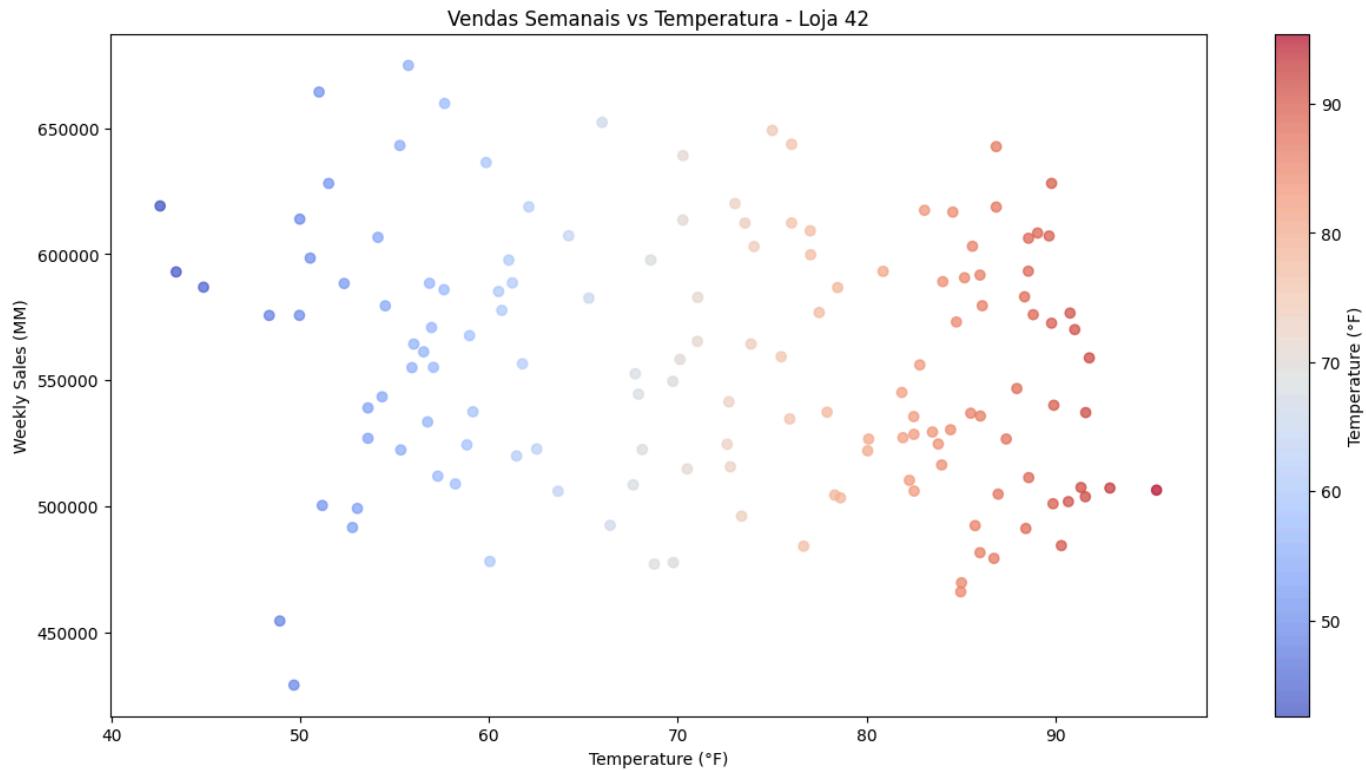
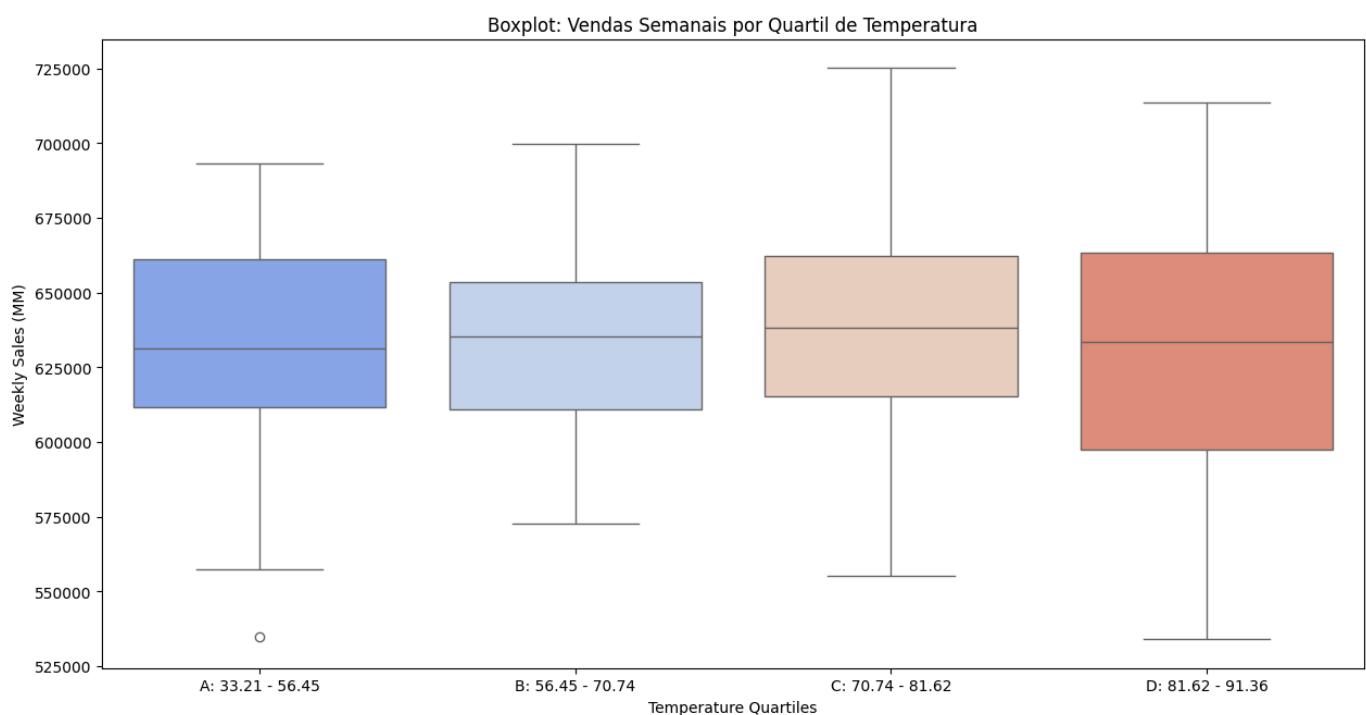


Gráfico de Vendas Semanais vs Temperatura - Loja 42 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 43 (Sem Outliers)



