



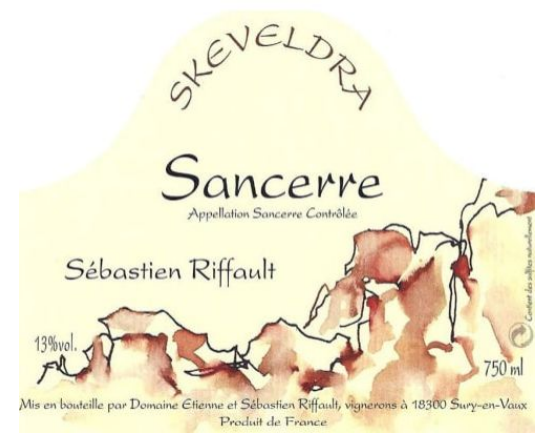
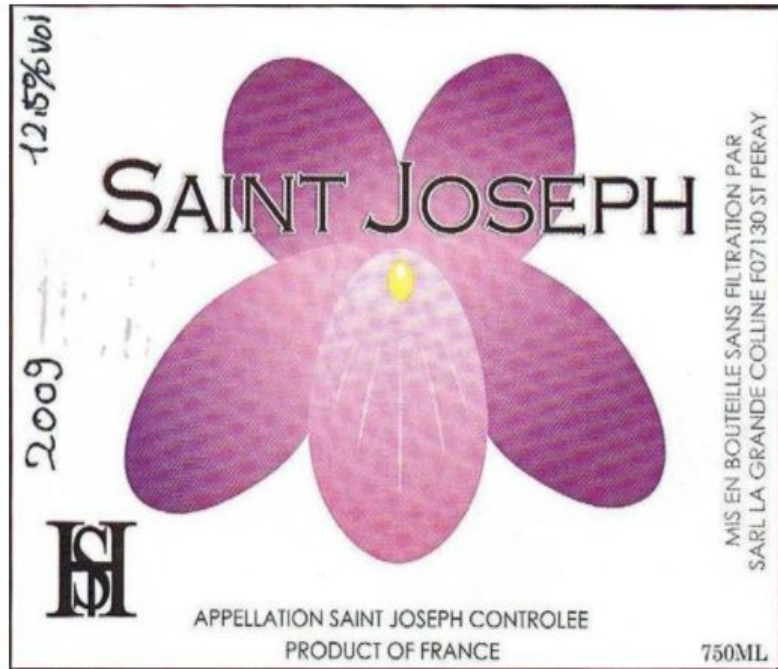
Pick me a Wine!

*Using Natural Language Processing to
select wines*

Jobeth Muncy

What can you learn from a Wine label?

