Importance Sampling Experiment 2

4/3/2022

Contents

Executive Summary

Note

This document differs from `imp_samp_exp_sum.pdf` in that the model averaging here is weighted differently. As briefly mentioned in the section titled "Fully Bayesian Approach", the importance weights are calculated using the sum of the ELBO across the individuals for each variable fixed as the response. This apparently helps with the issue observed in the case when the weights were calculated as specific to the individual, which led to the independence graph (precision matrix with all zero's) being estimated for many of the individuals.

Summary

In this experiment, I compare a fully Bayesian model averaging approach to grid search. I find that the results for the fully Bayesian model have a similar sensitivity to grid search, however, have significantly lower specificity.

Experiment Overview

This experiment compares three methods of hyperparameter specification for `covdepGE` across 10 trials.

In each of the trials, a 1,000 point grid of hyperparameters was generated as described in the *Grid generation* section.

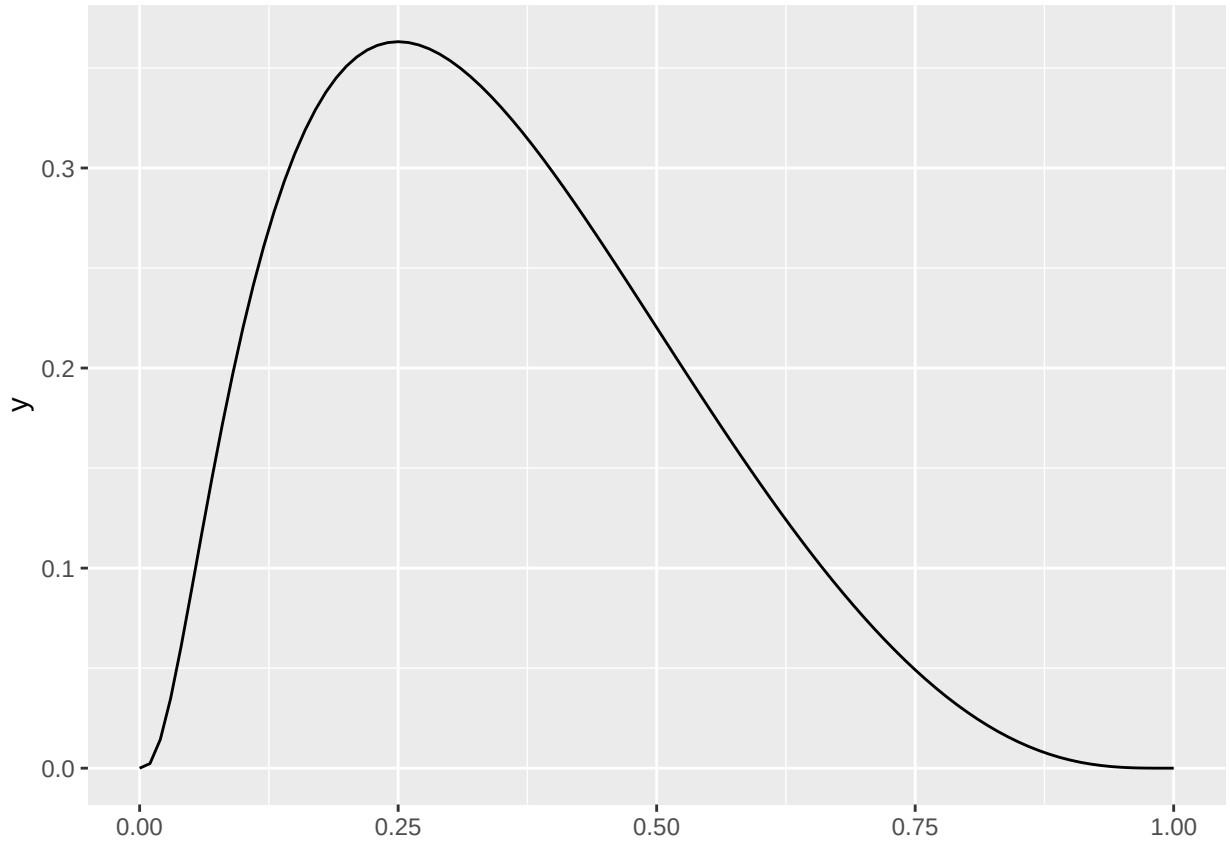
The first 2 methods (`prior` and `unif_prior`) employ a fully Bayesian approach of model averaging using importance weights, while the third method (`grid`) employs grid search.

`prior` placed non-uniform prior densities on each of the hyperparameters, while `unif_prior` used a uniform density for each. Both methods used a uniform importance density over each of the 1,000 grid points.

`prior` prior specification

Although the prior placed on σ^2 was a flat prior, the other two hyperparameters had proper, non-uniform priors.

First, the prior on π was induced by a normal prior on the log-odds of π. The mean, -1.1, corresponded to $\pi = 0.2$, while the standard deviation, 1.1, was chosen so that the log-odds of $\pi = 0.5$ was within one standard deviation of the mean. This prior is visualized below:



Next, the prior on σ^2_β was induced by the prior proportion of variance explained. First, a uniform(0, 1) prior is placed on the prior proportion of variance explained, $\hat{r}$, which is defined by:

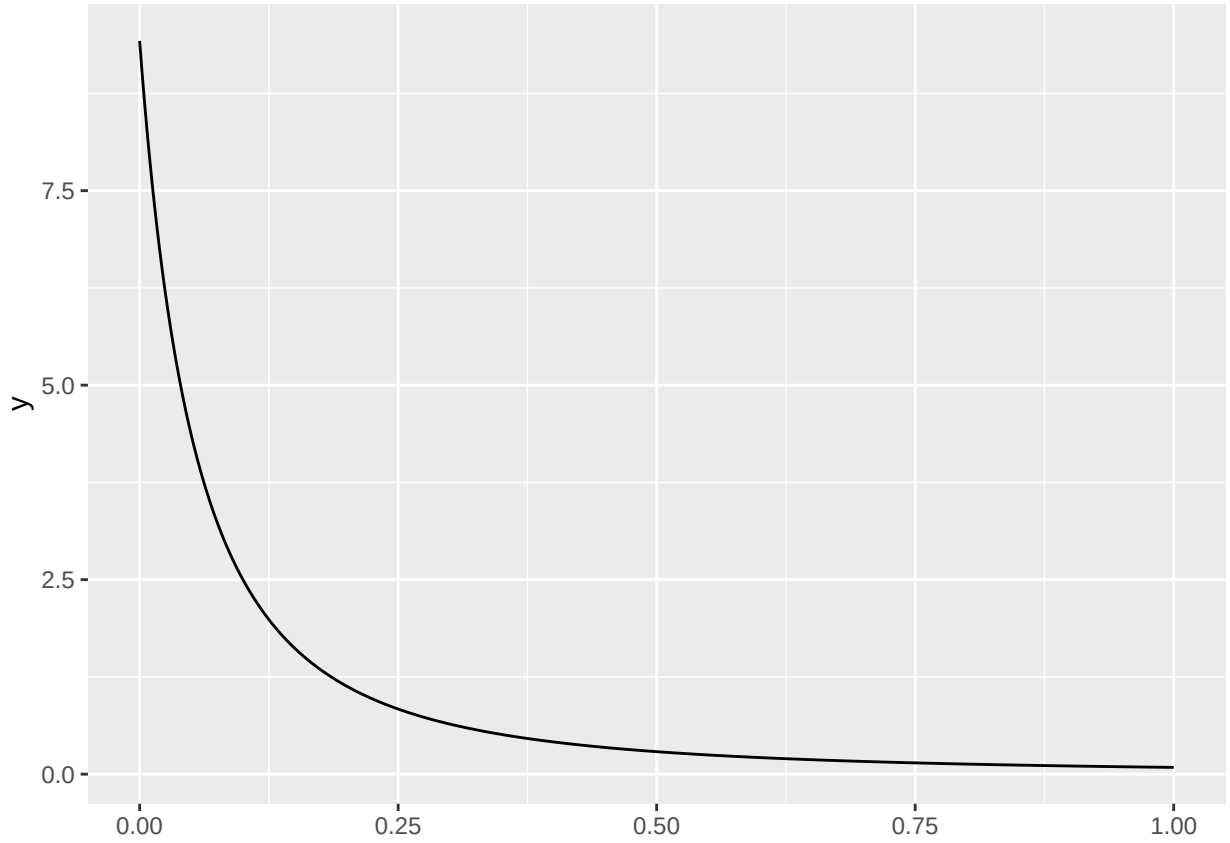
$$\hat{r} = \frac{\hat{s_z}}{1 + \hat{s_z}} \qquad \hat{s_z} = \sigma^2_\beta * \pi * T \qquad T = \sum_{k=1}^{p} s_k^2 \tag{1}$$

