



Types of wine

What factors help predict the type of a wine?

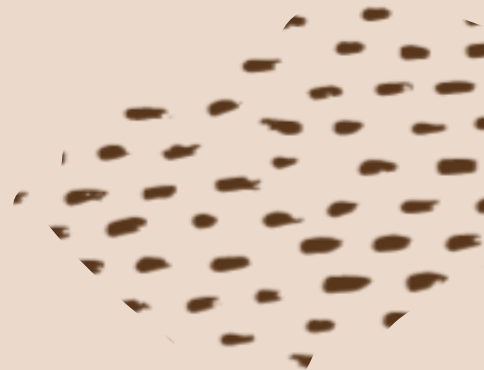
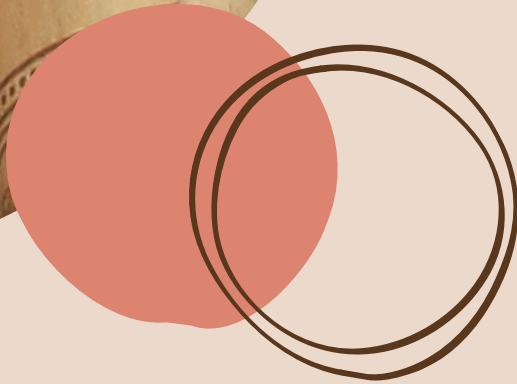


