Language Status Visualization

An Interactive Visualization of the Status of Dead, Dying, and Endangered Languages Worldwide

Grace Warren

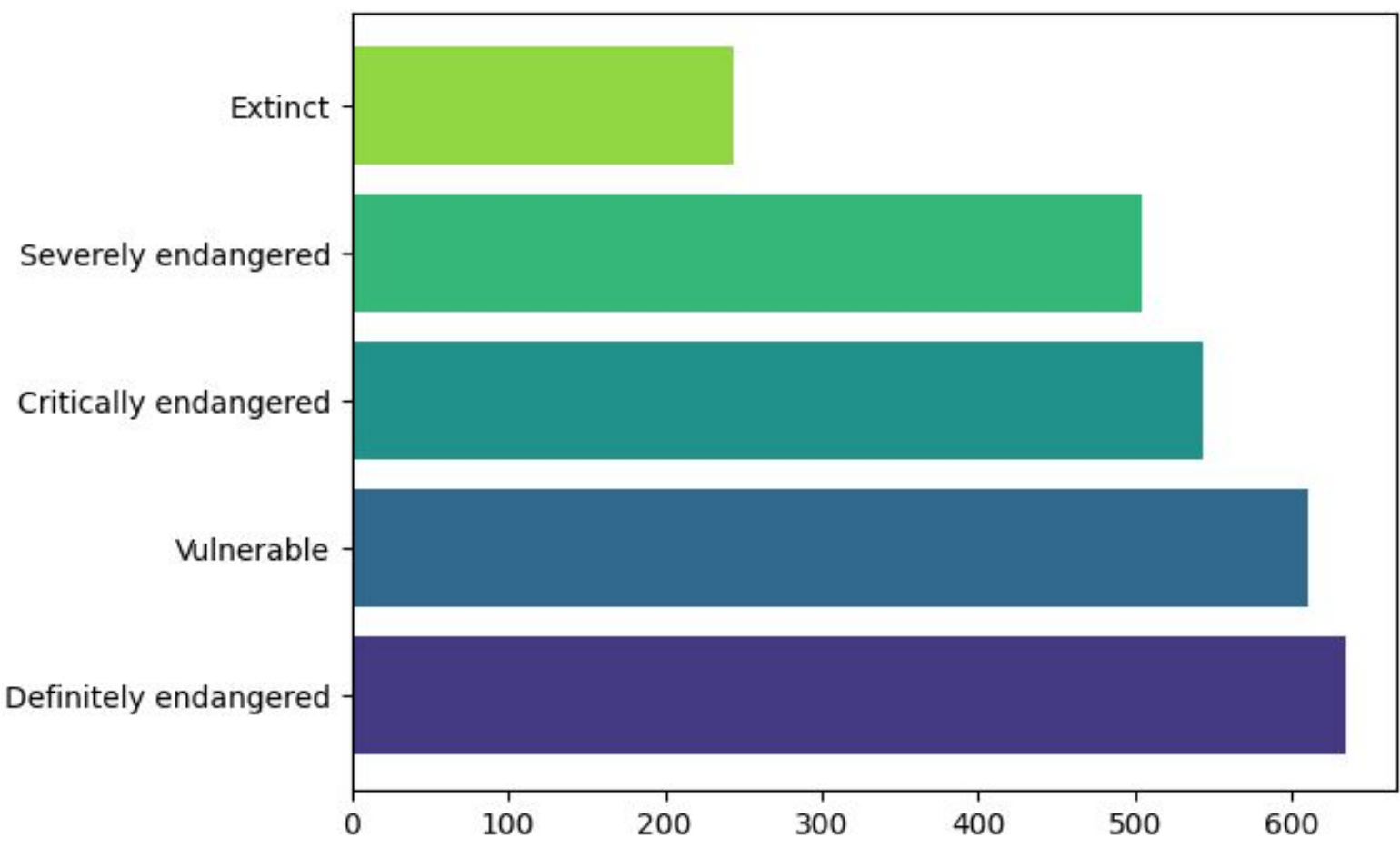**FURMAN** COMPUTER SCIENCE

ABSTRACT

Perhaps shared by distant genetic relationship, language endangerment is currently one of the most favored topics in the field of Linguistics. Despite this, comprehensive accounts of language status across world languages remains behind paywalls. This project seeks to consolidate publicly available information into a single user interface, which the layman might use to explore language endangerment in an accessible fashion.

