Data analysis in R assignment

Your group have been assigned a dataset to analyse.

I would like you to conduct some research on the data you have been assigned and add this to your presentation:

- Provide background information on the data set (e.g. what the data shows, why it's useful, etc).
- Explain each step of your data analysis

The data analysis steps I would like you to present:

1. Upload data set (read data into R)
2. Clean data (i.e. remove and NAs)
3. Create a new column(s) based on a loop and/or conditional statements
4. Select columns and assign to a new data frame.
5. Filter data and assign to a new data frame.
6. Write one of these new data frames into a csv file.
7. You must create at least the following plots of the data using ggplot2 (you can present any part of the filtered data):
   - Box plot
   - Scatter plot with a line of best fit
   - Histogram

You are welcome to make more plots if you like.

Plots must contain axis names, titles and colours.

At least 1 plot must be faceted.

Use head() where relevant to show changes made to data frame (if showing the whole data frame is impossible).

Groups and links for a starting point:

- Calista and Audrey: Alzheimer's dataset

https://r-packages.io/datasets/alzheimer

- Andy and Elaine: Pima Indians Diabetes Database

https://r-packages.io/datasets/PimaIndiansDiabetes

- Kate, Iris, Leco: BreastCancer

https://r-packages.io/datasets/BreastCancer