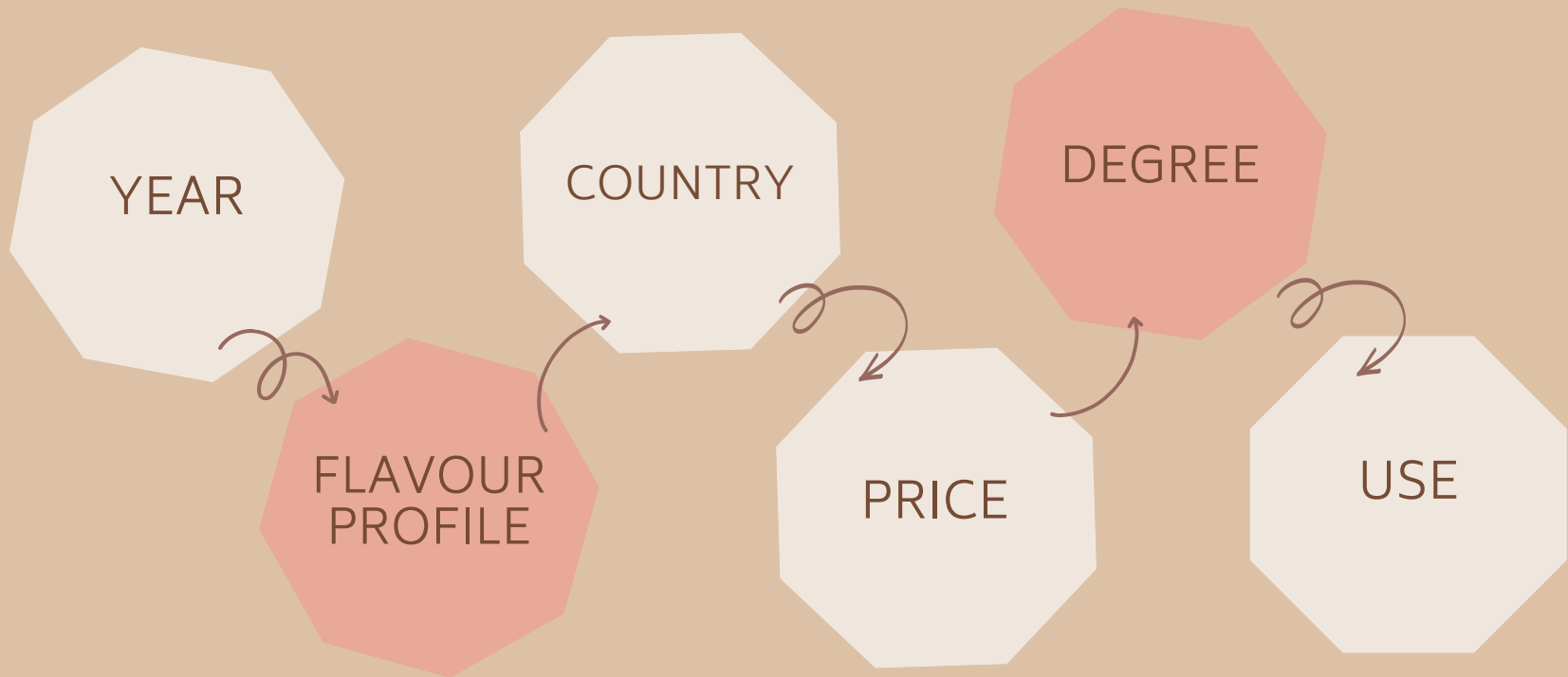




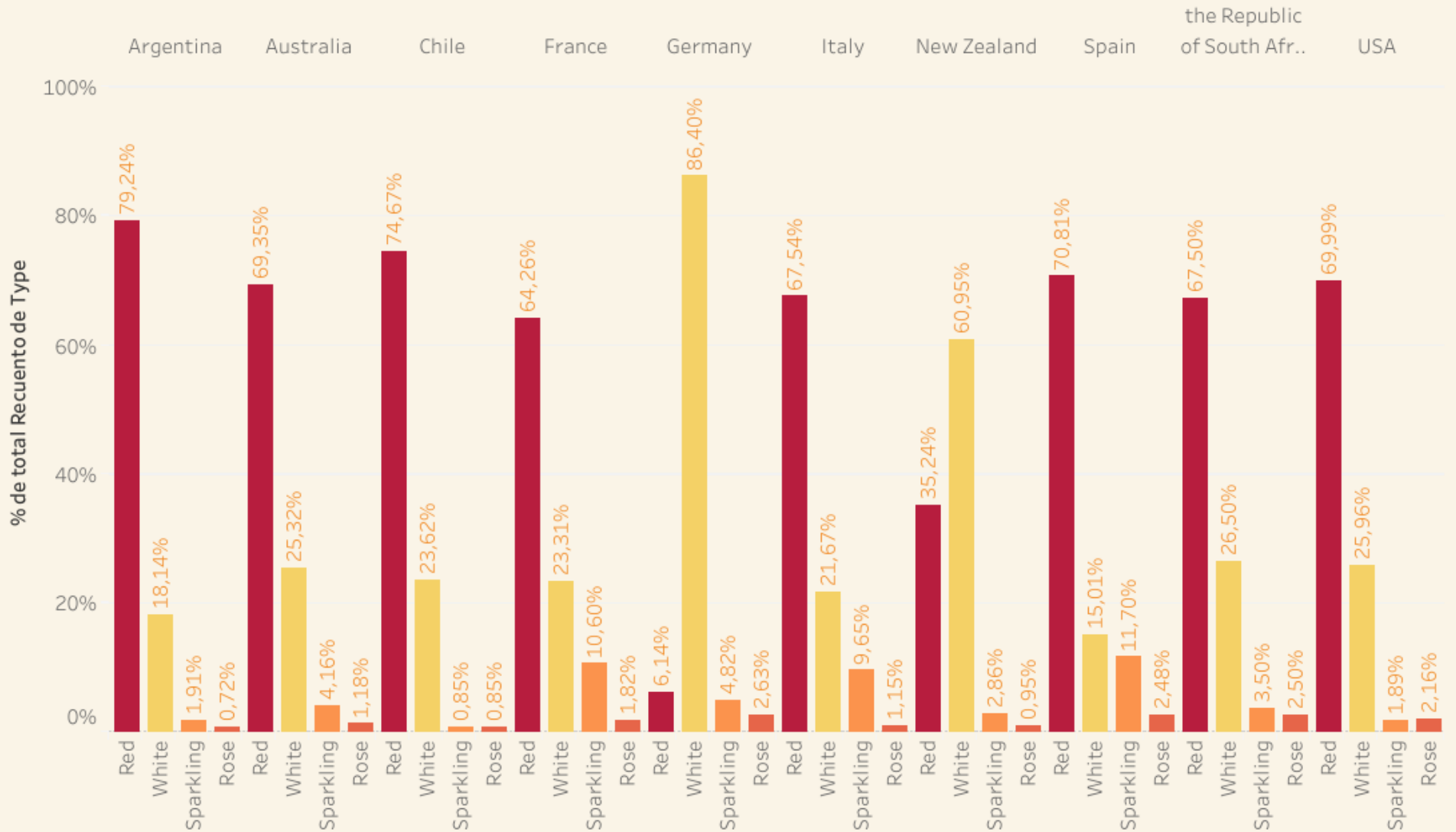
Table of Contents

- 
- Data exploration
 - Hypothesis testing
 - Predictive model
 - Next steps
- 

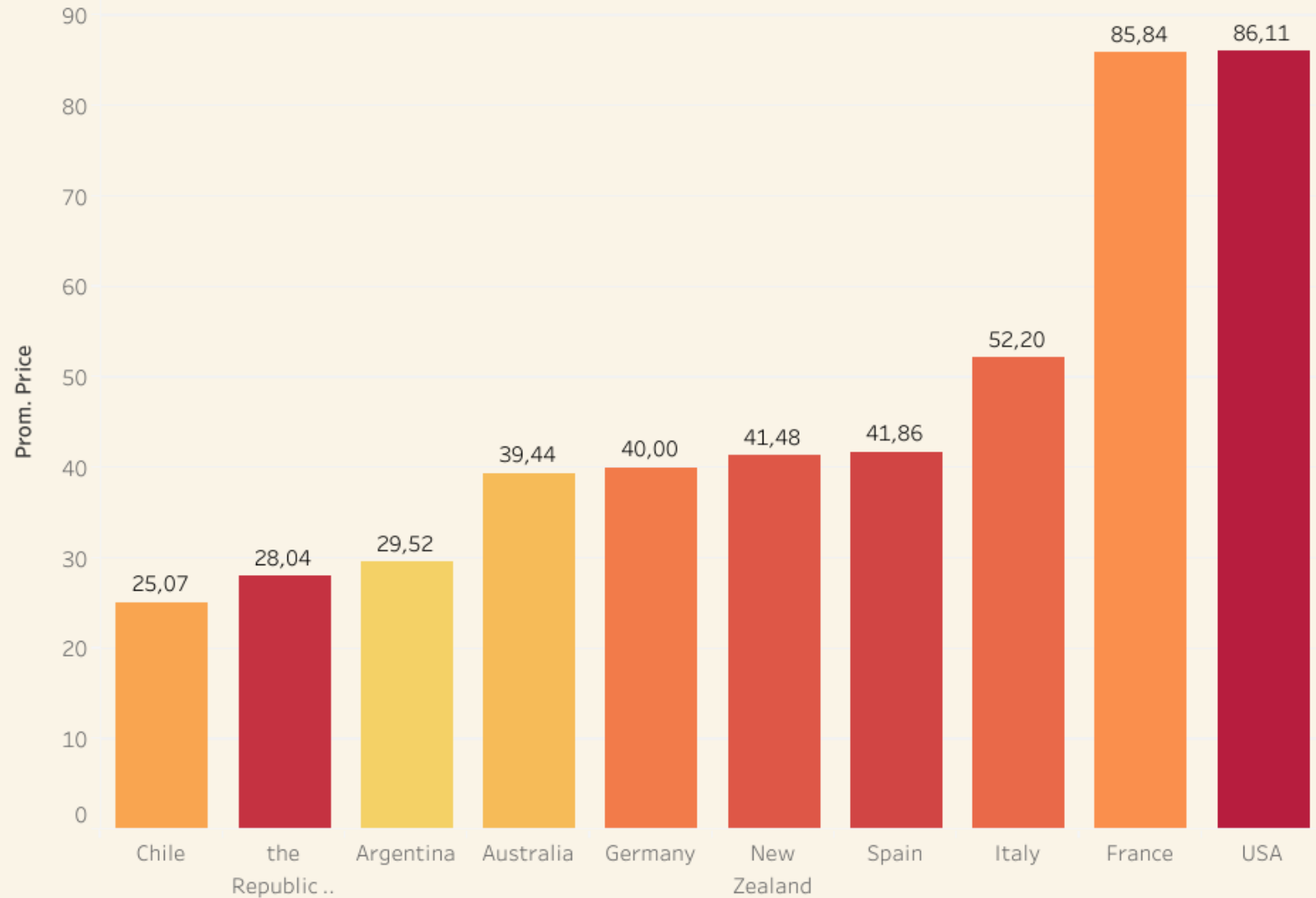
