

# *Wine Reviews Analysis*

*Wireframe Documentation*

As per the problem statement, we have performed EDA on Wine Reviews Dataset: -

## Data Preparation and Cleaning

- Total number of Records : 29665
- Total number of Attribute : 14
- df.head()

| Unnamed: 0 | country  | description                                       | designation  | points | price | province          | region_1          | region_2          | taster_name   | taster_twitter_handle | title   | variety        | winery              |
|------------|----------|---|--------------|--------|-------|-------------------|-------------------|-------------------|---------------|-----------------------|---|----------------|---------------------|
| 0          | Italy    | Aromas include tropical fruit, broom, brimston... | Vulkà Bianco | 87     | NaN   | Sicily & Sardinia | Etna              | NaN               | Kerin O'Keefe | @kerinokeefe          | Nicosia 2013 Vulkà Bianco (Etna)              | White Blend    | Nicosia             |
| 1          | Portugal | This is ripe and fruity, a wine that is smooth... | Avidagos     | 87     | 15.0  | Douro             | NaN               | NaN               | Roger Voss    | @vossroger            | Quinta dos Avidagos 2011 Avidagos Red (Douro) | Portuguese Red | Quinta dos Avidagos |
| 2          | US       | Tart and snappy, the flavors of lime flesh and... | NaN          | 87     | 14.0  | Oregon            | Willamette Valley | Willamette Valley | Paul Gregutt  | @paulgwine            | Rainstorm 2013 Pinot Gris (Willamette Valley) | Pinot Gris     | Rainstorm           |

- df.info()

Data columns (total 14 columns):

| #  | Column                | Non-Null Count | Dtype   |
|----|-----------------------|----------------|---------|
| 0  | Unnamed: 0            | 29665 non-null | int64   |
| 1  | country               | 29650 non-null | object  |
| 2  | description           | 29665 non-null | object  |
| 3  | designation           | 21216 non-null | object  |
| 4  | points                | 29665 non-null | int64   |
| 5  | price                 | 27540 non-null | float64 |
| 6  | province              | 29650 non-null | object  |
| 7  | region_1              | 24699 non-null | object  |
| 8  | region_2              | 11411 non-null | object  |
| 9  | taster_name           | 23544 non-null | object  |
| 10 | taster_twitter_handle | 22428 non-null | object  |
| 11 | title                 | 29665 non-null | object  |
| 12 | variety               | 29665 non-null | object  |
| 13 | winery                | 29665 non-null | object  |

dtypes: float64(1), int64(2), object(11)  
memory usage: 3.2+ MB

# Missing Values

```
df.isnull().sum()
```

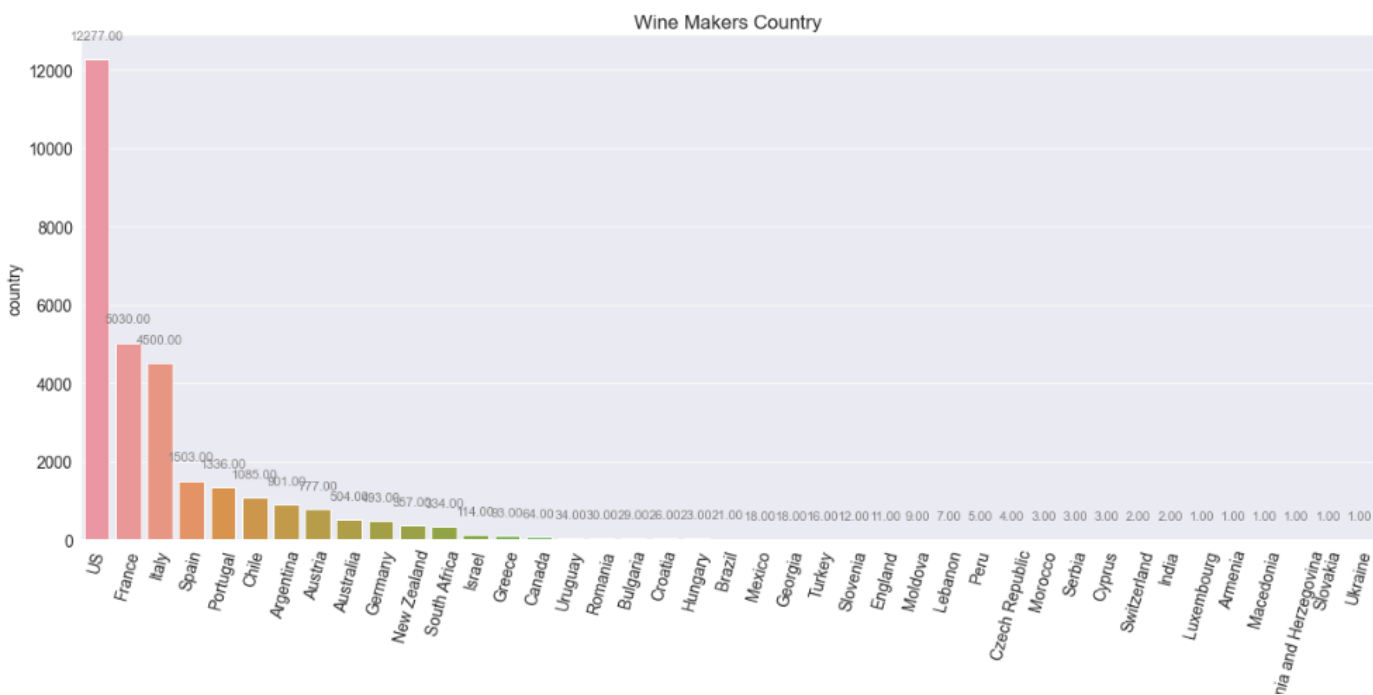
```
country          15
description       0
designation      8449
points           0
price            2125
province         15
region_1         4966
region_2        18254
taster_name      6121
taster_twitter_handle 7237
title            0
variety          0
winery           0
dtype: int64
```

