

# Data Appendix

## Section 1: MetMuseumData.csv

Unit of Observation: Each row in the data frame represents a single unique European art piece (Painting or Drawing) found within the MET Museum.

- In the dataset used there were **804** total art pieces. According to the type of art piece there were **399** paintings and **405** drawings.
- This joined table was created by merging the paintings and drawings dataframes together

### Subsection 1.1 Year

- Year from which a specific art piece was created. We used ~40 drawings and ~40 paintings for each decade between 1800 - 1900.
  - Type: Integer
  - Range: 1800 - 1899
  - Example: 1871

	Year
Mean	1846
Min	1800
Max	1899

Distribution: The distribution of dates is uniform as ~80 total art work pieces were used each decade across the longer 100 year time period.

### Subsection 1.2 URL

- Stores the URL that references the photograph of the artwork. Can be used to fetch the images associated with a given observation.
  - Type: Object

- Example: <https://images.metmuseum.org/CRDImages/ep/original/64090.jpg>



### Subsection 1.3 Type

- Stores information about the type of artwork. Each artwork within the dataframe can either be 'drawing' or 'painting'
  - Type: String
  - Example: 'drawing'

### Subsection 1.4 Label

- A binary label for each row in the dataframe that can be either 1 or 0.
  - 1 -> Drawing
  - 0 -> Painting
  - Type: Integer

### Subsection 1.5 Era

- Stores information about the Era from which each piece originates. Era's are organized into decade time frames. Era include 1800s, 1810s, 1820s ... 1890s.
  - Type: String
  - Example: 1830s

- Distribution: Uniform with equal counts across each Era (~40 for each Era for each type of artwork).

#### Subsection 1.6 Image\_path

- Stores the path to find the image associated with a specific artwork entry. All images are located in the images folder which can be found within the DATA folder.
  - Type: String
  - Example: “../DATA/images/img\_17.jpg”