Daniels Distributors:

Enhancing our Wine Portfolio through Data-Driven Insights

Hello!

I am Ryne Daniels, owner of Daniels Distributors.

Business Challenge:
- Charleston wine market is highly competitive
- Need to differentiate itself
- Opportunity to expand by offering international wines
- Selecting right wines and regions for expansion is expensive

Objective:

Leverage Data Analysis and Machine Learning to Expand Wine Portfolio of a Small Distributor.

Key Approaches:

- Utilize data analysis to identify trends and patterns to guide wine selection process
- Minimize risks associated with expanding our wine offerings
- Develop a ML model to predict wine ratings
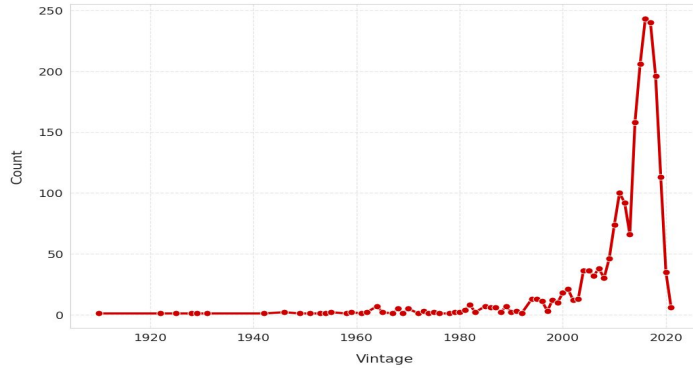- Create a lightweight marketing strategy based off the results

The Data

Our Dataset consists of 7,500 Spanish Wines, with valuable information on the following:

- Winery
- Wine
- Year
- Rating
- Number of Reviews
- Country

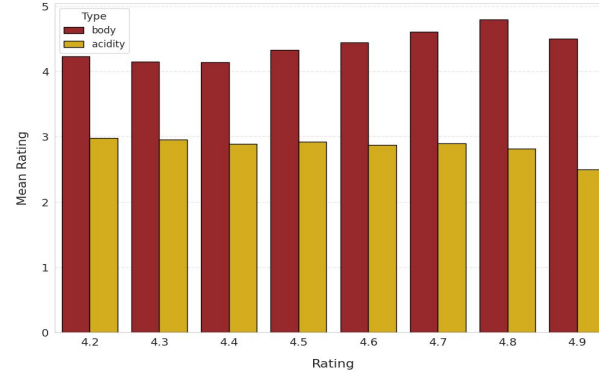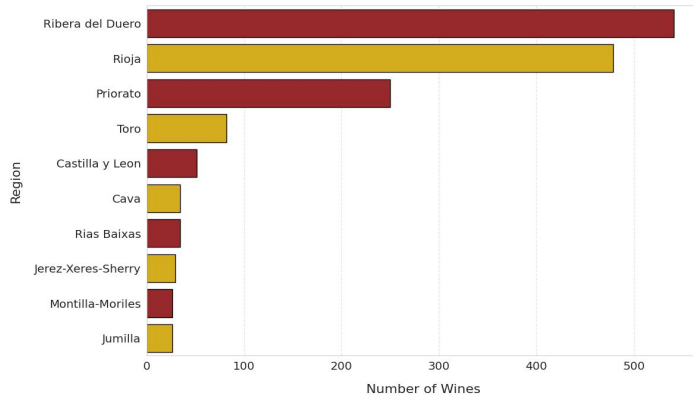- Region
- Price
- Type
- Body
- Acidity

Link to data: https://www.kaggle.com/datasets/fedesoriano/spanish-wine-quality-dataset
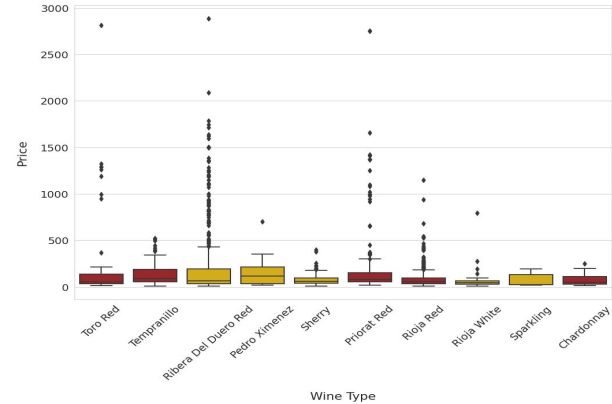
Targeting High-Potential Wine Regions

(and avoiding mistakes!!)