Where s_k^2 is the sample variance of the k-th column of X.

This prior, which is dependent on the prior value of π, induces the following prior on σ^2_β:

$$\sigma^2_\beta \sim f(\cdot|\pi) \qquad f(\sigma^2_\beta|\pi) = \frac{\pi \sum_{k=1}^{p} s_k^2}{(1 + \sigma^2_\beta \pi \sum_{k=1}^{p} s_k^2)} \mathbb{1}_{(0,\infty)}(\sigma^2_\beta) \tag{2}$$

This prior is visualized below for $\pi = 0.2$ and $T = 3$.



Fully Bayesian Approach

This section describes the model averaging approach used here for a spike-and-slab regression of $y \in \mathbb{R}^n$ on $X \in \mathbb{R}^{n \times p}$ to compute an importance-weighted-average of the posterior inclusion probabilities $P(\gamma_k = 1|y, X, \theta_j)$, where $\gamma_k = 1 \iff \beta_k = 1$.

Denote the importance weights by $\underset{\sim}{w} = w(\theta_1), ..., w(\theta_N)$, where $\theta_1, ..., \theta_N$ are points the points comprising the hyperparameter grid.

With the j-th column of the data as the response y and the data with the j-th column removed as the covariates X, an extraneous-covariate-weighted regression will be performed with respect to individual l for each $\theta \in \underset{\sim}{\theta}$. The importance-weighted sum of the posterior quantities $p(\gamma_k = 1|X, y, \theta_i)$ will then be computed using:

$$w(\theta_i) = \frac{p(y|X, \theta_i)p(\theta_i)}{q(\theta_i)} \tag{3}$$

Where $p(y|X, \theta_i), p(\theta_i)$, and $q(\theta_i)$ are the marginal log-likelihood, the prior density, and the importance density, respectively. The prior density and the importance density are defined as:

$$p(\theta_i) = p(\sigma^2{}_i) * p(\sigma^2_{\beta_i}) * p(\pi_i) \qquad q(\theta_i) = q(\sigma^2{}_i) * q(\sigma^2_{\beta_i}) * q(\pi_i) \tag{4}$$

In the case of `prior`, $p(\sigma^2_{\beta_i})$ should actually be written hierarchically as $p(\sigma^2_{\beta_i}|\pi_i)$. Also, in both `prior` and `unif_prior`, $q(\theta_i)$ is the uniform density on the hyperparameter grid. Then, the weighted importance sum coverges almost surely to the posterior probability of inclusion:

Proposition 1. *Let* $w(\theta_i) = p(y|X, \theta_i)$. *Then:*

$$\lim_{N\to\infty} \frac{\sum_{i=1}^{N} p(\gamma_k = 1|X, y, \theta_i) w(\theta_i)}{\sum_{i=1}^{N} w(\theta_i)} = p(\gamma_k = 1|X, y), \;\; a.s. \tag{5}$$

The approximation of $p(y|X, \theta_i)$ is made using $\exp(ELBO_j(\theta_i))$, where $ELBO_j(\theta_i)$ is defined as:

$$ELBO_j(\theta_i) = \sum_{l=1}^{n} ELBO_j^l(\theta_i) \tag{6}$$

And $ELBO_j^l(\theta_i)$ is the log-lower bound of the marginal likelihood $p(y|X, \theta_i)$ in the extraneous-covariate-weighted regression of y on X with respect to individual l.

Grid generation

Let X be the data with the j-th column removed, and y be the j-th column of the data. This section describes how a grid of the following hyperparameters for a spike-and-slab regression of y on X is generated.

σ^2: residual error variance
σ_β^2: regression coefficient variance
π: prior probability of inclusion

Each of the hyperparameters is assumed to be bounded in the intervals:

$$\sigma^2 \in [0.00001, \underset{max}{\sigma^2}] \qquad \sigma_\beta^2 \in [0.00001, \underset{max}{\sigma_\beta^2}] \qquad \pi \in [0.01, \underset{max}{\pi}]$$

The rough bound on σ_β^2 is induced by a bound on the signal to noise ratio of 25:

$$\underset{max}{\sigma_\beta^2} = \frac{25}{\hat{\pi} \sum_{k=1}^{p} s_k^2} \tag{7}$$

The bound on σ^2 is taken to be $\underset{max}{\sigma^2} = 10 s_y^2$.

Then the grid for the hyperparameter θ is generated by first creating a uniform grid on $[\log \underset{min}{\theta}, \log \underset{max}{\theta}]$ and then exponentiating the resulting values. This results in more densely packed grid points on the lower side of the interval. For example, for $\sigma_\beta^2 \in [1e^{-5}, 10]$, a grid of length 5 would be generated as:

```
exp(seq(log(1e-3), log(10), length.out = 5))
```

```
## [1] 1e-03 1e-02 1e-01 1e+00 1e+01
```

Then, the final grid is created by taking the cartesian product between each of the individual grids:

$$\mathcal{S} = \underset{\sim}{\sigma^2} \times \underset{\sim}{\sigma_\beta^2} \times \underset{\sim}{\pi} \tag{8}$$

