

03_normalize

March 21, 2022

1 Normalizing Data

Abid Ali

<https://github.com/abid-2362>

```
[ ]: import pandas as pd
data = pd.read_csv('./data/artwork_sample.csv')
```

```
[ ]: data.head()
```

```
[ ]:
      id accession_number      artist artistRole  artistId \
0  1035          A00001  Blake, Robert    artist        38
1  1036          A00002  Blake, Robert    artist        38
2  1037          A00003  Blake, Robert    artist        38
3  1038          A00004  Blake, Robert    artist        38
4  1039          A00005  Blake, William    artist        39

                                     title      dateText \
0  A Figure Bowing before a Seated Old Man with h...  date not known
1  Two Drawings of Frightened Figures, Probably f...  date not known
2  The Preaching of Warning. Verso: An Old Man En...    ?c.1785
3  Six Drawings of Figures with Outstretched Arms    date not known
4  The Circle of the Lustful: Francesca da Rimini... 1826-7, reprinted 1892

                                     medium \
0  Watercolour, ink, chalk and graphite on paper...
1                                     Graphite on paper
2  Graphite on paper. Verso: graphite on paper
3                                     Graphite on paper
4  Line engraving on paper

                                     creditLine  year  acquisitionYear \
0  Presented by Mrs John Richmond 1922    NaN        1922
1  Presented by Mrs John Richmond 1922    NaN        1922
2  Presented by Mrs John Richmond 1922  1785.0        1922
3  Presented by Mrs John Richmond 1922    NaN        1922
4  Purchased with the assistance of a special gra... 1826.0        1919
```

		dimensions	width	height	depth	units	inscription	\
0	support:	394 x 419 mm	394.0	419.0	NaN	mm	NaN	
1	support:	311 x 213 mm	311.0	213.0	NaN	mm	NaN	
2	support:	343 x 467 mm	343.0	467.0	NaN	mm	NaN	
3	support:	318 x 394 mm	318.0	394.0	NaN	mm	NaN	
4	image:	243 x 335 mm	243.0	335.0	NaN	mm	NaN	

	thumbnailCopyright	thumbnailUrl	\
0	NaN	http://www.tate.org.uk/art/images/work/A/A00/A...	
1	NaN	http://www.tate.org.uk/art/images/work/A/A00/A...	
2	NaN	http://www.tate.org.uk/art/images/work/A/A00/A...	
3	NaN	http://www.tate.org.uk/art/images/work/A/A00/A...	
4	NaN	http://www.tate.org.uk/art/images/work/A/A00/A...	

	url
0	http://www.tate.org.uk/art/artworks/blake-a-fi...
1	http://www.tate.org.uk/art/artworks/blake-two-...
2	http://www.tate.org.uk/art/artworks/blake-the-...
3	http://www.tate.org.uk/art/artworks/blake-six-...
4	http://www.tate.org.uk/art/artworks/blake-the-...

```
[ ]: data.dtypes
```

```
[ ]: id                int64
accession_number      object
artist                object
artistRole            object
artistId              int64
title                 object
dateText              object
medium                object
creditLine            object
year                  float64
acquisitionYear        int64
dimensions             object
width                  float64
height                 float64
depth                  float64
units                  object
inscription            float64
thumbnailCopyright     float64
thumbnailUrl           object
url                    object
dtype: object
```

normalize operation can only be performed on numeric values float or int

```
[ ]: data.height
```

```
[ ]: 0    419.0
      1    213.0
      2    467.0
      3    394.0
      4    335.0
      5    338.0
      6    334.0
      7    340.0
      8    335.0
      9    340.0
      Name: height, dtype: float64
```

```
[ ]: data.height.min()
```

```
[ ]: 213.0
```

```
[ ]: data.height.max()
```

```
[ ]: 467.0
```

```
[ ]: data.height.mean()
```

```
[ ]: 351.5
```

```
[ ]: data.height.std()
```

```
[ ]: 66.81857692455162
```

```
[ ]: height = data.height
      height
```

```
[ ]: 0    419.0
      1    213.0
      2    467.0
      3    394.0
      4    335.0
      5    338.0
      6    334.0
      7    340.0
      8    335.0
      9    340.0
      Name: height, dtype: float64
```

```
[ ]: # Standardization in statistics
      norm = (height - height.mean()) / height.std()
      norm
```

```
[ ]: 0    1.010198
      1   -2.072777
      2    1.728561
      3    0.636051
      4   -0.246937
      5   -0.202040
      6   -0.261903
      7   -0.172108
      8   -0.246937
      9   -0.172108
      Name: height, dtype: float64
```

```
[ ]: # normalize height with minimum 0, maximum 1
      minmax = (height - height.min()) / (height.max() - height.min())
      minmax
```

```
[ ]: 0    0.811024
      1    0.000000
      2    1.000000
      3    0.712598
      4    0.480315
      5    0.492126
      6    0.476378
      7    0.500000
      8    0.480315
      9    0.500000
      Name: height, dtype: float64
```

```
[ ]: minmax.min()
```

```
[ ]: 0.0
```

```
[ ]: minmax.max()
```

```
[ ]: 1.0
```

```
[ ]: data
```

```
[ ]:      id accession_number      artist artistRole  artistId \
0   1035          A00001  Blake, Robert    artist        38
1   1036          A00002  Blake, Robert    artist        38
2   1037          A00003  Blake, Robert    artist        38
3   1038          A00004  Blake, Robert    artist        38
4   1039          A00005  Blake, William  artist        39
5   1040          A00006  Blake, William  artist        39
6   1041          A00007  Blake, William  artist        39
7   1042          A00008  Blake, William  artist        39
8   1043          A00009  Blake, William  artist        39
```

