

Homework 7

This homework is due on Apr. 10, 2025 at 11:00pm. Please submit as a pdf file on Canvas.

For both problems in this homework, we will work with the `heart_disease_data` dataset, which is a simplified and recoded version of a dataset available from kaggle. You can read about the original dataset here: <https://www.kaggle.com/datasets/kamilpytlak/personal-key-indicators-of-heart-disease?resource=download>

The `heart_disease_data` dataset contains 9 variables: `HeartDisease` (whether or not the participant has heart disease), `BMI` (body mass index), `PhysicalHealth` (how many days a month was their physical health not good), `MentalHealth` (how many days a month was their mental health not good), `ApproximateAge` (participants age), `SleepTime` (how many hours of sleep do they get in a 24-hour period), `Smoking` (1-smoker, 0-nonsmoker), `AlcoholDrinking` (1-drinks alcohol, 0-does not drink), `PhysicalActivity` (1-did physical activity or exercise during the past 30 days, 0-hardly any physical activity). Compared to the original dataset, the columns `ApproximateAge`, `Smoking`, `AlcoholDrinking`, and `PhysicalActivity` have been converted into numeric columns so they can be included in a PCA.

Note: This homework is about the contents of the plots. Don't worry about styling. It's OK to use the default theme and plot labeling.

```
heart_data <- read_csv("https://wilkelab.org/SDS366/datasets/heart_disease_data.csv")
```

Problem 1: (5 pts)

Perform a PCA of the numerical columns of the `heart_disease_data` dataset. Then make two plots, a rotation plot of components 1 and 2 and a plot of the eigenvalues, showing the amount of variance explained by the various components.

Rotation plot:

```
# Add your code here
```

Variance explained plot:

```
# Add your code here
```

Problem 2: (5 pts) Make a scatter plot of PC 2 versus PC 1 and color by heart disease status. Then use the rotation plot from Problem 1 to describe the variables/factors by which we can separate the study participants with heart disease from the study participants without heart disease.

```
# Add your code here
```

Add your interpretation of the PCA results here.