

# Cardiovascular Disease Analysis

## Introduction

Cardiovascular diseases (CVD) are a leading cause of mortality worldwide. This report investigates whether cardiovascular disease (variable 'cardio') can be explained by other variables such as age, gender, blood pressure, BMI, and lifestyle factors like smoking, alcohol consumption, and physical activity. The analysis is based on a dataset containing various health metrics.

## **Chapter 01 Data Preparation**

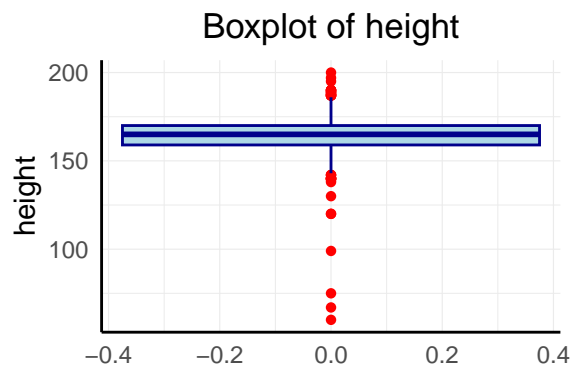
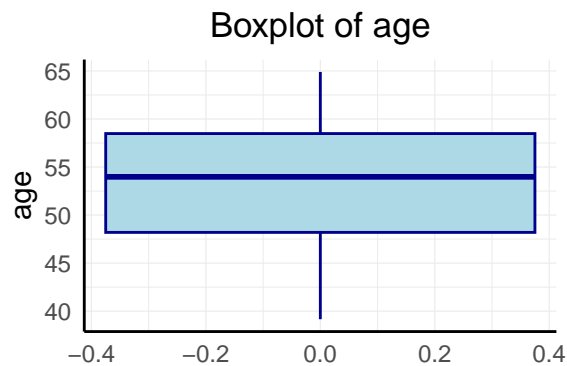
Task 1: Transform the variables of the data set to appropriate data types and assign factor labels for the categorical variables.

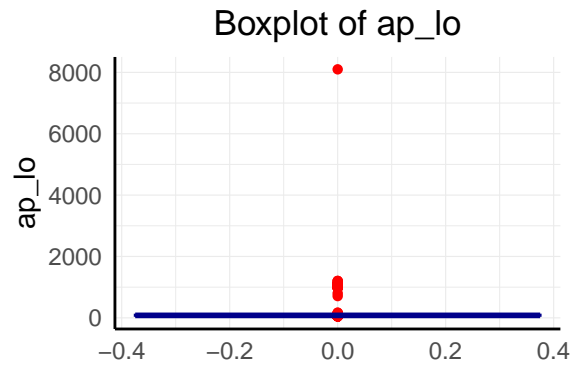
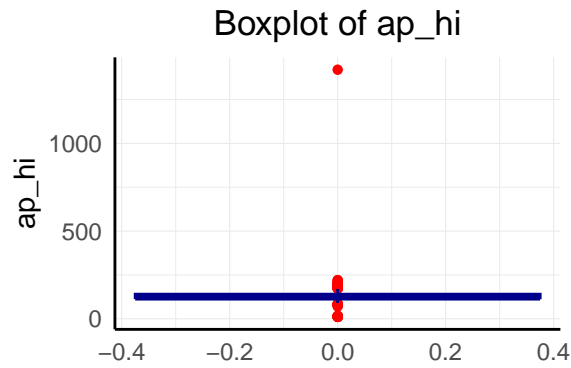
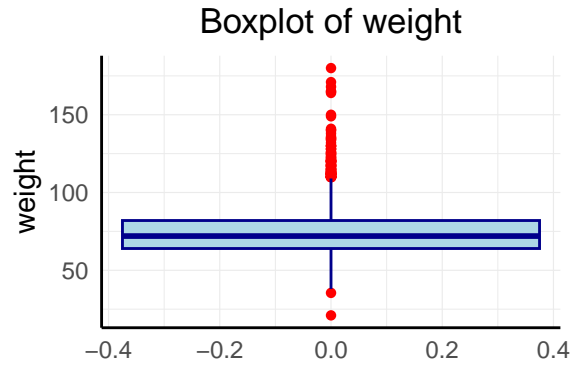
## Chapter 2: Outlier Detection

Task 2: Check the continuous variables for outliers and remove implausible values.

age	height	weight	ap_hi	ap_lo
Min. :39.16	Min. : 60.0	Min. : 21.00	Min. : 10.0	Min. : 40.0
1st Qu.:48.20	1st Qu.:159.0	1st Qu.: 64.00	1st Qu.: 120.0	1st Qu.: 80.0
Median :53.98	Median :165.0	Median : 72.00	Median : 120.0	Median : 80.0
Mean :53.31	Mean :164.3	Mean : 74.01	Mean : 126.7	Mean : 96.1
3rd Qu.:58.48	3rd Qu.:170.0	3rd Qu.: 82.00	3rd Qu.: 140.0	3rd Qu.: 90.0
Max. :64.90	Max. :200.0	Max. :180.00	Max. :1420.0	Max. :8099.0