Barplot: Vendas Totais por Quartil de Temperatura

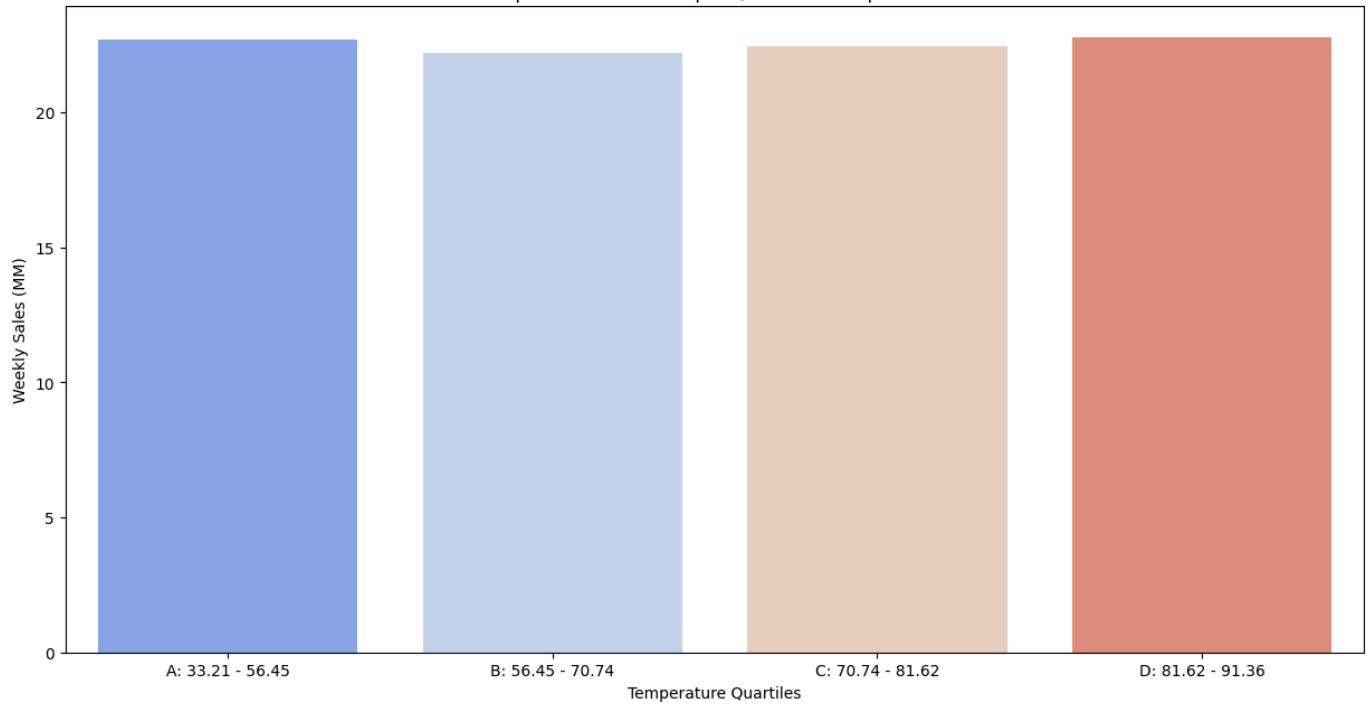
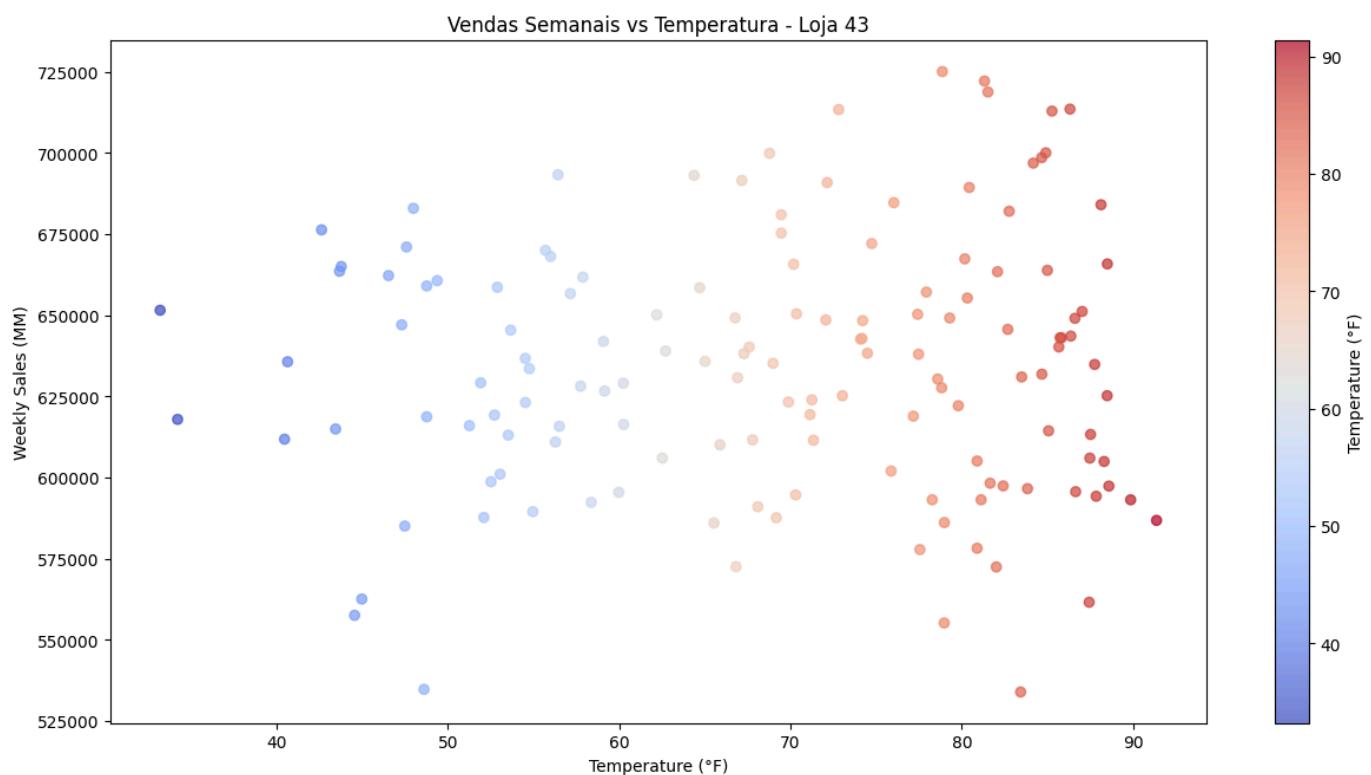
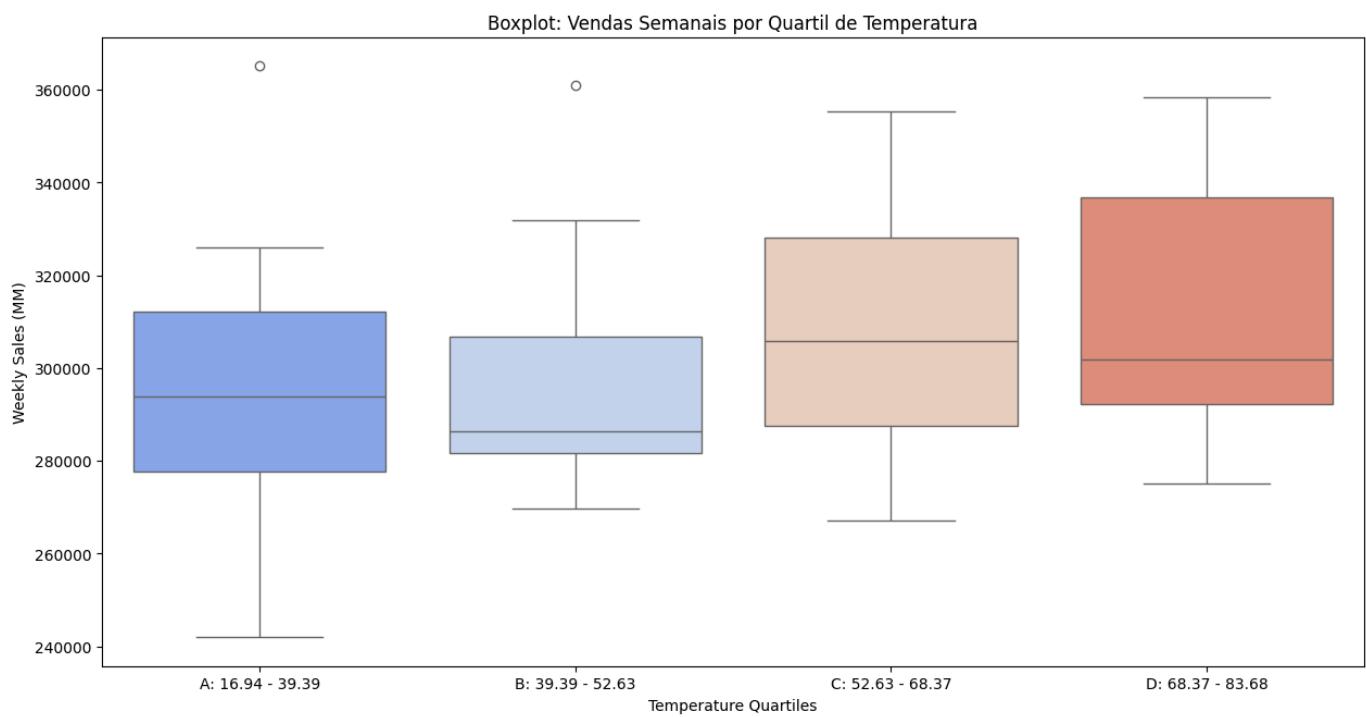


Gráfico de Vendas Semanais vs Temperatura - Loja 43 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 44 (Sem Outliers)



