

# Wines Reviews

Visualisation de données

Clément Brigliano, David Crevoiserat, Jihad Falfoul

# Sommaire

- Introduction
- Dataset
- Outils utilisés
- Représentations et choix
- Démonstration
- Limitations et perspectives
- Conclusion

# Introduction

- Application de visualisation de données liées au domaine du vin
  - Accent sur la qualité des vins
- Public cible : amateurs de vin
- Objectif : guider un amateur de vin dans son choix

# Dataset

- Trouvé sur Kaggle
- Données scrapées sur WineEnthusiast en 2017
- Différentes informations
  - Pays, note obtenue, prix, critique, etc.
- Note/points
  - Attribuées par des experts
  - De 80 à 100
  - 50% des vins sont notés entre 86 et 90

## Outils utilisés

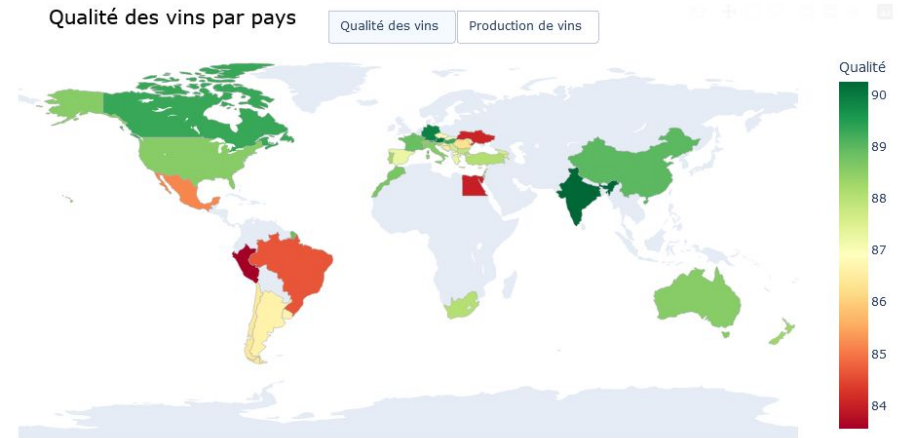


# Représentation et choix : informations générales

- Couleurs adaptées au daltonisme (palette de Paul Tol)
- Storytelling afin de guider l'utilisateur et faciliter la compréhension
- Graphiques interactifs (zoom, survol, filtrage)
  - *“Overview first, zoom and filter, then details-on-demand.”*

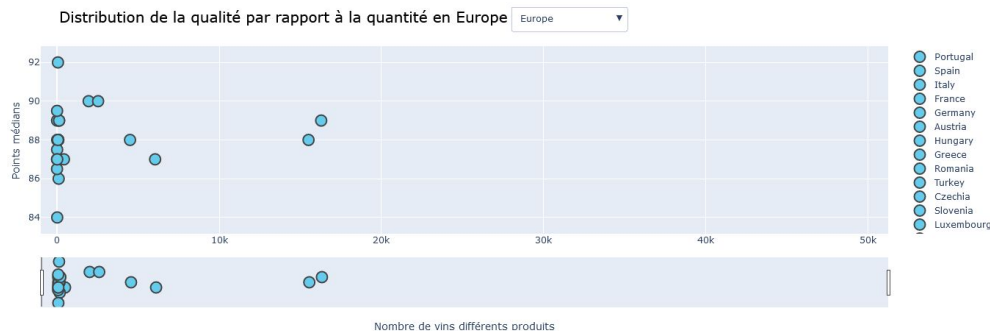
# Représentation et choix

- Carte
  - Adaptée à la nature des données
- Boutons permettant le changement des données représentées
- Zoom possible
- Couleurs intuitives
  - Vert = bon
  - Rouge = mauvais
- Objectif : introduire le sujet



