WINE DATA ANALYSIS
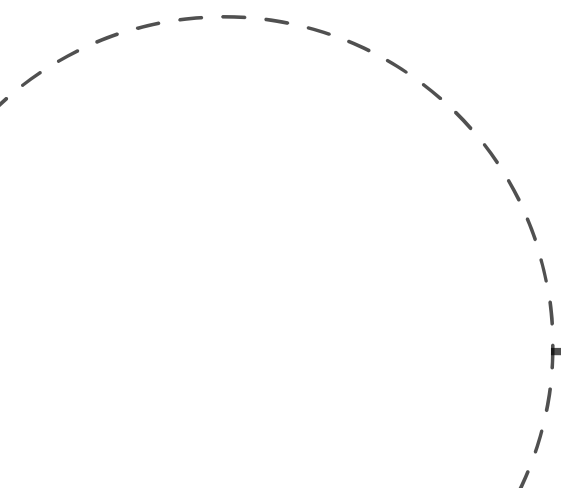
By Akash Singh

About The Wine Industry

The industry of wine is a very marketing driven business. Wine is now produced all over the world and the products are shipped globally. It is a natural product and the outcome depends heavily on the often on purpose indivualized production process. So as a result the products even from the same producer with the same name change every year. On the one hand the costumer is gaining more choice but on the other hand he is loosing the clear view and it is making the decision troublesome. So even sticking to one producers product can not guarantee you a standard if this season the grape has low quality. That is why in this opaque jungle reviews are so important. Nearly all buyers more or less rely on reviews because these are the only information they can get beside the metadata of the wine like name, origin etc. So in the past decades wine review itself became an industry. It produced so much data which can be mined for deductions and predictions which will be done here to make better decision when buying wine.

Overview

Objective

Discuss different types of wines which are best according to the testers and wines which are expensive and the the wines which are above average and below average and the teter who has done most reviews .

Benefits

- Identification of rating .
- Gives better insight of customers base .
- Help us to identify ratings distribution around the world .

Data Sharing Agreement

Sample File Name

Wine review_small.csv

Number Of Columns

13(exclusion:index col)

- **country** **object**
- **description** **object**
- **designation** **object**
- **points** **int64**
- **price** **float64**
- **province** **object**
- **region_1** **object**
- **region_2** **object**
- **taster_name** **object**
- **taster_twitter_handle** **object**
- **title** **object**
- **variety** **object**
- **winery** **object**

