BEER STYLES, ALCOHOL CONTENT, AND CONSUMER PROFILES:

STRATEGIC INSIGHTS FOR THE BREWING INDUSTRY

RAÍSSA ANDRÉ FRANCO

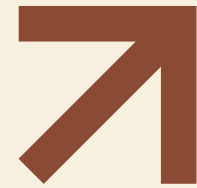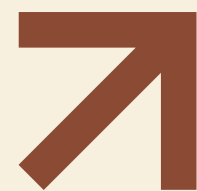# Project Overview

↗ **ANALYSIS OF BEER STYLES, ALCOHOL CONTENT, AND CONSUMER PROFILES**

↗ **COMPARISON BETWEEN MAINSTREAM AND PREMIUM BEER SEGMENTS**

↗ **STRATEGIC INSIGHTS FOR PRODUCT DEVELOPMENT AND MARKETING**

DATASET

Kaggle

Beer Reviews

	brewery_id	brewery_name	review_time	review_overall	review_aroma	review_appearance
0	10325	Vecchio Birraio	1234817823	1.5	2.0	2.5
1	10325	Vecchio Birraio	1235915097	3.0	2.5	3.0
2	10325	Vecchio Birraio	1235916604	3.0	2.5	3.0
3	10325	Vecchio Birraio	1234725145	3.0	3.0	3.5
4	1075	Caldera Brewing Company	1293735206	4.0	4.5	4.0

| review_profilename	beer_style	review_palate	review_taste	beer_name	beer_abv	beer_beerid
stcules	Hefeweizen	1.5	1.5	Sausa Weizen	5.0	47986
stcules	English Strong Ale	3.0	3.0	Red Moon	6.2	48213
stcules	Foreign / Export Stout	3.0	3.0	Black Horse Black Beer	6.5	48215
stcules	German Pilsener	2.5	3.0	Sausa Pils	5.0	47969
johnmichaelsen	American Double / Imperial IPA	4.0	4.5	Cauldron DIPA	7.7	64883

Quality Analysis
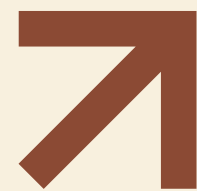
DATA TYPES

DUPLICATES AND UNIQUE VALUES
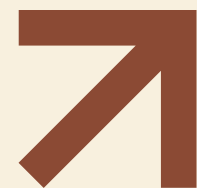
MISSING VALUES

