



SDAIA
الهيئة السعودية للبيانات
والذكاء الاصطناعي
Saudi Data & AI Authority

Coffee Bean Classification

Presented by

Nasser Alqahtani
Mohammed Alzabyedi

Submit to: Dr. Mejdal Alqahtani

Outline

We are here!



01 **Introduction**
Abstract a project

02 **Data**
Explore the data

03 **Algorithms**
Explain the algorithms

04 **Tools**
Show a tools

05 **EDA**
Visualization the data.

06 **Models**
Display the models

INTRODUCTION

171

cafe per million people in
KSA, 2018

1032

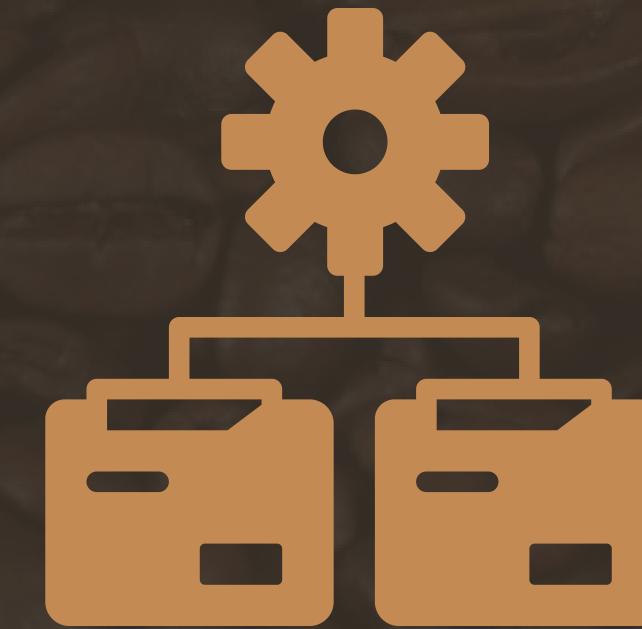
cafe per million people in
KSA, 2030



DATA



2560
crops



44
feature

Specialty coffee

commercial coffee

Algorithms

Preparing the
data.

Feature
Selection

Methods
LogisticRegression, KNN, Decision Tree,
Random Forest, Bagging, and Voting.

Presented by
Rosa Maria Aguado



Tools

- Python and Jupyter Notebook.
- Numpy and Pandas for data manipulation.
- Matplotlib and Seaborn for plotting visualization.
- Sklearn for ML algorithms.
- Imblearn to solve imbalance problem.



