**Data Type**: GPS  
**Data Source**: Beiwe System  
**How is the data received by the dashboard**: transferred directly from users’ mobile phones

**What type of preprocessing will be necessary:** Ensuring coordinates fall within certain limits – we might not care about coordinates that fall outside of a certain region (Austin, Texas, the contiguous US, etc.)  
**What measures can be calculated**: distance traveled, dwell time at significant locations, time spent travelling  
**What information should be stored in the database (including time interval)**: GPS flights that represent travel between two significant locations, which requires: start time stamp, source latitude, source longitude, destination latitude, destination longitude, distance travelled between, duration of flight  
**Example of how the data might be used in a query**: a user wanting to find the distance that they travelled between noon and 3pm on May 5th, we would pull out the database rows corresponding to the date range and add together the distances from the various flights.

**Data Type**: Accelerometer  
**Data Source**: Beiwe System  
**How is the data received by the dashboard**: transferred directly from users’ mobile phones

**What type of preprocessing will be necessary:** Ensuring proper limits on the values   
**What measures can be calculated**: vertical distance traveled, orientation of the user  
**What information should be stored in the database (including time interval)**: Timestamp and vertical distance climbed between timestamps  
**Example of how the data might be used in a query**: a user wanting to find the vertical distance they climbed between noon and 3pm on May 5th. We would pull out the database rows corresponding to the date range and add together the positive values (meaning they were climbing upwards).

**Data Type**: Audio Recordings  
**Data Source**: Beiwe System  
**How is the data received by the dashboard**: transferred directly from users’ mobile phones

**What type of preprocessing will be necessary:** Ensuring file size or video length is greater than and/or less than a certain duration  
**What measures can be calculated**: Whatever end use the user was intending for the audio recordings  
**What information should be stored in the database (including time interval)**: Timestamp and raw audio file  
**Example of how the data might be used in a query**: a user wanted to know what they recorded on May 5th. We would pull out the database row(s) the correspond to the date and make the audio available for the user to listen to.

**Data Type**: Battery Level  
**Data Source**: Beiwe System  
**How is the data received by the dashboard**: transferred directly from users’ mobile phones

**What type of preprocessing will be necessary:** Values must be greater than 0% and less than or equal to 100%  
**What measures can be calculated**: Amount of battery lost, the state of the user’s phone (charging/not charging primarily), screen time  
**What information should be stored in the database (including time interval)**: Timestamp, battery level, the difference in the battery level between the previous timestamp and the current, state of phone at that timestamp  
**Example of how the data might be used in a query**: a user might want to know how quickly they lost battery between noon and 3on on May 5th and whether or not the phone was charging during that time frame. We would pull out the rows that correspond to the time range and add the values for loss in battery level and check to see if the state of the phone was ever in “charging”

**Data Type**: Reachability  
**Data Source**: Beiwe System  
**How is the data received by the dashboard**: transferred directly from users’ mobile phones

**What type of preprocessing will be necessary:**   
**What measures can be calculated**: Amount of time spent in a certain reachability state (cellular, WiFi, airplane, etc.)  
**What information should be stored in the database (including time interval)**: Timestamp, reachability state, time difference between current and previous state (this will tell us how long the phone was in the previous state for).  
**Example of how the data might be used in a query**: a user might be curious why they used so much data on May 5th from 9am to 5pm when they should have been connected to their work’s WiFi network. We would pull out the rows that correspond to the time range and add the values for the time spent in each type of reachability state and return both the actual time and percentage spent in every state.

**Data Type**: Reachability  
**Data Source**: Beiwe System  
**How is the data received by the dashboard**: transferred directly from users’ mobile phones

**What type of preprocessing will be necessary:**  
**What measures can be calculated**: Amount of time spent in a certain reachability state (cellular, WiFi, airplane, etc.)  
**What information should be stored in the database (including time interval)**: Timestamp, reachability state, time difference between current and previous state (this will tell us how long the phone was in the previous state for).  
**Example of how the data might be used in a query**: a user might be curious why they used so much data on May 5th from 9am to 5pm when they should have been connected to their work’s WiFi network. We would pull out the rows that correspond to the time range and add the values for the time spent in each type of reachability state and return both the actual time and percentage spent in every state.