INTRODUCTION

In recent years, the emergence of comprehensive open-source databases like *Ethnologue* has begun to shed light on the realities of language endangerment. However, these databases remain behind paywalls, rendering them inaccessible to many.

To remedy the inaccessibility of this information, a compilation of pertinent information was necessary. Through the collection of publicly available information, a compilation of statistics regarding approximately 40% of all living languages (and counting) has been produced. This is accessible through an interactive publicly available website.
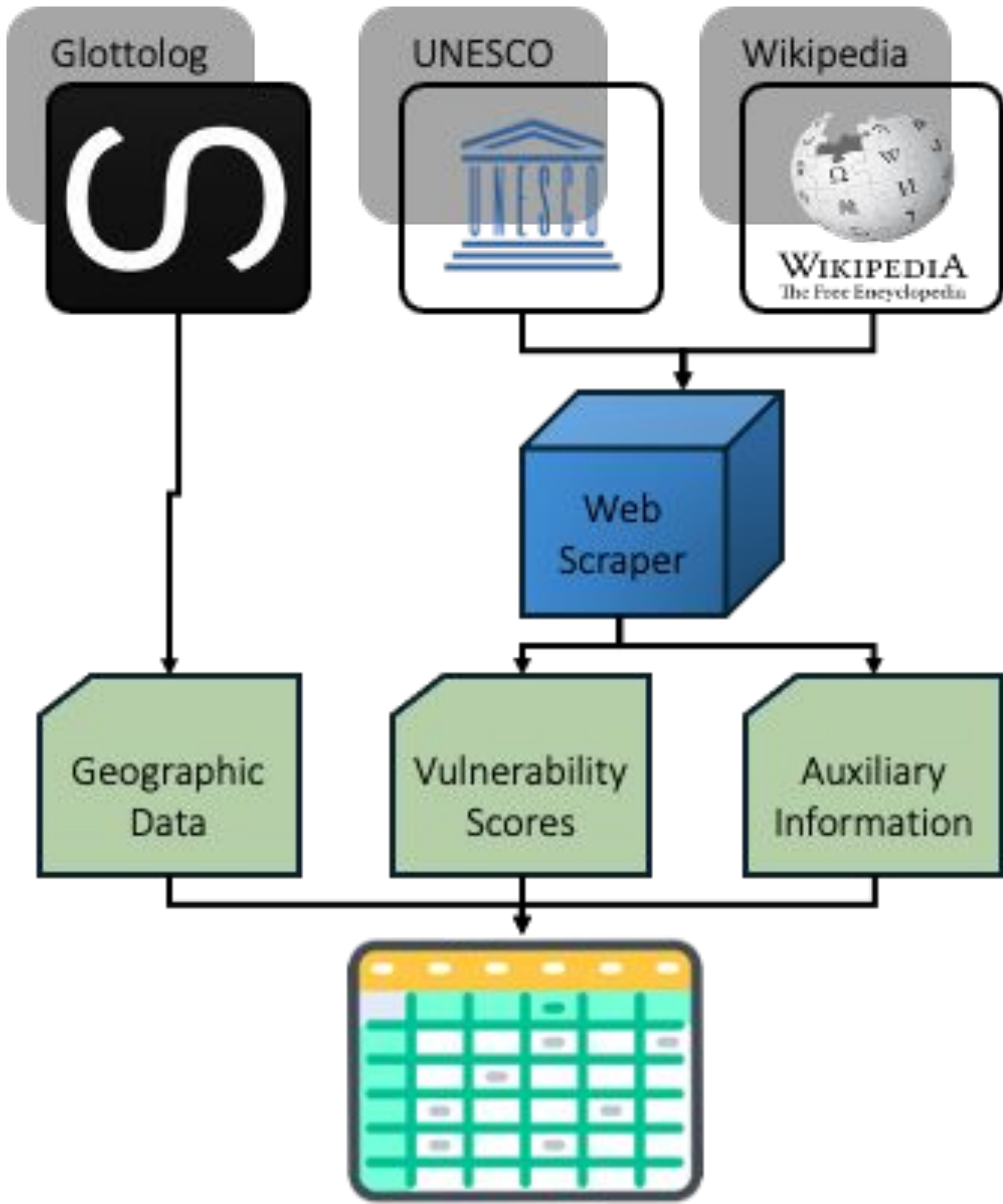
DATA

Data has been collected from Wikipedia, Glottolog, and UNESCO.

