# Heart Disease analysis with classification

**Authors:** Lola Sõukand, Robin Mürk

**Data:** <https://www.kaggle.com/datasets/alexteboul/heart-disease-health-indicators-dataset/discussion>

Aim is to build a binary classifier that predicts a person's risk of a heart attack and to analyse individual attributes to determine what lowers and what increases the risk of heart attack using regression analysis.

**Goal 1:** train a model to predict whether or not a patient is at high risk or at low risk of a heart attack.  
**Goal 2:** find factors, that lower the risk of heart attack  
**Goal 3:** find factors, that increase the risk of heart attack

## Tasks

1. Prodect planning- Robin, Lola (4 h)
2. Data describing and first visualitsion – Lola (8 h)
   1. Seaborn
   2. Math.plt
3. Data cleaning - Lola (1 h)
   1. Vajadusel
4. Training ja test splittid -Robin (2 h)
   1. Balazeering
   2. For every model?
5. Models (Goal 1) – Robin (20 h)
   1. RandomForest
   2. K-neighbours
   3. SVC
   4. Keras
   5. Ansamble
   6. Accurasi
   7. AOC
   8. Hyper parameter optimization for every model
   9. Random seed must be set
6. Visualistaion (Goal 2, 3) – Lola (12 h)
   1. Figures (outliers,
7. Summary -Robin, Lola (1-2 h)
8. Making poster – Robin, Lola (2 h)