Results

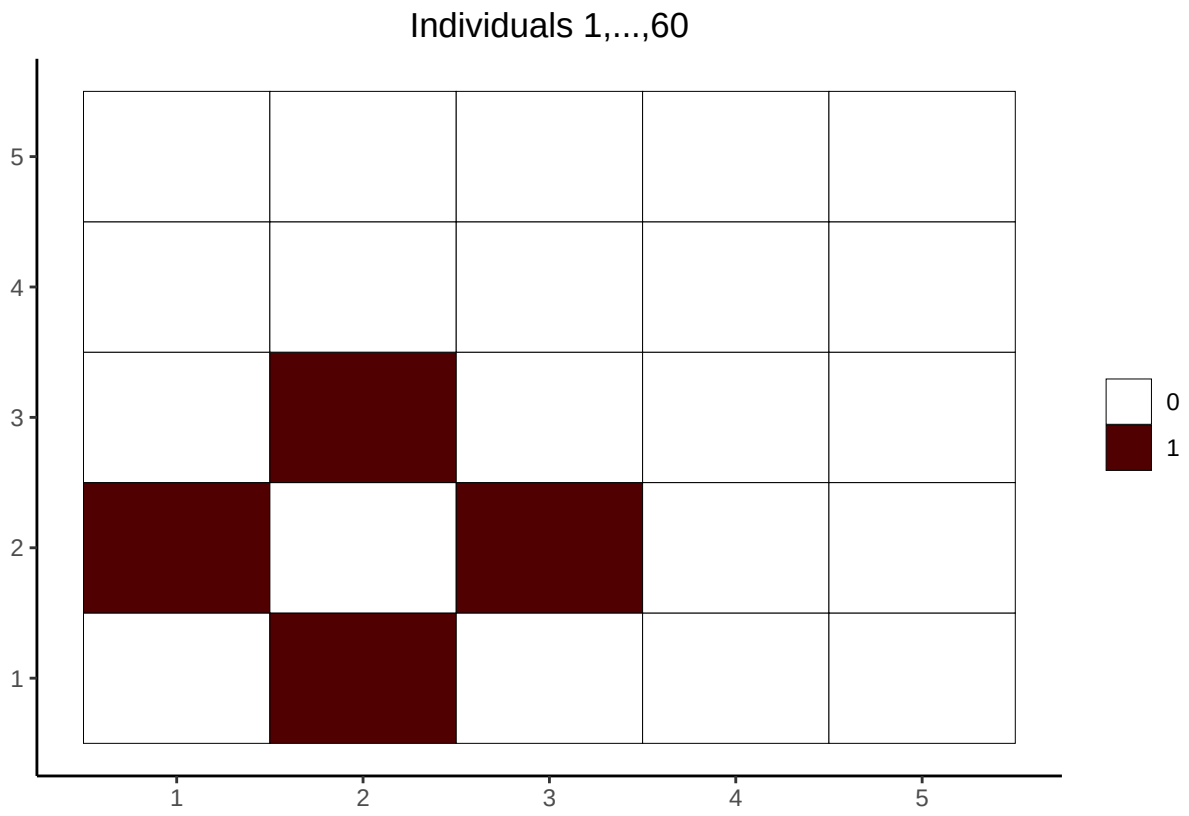
Detailed trial results are given below. Although the issue with independence graphs has been resolved, there is now an issue of too many false positives for both of the model averaging approaches.

Another potential issue is that it is unclear whether proposition 1 will still hold, since I am now approximating the marginal log-likelihood with the total ELBO across all individuals, instead of calculating weights to be individual-specific.

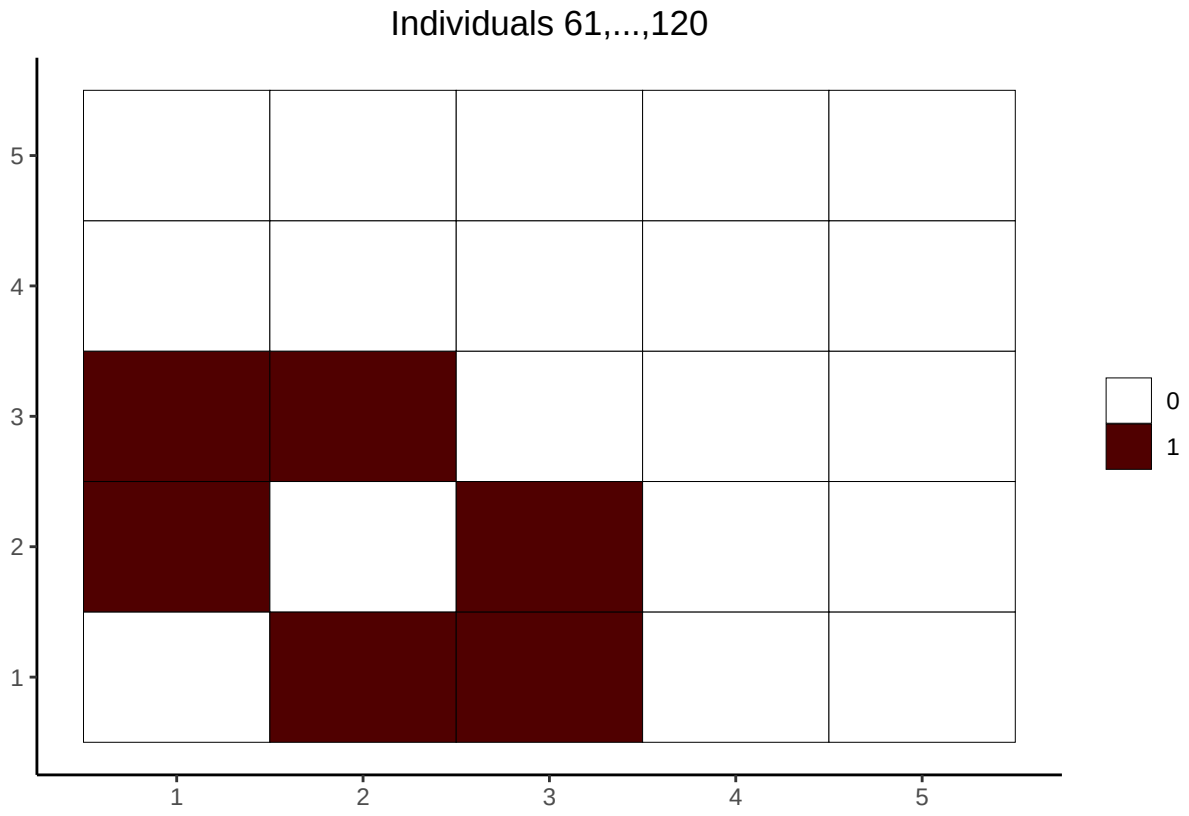
Detailed Results

I first visualize the true graphs:

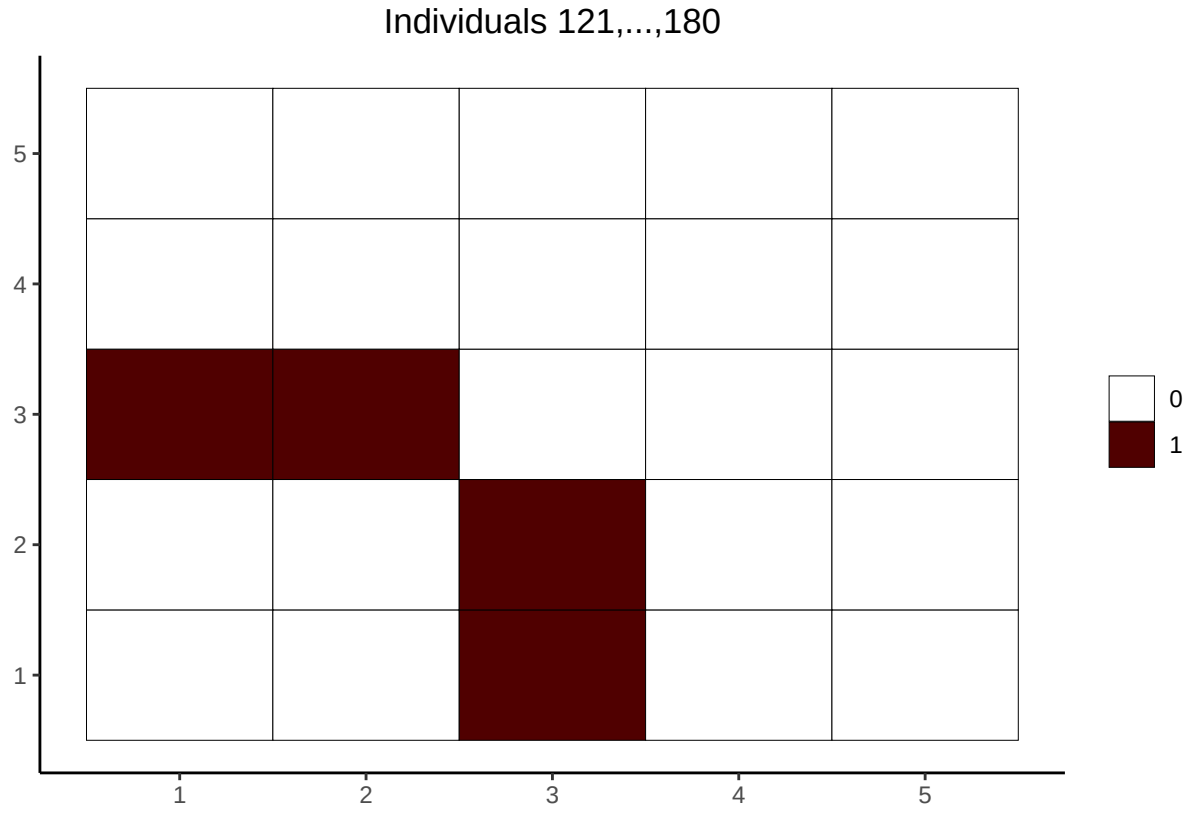
```
## [[1]]
```




```
##
## [[2]]
```




```
##
## [[3]]
```

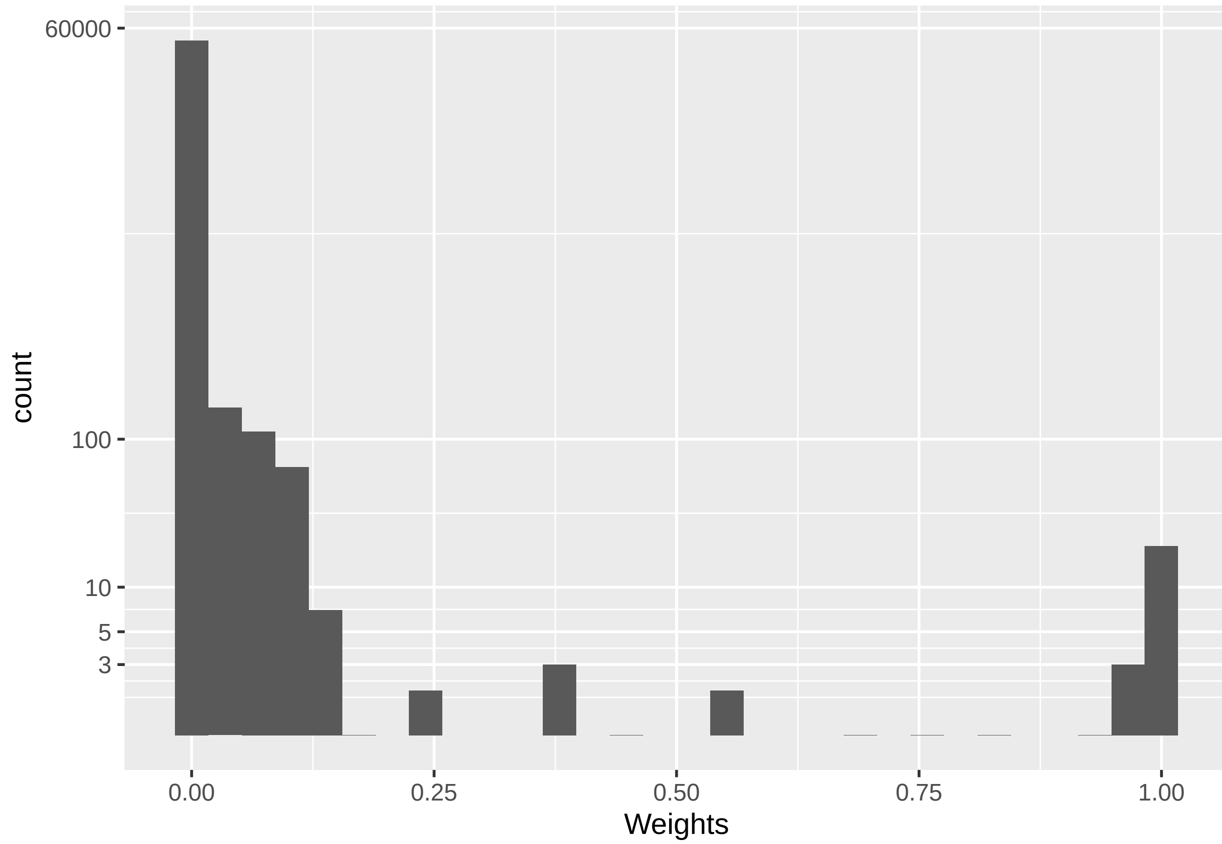


Next, I visualize the distribution of the weights in all trials. Since there are 1,000 hyperparameters, 10 trials, and 5 variables, there are 50,000 total weights in this histogram.