Vendas Totais por Quartil de Temperatura - Loja 44 (Sem Outliers)

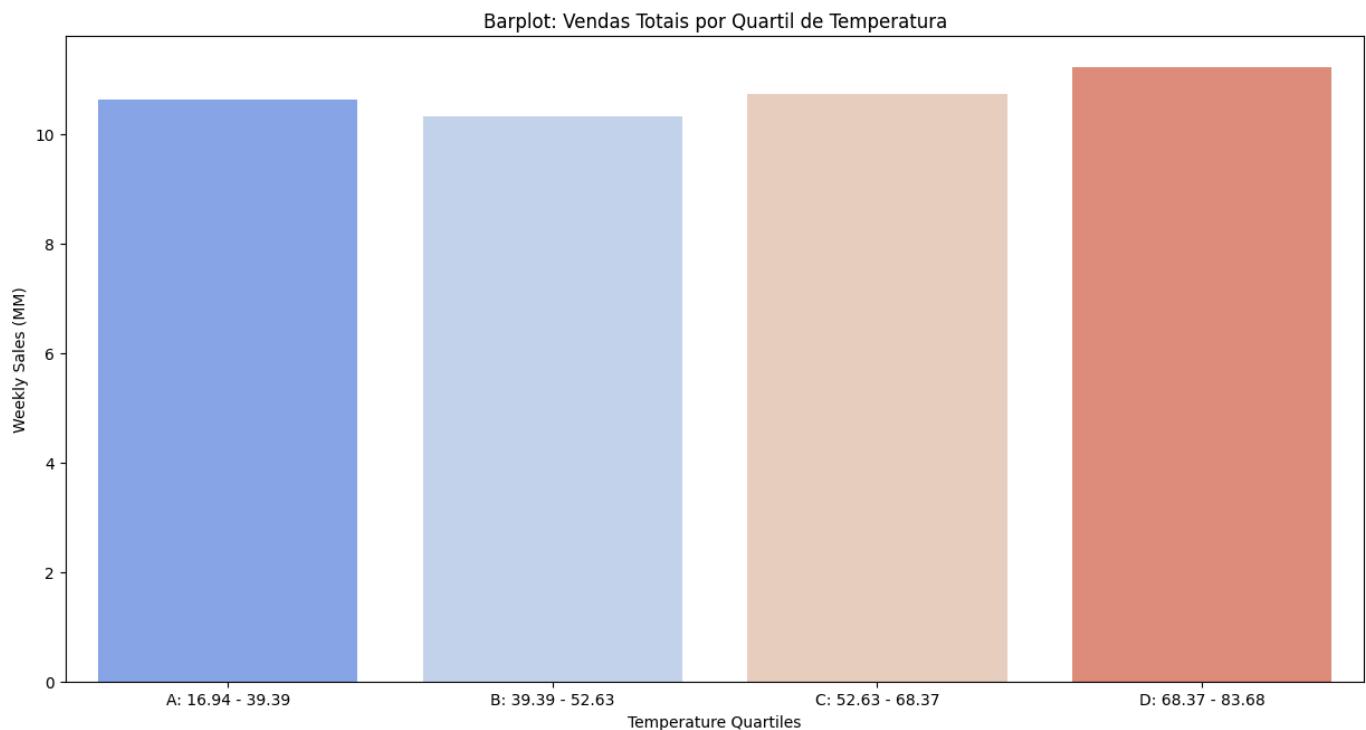
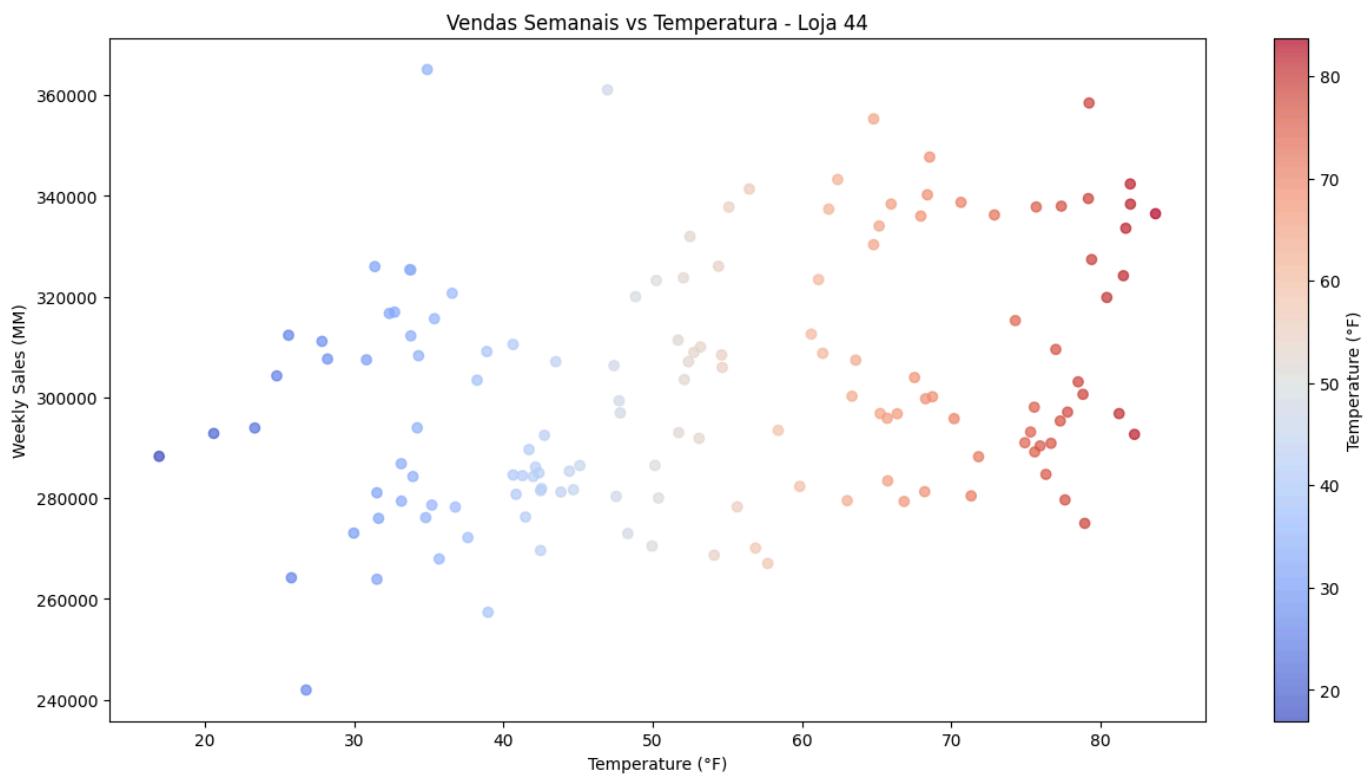
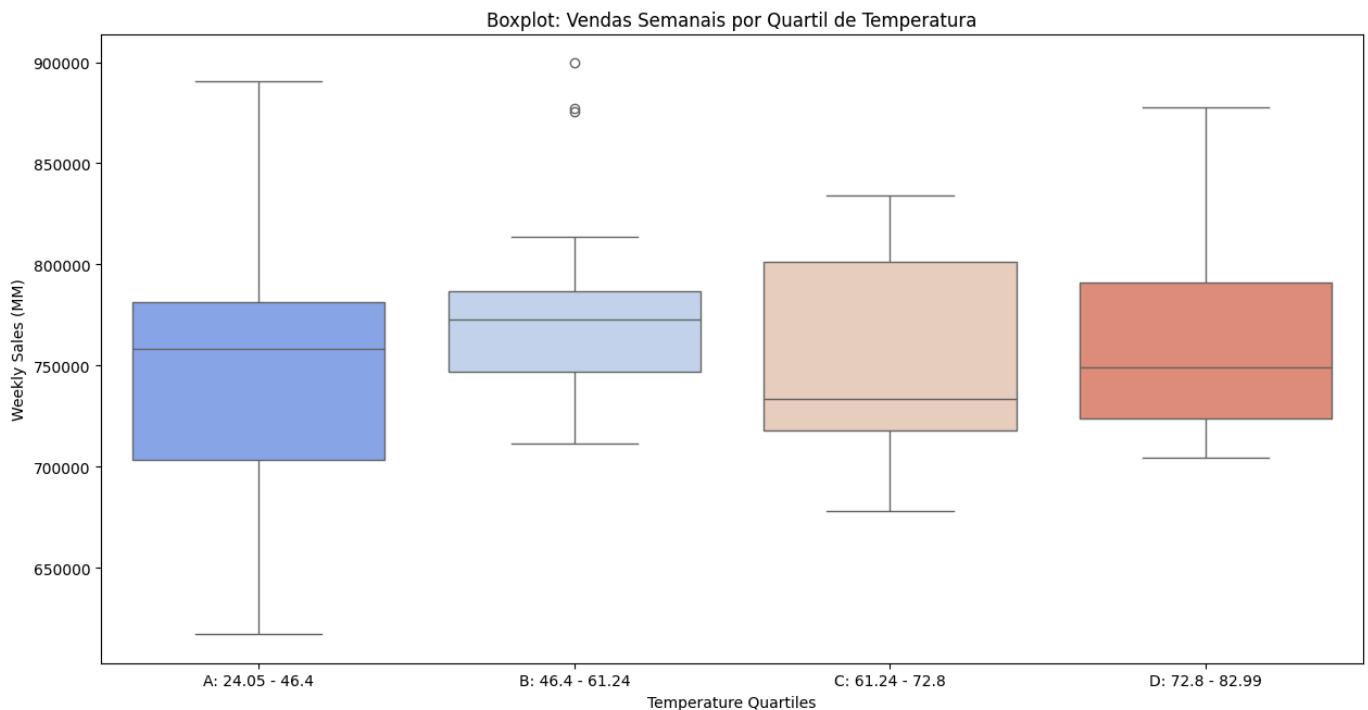