Architecture

Problem Statement

Data Mining

Data Preparation

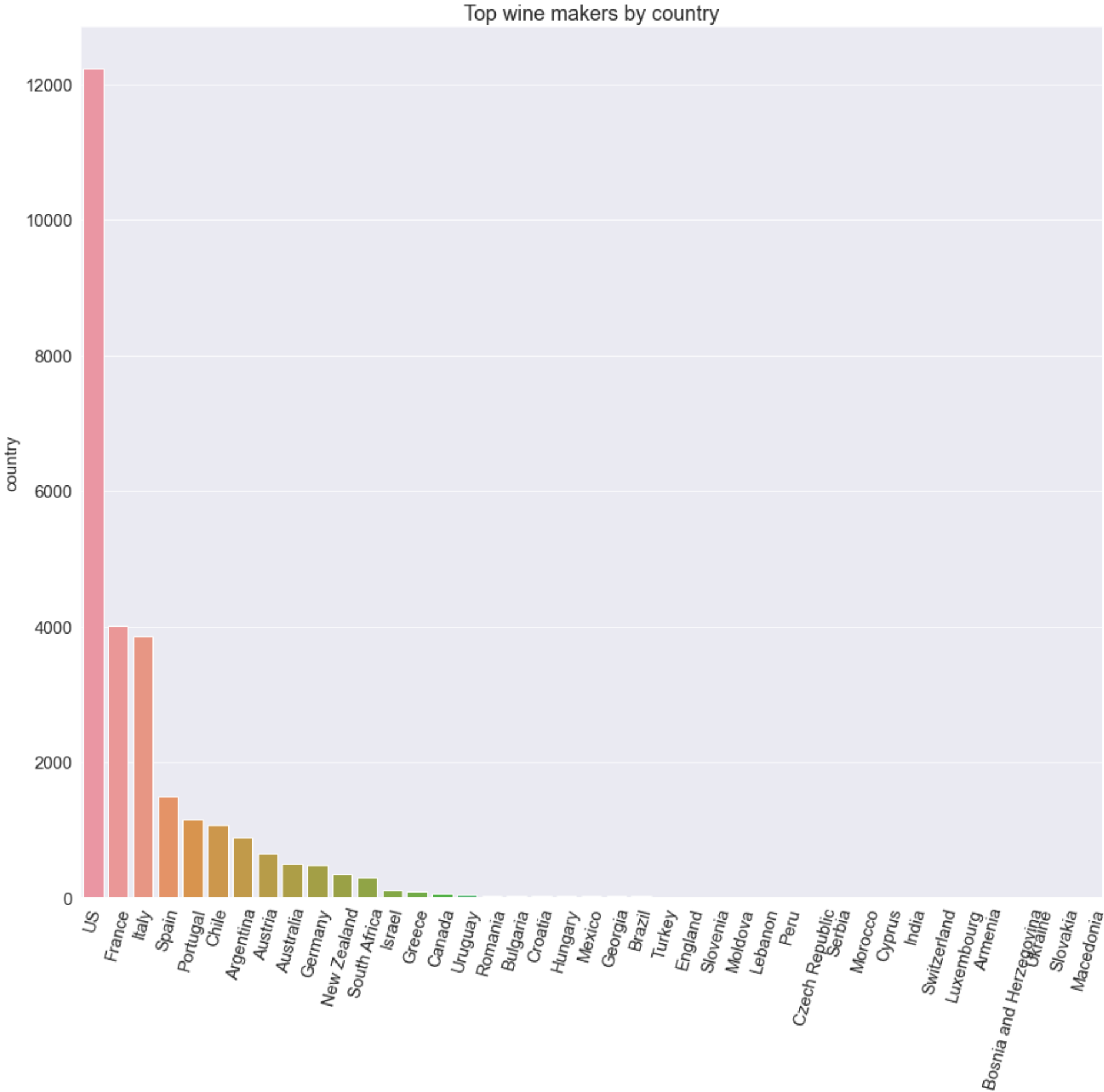**Exploratory Data Analysis**

Inferences

Data
Preparation

Removing The NAN Values From The Data Set

Removing The Duplicate Values From The Data Set

Removing The Unused Coloumns From the Data Set

