

Insights from the Data

1. Top 5 countries with the most wine reviews

- US 48153
- France 19932
- Italy 11370
- Portugal 4038
- Chile 3630

Code:

```
top_countries = df['country'].value_counts().head(5)
print("Top 5 countries with the most wine reviews:")
print(top_countries)
```

2. Most common wine varieties

- Pinot Noir 13272
- Chardonnay 11753
- Cabernet Sauvignon 9472
- Red Blend 8946
- Bordeaux-style Red Blend 6915

Code:

```
common_varieties = df['variety'].value_counts().head(5)
print("\nMost common wine varieties:")
print(common_varieties)
```

3. Correlation between points and price = 0.4156511636427041

Code :

```
correlation = df['points'].corr(df['price'])
print("\nCorrelation between points and price:")
print(correlation)
```

4. Top 5 varieties with the highest average rating

- Nebbiolo 90.251070
- Grüner Veltliner 89.980669
- Champagne Blend 89.663324
- Riesling 89.450183
- Pinot Noir 89.411468

Code:

```
top_varieties_rating = df.groupby('variety')['points'].mean().nlargest(5)
print("\ntop 5 varieties with the highest average rating:")
print(top_varieties_rating)
```

5. Top 10 wineries with the most wine varieties

- Testarossa 217
- Williams Selyem 203

- Louis Latour 199
- Georges Duboeuf 195
- Chateau Ste. Michelle 187
- Wines & Winemakers 181
- DFJ Vinhos 158
- Columbia Crest 144
- Concha y Toro 135
- Kendall-Jackson 128

Code :

```
top_wineres_varieties = df['winery'].value_counts().nlargest(10)
print("\ntop 10 wineries with the most wine varieties:")
print(top_wineres_varieties)
```

6. Average rating by province

- Südburgenland 93.000000
- Tokaji 92.000000
- Mittelrhein 92.000000
- Puente Alto 91.733333
- Wachau 91.723724
- England 91.581081
- Santa Cruz 91.500000
- Kamptal 91.466830
- Beira Atlantico 91.333333
- Traisental 91.255814

Code:

```
average_rating_by_province =
df.groupby('province')['points'].mean().sort_values(ascending=False).head(10)
print("\naverage rating by province:")
print(average_rating_by_province)
```