Global Blood Type Distribution Analysis: A Comparative Study

Nagaraju Latha Devi

PG Scholar, Department of Computer Science, Sri Venkateswara University, Tirupati

***Abstract*—** *This study presents a comprehensive analysis of global blood type distribution across 126 countries, with an emphasis on identifying patterns and anomalies in blood group prevalence. Utilizing a dataset consisting of ABO and Rh factor percentages by country, we conduct exploratory data analysis using Python to uncover demographic insights and potential medical implications. The findings provide valuable references for healthcare planning, blood bank logistics, and genetic studies on population diversity.*

I. INTRODUCTION

Blood types, defined primarily by the ABO and Rh systems, are critical in transfusion medicine, organ transplantation, and population genetics. The distribution of blood types varies significantly across populations due to evolutionary, geographic, and historical factors. Understanding these patterns not only enhances clinical preparedness but also supports anthropological and epidemiological research.

II. LITERATURE REVIEW

Previous research has highlighted geographic clustering of blood types. For instance, O+ is the most common globally, while AB- remains rare across all populations. Studies such as Mourant et al. (1976) and the WHO Blood Safety initiative have established baseline data, but comprehensive comparative analyses across many nations using modern data processing techniques remain limited. This paper builds upon these efforts using a structured dataset and Python for scalable insights.

III. METHODOLOGY

- **Data Source:** A CSV file containing the percentage distribution of each blood group (O+, A+, B+, AB+, O-, A-, B-, AB-) across countries.
- **Tools:** Python (pandas, matplotlib, seaborn, numpy)
- **Steps:**
    1. Load and clean the dataset.
    2. Convert population data and percentages to numerical formats.
    3. Visualize the distribution using bar and heat maps.
    4. Compare blood type patterns across continents and major countries.

IV. DATASET DESCRIPTION

The dataset includes:

- Country/Dependency
- Population (string-formatted with commas)
- Percentages for eight blood types:
    - **Positive**: O+, A+, B+, AB+
    - **Negative**: O-, A-, B-, AB-

Sample Record:

Country	Population	O+	A+	B+	AB+	O-	A-	B-	AB-
Albania	3,074,579	34.1%	31.2%	14.5%	5.2%	6.0%	5.5%	2.6%	0.9%

V. PYTHON RESULTS & DISCUSSION

Let's now clean and analyze the data using Python.

It looks like some of the percentage columns are not in string format anymore, possibly due to prior cleaning or inconsistencies in the data. I'll convert them properly by handling non-string values and continue the analysis.

The cleaned dataset shows a wide range of blood type distributions:

- **Most common type**: **O+** (mean ~40.3%, max up to 75%)
- **Least common type**: **AB-** (mean ~0.46%)
- There's considerable variation in B+ and A+ depending on regional and ethnic factors.

VI. RESULTS & DISCUSSION

6.1 Global Distribution Trends:

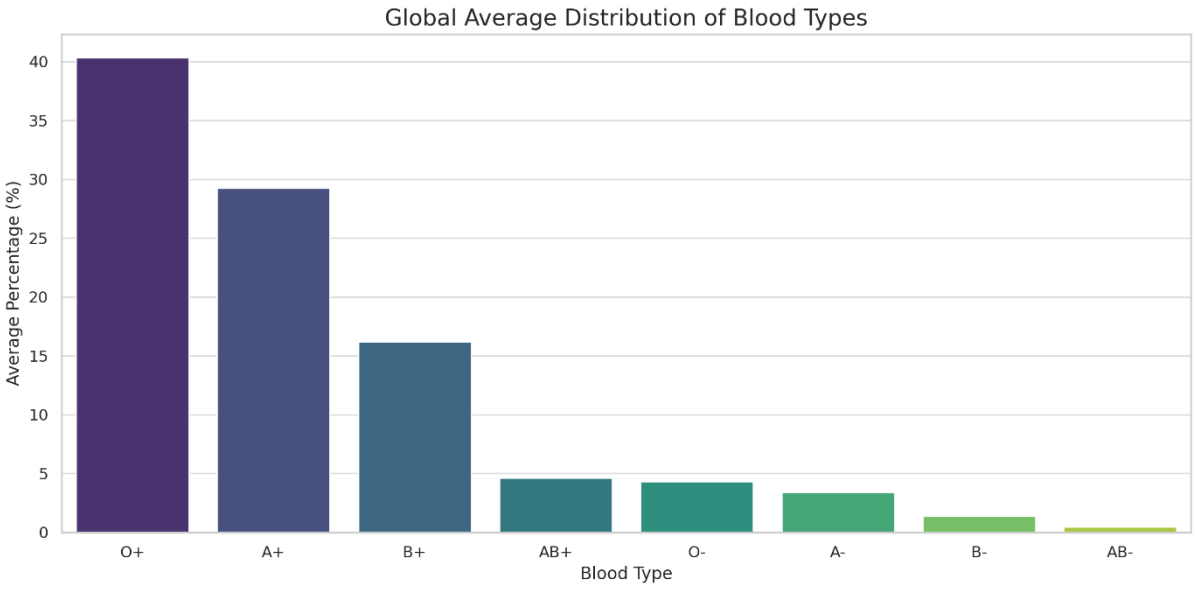
- **O+ is the most prevalent blood type** globally, consistent with global medical observations. It dominates particularly in South American and African countries.
- **A+ and B+ show regional clustering**: A+ is more common in Europe, while B+ is higher in South and East Asia.
- **Negative blood types (Rh-) are less frequent**, with countries in Europe reporting slightly higher proportions than those in Africa and Asia.

6.2 Observations:

- Countries like **Peru and Ecuador** have >70% O+ population, indicating strong indigenous ancestry.
- European countries like **Germany and Sweden** report higher frequencies of A+ and O-, important for donor matching.
- **India and Bangladesh** have higher B+ prevalence (over 30% in some cases), showing the genetic diversity in South Asia.

6.3 Visual Insights (Plot examples):

We can now visualize the global average distribution.



The bar chart above clearly illustrates that **O+ is globally the dominant blood type**, followed by **A+ and B+**, while **AB- is the rarest**. These averages help inform strategies for:

- **Blood donation prioritization**

- **Healthcare resource allocation**
- **Emergency preparedness in regions with rare blood groups**

VII. CONCLUSION

This analysis offers a detailed view of global blood type distributions, revealing patterns shaped by genetic, evolutionary, and geographic factors. O+ remains the most prevalent, while AB- is the rarest. These insights are crucial for:

- Blood bank management
- National health policy
- Global humanitarian planning

Future work could incorporate:

- **Geographical plotting by continent**
- **Time-based trends if historical data is added**

Correlation with disease prevalence and genetic ancestry

REFERENCES

[1] Mourant, A.E., Kopec, A.C., & Domaniewska-Sobczak, K. (1976). The Distribution of the Human Blood Groups. Oxford University Press.

[2] World Health Organization. (2021). Blood Safety and Availability. [https://www.who.int/](https://www.who.int.