Number Of Wine Makes By Countries

Top wine makers by country

USA is top of the list among the wine making countries at over 12000 units, followed by ITALY & FRANCE slighty bellow 4000 by volumn .
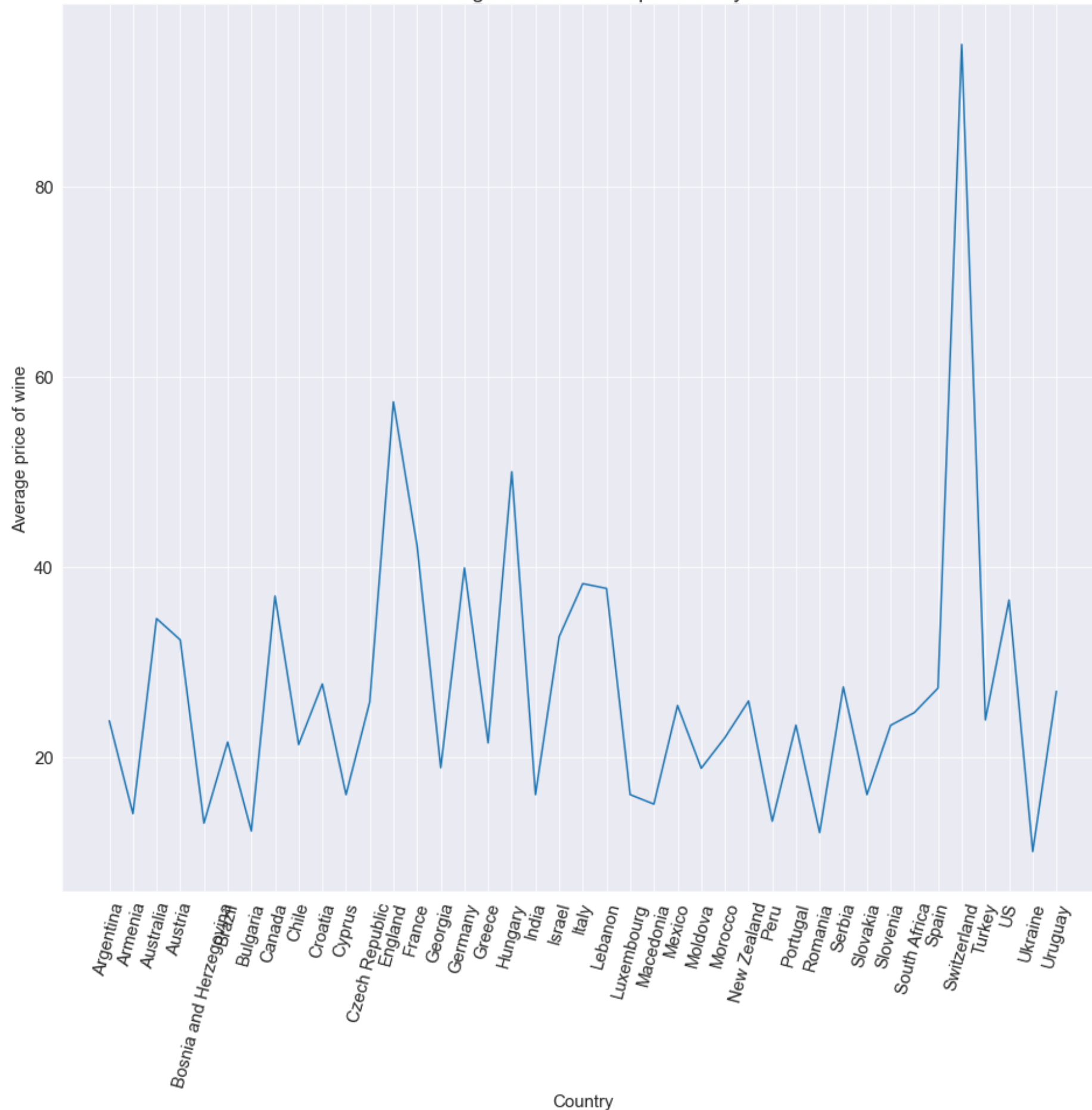
Wine Price Distribution

