Food Loss and Waste

Presented by:

Lisa Liang
Juliana Sahagun
Triston Crossland

Background

Around 1/3 of all food produced for human consumption each year - roughly 1.3 billion tons - is lost or wasted.*

At a time when nearly 10% of the global population, roughly 800 million people, are suffering from undernourishment or chronic food deprivation, governments all over the world are looking for solutions.

***Credit: United Nations**

Problem Statement

We are a team of data scientists contracted by the United Nations to develop models related to the UN Agenda 2030 Sustainable Development Goal (SDG) 12.3, which states: “by 2030, halve the per capita food waste at the retail and consumer level, and reduce food losses along production and supply chains including post-harvest losses.”

To meet our obligations, we must fulfill two primary objectives:

1) Develop a model that identifies the stages of food supply where loss occurs.
2) Develop a model that predicts the percentage of food wasted in different countries.

Methodology Overview

Dataset Description:

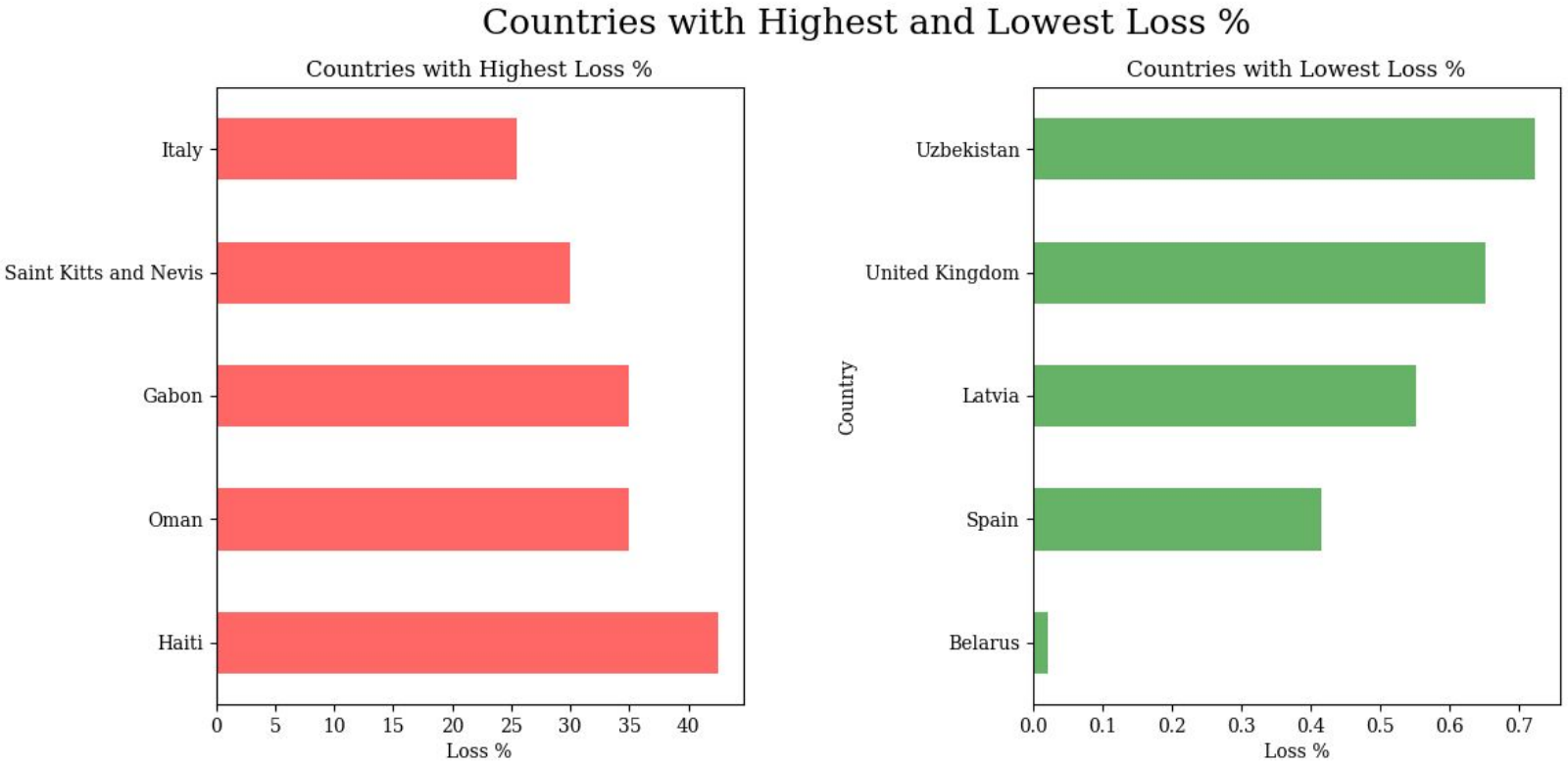
- **Source**: The Food and Agriculture Organization of the United Nations
- **Content**: Comprehensive data on food loss and waste across products, supply chain stages, and geographical areas
- **Size & Scope**: Dataset compiled from 700+ sources = 29k data points
- **Utility**: Provides structured and interactive data for deep understanding of food losses and waste, supporting efforts towards food security and waste reduction.