In an attempt to provide an impactful visualization, the focus was put on vulnerable, endangered, and extinct languages. The distribution of these languages is visualized above.
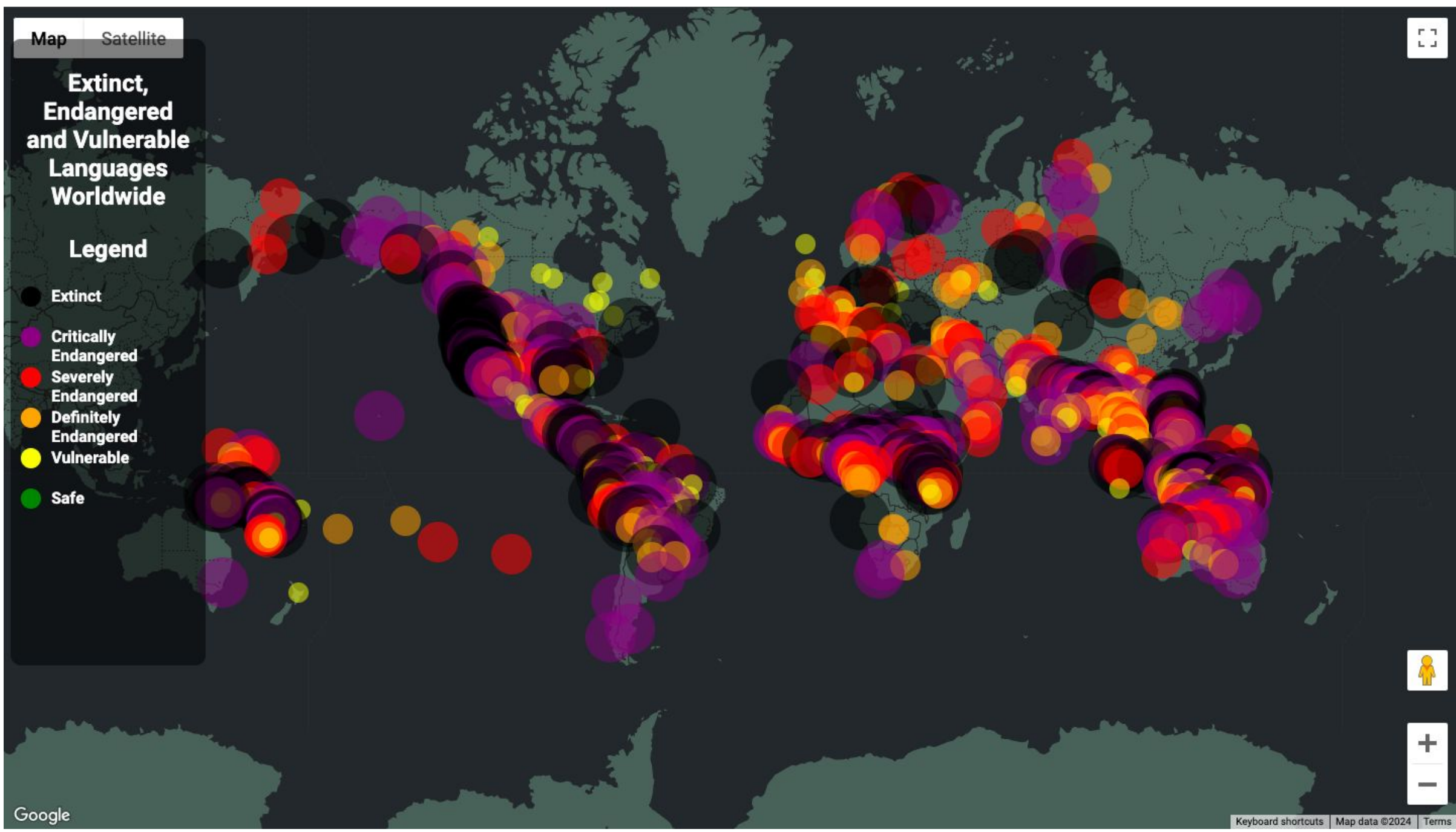
PIPELINE

The data for this project was acquired using web scrapers designed to gather the links for as many language