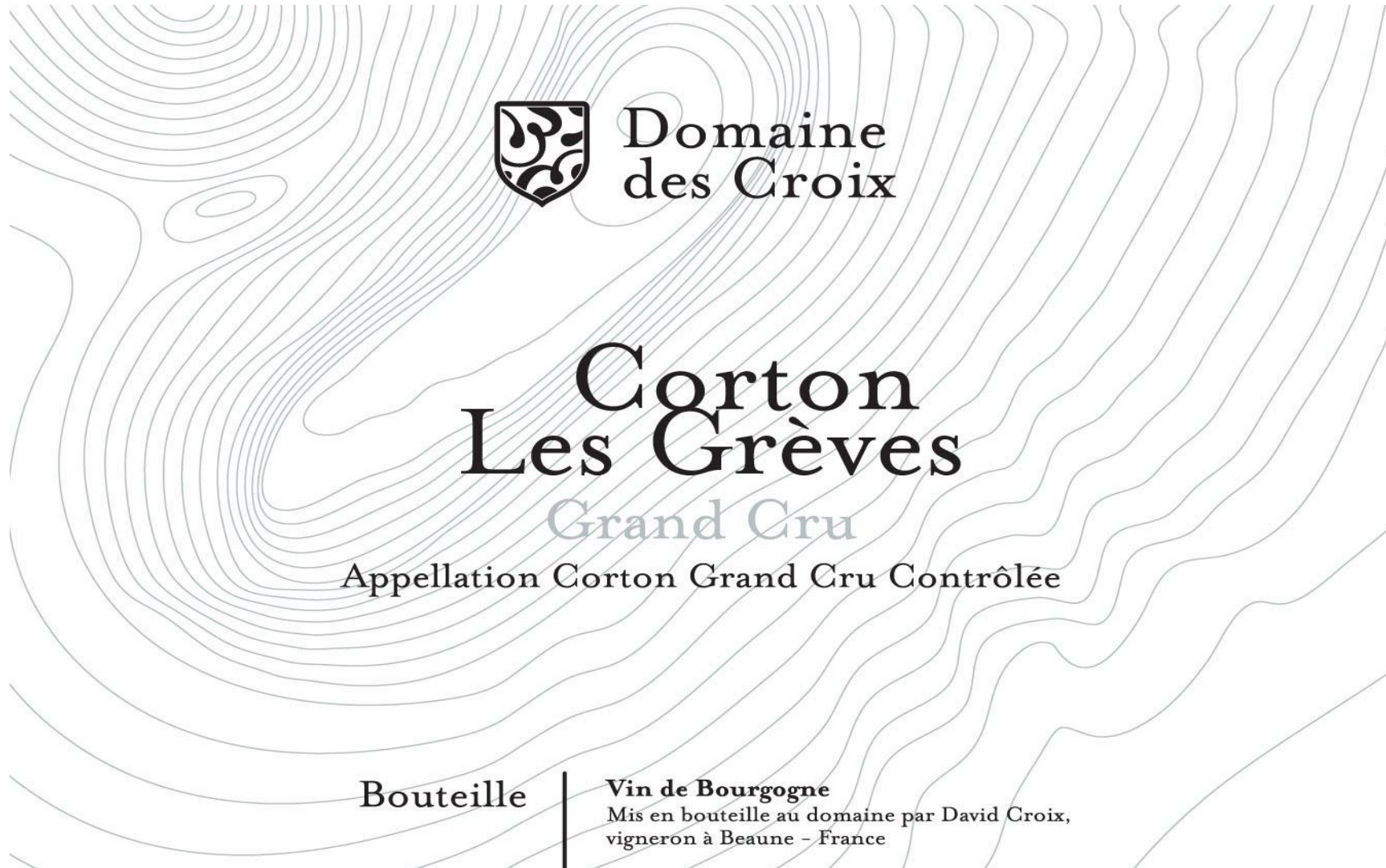


ETUDE DE MARCHÉ DES VINS AUX ETATS UNIS



ETUDE DE MARCHE DES VINS AUX ETATS UNIS

Domaine des Croix 2016 Corton Grèves

« Ferme et solide avec de grandes notes de fruits noirs ainsi que des tannins denses. Les fruits mûrs sont un élément majeur de ce vin concentré. Le vin a de la puissance et un noyau tannique ferme, et a certainement besoin d'être vieilli pendant sept ans. »



RAPPEL DU CONTEXTE ET DE LA PROBLEMATIQUE

MISSION

Réaliser une étude de marché sur les vins aux Etats-Unis

Objectif: fixer le prix d'une bouteille de vin du client

CLIENT

Le Domaine des Croix

PRODUIT

Domaine des Croix 2016 Corton Grèves



EXPLORATION DES DONNEES & METHODOLOGIE

Données disponibles

Un fichier de 130K références de vins fourni par le client
(cépages, pays et régions de production, millésimes et notes des produits)
Le descriptif de la bouteille pour laquelle le prix est à fixer



Retraitement et nettoyage des données

Outils utilisés:

- 1/ Python et ses librairies (pandas, matplotlib) pour nettoyer, explorer les données
- 2/ Power BI pour réaliser les visualisations



Créations de deux jeux de données

- 1/ un fichier de **11 K** enregistrements pour analyse de marché du Pinot Noir
- 2/ un fichier **2.8 K** enregistrements pour analyse du marché du « Burgundy »

EXPLORATION DES DONNEES & METHODOLOGIE

Données disponibles

Un fichier de 130K références de vins fourni par le client
(cépages, pays et régions de production, millésimes et notes des produits)

Récapitulatif des données à traiter, explorer et nettoyer

```
1 link = "https://github.com/WildCodeSchool/wilddata/raw/main/wine.zip"
2 df_wine = pd.read_csv(link)
3
4 df_wine
```

✓ 2.9s Open 'df_wine' in Data Wrangler

	country	description	designation
0	Italy	Aromas include tropical fruit, broom	Vulkà Bianco
1	Portugal	This is ripe and fruity, a wine that is s	Avidagos
2	US	Tart and snappy, the flavors of lime f	Missing value
3	US	Pineapple rind, lemon pith and oran	Reserve Late Harvest
4	US	Much like the regular bottling from 2	Vintner's Reserve Wild Child Block
5	Spain	Blackberry and raspberry aromas sho	Ars In Vitro
6	Italy	Here's a bright, informal red that op	Belsito
7	France	This dry and restrained wine offers s	Missing value
8	Germany	Savory dried thyme notes accent sur	Shine
9	France	This has great depth of flavor with it	Les Natures

129,957 rows x 13 cols 10 per page << < Page 1 of 12996 >

```
1 fn_exploration(df_wine)
```

✓ 1.8s

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 129957 entries, 0 to 129956
Data columns (total 13 columns):
#   Column                                Non-Null Count  Dtype
---  ---                                ---
0   country                              129894 non-null object
1   description                          129957 non-null object
2   designation                          92497 non-null  object
3   points                              129957 non-null int64
4   price                               120964 non-null float64
5   province                            129894 non-null object
6   region_1                            108710 non-null object
7   region_2                            50511 non-null  object
8   taster_name                         103713 non-null object
9   taster_twitter_handle               98744 non-null  object
10  title                               129957 non-null object
11  variety                             129956 non-null object
12  winery                              129957 non-null object
dtypes: float64(1), int64(1), object(11)
memory usage: 12.9+ MB
```

EXPLORATION DES DONNEES & METHODOLOGIE

Retraitement et nettoyage des données

Outils utilisés:

- 1/ Python et ses librairies (pandas, matplotlib) pour nettoyer, explorer les données
- 2/ Power BI pour réaliser les visualisations

Création d'une fonction d'exploration du jeu de données

```
1 # je commence mon exploration du dataset
2 # je crée une fonction d'exploration
3
4 def fn_exploration(df):
5     display(
6         df.info(),
7         f"valeurs manquantes",
8         df.isna().sum(),
9         f"describe",
10        df.describe(),
11        f"dtypes",
12        df.dtypes,
13        f"value_counts",
14        df.value_counts(),
15        f"duplicates",
16        df.duplicated().sum(),
17        f"describe corr",
18        df.describe().corr()
19    )
✓ 0.0s
```

Création d'une fonction d'exploration pour les colonnes