Type of Wine



Wine Type by Country

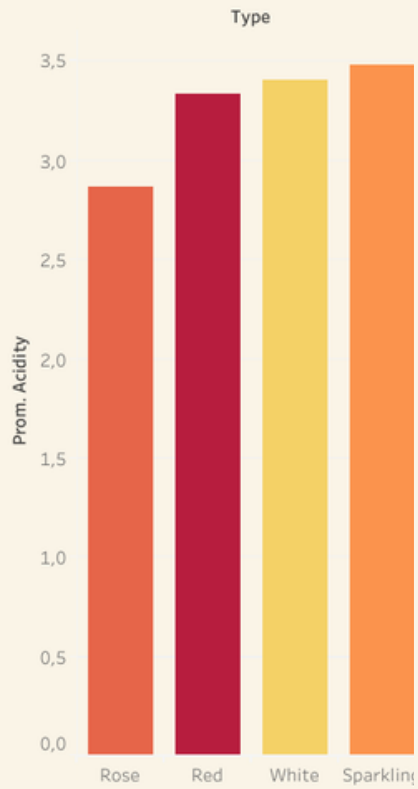


Average Price by Country

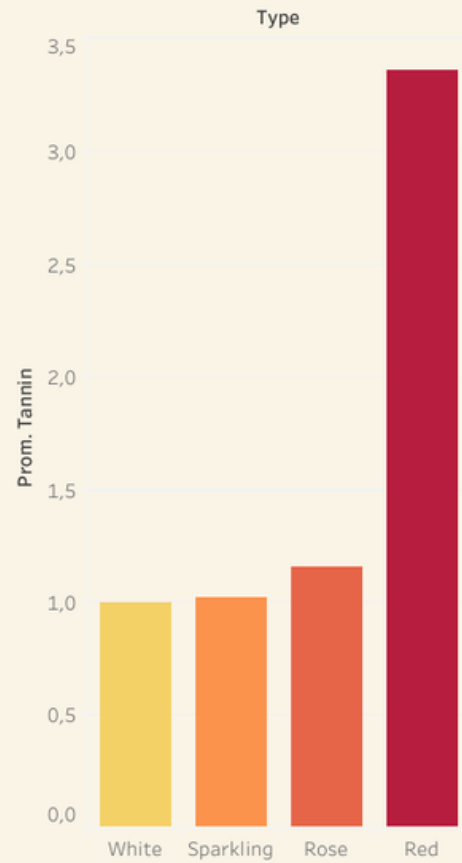


Flavour Profile by Type

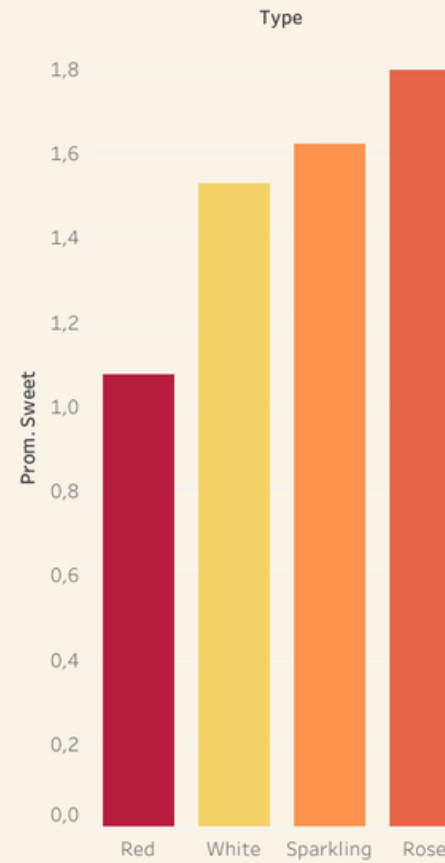