Top 10 Wine Regions by Rating:

- Montilla-Moriles: 4.54
- Tierra del Vino de Zamora: 4.50
- Dehesa del Carrizal: 4.50
- Condado de Huelva: 4.50
- Cigales: 4.50
- Jumilla: 4.50
- Dominio de Valdepusa: 4.47
- Manzanilla: 4.47
- Ribera del Duero: 4.45
- Cadiz: 4.44

Top 10 Wine Regions by Price:

- Bierzo: $265.34
- Ribera del Duero: $210.59
- Toro: $204.30
- Priorato: $203.55
- Montilla-Moriles: $178.69
- Conca de Barbera: $164.54
- Jerez Palo Cortado: $144.48
- Mentrida: $121.99
- Mallorca: $110.56
- Aragon: $108.51

There are only **2** common values in each list!!

What's next?

¿Qué sigue?

¡Machine Learning to the Rescue!

Why Machine Learning?

- Improve decision making process
- Reducing costs
- Customer Satisfaction
- Replicable Process

Deliverables:

- Stay ahead of market trends
- Wider Portfolio
- Targeted Marketing Efforts
- Better Supplier Relationships

ML Model Development Process

Data Preparation

Original data set was 7500 entries, after "cleaning" was reduced to ~1700 entries.

Model Selection

Use different type of model to predict "ratings" category in our dataset. We used a classification model.

Evaluation

Use metrics to select the best model to deploy and refine the model as we gather new data.

The Metrics:


```
Gradient Boosting Classifier
              precision    recall  f1-score   support

        high       0.50      0.30      0.37        20
         low       0.84      0.92      0.87       293
         med       0.60      0.48      0.53       106

    accuracy                           0.78       419
   macro avg       0.65      0.57      0.59       419
weighted avg       0.76      0.78      0.76       419
```
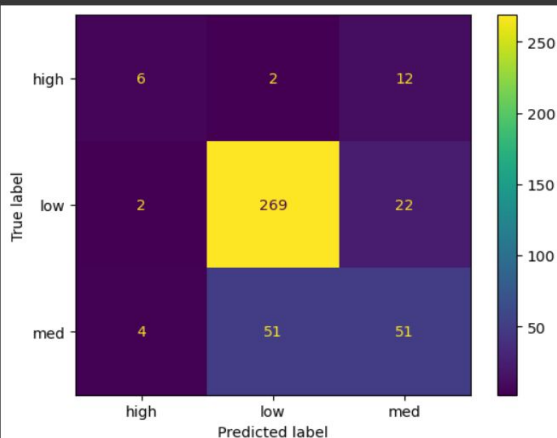
	Precision	Recall	F1 Score	Accuracy
Model Name				
Gradient Boosting Classifier	0.645135	0.566407	0.594609	0.778043

High class: False Pos = 14, False Neg = 6
Low class: False Pos = 24, False Neg = 54
Med class: False Pos = 56, False Neg = 34

Improvements can be made on this model!!

All of this Equals!!

Stronger Relationships

Data insights can identify high quality wineries and producers in key regions.

Joint Marketing

Combine our knowledge of customer targeting and acquisition with their knowledge of wine production.

Mutual Benefits

Collaboration can lead to increased sales for both companies.

Long-Term Relationships

Consistent supply of high quality wines.

Share Success

Promote each other's brand and story!

Cost Saving

Bulk purchasing , better distributions deal, which trickles down to the customer.

Gracias

Any questions?

You can find me at:

- ◦ @rynedaniels (GitHub)
- ◦ rynedaniels@gmail.com

Credits

Special thanks to all the people who made and released these awesome resources for free:

- ◦ Presentation template by SlidesCarnival
- ◦ Photographs by Unsplash