9 1044 A00010 Blake, William artist 39

	title	dateText \
0	A Figure Bowing before a Seated Old Man with h...	date not known
1	Two Drawings of Frightened Figures, Probably f...	date not known
2	The Preaching of Warning. Verso: An Old Man En...	?c.1785
3	Six Drawings of Figures with Outstretched Arms	date not known
4	The Circle of the Lustful: Francesca da Rimini...	1826-7, reprinted 1892
5	Ciampolo the Barrator Tormented by the Devils	1826-7, reprinted 1892
6	The Baffled Devils Fighting	1826-7, reprinted 1892
7	The Six-Footed Serpent Attacking Agnolo Brunel...	1826-7, reprinted 1892
8	The Serpent Attacking Buoso Donati	1826-7, reprinted 1892
9	The Pit of Disease: The Falsifiers	1826-7, reprinted 1892

	medium \
0	Watercolour, ink, chalk and graphite on paper...
1	Graphite on paper
2	Graphite on paper. Verso: graphite on paper
3	Graphite on paper
4	Line engraving on paper
5	Line engraving on paper
6	Line engraving on paper
7	Line engraving on paper
8	Line engraving on paper
9	Line engraving on paper

	creditLine	year	acquisitionYear \
0	Presented by Mrs John Richmond 1922	NaN	1922
1	Presented by Mrs John Richmond 1922	NaN	1922
2	Presented by Mrs John Richmond 1922	1785.0	1922
3	Presented by Mrs John Richmond 1922	NaN	1922
4	Purchased with the assistance of a special gra...	1826.0	1919
5	Purchased with the assistance of a special gra...	1826.0	1919
6	Purchased with the assistance of a special gra...	1826.0	1919
7	Purchased with the assistance of a special gra...	1826.0	1919
8	Purchased with the assistance of a special gra...	1826.0	1919
9	Purchased with the assistance of a special gra...	1826.0	1919

	dimensions	width	height	depth	units	inscription \
0	support: 394 x 419 mm	394.0	419.0	NaN	mm	NaN
1	support: 311 x 213 mm	311.0	213.0	NaN	mm	NaN
2	support: 343 x 467 mm	343.0	467.0	NaN	mm	NaN
3	support: 318 x 394 mm	318.0	394.0	NaN	mm	NaN
4	image: 243 x 335 mm	243.0	335.0	NaN	mm	NaN
5	image: 240 x 338 mm	240.0	338.0	NaN	mm	NaN
6	image: 242 x 334 mm	242.0	334.0	NaN	mm	NaN
7	image: 246 x 340 mm	246.0	340.0	NaN	mm	NaN

8	image: 241 x 335 mm	241.0	335.0	NaN	mm	NaN
9	image: 243 x 340 mm	243.0	340.0	NaN	mm	NaN

	thumbnailCopyright	thumbnailUrl \
0	NaN	http://www.tate.org.uk/art/images/work/A/A00/A...
1	NaN	http://www.tate.org.uk/art/images/work/A/A00/A...
2	NaN	http://www.tate.org.uk/art/images/work/A/A00/A...
3	NaN	http://www.tate.org.uk/art/images/work/A/A00/A...
4	NaN	http://www.tate.org.uk/art/images/work/A/A00/A...
5	NaN	http://www.tate.org.uk/art/images/work/A/A00/A...
6	NaN	http://www.tate.org.uk/art/images/work/A/A00/A...
7	NaN	http://www.tate.org.uk/art/images/work/A/A00/A...
8	NaN	http://www.tate.org.uk/art/images/work/A/A00/A...
9	NaN	http://www.tate.org.uk/art/images/work/A/A00/A...

	url
0	http://www.tate.org.uk/art/artworks/blake-a-fi...
1	http://www.tate.org.uk/art/artworks/blake-two-...
2	http://www.tate.org.uk/art/artworks/blake-the-...
3	http://www.tate.org.uk/art/artworks/blake-six-...
4	http://www.tate.org.uk/art/artworks/blake-the-...
5	http://www.tate.org.uk/art/artworks/blake-ciam...
6	http://www.tate.org.uk/art/artworks/blake-the-...
7	http://www.tate.org.uk/art/artworks/blake-the-...
8	http://www.tate.org.uk/art/artworks/blake-the-...
9	http://www.tate.org.uk/art/artworks/blake-the-...

```
[ ]: # replacing the existing columns with normalized data
data.height = minmax
data['standardaized_height'] = norm
data
```