EDA

Cupping
coffee

Aroma

Flavor

Aftertaste

Acidity

Body

Balance

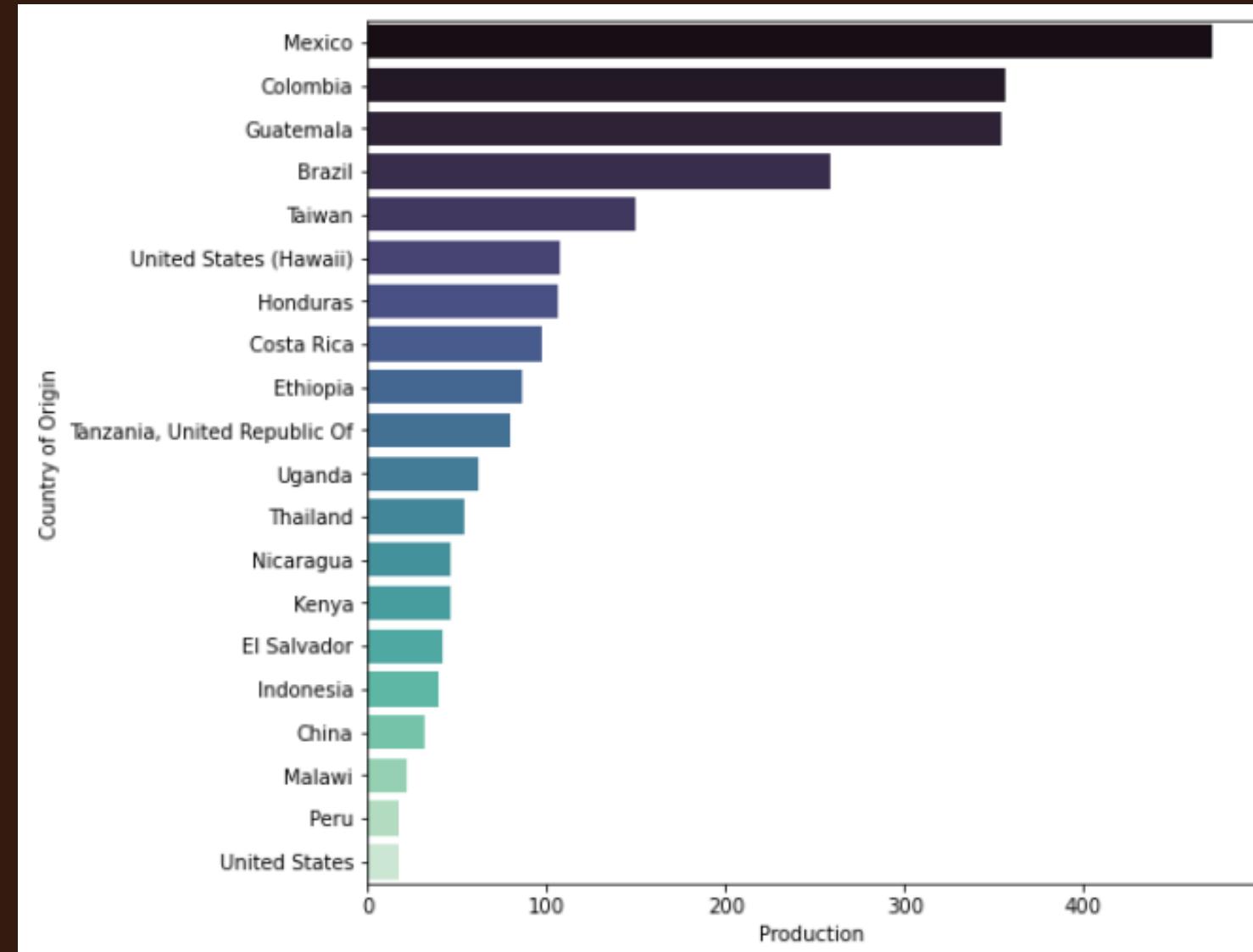
Uniformity

Sweetness