Methodology Overview

Data Science Process:

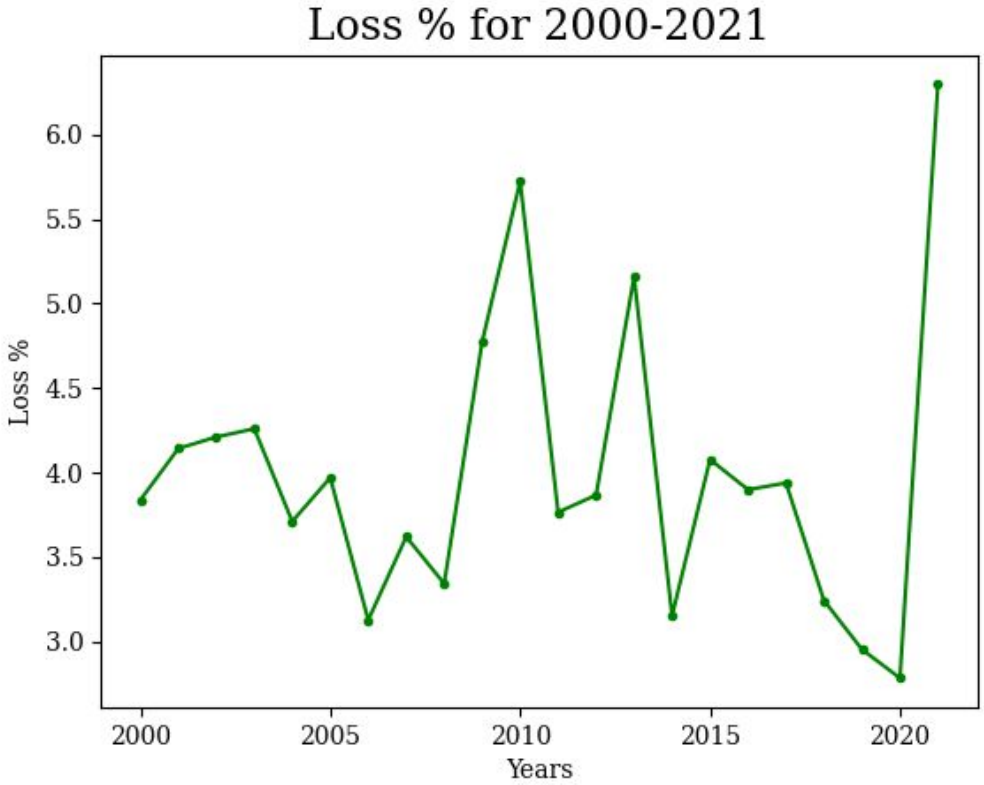
1. **Problem Understanding**
2. **Data Collection & Understanding (EDA)**
3. **Data Preprocessing**
4. **Modeling**
5. **Evaluation**