[]:	id	accession_number	artist	artistRole	artistId \
0	1035	A00001	Blake, Robert	artist	38
1	1036	A00002	Blake, Robert	artist	38
2	1037	A00003	Blake, Robert	artist	38
3	1038	A00004	Blake, Robert	artist	38
4	1039	A00005	Blake, William	artist	39
5	1040	A00006	Blake, William	artist	39
6	1041	A00007	Blake, William	artist	39
7	1042	A00008	Blake, William	artist	39
8	1043	A00009	Blake, William	artist	39
9	1044	A00010	Blake, William	artist	39

	title	dateText \
0	A Figure Bowing before a Seated Old Man with h...	date not known
1	Two Drawings of Frightened Figures, Probably f...	date not known

2	The Preaching of Warning. Verso: An Old Man En...	?c.1785
3	Six Drawings of Figures with Outstretched Arms	date not known
4	The Circle of the Lustful: Francesca da Rimini...	1826-7, reprinted 1892
5	Ciampolo the Barrator Tormented by the Devils	1826-7, reprinted 1892
6	The Baffled Devils Fighting	1826-7, reprinted 1892
7	The Six-Footed Serpent Attacking Agnolo Brunel...	1826-7, reprinted 1892
8	The Serpent Attacking Buoso Donati	1826-7, reprinted 1892
9	The Pit of Disease: The Falsifiers	1826-7, reprinted 1892

medium \

0	Watercolour, ink, chalk and graphite on paper...
1	Graphite on paper
2	Graphite on paper. Verso: graphite on paper
3	Graphite on paper
4	Line engraving on paper
5	Line engraving on paper
6	Line engraving on paper
7	Line engraving on paper
8	Line engraving on paper
9	Line engraving on paper

creditLine year ... \

0	Presented by Mrs John Richmond 1922	NaN	...
1	Presented by Mrs John Richmond 1922	NaN	...
2	Presented by Mrs John Richmond 1922	1785.0	...
3	Presented by Mrs John Richmond 1922	NaN	...
4	Purchased with the assistance of a special gra...	1826.0	...
5	Purchased with the assistance of a special gra...	1826.0	...
6	Purchased with the assistance of a special gra...	1826.0	...
7	Purchased with the assistance of a special gra...	1826.0	...
8	Purchased with the assistance of a special gra...	1826.0	...
9	Purchased with the assistance of a special gra...	1826.0	...

dimensions width height depth units inscription \

0	support: 394 x 419 mm	394.0	0.811024	NaN	mm	NaN
1	support: 311 x 213 mm	311.0	0.000000	NaN	mm	NaN
2	support: 343 x 467 mm	343.0	1.000000	NaN	mm	NaN
3	support: 318 x 394 mm	318.0	0.712598	NaN	mm	NaN
4	image: 243 x 335 mm	243.0	0.480315	NaN	mm	NaN
5	image: 240 x 338 mm	240.0	0.492126	NaN	mm	NaN
6	image: 242 x 334 mm	242.0	0.476378	NaN	mm	NaN
7	image: 246 x 340 mm	246.0	0.500000	NaN	mm	NaN
8	image: 241 x 335 mm	241.0	0.480315	NaN	mm	NaN
9	image: 243 x 340 mm	243.0	0.500000	NaN	mm	NaN

thumbnailCopyright thumbnailUrl \

0	NaN	http://www.tate.org.uk/art/images/work/A/A00/A...
---	-----	---

```

1      NaN http://www.tate.org.uk/art/images/work/A/A00/A...
2      NaN http://www.tate.org.uk/art/images/work/A/A00/A...
3      NaN http://www.tate.org.uk/art/images/work/A/A00/A...
4      NaN http://www.tate.org.uk/art/images/work/A/A00/A...
5      NaN http://www.tate.org.uk/art/images/work/A/A00/A...
6      NaN http://www.tate.org.uk/art/images/work/A/A00/A...
7      NaN http://www.tate.org.uk/art/images/work/A/A00/A...
8      NaN http://www.tate.org.uk/art/images/work/A/A00/A...
9      NaN http://www.tate.org.uk/art/images/work/A/A00/A...

```

```

                                url standardaized_height
0 http://www.tate.org.uk/art/artworks/blake-a-fi...      1.010198
1 http://www.tate.org.uk/art/artworks/blake-two-...     -2.072777
2 http://www.tate.org.uk/art/artworks/blake-the-...      1.728561
3 http://www.tate.org.uk/art/artworks/blake-six-...      0.636051
4 http://www.tate.org.uk/art/artworks/blake-the-...     -0.246937
5 http://www.tate.org.uk/art/artworks/blake-ciam...     -0.202040
6 http://www.tate.org.uk/art/artworks/blake-the-...     -0.261903
7 http://www.tate.org.uk/art/artworks/blake-the-...     -0.172108
8 http://www.tate.org.uk/art/artworks/blake-the-...     -0.246937
9 http://www.tate.org.uk/art/artworks/blake-the-...     -0.172108

```

[10 rows x 21 columns]

[]: