Quantitative Data Visualization

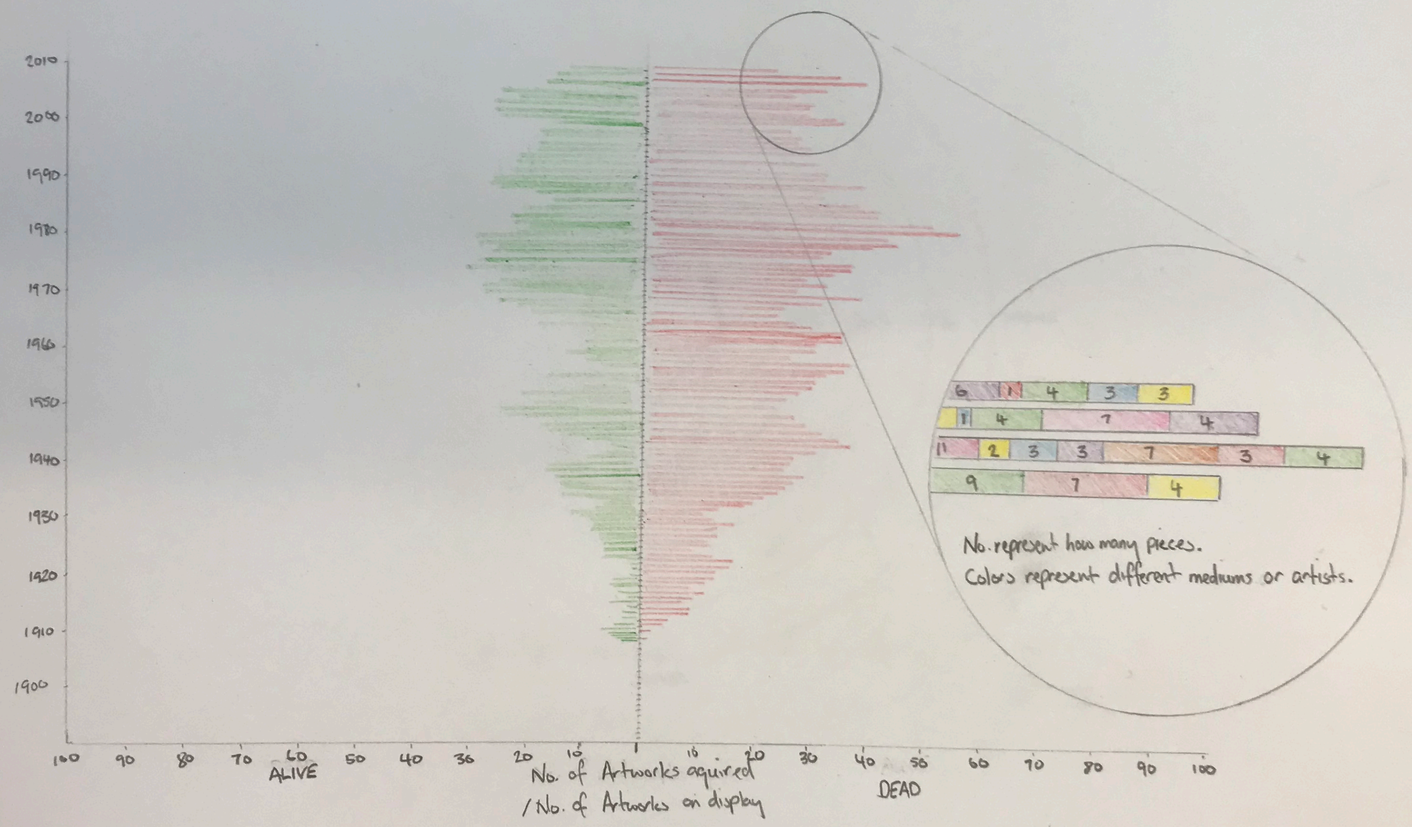
Neil Oliver

Getting an idea

- During his lifetime, Van Gogh was a failed and poor artist. He created more than 2,000 works of art, but sold just two when he was alive.

- In 1990, his portrait of Dr. Gachet sold for $82.5 million, making it one of the most expensive paintings ever sold.

Is artwork more likely to be considered ‘museum worthy’ only after an artist has died?