Wikipedia pages on as possible, followed by acquiring the following information from that page, where applicable:
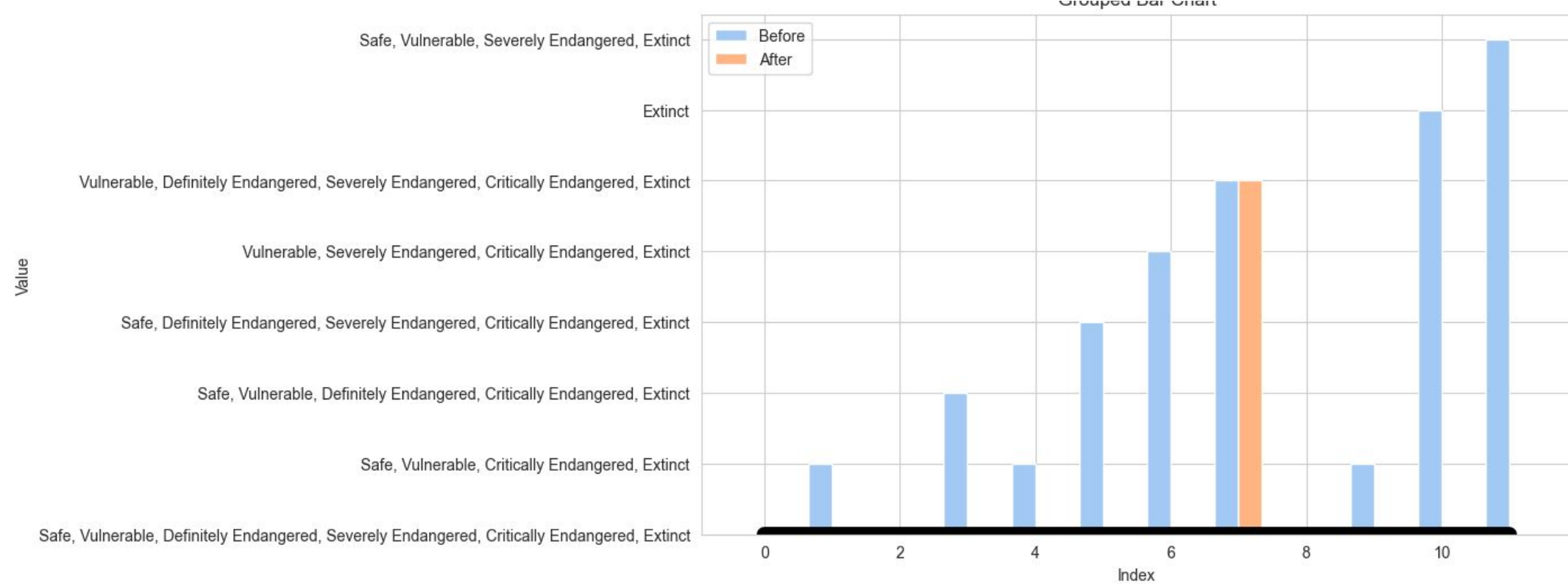