Price distribution of wines
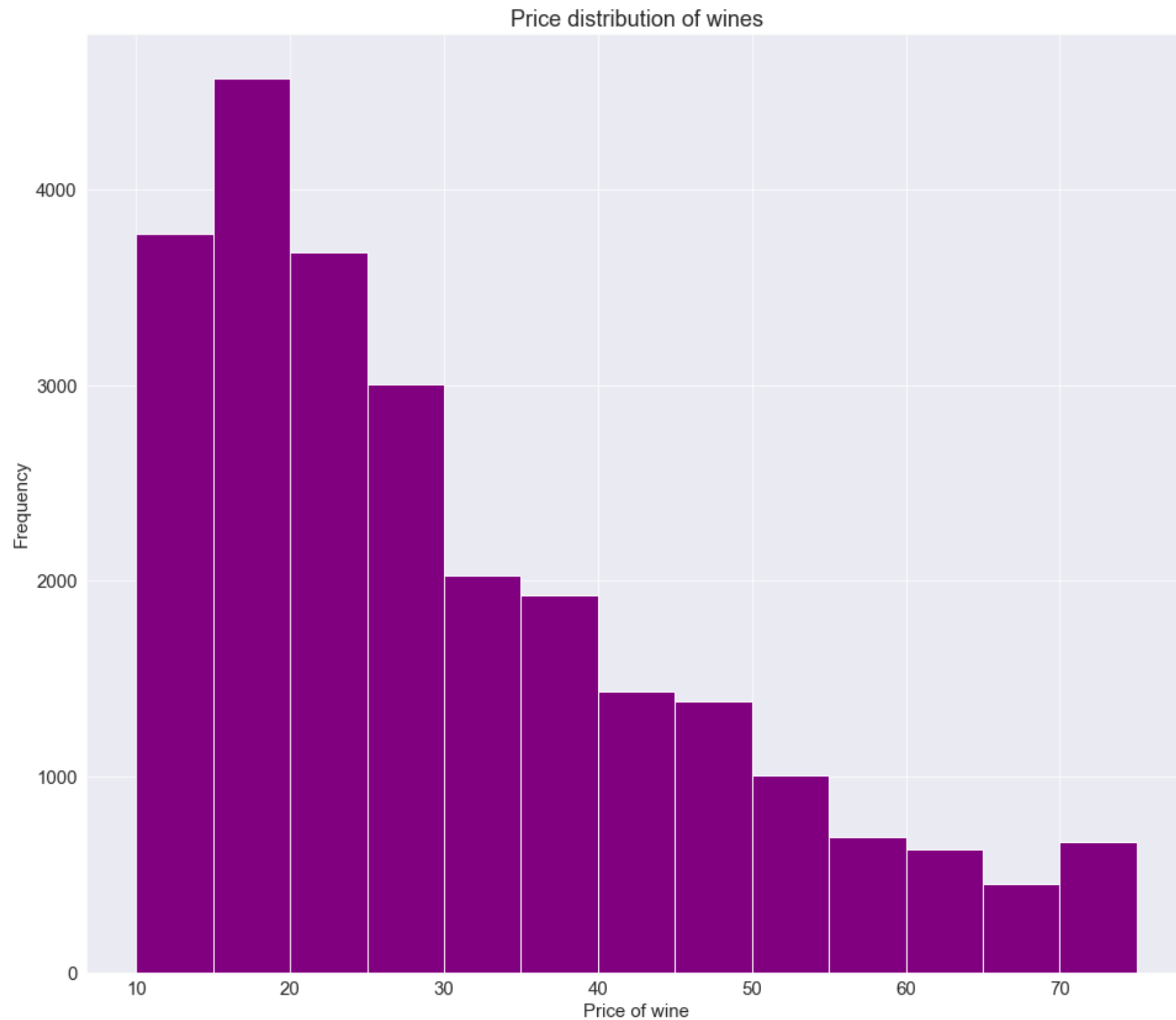
Most of the wines are on lowerside of the price range.

Average Price By Country

Average Price of wines per country.

Stzerland has The highest wine price which is just below 100 per unit while Armenia ,Bosnia and Herzegovina , Bulgaria , Peru , Romania have the lower price as compare to others which is just above 10 per unit and Ukrain has the lowest wine price below 10 per unit . Ther is no other countires which have unit price more than 60 except US. Hungary & England has uint price just below 60.Maximum Number Of coutries have the have there unit price between 20 to 40 .

Price Distribution of top 10 countries

	country	count	min	max	mean
38	US	12234	4	625	36.49109
13	France	4014	5	2500	42.205531
20	Italy	3863	5	540	38.217965
35	Spain	1486	4	770	27.219381
29	Portugal	1156	5	300	23.315744
8	Chile	1071	5	200	21.270775
0	Argentina	893	4	215	23.784994
3	Austria	654	7	150	32.295107
2	Australia	494	7	350	34.538462
15	Germany	482	9	775	39.86722

As you can see while all the wines unit despite their origins start around 5 ,only French wines have an actual unit of more than 2,000