2010
2000
1990
1980
1970
1960
1950
1940
1930
1920
1910
1900
100 90 80 70 60 50 40 30 20 10 10 20 30 40 50 60 70 80 90 100
ALIVE
DEAD
No. of Artworks aquired
/ No. of Artworks on display
6 1 4 3 3
1 4 7 4
1 2 3 3 7 3 4
9 7 4
No. represent how many pieces.
Colors represent different mediums or artists.

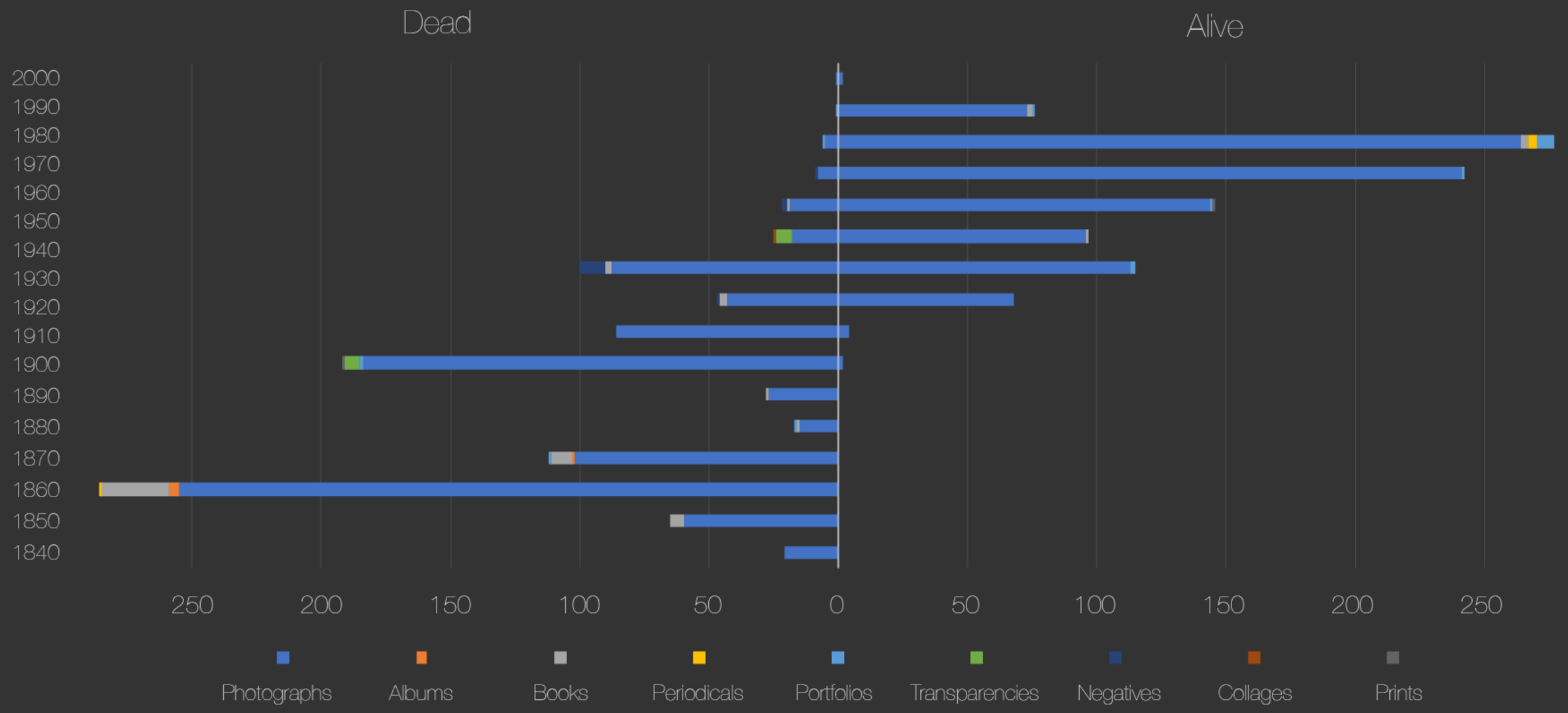
Reflections on the sketch

- With so many lines, the only way to see details would be through a zoom / interactivity.
- Representing each artist would take hundreds of different colors.
- Overall, need to focus the dataset.

One lesson with the MET API, some glue, sticky tape and glitter later…

Okay so I replaced the glue and glitter with Excel and Illustrator…

Acquisitions by the photography department in relation to the artists life time.



