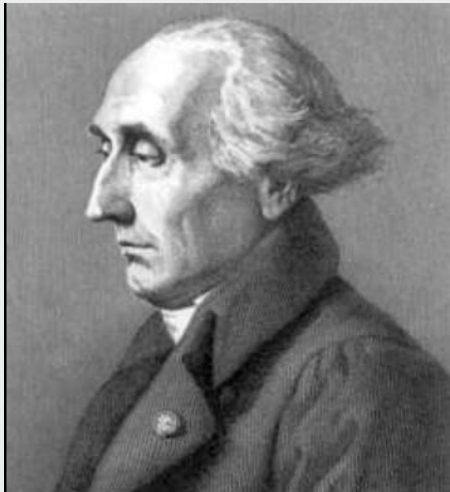
Information Visualization Final Project

Mathematician Genealogy



Weicheng Zhu & Fangyuan Liu

Introduction - motivation

- How did mathematics developed over the past centuries?
- How are mathematicians connected with each other?
- Is mathematical knowledge truly able to spread beyond the boundary of countries and regions, or mathematicians only communicate with their peers in the same country?
- Some Maths terms are named after the mathematicians. Learning their connection can provide us with a better understanding of the theorems and their connections.

Introduction - Tasks

1. Present the basic information of a mathematician with annotation (name, field of study, etc.)
2. Present the connections of mathematicians based on their nationalities
3. Present links between mathematicians as an advisor and his students
4. Present a mathematician in the scope of all mathematicians in the history
    a. Derive the importance of a mathematician and present it on the visualization
    b. Discover the ancestors of a mathematician
    c. Present the links between a mathematician and his ancestors
    d. Summarize the trend of change in research interest on any branch of mathematicians
5. Allow users to locate a mathematician, or to explore the visualization

Introduction - Dataset

Data type: HTML → Table data → Network data

Data size: > 200k mathematicians' information;

Attributes: {name, advisor, students, descendants, nationality, university, field, year}

Data Source: Mathematics Genealogy Project (https://genealogy.math.ndsu.nodak.edu/index.php)(North Dakota State University)

Preprocessing Methods: Scrapy according to xpath from html; clean data; transform into table and tree form

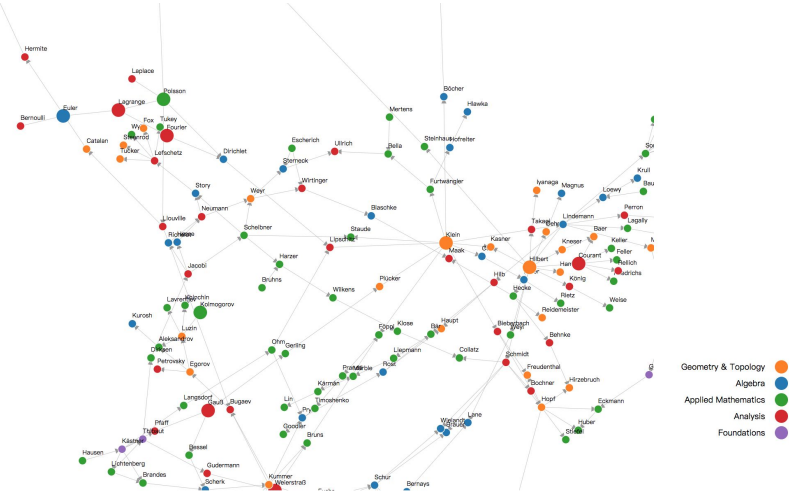
System Overview

Visualization System:

1. Multilevel Force-directed graph
2. Hierarchical treemap
3. Geographical map

Visualization Design - Forced Directed Graph

- Channel: Color, Position, Size
- Mark: Line, Point, Connection
- Rationale:
  - The user can clearly see the ancestor and direct students of a mathematician so he can know the origination of a certain school in Maths;
  - From colors, the user can easily tell the field alone the development of a family and the distribution
  - Force directed layout is efficient because it separates influential mathematicians farther
- Techniques: Force simulation layout, BFS in directed graph
- Result:
  - Less than 10% of mathematicians is neither a descendant of Poisson nor Hausen
  - Most mathematicians after 1850 is a descendant of both of them

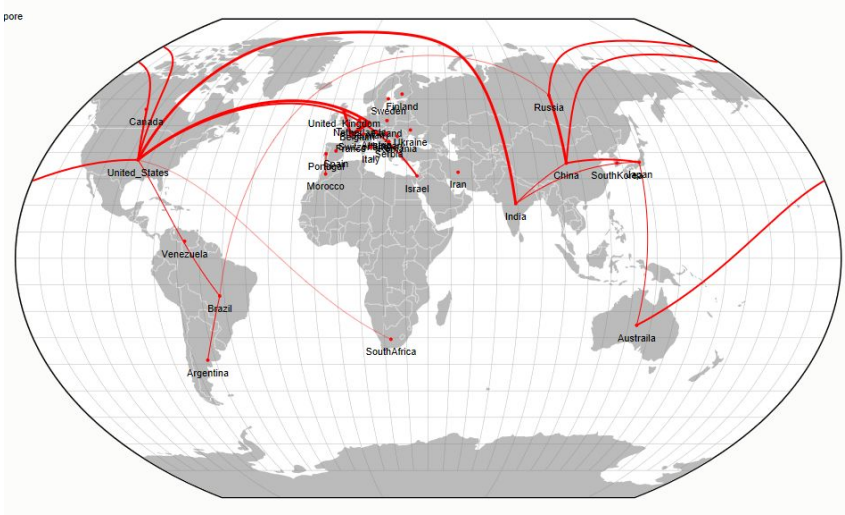


Visualization Design - Treemap on Maths fields

- Channel: Color, Size
- Mark: Area
- Techniques: Hierarchical Treemap
- Rationale:
  - The user can see how mathematical research fields are categorized, and how well are them being studied
  - From saturation of colors, the user would be able to recognize the most heated topics in mathematics in different times
  - A treemap layout depicts the inclusion relation of research fields and its subfields, and distinguishes them clearly
  - Have space to present more information
- Result:
  - More than half of all mathematicians devote themselves in applied mathematics research
  - The most studied pure mathematics subfield is differential equations

Visualization Design - Geographical Connection

- Channel: size, position
- Mark: line, point
- Techniques: Geographical Map
- Rationale:
  - The user can see the connections between mathematicians categorized in countries
  - From the menu above, the user can choose from the listed time periods to see the connection graph at that time
  - Geographical layout strengthens the sense of space



Result - Interesting Finding

- Most mathematicians can be traced back to several famous mathematicians in 18-19 Century, such as Euler, Gauss and etc.
- The families turned into applied Maths fields were not likely to switch back to pure Maths
- Analysis, topology and algebra are closely related, so a student need to learn all of them well
- The Maths world is really small, as some mathematicians have families that form a loop, even multiple loops sometimes
- Early mathematicians tend to have less students than mathematicians later