Final Project Proposal for Visualization (CS 6630)

Basic Info.
The project title: *The complete history of bids for summer Olympics*
Names: Ren Miao, Jiyuan Li
E-mail addresses: rmiao@cs.utah.edu, u0710273@utah.edu
UIDs: u0748747 (for Ren Miao), u0710273 (for Jiyuan Li)
The link to the project repository: https://github.com/VanillaCola/OlympicsBids

Background and Motivation.
The summer Olympic Game or the Game of the Olympiad is the largest international multi-sport event. The first summer Olympic Game was held in 1896 at Athens, Greece then hosted by a different city every four years. Hosting an Olympic Game oftentimes requires the host city to build new stadiums, expand mass transit system and renovate the public facilities. All these will bring challenges as well as tremendous development opportunities to the host cities, such as creating new jobs, receiving financial investments and attracting global attentions. Due to the benefits of hosting an Olympic Games, the major cities that are interested in hosting the game need to bid for it and rival for the hosting right. For example, 12 major cities around the world bid for the hosting right of 1936 Olympic Game and eventually Berlin, Germany was selected as the host city.

People tend to only pay attention to the host city and the game itself. For example, everyone knows that Rio hosted the 2016 summer Olympic Game; but probably only few people know that Chicago was also one of the candidate cities that bid for this game. Certainly this bidding process is an essential component of Olympic Games but often time overlooked by many people.

We therefore want to present the data for Olympic Game bid using interactive visualization techniques. This project can help people to have more comprehensive understanding of the Olympic Game.

Project Objectives.
As stated in the **Background and Motivation** section, people tend to pay attention to the Olympic Game host and overlook other candidate cities. we aim to present the Olympic Game bid data using interactive visualization techniques.
Our project will help answer the following questions:
1. How many cities participated the bids for each summer Olympic Game and what/where are they?
2. For each city that participated in the bids for Olympic Game, has it ever won the hosting right? How many times did it fail before won the right to host the game? What is the success rate?
3. What is the geographic distribution of the candidate cities? Which continent has more candidate cities?

The source webpages contain tables that tabulate the candidate city information for each game. However, the table is not interactive and lacks appropriate aggregate information. We will extract the data from the source webpages and visualize it using interactive techniques, including map and stack bar chart. For each summer Olympic Game, our design will interactively display the candidate city names, exhibit their geographic locations on the map, how many times they participated in the bids and how many times they won the hosting right, as well as some other well-designed aggregate information. We will also color the cities based on their continental location and our implementation might contain the option to zoom in to show each individual continent.

The benefits of our design:
1. The bids for Olympic Game is an integral component of this international event and our project will help people better understand the bidding process.
2. The original sources only contain tabulated data. We plan to use well-designed interactive techniques to visualize it and the users can view the complete candidate city information as well as their geolocations at a glance. Our design will also utilize carefully chosen color and transitions to better exhibit the data.

Data Source.
https://en.wikipedia.org/wiki/List_of_bids_for_the_Summer_Olympics
The data contained therein are presented as some HTML tables. Google Spreadsheet is the perfect tool [1] for us to easily scrape the tabulated data.
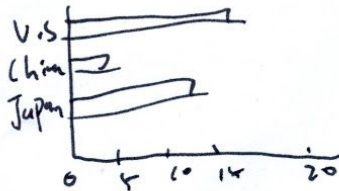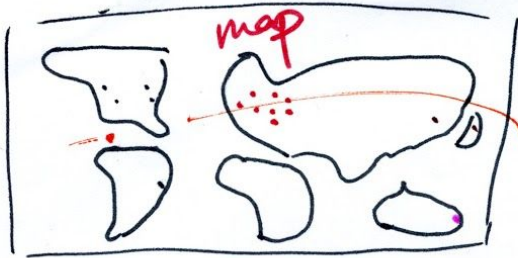
Data Processing.
- Mapping each country on a world map is a key part of our main features. However, the data doesn't come with the geolocation of the cities. We'll attach each city with an unique geolocation (longitude + latitude) in our later work.
- Secondly, we need to label the cities by continents, and this information hasn't been included in the original data. We'll add this information to the data set for the convenience of data labeling/filtering in the future implementation.
- Because it is needed to include a summary of the bid history for each city, which has not been included from the original data set. We'll aggregate these values, e.g. in which year this city participated the bids for summer olympic game? Did it win the bid? And what's its bid success rate?
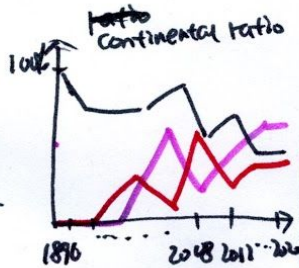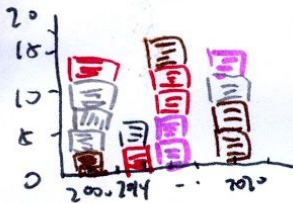
Visualization Design.
We followed the recommended "*Five Design Sheet Methodology*" to design our project. We first came up with a list of possible ideas and presented them in the Sheet #1. Then we organized and regrouped our ideas so when put together they can better represent the data. We generated three different initial layouts and they were presented in the Sheet #2/#3/#4. Finally we thought carefully about the merits/disadvantages of each possible combinations and finalized our design, which was presented in the Sheet #5.
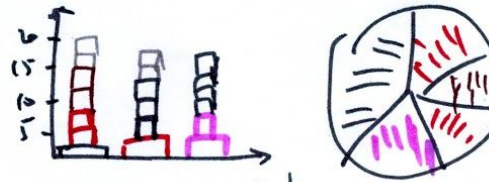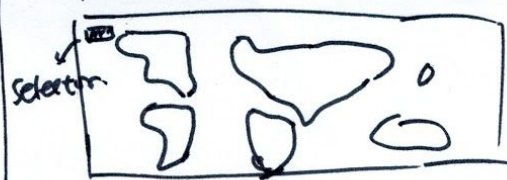
1. Ideas

