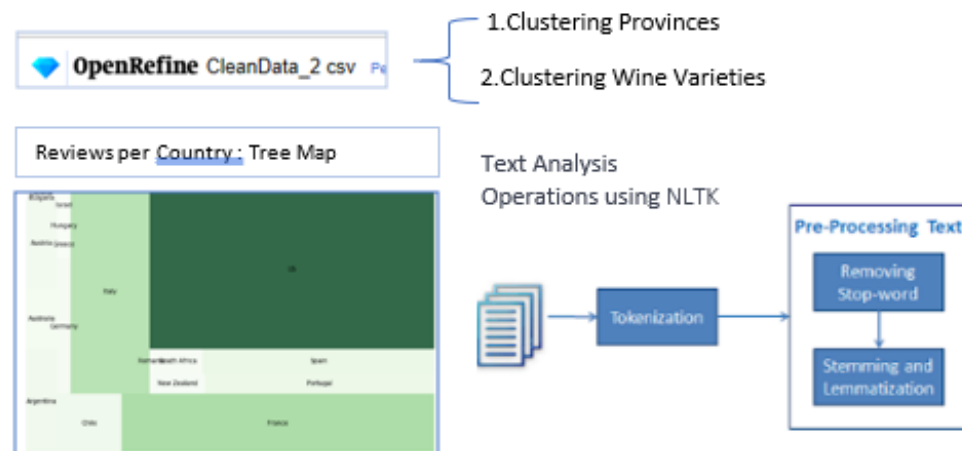


## Cleaning up !



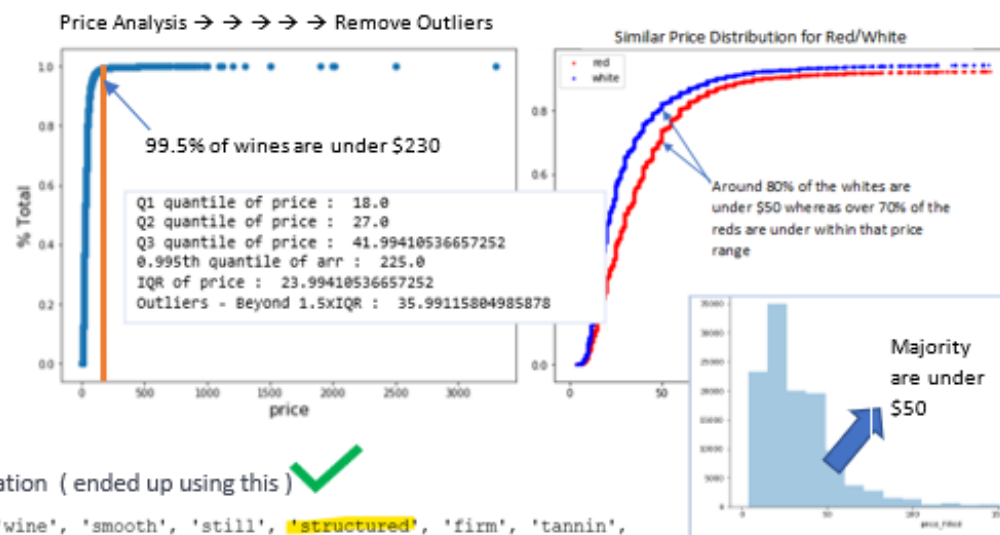
## SAMPLES: Stemming

['ripe', 'fruity', 'wine', 'smooth', 'still', 'structured', 'firm', 'tannin', 'fill', 'juicy', 'red', 'berry', 'fruit', 'freshen', 'acid', 'already', 'drinkable', 'although', 'certainly', 'better']

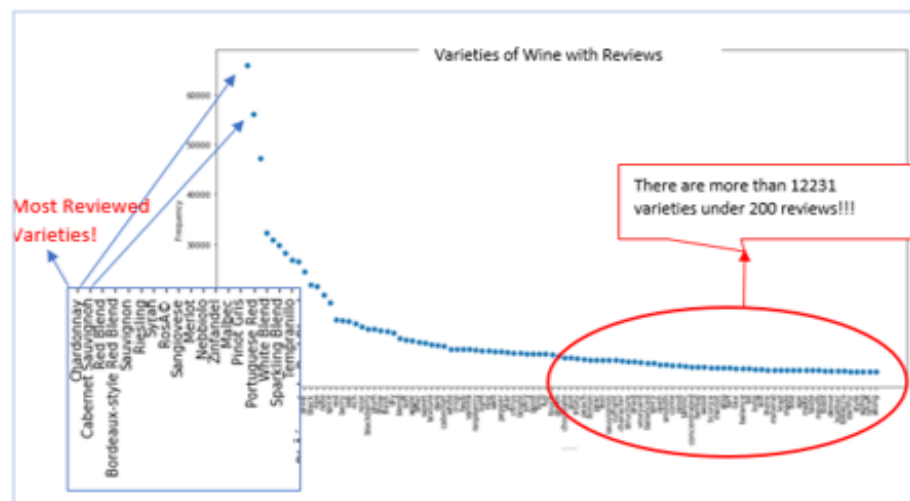
## SAMPLES: Lemmatization ( ended up using this )

['ripe', 'fruity', 'wine', 'smooth', 'still', 'structured', 'firm', 'tannin', 'filled', 'juicy', 'red', 'berry', 'fruit', 'freshened', 'acidity', 'already', 'drinkable', 'although', 'certainly', 'better']

## Preparing Data for Model ! Grouping, Imputations, Outliers Analysis



## Experimenting with varieties to Group them for Binary Classification → → → → End up with 2 classes ( Red and White )



## Classification Model : Random Forest