ABV

Alcohol by Volume

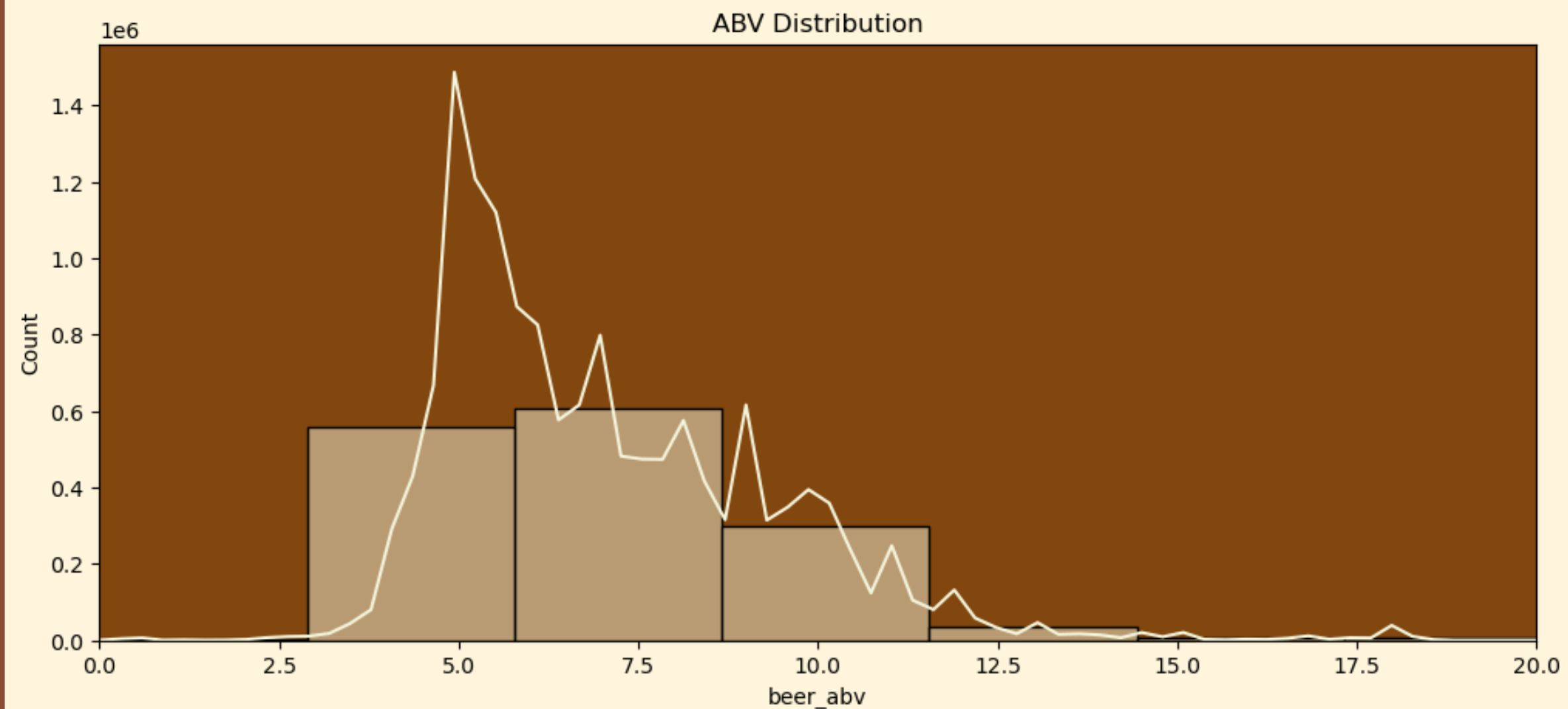↗ **Beer Styles**

↗ **Mean**

↗ **Mode**

ABV

AVERAGE ABV: APPROXIMATELY 7.04%

INTERQUARTILE RANGE (IQR): BETWEEN 5.2% AND 8.5%

MAXIMUM VALUE: UP TO 57.7%, INDICATING THE PRESENCE OF OUTLIERS

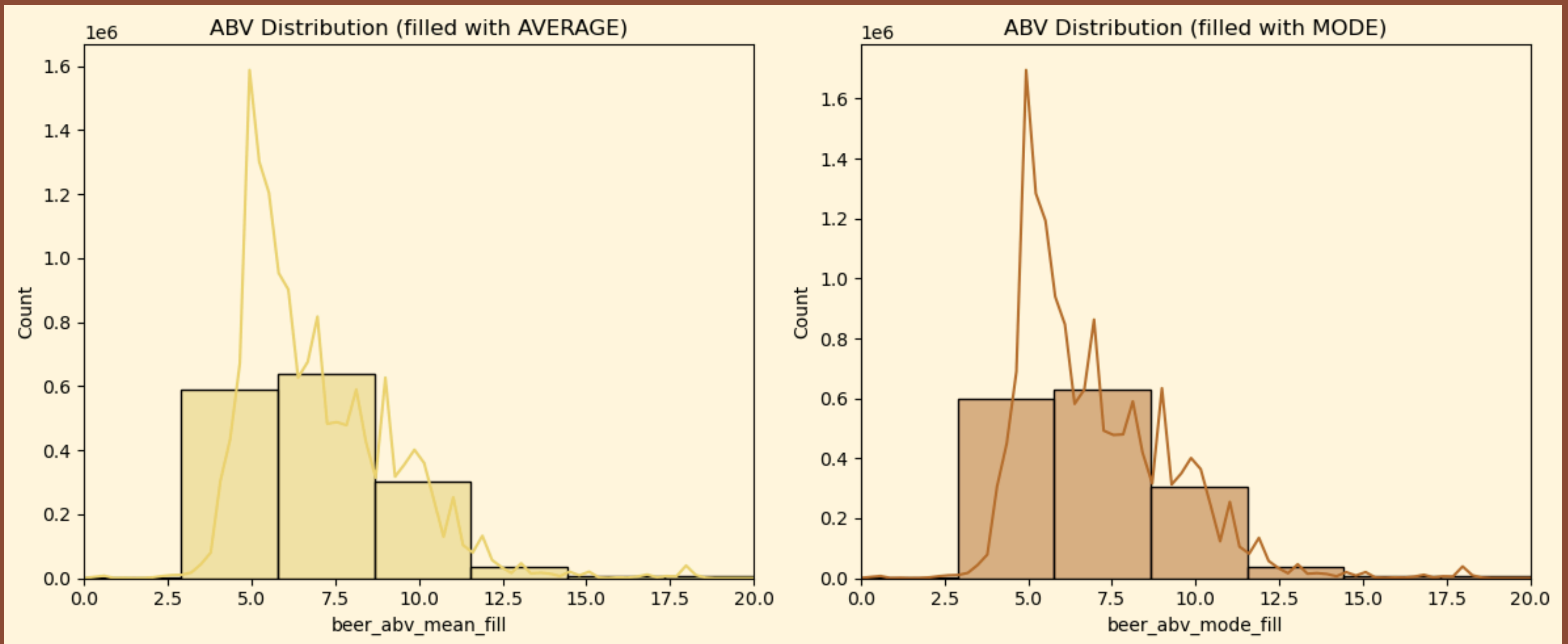```
ABV:
count   1.518829e+06
mean   7.042387e+00
std   2.322526e+00
min   1.000000e-02
25%   5.200000e+00
50%   6.500000e+00
75%   8.500000e+00