- df.describe()

|       | points       | price        |
|-------|--------------|--------------|
| count | 29665.000000 | 27540.000000 |
| mean  | 88.385539    | 34.990123    |
| std   | 3.010396     | 40.472213    |
| min   | 80.000000    | 4.000000     |
| 25%   | 86.000000    | 17.000000    |
| 50%   | 88.000000    | 25.000000    |
| 75%   | 91.000000    | 42.000000    |
| max   | 100.000000   | 2500.000000  |

# Data Visualization

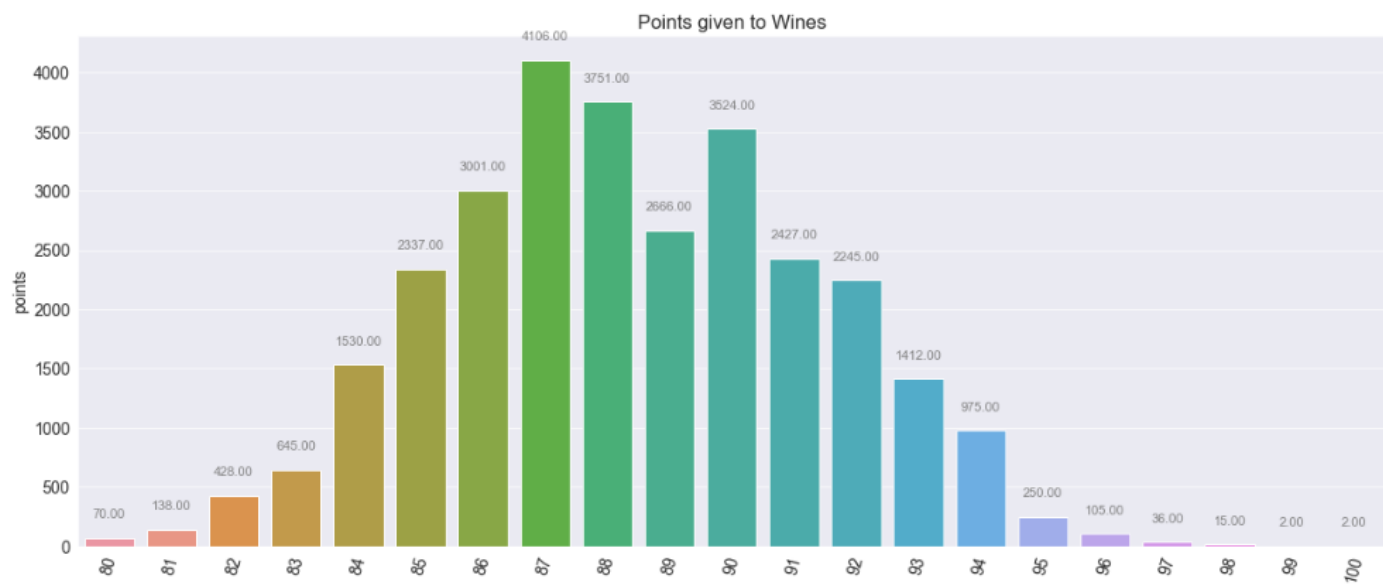
## ✓ Wine Makers Country



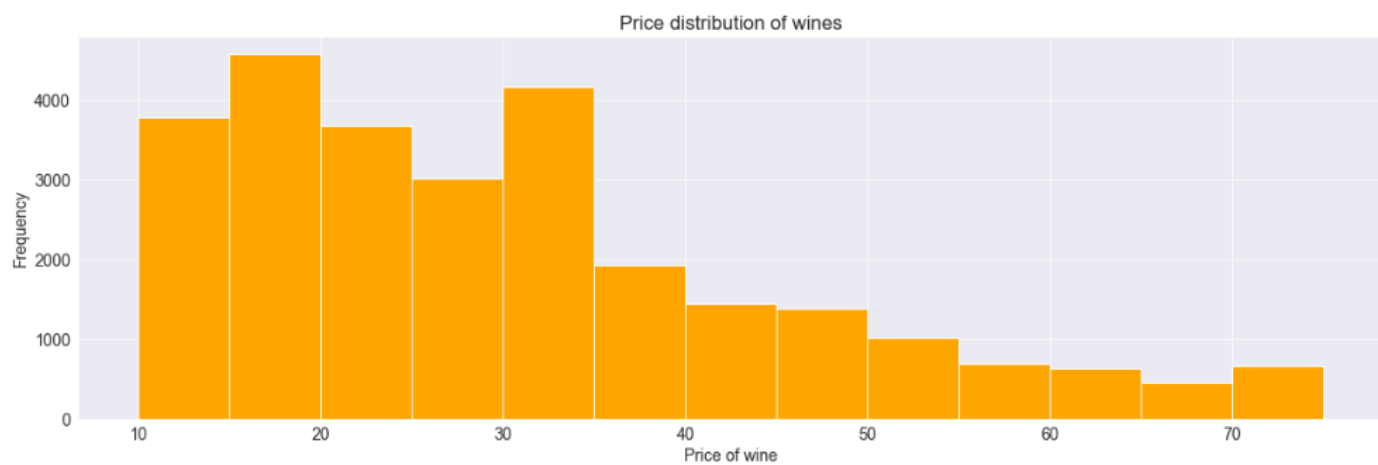
## ✓ Points given to Wines

|     |      |
|-----|------|
| 87  | 4106 |
| 88  | 3751 |
| 90  | 3524 |
| 86  | 3001 |
| 89  | 2666 |
| 91  | 2427 |
| 85  | 2337 |
| 92  | 2245 |
| 84  | 1530 |
| 93  | 1412 |
| 94  | 975  |
| 83  | 645  |
| 82  | 428  |
| 95  | 250  |
| 81  | 138  |
| 96  | 105  |
| 80  | 70   |
| 97  | 36   |
| 98  | 15   |
| 99  | 2    |
| 100 | 2    |