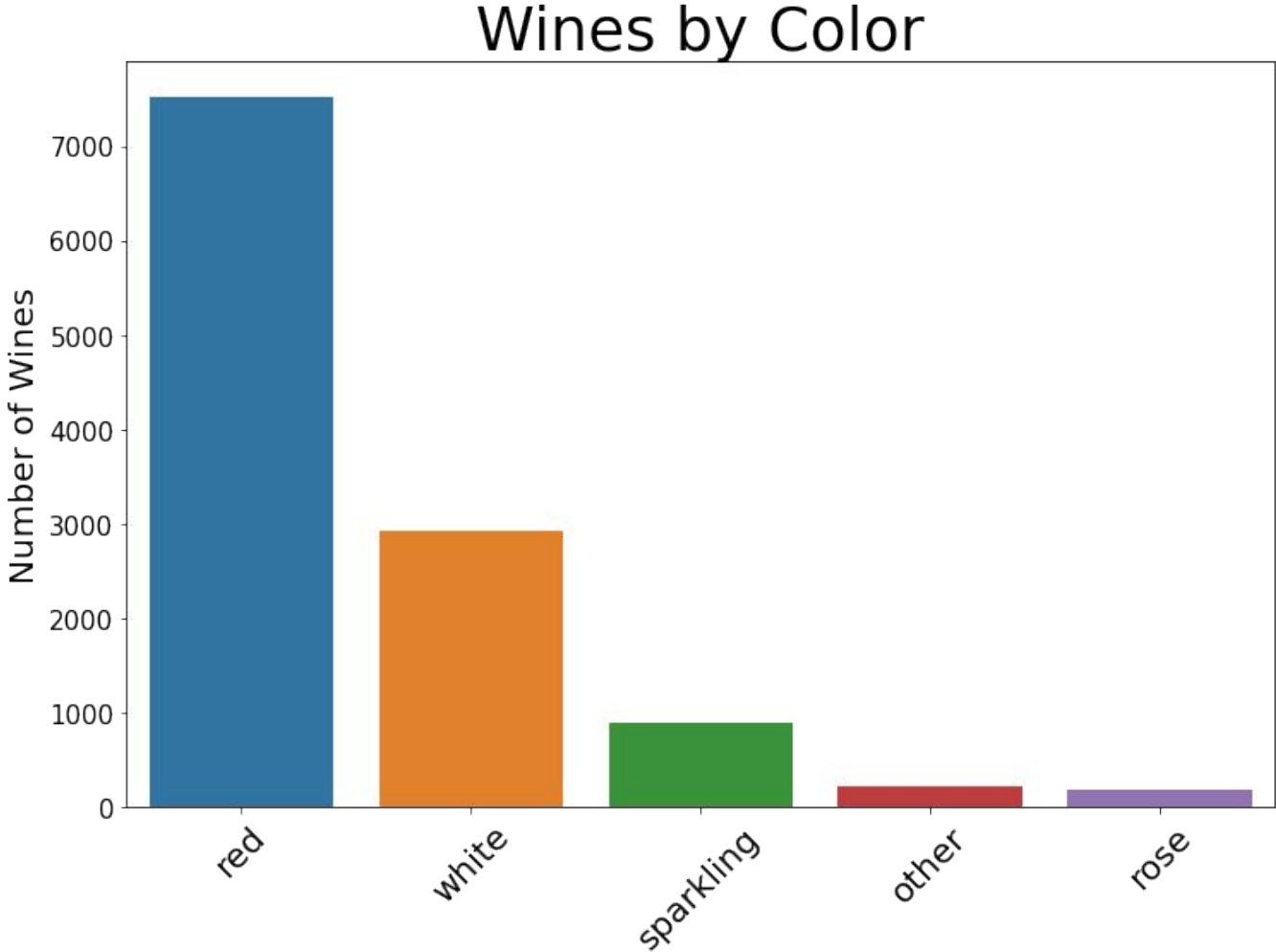
- ❖ *What does it taste like?*
- ❖ *Will I like it?*
- ❖ *Is it typical in style?*



Data Collection

	Company	Wines Collected	Pros / Cons
	Astor Wines	3182	Specialty and trendy
	Chambers Street	944	Descriptive, hard to pull data
	Twenty Twenty Wine	510	Primarily red, highly rated wines
	Wine.com	8273	More mainstream, largest collection

