

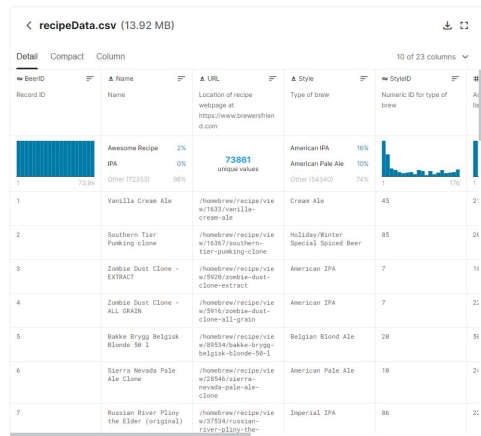
# Brewer's Friend Beer Recipes

Daniel Suárez 1566103

**Github:** [https://github.com/Danny-8/APC\\_Kaggle](https://github.com/Danny-8/APC_Kaggle)  
**Dataset:** <https://www.kaggle.com/jtrofe/beer-recipes>

# Introducció

## Dataset:



< styleData.csv (3.84 kB)

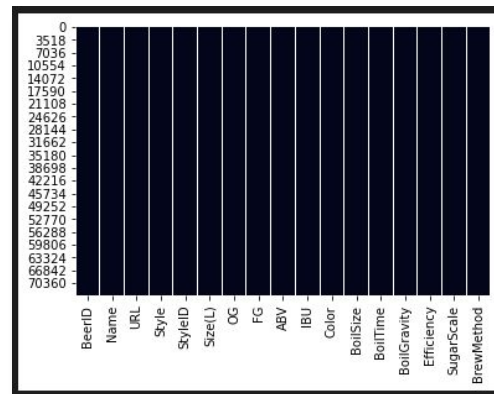
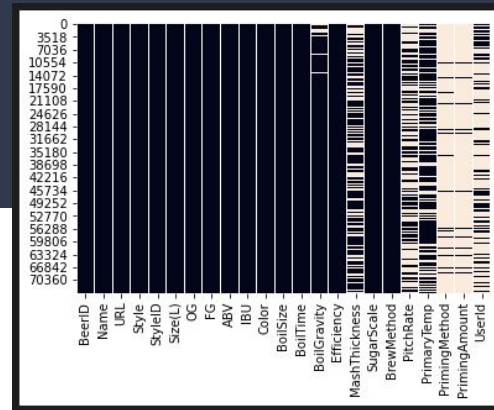
Detail Compact Column

Style

176 unique values

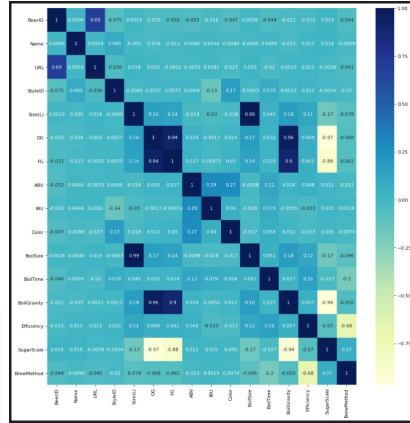
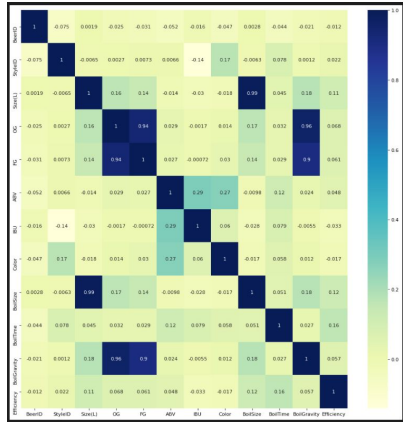
Style	StyleID
Altbier	1
Alternative Grain Beer	2
Alternative Sugar Beer	3
American Amber Ale	4
American Barleywine	5
American Brown Ale	6
American IPA	7
American Lager	8
American Light Lager	9
American Pale Ale	10
American Porter	11
American Stout	12
American Strong Ale	13
American Wheat Beer	14
American Wheat or Rye Beer	15
Ann's Wine	16