```
1 # je crée une fonction d'exploration pour les colonnes
2 def fn_explo_colonne(column_name):
3     display(
4         f"valeurs manquantes",
5         df_wine[column_name].isna().sum(),
6         f"value counts",
7         df_wine[column_name].value_counts(),
8         f"% value counts",
9         df_wine[column_name].value_counts(normalize=True)*100,
10        f"valeurs uniques",
11        df_wine[column_name].unique()
12    )
✓ 0.0s
```

Première approche statistique

```
1 df_wine['points'].describe().to_frame()
✓ 0.0s
```

# points	
count	119975.0
mean	88.44172535944989
std	3.092645845074776
min	80.0
25%	86.0
50%	88.0
75%	91.0
max	100.0

8 rows x 1 cols 10 per page

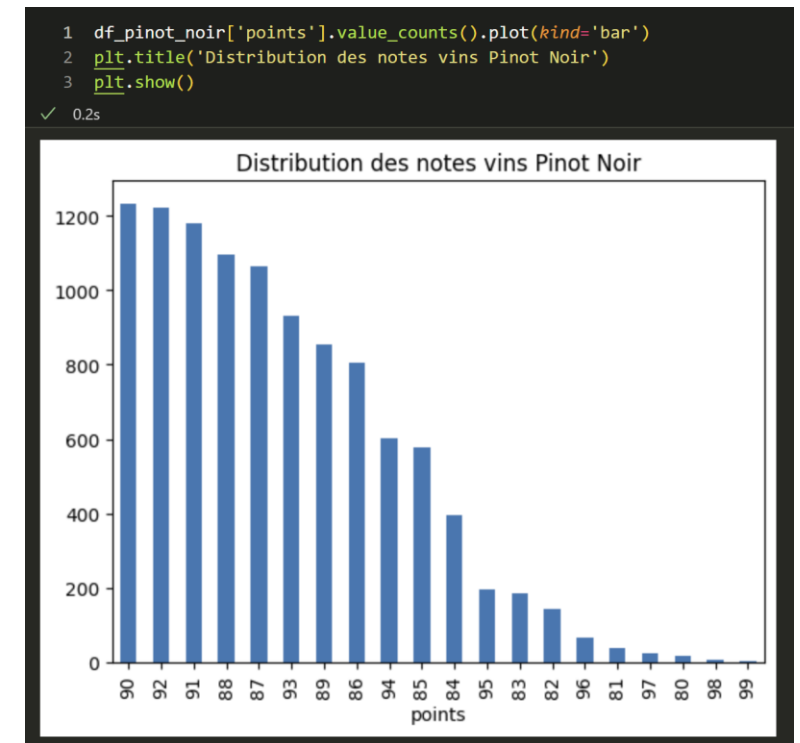
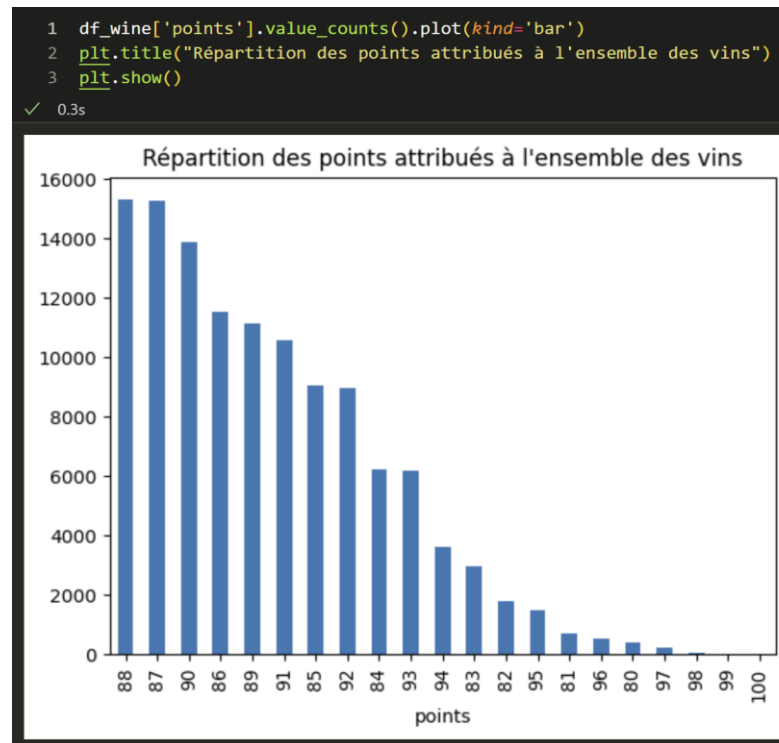
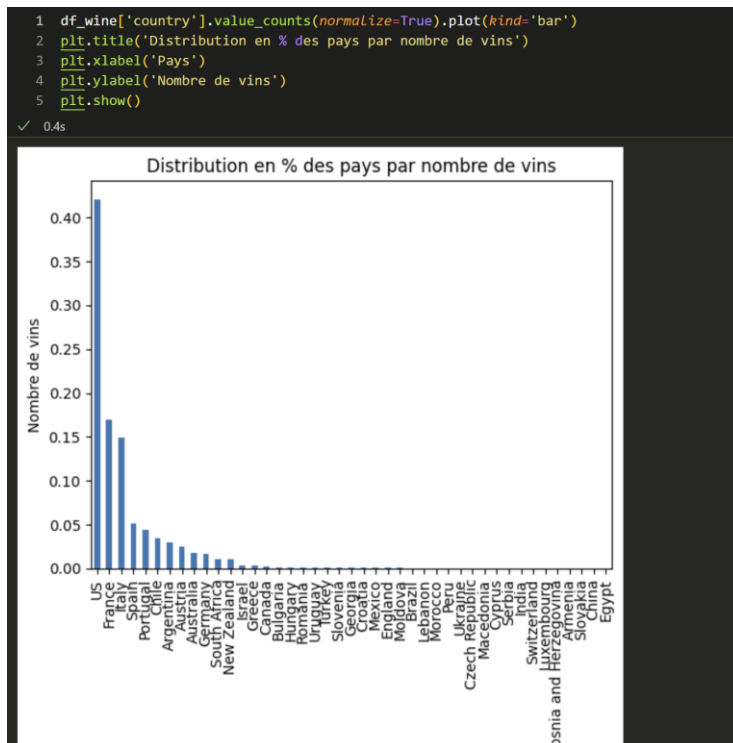
EXPLORATION DES DONNEES & METHODOLOGIE

Retraitement et nettoyage des données

Outils utilisés:

- 1/ Python et ses librairies (pandas, matplotlib) pour nettoyer, explorer les données
- 2/ Power BI pour réaliser les visualisations

Réalisation de graphiques pour explorer les données: cette première approche permet d'appréhender la problématique et donne des clés de lecture pertinentes



EXPLORATION DES DONNEES & METHODOLOGIE

Créations de deux jeux de données

- 1/ un fichier de **11 K** enregistrements pour analyse de marché du Pinot Noir
- 2/ un fichier **2.8 K** enregistrements pour analyse du marché du « Burgundy »

Création du jeu de données Pinot Noir

```
1 df_wine[df_wine['variety'] == 'Pinot Noir']
```

✓ 0.0s

	country	description	designation
4	US	Much like the regular bottling from :	Vintner's Reserve Wild Child Block
21	US	A sleek mix of tart berry, stem and h	Missing value
25	US	Oak and earth intermingle around rc	King Ridge Vineyard
35	US	As with many of the Erath 2014 vine	Hyland
41	US	A stiff, tannic wine, this slowly opens	Missing value
58	Chile	Lightly herbal strawberry and raspbe	Reserve
78	US	Some rosés are made simply by blee	Rosé of
145	US	Cooked cranberry is spiced with anis	Guidotti Vineyard
152	US	From a promising new project, the n	Estate Reserve
169	US	This wine is put together from multi	Dutton Ranch

11,826 rows x 14 cols 10 per page << < Page 1 of 1183 >