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

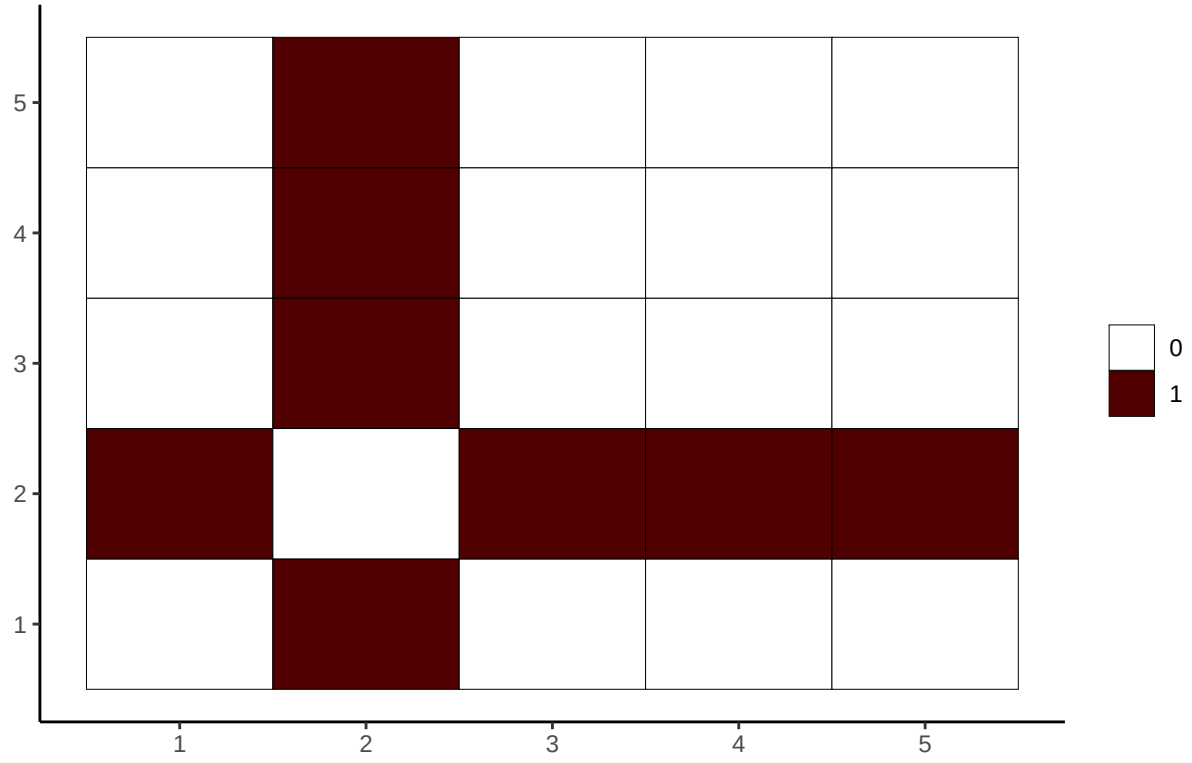
```
## Warning: Transformation introduced infinite values in continuous y-axis
```

```
## Warning: Removed 14 rows containing missing values (geom_bar).
```

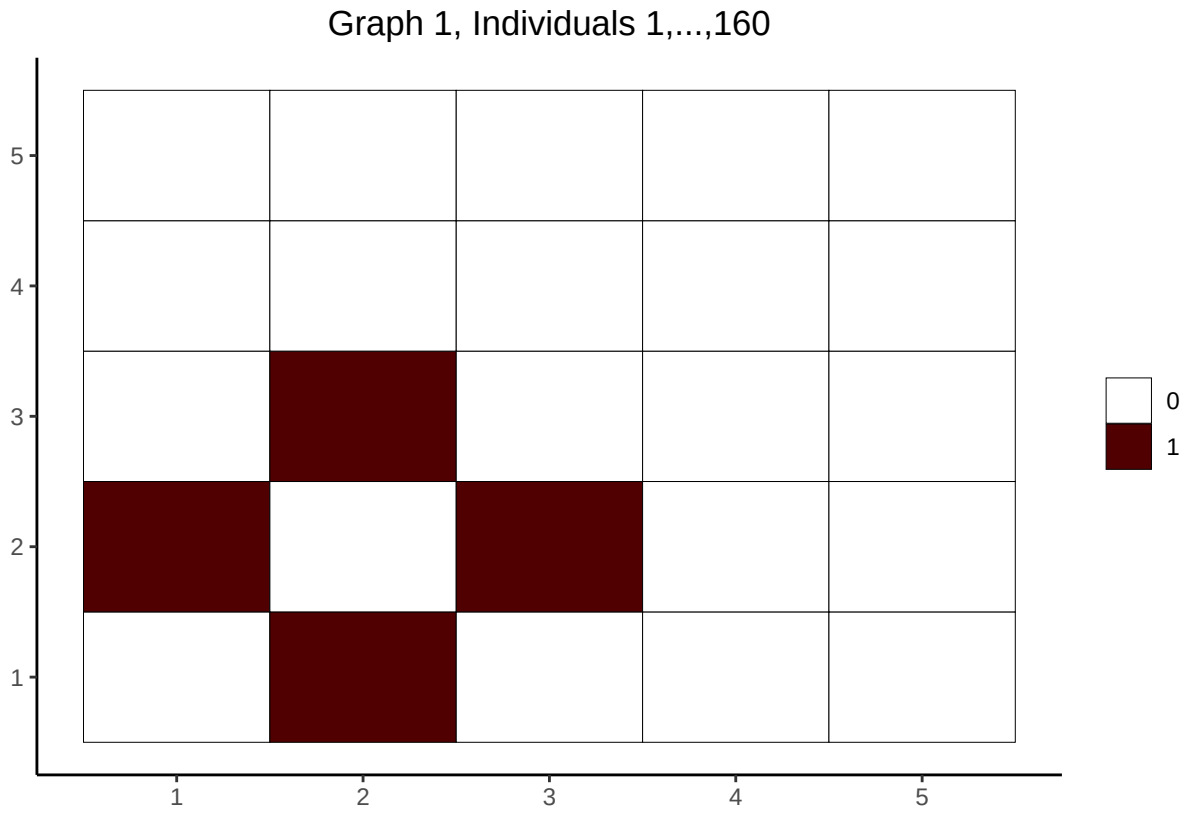


Now, the graphs for each of the trials:

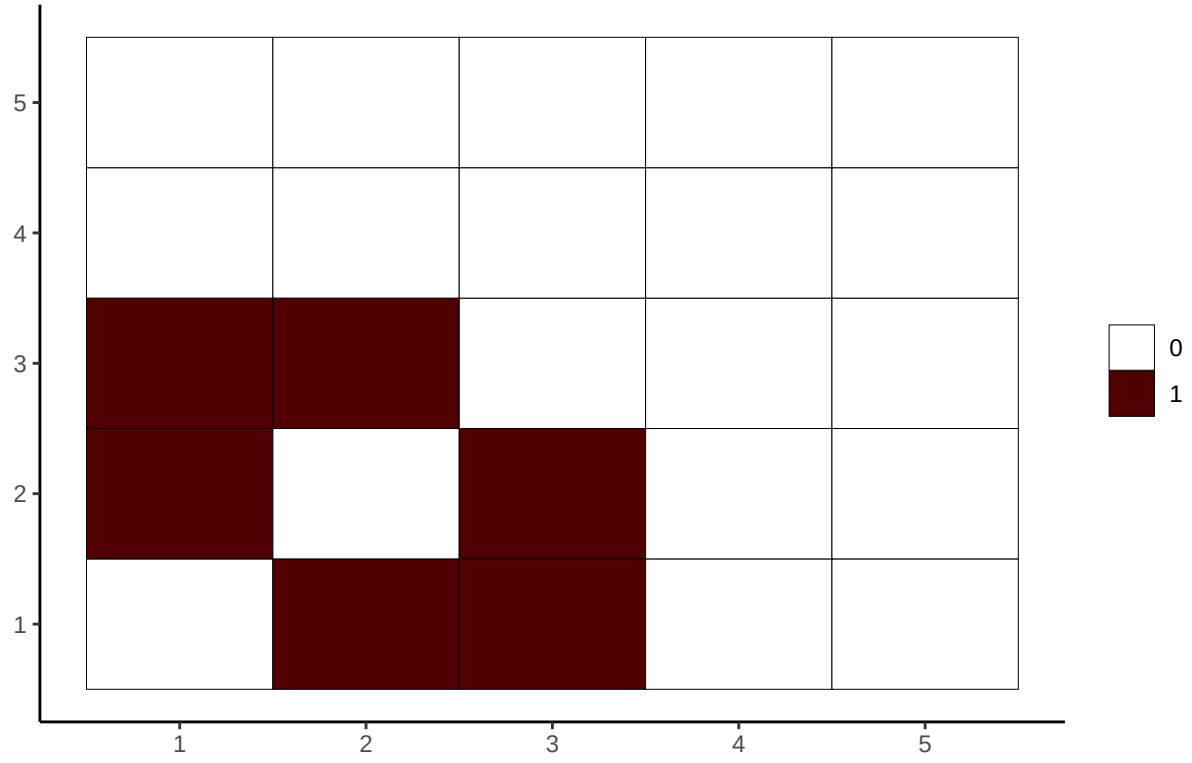
```
## ----------------------------------------------------------------
## TRIAL  1
##
## prior
##                       Covariate Dependent Graphical Model
##
## Model ELBO: -9334.04                Unique conditional dependence structures: 1
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 1.893 mins
## [[1]]
```