## Nulls:

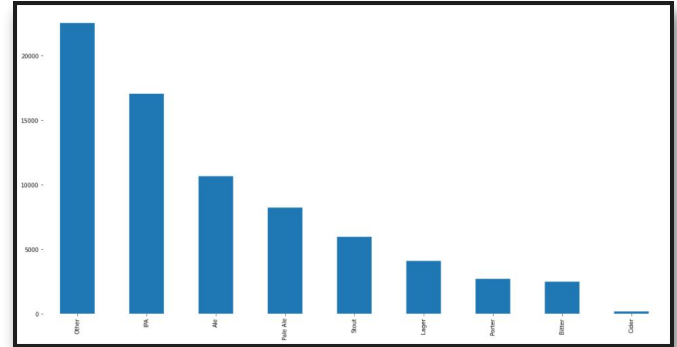
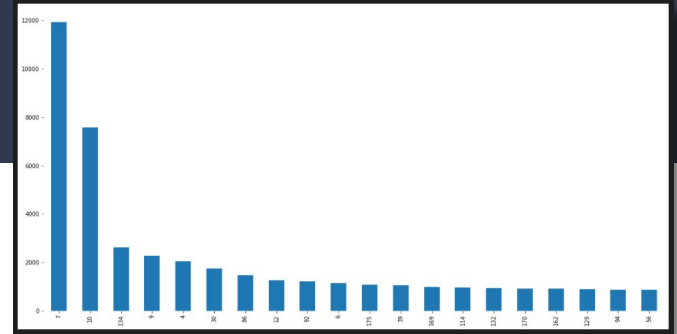


# Preprocessat

Correlació:



Classificació:



['Other', 'Ale', 'IPA', 'Pale Ale', 'Lager', 'Stout', 'Bitter', 'Cider', 'Porter']

# Mètodes Utilitzats

```
X = recipe_dataset.drop(['StyleID', 'GenericStyleID'], axis=1)
Y_s = recipe_dataset['StyleID']
Y_gs = recipe_dataset['GenericStyleID']
```

GridSearch:

```
hyp_c = [0.01, 0.1, 1.0, 10]
hyp_tol = [0.0001, 1, 1000]
hyp_criterion = ['gini', 'entropy']
hyp_n_estimators = [100, 200, 300]
hyp_min_samples_leaf = [1, 2, 3]
```

Regresió Logística

Random Forest

Decision Tree

Ada Boost

Ex:

```
Decision Tree
F1 score Style:      0.1123056073685654
F1 score GenericStyle: 0.3716018062432864
Criterion: gini

F1 score Style:      0.09895491662789191
F1 score GenericStyle: 0.39268361931125084
Criterion: entropy
```

# Resultats i Exemples

Millor model:

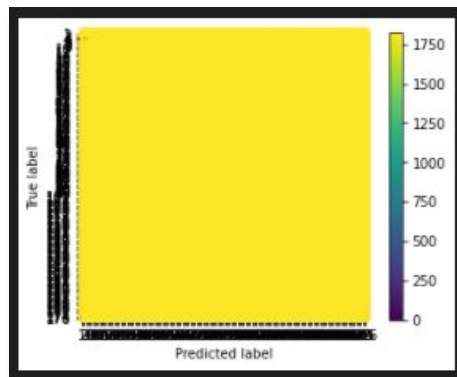
	model	mean	std
1	Random Forest	0.157530	0.006265
2	Decision Tree	0.108205	0.002877
0	Regresió Logística	0.059266	0.003953
3	Ada Boost	0.008849	0.001686

Style

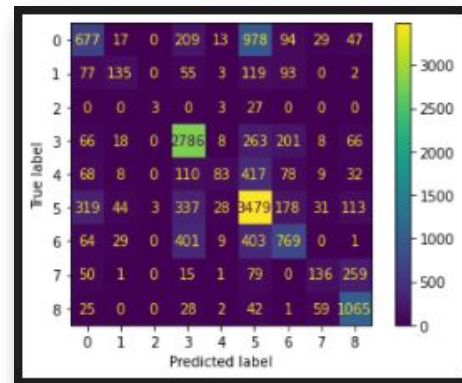
	model	mean	std
1	Random Forest	0.458610	0.019237
2	Decision Tree	0.381387	0.014562
3	Ada Boost	0.361104	0.020857
0	Regresió Logística	0.271721	0.015013

GenericStyle

Confusion Matrix:



Style



GenericStyle

Classification Report:

accuracy			0.39	14773
macro avg	0.27	0.15	0.17	14773
weighted avg	0.38	0.39	0.35	14773

Style

accuracy			0.62	14773
macro avg	0.57	0.44	0.46	14773
weighted avg	0.60	0.62	0.59	14773

GenericStyle

# Conclusió i Treball a futur