```
1 df_pinot_noir_final = df_wine[df_wine['variety'] == 'Pinot Noir']
```

✓ 0.0s

Création du jeu de données « Burgundy »

```
1 df_wine[df_wine['province'] == 'Burgundy']
```

✓ 0.0s

	country	description	designation
65	France	From the warm 2019 vintage, this is	Missing value
66	France	This soft, rounded wine is ripe with c	Missing value
179	France	The impact of 50-year old vines is se	Les Vénérables Vieilles Vignes
193	France	A citrus-dominated wine, lime and le	Missing value
307	France	Rich and full in the mouth, packed w	Les Dames Huguettes
353	France	This beautiful, rich wine has yellow fr	Le Montrachet
357	France	Produced from vines more than 100	Vieilles Vignes
359	France	Up the slope above the Grand Cru o	Les Chenevottes Premier Cru
363	France	This is an opulent wine from one of	Missing value
462	France	This big, ripe and fruity wine shows	Blagny Premier Cru

2,862 rows x 14 cols 10 per page << < Page 1 of 287 > + Code + Markdown

```
1 df_burgundy = df_wine[df_wine['province'] == 'Burgundy']
```

✓ 0.0s

LE MARCHE DES VINS AUX ETATS UNIS

ETUDE EN GRANULARITE

La marché US au global

Les producteurs

Le marché du Pinot Noir

Le marché du Pinot Noir, millésimé 2016

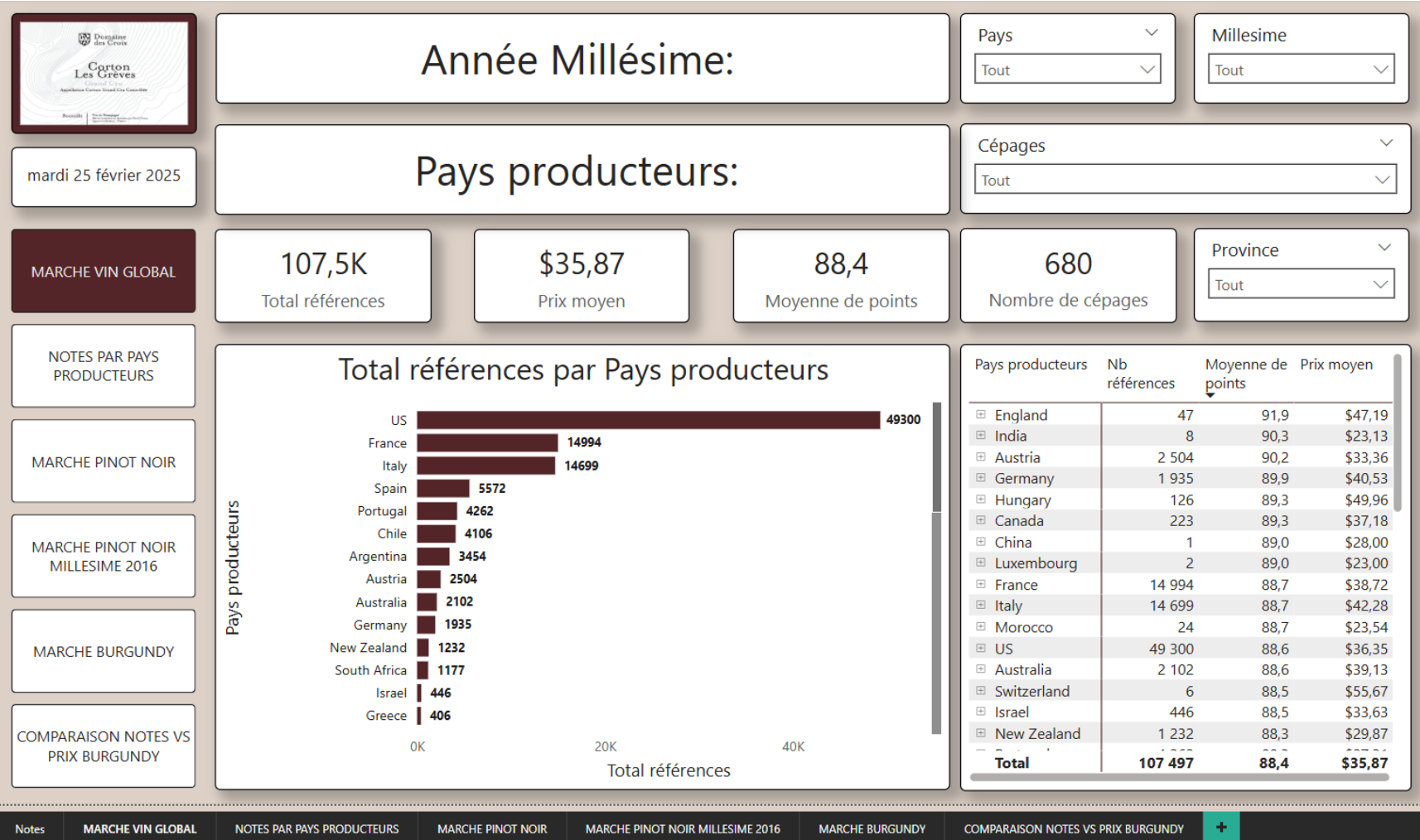
L'évolution des prix du « Burgundy »

Conclusion

Analyse de sentiment

LE MARCHE DES VINS AUX ETATS UNIS

Approche globale du marché US



Nombre de références

Prix moyen

Moyenne de points

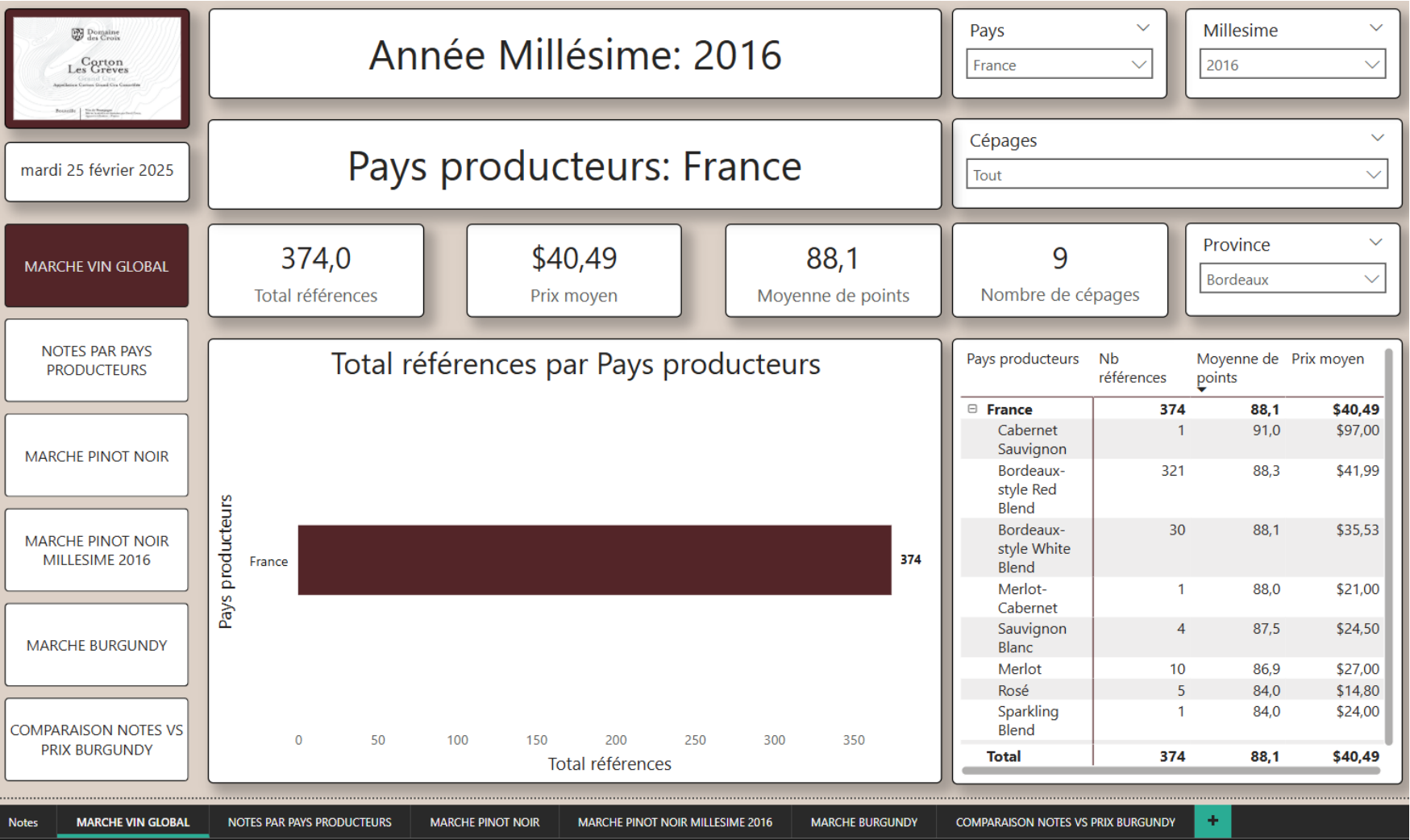
Nombre de cépages

Filtres sur pays producteurs, millésimes, province et cépages

Une table pour visualiser les pays producteurs et les cépages associés


LE MARCHE DES VINS AUX ETATS UNIS

Approche globale du marché US



LE MARCHE DES VINS AUX ETATS UNIS

Notes par pays producteurs



mardi 25 février 2025

MARCHE VIN GLOBAL

NOTES PAR PAYS PRODUCTEURS

MARCHE PINOT NOIR

MARCHE PINOT NOIR MILLESIME 2016

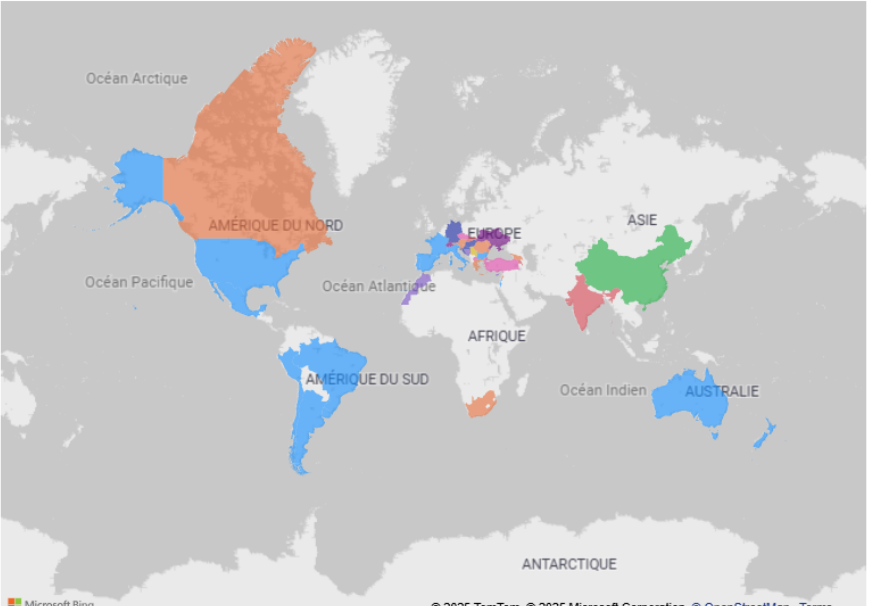
MARCHE BURGUNDY

COMPARAISON NOTES VS PRIX BURGUNDY

Année Millésime:

Pays producteurs:

Moyenne de points par pays producteurs



Pays producteurs

Tout

Millesime

