What is White Supremacism?

Belief systems with the core ideas that white people should:
- have dominance over people of other backgrounds
- live by themselves in a “superior” whites-only society
- and the belief that they are genetically superior to other people

The Task:

In a White Supremacist forum scraped from the Dark Web, some users mention where they live.

This project consisted of the following two tasks:

1) Extracting geographical markers that signify where a user currently lives
2) Plotting and analysis of the geographical marker distribution

The Dataset:

Stormfront, the forum:
- The Internet’s first major racial hate site
- Comprised of posts made from 2001-2015
- Totals in 23,193,784 posts and ~24GB in size

How it was done:

After filtering out all HTML and other markup, the following algorithm was applied:

1) Determine if the user states where they currently live
2) If so, use the Natural Language Tool Kit and the Geopy Geocoder to verify the location

For every message in the dataset:

- “... I live in ✓ Toronto” ✓ ✓
- “... we reside in ✓ a small house” × ×
- “... if I live in × Texas” ✓ ×

Please Note:

- The technique (completed in Python 3) currently only considers English and some French speaking posts