map

Continental ratio for each game

V.S
China
Japan

0   5   10   15   20

ratio
Continental ratio
100%

1896   2048 2012 ... 2020

200..294 -- 2020

20
18
10
8
0

2. Filter   sheet #1.

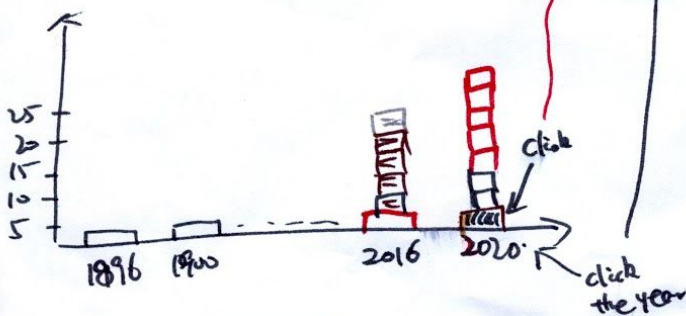Selection

5
15
10
5

3. Categorize

Category #1 : map,

Category #2 : bar chart.
                   pie chart.

4. Combine and Refine

Continent
Selector

Tokyo

click

1896 1900   2016   2020
click
the year

25
20
15
10
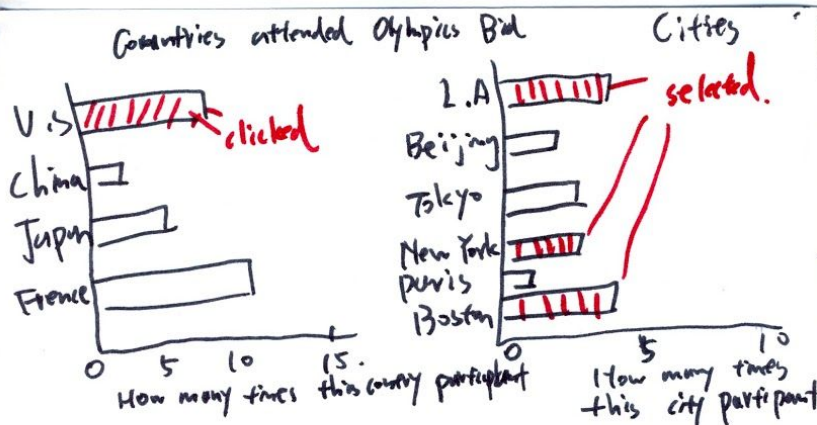5
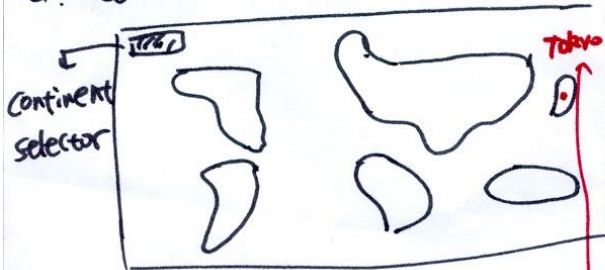
continental ratio for 2020 game.