Correlation Between Price & Point
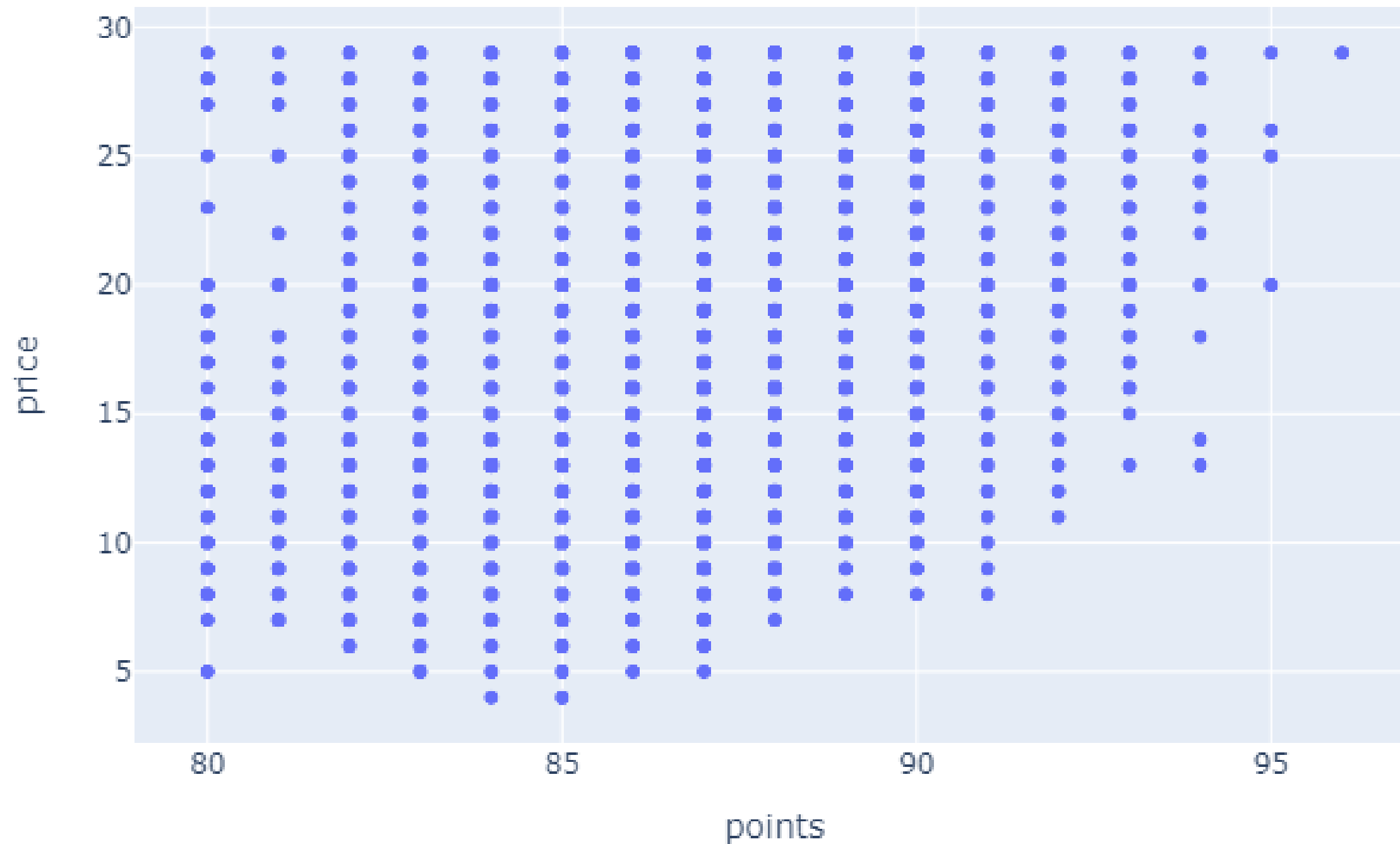
Consider paying max 30/unit for a bottle which is quite acceptable, so here we can see the points begin at 80.

You can find a bottle at any price but to have a bottl\e at min 94 pts, you have to pay at least 15$

Which countries have the top 5 highly rated wine?

	country	description	points	price	province	title	variety	winery
7335	Italy	Thick as molasses and dark as caramelized brow...	100	210	Tuscany	Avignonesi 1995 Occhio di Pernice (Vin Santo ...	Prugnolo Gentile	Avignonesi
345	Australia	This wine contains some material over 100 year...	100	350	Victoria	Chambers Rosewood Vineyards NV Rare Muscat (Ru...	Muscat	Chambers Rosewood Vineyards
1557	US	A stupendous Pinot Noir, showing how beautiful...	99	94	California	Williams Selyem 2009 Precious Mountain Vineyar...	Pinot Noir	Williams Selyem
1556	US	The flagship wine from Quilceda Creek offers e...	99	125	Washington	Quilceda Creek 2008 Cabernet Sauvignon (Columb...	Cabernet Sauvignon	Quilceda Creek
1561	Italy	Here's a "wow" wine on every level, with incre...	98	102	Tuscany	Marchesi Antinori 2008 Guado al Tasso (Bolghe...	Red Blend	Marchesi Antinori

