# Data-driven purchasing strategy for a wine marketplace

*Federico Trotta*

# Description of the project

**THE PROJECT:**

Starting from a dataset of wine reviews, I created a notebook with the analyzes for a wine marketplace for small local producers.

In the dataset, we find the type of wine associated with a description, a price in $ (presumed per bottle), the state of origin of the wine, a review with a score from 0 to 100, and the name and reference of those who reviewed it.

**THE IDEA:**

I hypothesized the opening of a possible wine marketplace, researching for wines in a certain price range, and that they had obtained at least a certain score in the rating. In this way, I calculated the number of bottles needed and the total price (to be spent on wine alone) to launch the business.

# Correlation analysis between price and score
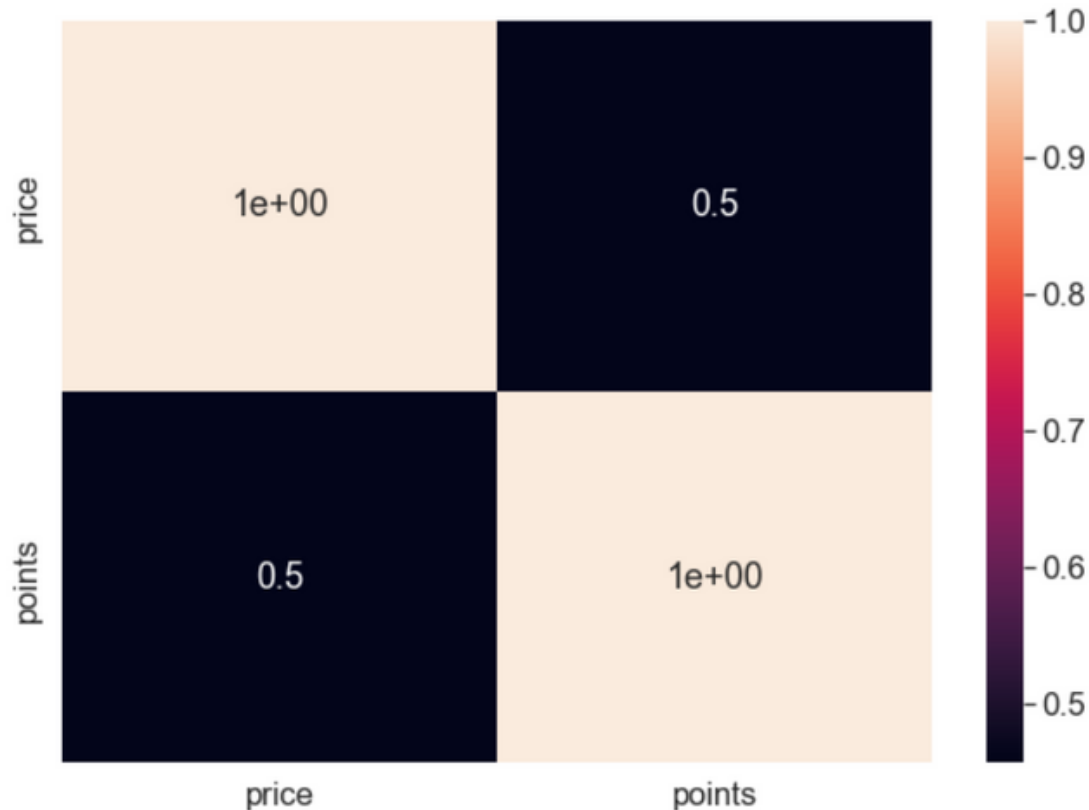
In general, we tend to think that the higher the price, the higher the quality of a product.

In this case, it is natural to ask if a certain price corresponds to a certain score (high price = high score?).

I calculated the correlation matrix which, however, tells us that there is no correlation between the price and the score received.

