Heart Disease Dataset

# Kontext

V tomto projekte som použil dataset Heart Disease (z Kaggle), ktorý obsahuje informácie o pacientoch a ich zdravotných ukazovateľoch, napríklad vek, cholesterol či krvný tlak. Cieľom bolo vykonať základné čistenie dát, zobraziť grafy a zhrnúť výsledky.

[Link na databázu](https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset/data)

## Kód by Byomars

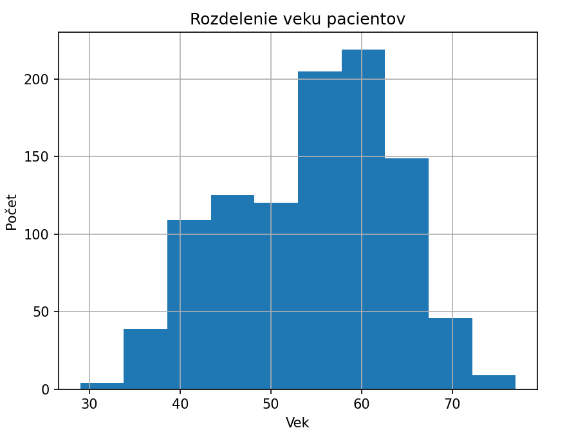
Importuje rôzne knižnice ako napríklad Axes z mpl\_toolkits.mplot3d,

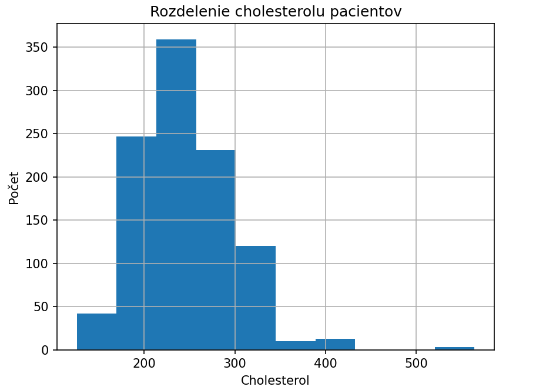
Ďalej numpy – táto knižnica nám umožňuje prácu s číselnými poliamy a lineárnou algebrou,

Pandas – použitá je na prácu s tabuľkovými dátami

Tento kód načíta prvých 1000 riadkov zo súboru **heart.csv**, uloží ich do DataFrame df1 a vypíše počet riadkov a stĺpcov. Okrem toho definuje pomocné funkcie na vizualizáciu dát (histogramy, korelačná matica, scatter matrix).

### Vizualizácia dát

Vytvoril som histogram veku pacientov, ktorý ukazuje, že väčšina pacientov je vo veku 40 až 60 rokov. Ďalší histogram zobrazuje rozdelenie cholesterolu, kde vidno, že hodnoty sa najčastejšie pohybujú v rozmedzí 200 až 300.



#### Záver

Dataset Heart poskytuje zaujímavé údaje o pacientoch a faktoroch, ktoré môžu súvisieť so srdcovými ochoreniami. Z analýzy vyplýva, že väčšina pacientov má zvýšený cholesterol a vek medzi 40–60 rokmi, čo môže predstavovať rizikovú skupinu. Projekt splnil cieľ: údaje boli načítané, vyčistené, analyzované a vizualizované.

[Link na kód](https://www.kaggle.com/code/bayomars12/starter-heart-disease-dataset-7fd7459c-d)