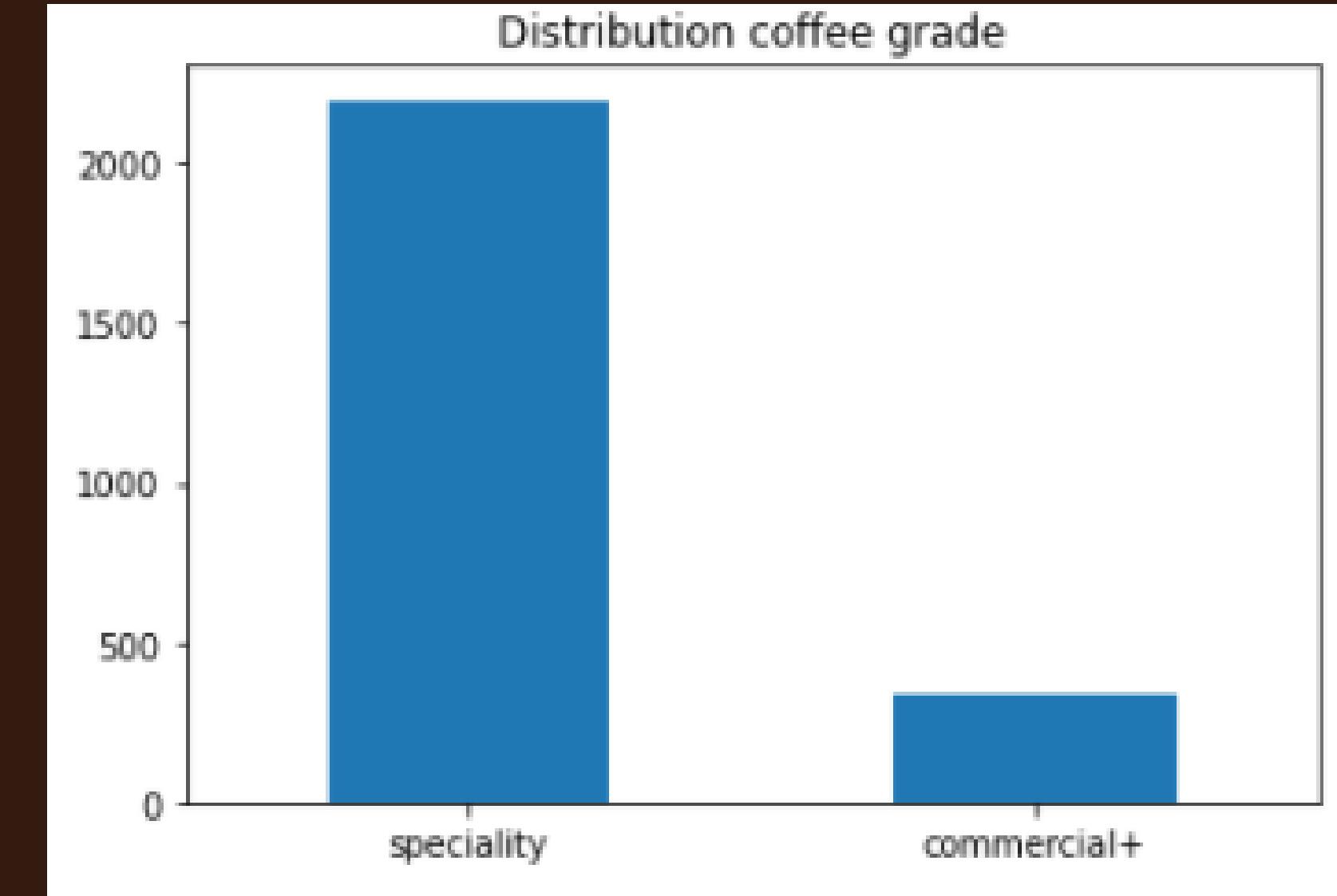
Cleanliness

Cupper
Points

EDA

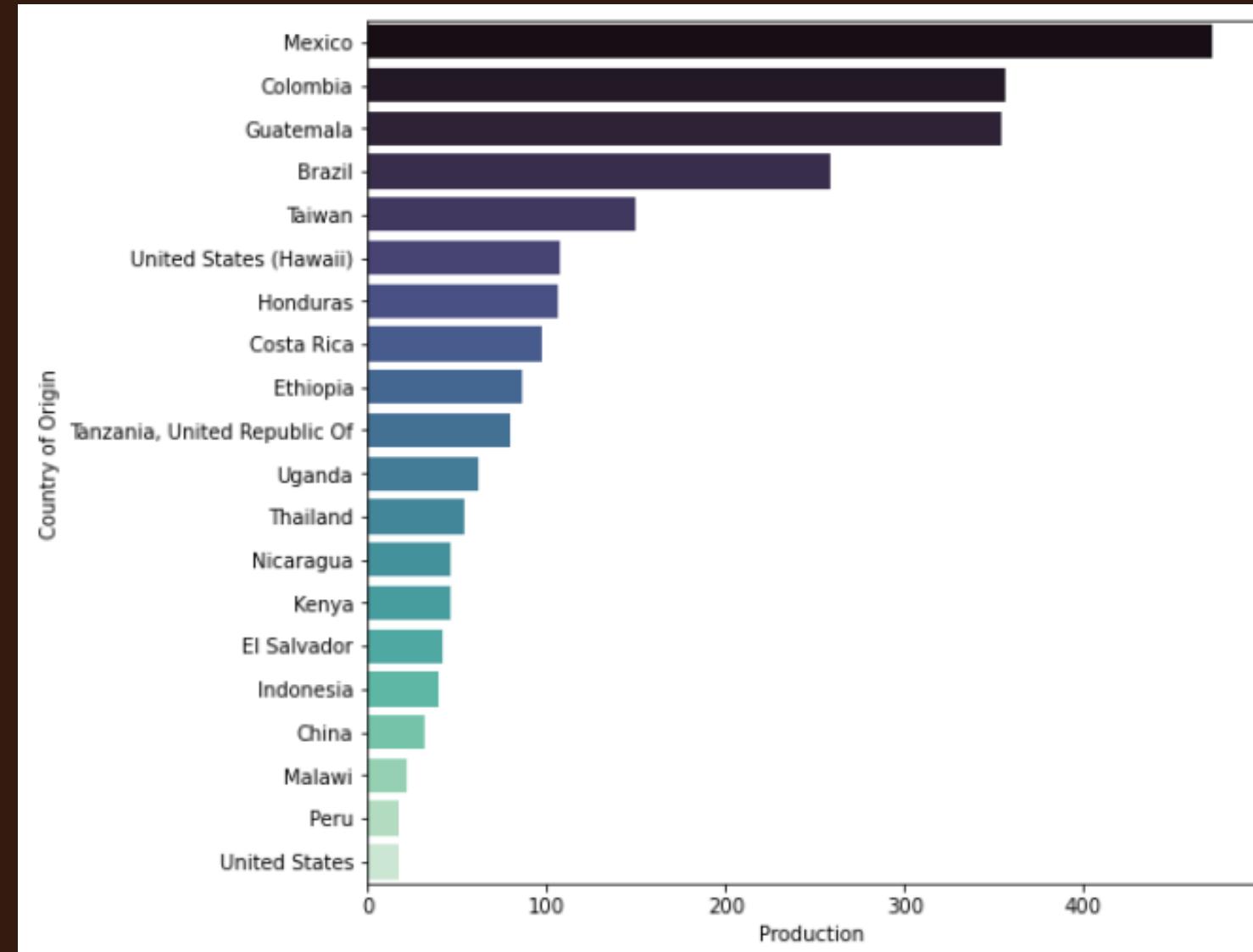


Distribution of coffee to
