

Final Project Proposal

(due March 10th 7:00p.m)

This document outlines the guidelines for the project proposal. You can start working on the project once your proposal is accepted and graded by your TA on gradescope. The entire final project is worth 25% of your final grade and the proposal takes account for **5%**. There is no late-submission on the proposal.

Submission Guideline

Download this google doc, fill the table and 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](#)

Project Proposal

	Description
Project Topic	Artwork Analytics
Dataset Description	<p>Provide 1) the list of attributes and 2) a single item in the dataset as an example.</p> <p>This project will utilize multiple datasets composed in .csv format, including objects.csv, objects_dimensions.csv, constituents.csv, objects_constituents.csv, etc.</p> <p>objects.csv:</p> <ol style="list-style-type: none">1) ['objectid', 'accessioned', 'accessionnum', 'locationid', 'title', 'displaydate', 'beginyear', 'endyear', 'visualbrowsertimespan', 'medium', 'dimensions', 'inscription', 'markings', 'attributioninverted', 'attribution', 'provenancetext', 'creditline', 'classification', 'subclassification', 'visualbrowserclassification', 'parentid', 'isvirtual', 'departmentabbr', 'portfolio', 'series', 'volume', 'watermarks', 'lastdetectedmodification', 'wikidataid', 'customprinturl']. Brief explanation of each attribute can be found in Data Dictionary.txt.2)

```

objectid 0
accessioned 1
accessionnum 1937.1.2.c
locationid NaN
title Saint James Major
displaydate c. 1310
beginyear 1310.0
endyear 1310.0
visualbrowsertimespan 1300 to 1400
medium tempera on panel
dimensions painted surface (top of gilding): 62.2 x 34.8 ...
inscription NaN
markings NaN
attributioninverted Grifo di Tancredi
attribution Grifo di Tancredi
provenancetext By 1808 in the collection of Alexis-François A...
creditline Andrew W. Mellon Collection
classification Painting
subclassification NaN
visualbrowserclassification painting
parentid 34.0
isvirtual 0
departmentabbr CIS-R
portfolio NaN
series NaN
volume NaN
watermarks NaN
lastdetectedmodification 2022-05-16 22:01:19-04
wikidataid Q20172973
customprinturl NaN

```

objects_dimensions.csv:

- 1) ['objectid', 'element', 'dimensiontype', 'dimension', 'unitname']
- 2)

```

objectid 5
element painted surface
dimensiontype width
dimension 45.2
unitname centimeters

```

constituents.csv:

- 1) ['constituentid', 'ulanid', 'preferreddisplayname', 'forwarddisplayname', 'lastname', 'displaydate', 'artistofnngaobject', 'beginyear', 'endyear', 'visualbrowsertimespan', 'nationality', 'visualbrowsernationality', 'constituenttype', 'wikidataid']
- 2)

constituentid	9
ulanid	NaN
preferreddisplayname	Anonymous
forwarddisplayname	Anonymous
lastname	NaN
displaydate	Anonymous
artistofngaobject	0
beginyear	NaN
endyear	NaN
visualbrowsertimespan	NaN
nationality	NaN
visualbrowsernationality	Other
constituenttype	anonymous
wikidataid	NaN

objects_constituents.csv:

- 1) ['objectid', 'constituentid', 'displayorder', 'roletype', 'role', 'prefix', 'suffix', 'displaydate', 'beginyear', 'endyear', 'country', 'zipcode']
- 2)

objectid	8
constituentid	12
displayorder	7
roletype	owner
role	current owner
prefix	NaN
suffix	NaN
displaydate	1937 -
beginyear	1937.0
endyear	NaN
country	United States
zipcode	20565

object_associations.csv:

- 1) ['parentobjectid', 'childobjectid', 'relationship']
- 2)

parentobjectid	75981
childobjectid	76031
relationship	inseparable

Update Data Set Usage:

dimension.csv:

- 1) ['objectid', 'element', 'width', 'title', 'beginyear', 'attribution',

'classification', 'isvirtual', 'height']

2)

objectid	6
element	painting surface
width	33.6
title	Saint Anthony Abbot [left panel]
beginyear	1354.0
attribution	Puccio di Simone and Allegretto Nuzi
classification	Painting
isvirtual	False
height	89.3

constituent.csv

1) ['objectid', 'classification', 'constituenttype', 'beginyear']

2)

objectid	33
classification	Painting
constituenttype	individual
beginyear	1310.0

object.csv

1) ['objectid', 'classification', 'beginyear']

2)

objectid	1
classification	Painting
beginyear	1333.0

country.csv

1) ['objectid', 'country', 'latitude', 'longitude']

2)

objectid	11
country	United States
latitude	37.09024
longitude	-95.712891

assocaitions.csv

1) ['childtitle', 'parenttitle', 'parentobjectid', 'childobjectid', 'relationship']

2)

childtitle	The Crucifixion with the Virgin, Saint John, S...
parenttitle	The Crucifixion with the Virgin, Saint John, S...
parentobjectid	206127
childobjectid	29
relationship	separable

