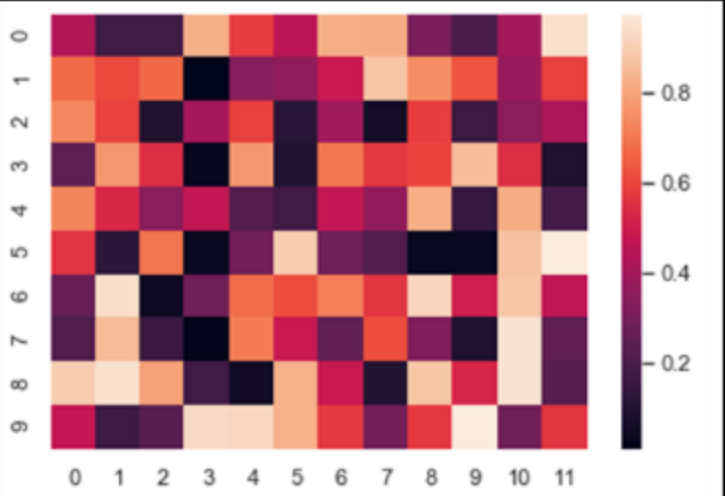


Mapas de calor

Los **mapas de calor** se utilizan para representar datos utilizando diferentes variaciones de colores, se aplican a datos almacenados en formato tabular y los valores numéricos contenidos en cada una de las celdas son sustituidos por colores establecidos previamente por el investigador en una escala que se incluye junto al mapa de calor para que cualquier observador pueda entender la información mostrada.

Elementos que componen a los mapas de calor:

- Representación tabular** compuesta por dos subconjuntos de variables divididas en las variables de las filas y las variables de las columnas.
- Escala de colores**, la cual asocia a cada color con un diferente valor numérico.
- El color asignado a cada celda**, el cuál representa la relación entre la variable de la fila y la variable de la columna.

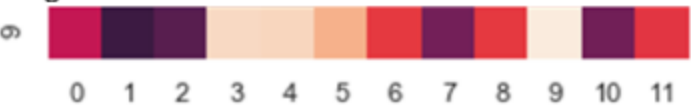


En nuestra imagen se observa que nuestro mapa de calor contiene diez variables de filas (*Variables 0 hasta 9*) y que tiene doce variables de renglón (*Variables 0 hasta 11*).

Nuestra escala de colores asigna colores oscuros a las celdas que contengan un valor cercano a cero y le asigna colores claros a las celdas que contengan un valor cercano a uno.

Para continuar con el análisis de nuestra imagen, aislaremos una fila y la describiremos para entender como funciona el mapa de calor. En esta fila se analiza a la variable nueve (del conjunto de filas) con respecto a cada una de las doce variables del conjunto de columnas. Cada una de las celdas es equivalente a asignarle un valor numérico en una tabla.

9



01234567891011

↓

9

0.5	0.0	0.2	0.9	0.9	0.8	0.6	0.3	0.8	1.0	0.3	0.5
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

01234567891011

La variable de fila nueve tiene valores más cercanos a uno con respecto a las variables tres, cuatro y nueve del conjunto de columnas, pero presenta valores más cercanos a cero para las variables uno, siete y diez del conjunto de columnas.


Los mapas de calor son buenos para mostrar la varianza a través de múltiples variables, revelando cualquier patrón, mostrando si las variables son similares entre sí y para detectar si existen correlaciones entre ellas. Una desventaja que presentan es que, debido a su dependencia del color para comunicar los valores, muchas veces es muy difícil distinguir con precisión los valores exactos que se representan en cada una de las celdas, por lo que los mapas de calor se utilizan en mayor medida para darnos una idea generalizada de la información.

