

Reviewed Dataset

642 voters

Wine Reviews

130k wine reviews with variety, location, winery, price, and description

zackthoult

last updated 9 months ago

Data (51 MB)

API

kaggle datasets download -d zynicide/wine-reviews

Download All

Data Sources

winemag-data-130k-... 130k x 14

winemag-data_first1... 151k x 11

winemag-data-130k-...

About this file

Here is a CSV version of the data I scraped. This dataset has three new fields

--Title (which you can parse the vintage from), Taster Name, and Taster Twitter Handle. This should also fix the duplicate entries problem in the first version of the dataset and add ~25k unique reviews to play with.

Columns

#

country The country that the wine is from

description A few sentences from a sommelier describing the wine's taste, smell, look, feel, etc.

designation The vineyard within the winery where the grapes that made the wine are from

winemag-data-130k-v2.csv (50.46 MB)

✂

📄

📁

📥

🔗

Name of the person who tasted and reviewed the wine

Michael Schachner 12%

Kerin O'Keefe 8%

Virginie Boone 7%

Other (15) 53%

taster_twitter_handle

String

@vossroger 20%

@wineschach 12%

@kerinokeefe 8%

@vboone 7%

Other (11) 53%

Missing 0 0%

Unique 19

Most Common Roger Voss 20%

Valid 98.8k 76%

Mismatched 31.2k 24%

Missing 0 0%





Unique 15

Most Common @vossroger 20%

https://www.kaggle.com/zynicide/wine-reviews

1/2

8/16/2018		Wine Reviews Kaggle			
title The title of the wine review, which often contains the vintage if you're interested in extracting that feature	A String	Gloria Ferrer NV Sonoma Brut Spa...	0%	Valid 130k 100% Mismatched 0 0% Missing 0 0% Unique 119k Most Common Gloria Fer 0%	
		Korbel NV Brut Sparkling (Californ...	0%		
		Segura Viudas NV Extra Dry Spark...	0%		
		Ruinart NV Brut Rosé (Champagne)	0%		
		Other (118836)	100%		
variety The type of grapes used to make the wine (ie Pinot Noir)	A String	Pinot Noir	10%	Valid 130k 100% Mismatched 1 0% Missing 0 0% Unique 707 Most Common Pinot Noir 10%	
		Chardonnay	9%		
		Cabernet Sauvignon	7%		
		Red Blend	7%		
		Other (703)	67%		
winery The winery that made the wine	A String	Wines & Winemakers	0%	Valid 130k 100% Mismatched 0 0% Missing 0 0% Unique 16.8k Most Common Wines & Win 0%	
		DFJ Vinhos	0%		
		Williams Selyem	0%		
		Other (16753)	99%		

Versions		
	Version 4	9 months ago Updated and re-scraped data with new fields.
	Version 3	9 months ago Added title (you can extract vintage), taster name, and taster twitter handle. The duplicates should also be removed.
	Version 2	a year ago A CSV version of the data
	Version 1	a year ago Initial release