Reflections on the initial design

- Focused on a single department with a good number of different mediums
- Used real data which did start to tell a story
- Dominated by one medium

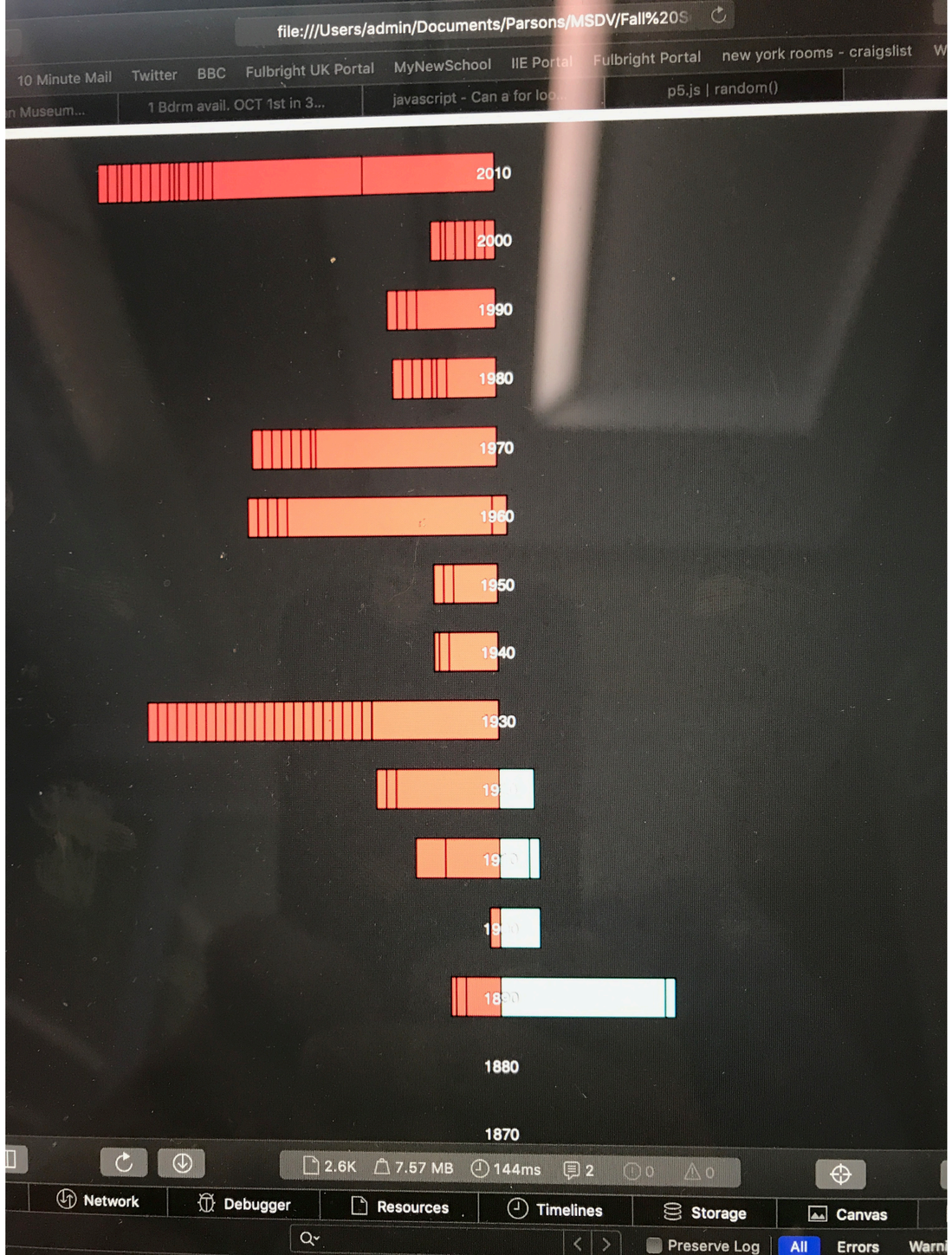
One lesson on .P5, a few late nights and the occasional beer later…

