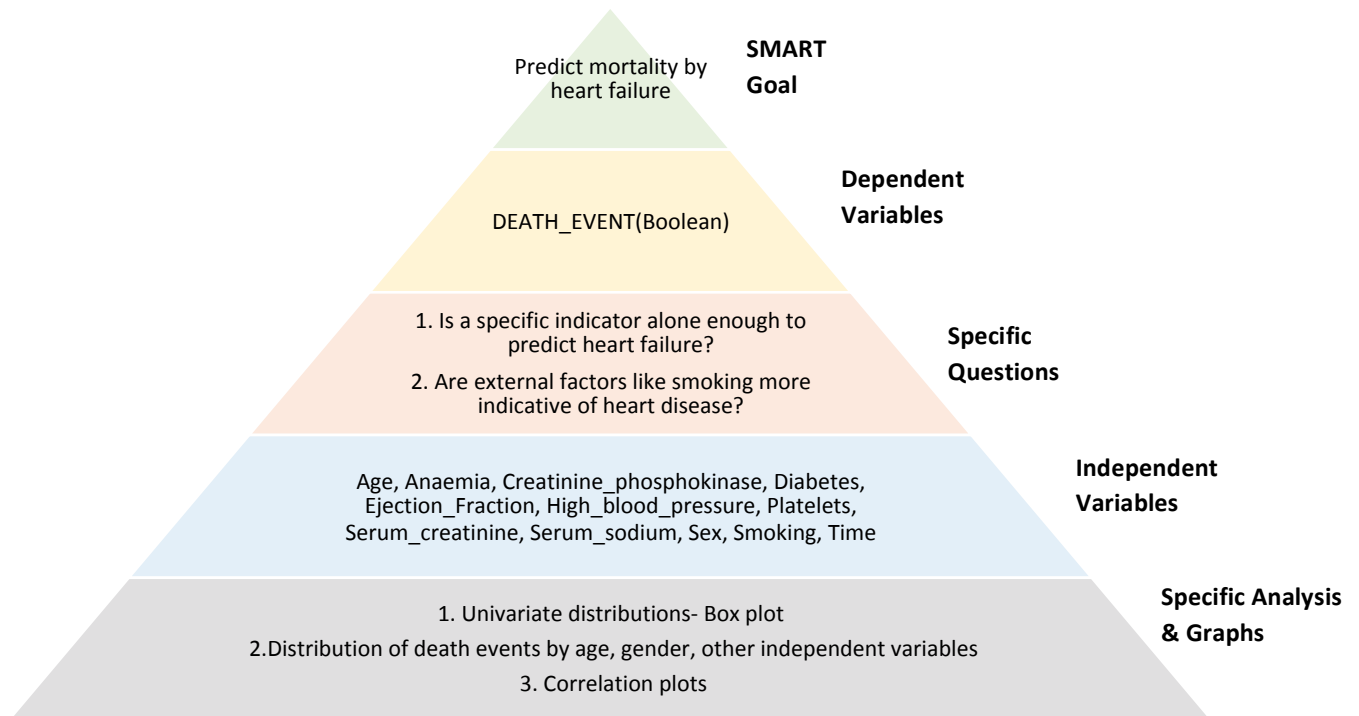


## 26:630:670 Multivariate Analysis Fall 2020

**Github Link** - <https://github.com/ag77in/HeartFailurePrediction-MVA>

### Initial Draft of SPAP



# Heart Failure Prediction

Problem Statement - Predict mortality by heart failure (Classification)

People with cardiovascular disease or who are at high cardiovascular risk (due to the presence of one or more risk factors such as hypertension, diabetes, hyperlipidaemia or already established disease) need early detection and management for which a dataset will be used to predict who are at high risk.

## Project Team

- Aman Goswami
- Tanvi Rudrashetty

## About the Data

Cardiovascular diseases (CVDs) are the number 1 cause of death globally, taking an estimated 17.9 million lives each year, which accounts for 31% of all deaths worldwide. Heart failure is a common event caused by CVDs and this dataset contains 12 features that can be used to predict mortality by heart failure.

Most cardiovascular diseases can be prevented by addressing behavioural risk factors such as tobacco use, unhealthy diet and obesity, physical inactivity and harmful use of alcohol using population-wide strategies.

Data Source - <https://www.kaggle.com/andrewmvd/heart-failure-clinical-data>

## Data Dimension

- 300 rows or patient details
- 13 columns or features (behavioral)

## Data Dictionary

Variable	Definition
age	Age of the patient
anaemia	Decrease of red blood cells or hemoglobin (boolean)
creatinine_phosphokinase	Level of the CPK enzyme in the blood (mcg/L)

Variable	Definition
diabetes	If the patient has diabetes (boolean)
ejection_fraction	Percentage of blood leaving the heart at each contraction (percentage)
high_blood_pressure	If the patient has hypertension (boolean)
platelets	Platelets in the blood (kiloplatelets/mL)
serum_creatinine	Level of serum creatinine in the blood (mg/dL)
serum_sodium	Level of serum sodium in the blood (mEq/L)
sex	Woman or man (binary)
smoking	If the patient smokes or not (boolean)
time	Follow-up period (days)
DEATH_EVENT	If the patient deceased during the follow-up period (boolean) - Target variable

## Credit to Authors

Citation Davide Chicco, Giuseppe Jurman: Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone. BMC Medical Informatics and Decision Making 20, 16 (2020). ([link](#))

License CC BY 4.0

Splash icon Icon by Freepik, available on Flaticon.

Splash banner Wallpaper by jcomp, available on Freepik.

## MVA Course details

26:630:670 - Multivariate Analysis Fall 2020