I will be exploring one of Kaggle’s competitions: Instacart Market Basket Analysis. The objective is to predict the products in a user’s next order. I collected the data from Kaggle site, which is provided and open sourced for whomever wants to compete. The data is setup relationally with historical customers’ orders in the past. The data provided has aisles and departments descriptions. Also provided, are the products that were ordered in a specific order\_id and whether the item was reordered or not in boolean logic. The order hour of the day and days since prior order are also provided in the dataset. The days since prior order is especially important to this model to predict in the model if a user will need to order a certain product because it’s reaching this number. Days since prior order is also very subjective since each customer’s usage rate for the item can fluctuate. From the test set data, I will need to predict the products for the respective order\_id. I believe I have what I will need to answer this question.