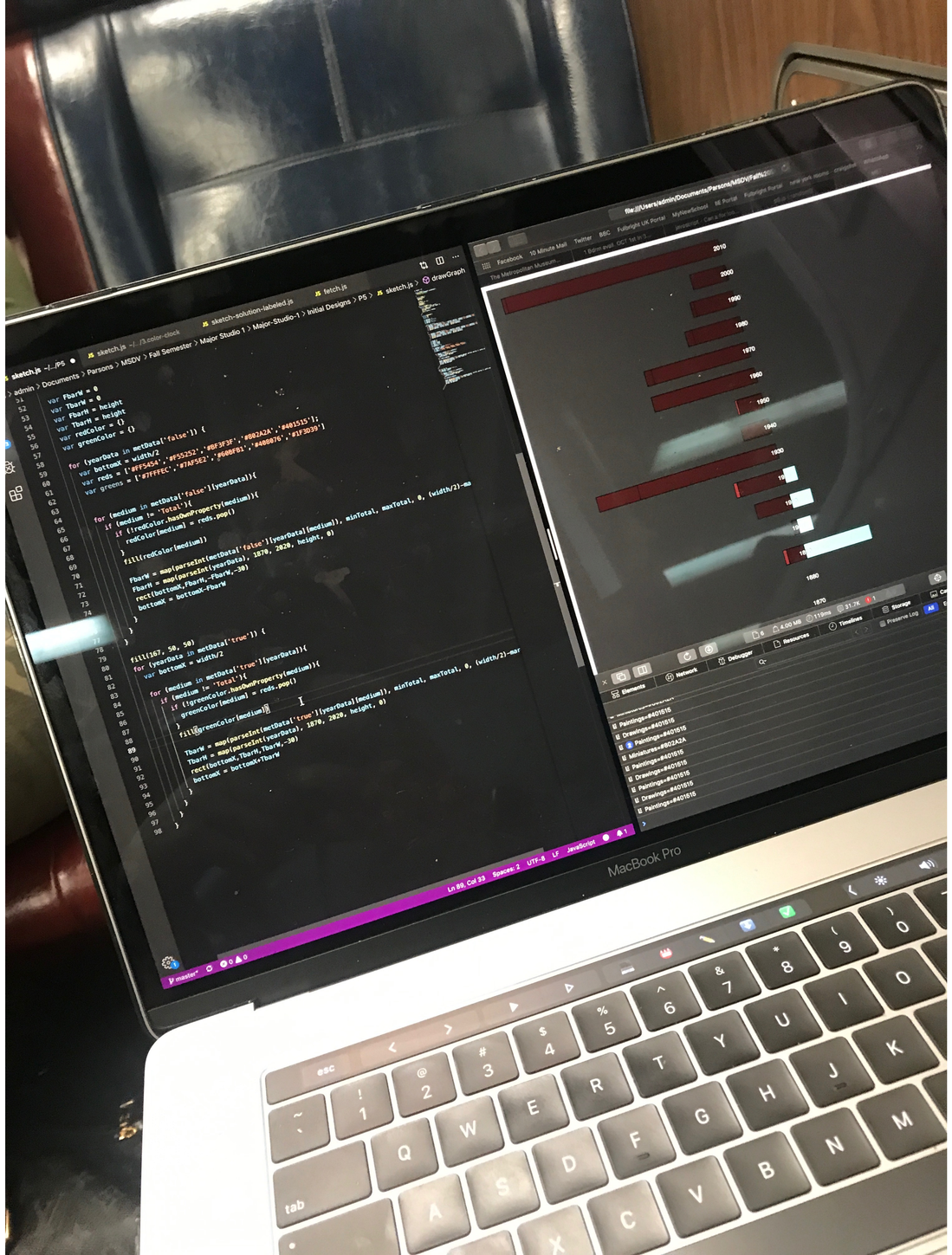
```
jectBaseUrl + n;

)
- data.json())
- {
ata); //for live use
g(data)
Data.push(data) //for offline use
.log(requestData.length)

uestData.length == objectIDs.length) {
sole.log("saving file")
nload(JSON.stringify(requestData), "METPaintings.json", '"application/json"');
{
sole.log("not finished getting objects")



lly
content, fileName, contentType) {
nt.createElement("a");
Blob([content], {type: contentType});
reateObjectURL(file);
ileName;


File(){
n METPaintings) {
ETPaintings[painting]);


t(objectDate){
ull;

Int(objectDate) // need a more robust method
eInt(bracket/10, 10)+1)*10


array using just the data you need
Data(data){
ta.medium;
= data.objectDate;
Date = data.artistBeginDate;
te = data.artistEndDate;
Date = data.creditLine;
it aliveAtAcquisition(artistBeginDate, artistEndDate, acquisitionDate, objectDate)

= acquisitionDate.split(',')
wait dateBracket(acquisitionDate[acquisitionDate.length-1])
```

