Proposal : Heart Disease dataset

Dataset Overview:

All the values in the dataset were collected at the time of medical examination and provided information about the signs of Cardiovascular Disease. This Dataset has 70,000 observations(rows) and 12 features(columns).

Data description

There are 3 types of input columns:

1. **Objective:** factual information ,
2. **Examination:** results of medical examination;
3. **Subjective:** information was given by the patient.

   ### Columns description

1. **Objective Feature**
   - Age | age | int (days)
   - Height | height | int (cm)
   - Weight | weight | float (kg)
   - Gender | gender | categorical

1. **Examination Feature**
   - Systolic blood pressure | ap_hi | int
   - Diastolic blood pressure | ap_lo | int
   - Cholesterol | cholesterol | 1: normal, 2: above normal, 3: well above normal
   - Glucose | gluc | 1: normal, 2: above normal, 3: well above normal

1. **Subjective Feature**
   - Smoking | smoke | binary
   - Alcohol intake alco | binary
   - Physical activity | Subjective Feature | active | binary
   - Presence or absence of cardiovascular disease | Target Variable | cardio | binary |

MVP Goal :

The goal is to do exploratory data analysis (EDA) to answer these questions :

1. What is the most gender is exposed to be Cardiovascular Disease that has?
2. Is there a relationship between smoking and getting Cardiovascular Disease?
3. Is there a relationship between age and getting Cardiovascular Disease?
4. what is the age that has the most patient number of Cardiovascular Disease?
5. How many Smokers and Non-Smokers do not have a Cardiovascular Disease?
6. How many Smokers and Non-Smokers have a Cardiovascular Disease?
7. Is their a relationship between doing physical activity and not having a Cardiovascular Disease?

8. Is there a relationship between being overweight and having Cardiovascular Disease?
9. How many cholesterol patients have a Cardiovascular and at what level they are?
10. What is the cholesterol level that has a large number of Cardiovascular Disease?
11. How many alcoholics are Cardiovascular Disease?

Dataset sourec:

Kaggle.com :

- To download the dataset Click_here

Tools :

- VSCode
- Jupyter notebook
- Python lang
- Github & Git