Wine price for Country
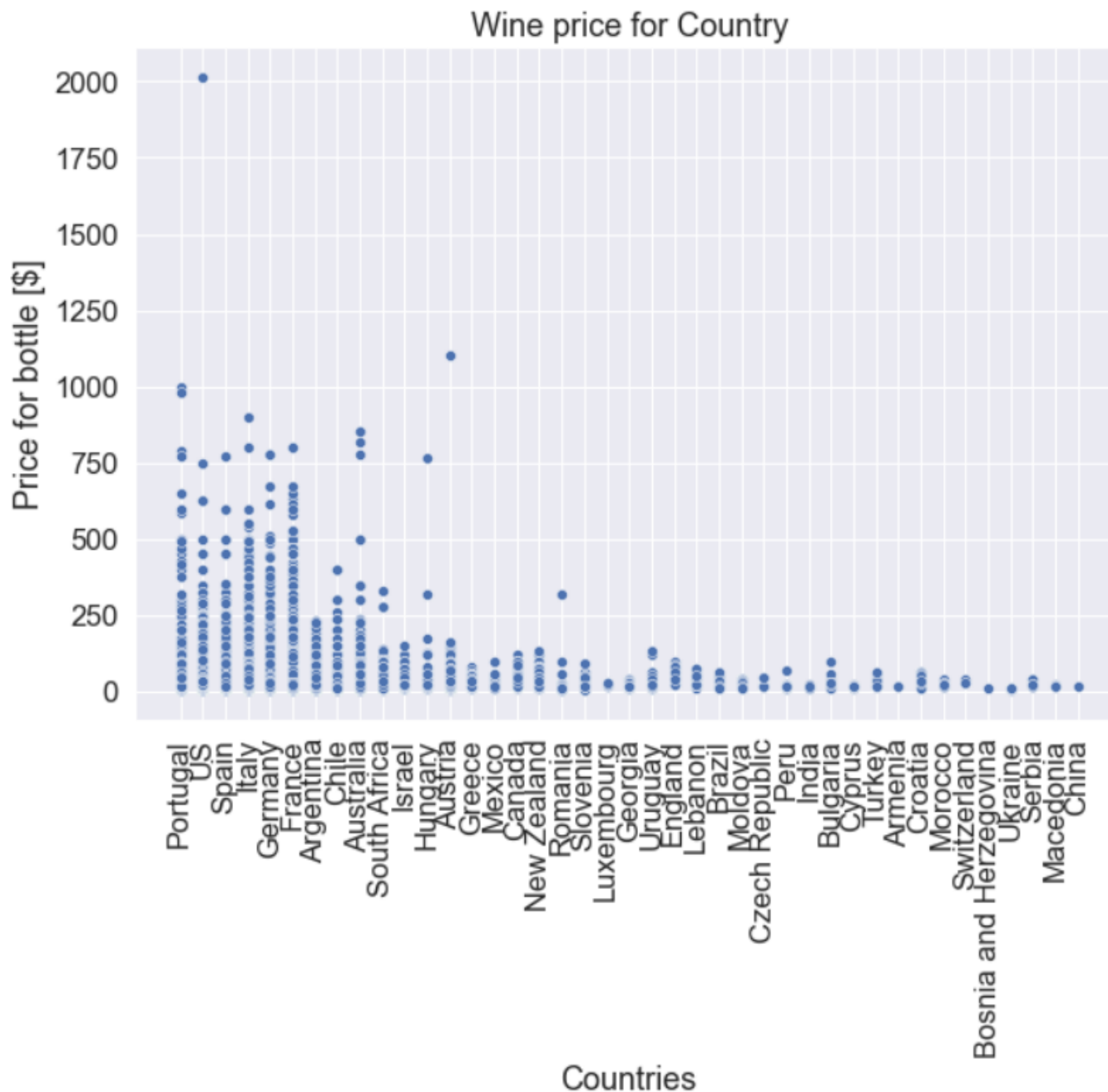
## Price analysis by state

I would have expected to see the highest prices in Italy and France; while, on the other hand, as you can see, high prices are quite common in Europe but, to my great surprise, the maximum price is in the USA.