max   5.770000e+01
Name: beer_abv, dtype: float64
```
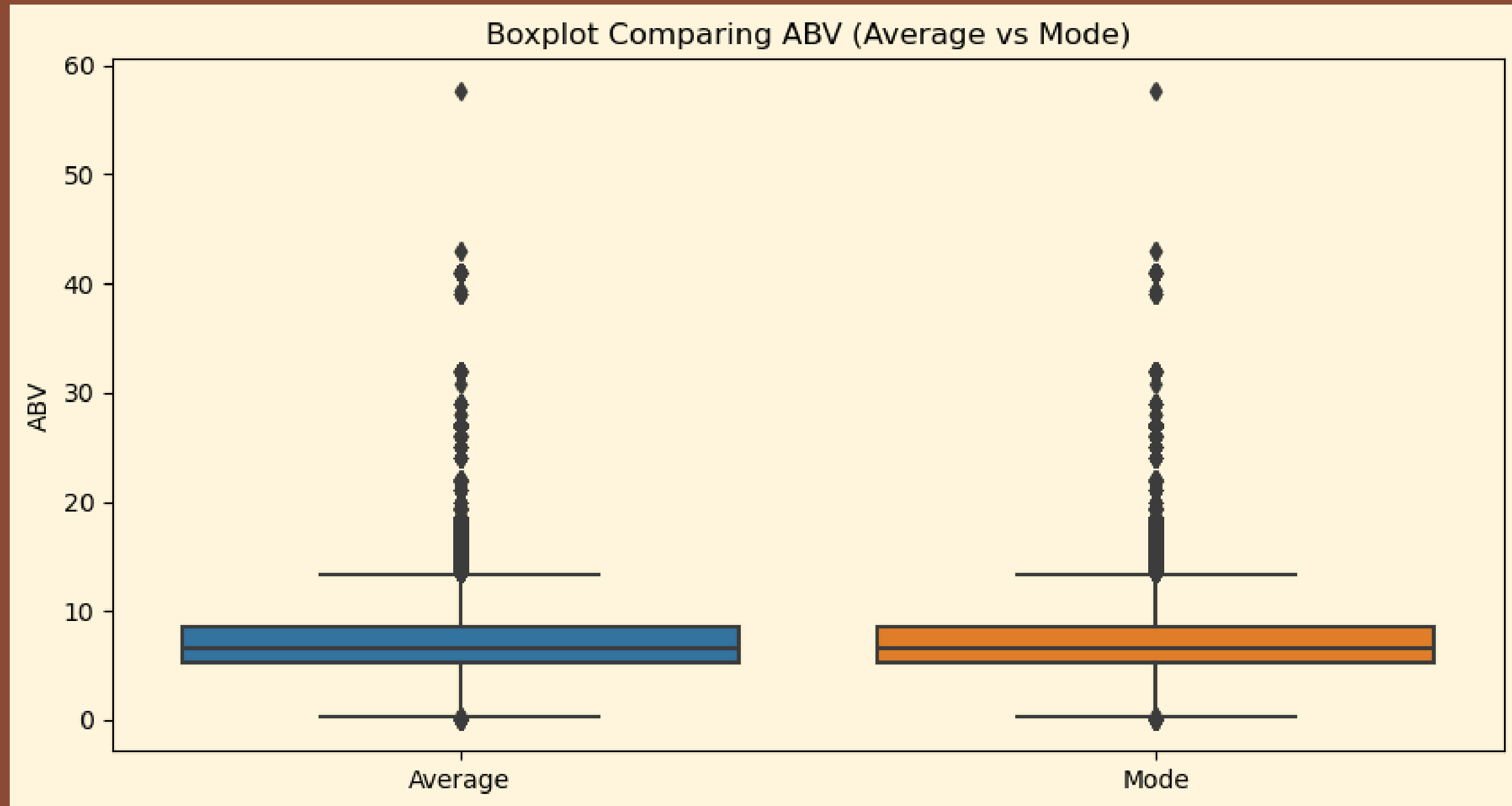


ABV Distribution

# ABV

## Mean

## Mode

# ABV

## Mean

## Mode



Boxplot Comparing ABV (Average vs Mode)

# Insights

```
ABV filled with mean:
count    1.586614e+06
mean     7.001855e+00
std      2.297935e+00
min      1.000000e-02
25%      5.200000e+00
50%      6.50000e+00
75%      8.500000e+00
max      5.770000e+01
Name: beer_abv_mean_fill, dtype: float64

