

Visualization with Heat Maps

This lesson explains what a heat map is, why it is used, and how to visualize data with a heat map using Python libraries.

WE'LL COVER THE FOLLOWING ^

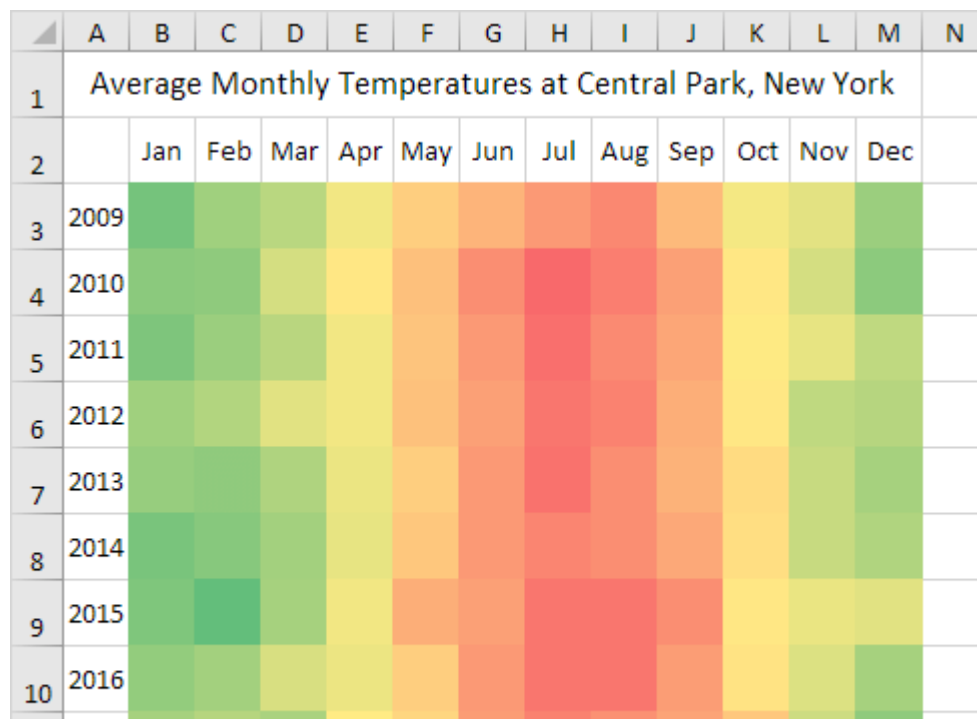
- Introduction to heat maps
- Heat maps in Python

Introduction to heat maps

Heat maps are a plotted data matrix where each value in the matrix is a category.

As an example, consider the **heat map** shown below.

This heat map represents the average monthly temperatures at Central Park, New York for a given year (**rows**) and month (**columns**). The coloring of the heat map makes it easy to see that the warmest months are consistently from Jun to Sep.



11	2017														
12															

Heat maps in Python

You can create heat maps in Python by using the `heatmap()` function from Seaborn. Some of the core parameters:

- `data` is the first parameter and is your matrix of data represented as a Panda's dataframe.
- `annot` if `True` will plot the actual data values in each cell.
- `fmt` lets you control the string formatting. A value of `"d"` uses a decimal integer.
- `linewidths` lets you set the width of the lines which separate each cell.
- `ax` allows you to pass a custom Matplotlib Axes.
- `cmap` allows you to pick the colormap to use when plotting.

Let's look at an example using the same flight data from the Line Graphs lesson:

```
import seaborn as sns
import matplotlib.pyplot as plt

# Load dataset
flights_long = sns.load_dataset("flights")
# Pivot the dataset from long to wide format
flights = flights_long.pivot("month", "year", "passengers")
# Create a larger figure size to plot on
f, ax = plt.subplots(figsize=(12, 6))
# Create the heat map
sns.heatmap(flights, annot=True, fmt="d", linewidths=.5, ax=ax, cmap='Blues')
```



Above, we plotted the number of airline passengers by month and year. In our code, we used Matplotlib to adjust the figure size to make it larger, (12,6). That returns a value we called `"ax"` and we pass this to Seaborn with the `ax` parameter. We also use the `annot = True` parameter to add the actual values to our cells.

The other option that I want to call out is the `cmap` parameter. This allowed us to adjust the color used for the cells by passing `"Blues"`.

I like heat maps because the color on top of the values allows us to see the trends much easier. We can see over the years the number of passengers has increased, and that July and August are consistently the most popular months. There is a lot of value from a single plot.

That's how the low and high values can be tracked down with heat maps. Next, we'll look at multi-plot grids in Python.