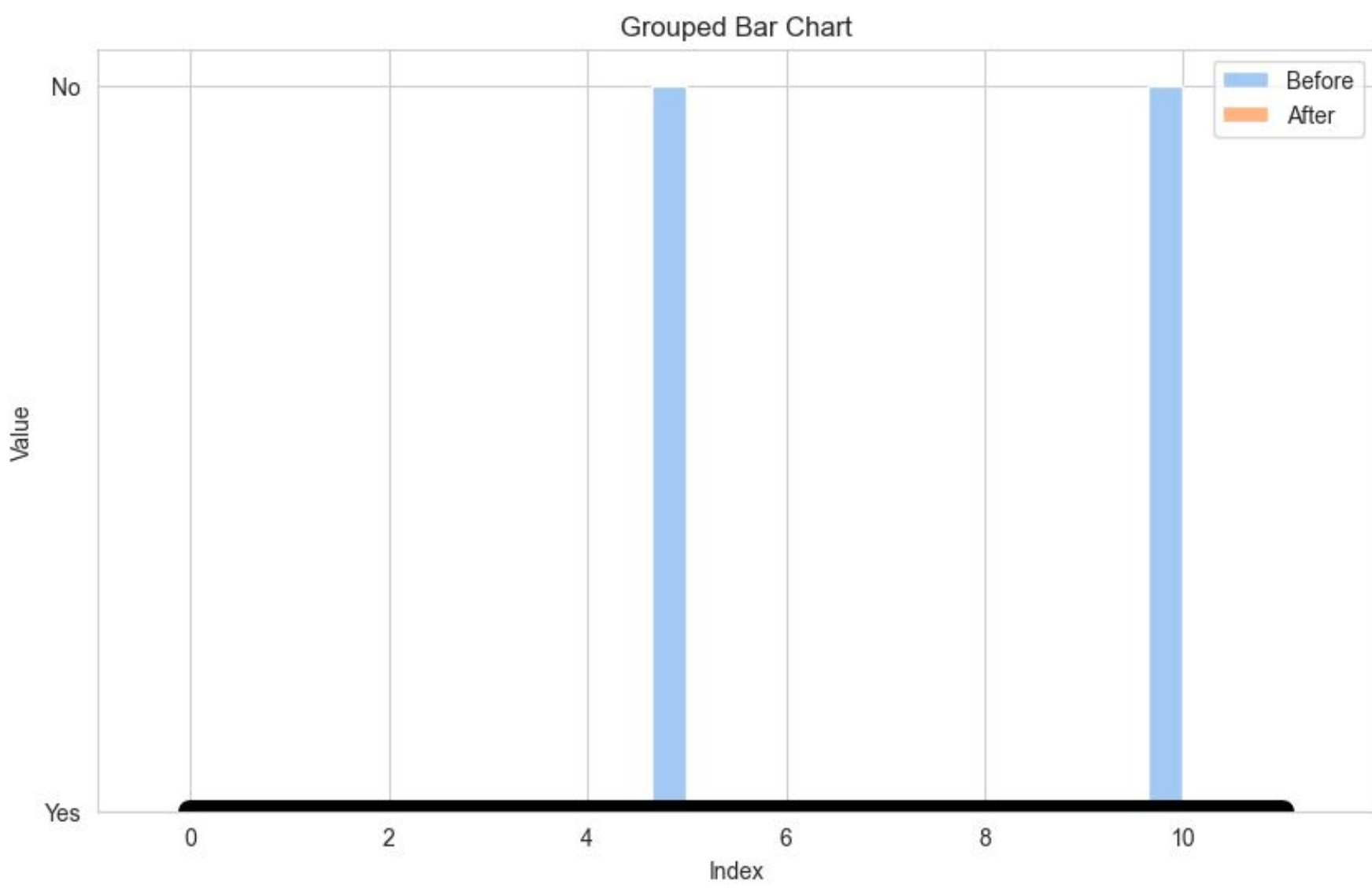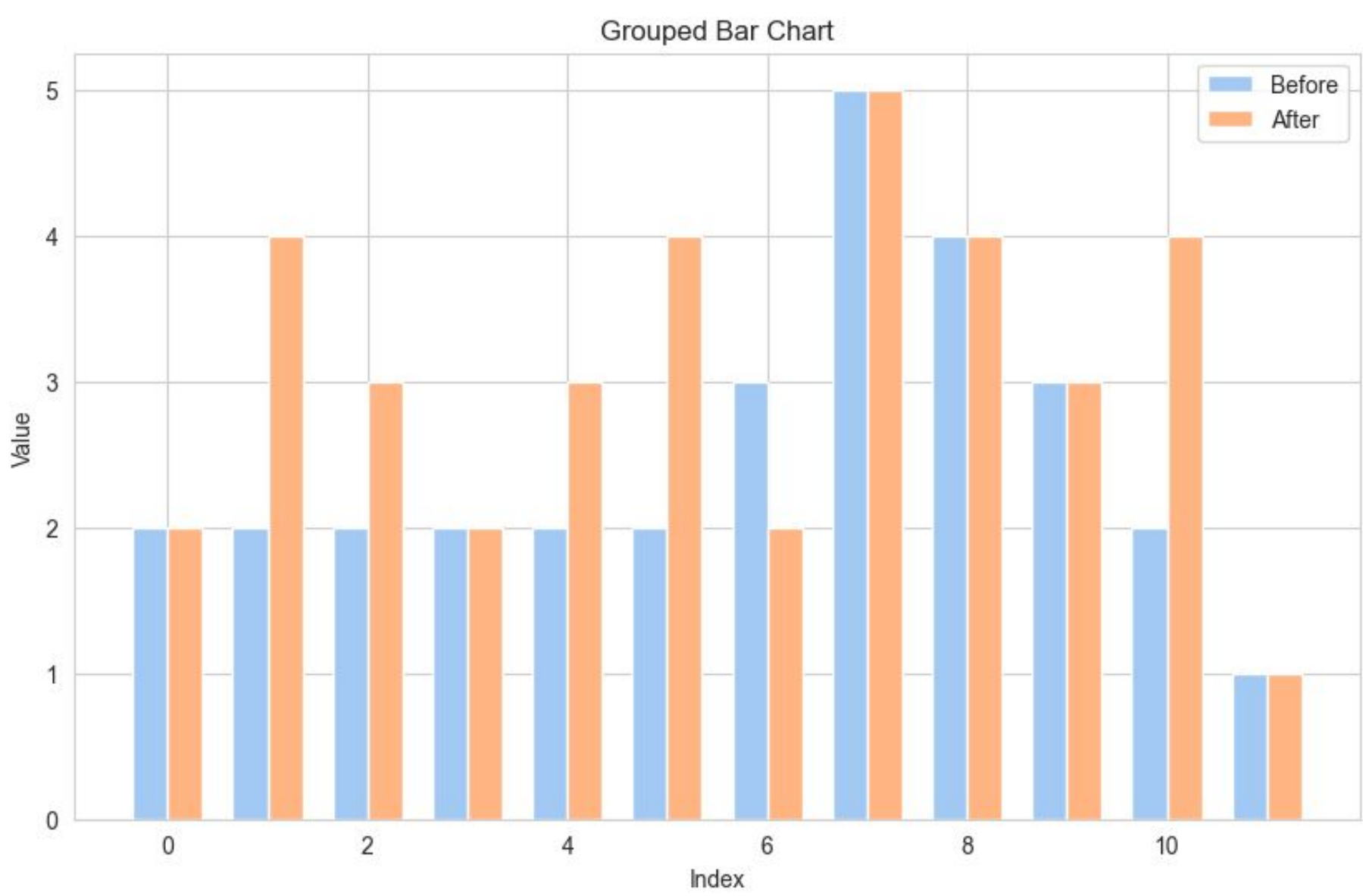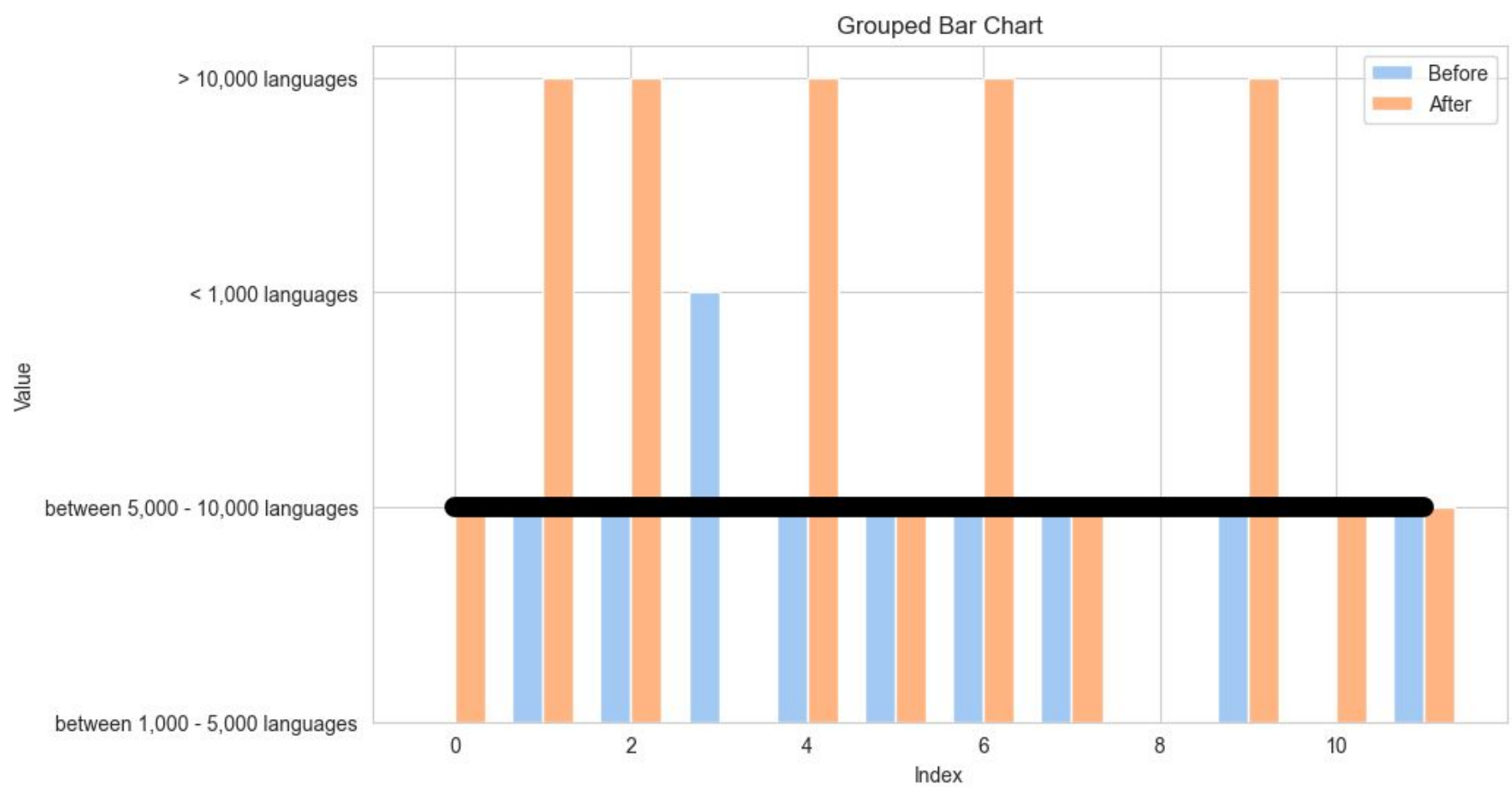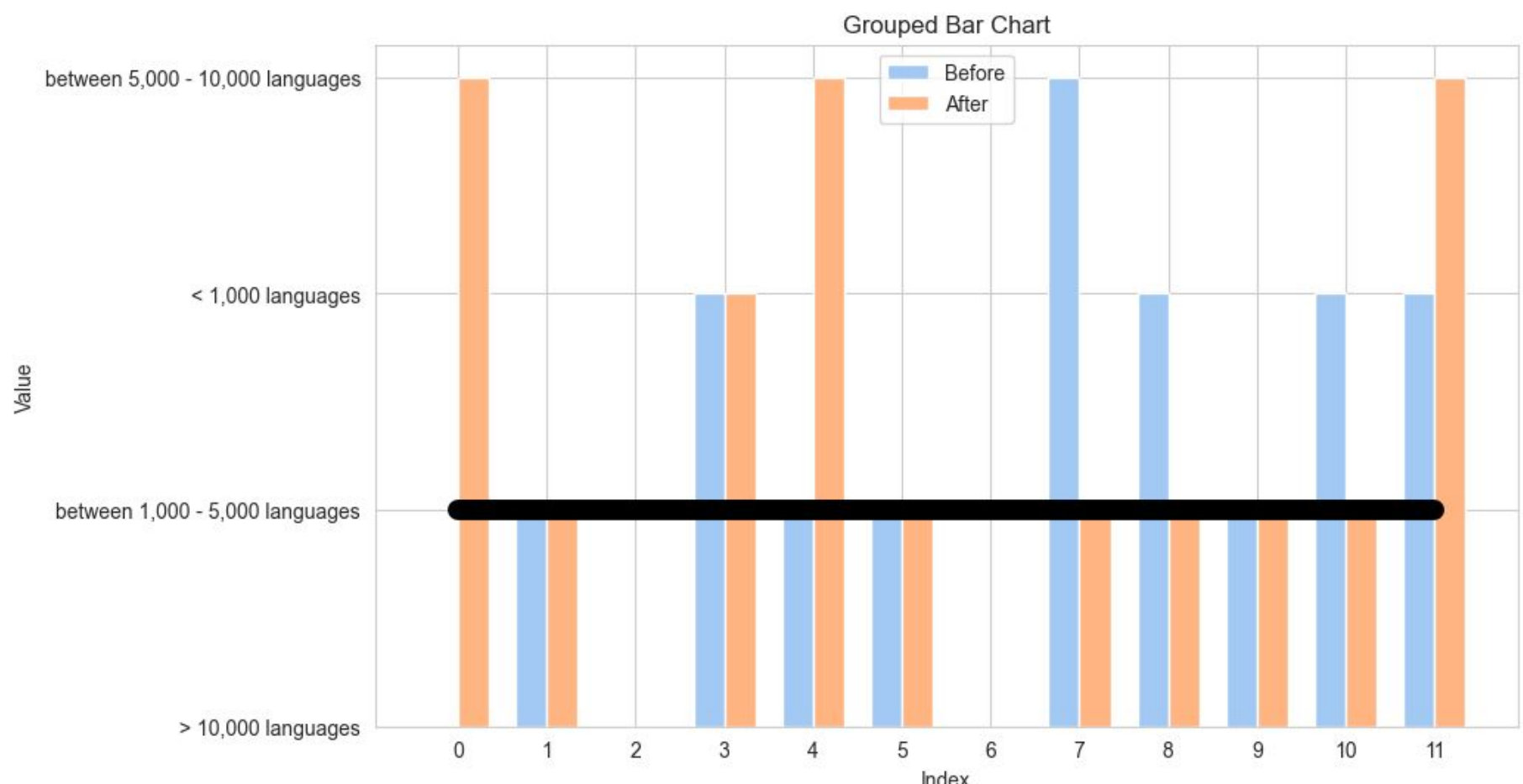
- Language name
- Language family
- Dialects
- ISO 3 Code
- Number of speakers
- Regions spoken
- Official Language
- Recognized minority language

USER INTERFACE

To create language markers which draw proportional attention to the status of a language, markers were assigned a color and size, for which larger sizes correspond to a larger level of endangerment.

IMPACT