ABV filled with mode:
count    1.586614e+06
mean     6.992876e+00
std      2.305957e+00
min      1.000000e-02
25%      5.200000e+00
50%      6.500000e+00
75%      8.500000e+00
max      5.770000e+01
Name: beer_abv_mode_fill, dtype: float64
```

↗ BOTH IMPUTATION METHODS SHOWED MINIMAL EFFECT ON THE DATA DISTRIBUTION BECAUSE OF THE LARGE SAMPLE SIZE

↗ MEAN IMPUTATION BETTER PRESERVED KEY STATISTICS LIKE THE MEAN AND STANDARD DEVIATION.

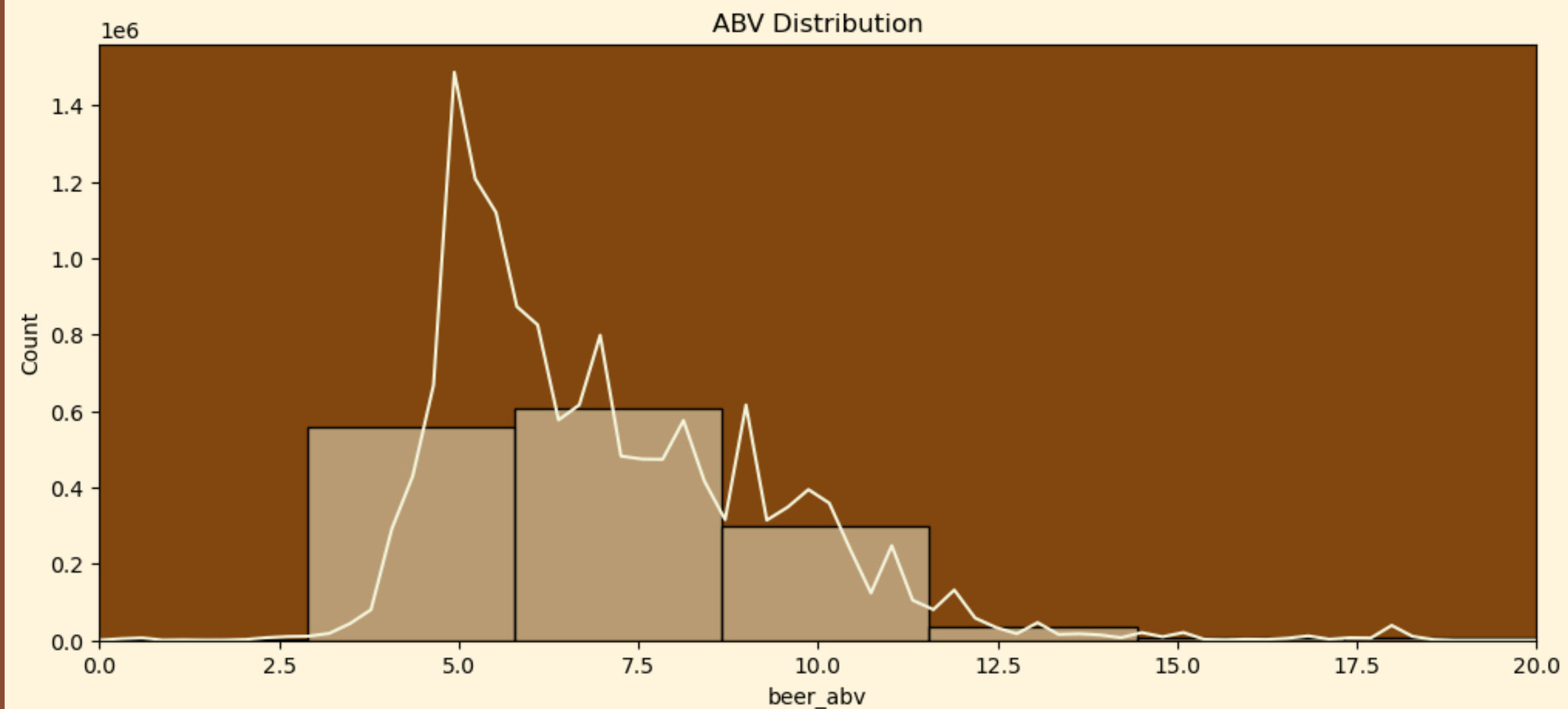↗ FOR MAINTAINING THE ORIGINAL DISTRIBUTION'S INTEGRITY, MEAN IMPUTATION IS THE PREFERRED METHOD.

ABV

