

The Data

For both challenges, you will need to assign students to stops, and link stops through routes, all while operating within our current fleet of buses. Posted on our website you will find a fake dataset in a .xls file. We have created fictitious students, assigned them to random households, and then assigned those fictitious students to fictitious schools (all named after players in the Red Sox Hall of Fame). The first tab, "Student Information," includes the same headers that we use in the real file that we will share with you later in the process. Those headers include:

- **Student ID**
- **Street Number**
- **Street Name**
- **Zip Code**
- **Full Address**
- **Latitude**
- **Longitude**
- **Pickup Type: Corner or Door-to-Door**
- **Grade**
- **Geocode:** BPS tracks various neighborhoods by census 'geocodes' which we use to analyze student neighborhoods.
- **Neighborhood Safety Score:** This is based on the amount of reported violent crime in the neighborhood -- 7 reflects a high level of reported violent crime and 1 represents a low level of reported violent crime.
- **Proposed Maximum Walk to Stop Distance:** This is the maximum distance that this student should walk to a stop. This is not straight line distance but rather actual walking distance along a common path. Each student will have a specific maximum walking distance which varies based on their grade and the rates of reported crime within their neighborhood.
- **Assigned School:** This is the school that the student is required to be transported to and from each day.
- **Current School Start Time:** Students need to arrive at school between 10-15 minutes before the start of school (no earlier and no later). While the bus needs to arrive 10-15 minutes before the start of school, the bus should plan to leave at the start of school. This allows time for buses to unload and builds in "recovery time" to account for unexpected delays on the route. (More detail in *Appendix 1: Rules of the Road*)
- **Current School End Time:** The time between school start time and school end time is the length of the school day for that school. Assume all students go home at the end of the day (i.e., they attend no after school activities). As detailed in *Appendix 1: Rules of the Road* buses should arrive by the school's end time and leave no later than 15 minutes afterward.
- **School Address**
- **School Latitude**
- **School Longitude**

Below these headers is a series of rows, in which each row represents exactly one student. Assume all students need to be transported to and from school.

There is a second tab, “Bus Yards,” that provides real information about our buses and their assigned bus yards. This information will not change as the competition proceeds, and it includes the following headers:

- **Bus Number**
- **Bus Yard:** An abbreviated version of the name of the bus yard.
- **Bus Yard Address:** The location at which each bus must start and end the day. Be sure to include the drive time distances between this bus yard and the first and last trip for this bus.
- **Bus Yard Latitude**
- **Bus Yard Longitude**

The real dataset will include two additional datasets that are not fully included in the fictional datasets: the list of our current corner bus stops and the list of our current routes. We are not including this in the fictional dataset because it is very difficult to meaningfully assign fictional students to fictional stops and fictional stops to routes in a way that mirrors reality. However, you will be able to see the headers that we will share with you in the third (“Stop Locations”) and fourth (“Route Assignments”) tabs.