Name: points, dtype: int64

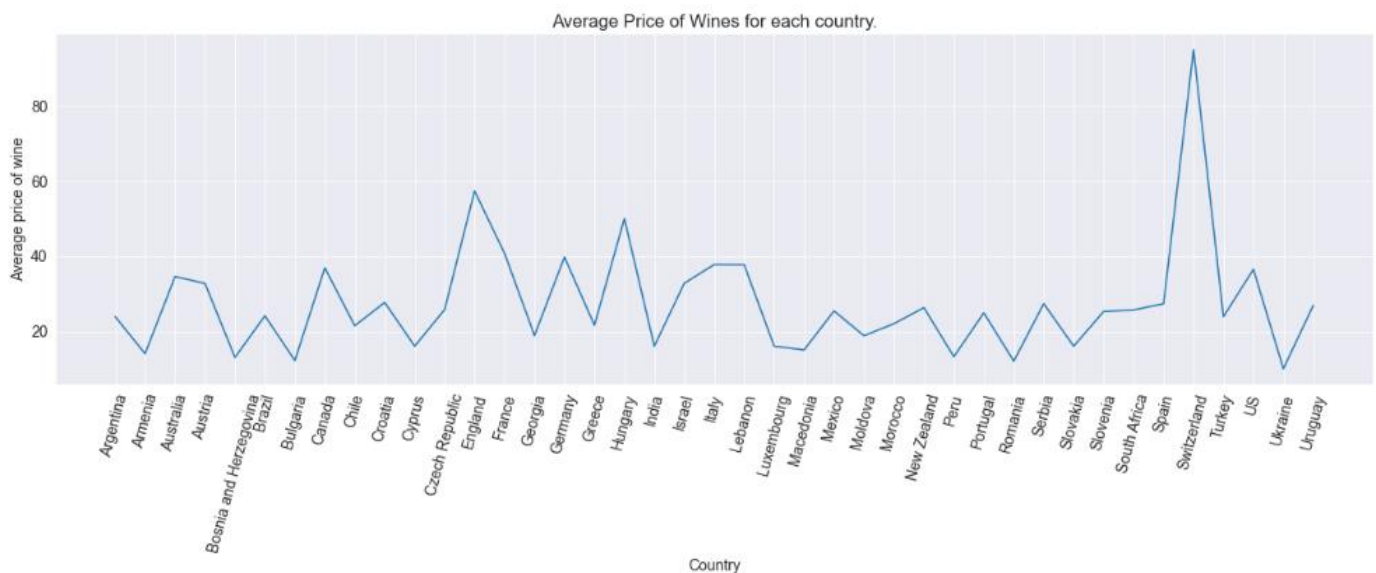


✓ Price distribution of wines

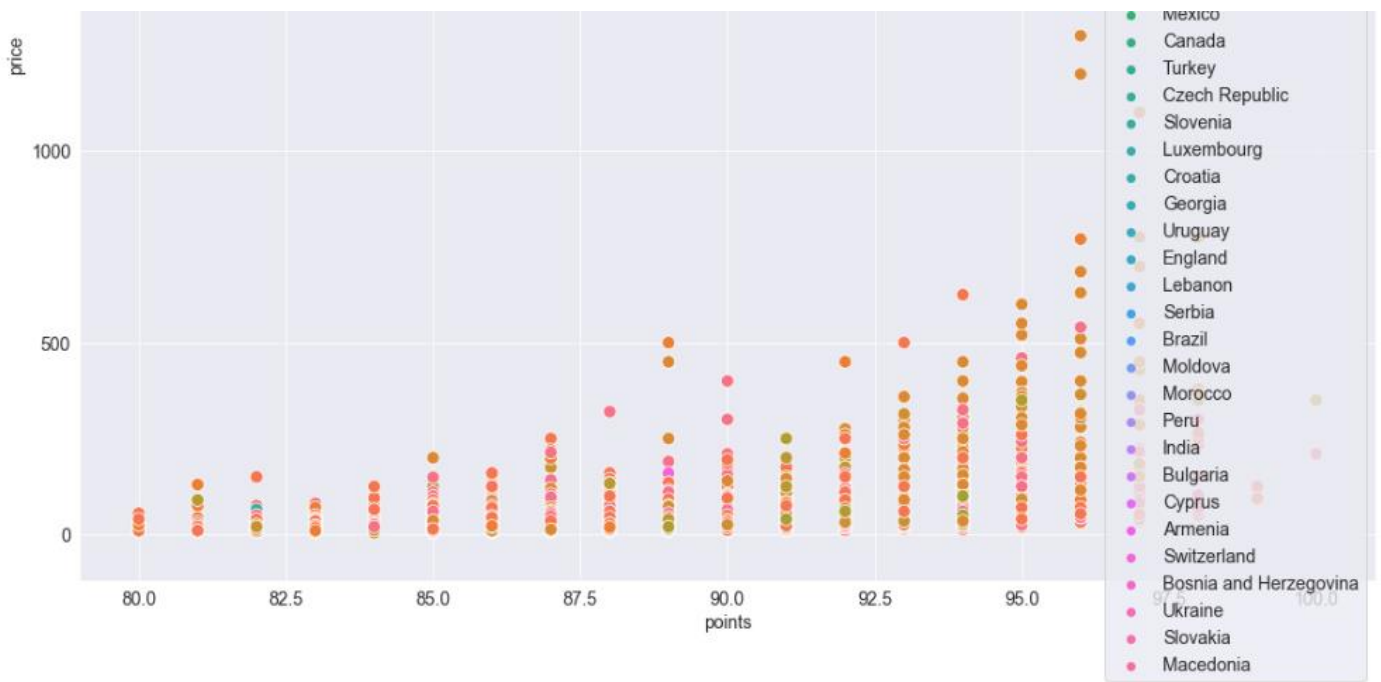


## ✓ Average Price of Wines for each country

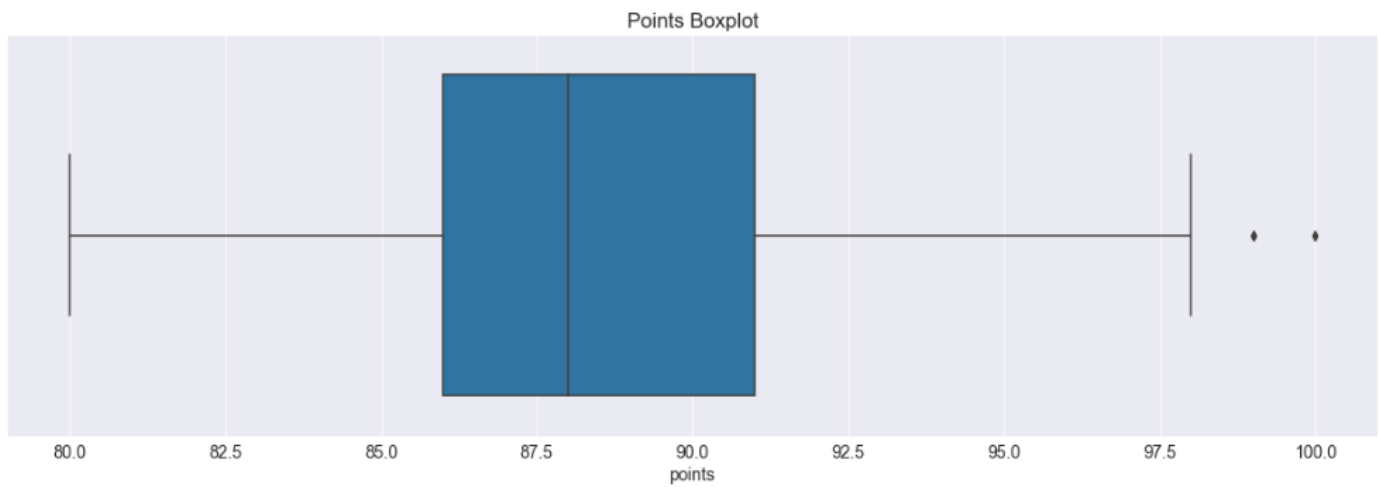
|                        | points    | price     |
|------------------------|-----------|-----------|
| country                |           |           |
| Argentina              | 86.618202 | 23.884485 |
| Armenia                | 87.000000 | 14.000000 |
| Australia              | 88.496032 | 34.547423 |
| Austria                | 90.248391 | 32.721731 |
| Bosnia and Herzegovina | 85.000000 | 13.000000 |
| Brazil                 | 83.904762 | 24.093357 |
| Bulgaria               | 87.551724 | 12.172414 |
| Canada                 | 89.390625 | 36.874846 |
| Chile                  | 86.593548 | 21.447799 |
| Croatia                | 87.615385 | 27.653846 |
| Cyprus                 | 85.666667 | 16.000000 |
| Czech Republic         | 86.500000 | 25.750000 |
| England                | 91.636364 | 57.363636 |
| France                 | 88.802783 | 40.748104 |
| Georgia                | 87.333333 | 18.833333 |
| Germany                | 89.813387 | 39.758400 |
| Greece                 | 87.387097 | 21.591292 |
| Hungary                | 89.260870 | 50.000000 |
| India                  | 91.000000 | 16.000000 |