Tout

88,4

Moyenne de points

680

Nombre de cépages

Pays producteurs	Nb références	Moyenne de points	Prix moyen
England	47	91,9	\$47,19
India	8	90,3	\$23,13
Austria	2 504	90,2	\$33,36
Germany	1 935	89,9	\$40,53
Hungary	126	89,3	\$49,96
Canada	223	89,3	\$37,18
China	1	89,0	\$28,00
Luxembourg	2	89,0	\$23,00
France	14 994	88,7	\$38,72
Italy	14 699	88,7	\$42,28
Morocco	24	88,7	\$23,54
US	49 300	88,6	\$36,35
Australia	2 102	88,6	\$39,13
Switzerland	6	88,5	\$55,67
Israel	446	88,5	\$33,63
New Zealand	1 232	88,3	\$29,87
Portugal	4 262	88,3	\$27,31
Total	107 497	88,4	\$35,87

Moyenne de points

Nombre de cépages

Une carte pour visualiser les pays producteurs

Filtres sur pays producteurs et millésimes

Une table pour visualiser les pays producteurs et les cépages associés

Notes

MARCHE VIN GLOBAL

NOTES PAR PAYS PRODUCTEURS

MARCHE PINOT NOIR

MARCHE PINOT NOIR MILLESIME 2016


MARCHE BURGUNDY

COMPARAISON NOTES VS PRIX BURGUNDY

+

LE MARCHE DES VINS AUX ETATS UNIS

Notes par pays producteurs



mardi 25 février 2025

