

## HEART DISEASE

### DATA

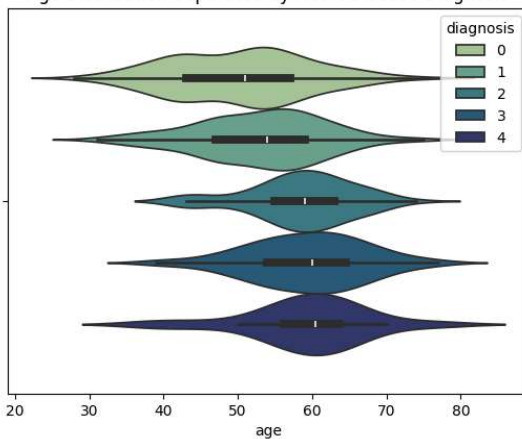
Objective:

Understand the relationship between heart disease and biological traits (such as age, cholesterol, and maximum heart rate)

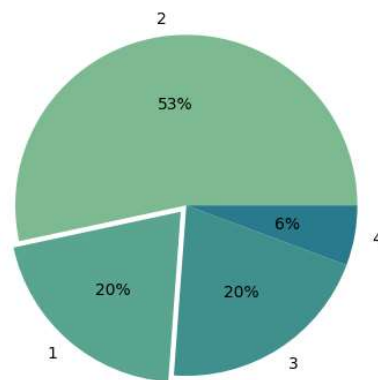
	age	sex	restingbp	chol	blood_sugar	max_hr	diagnosis
count	745.000000	745.000000	745.000000	745.000000	745.000000	745.000000	745.000000
mean	53.130201	0.766443	132.771812	219.923490	0.153020	138.817450	0.924832
std	9.392743	0.423378	18.599746	93.713004	0.360249	25.825905	1.128677
min	28.000000	0.000000	0.000000	0.000000	0.000000	60.000000	0.000000
25%	46.000000	1.000000	120.000000	197.000000	0.000000	120.000000	0.000000
50%	54.000000	1.000000	130.000000	231.000000	0.000000	140.000000	1.000000
75%	60.000000	1.000000	140.000000	271.000000	0.000000	160.000000	1.000000
max	77.000000	1.000000	200.000000	603.000000	1.000000	202.000000	4.000000

No Heart Disease	Heart Disease
<ul style="list-style-type: none"> <li>Age: 50.32</li> <li>Resting Blood Pressure: 129.89</li> <li>Cholesterol: 233.112</li> <li>Blood Sugar: 0.109</li> <li>Max HR: 149.41</li> </ul>	<ul style="list-style-type: none"> <li>Age: 55.74</li> <li>Resting Blood Pressure: 135.46</li> <li>Cholesterol: 207.65</li> <li>Blood Sugar: 0.194</li> <li>Max HR: 128.97</li> </ul>

Age Distribution Separated by Heart Disease Diagnosis



Percentage of Different Diagnosis



### Commands Used:

- **sed**: to replace question marks for missing values with -
- **awk**: filtering out relevant columns for data analysis
- **cat**: used to display individual datasets and merge them into one big dataset
- **Python**: Plotting and Dataframe manipulation

### Key Takeaways

- On average, the older one is, the more likely you are to have heart disease
- Cholesterol levels were *unexpectedly* inversely related to diagnosis
- Heart Disease correlates to a *lower average* Max HR
- But, Blood Pressure is increased.

3D Scatter of Heart Disease Features