Gráfico de Vendas Semanais vs Temperatura - Loja 44 (Sem Outliers)



Boxplot de Vendas Semanais por Quartil de Temperatura - Loja 45 (Sem Outliers)



Barplot: Vendas Totais por Quartil de Temperatura

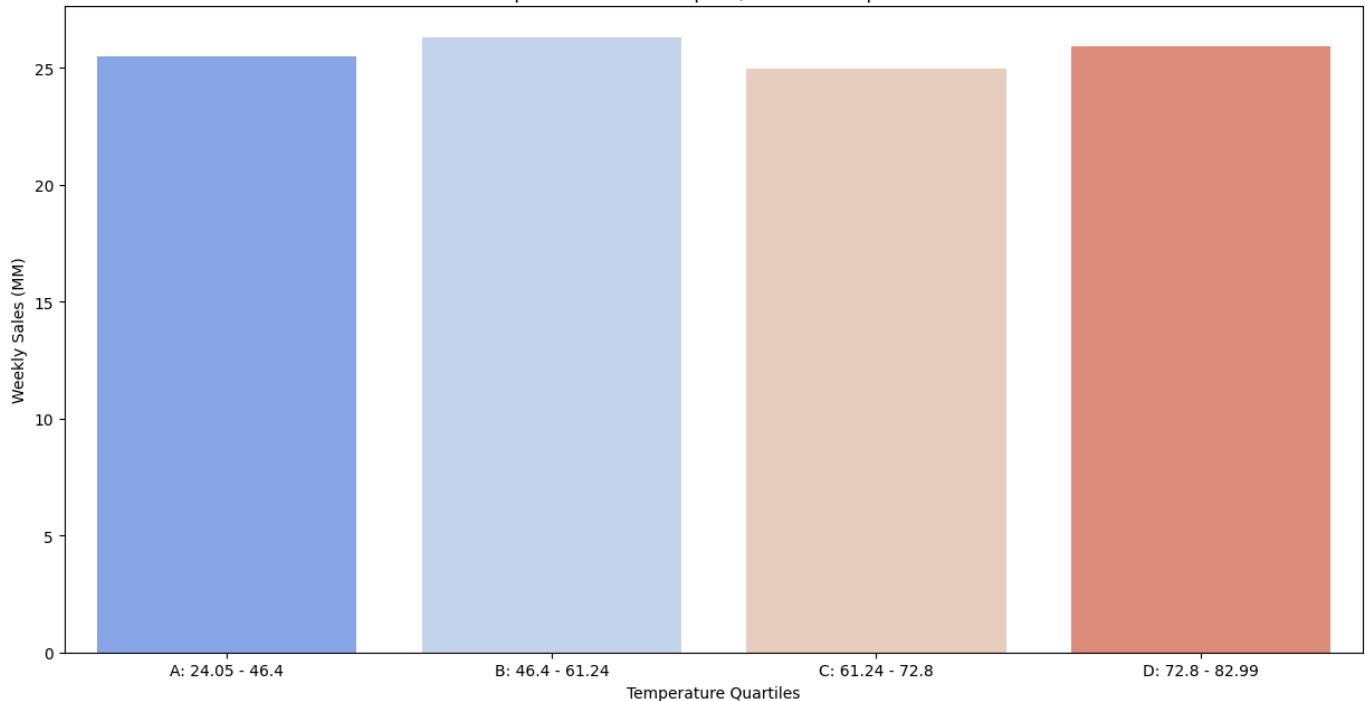
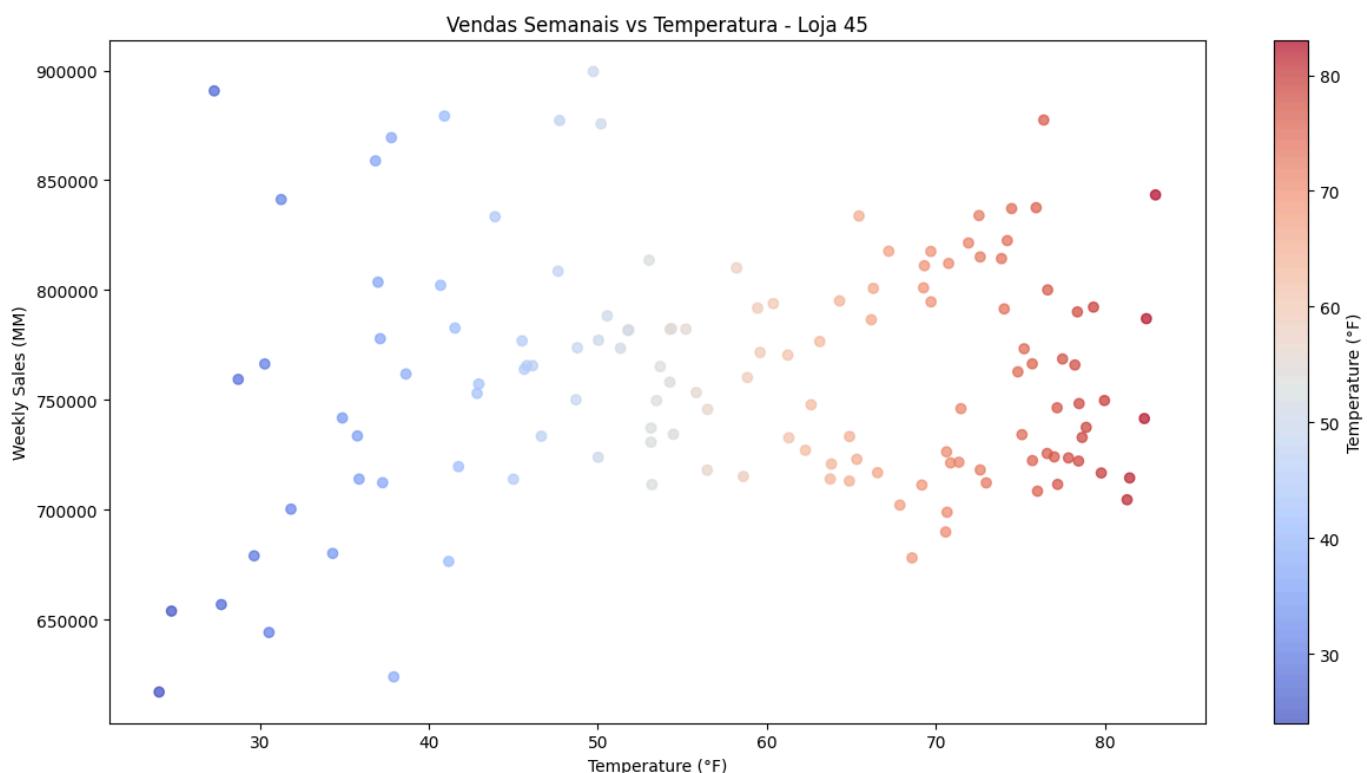


Gráfico de Vendas Semanais vs Temperatura - Loja 45 (Sem Outliers)



```
In [2]: # Importar a biblioteca pandas
import pandas as pd

# Ler os dados
data_path = 'Walmart_Sales.csv'
df = pd.read_csv(data_path)

# Identificar as 10 piores semanas
worst_dates = df.nsmallest(5, 'Weekly_Sales')[['Date', 'Weekly_Sales', 'Temperature']]
print("As 5 piores semanas (menores vendas e temperaturas):")
print(worst_dates)
```

As 5 piores semanas (menores vendas e temperaturas):

	Date	Weekly_Sales	Temperature
4619	03-12-2010	209986.25	52.82
4614	29-10-2010	213538.32	71.34
4675	30-12-2011	215359.21	51.60
4623	31-12-2010	219804.85	52.91
4671	02-12-2011	220060.35	59.12

As 5 piores semanas mostram a necessidade de um estudo mais aprofundado sobre o impacto da temperatura nas vendas semanais.