MARCHE VIN GLOBAL

NOTES PAR PAYS PRODUCTEURS

MARCHE PINOT NOIR

MARCHE PINOT NOIR MILLESIME 2016

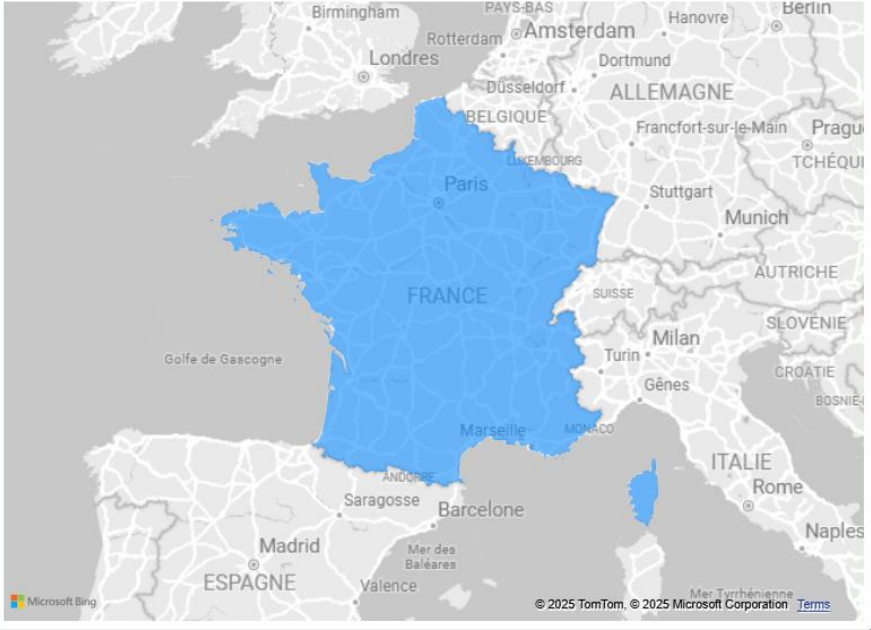
MARCHE BURGUNDY

COMPARAISON NOTES VS PRIX BURGUNDY

Année Millésime: 2015

Pays producteurs: France

Moyenne de points par pays producteurs



Pays producteurs

France

Millesime

2015

89,3

Moyenne de points

58

Nombre de cépages

Pays producteurs	Nb références	Moyenne de points	Prix moyen
France	979	89,3	\$54,71
Alsace white blend	2	94,5	\$64,00
Tannat-	1	94,0	\$51,00
Malbec-Tannat	1	93,0	\$39,00
Merlot-Tannat	1	93,0	\$39,00
Pinot Meunier	1	93,0	\$81,00
Provence red blend	1	93,0	\$45,00
Champagne Blend	13	92,4	\$103,92
Gros and Petit Manseng	1	92,0	\$38,00
Gros Manseng	1	92,0	\$38,00
Loin de l'Oeil	1	92,0	\$44,00
Tannat	9	91,6	\$41,44
Syrah-	1	91,0	\$33,00
Mourvèdre			
Total	979	89,3	\$54,71

Notes

MARCHE VIN GLOBAL

NOTES PAR PAYS PRODUCTEURS

MARCHE PINOT NOIR

MARCHE PINOT NOIR MILLESIME 2016

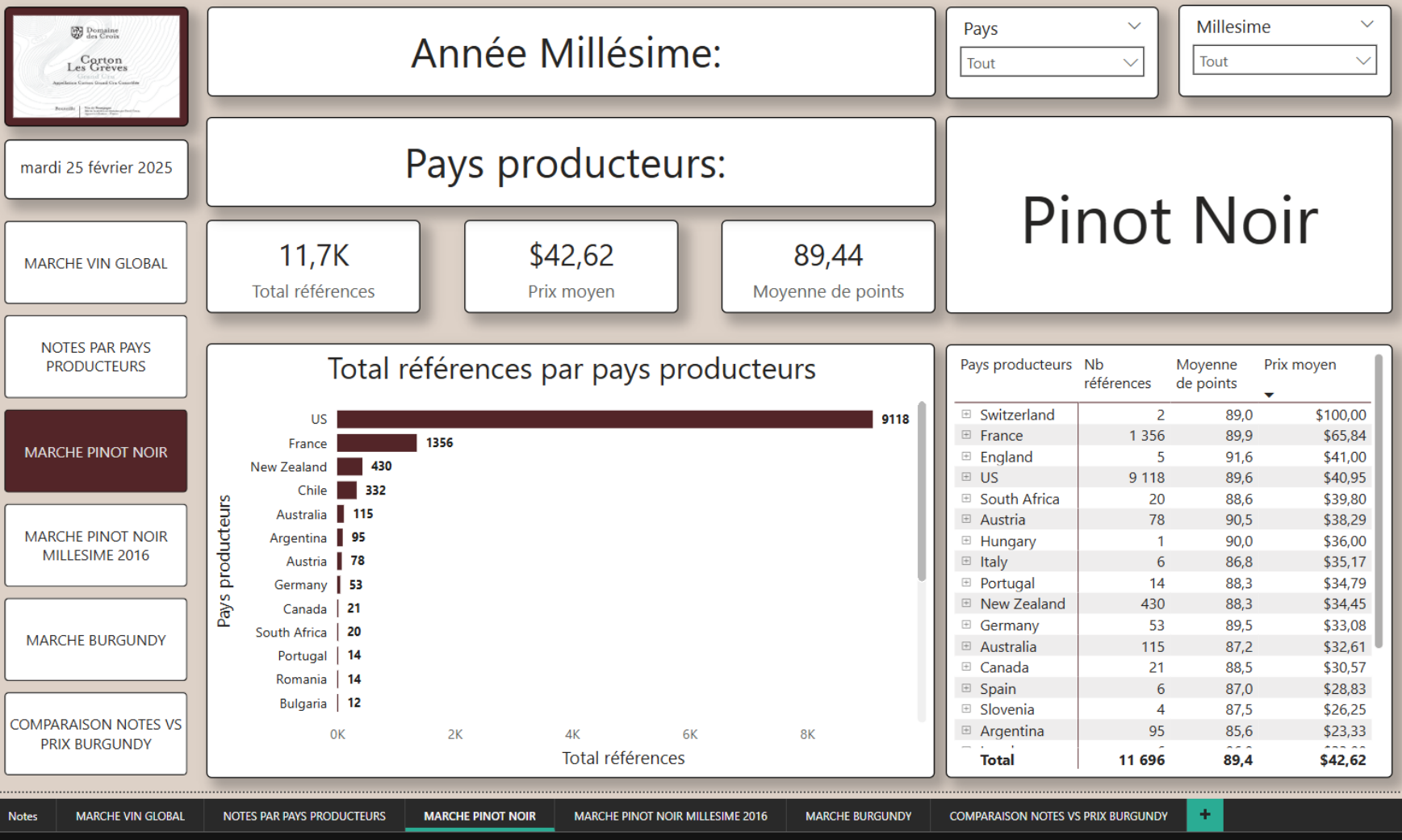
MARCHE BURGUNDY

COMPARAISON NOTES VS PRIX BURGUNDY

+