5. Question.
The ideas fulfill
the needs of the
clients.

Countries attended Olympics Bid   Cities

Countries attended Olympics Bid

L.A ▭▭▭ ← selected.
Beijing ▭
Tokyo ▭▭
New York ▭
Paris ▭
Boston ▭▭

U.S ▭▭▭▭ ← clicked
China ▭
Japan ▭▭
France ▭▭

0   5   10   15.
How many times this country participant

0   5   10
How many times this city participant

Japan | 0.6 | China | France
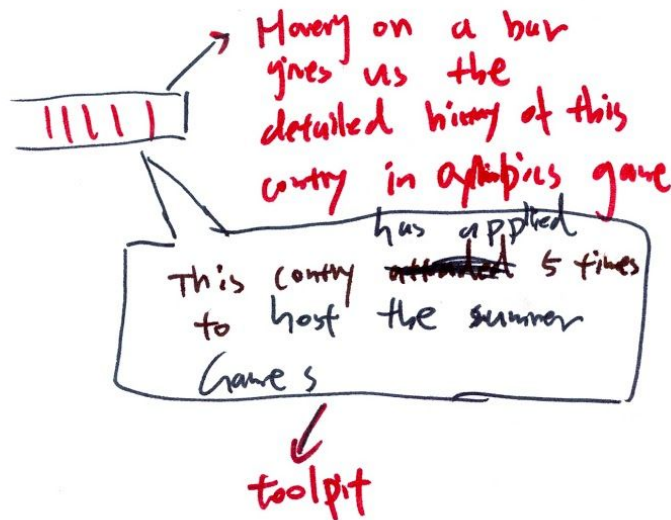
participation ration by country
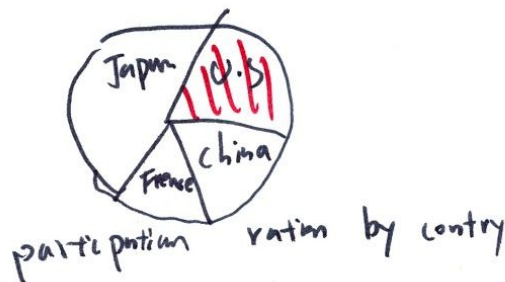
Title: Design #1
Author: Jiyuan Li, Ren Miao
Date: 10/16/2016
Sheet: #2
Task: final project proposal.

Operations:

- Clicking on bar gives the detail of this country of city, it includes in which year this city has applied to host the game and its success ratio
- Hovery on a bar gives us the detailed history of this country or city in olympic games

Focus:

Hovery on a bar gives us the detailed history of this country in olympics game

▮ | ▮ | ▮ |

This country attended has applied 5 times to host the summer Games
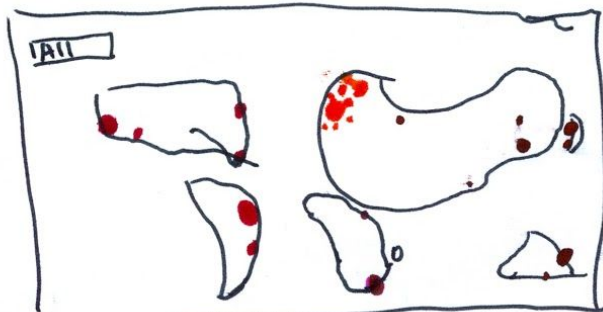
↓
toolpit

Discussion:
This design is fairly simple, and doesn't take advantage of time series from the original data

Visual Encoding:
We chose bar and pie charts because they are simple and easy to interpret the data.

Title: Design #2
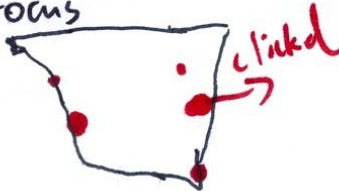Author: Jiyuan Li,
Ren Miao
Date: 10/16/2016
Sheet: #3
Task: final project proposal

All

Host ☐   Candidates ☐   all ☒

1984

Toronto L.A cape Town
Paris Tokyo

word cloud

Euro Asia
America
Australia

Participation ratio
by continent
A bigger circle
as more host cities
this continent has

Operations
- Continent selector to choose different region
- Year selector
- Filter cities by host, candidates or all

Focus

← clicked

New York
1984   2000   2004
☒   ☑   ☒
success rate
33.3%

hovered →

Euro Asia

shows all cities from this country that participat bids in this specific year

Discussion
This Design is informative
but the visualizations
are not creative.
Visual encoding:
- Words are informative and explicit to show the data.
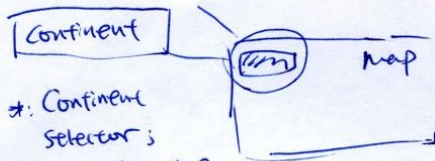- Map is interactive

Layout: → Continent selector.