countries

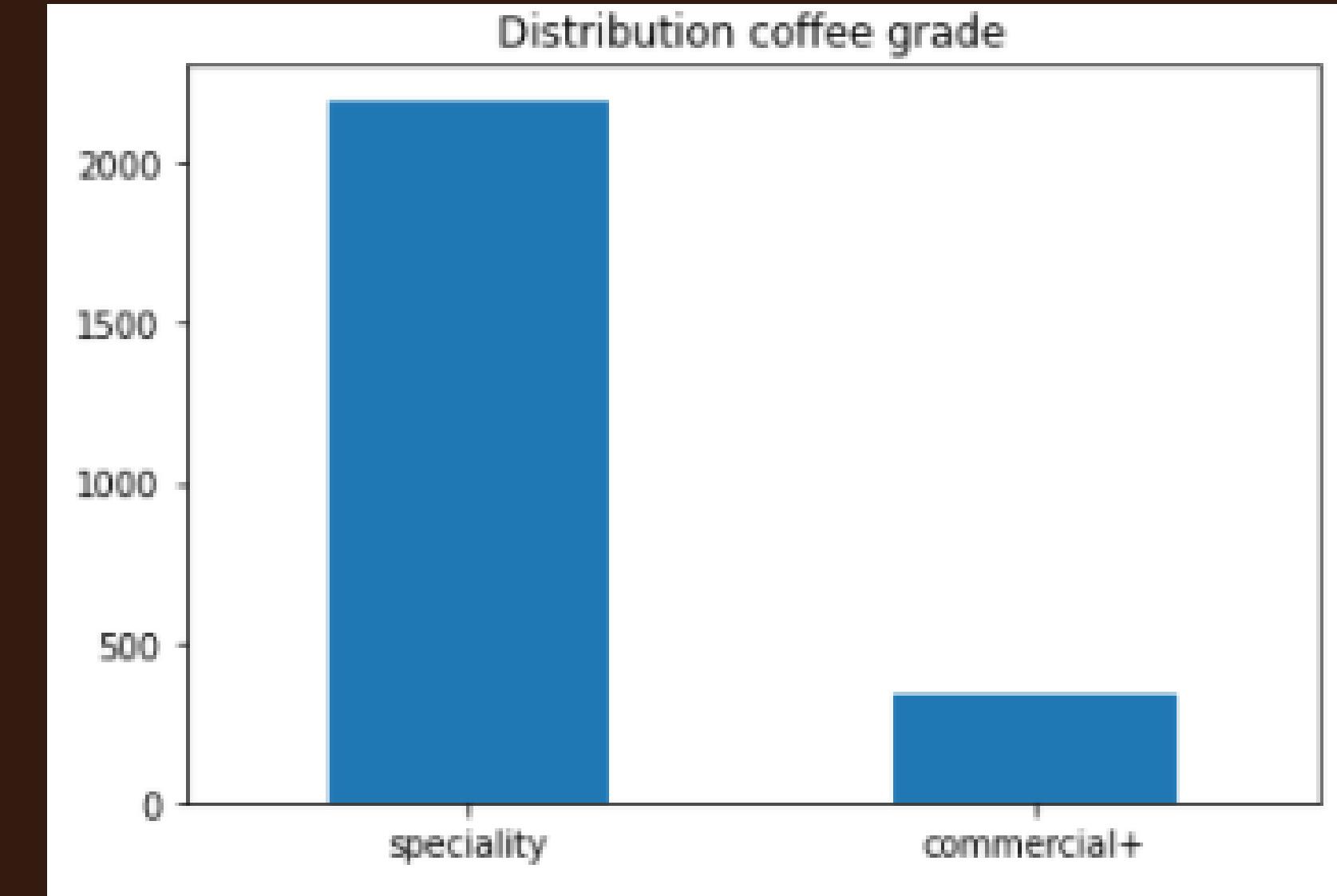


Distribution of coffee to the
grades

EDA



Distribution of coffee to
countries