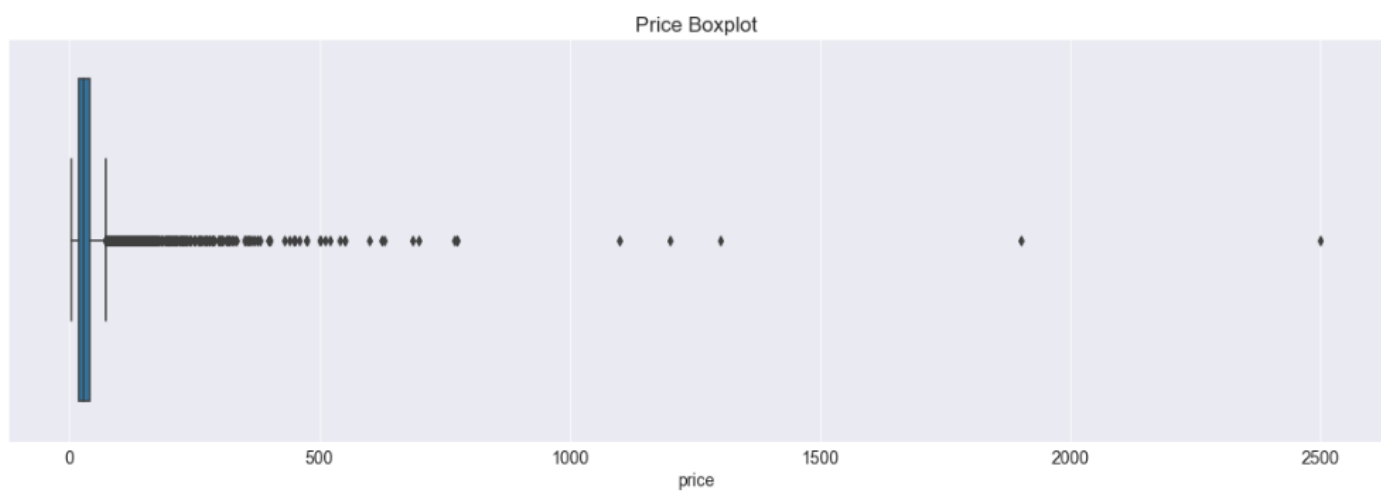
## ✓ Scatter Plot of Points and Price



## ✓ Outliers in Points

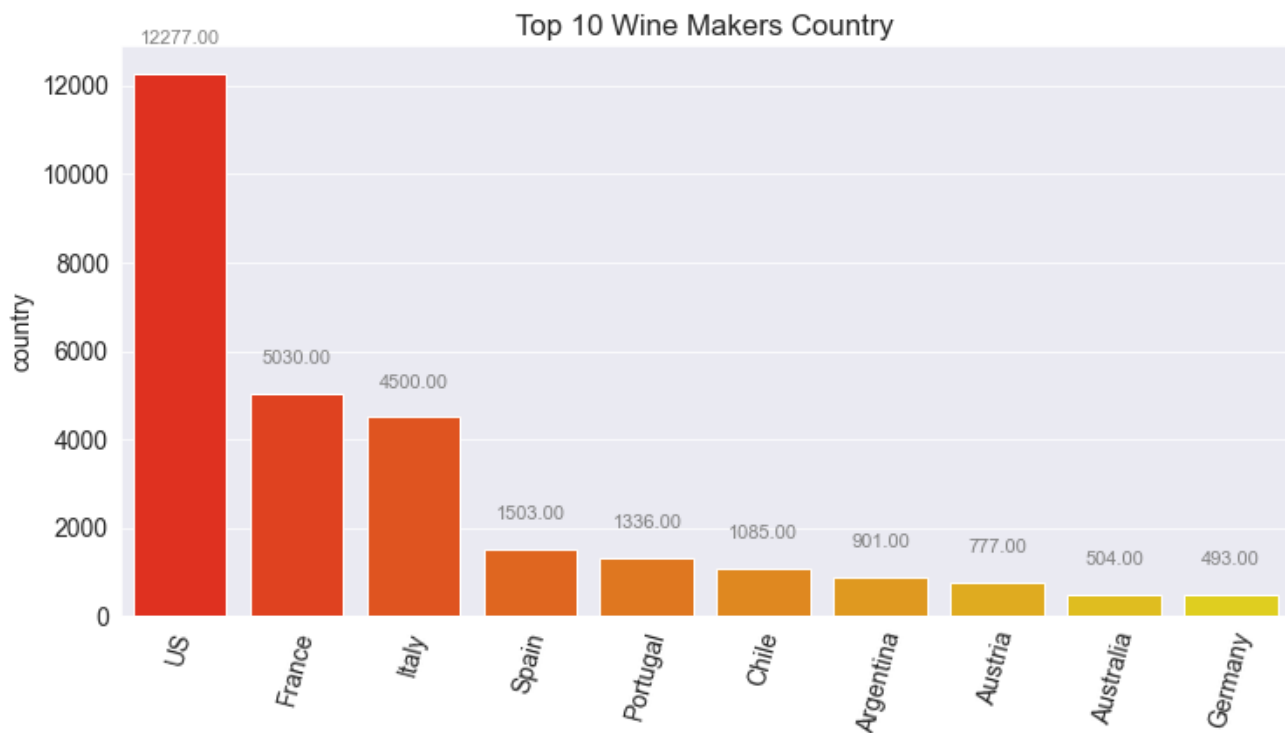


## ✓ Outliers in Price



# Answering Problem Statement

## Q1. Top 10 Wine makers Country



## Q2. Top 10 wine varieties are rated the best

|       | country   | variety            | points | price |
|-------|-----------|--------------------|--------|-------|
| 7335  | Italy     | Prugnolo Gentile   | 100    | 210.0 |
| 345   | Australia | Muscat             | 100    | 350.0 |
| 1556  | US        | Cabernet Sauvignon | 99     | 125.0 |
| 1557  | US        | Pinot Noir         | 99     | 94.0  |
| 16437 | Italy     | Nebbiolo           | 98     | 300.0 |
| 346   | Australia | Muscadelle         | 98     | 350.0 |
| 16186 | Germany   | Riesling           | 98     | 775.0 |
| 16435 | Italy     | Nebbiolo           | 98     | 70.0  |
| 16436 | Italy     | Nebbiolo           | 98     | 151.0 |
| 16434 | Italy     | Nebbiolo           | 98     | 60.0  |



Q3. 10 wine varieties are rated the least

|       | country   | variety            | points | price |
|-------|-----------|--------------------|--------|-------|
| 23101 | Argentina | Malbec             | 80     | 5.0   |
| 12689 | US        | Cabernet Sauvignon | 80     | 28.0  |
| 16510 | Italy     | Sangiovese         | 80     | 11.0  |
| 28234 | US        | Petit Verdot       | 80     | 40.0  |
| 28233 | US        | Red Blend          | 80     | 55.0  |
| 28231 | Spain     | Sparkling Blend    | 80     | 27.0  |
| 28230 | US        | Sauvignon Blanc    | 80     | 10.0  |
| 28229 | US        | Sauvignon Blanc    | 80     | 18.0  |
| 28228 | US        | Sauvignon Blanc    | 80     | 18.0  |
| 28227 | Spain     | Garnacha           | 80     | 8.0   |