Candidate Cities for each game. (chart #1)

1896 1900 1904   2008 2012 2016 2020

continental (ratio (percentage) (chart #2)
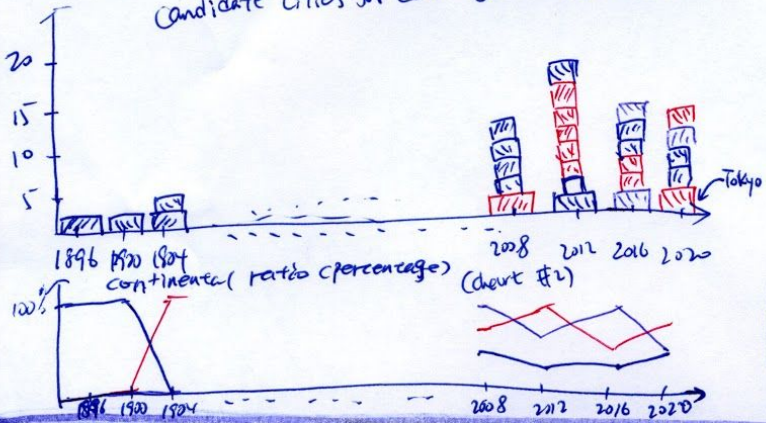
Focus:

*: Continent
Selector;
change the value
to see the candidate cities within each continent.
highlight

← chart #1.

* Hovering over a rectangle
a tooltip shows up to
show the city name
and country name.

Toyko.
Japan

?chang to Europe

Continent
selector
change value.

↳ highlight the
corresponding
line.

map
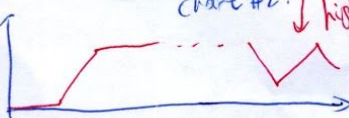
chart #2. → highlight Europe data.

Operations:

1. ~~Click~~ Hovering on a rectangle in chart #1, it will highlight the country in the world map;

2. Click on each year, ~~the~~ it will highlight the candidate countries in the map;

3. ~~change~~ change the continent selector value, it will zoom in and highlight the candidate cities located within the ~~con~~ continent. Also it will highlight the line in the line chart.

Discussion:

+: map is an essential part of the design; it can clearly show the geolocation of one or several cities.

+: ~~the~~ stack bar chart and the line chart are simple yet clear representation of the data; easy to follow and straightforward.

—: The map will show uneven city distribution. Historically more European and North American ~~cities~~ cities participated the bids; therefore Europe and North America will contain busier information. The user might need to zoom in to clearly see the cities in Europe and North America.
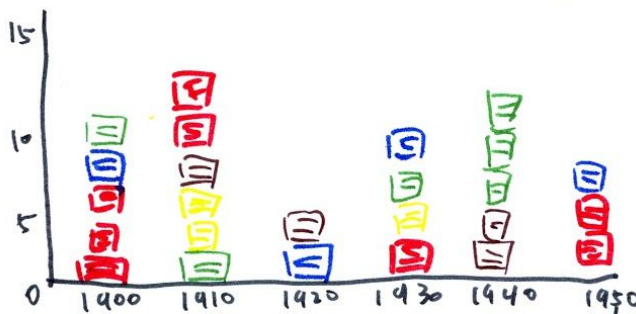
→ Continent Selector
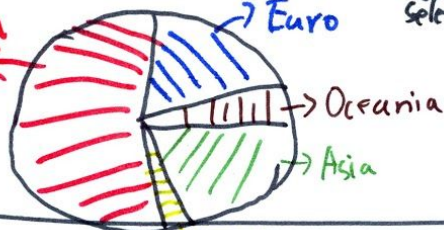
[All]

Clicked
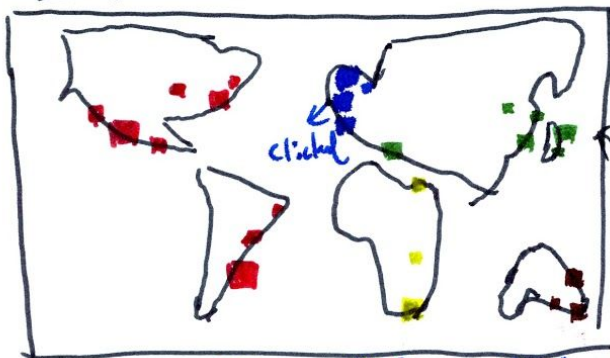
Paris   1900   1920   1950   success rate
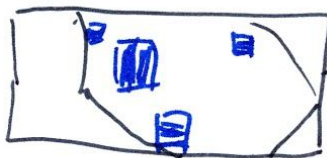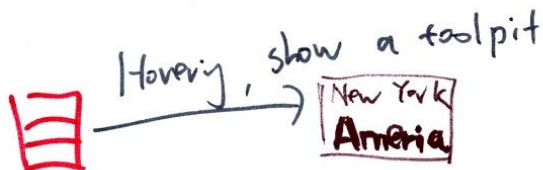   ☒   ☒   ☒   33.3%