Which wine has the highest unit price per wine rating?

	country	description	points	price	province	title	variety	winery	price_value_per_rating
20148	US	Fruity, soft and rather sweet, this wine smell...	85	4	California	Dancing Coyote 2015 White (Clarksburg)	White Blend	Dancing Coyote	21.25
1987	Spain	Berry and cherry aromas are surprisingly sturd...	85	4	Central Spain	Felix Solis 2013 Flirty Bird Syrah (Vino de la...	Syrah	Felix Solis	21.25
29217	Argentina	Crimson in color but also translucent, with a ...	84	4	Mendoza Province	Broke Ass 2009 Red Malbec-Syrah (Mendoza)	Malbec-Syrah	Broke Ass	21
24256	Chile	More than decent; this is good SB with no bell...	87	5	Aconcagua Valley	In Situ 2008 Reserva Sauvignon Blanc (Aconcagu...	Sauvignon Blanc	In Situ	17.4
3167	Italy	Packaged in a pocket-friendly 187-ml bottle, t...	86	5	Veneto	Anna Spinato NV Mini (Prosecco)	Glera	Anna Spinato	17.2

Which 5 wine varieties are rated the least?

	country	description	points	price	province	title	variety	winery	price_value_per_rating
18786	France	Bold and very clumsy. There are leathery eleme...	80	8	Languedoc-Roussillon	Château de Pennautier 2001 Cabernet Sauvignon ...	Cabernet Sauvignon	Château de Pennautier	10
4557	Spain	With dusty, candied aromas, the bouquet on thi...	80	14	Catalonia	Paul Cheneau NV Demi Sec Sparkling (Cava)	Sparkling Blend	Paul Cheneau	5.714286
4556	Italy	There's a thorny, almost raw quality of fruit ...	80	12	Central Italy	Fattoria Giuseppe Savini 2011 Rondineto Merlot...	Merlot	Fattoria Giuseppe Savini	6.666667
5905	Argentina	Ashy, green and funky smelling, then sweet, ch...	80	17	Mendoza Province	Bodega Carmine Granata 2009 Sémillon (Mendoza)	Sémillon	Bodega Carmine Granata	4.705882
5906	Argentina	A rough, pinchy nose with leather and funk is ...	80	20	Mendoza Province	Cruz Alta 2007 Grand Reserve Malbec (Mendoza)	Malbec	Cruz Alta	4

