

## Problem Statement

**Predict the type of beer given the ABV & IBU.**

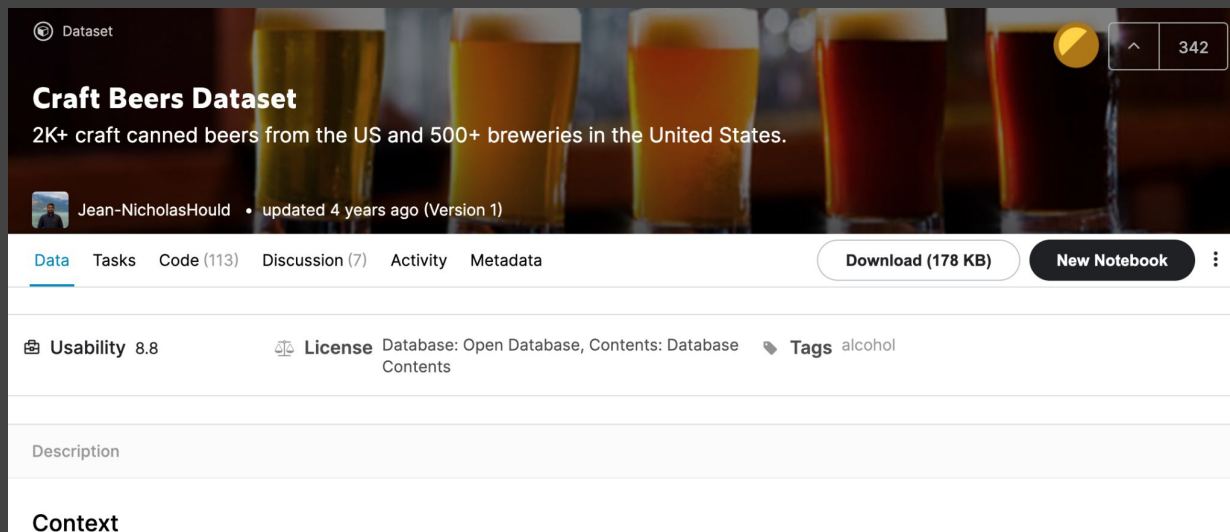
ABV: alcohol by volume

IBU: a gauge of beer's bitterness

# Dataset

## Kaggle Craft Beers Dataset

This dataset contains a list of 2,410 US craft beers and 510 US breweries.



The screenshot shows the Kaggle interface for the 'Craft Beers Dataset'. The header features a background image of several glasses of beer. The dataset title 'Craft Beers Dataset' is prominently displayed, followed by a description: '2K+ craft canned beers from the US and 500+ breweries in the United States.' Below this, the creator's name 'Jean-NicholasHould' and the update status 'updated 4 years ago (Version 1)' are shown. A navigation bar includes links for 'Data', 'Tasks', 'Code (113)', 'Discussion (7)', 'Activity', and 'Metadata'. On the right side of the navigation bar, there are buttons for 'Download (178 KB)' and 'New Notebook', along with a user icon and a count of '342'. Below the navigation bar, the 'Usability' score is 8.8, and the 'License' is 'Database: Open Database, Contents: Database Contents'. The 'Tags' section lists 'alcohol'. The 'Description' and 'Context' sections are currently empty.

Dataset

### Craft Beers Dataset

2K+ craft canned beers from the US and 500+ breweries in the United States.

Jean-NicholasHould • updated 4 years ago (Version 1)

[Data](#) [Tasks](#) [Code \(113\)](#) [Discussion \(7\)](#) [Activity](#) [Metadata](#) [Download \(178 KB\)](#) [New Notebook](#) 342

**Usability** 8.8 **License** Database: Open Database, Contents: Database Contents **Tags** alcohol

Description

Context

# Assumptions

"There are two main styles of beer—**lagers** and **ales**. All beers fall under one of these two categories. In other words, a pilsner is a lager, and porters and stouts are ales."

- Noble Hops Brewery



# Model

## KNN Classification

This classification problem involved predicting whether a beer was a lager or an ale. As a baseline, I used the majority class, which was ales.

## Metrics

Can the model improve on the majority class?

```
In [13]: y.value_counts(normalize=True)
```

```
Out[13]: 0.0    0.874021  
         1.0    0.125979  
         Name: type, dtype: float64
```