Wine points for Country
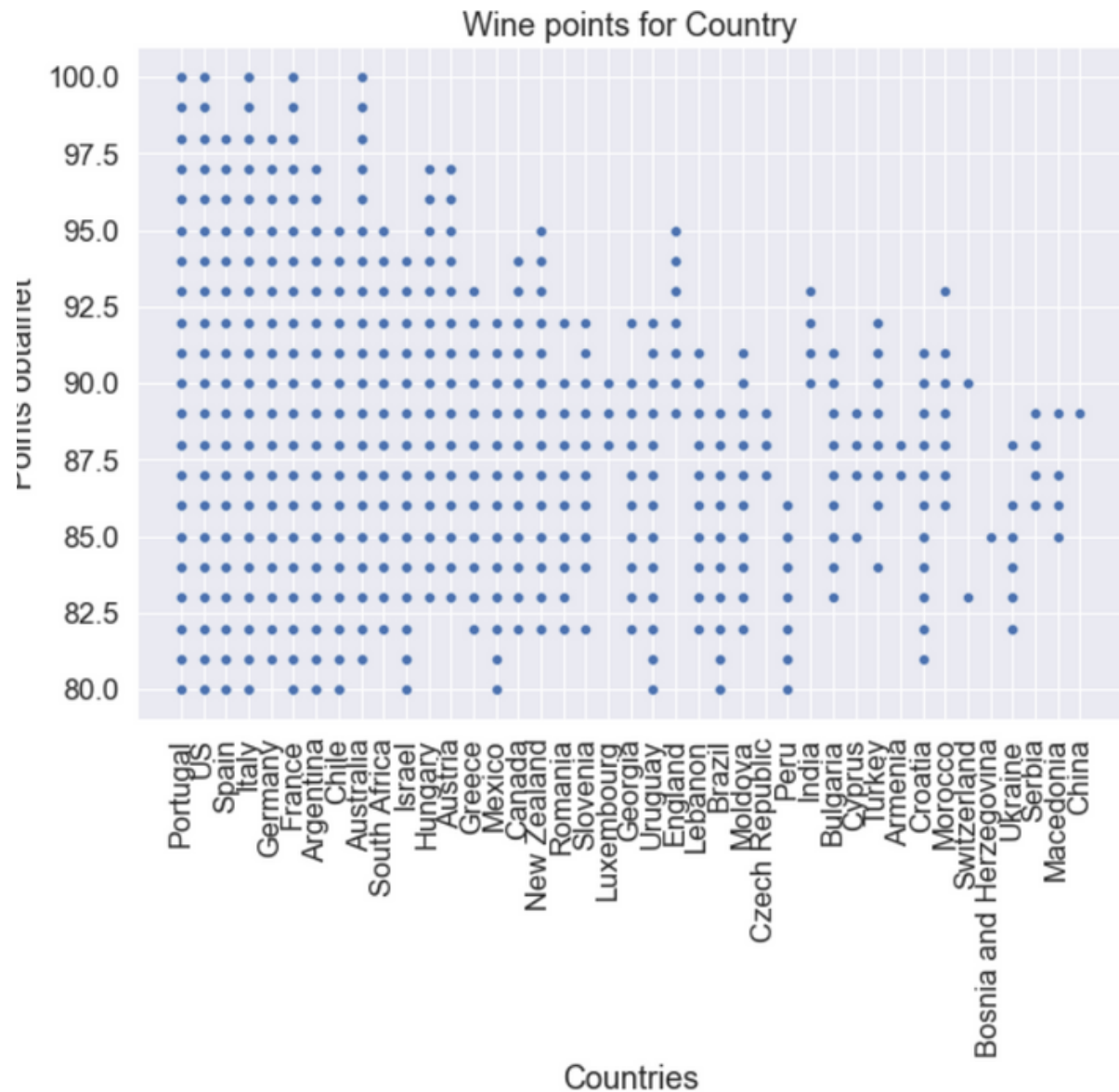
# Score analysis by state

There is no particular trend in the scores, even if it is seen that several European countries have the whole score range 80-100

Wine price for Country

## Strategy setting

I hypothesized a strategy for a small-medium marketplace that wants to target customers who want good wines, in a medium-high price range. For this, I went to look for wines with a minimum price of $50 and a maximum of $500 that at the same time had a rating greater than 95 points.

There are 679 wines with these characteristics and they are mainly found in Europe. Assuming a physical marketplace in the Milan area, this also reduces transport costs.

# Calculation of the investment for wine only

679 different types of wine are however many; to reduce the number (and, therefore, the physical space in the marketplace) and to reduce transport costs, I chose to select only wines in Italy, France, Spain, and Portugal.

Assuming buying 2 bottles of each type, I calculated the total number of bottles to be purchased and the initial investment (for the bottles only).

```python
#total numbers of bottles
bottles = 2*wine.shape[0]

#initial investment (just bottles)
initial_investment = investment*2

print(f'the number of bottles to buy is: {bottles}')
print(f'the initial investment for only the wine bottles is of: {initial_investment}[$]')
```

```
the number of bottles to buy is: 426
the initial investment for only the wine bottles is of: 71926.0[$]
```