Daniel Huang, SMU UbiComp Lab, June-August 2016
Contact: danielhuang74@gmail.com

List of files

csv	Alletaev.csv, Anomali.csv, and London.csv are the csv equivalents of the original .dbf files. The other files are just temporary csv files I used while coding and can probably be deleted
figs	BBT graphs, file name for each picture is the cycle ID, I believe. Not every cycle's graph has been generated (you can do it on your own in the BBT Graphs.ipynb file)
BBT Graphs.ipynb	create BBT vs Time graph
Calculate Day of Ovulation.ipynb	calculate day of ovulation based on the three over six rule, compare it to the recorded day of ovulation
Correlation Matrix.ipynb	correlation and scatter matrices
Graph Consecutive Cycles-General.ipynb	graph the consecutive cycles for a participant given her ID
Graph Consecutive Cycles.ipynb	original version that only worked for participant with ID of 9
Intro.ipynb	Introduction to the project (Start here)
Linear Regression-Many Participants?-Folding on each participant.ipynb	Model based on data for many participants (not just one), each participants' cycles are assigned to a unique fold
Linear Regression-Many Participants?.ipynb	Model based on data for many participants (not just one), 10 folds are used (you get similar results as when you fold on each participant, but it's less resource intensive)

Linear Regression-One Participant.ipynb	Model based on data for only one participant
Random Forest-OOB.ipynb	Graph OOB error rate vs n_estimators for a Random Forest Regressor using three different max_features
Random Forest-Process Residuals SDs.ipynb	Graph residuals standard deviation vs n_estimators (the data is read in from an external file random_forest_residuals_sds in the same directory)
Random Forest-SD graph.ipynb	Train Random Forest Regressors on a range from 20 to 80 n_estimators and calculate the SD of the residuals (takes a couple hours to run), the resulting SDs are stored in file random_forest_residuals_sds so you don't have to run it every time
Visualization.ipynb	Minor visualization of the data. This file also provides an example on how to import and process one of the csv data files before you can use it (as is done at the beginning of in every .ipynb file)
alletaev.dbf	
ALLETAEV.TXT	
ANOMALI.DBF	
daniel_utils.py	utilities file
figure_1.png	scatter matrix (unhelpful)
london.dbf	
LONDON.TXT	
london_datetime.csv	
London_ReadMe.pdf	
random_forest_residuals_sds	file with the residuals for various Random Forest Regressors trained with 20 to 80 n_estimators (this takes a

	few hours and is done in Random Forest-SD graph.ipynb), the file is used in Random Forest-Process Residuals SDs.ipynb to create a graph (see that file description for more details)
VARALLETAEV.pdf	
VARLONDON.pdf	