In [3]:

```
# Importar a biblioteca pandas
import pandas as pd

# Ler os dados
data_path = 'Walmart_Sales.csv'
df = pd.read_csv(data_path)

# Identificar as 5 melhores semanas
best_dates = df.nlargest(5, 'Weekly_Sales')[['Date', 'Weekly_Sales', 'Temperature']]
print("As 5 melhores semanas (maiores vendas e temperaturas):")
print(best_dates)
```

As 5 melhores semanas (maiores vendas e temperaturas):

	Date	Weekly_Sales	Temperature
1905	24-12-2010	3818686.45	30.59
2763	24-12-2010	3766687.43	25.17
1333	24-12-2010	3749057.69	57.06
527	23-12-2011	3676388.98	35.92
1762	24-12-2010	3595903.20	34.90

Quando fechamos o resultado para as melhores 5 semanas ficam evidentes a temperatura baixa , só que devido as datas sempre serem no natal podemos descartar esse fato (outliers), 23 ou 24/12

In [6]:

```
import pandas as pd

# Ler os dados
data_path = 'Walmart_Sales.csv'
df = pd.read_csv(data_path)

# Converter a coluna 'Date' para o formato datetime
df['Date'] = pd.to_datetime(df['Date'], format='%d-%m-%Y')

# Função para determinar a estação com base no mês
def get_season(month):
    if month in [12, 1, 2]:
        return 'Inverno'
    elif month in [3, 4, 5]:
        return 'Primavera'
    elif month in [6, 7, 8]:
        return 'Verão'
    else:
        return 'Outono'

# Adicionar coluna 'Season' para a estação
df['Season'] = df['Date'].dt.month.apply(get_season)

# Agrupar por estação e calcular a soma das vendas
sales_by_season = df.groupby('Season')['Weekly_Sales'].sum()

print("Vendas totais por estação do ano:")
print(sales_by_season)
```

```
Vendas totais por estação do ano:  
Season  
Inverno      1.478165e+09  
Outono       1.576562e+09  
Primavera    1.796771e+09  
Verão        1.885721e+09  
Name: Weekly_Sales, dtype: float64
```

In []: Os resultados mostraram a soma total das vendas para cada estação do ano, e agora podemos observar que Verão tem a maior soma de vendas, seguido por Primavera, Outono e Inverno.

As vendas mais altas no Verão são esperadas, já que é uma estação que pode envolver promoções sazonais.

Aqui estão algumas conclusões a partir dos dados:

Tendência Observada:

Verão e Primavera lideram em termos de vendas totais, o que pode indicar que essas estações são períodos com mais tempo livre para ir às lojas.

Outono e Inverno têm vendas um pouco mais baixas. Embora o inverno envolva feriados como o Natal, isso possivelmente devido a condições climáticas mais severas ou outros fatores que influenciam as compras.

Potenciais Análises:

Fatores Climáticos: Como as vendas são mais altas no verão, o clima mais quente pode incentivar o consumo.

Eventos Sazonais: Durante a primavera e o verão, pode haver mais eventos promocionais, feriados e outras celebrações que estimulam as vendas.

Comportamento de Consumo: O comportamento do consumidor pode ser mais impulsionado por fatores sazonais, considerando o aumento de compras online durante períodos mais frios.

```
In [11]: import pandas as pd  
  
# Ler os dados  
data_path = 'Walmart_Sales.csv'  
df = pd.read_csv(data_path)  
  
# Converter a coluna 'Date' para o formato datetime  
df['Date'] = pd.to_datetime(df['Date'], format='%d-%m-%Y')  
  
# Definir as faixas de temperatura conforme solicitado  
bins = [-float('inf'), 30, 50, 70, 100] # Ajuste os intervalos, excluindo >100°C  
labels = ['<30°C', '30-50°C', '50-70°C', '70-100°C'] # Rótulos para as faixas  
  
# Criar uma nova coluna 'Temp_Range' com as faixas de temperatura  
df['Temp_Range'] = pd.cut(df['Temperature'], bins=bins, labels=labels)  
  
# Calcular a média de vendas por faixa de temperatura, com o parâmetro 'observed=False' para evitar zeros  
sales_by_temp_range = df.groupby('Temp_Range', observed=False)[['Weekly_Sales']].mean()  
  
print("Média de vendas por faixa de temperatura:")  
print(sales_by_temp_range)
```

Média de vendas por faixa de temperatura:

Temp_Range	Mean Sales
<30°C	1.017733e+06
30-50°C	1.118767e+06
50-70°C	1.047742e+06
70-100°C	1.006549e+06

Name: Weekly_Sales, dtype: float64

Temperaturas abaixo de 30°F têm vendas médias mais baixas, enquanto temperaturas entre 50°F e 70°F apresentam vendas médias mais altas. A faixa de temperatura entre 30°F e 50°F mostra um aumento nas vendas médias, mas não tão acentuado quanto a faixa de 50°F a 70°F.

In [32]:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

# Carregar os dados
data_path = 'Walmart_Sales.csv'
df = pd.read_csv(data_path)

# Converter a coluna 'Date' para o formato de data, caso não esteja
df['Date'] = pd.to_datetime(df['Date'], errors='coerce')

# Remover outliers usando o IQR (Intervalo Interquartil) para todas as lojas
Q1 = df['Weekly_Sales'].quantile(0.25)
Q3 = df['Weekly_Sales'].quantile(0.75)
IQR = Q3 - Q1
lower_bound = Q1 - 1.5 * IQR
upper_bound = Q3 + 1.5 * IQR
df = df[(df['Weekly_Sales'] > lower_bound) & (df['Weekly_Sales'] < upper_bound)]

# Identificar temperaturas extremas (abaixo de 30°F ou acima de 90°F) para todas as lojas
extreme_temperatures = df[(df['Temperature'] < 30) | (df['Temperature'] > 90)]

# Criar o gráfico de dispersão com linha de tendência para todas as lojas
def create_scatter_plot_with_trend(df, extreme_temperatures):
    if df.empty:
        print("Sem dados suficientes para gerar o gráfico.")
        return

    # Criar o gráfico para todas as lojas
    plt.figure(figsize=(15, 7.5))

    # Gráfico de dispersão
    plt.scatter(df['Temperature'], df['Weekly_Sales'], color='blue', alpha=0.5, label='Dados de Vendas Semanais')