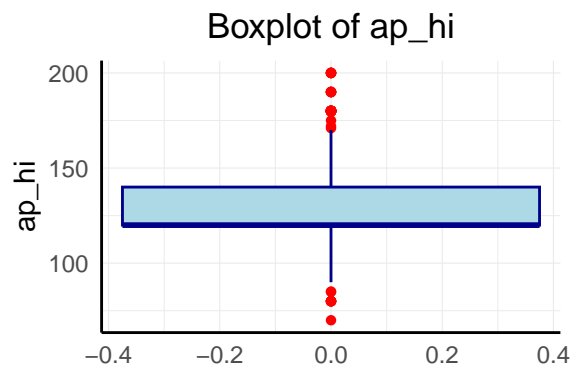
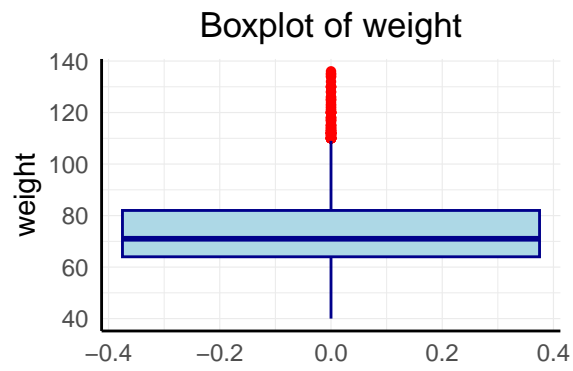
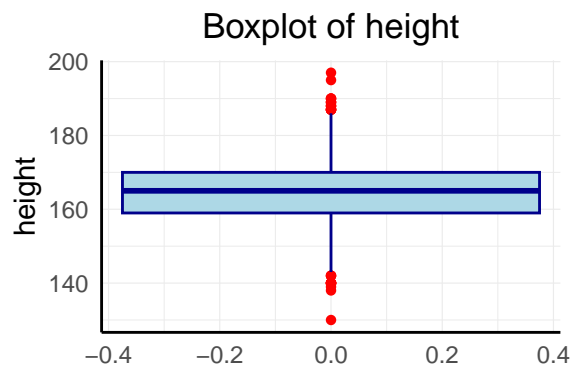
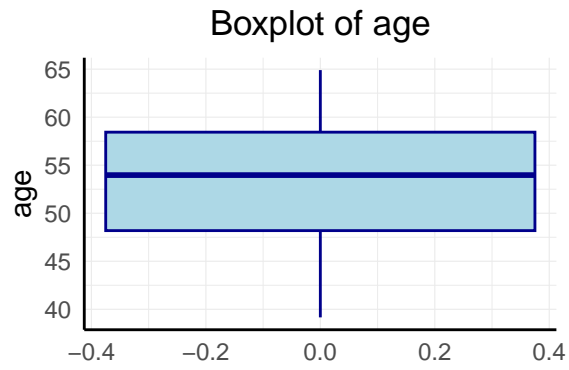
Boxplot before outliers are removed:

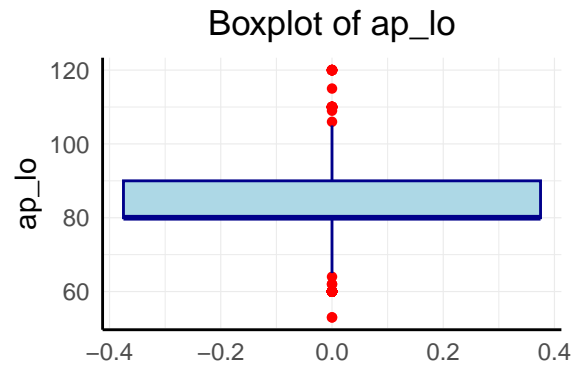




**Boxplot after outliers are removed**

age	height	weight	ap_hi	ap_lo
Min. :39.16	Min. :130.0	Min. : 40.00	Min. : 70.0	Min. : 53.00
1st Qu.:48.19	1st Qu.:159.0	1st Qu.: 64.00	1st Qu.:120.0	1st Qu.: 80.00
Median :53.97	Median :165.0	Median : 71.00	Median :120.0	Median : 80.00
Mean :53.30	Mean :164.4	Mean : 73.71	Mean :126.4	Mean : 81.25
3rd Qu.:58.44	3rd Qu.:170.0	3rd Qu.: 82.00	3rd Qu.:140.0	3rd Qu.: 90.00
Max. :64.90	Max. :197.0	Max. :136.00	Max. :200.0	Max. :120.00





### Chapter 3: BMI Calculation

Task 3: Create a new variable BMI and provide a summary table for the variable BMI for both cardio groups.

```
# A tibble: 2 x 7
  cardio Count Mean_BMI Median_BMI SD_BMI Min_BMI Max_BMI
  <fct>  <int>   <dbl>     <dbl>  <dbl>  <dbl>   <dbl>
1 absent   2495    26.3      25.3   4.62   15.8    50.4
2 present  2384    28.3      27.2   5.23   16.7    54.4
```