**Data Type**: Survey Answers – Qualitative Sleep Metrics  
**Data Source**: Beiwe System  
**How is the data received by the dashboard**: transferred directly from users’ mobile phones

**What type of preprocessing will be necessary:** Omit responses that are left blank  
**What measures can be calculated**: Amount of sleep user got over certain day-range, the quality of sleep they got over a certain day-range  
**What information should be stored in the database (including time interval)**: Day, number of hours slept, numeric rating for how restful the sleep was (1-4), numeric rating for how refreshed the user felt after sleeping (1-4)  
**Example of how the data might be used in a query**: a user might want to know whether getting 7 versus 8 hours of sleep really makes them feel more refreshed in the mornings. We would pull out all the days that correspond to 7 hours or sleep and average the numeric rating of the restful question and do the same for all days when the user reported 8 hours of sleep and return the numeric rating for each situation.

**Data Type**: Survey Answers – Behavior Questions  
**Data Source**: Beiwe System  
**How is the data received by the dashboard**: transferred directly from users’ mobile phones

**What type of preprocessing will be necessary:** Omit responses that are left blank  
**What measures can be calculated**: How often students were engaging in a certain activity

(from a dropdown menu), in a certain place (from a dropdown menu), with certain people (from a dropdown menu), interacting in a certain way (from a dropdown menu)  
**What information should be stored in the database (including time interval)**: Day, number of times user chose a specific answer  
**Example of how the data might be used in a query**: a user might want to know how often they were in class versus at home or spending time with friends versus being alone in a certain week. We would pull out the rows from each day in the time range and sum the numbers in column that correspond to a specific answer choice and return a percentage for each activity in activities, place in places, group in people, and interaction in interactions.

**Data Type**: Survey Answers – Mood Questions  
**Data Source**: Beiwe System  
**How is the data received by the dashboard**: transferred directly from users’ mobile phones

**What type of preprocessing will be necessary:** Omit responses that are left blank  
**What measures can be calculated**: How often a student felt lonely, stressed, content, and sad on a scale of 1 to 4 and an aggregate mood score  
**What information should be stored in the database (including time interval)**: Day, number of times user chose a specific value for their mood  
**Example of how the data might be used in a query**: a user might be curious about their mental health and want to know how often they felt sad, lonely, and stressed during a particular week. We would pull rows from the days in that time range and report a percentage for how often a user felt a certain way.

**Data Type**: Survey Answers – Energy Level  
**Data Source**: Beiwe System  
**How is the data received by the dashboard**: transferred directly from users’ mobile phones

**What type of preprocessing will be necessary:** Omit responses that are left blank  
**What measures can be calculated**: How a user’s energy level varies throughout the day  
**What information should be stored in the database (including time interval)**: Timestamp, time of day (early morning, morning, midday, afternoon, evening), and energy level (1-5)   
**Example of how the data might be used in a query**: A user might want to know at what time during the day they have enough energy to perform complex tasks. We would average the energy levels from each time of day category (which would be based on the timestamp), and report the values to the user.

**Data Type**: Survey Timings  
**Data Source**: Beiwe System  
**How is the data received by the dashboard**: transferred directly from users’ mobile phones

**What type of preprocessing will be necessary:** Omit answers that are received a certain amount of time after distribution  
**What measures can be calculated**: Time of day user typically answers surveys, typical time it takes for users to complete each survey  
**What information should be stored in the database (including time interval)**: Survey ID, Beginning timestamp, ending timestamp, and the time difference between the two  
**Example of how the data might be used in a query**: A user might like to know how long they spend answering surveys in a day. We would pull the data from the rows within a certain time frame and report the average time spent answering the survey based on the survey ID.