    # Ordenar os dados por temperatura
    df_sorted = df.sort_values('Temperature')

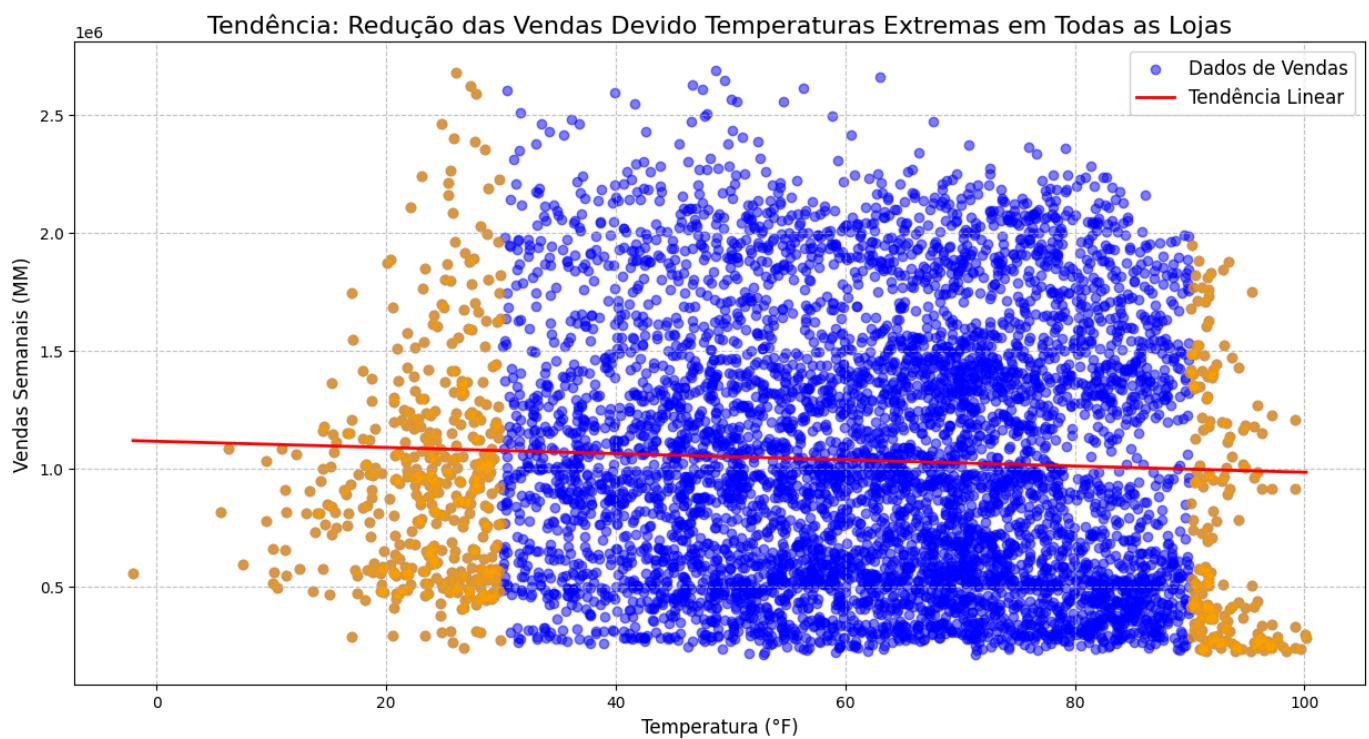
    # Linha de tendência (Regressão Linear)
    z = np.polyfit(df_sorted['Temperature'], df_sorted['Weekly_Sales'], 1) # Ajustar uma reta
    p = np.poly1d(z) # Criar a função da reta
    plt.plot(df_sorted['Temperature'], p(df_sorted['Temperature']), color='red', linewidth=2, label='Tendência')

    # Detalhes do gráfico
    plt.title('Tendência: Redução das Vendas Devido Temperaturas Extremas em Todas as Lojas', fontsize=14)
    plt.xlabel('Temperatura (°F)', fontsize=12)
    plt.ylabel('Vendas Semanais (MM)', fontsize=12)
    plt.grid(True, linestyle='--', alpha=0.7)
    plt.legend(fontsize=12)
    plt.suptitle('Evidência da Redução de Vendas Devido Temperaturas Extremas', fontsize=14)

    # Destacar as temperaturas extremas no gráfico
    plt.scatter(extreme_temperatures['Temperature'], extreme_temperatures['Weekly_Sales'],
               color='orange', alpha=0.7, label='Temperaturas Extremas')

    plt.show()

# Gerar o gráfico para todas as lojas
create_scatter_plot_with_trend(df, extreme_temperatures)
```



Análise das Temperaturas Extremas para todas as lojas:

Temperaturas Extremas (abaixo de 30°F ou acima de 90°F): Temperature Weekly_Sales 910 -2.06 558027.77
3626 5.54 817485.14 2336 6.23 1083071.14 959 7.46 593875.46 5628 9.51 775910.43 1597 99.22
912403.67 3885 99.22 1205884.98 4657 99.66 237095.82 4707 100.07 297753.49 4599 100.14 280937.84

[587 rows x 2 columns]

Total de Vendas nas Temperaturas Extremas: R

543,418,702.43 Vendas Médias nas Temperaturas Extremas : R925,755.88

In [30]:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

# Carregar os dados
data_path = 'Walmart_Sales.csv'
df = pd.read_csv(data_path)

# Converter a coluna 'Date' para o formato de data, caso não esteja
df['Date'] = pd.to_datetime(df['Date'], errors='coerce')

# Filtrar apenas os dados da Loja 34
nstore = 34
dfplot = df[df['Store'] == nstore].copy()

# Remover outliers usando o IQR (Intervalo Interquartil)
Q1 = dfplot['Weekly_Sales'].quantile(0.25)
Q3 = dfplot['Weekly_Sales'].quantile(0.75)
IQR = Q3 - Q1
lower_bound = Q1 - 1.5 * IQR
upper_bound = Q3 + 1.5 * IQR
dfplot = dfplot[(dfplot['Weekly_Sales'] > lower_bound) & (dfplot['Weekly_Sales'] < upper_bound)]

# Identificar temperaturas extremas (abaixo de 30°F ou acima de 90°F)
extreme_temperatures = dfplot[(dfplot['Temperature'] < 30) | (dfplot['Temperature'] > 90)]

# Criar o gráfico de dispersão com Linha de tendência
def create_scatter_plot_with_trend(dfplot, extreme_temperatures, nstore):
    if dfplot.empty:
```

```

print(f"Loja {nstore}: Sem dados suficientes para gerar o gráfico.")
return

plt.figure(figsize=(15, 7.5))

# Gráfico de dispersão
plt.scatter(dfplot['Temperature'], dfplot['Weekly_Sales'], color='blue', alpha=0.5, label='Tendência')

# Ordenar os dados por temperatura
df_sorted = dfplot.sort_values('Temperature')

# Linha de tendência (Regressão Linear)
z = np.polyfit(df_sorted['Temperature'], df_sorted['Weekly_Sales'], 1) # Ajustar uma reta
p = np.poly1d(z) # Criar a função da reta
plt.plot(df_sorted['Temperature'], p(df_sorted['Temperature']), color='red', linewidth=2, label='Tendência')

# Detalhes do gráfico
plt.title(f'Tendência: Redução das Vendas Devido Temperaturas Extremas - Loja {nstore}', fontsize=14)
plt.xlabel('Temperatura (°F)', fontsize=12)
plt.ylabel('Vendas Semanais (MM)', fontsize=12)
plt.grid(True, linestyle='--', alpha=0.7)
plt.legend(fontsize=12)
plt.suptitle(f'Evidência da Redução de Vendas Devido Temperaturas Extremas- Loja {nstore}', fontsize=14)

# Destacar as temperaturas extremas no gráfico
plt.scatter(extreme_temperatures['Temperature'], extreme_temperatures['Weekly_Sales'],
           color='orange', alpha=0.7, label='Temperaturas Extremas')

plt.show()

# Gerar o gráfico para a Loja 34
create_scatter_plot_with_trend(dfplot, extreme_temperatures, nstore)

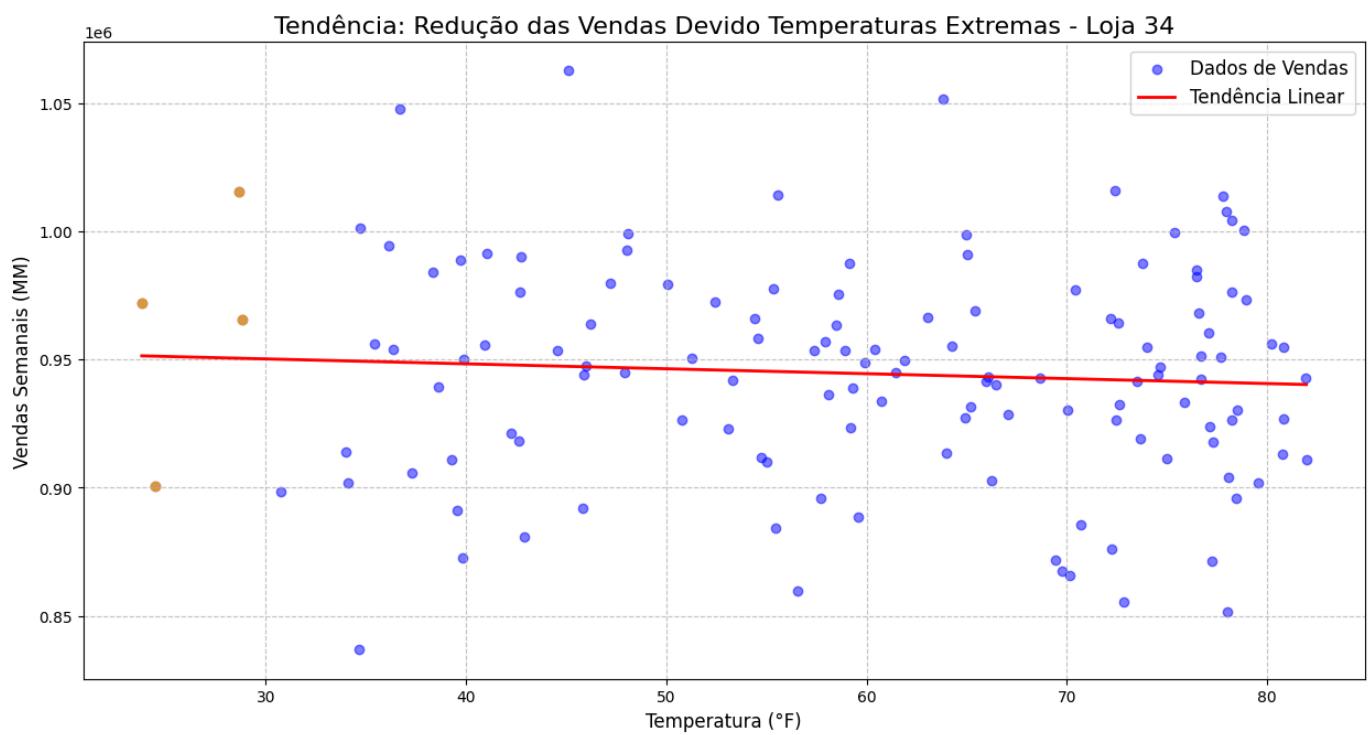
# Texto com informações extraídas
extreme_sales = extreme_temperatures[['Temperature', 'Weekly_Sales']].sort_values(by='Temperature', ascending=False)

print(f"Análise das Temperaturas Extremas para a Loja {nstore}:")

if not extreme_temperatures.empty:
    print("\nTemperaturas Extremas (abaixo de 30°F ou acima de 90°F):")
    print(extreme_sales)

    # Insights sobre as vendas nas temperaturas extremas
    print("\nAnalizando as Vendas nas Temperaturas Extremas:")
    for index, row in extreme_sales.iterrows():
        temp = row['Temperature']
        sales = row['Weekly_Sales']
        print(f"Temperatura: {temp}°F - Vendas Semanais: R${sales:,.2f}")
else:
    print("Não há dados de temperaturas extremas para essa loja.")