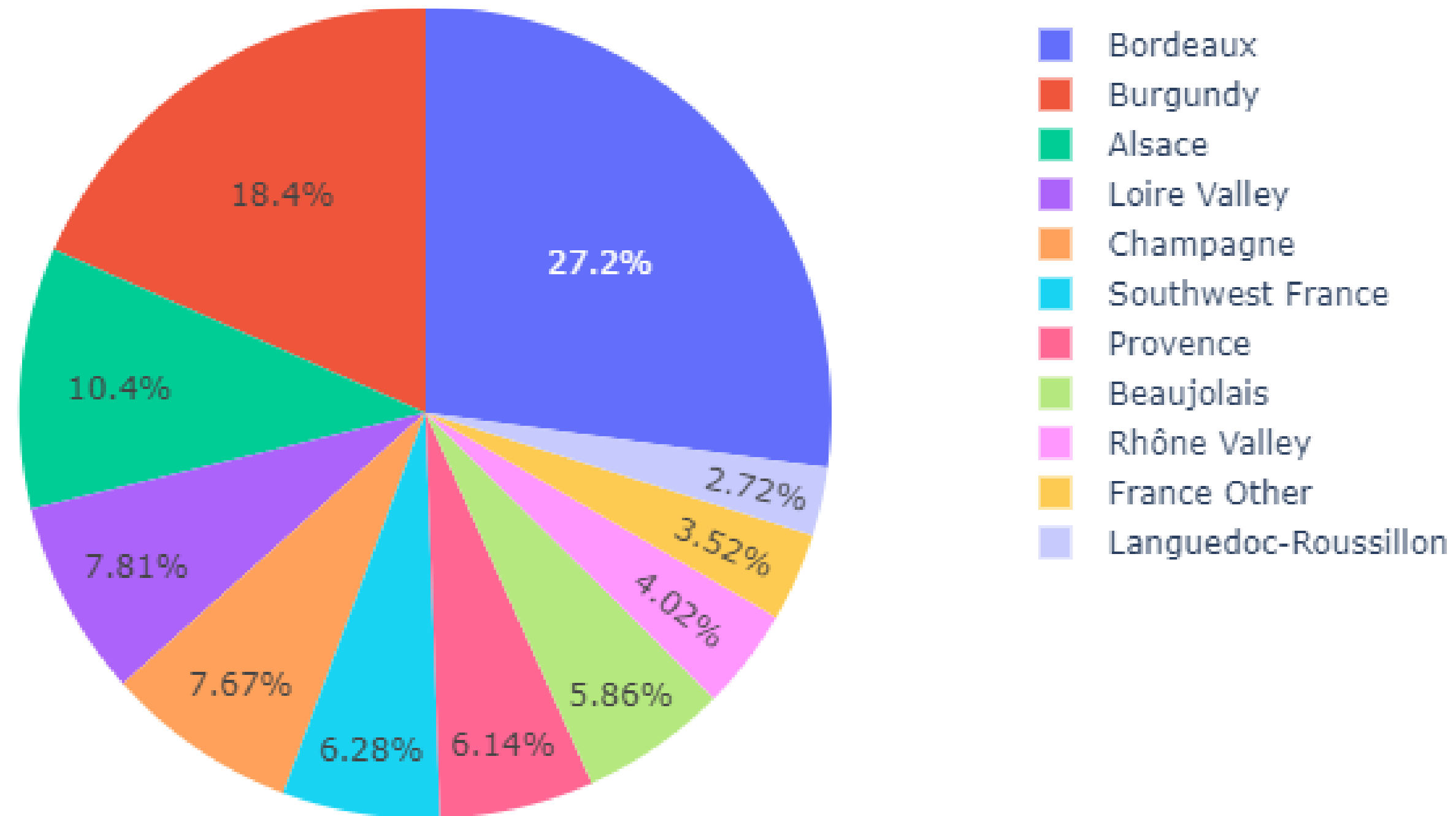# How many varities of wine are there in production all over the world ?

There are 465 different varieties of wines in the world.

Which wines are rated the highest in the world?

	country	description	points	price	province	title	variety	winery	price_value_per_rating
7335	Italy	Thick as molasses and dark as caramelized brow...	100	210	Tuscany	Avignonesi 1995 Occhio di Pernice (Vin Santo ...	Prugnolo Gentile	Avignonesi	0.47619
345	Australia	This wine contains some material over 100 year...	100	350	Victoria	Chambers Rosewood Vineyards NV Rare Muscat (Ru...	Muscat	Chambers Rosewood Vineyards	0.285714
1557	US	A stupendous Pinot Noir, showing how beautiful...	99	94	California	Williams Selyem 2009 Precious Mountain Vineyar...	Pinot Noir	Williams Selyem	1.053191
1556	US	The flagship wine from Quilceda Creek offers e...	99	125	Washington	Quilceda Creek 2008 Cabernet Sauvignon (Columb...	Cabernet Sauvignon	Quilceda Creek	0.792

Zooming in to French province

Legend:
- Bordeaux — 27.2%
- Burgundy — 18.4%
- Alsace — 10.4%
- Loire Valley — 7.81%
- Champagne — 7.67%
- Southwest France — 6.28%
- Provence — 6.14%
- Beaujolais — 5.86%
- Rhône Valley — 4.02%
- France Other — 3.52%
- Languedoc-Roussillon — 2.72%

Although Bordeaux is the most popular province for wine, France has quite a list of good wines through the country.