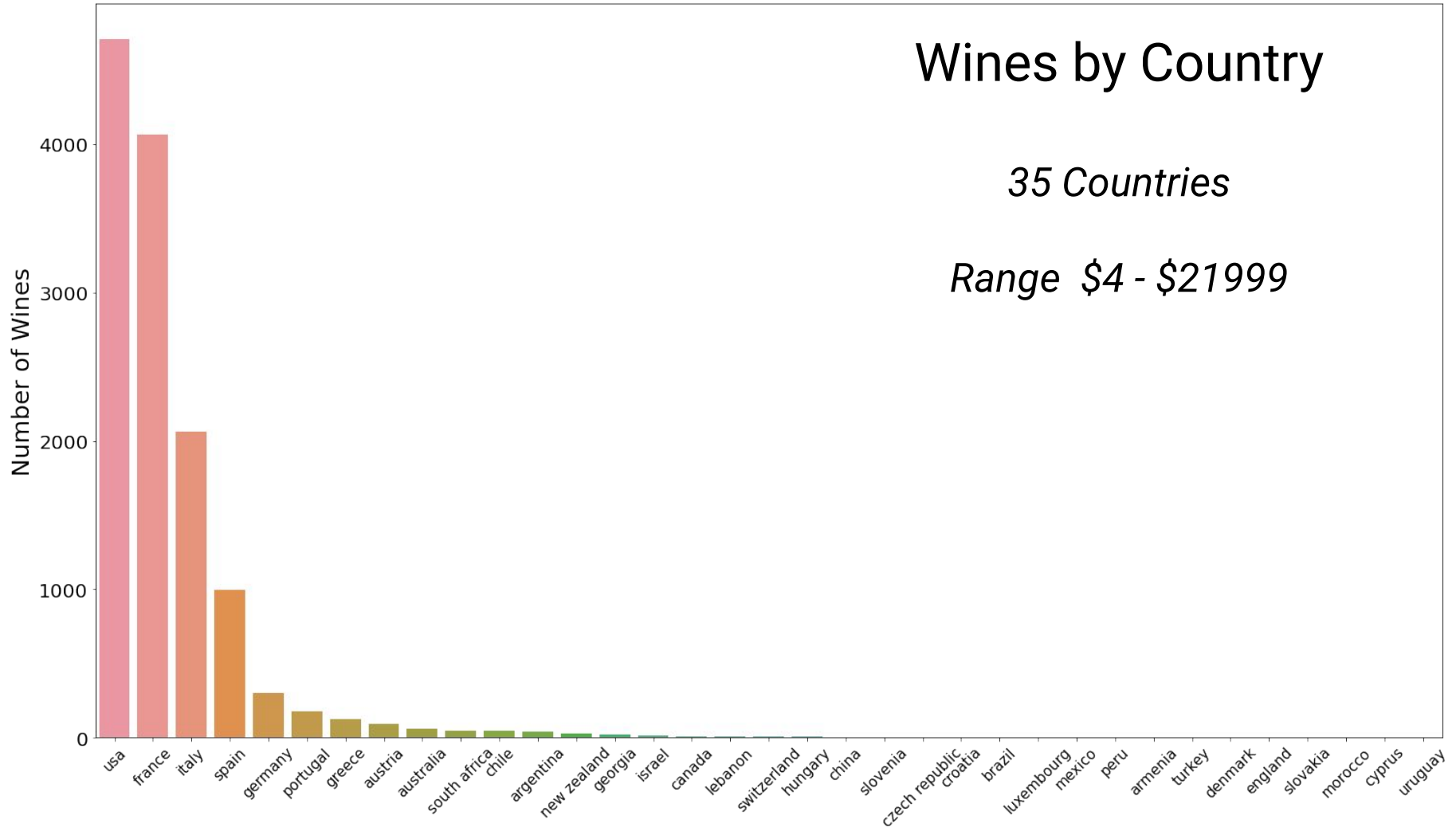
Red	7515
White	2930
Sparkling	903
Other	223
Rose	196
Total	11767



Wines by Country

35 Countries

Range \$4 - \$21999



*fast*Text

Library for efficient text classification and representation learning

- Gensim (Generate Similar) library
- Pretrained on 1 million words
- Character embeddings allows for it to find vectors even for made-up words
- Trained Skip-Grams on Region, Country, Color, Description

Mosel Wine Region



Conclusions / Further Research:

- Pull more wine (sparkling, rose, other)
- Categories for Organic, Natural, Orange
- Incorporate TFIDF recommender to pull similar wines to selected wines
- Pull data from local restaurants and wine shops into the model



Cheers!

Questions?