

Project 1. Solution

Santiago Castro Dau, June Monge, Rachita Kumar, Sarus

2/25/2022

Problem 1: Conditional independence and BNs

Problem 2: Markov blanket

Problem 3: Learning Bayesian networks from protein data

a)

```
## Set seed
set.seed(2022)

## Read data
data <- vroom(file = "sachs.data.txt", show_col_types = FALSE)

## Visualize data dimensions.
dim(data)

## [1] 853  11
•  $N = 853$ 
•  $n = 11$ 

## Split data
data = data %>% rowid_to_column(var = 'row_id')
train = data %>% sample_frac(0.8, replace = FALSE)
test = anti_join(data, train, by = "row_id")

## Initializing socre object
score_object = scoreparameters(scoretype = 'bge', train)
```