LE MARCHE DES VINS AUX ETATS UNIS

Le marché du Pinot noir



Nombre de références

Moyenne de points

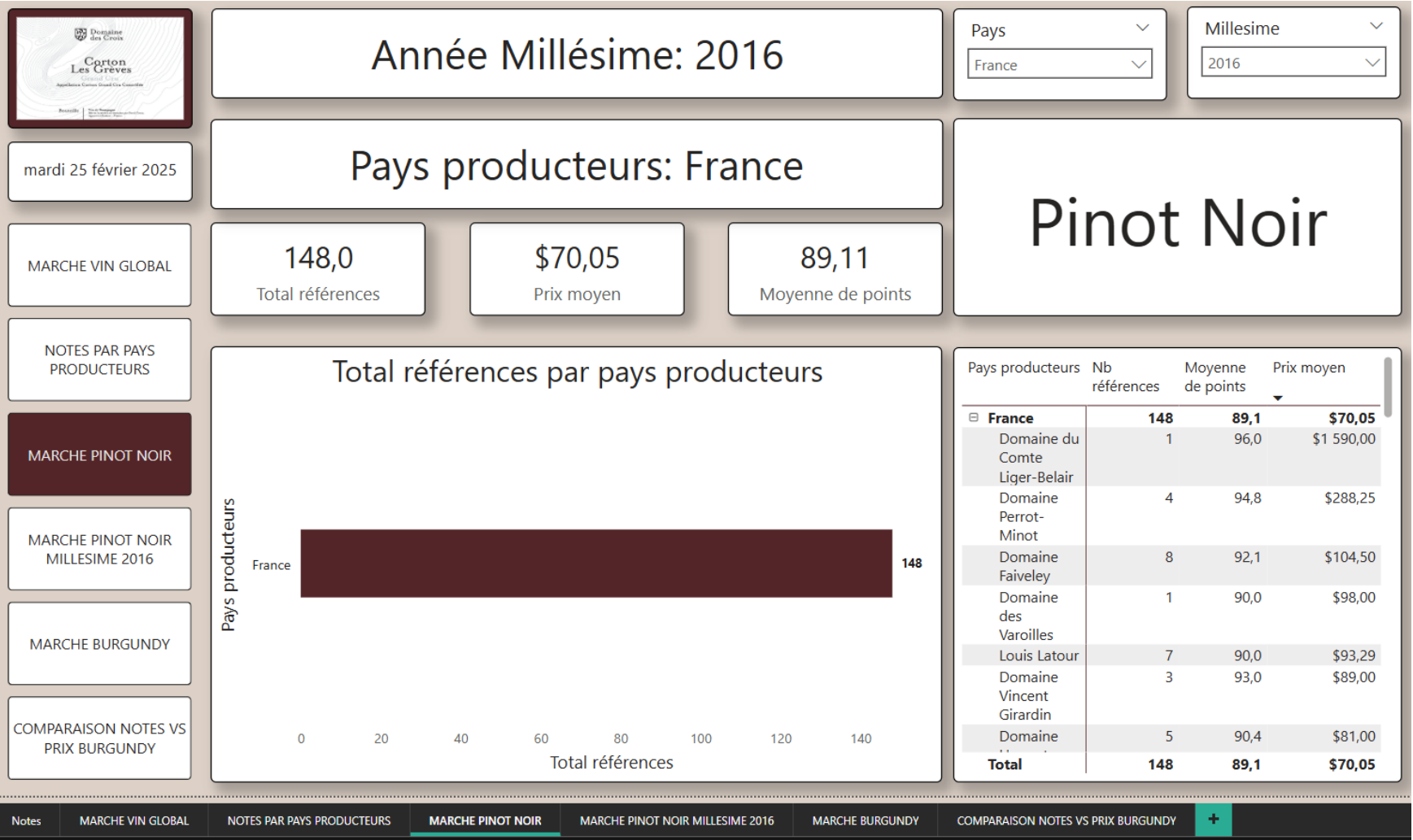
Prix moyen

Filtres sur pays producteurs et millésimes

Une table pour visualiser les pays producteurs et les cépages associés

LE MARCHE DES VINS AUX ETATS UNIS

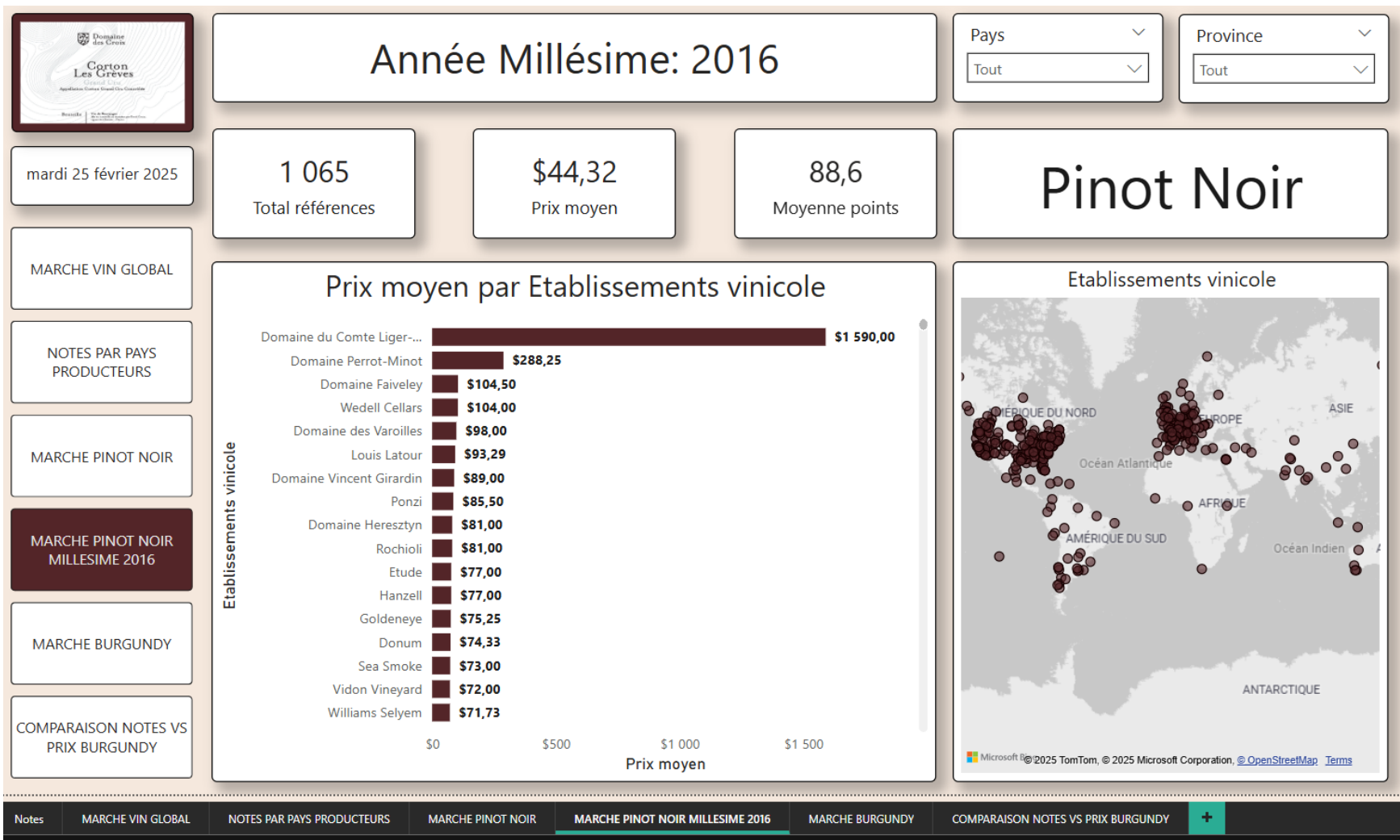
Le marché du Pinot noir



Exemple d'utilisation

LE MARCHE DES VINS AUX ETATS UNIS

Le marché du Pinot noir, Millésime 2016



Nombre de références

Moyenne de points

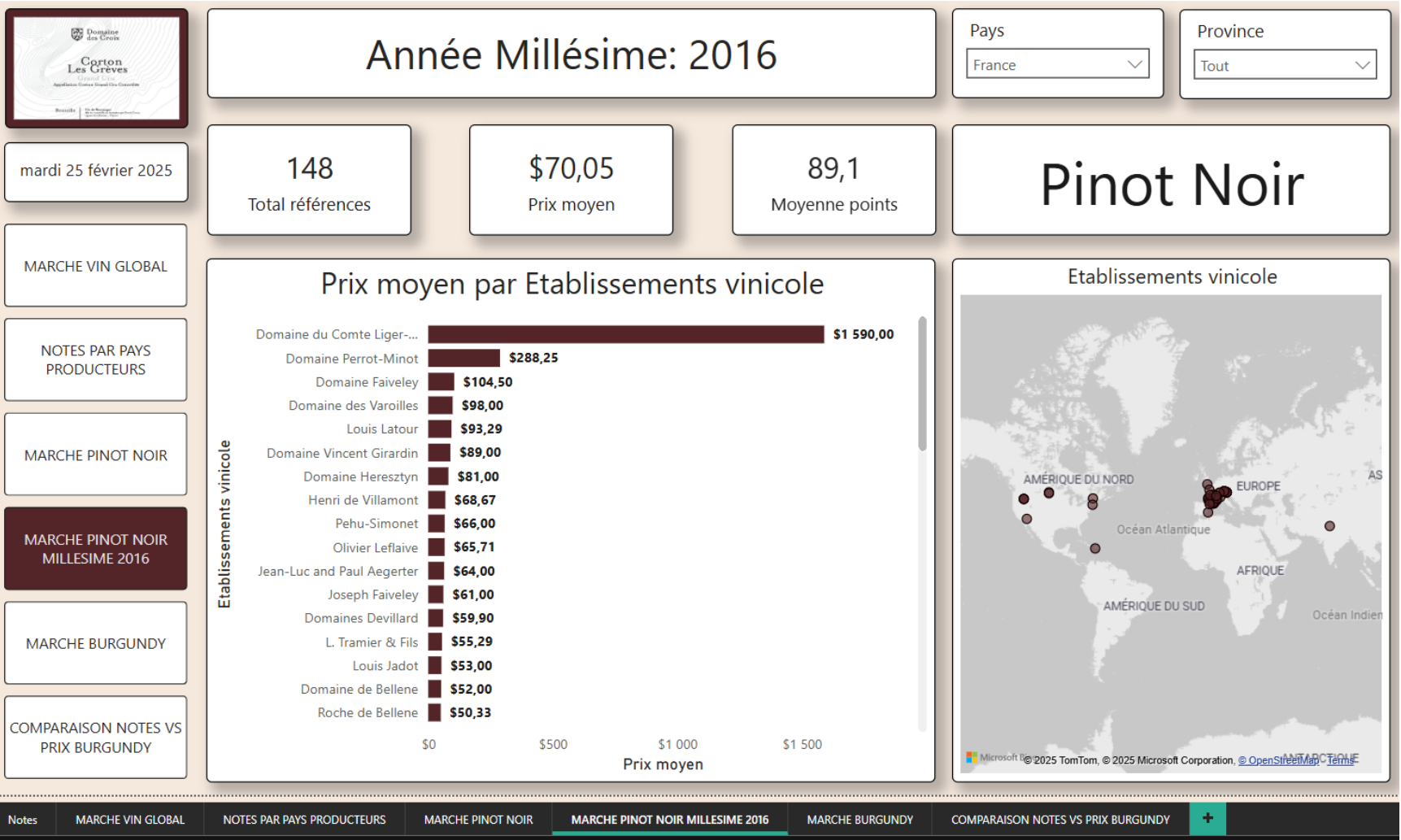
Prix moyen

Une carte pour visualiser les établissements
vinoles

Filtres sur pays producteurs et
province

LE MARCHE DES VINS AUX ETATS UNIS

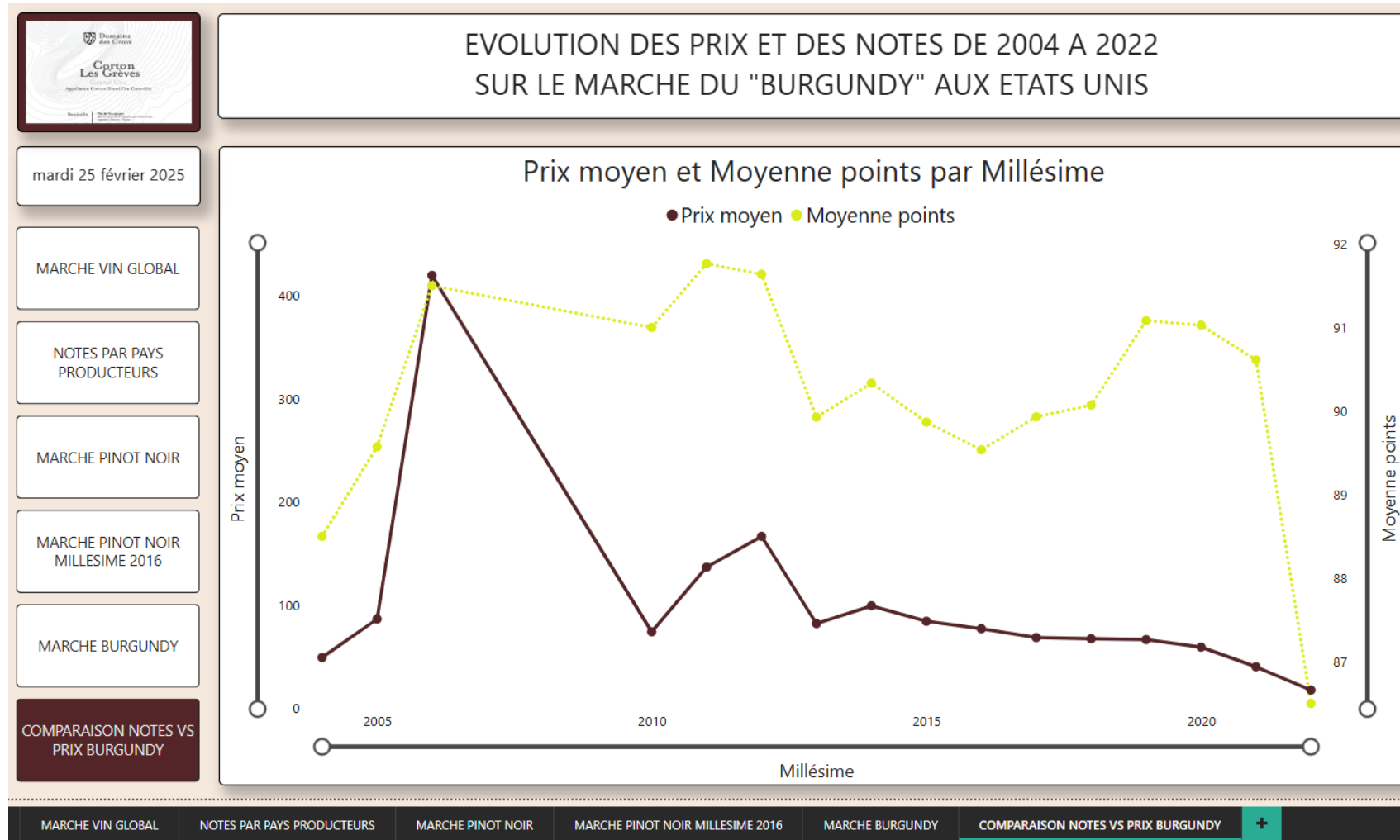
Le marché du Pinot noir, Millésime 2016



Exemple d'utilisation

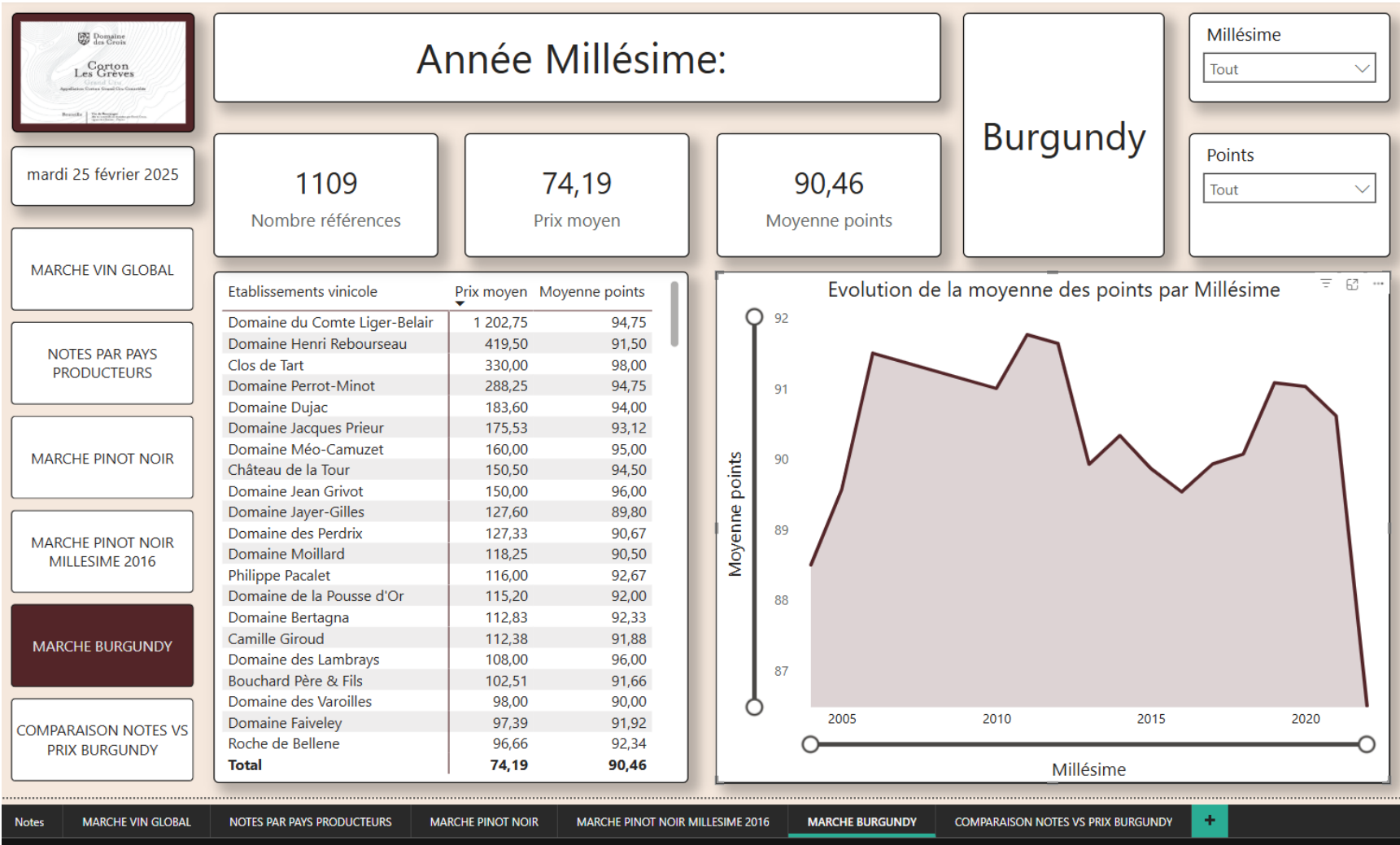
LE MARCHE DES VINS AUX ETATS UNIS

Evolution des prix sur le marché du « Burgundy »



LE MARCHE DES VINS AUX ETATS UNIS

Le marché du « Burgundy »



Moyenne de points

Filtres sur millésime et points