Acidity by Type



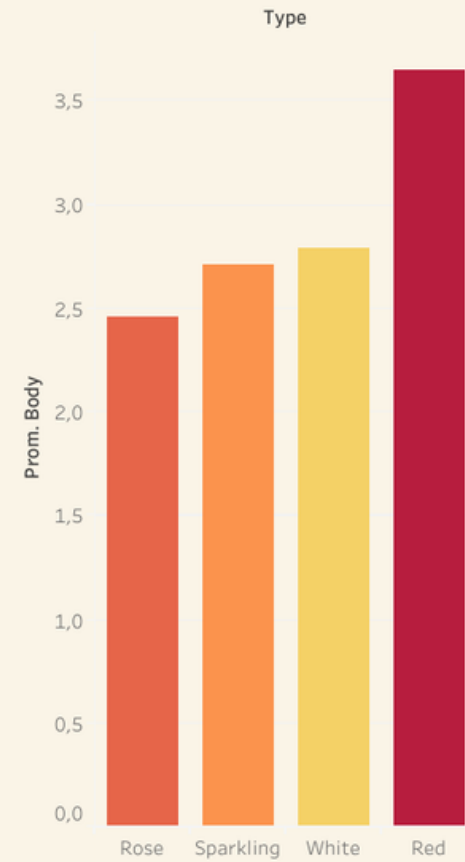
Tannin by Type



Sweetness by Type



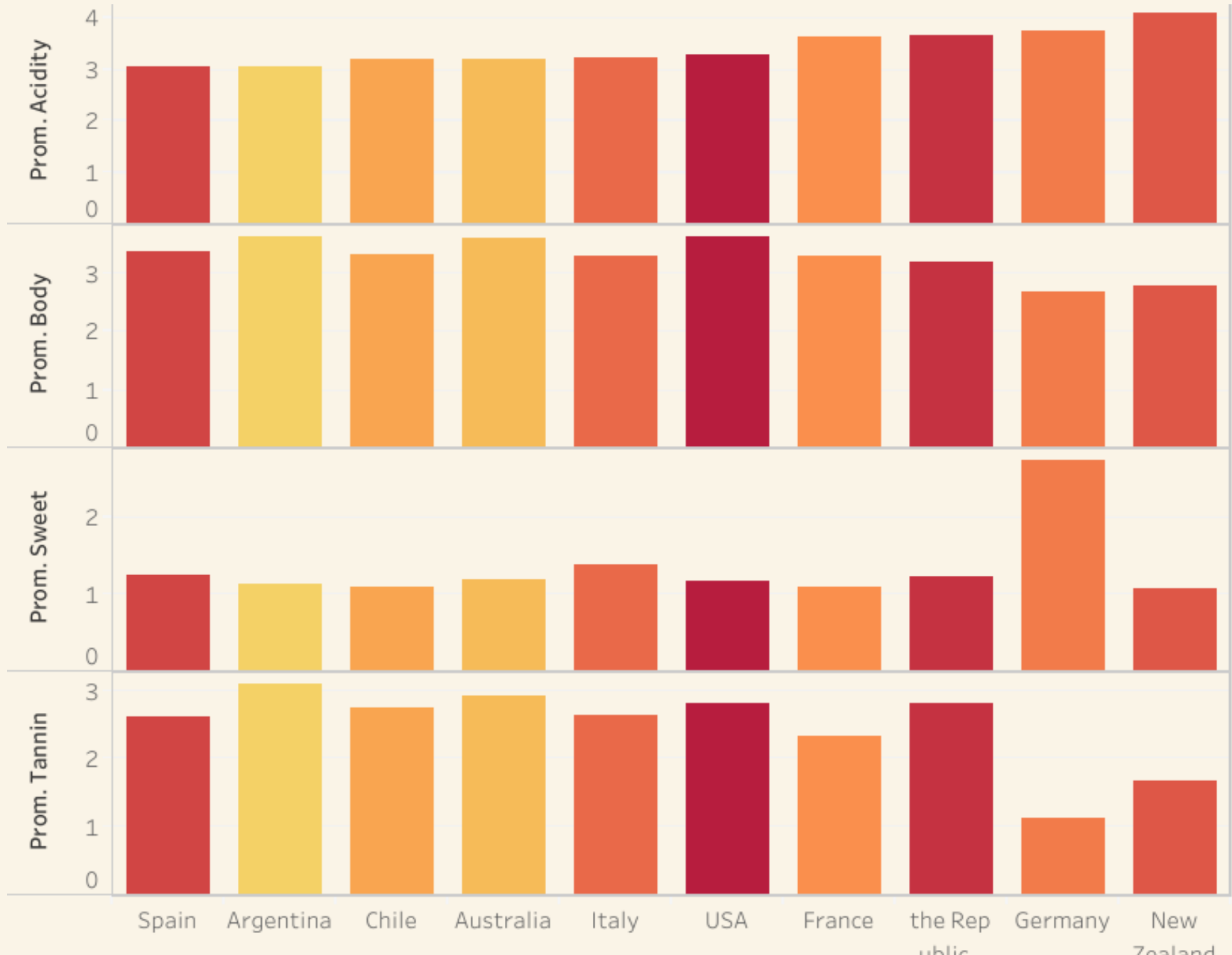
Body by Type



Degrees of Alcohol by Type



Flavour Profile by Country



Two Sample t-test



Pooled variance

$$s_p^2 = \frac{((n_1-1)s_1^2) + ((n_2-1)s_2^2)}{n_1 + n_2 - 2}$$

Calculate t

$$t = \frac{\text{difference of group averages}}{\text{standard error of difference}}$$