Ln 37, Col 26 Spaces: 4 UTF-8 LF JavaScript 1

MacBook Pro

Boring my classmates with random photo updates…


file:///Users/admin/Documents/Parsons/MSDV/Fall%20S
10 Minute Mail
Twitter
BBC
Fulbright UK Portal
MyNewSchool
IIE Portal
Fulbright Portal
new york rooms - craigslist
1 Bdrm avail. OCT 1st in 3...
javascript - Can a for loo...
p5.js | random()
2010
2000
1990
1980
1970
1960
1950
1940
1930
1880
1870
2.6K
7.57 MB
144ms
Network
Debugger
Resources
Timelines
Storage
Canvas
Preserve Log
All
Errors

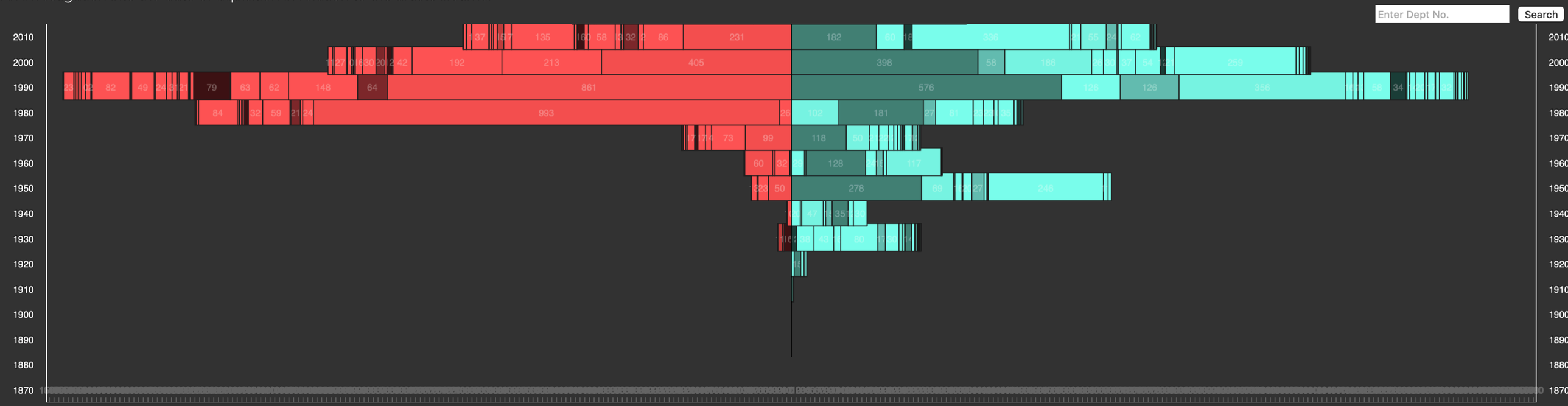

MacBook Pro

Design V.1

The MET Museum

Do we only value art after the artist is deceased?

An investigation into the date of aquisition in relation to the artists lifetime.



Reflections on V.1

- The design is a lot more interesting to look at and explore with a larger dataset (modern art department)
- Investigating which mediums were acquired each decade doesn’t really add value (although it looks nice).
- Different departments have very different stories to tell.
- The loading of such a huge dataset takes ages (over 20 minutes!).
- Most people are still unclear on what the graph actually shows.

Decisions based on reflections

- The larger data set looks good
- Displaying which departments purchase art from dead/alive artists would be much better than which mediums
- Offline data is needed if it is to be used in a website format.