America → Euro → Oceania → Asia → Africa (pie chart)

15
10
5
0
1900 1910 1920 1930 1940 1950

Selected

click
show

Focus/Zoom

→ Zoom In if a specific continent selected

Hovering, show a toolpit → New York America

Title: Final Design
Author: Ren Miao.
        Jiyuan Li
Date: 10/17/2016
Sheet: #5
Task: final project proposal.

Operations:
1. Hovering over a country on the map. tooltip shows aggregate info; (city)
2. clicking on a country on the map, it shows its success rate;
3. hovering over the rectangle highlights the city on the map and shows the city name in a tooltip;
4. Click the game year in the bar chart (x-axis), the pie chart shows the percentage of the candidate cities from each continent.

Detail
• Import world.json as map data;
• D3.js will be used to visualize the data;
• we will use approximately 3-4 weeks to build the project;
• . . . .

AS mentioned earlier, *Five Design Sheet Methodology* was followed to design our visualization project. Please refer to the attached design sketches for the details. The five pages of sketches attached are: sheet 1, brainstorm, in which we described many of our initial ideas; sheet 2, 3, 4, initial designs, in which we sketched three prototypes using different combinations of our initial ideas; sheet 5, realization design, in which we incorporated the best of our three designs.

The advantages of our choices of visual encoding:
- Map is the key part of our visualization design. To depict the geolocation information, map is the most straight-forward and convincing technique to use. On the map the visual channels we will use are: position (based on the geographic information of the cities), area in 2D (the size of the city token will be proportional to the number of the bids the city has participated in), color hue (the cities from the same continent will be rendered with the same color but different from other continents).
- We will use a stack bar chart to illustrate the candidate cities for each game. The x-axis will be the game year and the y-axis column for each game year will be a stack of rectangles, each of which represents a candidate city. The visual channels we will use are: position on common scale (on x-axis, different year will be shown on a different position; on a stack bar, different country will be depicted with a different position), color hue (on a stack bar, the rectangles representing the cities from the same continent will be rendered with the same color but different from other continents).
- We will use a pie chart to show the percentages of the candidate cities from each continents. In the pie chart we will use the following visual channels: color hue (different continents will be rendered in different color); area in 2D (the area of the continent will be proportional to the number of the candidate cities from this continent).

Must-Have Features.
- Display cities by their geolocation on a world map.
- Cities are presented as circles, the size of the circle is determined by the number of time participating in the bids for summer Olympics games.
- Filter cities by their regions, e.g. continent. The cities from the same region are rendered with the same color.
- Clicking on the city on the map give us the detail of their bids, including the years this city participated in the bids, how many times it won the right to host the game and the success rate.
- A stack chart that shows the complete history of bids for summer Olympics. X-axis will be the years. Each stack depicts the candidates and the host as stacking rectangles with different color based on the region(continent) they belong to.
- Hovering on the rectangle will display the city name and its country (and other aggregate information) in a tooltip.
- Host and candidates filter. Switching this filter on or off shows host or candidate cities only. This give the users more interactions with the data.

Optional Features.
- Hovering over a city on the map displays a tooltip that shows the number of times this city as a candidate since 1896.

- Hovering over a country on the map displays a tooltip showing the cities within the country that ever participated in the bids.
- A pie chart that shows the participation ratios of each continent for each Olympic Game. Appropriate transition will be applied if the user switches between different year.

Project Schedule.

Week	Plan
10/24 - 10/30	**Data Processing: Jiyuan, Ren**
10/31 - 11/06	**Interactive Map: Jiyuan** **Bids summary: Ren**
11/07 - 11/13	**Stack bar chart: Jiyuan** **Data filtering: Ren**
11/14 - 11/20	**Optional features if time permits:** **Pie Chart: Ren,** **Transition: Jiyuan**
11/21 - 11/27	**Process Book: Jiyuan, Ren**
11/28 - 12/01	**Wrap-up**

Reference
1. https://blog.ouseful.info/2008/10/14/data-scraping-wikipedia-with-google-spreadsheets/