# Représentation et choix

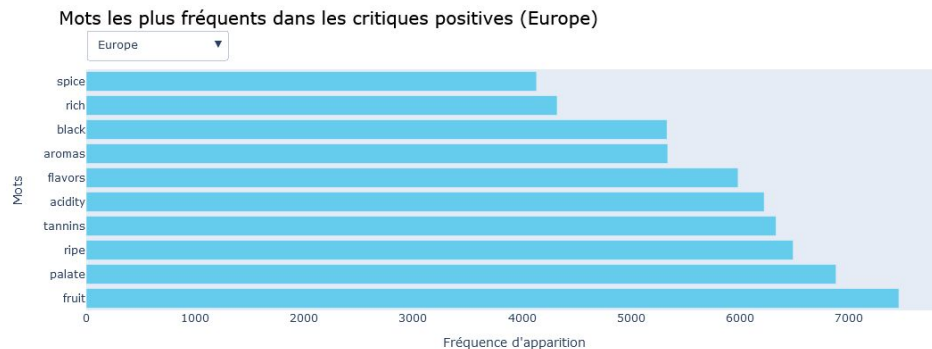
- Nuage de points
  - Met en évidence une potentielle corrélation
- Filtrage par continent et pays
- Zoom possible
- Objectif : un pays qui produit beaucoup est-il à privilégier (ou inversement) ?





# Représentation et choix

- Histogramme horizontal
  - Choix de ne pas utiliser un wordcloud par soucis de visibilité
- Liste déroulante permettant le filtrage
- Objectif : trouver les mots qui définissent les meilleurs vins

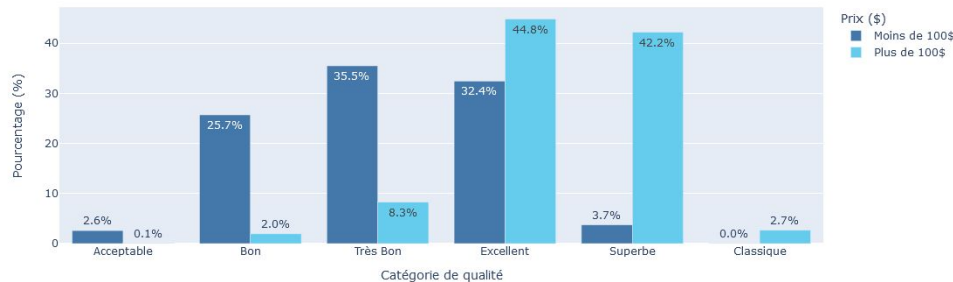


# Représentation et choix

- Histogramme empilé
  - Adapté aux données (deux pourcentage par catégorie)
- Champ texte pour définir le prix
  - La notion de “cher” n’est pas la même pour tout le monde
- Objectifs :
  - Montrer que le prix est indicateur de la catégorie
  - Permettre à l'utilisateur de “prédire” la qualité en fonction du prix qu’il souhaite mettre

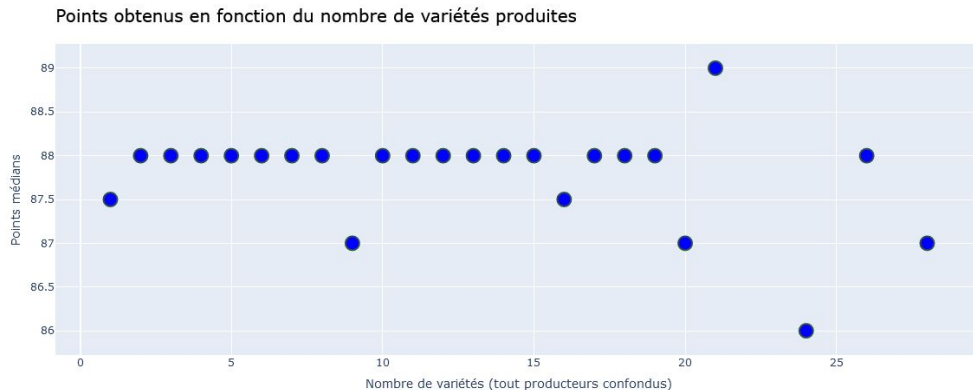
Limite de prix (\$) :

Pourcentage de chance qu'un vin appartienne à une catégorie en fonction de son prix (100\$)



# Représentation et choix

- Nuage de points
  - Met en évidence la potentielle corrélation
- Objectif : un vigneron qui produit beaucoup est-il à privilégier (ou inversement) ?



# Démonstration

# Limitations et perspectives

- “Pourcentage de chance qu’un vin appartienne à une catégorie en fonction de son prix” : ajouter un nuage de points lors du clic sur une barre.
- Utiliser Dask à la place de pandas

# Conclusion

- Toutes les questions définies ont été répondues via une visualisation
- Application complète et fonctionnelle
- Applique les concepts vu en cours

Questions ?