

Historical data visualization

November 11, 2015

Abstract

In the contemporary humanities, datasets are not just evidence but archives, demanding reinterpretation; visualization provides one of the richest and most widespread ways facilitating this. This talk will describe the reception and remarkable misrepresentations of the most influential single data visualization in the historical profession, the US Census's maps of the frontier line from the late 19th century; and then describe an agenda of web-based (D3) data visualizations geared towards exploratory analysis that can make allow freer exploration of data archives as evidence. These platforms—for exploring census data, historical shipping routes, and text collections with metadata—embody an approach towards humanities data visualization not simply as presenting single views, but as creating weak domain-specific-languages for sharing data archives with scholars and a wider public.

1. Overview
 - (a) Core questions
 - (b) Description of the project
2. Visualizing the frontier
 - (a) Famous Atlases
 - (b) At the heart of the historical profession
the frontier line itself is restored
 - (c) Baking in the inability to criticize
 - i. Center of population as an example
 - A. Most likely skip all of this.
 - ii. Case studies of people marvelling at maps
3. DSLs for datasets
embedded data visualizations underwritten by a rich API.

(a) This is what I rely on. But it does not allow argumentation directly.

web sites driven by a public facing data API.

(b) Domain-specific languages.

4. Visualizing the census

5. Visualizing Shipping

6. Visualizing Texts

platform for the analysis of large textual collections

general users,

(a) Family resemblance to Google Ngrams

This is not an advance in visualization

i. Our other browsers

A. Yale University Libraries

B. Medical Heritage Library

C. US State Department

D. Hathi

(b) A generative grammar for textual analysis

(c) Core functions of the grammar.

i. 1. Creating a corpus

ii. 2. Setting a working definition of texts and tokens

A. First, Texts.

B. Gender and language

C. Different kinds of multivariate approaches

useful even without words:

iii. 3. Performing an operation that compares counts in two corpora against each other.

A. Log-Likelihood

iv. Returning and comparing results

v. Geocoding

(d) Ending