How to Color a Pandas DataFrame

A short tutorial on how to set the colors on a pandas DataFrame.

Roman Orac

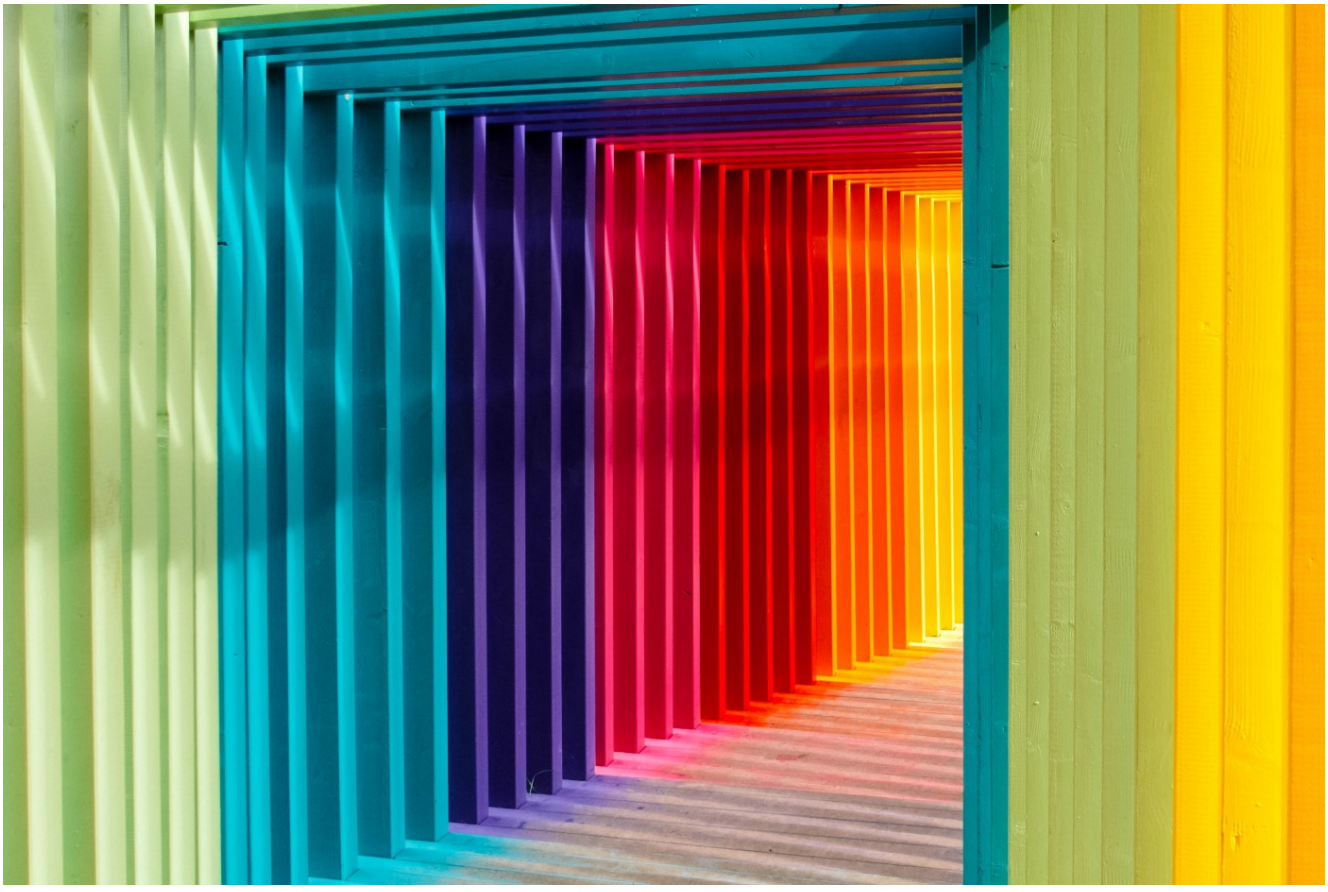


Photo by Robert Katzki on Unsplash

Pandas needs no introduction as it became the de facto tool for Data Analysis in Python. As a Data Scientist, I use pandas daily and it never ceases to amaze me with better ways of achieving my goals.

Another useful feature that I learned recently is how to color a pandas Dataframe.