```
ABV:
count       1.518829e+06
mean        7.042387e+00
std         2.322526e+00
min         1.000000e-02
25%         5.200000e+00
50%         6.500000e+00
75%         8.500000e+00
max         5.770000e+01
Name: beer_abv, dtype: float64
```

AVERAGE ABV: APPROXIMATELY 7.04%

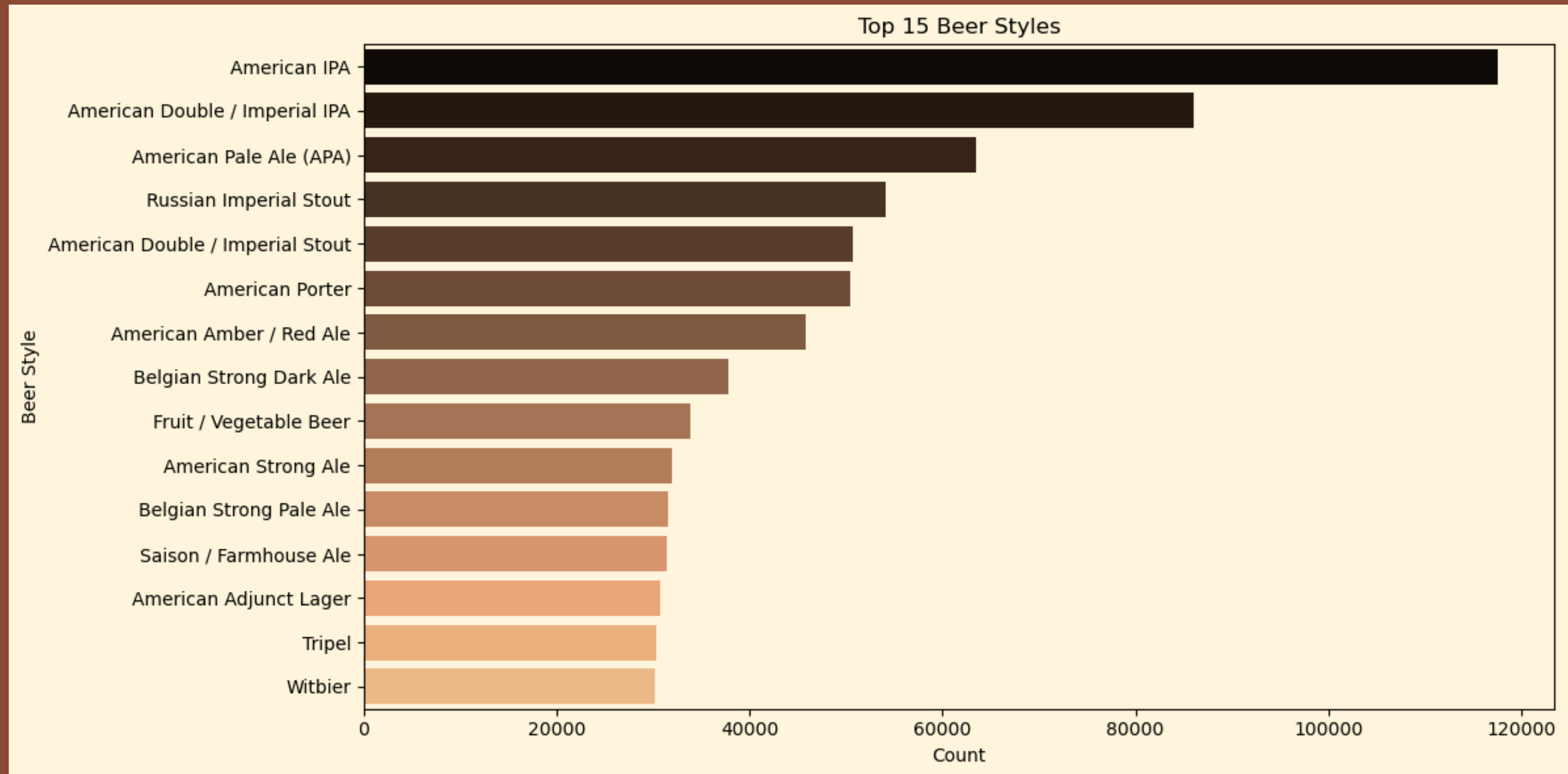**INTERQUARTILE RANGE (IQR): BETWEEN 5.2% AND 8.5%**

MAXIMUM VALUE: UP TO 57.7%, INDICATING THE PRESENCE OF OUTLIERS

ABV Distribution

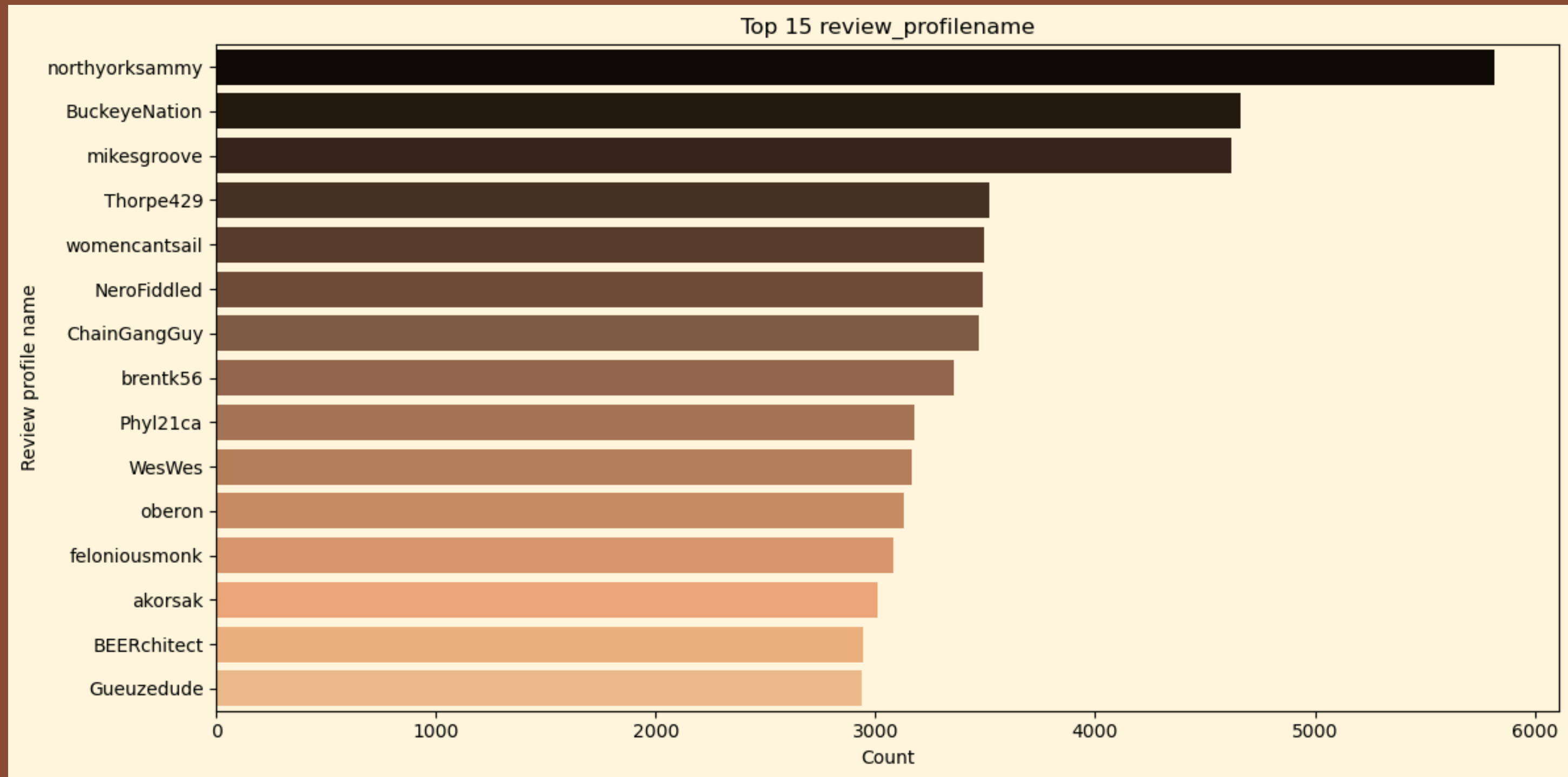# Beer Styles

Top 15 Beer Styles

Reviewers

Top 15 review_profilename

Insights

↗ The most reviewed beer styles likely reflect the most consumed styles.

↗ The most active reviewers are likely the users who consume beer most frequently.

Style ❌ ABV

ABV Distribution by Beer Style (Top 15 Most Frequent Styles)

Alcohol by Volume (ABV) vs Beer Style box plot. X-axis categories: American IPA, American Double / Imperial IPA, American Pale Ale (APA), Russian Imperial Stout, American Double / Imperial Stout, American Porter, American Amber / Red Ale, Belgian Strong Dark Ale, Fruit / Vegetable Beer, American Strong Ale, Belgian Strong Pale Ale, Saison / Farmhouse Ale, American Adjunct Lager, Tripel, Witbier.

STD