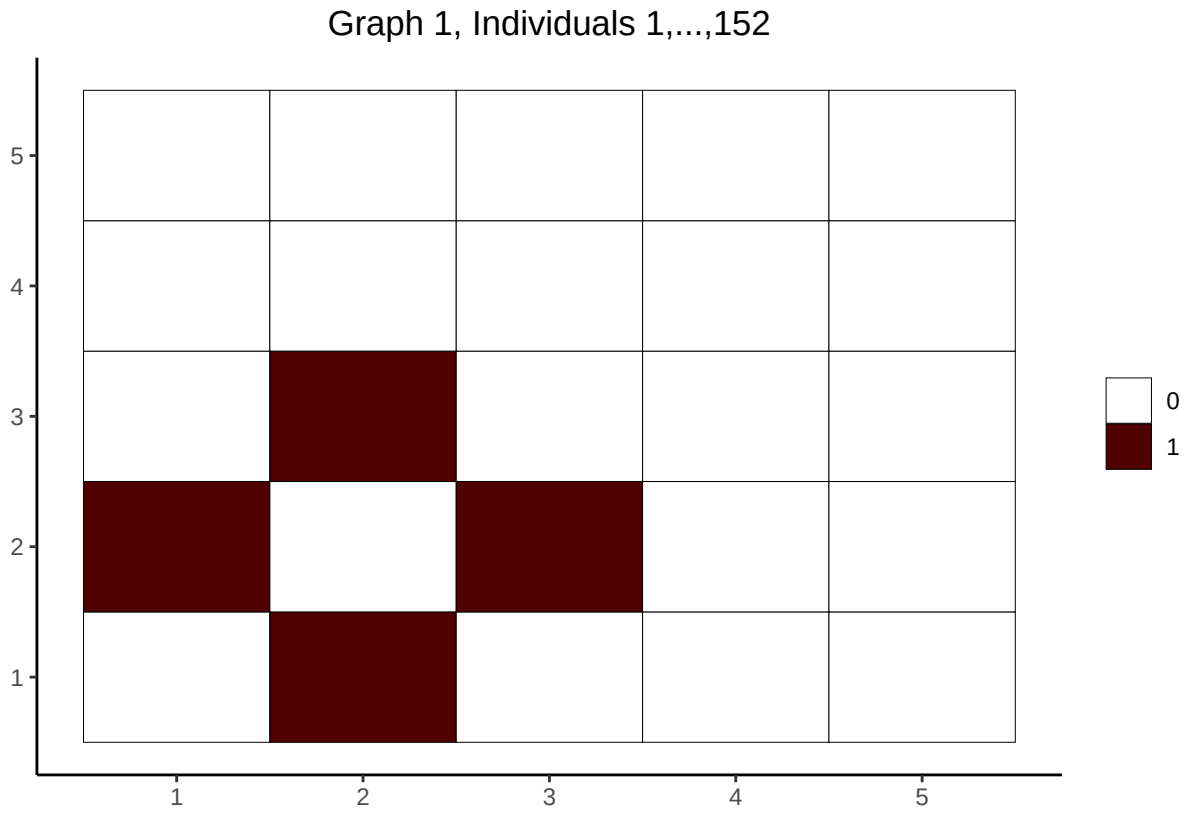

```
##
##
##
## TRIAL  1
##
## unif_prior
##                        Covariate Dependent Graphical Model
##
## Model ELBO: -9333.74                Unique conditional dependence structures: 2
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 1.867 mins
## [[1]]
```