85.8%

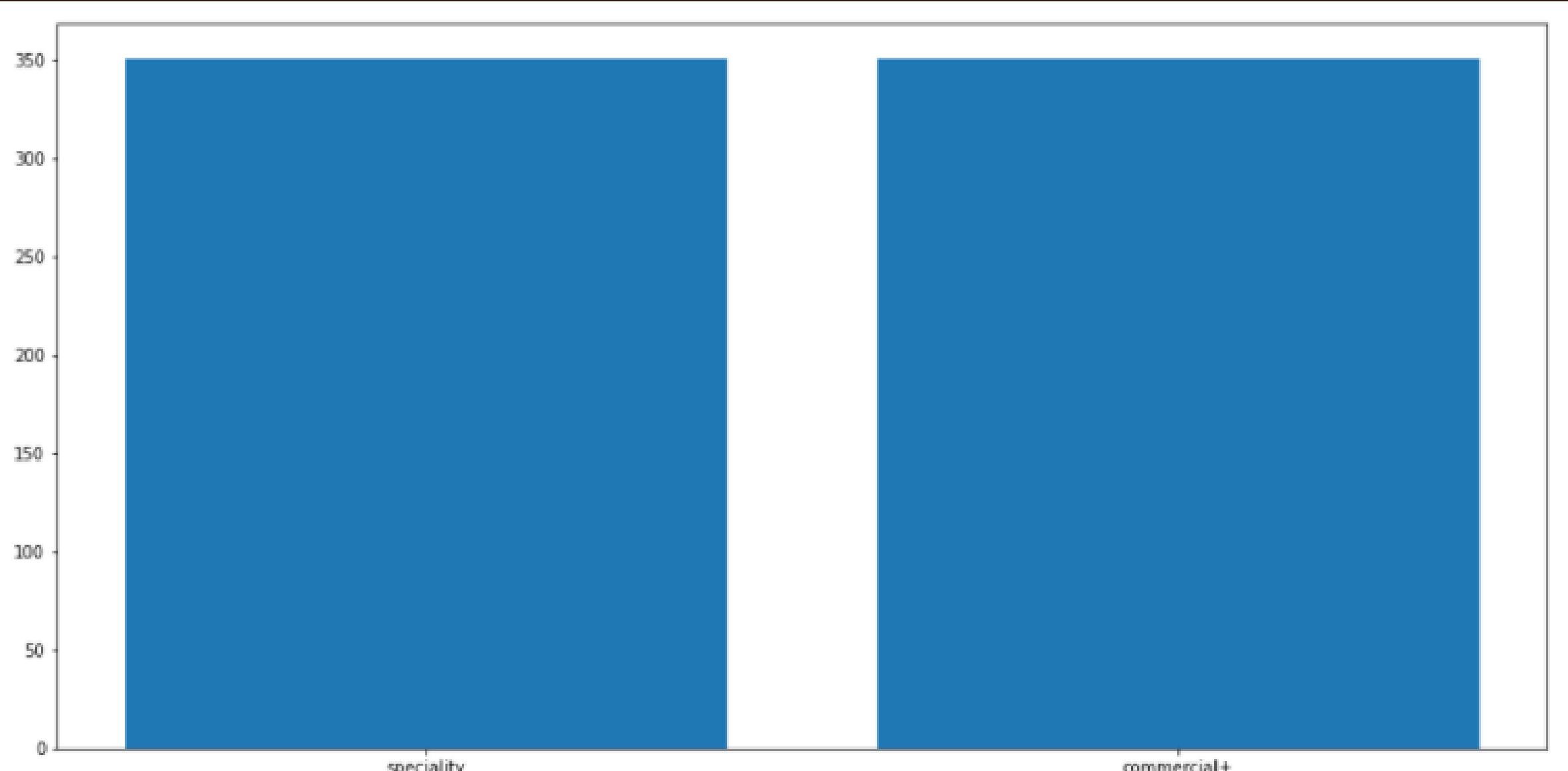
Distribution of coffee to the
grades

14.2%

EDA



Apply SMOTE to
balance the classes.