```
beer_style
American IPA                          0.566018
American Double / Imperial IPA        1.888820
American Pale Ale (APA)               0.565016
Russian Imperial Stout               1.584972
American Double / Imperial Stout     2.769471
American Porter                      1.305109
American Amber / Red Ale             1.142649
Belgian Strong Dark Ale              1.809967
Fruit / Vegetable Beer               2.169462
American Strong Ale                  2.907647
Belgian Strong Pale Ale              1.323124
Saison / Farmhouse Ale               1.240895
American Adjunct Lager               0.410008
Tripel                               0.894779
Witbier                              1.252107
Name: beer_abv_mean_fill, dtype: float64
```


std < 0.5: low → the style has a very consistent ABV

0.5 < std < 1.0: medium → natural variations between brands

std > 1.0: high → the style includes beers with quite different ABVs

std > 2.0: very high → the style might be grouping very different kinds of beers

Insights

↗ Strong styles vary a lot because they include everything from "standard" versions to highly concentrated and alcoholic ones.
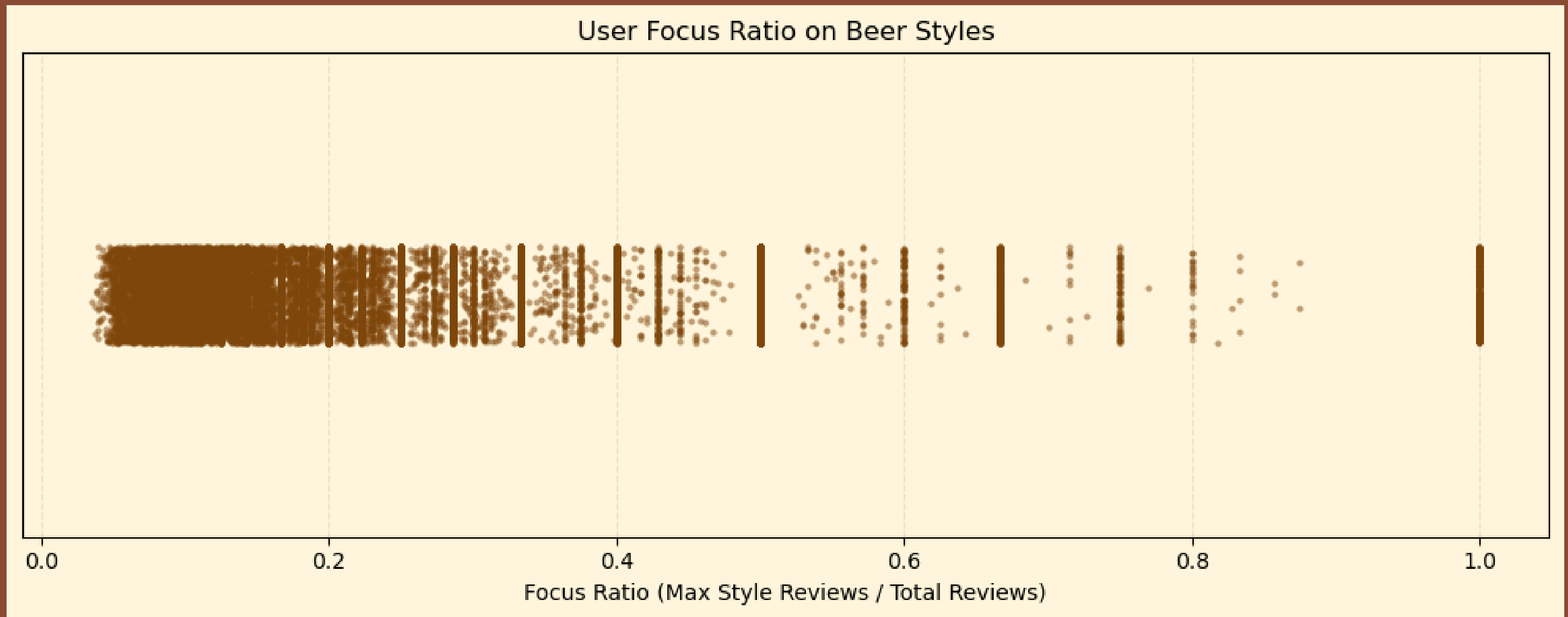
↗ Beers like American IPA and Pale Ale are very consistent in ABV

↗ Which styles to standardize for the mass market and which to explore creatively for niche or premium segments.

Style ❌ Reviewers

User Focus Ratio on Beer Styles

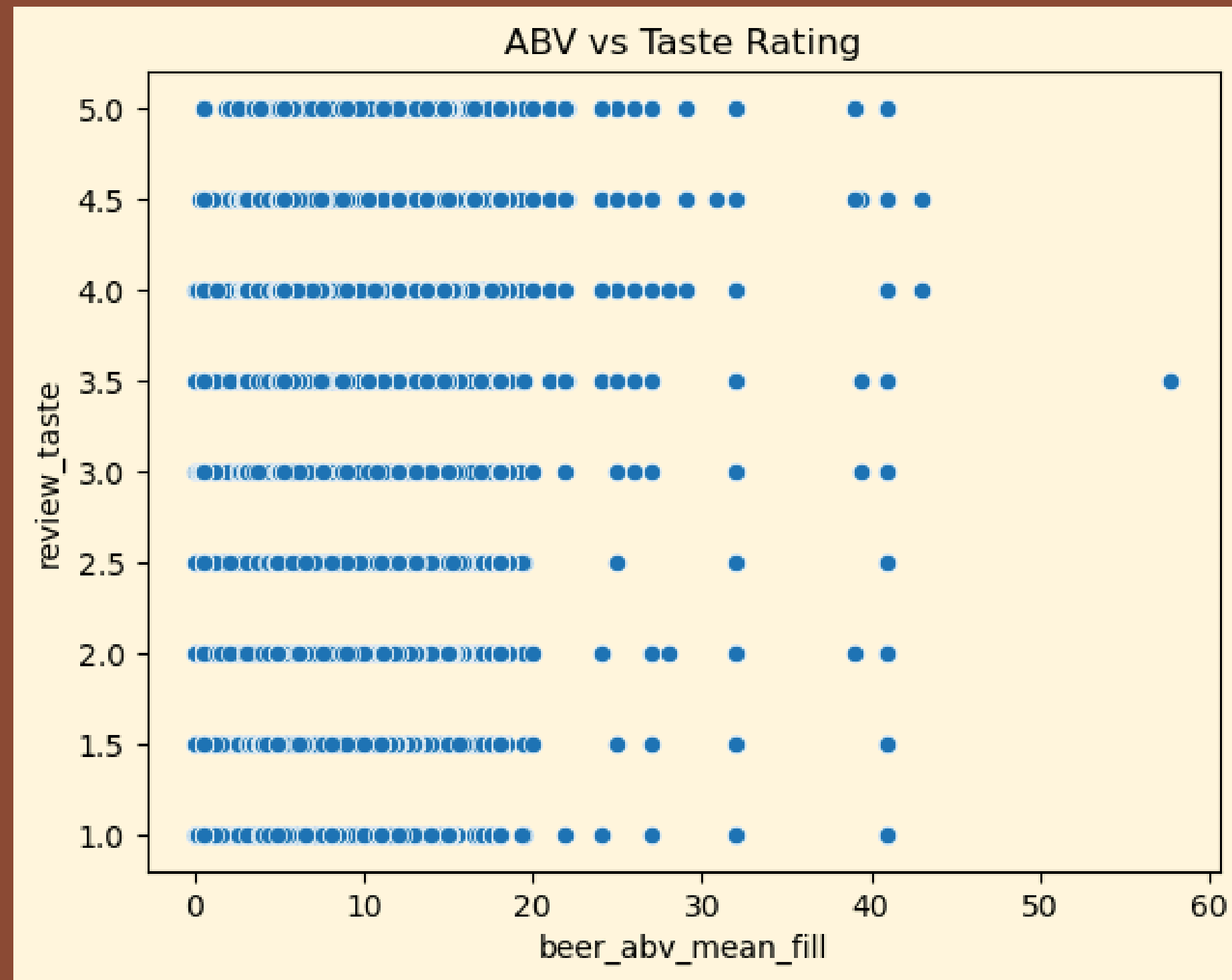Focus Ratio (Max Style Reviews / Total Reviews)

Insights

User focus ratios range from near 0 to 1, showing both style specialists and generalists.

Most users have low focus ratios (below 0.2), suggesting broad and diverse beer preferences.

Users with low focus, like yelterdow and jdhilt, are more open to varied recommendations, while high-focus users may prefer niche suggestions.