- The average price of the wine types is close enough to be equivalent
- T_student to calculate p-value and compare the two

Chi2 Test

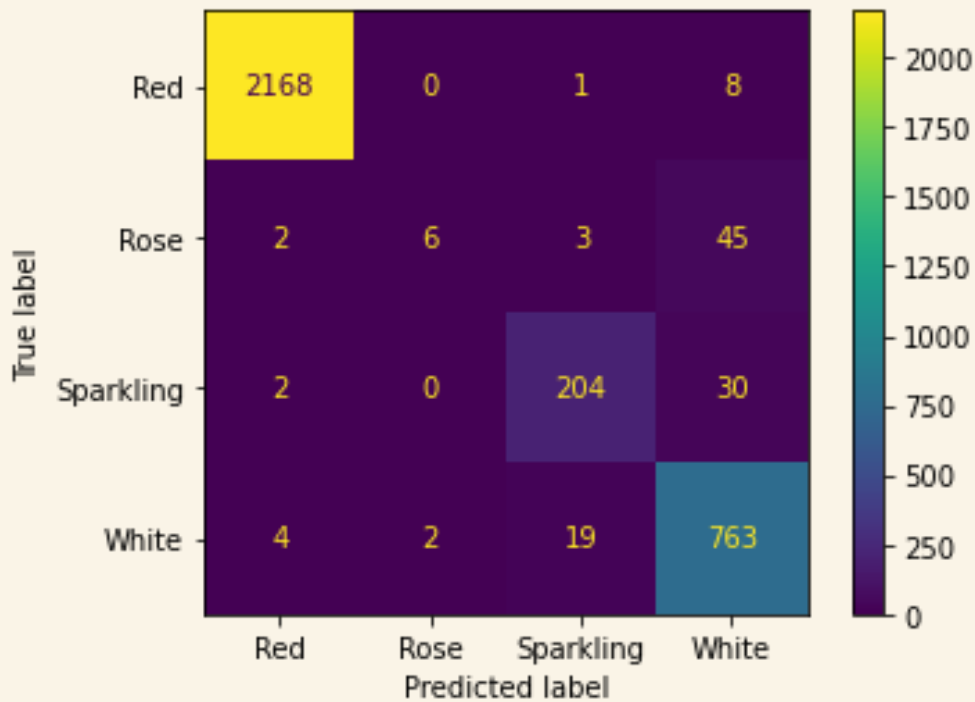
- Testing the relationship of categorical columns
- There is a relationship between the different flavour characteristics and the type of wine

Prediction Model

- Predicting the type of wine using logistic regression
- First results are really good but is it too good to be true?
- Mitigating class Imbalance



Prediction Model Results



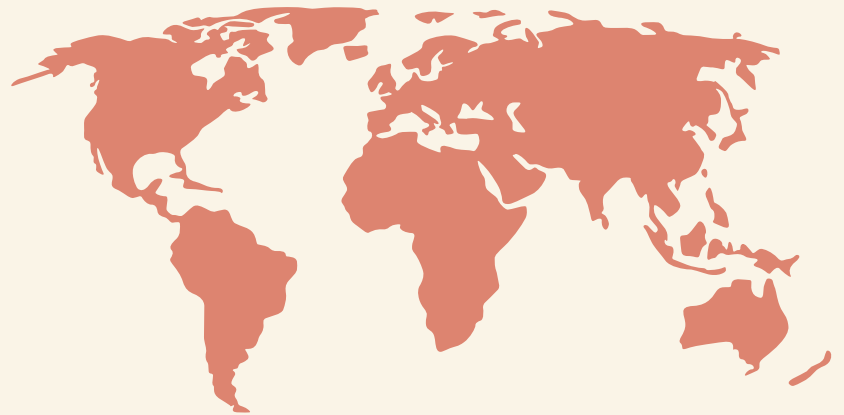
The accuracy in the TEST set is: 0.96
The Kappa in the TEST set is: 0.92

The accuracy in the TEST set is: 0.92
The Kappa in the TEST set is: 0.83



Next steps

- Create a wine recommender
- Use natural language processing to extract information from reviews
- Include ratings for a more complete recommender
- Get data that better reflects real world numbers



Cheers!