Key Findings



Description: This illustration identifies the countries with the highest and lowest food loss percentage. “Loss %” represents the average percentage of food lost annually from 2000 to

Key Findings

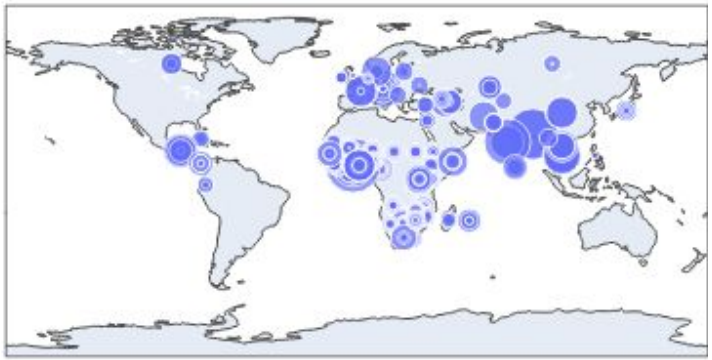


Description: This illustration shows the global average food loss percentage from 2000 to 2021.

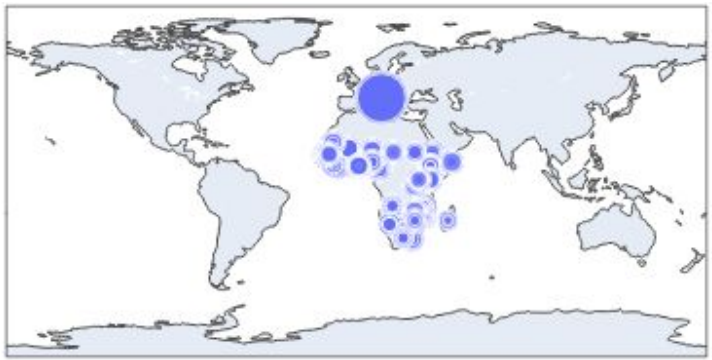
Key Findings

Geographical Distribution of Loss Percentage

2008



2020



Description: This illustration shows the geographical distribution of loss percentage in the years 2008 and 2020.

Classification Models Tested

- Neural Network
- Random Forest Classifier
- Decision Tree

Classification Model Performance

	Recall	Precision	F1	Accuracy
DecisionTree	0.9641	0.9643	0.9641	0.9641
RandomForestClassifier	0.9628	0.9617	0.9619	0.9628
NeuralNetwork	0.9246	0.9146	0.9176	0.9246

Regression Models Tested

- Stacking
- Random Forest
- Decision Tree
- Gradient Boost

Regression
Model Performance

	R2	MSE	RMSE
Stacking	0.7224	7.4384	2.7273
RandomForest	0.7174	7.5719	2.7517
DecisionTree	0.6063	10.5502	3.2481
GradientBoost	0.4874	13.7363	3.7063

Limitations

Imbalanced Classes

Data Collection

Missing Values

Modeling Limitations

Encoding Choices

Next Steps

1

Addressing Class Imbalance

2

Feature Engineering

3

Model Tuning

4

Advanced Imputation Techniques

5

Expand Data Sources

Real-World Application

Approximately 4.7 million people in Haiti are experiencing food insecurity, attributed to ‘poor performance on agricultural productions and heavy dependence on food imports’

- Model initiative to apply modeling to support Haitian producers and farmers
- Recommendations for improving models and features

Conclusion & Recommendations

- **Further Research**
- **Focus on High Loss Stages**
- **Policy Making**
- **Investment in Technology**
- **Awareness and Education**

Thank you !

Any questions?

This concludes our presentation!

Presentation by:

- Lisa Liang
- Juliana Sahagun
- Triston Crossland