```
review_profilename
yelterdow          0.035330
jdhilt             0.037159
elricorico         0.037383
DaPeculierDane     0.038501
clvand0            0.038552
SilentSabre        0.039074
hyuga              0.039216
ADR                0.039660
jar2574            0.040000
TheDM              0.040107
Name: count, dtype: float64
```

1.0 → the user only reviewed one beer style (or almost exclusively one).

~0.6 → around 60% of the user's reviews are for the same style.

<0.3 → the user has a diverse set of reviews across multiple styles.

Taste review ❌ ABV

ABV vs Taste Rating

Pearson and Spearman

```
Pearson correlation: 0.2923
Spearman correlation: 0.3475
```

There is a weak but positive relationship between ABV and taste ratings.

Spearman's correlation (0.3475) is slightly stronger than Pearson's (0.2923), reflecting a better fit for the non-linear pattern.

Higher ABV tends to correlate with higher taste ratings, but the effect is moderate and not strictly linear.

Insights

The regression line shows a moderate positive association between ABV and taste rating.

Taste ratings vary widely at all ABV levels, suggesting other factors also influence perception.

Prediction uncertainty increases at higher ABV values, as shown by the widening confidence interval.

Relationship Between ABV and Taste Rating (Sample)

Final observations

↗ SPECIALIST AND GENERALIST CONSUMERS

↗ MAINSTREAM AND PREMIUM

↗ MARKETING STRATEGIES

↗ IMPROVE PRODUCTS AND REACH MORE CUSTOMERS

THANK YOU

QUESTIONS

RAÍSSA ANDRÉ FRANCO