

World Population Data - EDA Findings

Dataset Structure

- The dataset includes population, area, continent, and country-related information.
- Key columns: Country Name, Continent, 2022 Population, 2020 Population, Area (km²), Density, Growth Rate, and World Share.

Descriptive Statistics

- Population (2022) ranges from a few thousand (small nations) to over 1.4 billion (China, India).
- Growth Rate is generally positive, but some countries exhibit negative growth.
- Density varies widely: Bangladesh is highly dense; Australia is very sparse.

Missing or Inconsistent Data

- No major missing values were observed.
- Some columns (like World Share) may require type conversion (percentage format).

Correlation Heatmap Findings

- Area and Population have a positive but weak correlation.
- Density and Area have a strong negative correlation (larger countries generally have lower population densities).

Histograms

- Population distribution is extremely right-skewed.
- Area distribution is also highly skewed.

Scatterplot (Area vs Population)

- No strong direct relationship: A large area does not guarantee a large population (e.g., Australia).
-

Final Insights

- Asia dominates the world's population.
- Population growth is slowing down in several regions.
- Area size does not predict population size reliably.
- Population distribution is highly uneven across the world.

Summary Conclusion

The world's population distribution is heavily skewed towards a few countries. Area size does not guarantee a high population, and growth rates are slowing in several parts of the world. Asia remains the dominant continent in terms of population spread and density. Overall, the data highlights stark differences in population trends across different regions.