Data Assumptions and Mock Data

Our team brainstormed how the prototype would address the needs expressed through the caseworker Q&A session and prospective user interviews. The recurring themes that emerged were for enhanced search criteria. In particular, users were interested in the ability to find providers by:

- Location/radius
- Age of children served
- Gender
- Hours and days of operation
- Behavioral and medical needs
- USDA Food program/Subsidy

To address this in the prototype, we took into account the feasibility of obtaining the data. In particular, we chose variables that could be obtained easily, such as fields currently captured in the state child care license application.

Additionally, we used publicly available location data. Since the vendor challenge dataset did not include provider addresses, we used Google services to identify the latitude and longitude associated with the zip codes for each provider.

We augmented the vendor challenge data to include:

Variable	**Description**
MinAge	Minimum age of children served
MaxAge	Maximum age of children served
Gender	Genders served
Longitude	Longitude of the location
Latitude	Latitude of the location
HoursofOperation	Provider hours of operation
DaysofOperation	Provider days of operation
CanTakeChildrenWithBehavioralProblems	Provider serves children with special behavioral needs
CanTakeChildrenWithMedicalProblems	Provider serves children with special medical needs
USDAFoodPrograms	Provider accepts USDA food program

Our team developed a tool to populate the file with mock data records for each provider. The tool read the .csv file, added the fields, calculated values for the additional variables, and generated a new file in JSON format for the prototype.

For the purposes of the vendor challenge, we decided to proceed with assumptions about the data and its availability. For future task orders, we would work closely with the Mississippi Department of Information Technology Services (ITS) to identify the necessary data elements and use real data instead of generating sample data.