Accidental Drug Related Death Azure Explanation

1. The first step is to upload the dataset. We pulled our dataset from the accidental drug related death.
2. For our experiment we selected specific columns that we wanted to focus on, such as age, sex, race, residence city, death city, and location.
3. After selecting the columns, we introduce split data. The purpose of split data is to prevent us from overfitting the data. We are not sure if the data we have is a prefect random sample.
4. We use linear regression to establish a relationship between our variables.
5. In our next step we used train model to select variables that we wish to predict. In our first experiment we focused on age. In our second experiment we focused on gender.
6. After the model is trained, we can then score it. For our next step we must include score model which should be connect to split data in order to generate a predicted score.
7. For our last step, we need to evaluate the model. The purpose of evaluating the model is to see how well our prediction was.