Project 2

* Chandan, Nishank, Telly

## **Overview:**

Visual art provides a rich source of data. We can study the composition of a work of art (e.g., the structure, contents, colors, sizes, etc) as well as more qualitative components of the work (e.g., the emotions it allows us to feel, our preference for a piece, etc.). While many components of composition can be extracted using computer vision, others require human expertise and experience to measure.

Mohammad & Kiritchenko ran a large scale study to understand the emotions and preferences viewers have for different pieces of art using the WikiArt dataset (<https://www.wikiart.org/>). They asked people to provide different ratings about their subjective experiences with 4,105 pieces in this corpus and collected those experiences through a series of measures. More details of their experiment are at <http://saifmohammad.com/WebPages/wikiartemotions.html>.

I have cleaned up the data from this experiment in WikiDataClean.csv, reducing it to 3,646 pieces with full data and numeric years (note that you are welcome to download and use the full dataset for an additional challenge). In this file, you will see a series of columns containing the following data:

* Style: What stylistic period was the piece from?
* Category: What artistic movement is the piece a part of?
* Artist: Who created the piece?
* Title: What is the piece called?
* Year: When was the piece created?
* Image URL: Link to the piece in WikiArt
* Painting Info URL: Link to details about the piece
* Artist Info URL: Link to details about the artist
* Is painting: Is the piece a painting or something else (e.g., a sculpture)?
* Face or body: Does the piece contain a face, a body, or no human figures (note that face takes priority over body)?
* Mean rating: How did the annotators like the piece on average (-3 (strongly dislike) to 3 (strongly like))?

The remaining columns represent terms that annotators could provide about the piece. In each cell of the table, you will find the average annotation for that particular term. For example, if 4 of 10 annotators labeled an image as evoking trust, Trust would contain the value 0.4 for that image.

Your goal in this project is to create visualizations that illustrate interesting features of the dataset. You can do this in one of two ways:

1. Create at least two interactive visualizations that let users explore the data in interesting ways.
2. Create a website that tells a story about this data through at least four different visualizations paired with explanatory text.

## **Option 2: Storytelling with Data**

### **Minimum Requirements:**

This option will allow you to create a website that illustrates key aspects of the data to tell a story about visual art using the results from Mohammad & Kiritchenko's experiment. The story should include a series of visualizations and texts that together form a cohesive narrative about the data.

The visualizations you create as part of this option can be static: you do not have to render them in the browser, but can create visualizations using any combination of tools (e.g., Photoshop, Excel, Tableau, Illustrator, D3, etc) and display them as static images. Your website could be a single page or multiple pages, but should allow for intuitive navigation and a natural story flow.

Some examples of professional storytelling visualization projects include:

* <https://www.newyorker.com/tech/annals-of-technology/mapping-new-york-noise-complaints>
* <http://stars.chromeexperiments.com/>
* <http://archive.nytimes.com/www.nytimes.com/interactive/2012/08/04/sports/olympics/bob-beamons-long-olympic-shadow.html?_r=2>
* <http://www.r2d3.us/visual-intro-to-machine-learning-part-1/>

<https://twooctobers.com/blog/8-data-storytelling-concepts-with-examples/> has a nice set of principles for good data storytelling.

If you choose this option, your project must:

* Create a website containing four visualizations that illustrates interesting features of the data.
* Include at least four unique visualizations:
  + One visualization must include some quantitative data
  + One visualization must include categorical data
  + Your visualizations should support at least one meaningful comparison between related data attributes
  + Your visualizations should each be accompanied by text explaining what the visualization shows. Note what tasks the visualization allows you to accomplish to derive this insight and how your design is tailored to support these tasks.
* Include as a part of your website a section or page with the following details:
  + Information about your data preprocessing and design as appropriate.
  + Your design process (e.g., how did you go about designing, building, and refining your system? Why did you choose these representations?)
  + Any additional elements of the project you consider "above and beyond" the basic requirements.
  + Your team roles for each individual

### **Above and Beyond:**

The above requirements are the minimum for a passing grade on this project. Some ideas to improve your project include (but are not limited to):

* Unusual Representations: Draw on some of the examples from class to represent data in ways beyond a typical scatterplot or bar chart.
* Style: Keep the style consistent across all your views, with an eye towards intelligently applying visual design.
* Interesting Tasks: Derive insight into the data. Highlight these insights in your readme and describe how the visualization enables them. What do these insights tell us about the data?
* Perceptually-Informed Design: Integrate perceptual concepts into your visualization design and discuss how you've integrated those concepts.
* Integrate Imagery: Use data from the provided URLs to add more context to your visualization. For example, you could link to or include text about individual paintings or artists or integrate images of the art into your visualizations.
* Coordinated Views: Use visualizations that complement one another in meaningful ways to tell a deeper story than they otherwise could on their own.

### **Platforms:**

You can use any tools or platforms you'd like to create your visualizations. Your webpage should run in the browser without me having to install any additional software.

## **Submissions:**

All submissions must be made through GitHub with a timestamp by 11:59pm on 12.2. Your submission files should include:

* Your code and/or website
* Your README (for Option 1)

### **Plan:**

A cohesive narrative about the data.

Intuitive navigation and a natural story flow.

Find interesting features of the data.

1. Over the year, category dist, or counts of pieces
2. Within the year, show the mean values for the year
3. Split the data into the paining or not?
4. Rating values for the year?