```



Análise das Temperaturas Extremas para a Loja 34:

Temperaturas Extremas (abaixo de 30°F ou acima de 90°F):

	Temperature	Weekly_Sales
4771	23.82	971932.87
4767	24.50	900646.94
4772	28.66	1015654.60
4818	28.84	965512.36

Analizando as Vendas nas Temperaturas Extremas:

Temperatura: 23.82°F - Vendas Semanais: R\$971,932.87

Temperatura: 24.5°F - Vendas Semanais: R\$900,646.94

Temperatura: 28.66°F - Vendas Semanais: R\$1,015,654.60

Temperatura: 28.84°F - Vendas Semanais: R\$965,512.36

In [5]: # Criar tabela das 10 lojas que menos vendem com redução de temperatura

```

def get_top_10_negative_corr(df):
    corr_list = []
    for store in df['Store'].unique():
        temp_sales_corr = df[df['Store'] == store]['Temperature'].corr(df[df['Store'] == store])
        corr_list.append((store, temp_sales_corr))

    # Ordenar por correlação negativa
    sorted_corr = sorted(corr_list, key=lambda x: x[1])[:10]

    # Converter em DataFrame para visualização
    top_10_negative_corr = pd.DataFrame(sorted_corr, columns=['Store', 'Correlation'])
    return top_10_negative_corr

# Gerar tabela
top_10_stores = get_top_10_negative_corr(df)
print(top_10_stores)

```

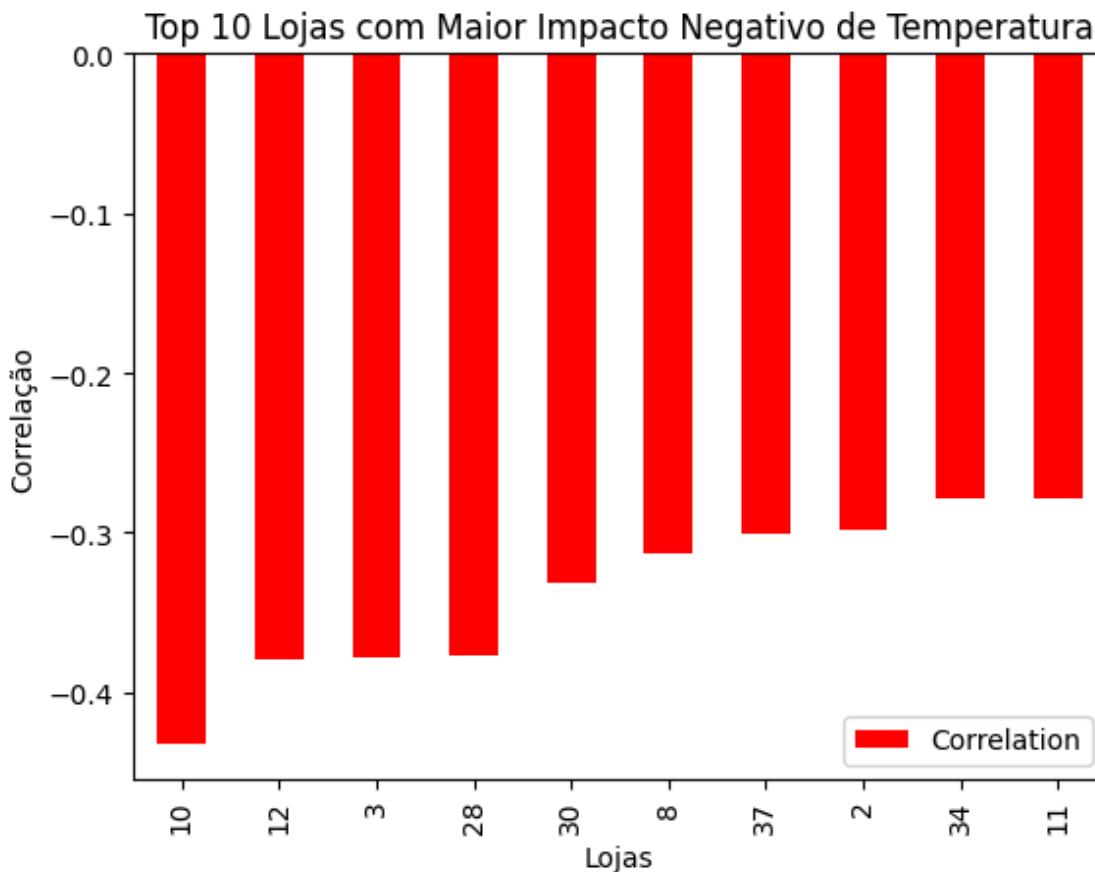
```

Store Correlation
0    10   -0.432569
1    12   -0.379416
2     3   -0.377524
3    28   -0.376506
4    30   -0.330816
5     8   -0.312324
6    37   -0.300493
7     2   -0.297673
8    34   -0.278523
9    11   -0.278503

