Scenario 1: Task 3 Medical Market Basket (Readmission)

In the medical industry, readmission of patients is such a problem that an external organization penalizes hospitals for excessive readmissions (Centers for Medicare and Medicaid Services or CMS). When it comes to readmission penalties, studies show that many hospitals are overconfident and underprepared. The percentage of hospitals penalized for readmissions has increased each year since CMS began imposing penalties, and according to the CMS reporting, as much as 78 percent of hospitals were fined in fiscal year 2015. However, three-quarters of hospitals feel confident in their ability to reduce readmissions, and only 55 percent of them anticipate receiving a penalty this year. Given the historical trend and the addition of COPD and Hip and Knee replacement to the list of medical conditions measured, the percentage of hospitals penalized will likely be much higher than 55 percent. Additionally, although hospitals are applying various reduction strategies, fewer than 1 in 5 utilize technology that is specific to reducing their readmissions, so they may not be doing all that they can.

You are an analyst on a team of analysts for a popular medical hospital chain with patients in almost every state in the United States. Executives are interested in learning more about the typical prescriptions that are associated with patients. You have been asked to perform a market basket analysis on historical prescription data of patients for this chain of hospitals.

Data File being used:
medical_market_basket.csv

Data Dictionary:

The data set includes the following information:

- Prescriptions from previous patients (prior 2 years) as based on historical data
- Prescription history reflects all historical prescriptions prescribed at least once (not just at a single moment; also, history might not reflect current prescription status)
- Data may include both drugs, prescribed vitamins, etc. The reason for the prescription is not provided.
- Data might need some cleaning and preparation (depending on tools used)

The data set consists of 7,501 patient prescription history and 20 columns/variables:

- **Presc01–Presc20:** Prescription history of patients in historical data