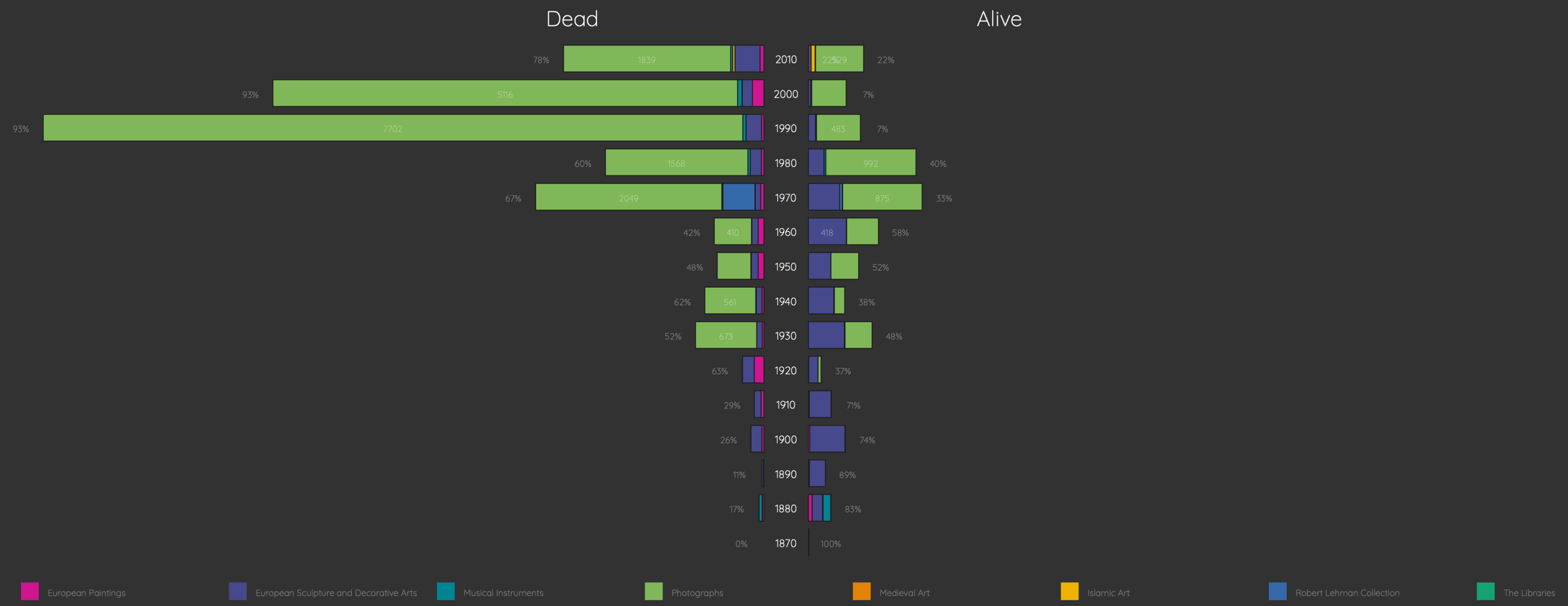
2 days and over
1,000,000 requests
later…

Do we only value art after the artist has died?

An investigation into the date of aquisition in relation to the artists lifetime.

From the opening on the MET Mueseum in 1870, have the acquisition habbits of the museum shifted between buying historical art to buying art from current artists? The visualization below maps the museums acquisition per decade in relation to if the artist was alive or deceased. The data used only includes acquisions of artists that were still alive at the time of the opening of the museum.

Display Percentages



Disclaimer: Art data isn't perfect. Sometimes we dont know the exact date an object was created or when an artist was born or died. Out of the 7530 items purchased by the MET museum since 1870, a total of 930 were unable to be put into a category.

Reflections on Final Design

- Removing the side and bottom axis and labeling the individual data points looks much cleaner while maintaining the same level of detail.

- Using the same colors for each side help show the same departments easily while not obstructing the difference between dead or alive

