

# Phân tích dữ liệu sức khỏe

Trần Anh Quân

Ngày 6 tháng 3 năm 2025

## 1 Giới thiệu

Quản lý cân nặng: `weight_kg`, `bmi` (cân nặng và chỉ số khối cơ thể)

Thói quen luyện tập: `activity_type`, `duration_minutes`, `intensity`, `calories_burned`, `daily_steps`

Thói quen sinh hoạt: `hours_sleep`, `stress_level`, `hydration_level`, `smoking_status`

## 2 Giới thiệu chung về các phương pháp phân tích dữ liệu

### 2.1 Thống kê mô tả

Thống kê mô tả là bước phân tích ban đầu nhằm tóm tắt và mô tả các đặc điểm cơ bản của dữ liệu bằng các số liệu (như trung bình, trung vị, độ lệch chuẩn) và biểu đồ (như histogram, boxplot).

Mục tiêu: Cung cấp cái nhìn tổng quan về dữ liệu, phát hiện xu hướng, và xác định các đặc điểm quan trọng trước khi đi sâu vào các phương pháp phức tạp hơn.

### 2.2 Hồi quy tuyến tính đa biến

Mục tiêu: Tìm ra sự ảnh hưởng của từng biến lên chỉ số cân nặng

#### 2.2.1 Hệ số tương quan