

## Project 4: Association Rule Mining

Due: see Canvas  
Points: 100

**Please submit your report in PDF format.**



### ***Introduction***

For this project, we will look for associations in hospitalization data. We will use one last time the same data.

### ***Follow the CRISP-DM Framework for your Report***

#### **3. Data Preparation [30 points]**

- Describe different ways to discretize your continuous variables? What are the advantages of each approach?
- Construct the needed transaction data set(s) using the most advantageous discretization approach for continuous variables.

#### **4. Modeling [40 points]**

- Create frequent itemsets **and** association rules.
- Use tables/visualization to discuss the found results.

#### **5. Evaluation [10 points]**

- What findings are the most interesting? Do they agree with your findings using classification and clustering?

#### **6. Deployment [10 points]**

- How useful are your results to the health care industry?

#### **Exceptional Work [10 points]**

- Example: Create separate transaction sets for subsets of the data (e.g., gender differences, differences for some primary condition groups) and compare the found rules.