
Sugarcane Production EDA – Executive Summary

This report summarizes the key insights derived from the exploratory data analysis (EDA) of sugarcane production data across various countries.

1. Data Overview

- The dataset consists of sugarcane production values reported for multiple countries.
 - Key columns in the dataset include:
 - **Country**
 - **Production (in tonnes)**
 - **Year** (if available)
 - The data required minor cleaning such as formatting production values and ensuring consistent country naming.
-

2. Top Producers

- The top 5 countries in sugarcane production are:
 1. **India**
 2. **Brazil**
 3. **China**
 4. **Thailand**
 5. **Pakistan**
 - These countries contribute a large portion of the global sugarcane output, with India often leading by volume.
-

3. Production Distribution

- A relatively small group of countries dominate global production.
 - Bar charts and pie charts revealed clear disparities between high-producing and low-producing countries.
 - Most of the sugarcane is produced in tropical and subtropical regions, ideal for cane growth.
-

4. Observations

- Production data showed high skewness, where a few countries produced over 70% of the global total.
 - Many countries had minimal production, either for local use or due to climate limitations.
 - Outliers and missing values were minimal and had little impact on trends.
-

5. Conclusion

- The sugarcane industry is concentrated in a few major producing countries.
 - This centralization may influence global trade patterns and policy.
 - Further analysis could explore:
 - Year-over-year trends
 - Correlation with GDP or agricultural land
 - Impact of climate change on production
-