Let’s add colors

Let’s create a pandas DataFrame with random numbers:

```
import numpy as np
import pandas as pddf = pd.DataFrame(np.random.ra
```

	A	B	C	D
0	68	11	17	12
1	96	36	82	61
2	90	86	58	77
3	21	33	18	49
4	21	40	27	17
5	87	75	73	47
6	33	86	1	77
7	31	23	43	42
8	61	34	26	87
9	65	22	0	31
10	39	16	10	44
11	82	57	56	34
12	61	86	65	64
13	91	30	85	81
14	0	75	29	87

Pandas Dataframe with random numbers (image made by author)

Coloring cells (by frequencies) is as simple as:

```
df.style.background_gradient(cmap="Blues")
```

	A	B	C	D
0	68	11	17	12
1	96	36	82	61
2	90	86	58	77
3	21	33	18	49
4	21	40	27	17
5	87	75	73	47
6	33	86	1	77
7	31	23	43	42
8	61	34	26	87
9	65	22	0	31
10	39	16	10	44
11	82	57	56	34
12	61	86	65	64
13	91	30	85	81
14	0	75	29	87

Colored Pandas Dataframe with random numbers (image made by author)

Coloring is column-based

If we increase column B by 1000, it won't interfere with other column colors.

```
df["B"] *= 1000
```

As we see in the image below, maximums of columns A, C, D retained their color.

	A	B	C	D
0	68	11000	17	12
1	96	36000	82	61
2	90	86000	58	77
3	21	33000	18	49
4	21	40000	27	17
5	87	75000	73	47
6	33	86000	1	77
7	31	23000	43	42
8	61	34000	26	87
9	65	22000	0	31
10	39	16000	10	44
11	82	57000	56	34
12	61	86000	65	64
13	91	30000	85	81
14	0	75000	29	87

Colored Pandas Dataframe with random numbers (image made by author)

Change the color map

You can set any colormap supported in matplotlib. Just be careful to select the sequential colormap if your goal is to visualize the frequencies.

Sequential

For the Sequential plots, the lightness value increases monotonically through the colormaps. This is good. Some of the L^* values in the colormaps span from 0 to 100 (binary and the other grayscale), and others start around $L^* = 20$. Those that have a smaller range of L^* will accordingly have a smaller perceptual range. Note also that the L^* function varies amongst the colormaps: some are approximately linear in L^* and others are more curved.

```
cmaps['Perceptually Uniform Sequential'] = [
            'viridis', 'plasma', 'inferno', 'magma', 'cividis']

cmaps['Sequential'] = [
            'Greys', 'Purples', 'Blues', 'Greens', 'Oranges', 'Reds',
            'YlOrBr', 'YlOrRd', 'OrRd', 'PuRd', 'RdPu', 'BuPu',
            'GnBu', 'PuBu', 'YlGnBu', 'PuBuGn', 'BuGn', 'YlGn']
```

Sequential2

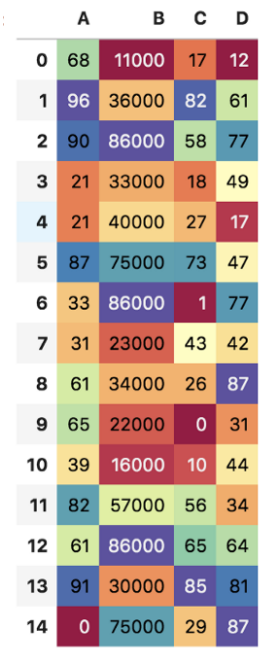
Many of the L^* values from the Sequential2 plots are monotonically increasing, but some (autumn, cool, spring, and winter) plateau or even go both up and down in L^* space. Others (afmhot, copper, gist_heat, and hot) have kinks in the L^* functions. Data that is being represented in a region of the colormap that is at a plateau or kink will lead to a perception of banding of the data in those values in the colormap (see [mycarta-banding] for an excellent example of this).

```
cmaps['Sequential (2)'] = [
            'binary', 'gist_yarg', 'gist_gray', 'gray', 'bone', 'pink',
            'spring', 'summer', 'autumn', 'winter', 'cool', 'Wistia',
            'hot', 'afmhot', 'gist_heat', 'copper']
```

Sequential colormaps from matplotlib (image made by author)

A bad example of visualizing frequencies is with a non-sequential colormap (so make sure you use a sequential colormap):

```
df.style.background_gradient(cmap="Spectral")
```



	A	B	C	D
0	68	11000	17	12
1	96	36000	82	61
2	90	86000	58	77
3	21	33000	18	49
4	21	40000	27	17
5	87	75000	73	47
6	33	86000	1	77
7	31	23000	43	42
8	61	34000	26	87
9	65	22000	0	31
10	39	16000	10	44
11	82	57000	56	34
12	61	86000	65	64
13	91	30000	85	81
14	0	75000	29	87

A bad example of visualizing frequencies is with a non-sequential colormap (image made by author)

Before you go

– Intro to Machine Learning with PyTorch [Course]

Some of the links above are affiliate links and if you go through them to make a purchase I'll earn a commission. Keep in mind that I link courses because of their quality and not because of the commission I receive from your purchases.

Follow me on Twitter, where I regularly tweet about Data

Science and Machine Learning.

Photo by Courtney Hedger on Unsplash