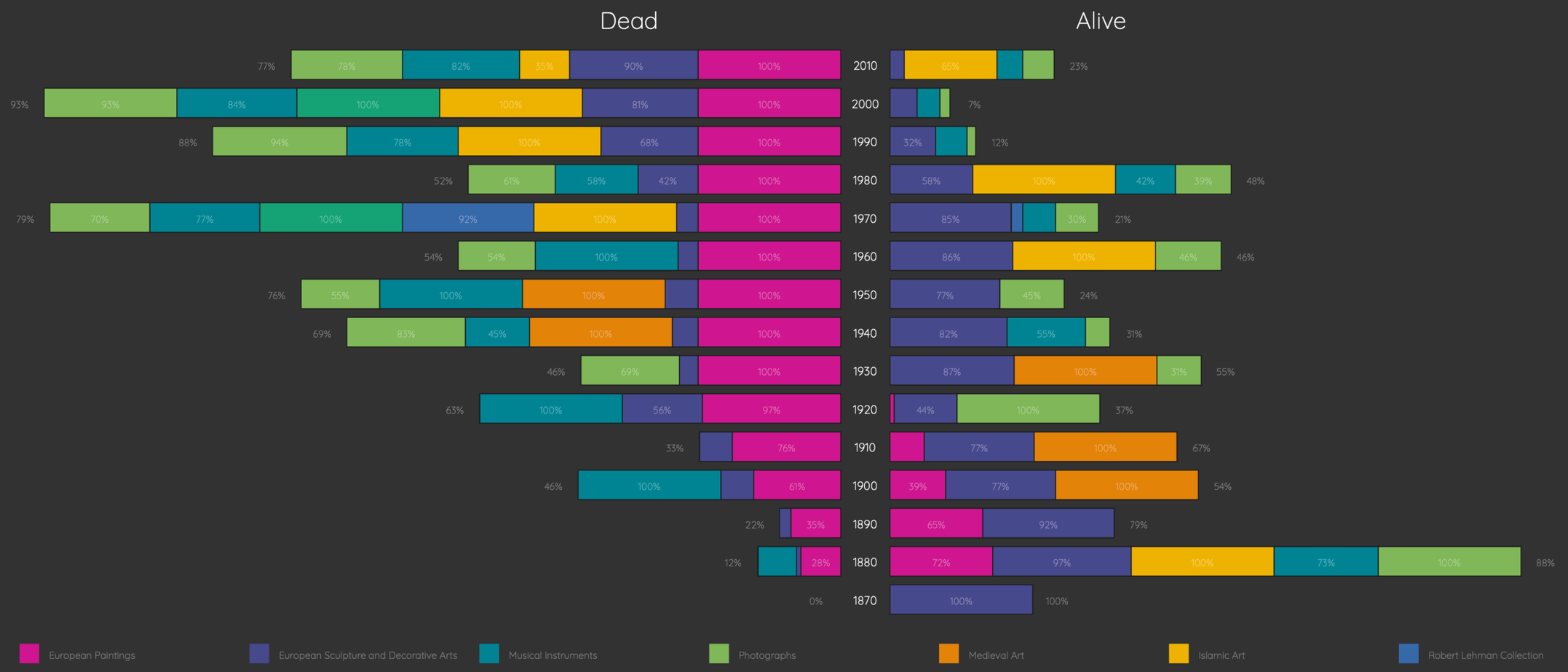
- One department (photography) dominates the number of items acquired. A better representation would have been the percentage values for each department.

Do we only value art after the artist has died?

An investigation into the date of aquisition in relation to the artists lifetime.

From the opening on the MET Mueseum in 1870, have the acquisition habbits of the museum shifted between buying historical art to buying art from current artists? The visualization below maps the museums acquisition per decade in relation to if the artist was alive or deceased. The data used only includes acquisions of artists that were still alive at the time of the opening of the museum.