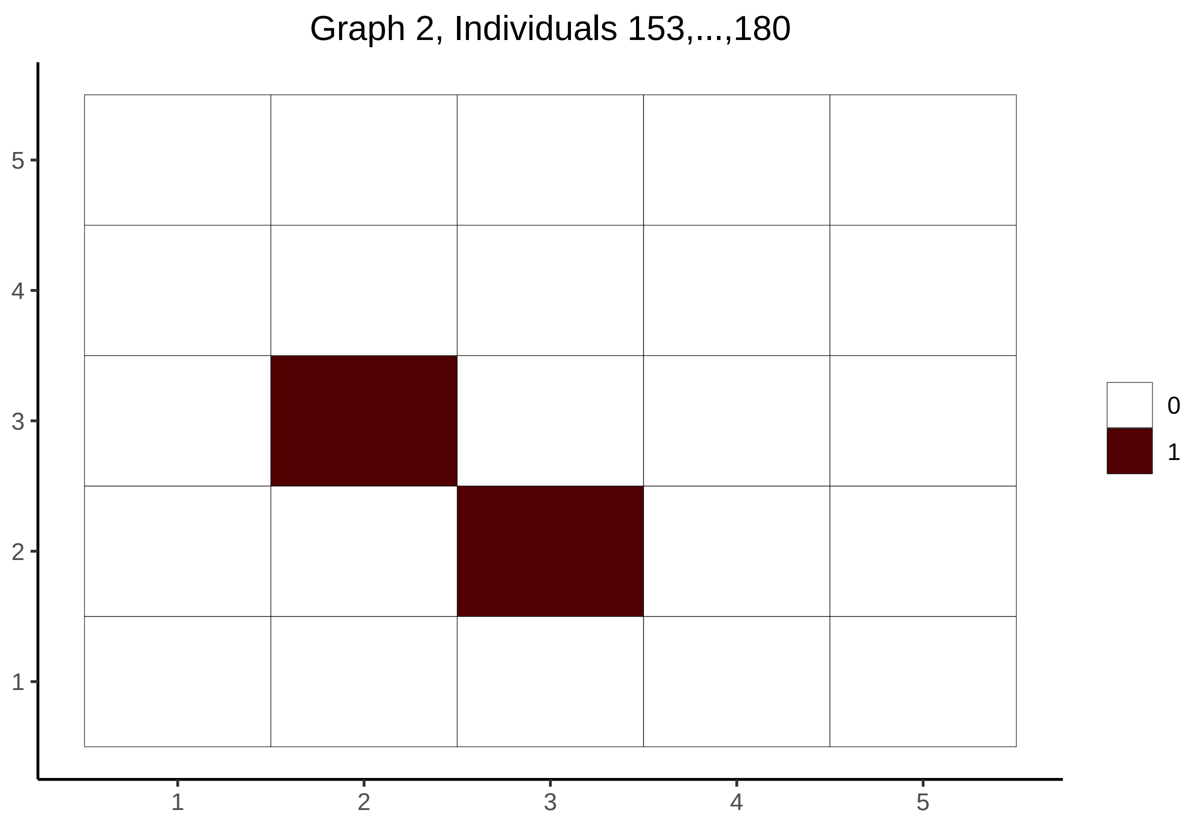

```
##
## [[2]]
```



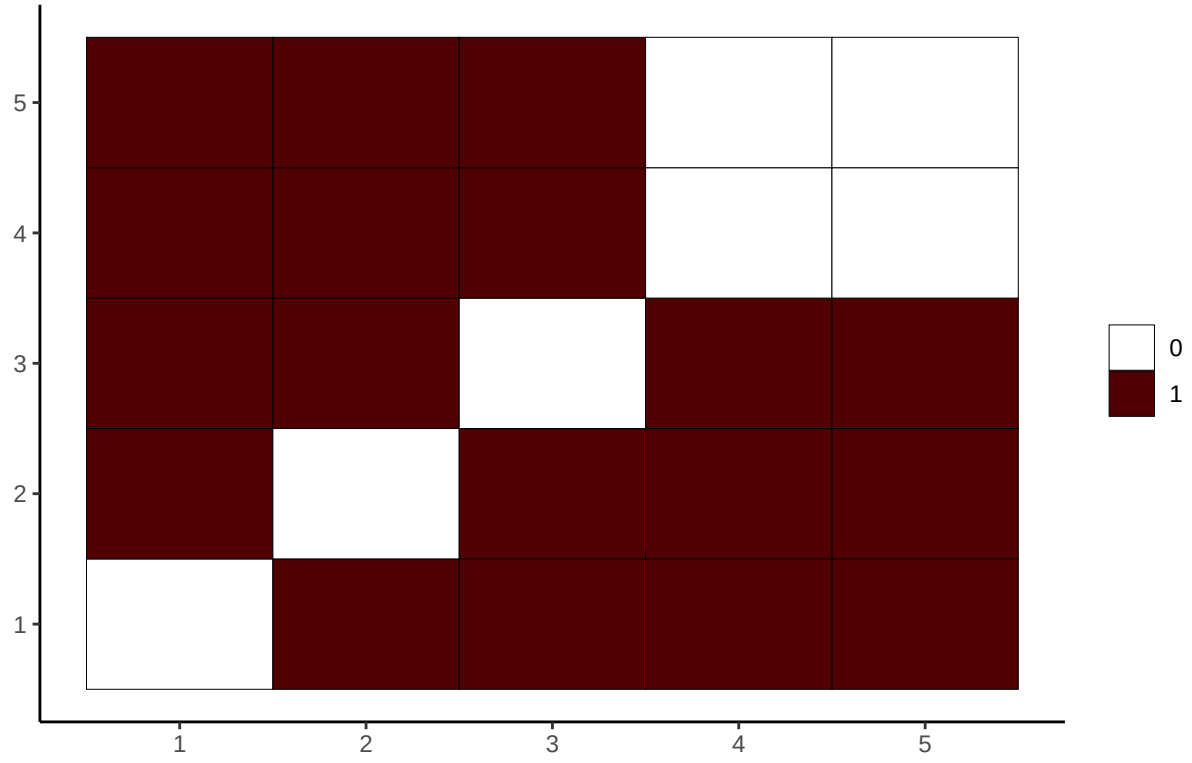

```
##
##
##
## TRIAL  1
##
## grid
##                        Covariate Dependent Graphical Model
##
## Model ELBO: -9332.56                Unique conditional dependence structures: 2
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 54.577 secs
## [[1]]
```



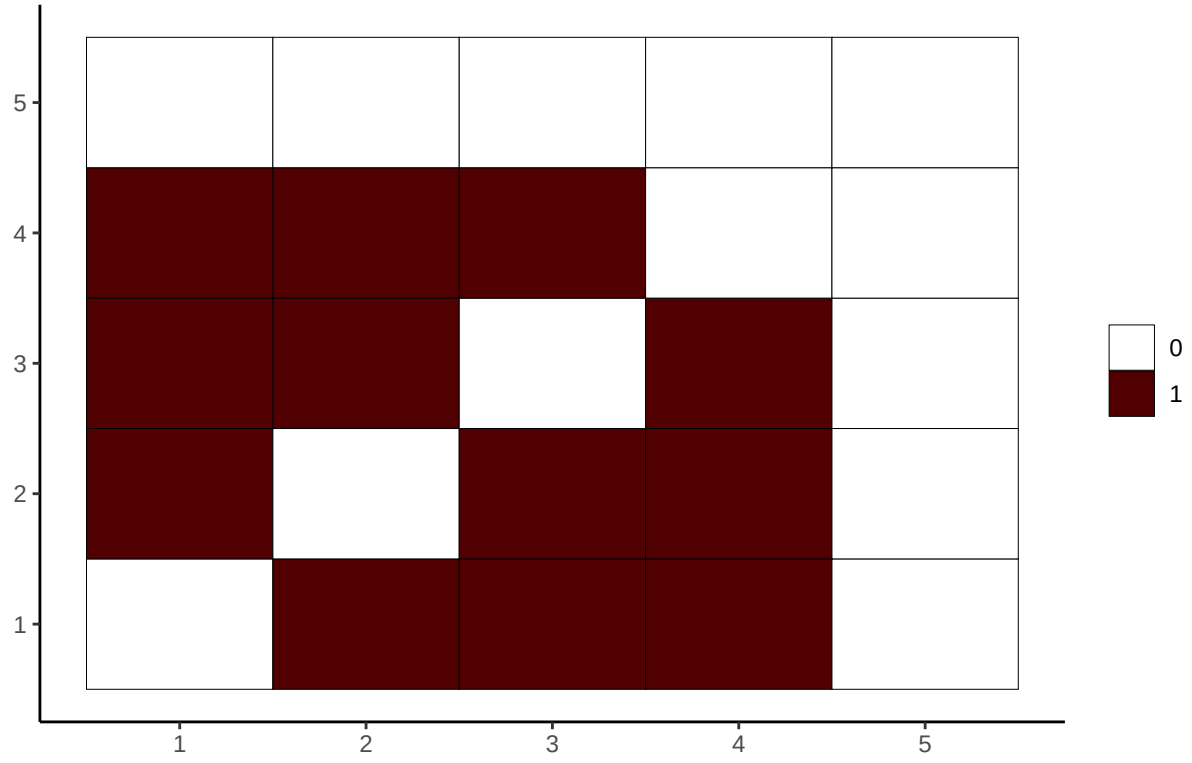

```
##
## [[2]]
```




```
##
## - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
## TRIAL  2
##
## prior
##                       Covariate Dependent Graphical Model
##
## Model ELBO: -9129.6                Unique conditional dependence structures: 1
## n: 180, variables: 5                     Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 2.166 mins
## [[1]]
```