Nombre de références

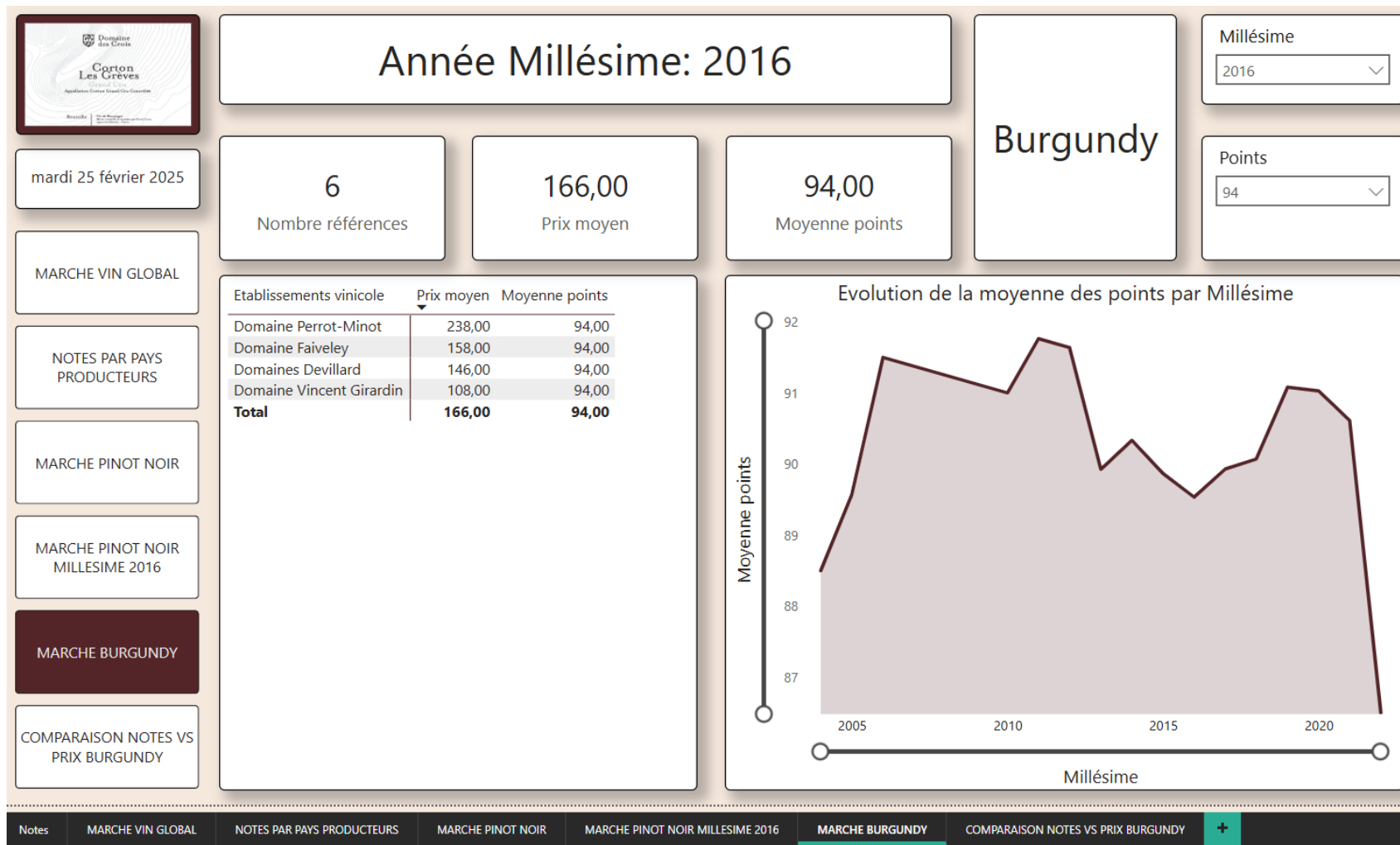
Prix moyen

Une table pour visualiser les établissements
vinicoles avec les prix moyens et la
moyenne de points

Un graphique montrant l'évolution des
points par millésime

LE MARCHE DES VINS AUX ETATS UNIS

Le marché du « Burgundy »



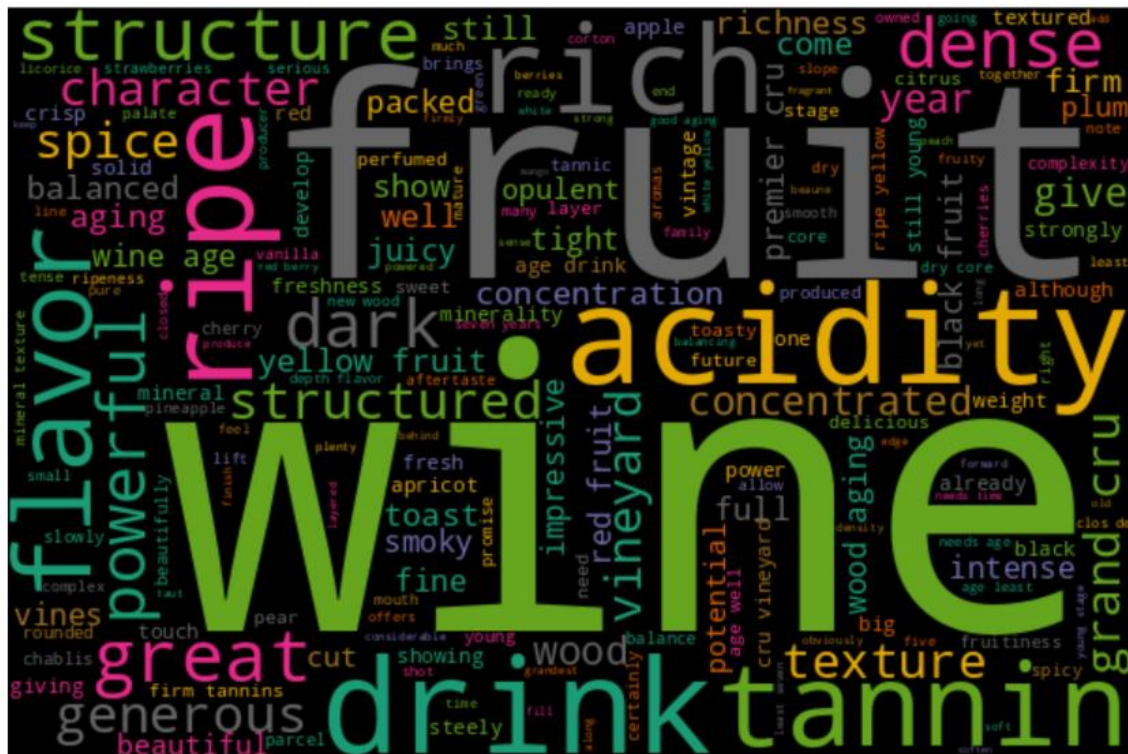
CONCLUSION

Ici, nous remarquons que les prix des bouteilles de vin « Burgundy », millésimée 2016, sont en moyenne de 166 dollars.

D'après cette analyse, je peux conclure que le prix de cette bouteille de vin pourrait être dans une fourchette de 120/ 170 dollars US.

LE MARCHE DES VINS AUX ETATS UNIS

WORDCLOUD SELON LA NOTE MOYENNE ATTRIBUEE AU "BURGUNDY"



points

94

L'utilisation de l'algorithme de « Natural Language Processing », permet d'afficher les mots qui ressortent le plus sur l'appellation « Burgundy », le tout au regard des notes obtenues pour chaque cuvée.

Il est très utile pour effectuer des analyses de sentiment, et nous renvoie une juste perception des commentaires.