

# Project Progress Report 2

(due May 24th 11:59p.m)

You can start working on the project once your report is accepted and graded by your TA. The entire final project is worth **35%** of your final grade and this report accounts for **10%**. This project is done individually.

## **Submission Guideline**

Download this google doc, fill the table. **Type** your answers, no handwritten answers will be accepted (except for the very last question). Submit it in **PDF** format on Gradescope.

If you need some inspirations please feel free to take a look at:

[Showcase of Information is Beautiful Awards](#)

[Bloomberg Year In Graphics Review](#)

[The Pudding](#)

[The New York Times](#)

## Project Guidelines

**Note: The guideline has been further clarified from Progress Report 1, so double-check whether your dataset choice still satisfies the updated guideline below.**

1. You may use more than one dataset, however, regardless if you use one or multiple datasets, your visualizations must make use of at least three following data types - **link, position, and attribute**.
2. You cannot use any dataset from the class (Labs, Assignments, Lecture Exercises)
3. You can make your own dataset (Web scrape etc.) provided point 1. is satisfied.

## Part 1 - Story and Narrative

[illegible]

<p>Story you want to deliver</p>	<p>Language is a huge element of personal and cultural identity, and I want to demonstrate the relationships between English and other world languages. With my project, I want to challenge notions of national differences and language-based xenophobia, and I especially want to emphasize the way that language is constantly changing through rich cultural exchange. Additionally, I hope to bring some attention to linguistics and present etymology as something interesting and accessible for people to engage with.</p>
<p>Describe your target audience.</p>	<p>(using the questions the lecture slide listed)</p> <p>My target audience is going to be English-speakers, since my dataset primarily focuses on English words.</p> <ul style="list-style-type: none"> <li>● Familiarity with your topic? If not, how do we catch them up? <ul style="list-style-type: none"> <li>○ I don't expect my audience to have any familiarity with my topic, aside from being a speaker of English. I plan to give a brief introduction into the general field of etymology (i.e. introducing that languages can borrow from each other), but otherwise I am going to keep my project very non-technical.</li> </ul> </li> <li>● Do they care? Why? Why not? <ul style="list-style-type: none"> <li>○ The project should be relevant (and hopefully interesting) for any English speaker, and I hope to make people a little interested in etymology/linguistics, but I don't expect this project to be life-changing for anyone.</li> </ul> </li> <li>● What do you want them to take away? Key points? <ul style="list-style-type: none"> <li>○ I hope people become more curious/aware of the words they use in daily life, which they often take for granted. I also hope people leave with a greater appreciation of the global, changing nature of our language.</li> </ul> </li> <li>● What do they know about visualization? Are your techniques standard? <ul style="list-style-type: none"> <li>○ I want to make my visualizations accessible for everyone, so I plan to use very standard techniques and otherwise have informative labels to make my visualizations easy to understand and use.</li> </ul> </li> <li>● How do they encounter your visualization?</li> </ul>

	<ul style="list-style-type: none"> <li>○ On a website or potentially a blog (digital, online format with potential for interaction).</li> <li>● Mathematical background? Are you assuming too much? Too little? <ul style="list-style-type: none"> <li>○ I don't plan on featuring any heavy mathematical details in my visualizations, so I won't really be making any assumptions about viewers' math background. Because math won't be a super heavy feature of my visualizations, someone with no background should be able to enjoy my project, but I also don't plan on over-explaining any math for people who already have a thorough math background.</li> </ul> </li> <li>● Device? <ul style="list-style-type: none"> <li>○ I am going to plan my layout for a computer screen, but if I have time, I may do some adaptations for mobile viewing.</li> </ul> </li> </ul>
The goal of your project outcome. And why?	<p>(exploratory vs. explanatory)</p> <p>The focus of my project will primarily be explanatory, as my goal is more centered about communication, and I expect my general audience to not have a lot of experience with linguistics or etymology. However, I hope to provide some interesting interactive elements for users to explore more freely and engage with the topic more deeply, if they want.</p>
Narrative structure you plan to use	Martini Glass Structure
Elaborate your choice of narrative structure.	<p>I want to start by presenting some curated overview graphics, to introduce the user to the dataset, and then I will present my primary graphic for communicating my message about globalism and the development of the English language. This will be the "guided, structured narrative" section, after which I will provide some more open, exploratory tools for the user to investigate the dataset (e.g. by searching for specific words and/or using interactive tools to modify the visualization/data settings).</p>

Narrative genre you plan to use	I expect that my project will be most similar to the “annotated chart” and “partitioned poster” genres from lecture.
Elaborate your choice of narrative genre.	My individual visualizations will be accompanied by annotations, but my overall structure will feature multiple projects/sections, which matches the "partitioned poster" genre. I think this structure will provide a good balance of informativity and accessibility for users without overwhelming them with information. I also think that this structure is most appropriate for the scope of content I want to present.