☑ Display Percentages



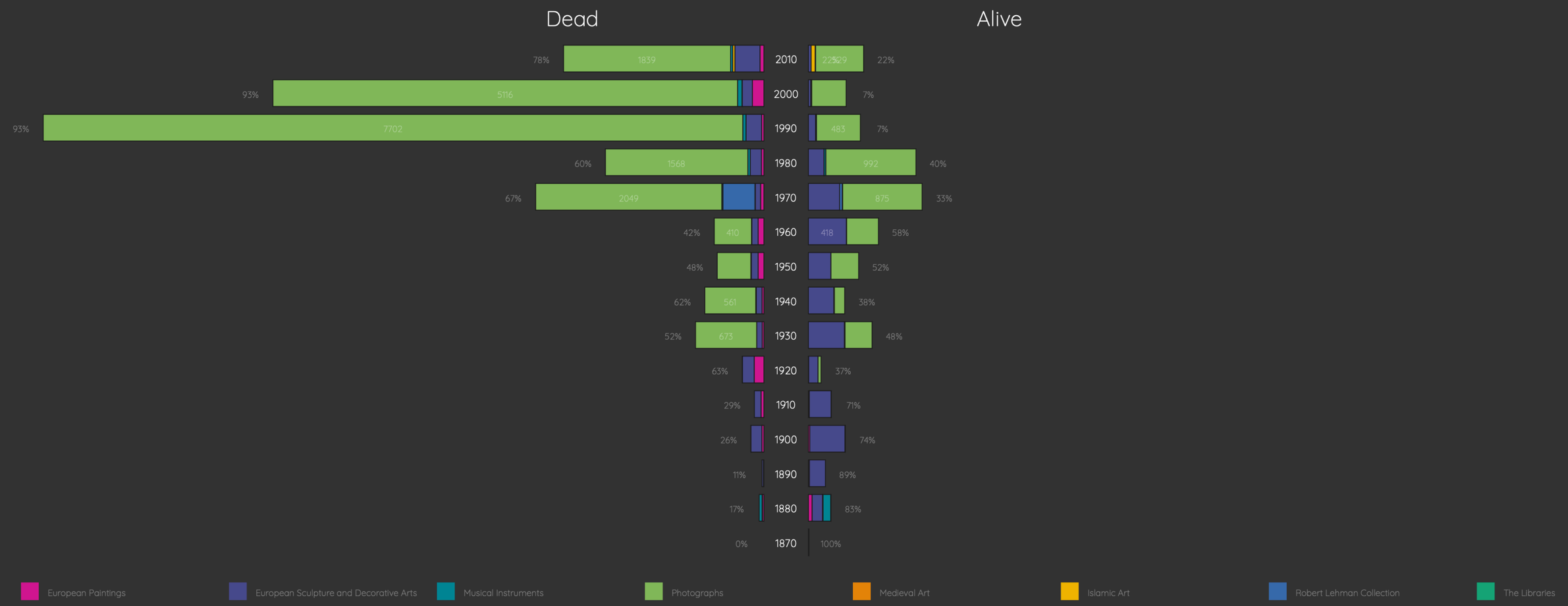
Disclaimer: Art data isn't perfect. Sometimes we dont know the exact date an object was created or when an artist was born or died. Out of the 7530 items purchased by the MET museum since 1870, a total of 930 were unable to be put into a category.

Do we only value art after the artist has died?

An investigation into the date of aquisition in relation to the artists lifetime.

From the opening on the MET Mueseum in 1870, have the acquisition habbits of the museum shifted between buying historical art to buying art from current artists? The visualization below maps the museums acquisition per decade in relation to if the artist was alive or deceased. The data used only includes acquisions of artists that were still alive at the time of the opening of the museum.

Display Percentages



Disclaimer: Art data isn't perfect. Sometimes we dont know the exact date an object was created or when an artist was born or died. Out of the 7530 items purchased by the MET museum since 1870, a total of 930 were unable to be put into a category.

The process

- The coding process changed dramatically throughout each stage of the development.

- The next slides contain an brief overview of the three main areas (scraping, processing & design) of the final solution; with more in depth details included on GitHub.

Scraping

- Using the browser to get all of the data and then save it was problematic as it had to receive all of the information before creating a save file.

- Node.js was used as it could incrementally save each object that was received. However even the server side solution would time out regularly and the process would have to start again.

- An adapted approach was used to keep track of which objects had been requested and received, and combine with setTimeOut() the process was managed and successfully saved all of the data.

- This approach also also the saved data to be updated very quickly, requesting only new unsaved items.

- It took over 12 hours non stop to receive the data and save it to a JSON file.

Processing

- The processing of the MET collection was done using vanilla javascript. It processes the whole database and produces a smaller ‘processed’ javascript file to be used with P5.

- The processing extracts the artists information (known as artist start and end date) and make assumptions for missing data (such as if someone without an ‘end’ date was born more than 120 years ago, we can be sure that they are dead.

- All items where the artist had died before the opening of the MET in 1870 were discounted to avoid a large skew of the data.

- The acquisition date is compared to the artist start and end date and totaled and sorted into two categories within an object. The information is further broken down into each decade (which had to be calculated) and then department.

- The design is totally adaptable, changing a single line of code would let it display a single department and either the individual artists or mediums (I would like for this to be a future interactive feature).

- The process is repeated and the split percentage (between the ‘dead’ and ‘alive’ category) are are calculated for each department.

- The information is saved into a separate file so the processing is not done every time the HTML file is loaded.

Visuals

- The design was created in P5 without any templates or examples. It was designed it from the ground up only using the official P5 documentation.
- It is designed as two stacked bar charts displayed back to back.
- The colors and labels are assigned dynamically and will adapt to different data sources.
- The visual is instantly switchable using the checkbox to switch between the total values and percentage values.
- Additional dynamic details are inserted into the DOM outside of the P5 canvas.