Background

Omada Health is on a mission to use data to improve weight-loss outcomes and engagements of its participants. Our pricing structure reflects this mission - we only charge participants if they engage in the program or lose weight! This is one of many reasons that it is important for us to predict how much weight our current and future participants will lose!

Using the simulated data provided - **your task is to predict future weight loss, based on historical data.**

(Bonus Challenge) At Omada, we use data to steer product development toward features which facilitate lasting behavior change that lead to weight loss and the reduction in the risk of chronic disease. Using the data (and/or the results of your model), **recommend a course of action for the next feature to be developed to maximize participant engagement and outcomes**. The options for potential future features to be developed: 1) improvement in the online weight tracker, 2) improvement in the online activity tracker, 3) improvement in the online meal tracker, or 4) improvement in the weekly lessons. Please provide a recommendation of the above to the product team, with a clear explanation of why the data supports that decision.

Dataset

Attached is a CSV containing simulated Prevent participant data. For each participant their weekly activities (e.g. # meals tracked, # of weigh-ins, # group messages, # lessons completed, etc), are included along with demographic information (e.g. age, gender, etc) and a few metrics based on their self-reported personality questionnaire.

Data is provided for each participant through week 16. If the participant has not reached week 16, only the weekly data up to their current week is included. See the attached data dictionary for exact definitions.

Challenge

- For those participants who have not yet reached week-16 in the program, predict their % weight loss at week 16. Please include uncertainties in this prediction.
- The prediction model should be able to predict *individual* weight-losses, not just in aggregate.
- **(Bonus Challenge)** Recommending which feature development path (see options above) the product team should move forward with. This recommendation can be motivated from the output of the predictive model itself, or any other learnings that can be gleaned from the data.

Solution

- The deliverable should be a description of the model itself and any plots or summary statistics you feel necessary to describe the model, its predictions, and how it was built (including R scripts).
- **(Bonus Challenge)** Create a short (1-2 slide) presentation with your recommendation. Use the data to support your recommendation as you see fit.

Notes

- You may use any programming language, but R is preferred.
- Please include your code. Feel free to visualize or otherwise point out any interesting patterns you notice in the dataset - we're always excited about exploring our data in useful ways and give major bonus points to anyone for doing so.
- The target schedule for this project is ideally a day or two (and should only be a few hours of work); for a short-term coding challenge we're less interested in cutting edge research and more on solid approach and solid implementation.

Please don't hesitate to reach out with any questions!