Dataset Link	https://github.com/NationalGalleryOfArt/opensource Update on Actual Project: Data cleaning and Merging is done on a separable notebook file to output the imported csv file is d3.js.
Why did you choose this particular dataset? What kind of story you aim to deliver (e.g “Sales analysis of company xyz”)	Hint) You can refer to the storytelling lecture slides. <ul style="list-style-type: none"> - The reason for selecting this particular dataset was simply because of my personal fascination towards art pieces. - Perform an exploratory analysis to find out whether there exists a trend in the amount of artwork across different categories and countries.
1 plot with 0 Key and 2 values	i) Question you are asking from this graph. <ul style="list-style-type: none"> - What is the biggest artwork in terms of physical dimensions? And when was it produced? - A condition(i.e. Virtual artwork, Painting Artwork, Print Artwork, etc.) may be added due to scalability limit in scatter plot along with buttons to switch between. ii) Columns you are going to use <ul style="list-style-type: none"> - ['objectid', 'beginyear', 'dimension', 'classification', 'isvirtual'] from objects.csv. - ['objectid', 'element', 'dimensiontype', 'dimension', 'unitname'] from objects_dimensions.csv - Update: ['width', 'title', 'beginyear', 'classification', 'height'] from dimension.csv. iii) Type of graph <ul style="list-style-type: none"> - Scatter Plot

<p>1 plot with 1 key and 1 value</p>	<p>i) Question you are asking from this graph.</p> <ul style="list-style-type: none"> - What is the quantity of artwork produced anonymously for each category? <p>ii) Columns you are going to use</p> <ul style="list-style-type: none"> - ['constituentid','constituenttype'] from constituents.csv. - ['constituentid', 'objectid', 'beginyear'] from objects_constituents.csv. - ['objectid', 'classification'] from objects.csv. - Merging and Aggregation will be performed to compute the value. - Key: year ('beginyear') - Update: ['classification', 'constituenttype', 'beginyear'] from constituent.csv. <p>iii) Type of graph</p> <ul style="list-style-type: none"> - Bar Chart <p>iiii) Backup for myself</p> <ul style="list-style-type: none"> - View the distribution of artists' ages across different categories. - Use a line graph to capture the trend in the amount of artwork created.
<p>1 plot with 2 keys and 1 value</p>	<p>i) Question you are asking from this graph.</p> <ul style="list-style-type: none"> - Which country produces the most artwork in terms of paintings, sculpture, drawings, and photographs? <p>ii) Columns you are going to use</p> <ul style="list-style-type: none"> - ['objectid', 'classification'] from objects.csv. - ['objectid', 'country'] from objects_constituents.csv. - Update: ['objectid', 'classification', 'beginyear'] from object.csv. <p>iii) Type of graph</p> <ul style="list-style-type: none"> - Stacked Bar Chart
<p>1 geometric visualization</p>	<p>i) Question you are asking from this graph.</p> <ul style="list-style-type: none"> - Which country holds the majority of the artworks present in this exhibition? <p>ii) Columns you are going to use</p> <ul style="list-style-type: none"> - ['objectid', 'constituentid', 'country', 'beginyear'] from objects_constituents.csv; attributes from world.json or

	<p>anything similar to that to outline the world map.</p> <ul style="list-style-type: none"> - Value: count of art work per country. - Update: ['objectid', 'country', 'latitude', 'longitude'] from country.csv. <p>iii) Type of graph</p> <ul style="list-style-type: none"> - Choropleth
1 visualization from - box plot, node-link diagram, adjacency matrix	<p>i) Question you are asking from this graph.</p> <ul style="list-style-type: none"> - Which artwork is most closely related to other pieces, and what are those connections? <p>ii) Columns you are going to use</p> <ul style="list-style-type: none"> - ['objectid', 'title', 'beginyear'] from objects.csv - ['parentobjectid', 'childobjectid', 'relationship'] from object_associations.csv. - Update: ['parenttitle', 'childtitle', 'relationship'] from association.csv. <p>iii) Type of graph</p> <ul style="list-style-type: none"> - Node-Link Diagram - Update: Adjacent Matrix
1 interactivity using Buttons	<p>Describe in which visualization you plan to add the button-related interactivity</p> <ul style="list-style-type: none"> - Scatter Plot. The buttons (left and right) are added so that the audience can switch between the dimension distributions of different categories. - Update: Scatter Plot, Bar Plot - Use to switch graphs between selected categories.
1 interactivity using Tooltips (Display data on hover).	<p>Describe in which visualization you plan to add a tooltip.</p> <ul style="list-style-type: none"> - Choropleth. Mouse over around a country will display the number of artwork produced in this country and the year of the earliest produced artwork. - Update: Scatter Plot, Choropleth - Display the actual value.

1 interactivity using Animation.	<p>Describe 1) what type of animation you plan to add and 2) in which visualization you plan to add.</p> <ul style="list-style-type: none"> - Click to expand the tree, similar to the example provided in lecture 16. Click on the un-expanded node to expand and visualize its related node, and click on an expanded node to collapse the expansion. - Node-Link Diagram - Update: Bar Chart, Stack Bar Chart - Animation in initializing the bar charts.
1 interactivity not learned in class	<p>Describe 1) what type of animation you plan to add and 2) in which visualization you plan to add.</p> <ul style="list-style-type: none"> - The audience can use checkboxes or dropdown menus to indicate their preferred artist for obtaining additional information. For instance, change “anonymously” to a specific artist or change “2020” to another year. - Bar Chart - Update: Scatter Plot - Drop Down button to switch between graphs.
Any creative form of plot you want to try for the five you selected above? (e.g. pictogram)	<p>Hint) You can refer to the storytelling lecture slides. Note) This is going to be for extra credit.</p> <ul style="list-style-type: none"> - Pictogram in Bar Chart to represent each category.

Final Project:

- Update on data set usage are marked as **Update** accordingly in the above table.