



THE BUSINESS ISSUE

What can coffee roasters do to maximize the ratings and pricing potential of their coffees? Should they source their beans from a certain area? Or favor dark or light roasts?

Having had a decades-long love affair with coffee and even spending some of the best years of my life as a barista, I've always been paying attention to what makes a cup of coffee great. But what does the data have to say about it?

THE DATASETS

Thousands of coffees have been rated, described, and priced.

[This](#) dataset from Kaggle contains over 7000 coffee ratings, with separate ratings for aspects like aroma and aftertaste, but not much information on bean sourcing.

[This](#) dataset from Kaggle contains over 2000 coffee ratings, with more information on bean sourcing, but just one overall rating for each entry.

Both were web-scraped from coffeereview.com and I'd like to look at both and combine them where possible.