## Part 2 - Outline

Story you want to deliver	Language is a huge element of personal and cultural identity, and I want to demonstrate the relationships between English and other world languages. With my project, I want to challenge notions of national differences and language-based xenophobia, and I especially want to emphasize the way that language is constantly changing through rich cultural exchange. Additionally, I hope to bring some attention to linguistics and present etymology as something interesting and accessible for people to engage with.
Specifications on each plot <b>in the order of how you lay out on your project</b>	<ol style="list-style-type: none"> <li>1. Etymology language-relationship stacked bar chart <ol style="list-style-type: none"> <li>a. Tasks: <ol style="list-style-type: none"> <li>i. Present the different etymological relationships represented in the dataset.</li> <li>ii. Summarize the languages represented in the dataset.</li> <li>iii. Allow users to compare the impact and prevalence of different languages and etymological relationships.</li> </ol> </li> <li>b. Attributes: origin language (categorical) etymological relationship (categorical), number of terms (numerical discrete)</li> <li>c. Marks: line mark</li> <li>d. Channels: <ol style="list-style-type: none"> <li>i. Vertical position (aligned at bottom group, unaligned for others) for number of terms</li> </ol> </li> </ol> </li> </ol>

- ii. Aligned horizontal position for term origin language
  - iii. Color hues (in stacked bars) for type of etymological relationship (e.g. borrowing, cognate, compound)
- e. How this plot adds to the story: It will introduce the languages and etymological relationships presented in the dataset. This should prepare the viewers for the other, more impactful visuals.
- 2. Map-network plot
  - a. Task:
    - i. Allow users to lookup countries/languages of interest (by position on the map) and view their connections to other countries/languages
    - ii. Allow users to compare the etymological impactfulness and connectivity of different countries/languages.
    - iii. Allow users to identify geographical trends in language interactions.
  - b. Attributes: country location/shape (position), number of terms (numerical discrete)
  - c. Marks: area marks (map shapes) and line marks (links)
  - d. Channels
    - i. Shape/position channels to show geographical location of countries/languages on the map
    - ii. Line width channel to show number of terms in the dataset connecting the two languages
  - e. How this plot adds to the story: This plot will directly demonstrate which languages English terms show the most etymological relationships with. This will demonstrate the global relationship between different languages and emphasize the main idea of my project - that languages are deeply connected.
- 3. Explorative etymology network visualizer
  - a. Task:
    - i. Allow users to locate/lookup the relationships and related terms for a specific target term.
    - ii. Allow users to explore the etymology dataset on a more granular level than the other visuals.
    - iii. Allow users to identify etymological connections between target words.
  - b. Attributes: term (categorical), term language (categorical), term connectivity (number of related terms; numerical discrete), etymological relationship type

	<ul style="list-style-type: none"> <li>c. Marks: circle for terms, line for etymological relationships (links)</li> <li>d. Channels <ul style="list-style-type: none"> <li>i. Circle color (hue) for term language</li> <li>ii. Circle radius (size) for term connectivity</li> <li>iii. Line style for relationship type</li> </ul> </li> <li>e. How this plot adds to the story: This will allow users to explore the etymological data for any term they want to investigate. This network-style visual will provide more details for a specific term and its connected terms, whereas the previous visualizations are more general and language-level. In contrast, this visual will allow the user to do self-guided exploration of the data at a term-level.</li> </ul>
Elaborate the choice of their marks and channels for each vis	I generally opted for more salient visual channels (e.g. color hue, length) first. For categorical attributes, I generally opted to use color (hue) where possible, and with numerical attributes, I generally opted to use size/area where possible. I don't generally plan on double-encoding the attributes in my visuals, but I may do some double-encoding to provide more visual interest and consistency across visuals (e.g. adding color hue to the map-network plot to correspond with the color hue encoding of one of the other visuals).

## Part 3 - Prototype

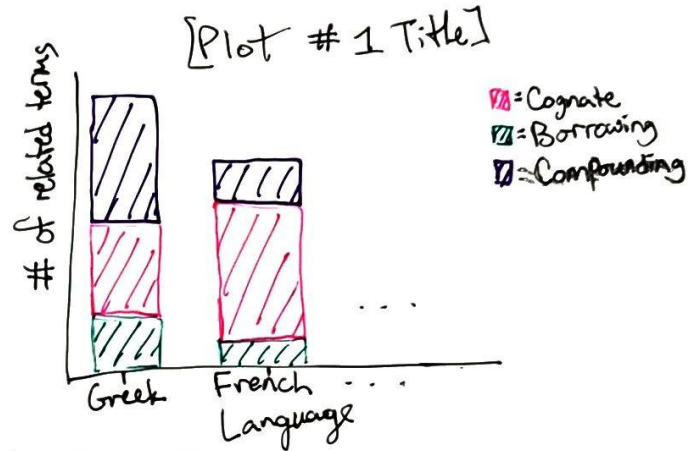
Provide a photo or screenshot of your prototype. A prototype should depict how you place different components of your visualization. You may use pen-paper, or using tools like excalidraw, figma etc.

A basic, barebones sample prototype for this project

# Etymology & Globalization of Language

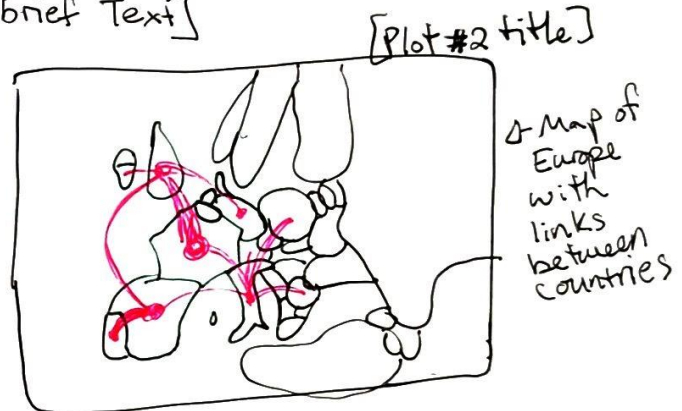
[Text-introduction to etymology & project]

[Description  
of Etymological  
Relationships]



[Plot #1 Debrief Text]

[Message about global  
interaction & language  
development  
(Storytelling section)]



[Plot #2 Debrief & Main message]

[Plot #3  
Description  
& instructions  
for  
interaction]

Plot # 3 Title

Word: cat

Languages:

- ☒ Greek
- ☒ French
- ☒ Latin
- ☒ English