Continuaremos con nuestro estudio de los mapas de calor, generando el mapa de calor correspondiente a uno de los ejemplos más populares de la literatura. En este ejemplo se nos presenta una tabla que contiene la correlación entre cuatro variables observadas en las flores de iris: Largo del sépalo de la flor, ancho del sépalo, largo del pétalo y ancho del pétalo.

	sepal_length	sepal_width	petal_length	petal_width
sepal_length	1	-0.11757	0.871754	0.817941
sepal_width	-0.11757	1	-0.42844	-0.36613
petal_length	0.871754	-0.42844	1	0.962865
petal_width	0.817941	-0.36613	0.962865	1

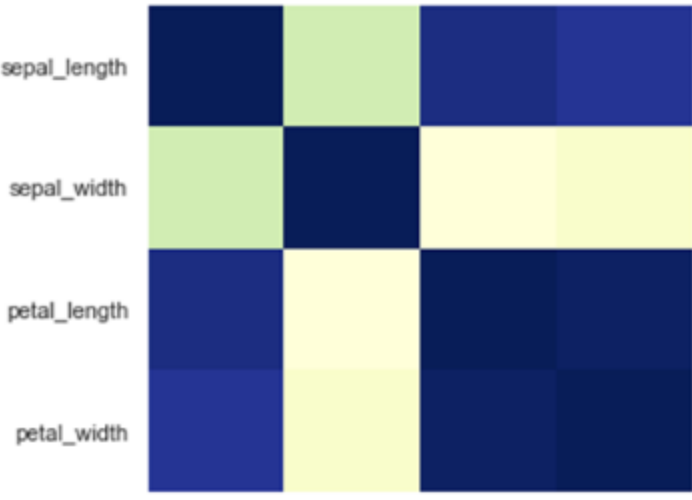
Utilizando los valores obtenidos en la tabla se propone una escala de colores que se utilizará para asignarle a cada celda un color que represente su valor numérico.

1.000.750.500.250.00-0.25



Por último, se sustituye el valor de cada celda por el color correspondiente en la escala propuesta para terminar de fabricar el mapa de calor.

sepal_lengthsepal_widthpetal_lengthpetal_width



sepal_lengthsepal_widthpetal_lengthpetal_width

```
In [ ]: import pandas as pd

Heart Disease UCI Cleveland
https://www.kaggle.com/ronitf/heart-disease-uci

In [ ]: df = pd.read_csv('data/heart/heart.csv')
df.head(7)

In [ ]: df.corr() # Hallar la correlación

In [ ]: import matplotlib.pyplot as plt
%matplotlib inline
plt.imshow(df.corr()) # imshow: Muestra una imagen construida

In [ ]: corr=df.corr()

fig, ax = plt.subplots()
plt.imshow(corr)
plt.xticks(range(len(corr)), corr.index, rotation=75) # Los mismos nombres de las columnas se asignan a los ejes 75 es la inclinación
plt.yticks(range(len(corr)), corr.index)
plt.show()

In [ ]: corr=df.corr()

fig, ax = plt.subplots()
plt.imshow(corr,cmap='PuOr') # Visitar y cambiar el color: https://matplotlib.org/stable/tutorials/colors/colormaps.html
plt.xticks(range(len(corr)), corr.index, rotation=75)
plt.yticks(range(len(corr)), corr.index)
plt.colorbar() # Barra de colores o nueva escala
plt.show()

In [ ]: import seaborn as sns
import matplotlib.pyplot as plt
plt.figure(figsize=(12,10))
sns.heatmap(corr)
#sns.heatmap(corr, linewidth=0.5) # Con un ancho de línea de 0.5
plt.show()

In [ ]: import seaborn as sns
import matplotlib.pyplot as plt
plt.figure(figsize=(12,10))
sns.heatmap(corr, linewidth=0.5,annot=True,cmap="YlGnBu") #<- https://matplotlib.org/stable/tutorials/colors/colormaps.html

#Guardar La imagen
plt.savefig('img\mapa.png',dpi=400)
# Muestra La imagen en pantalla
plt.show()
```

Se solicita leer sobre mejores gráficas o elementos para hacer visualizaciones

<https://matplotlib.org/stable/tutorials/colors/colormaps.html>

```
In [ ]:
```