```

A análise das 10 lojas que apresentam maior correlação negativa entre temperatura e vendas semanais evidencia uma tendência nas regiões com temperaturas severas. Essas áreas são impactadas por condições climáticas extremas, que reduzem a circulação de clientes nas lojas e afetam diretamente o volume de vendas. Esse padrão demonstra como a sazonalidade e fatores regionais influenciam o desempenho comercial, reforçando a importância de estratégias personalizadas para atender às necessidades específicas dos consumidores em locais com temperaturas extremas.

```
In [26]: top_10_stores.plot.bar(x='Store', y='Correlation', color='red', title='Top 10 Lojas com Maior Impacto Negativo de Temperatura')
plt.xlabel('Lojas')
plt.ylabel('Correlação')
plt.show()
```



```
In [32]: import os

# Certificar-se de que a pasta para salvar gráficos existe
output_dir = "gráficos"
os.makedirs(output_dir, exist_ok=True) # Cria a pasta se ela não existir
```

```
In [29]: import os
import pandas as pd
import matplotlib.pyplot as plt

# Certificar-se de que a pasta para salvar gráficos existe
output_dir = "gráficos"
os.makedirs(output_dir, exist_ok=True) # Cria a pasta se ela não existir
```

```

# Carregar os dados
data_path = 'Walmart_Sales.csv' # Substitua pelo caminho correto do seu arquivo
df = pd.read_csv(data_path)

# Converter a coluna 'Date' para o formato de data e ordenar por data
df['Date'] = pd.to_datetime(df['Date'], format='%d-%m-%Y')
df.sort_values('Date', inplace=True)

# Agrupar as vendas por semana e por temperatura
df_grouped = df.groupby(['Date']).agg({'Weekly_Sales': 'sum', 'Temperature': 'mean'}).reset_index()

# Função para criar o gráfico combinado de linha
def plot_sales_temperature_comparison(df, save_path=None):
    plt.figure(figsize=(15, 7.5))

    # Plotar vendas semanais no eixo primário
    plt.plot(df['Date'], df['Weekly_Sales'], color='dodgerblue', label='Vendas Semanais', marker='o')
    plt.ylabel('Vendas Semanais (MM)', fontsize=12, color='dodgerblue')
    plt.xlabel('Data', fontsize=12)

    # Configurar segundo eixo-y para temperatura
    ax = plt.gca()
    ax2 = ax.twinx()
    ax2.plot(df['Date'], df['Temperature'], color='coral', linestyle='--', label='Temperatura (' + str(df['Temperature'].min()) + ' a ' + str(df['Temperature'].max()) + ' °F)')
    ax2.set_ylabel('Temperatura (°F)', fontsize=12, color='coral')
    ax2.tick_params(axis='y', labelcolor='coral')

    # Adicionar título e legendas
    plt.title(f'Comparação de Vendas Totais e Temperatura', fontsize=16)
    ax.legend(loc='upper left', fontsize=10)
    ax2.legend(loc='upper right', fontsize=10)
    plt.grid(True, linestyle='--', alpha=0.7)

    # Salvar gráfico
    if save_path:
        plt.savefig(f"{save_path}/sales_temperature_total.png")
        print(f"Gráfico salvo em: {save_path}/sales_temperature_total.png")

    plt.show()

# Gerar o gráfico
plot_sales_temperature_comparison(df_grouped, save_path=output_dir)

# Resumo textual dos melhores indicadores
max_sales = df_grouped.loc[df_grouped['Weekly_Sales'].idxmax()]
max_temp = df_grouped.loc[df_grouped['Temperature'].idxmax()]

melhor_periodo_vendas = max_sales['Date'].strftime('%d-%m-%Y')
melhor_venda = max_sales['Weekly_Sales']
melhor_periodo_temp = max_temp['Date'].strftime('%d-%m-%Y')
melhor_temperatura = max_temp['Temperature']

# Exibir o resumo
resumo = f"""
Resumo dos Melhores Indicadores:
- Melhor pico de vendas: {melhor_venda:.2f} MM em {melhor_periodo_vendas}
- Melhor pico de temperatura: {melhor_temperatura:.2f}°F em {melhor_periodo_temp}
"""

print(resumo)

# Análise da relação entre Temperatura e Vendas
correlacao = df_grouped[['Weekly_Sales', 'Temperature']].corr().iloc[0, 1]

# Conclusão
conclusao = f"""
Análise da Relação Temperatura x Vendas:

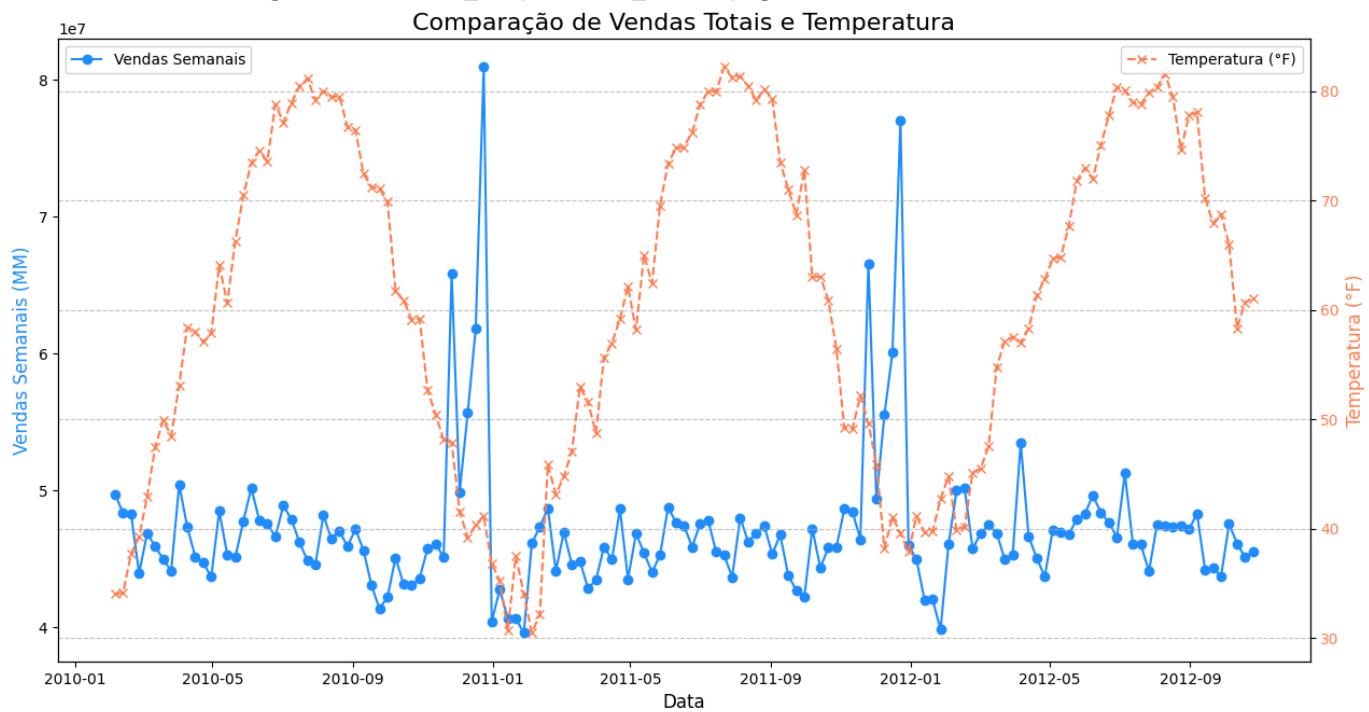
```

```

A correlação entre as vendas semanais e a temperatura é de {correlacao:.2f}.
Isso sugere uma { 'correlação positiva' if correlacao > 0 else 'correlação negativa' if correlacao
Ou seja, um aumento (ou diminuição) na temperatura parece (ou não) estar associado a um aumento
"""
print(conclusao)

```

Gráfico salvo em: gráficos/sales_temperature_total.png



Resumo dos Melhores Indicadores:

- Melhor pico de vendas: 80931415.60 MM em 24-12-2010
- Melhor pico de temperatura: 82.18°F em 22-07-2011

Análise da Relação Temperatura x Vendas:

A correlação entre as vendas semanais e a temperatura é de -0.16.

Isso sugere uma correlação negativa entre essas duas variáveis.

Ou seja, um aumento (ou diminuição) na temperatura parece (ou não) estar associado a um aumento (ou diminuição) nas vendas.

Conclusão: A variação de temperatura pode ter um impacto importante sobre as vendas, com certos picos de vendas ocorrendo durante temperaturas mais altas ou mais baixas, dependendo do comportamento do consumidor e da natureza dos produtos vendidos.