Zooming in to US province
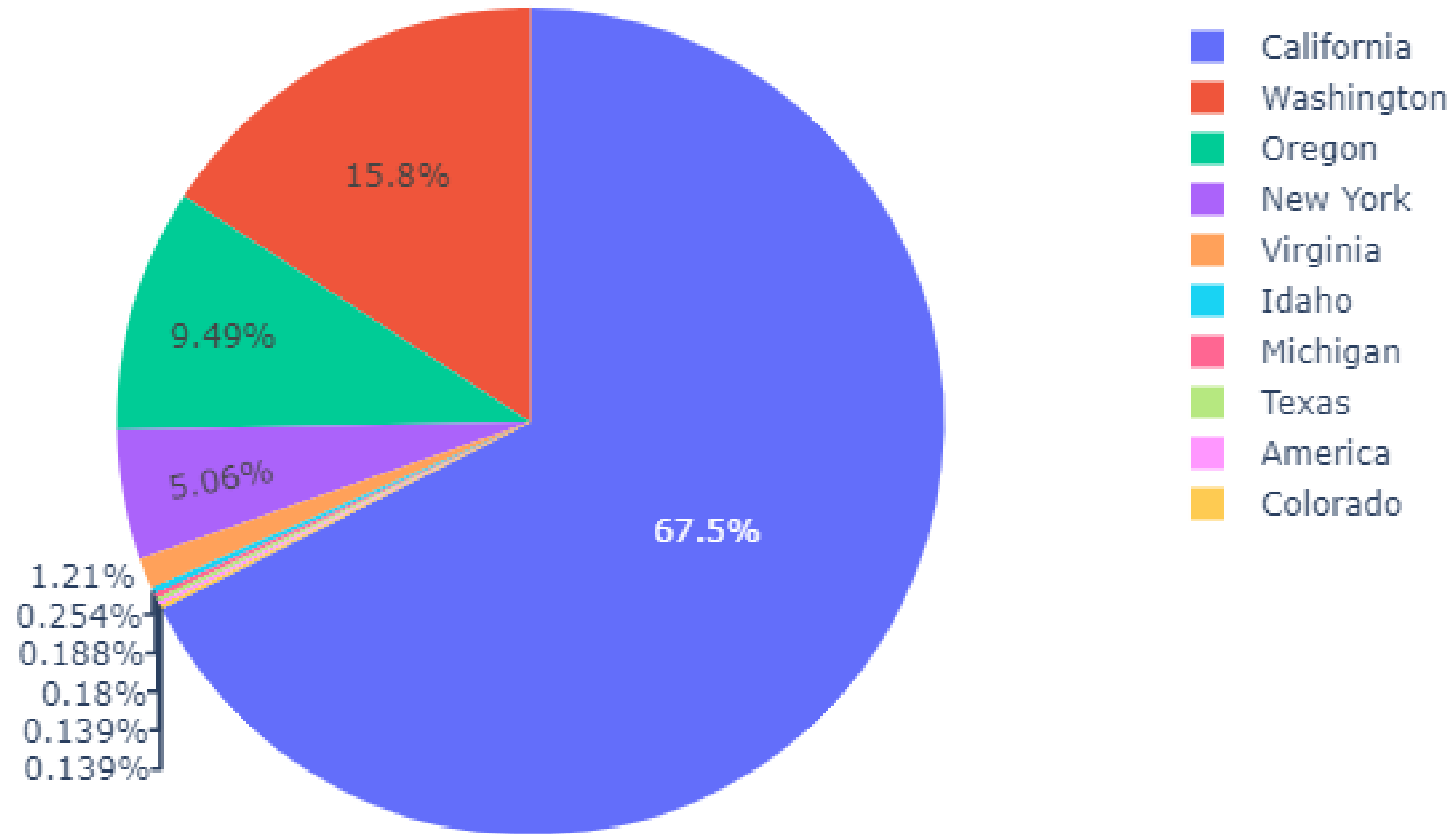
California is the place-to-be if you want to produce wine in US with over 67.5%, following with Washington and Oregon who cover already over 90% of the market.

Which Variety of wine is mentioned here the most ?

	variety	count
291	Pinot Noir	2873
82	Chardonnay	2440
49	Cabernet Sauvignon	2141
314	Red Blend	1902
38	Bordeaux-style Red Blend	1215

Thank You