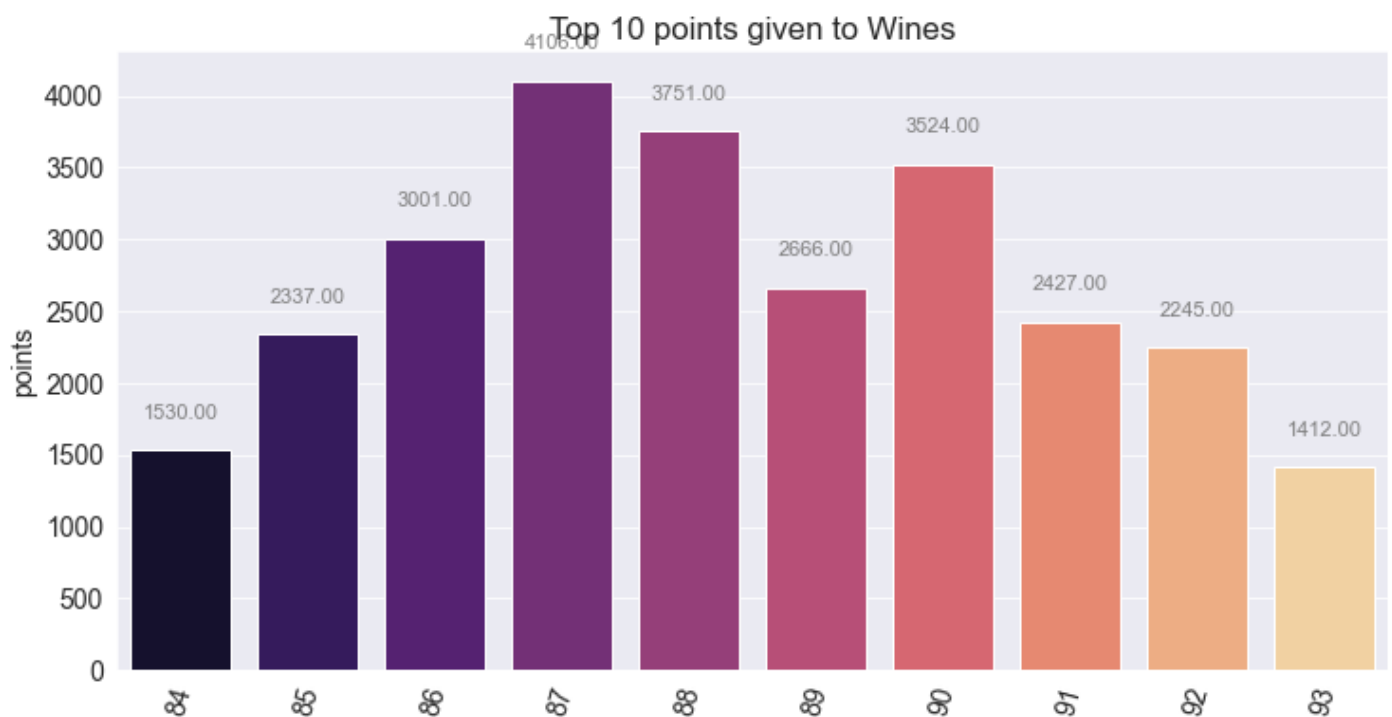
Q4. Which is costliest Wine Variety?

|       | country | variety                    | points | price  |
|-------|---------|----------------------------|--------|--------|
| 15504 | France  | Bordeaux-style Red Blend   | 96     | 2500.0 |
| 1558  | France  | Bordeaux-style Red Blend   | 98     | 1900.0 |
| 1575  | France  | Bordeaux-style Red Blend   | 96     | 1300.0 |
| 1571  | France  | Bordeaux-style Red Blend   | 96     | 1200.0 |
| 1566  | France  | Bordeaux-style Red Blend   | 97     | 1100.0 |
| 347   | Germany | Riesling                   | 97     | 775.0  |
| 16186 | Germany | Riesling                   | 98     | 775.0  |
| 15510 | Spain   | Carignan                   | 96     | 770.0  |
| 15497 | France  | Bordeaux-style White Blend | 97     | 698.0  |
| 20743 | France  | Bordeaux-style Red Blend   | 96     | 685.0  |

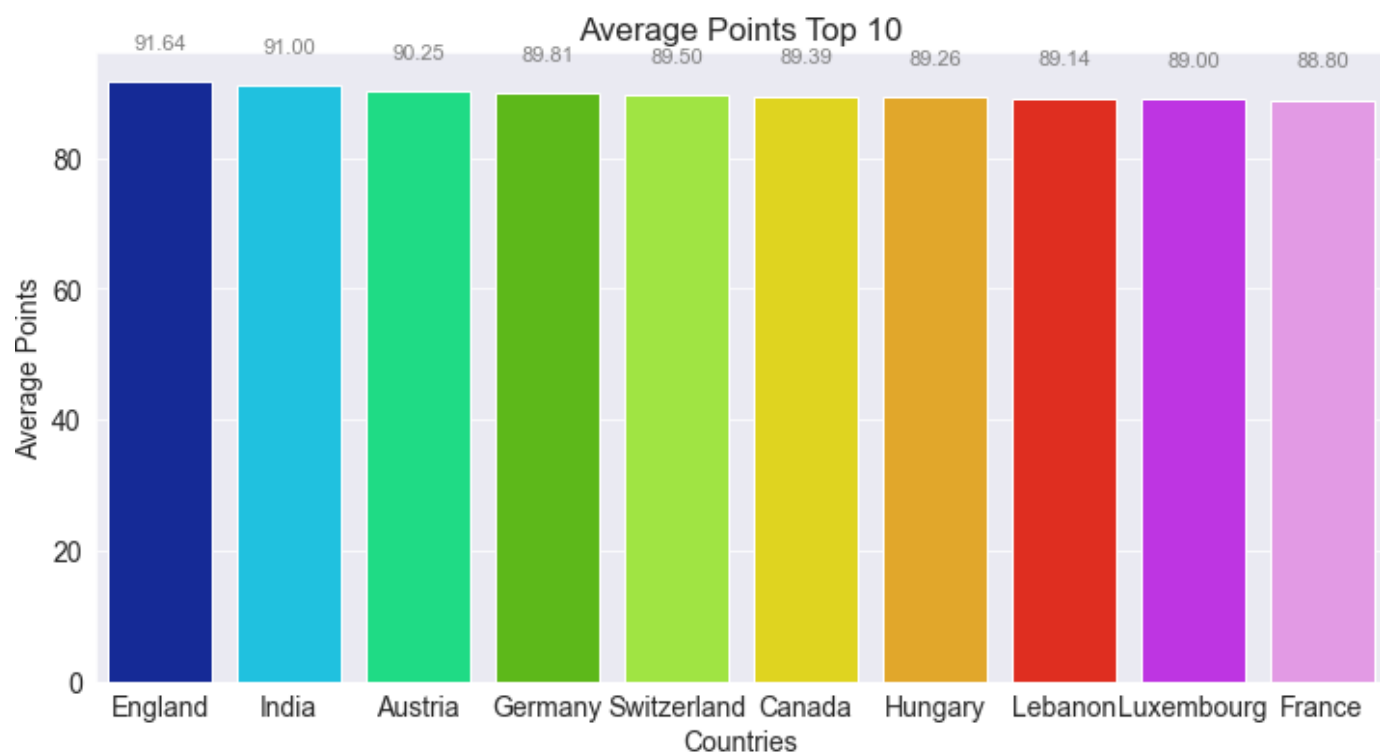
## Q5. Which is Cheapest Wine Variety?

