

Project Proposal for Advanced Visualization in R

Visualizing Global Religious and Population Trends Over Time

Poojitha Annabathula (468459)

Muhammad Ibrahim Hanif (468257)

November 2024

Introduction

Religion has been an integral force shaping societies, cultures, and conflicts throughout history. Understanding global religious trends provides unique insights into human behavior, societal evolution, and the interconnection between religion and population dynamics. This project leverages the World Religion Data (WRD) from the Correlates of War Project, offering a comprehensive perspective on global religious affiliations and their changes over time. Through advanced visualization techniques in R, this project aims to uncover patterns, shifts, and correlations in religious demographics from 1945 onward.

Objectives

1. Analyze the historical growth, decline, and distribution of major religions globally over time.
2. Examine regional religious dominance and identify significant shifts in affiliations due to socio-political factors.
3. Highlight the growth of non-religious populations and their correlation with migration, wars, and decolonization events.
4. Explore the reliability of data across regions to assess trends in reporting and accuracy.
5. Produce high-quality, innovative, and aesthetically pleasing R visualizations to effectively communicate key insights and patterns from the data.

Dataset

The dataset for this project is the **World Religion Data (WRD)** provided by the *Correlates of War Project*. It offers a comprehensive overview of religious populations across 194 countries from 1945 to 2010. The dataset consists of over 18,000 entries and includes variables such as country codes, year, total population, and specific numeric values for each religious group. Data is stored in a structured tabular format. It can be accessed at the following link: <https://correlatesofwar.org/data-sets/world-religion-data/>

Key Features:

- **Religious Groups:** Includes major religions such as Christianity (subcategories like Catholic, Protestant), Islam (Sunni, Shia), Judaism, Buddhism, Hinduism, and non-religious populations.
- **Population Metrics:** Provides both percentage and absolute population data for each religious group in each country, along with total population estimates.
- **Temporal Range:** Annual data from 1945 to 2010.
- **Reliability Indicators:** Variables (sourcereliab, recreliab, reliabilevel) assess data accuracy by region.

Methodology

Data Exploration and Cleaning

- Preprocessing the data.
- Address missing data and inconsistencies in population and percentage values.
- Summarize key descriptive statistics to understand global and regional trends.

Visualization Techniques

We plan to implement various advanced visualization techniques, including line plots, area charts, treemaps, and interactive visualizations using R. The exact techniques may change as we refine our analysis.

Timeline

- **End of November:** Finalize dataset selection, project topic, and submit the proposal.
- **December:** Perform data exploration, cleaning, and create initial visualizations.
- **January:** Develop and refine advanced visualizations, iterating based on findings. Complete documentation, and deliver the final submission.