

# Project 2 Data analysis plan

Shuai Zhu

2024-11-04

## 1 Introduction

The National Health and Nutrition Examination Survey (NHANES) is a program of studies aimed at assessing the health and nutritional status of adults and children in the United States. This project leverages the NHANES 2011-2014 dataset, which provides data on physical activity, mortality, and various demographic factors. The primary objective is to identify risk factors associated with all-cause mortality and to evaluate the relative predictive and discriminative value of several activity and circadian features derived from wearable accelerometers.

## 2 Method

### 2.1 Data cleaning

After filtering for participants aged 40 to 80 with at least three valid days of accelerometer data, the dataset was further narrowed down to complete cases for subsequent analysis. This resulted in a reduction from 19,141 subjects to 5,564 subjects.

## 2.2 Data analysis