|       | country   | variety            | points | price |
|-------|-----------|--------------------|--------|-------|
| 1987  | Spain     | Syrah              | 85     | 4.0   |
| 20148 | US        | White Blend        | 85     | 4.0   |
| 29217 | Argentina | Malbec-Syrah       | 84     | 4.0   |
| 24256 | Chile     | Sauvignon Blanc    | 87     | 5.0   |
| 5789  | France    | Cabernet Sauvignon | 83     | 5.0   |
| 7605  | Chile     | Cabernet Sauvignon | 83     | 5.0   |
| 23101 | Argentina | Malbec             | 80     | 5.0   |
| 3950  | Portugal  | Rosé               | 83     | 5.0   |
| 24547 | Portugal  | Portuguese Red     | 85     | 5.0   |
| 21864 | US        | Pinot Grigio       | 85     | 5.0   |

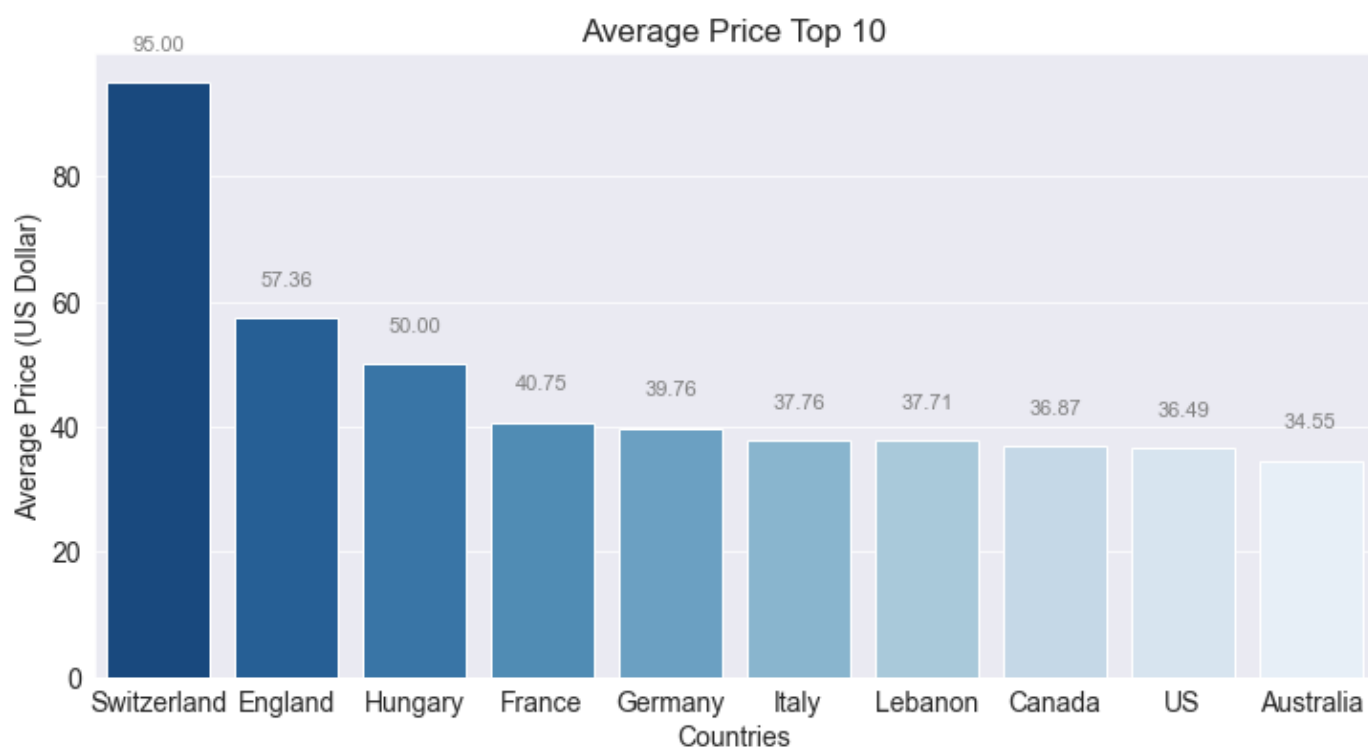
## Q6. Top 10 points given to Wines



## Q7. Average points scored by Top 10 Countries

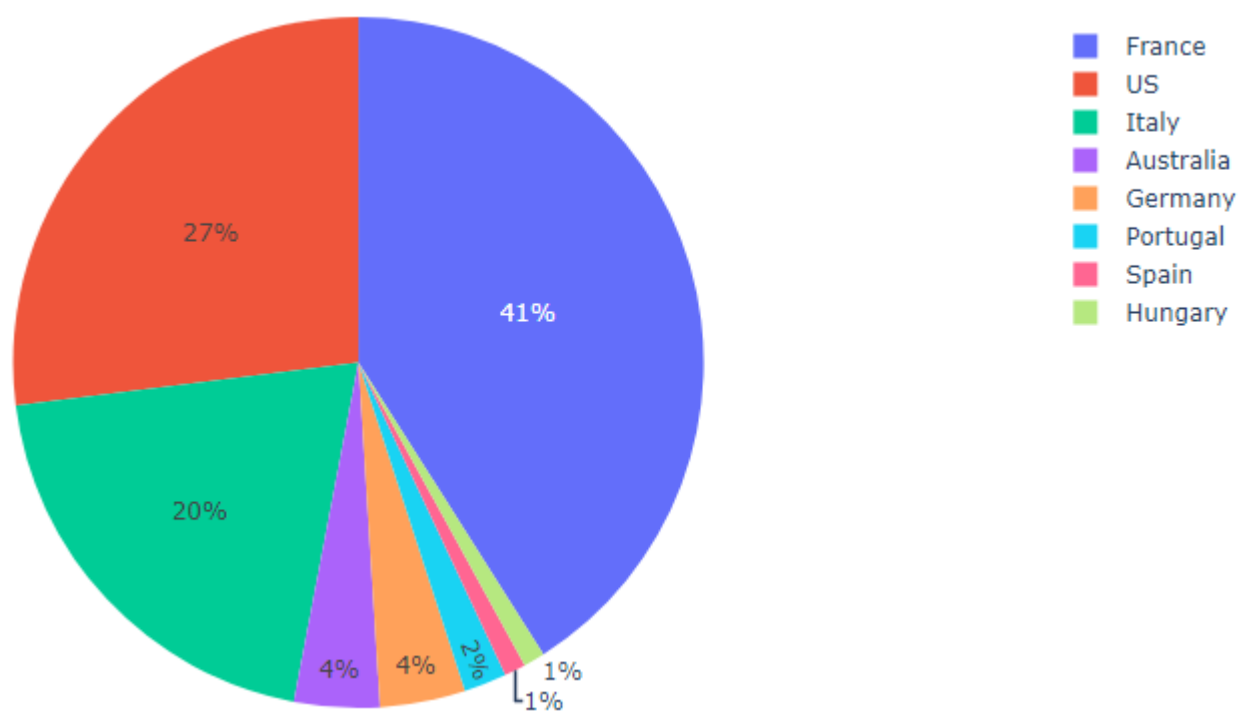


## Q8. Average Wine Price by Top 10 Countries

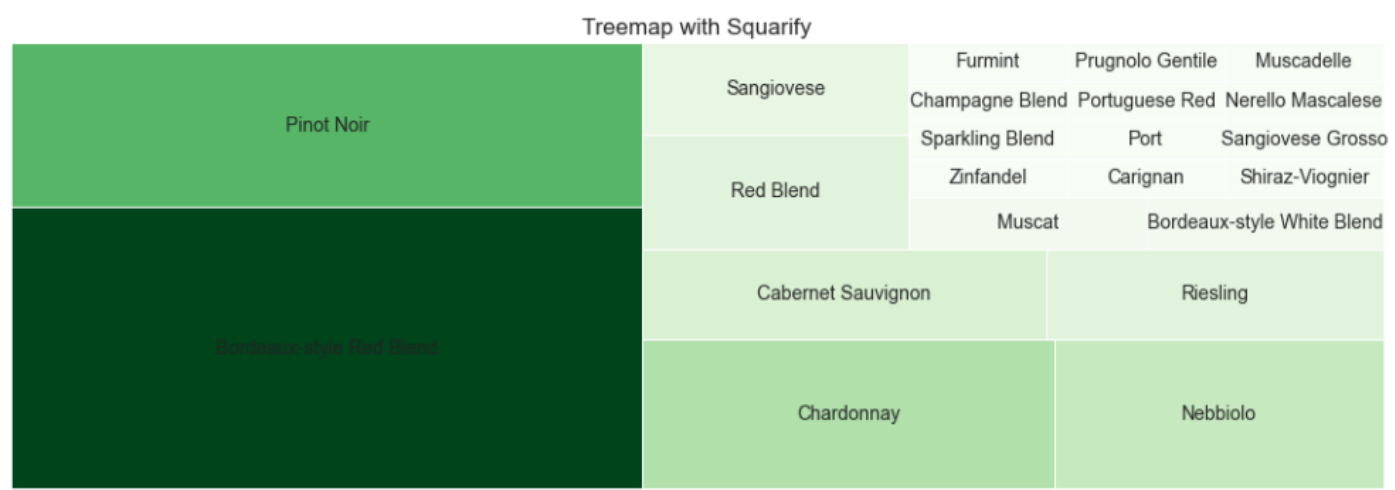


# Conclusion

✓ Acquisition by each Country according Points given



✓ Best Wine Variety



## ✓ Best 100 Wines Rated Worldwide

- |   |                    |                      |                |
|---|--------------------|----------------------|----------------|