50%

50%

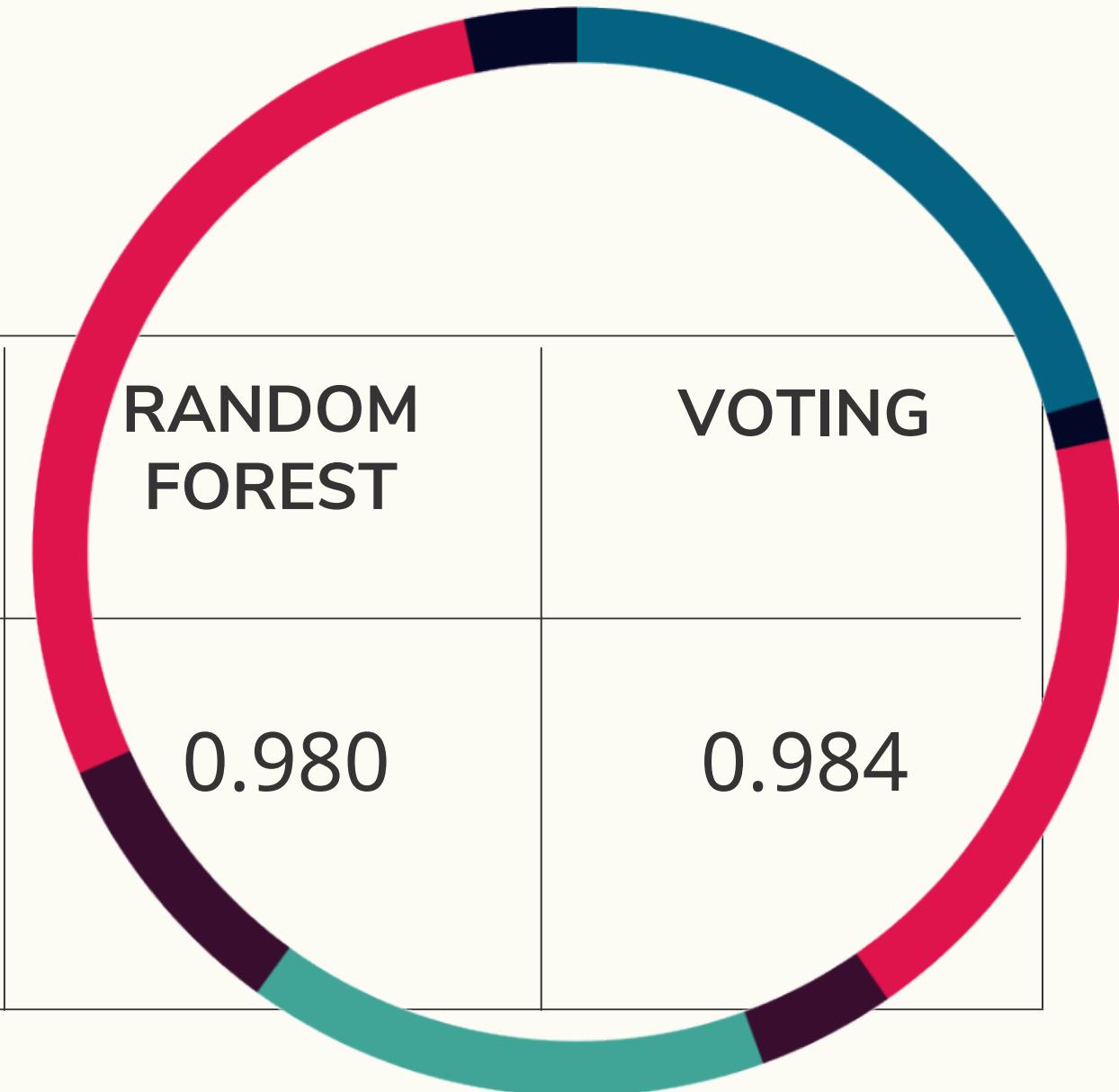
Models

NAME OF MODEL	LOGISTIC REGRESSION	KNN	DECISION TREE	RANDOM FOREST	VOTING
Recall	0.972	0.977	0.975	0.980	0.984



Models

NAME OF MODEL	LOGISTIC REGRESSION	KNN	DECISION TREE	Bagging	RANDOM FOREST	VOTING
Recall	0.972	0.977	0.975	0.972	0.980	0.984





THANK YOU

سولي الكيف وارهولي من الدله
البن الاشقر يداوي الراس فنجاله
كيف نـ لـ نـ حـ رـ قـهـ بـ الـ تـ اـرـ وـ نـ زـ لـهـ
ولـا انـ قـطـعـ لوـ وـ دـ رـىـ صـنـعـاءـ عـنـيـناـ لـهـ