| 1. Title: Avignonesi 1995 Occhio di Pernice (Vin Santo di Montepulciano)            | Country: Italy     | Province: Tuscany    | Price: \$210.0 |
| 2. Title: Chambers Rosewood Vineyards NV Rare Muscat (Rutherglen)                   | Country: Australia | Province: Victoria   | Price: \$350.0 |
| 3. Title: Quilceda Creek 2008 Cabernet Sauvignon (Columbia Valley (WA))             | Country: US        | Province: Washington | Price: \$125.0 |
| 4. Title: Williams Selyem 2009 Precious Mountain Vineyard Pinot Noir (Sonoma Coast) | Country: US        | Province: California | Price: \$94.0  |
| 5. Title: Giuseppe Rinaldi 2013 Brunate (Barolo)                                    | Country: Italy     | Province: Piedmont   | Price: \$300.0 |
| 6. Title: Chambers Rosewood Vineyards NV Rare Muscadelle (Rutherglen)               | Country: Australia | Province: Victoria   | Price: \$350.0 |

## ✓ Best Wines Rated in India

- |   |                |                  |               |
|---|----------------|------------------|---------------|
| 1. Title: Sula 2011 Dindori Reserve Shiraz (Nashik) | Country: India | Province: Nashik | Price: \$20.0 |
| 2. Title: Sula 2015 Sauvignon Blanc (Nashik)        | Country: India | Province: Nashik | Price: \$12.0 |