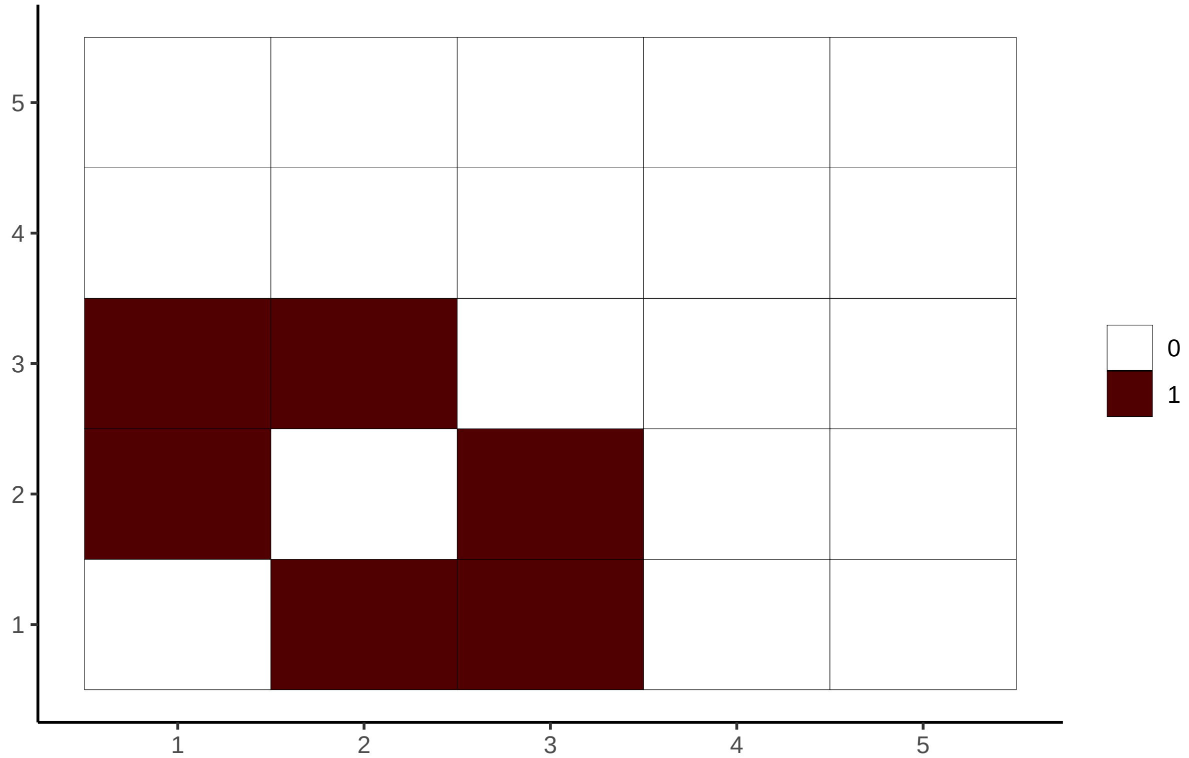



```
##
##
##
## TRIAL  2
##
## unif_prior
##                        Covariate Dependent Graphical Model
##
## Model ELBO: -9129.49                Unique conditional dependence structures: 1
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 2.101 mins
## [[1]]
```

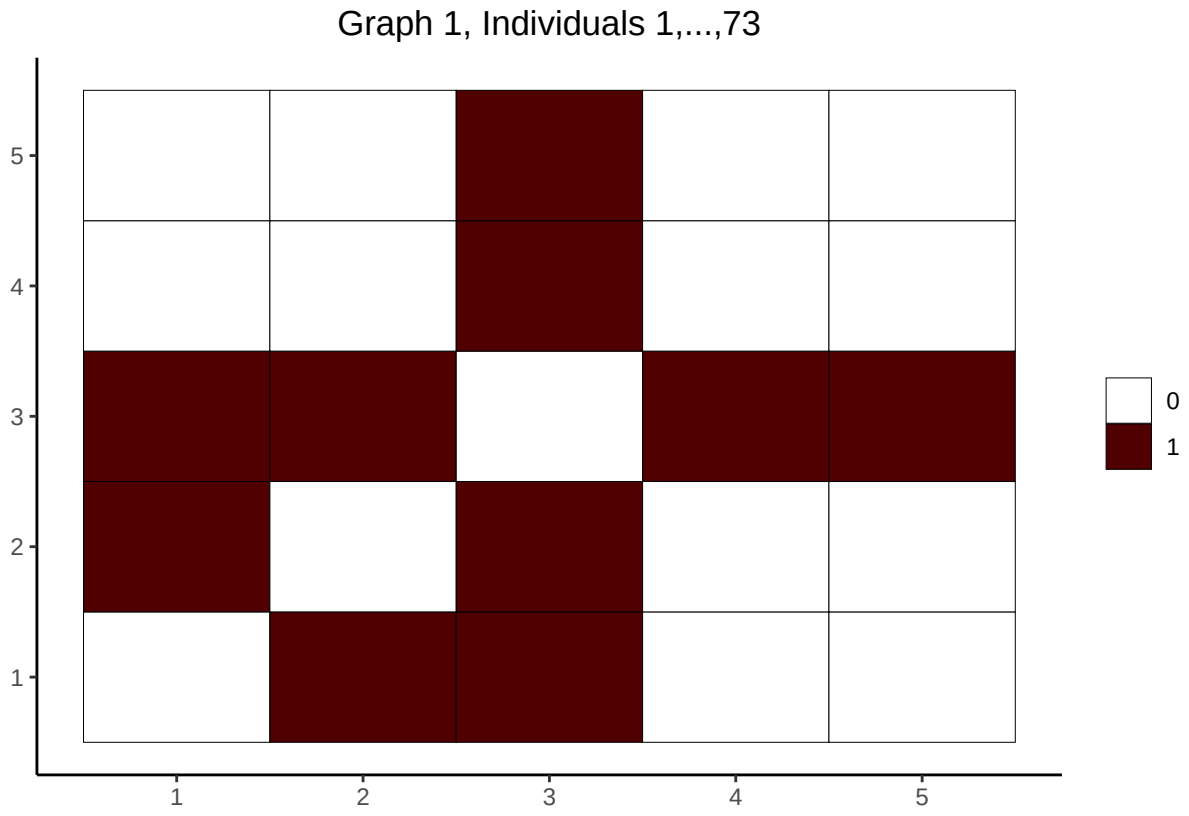



```
##
##
##
## TRIAL  2
##
## grid
##                        Covariate Dependent Graphical Model
##
## Model ELBO: -9128.94                Unique conditional dependence structures: 1
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 50.893 secs
## [[1]]
```

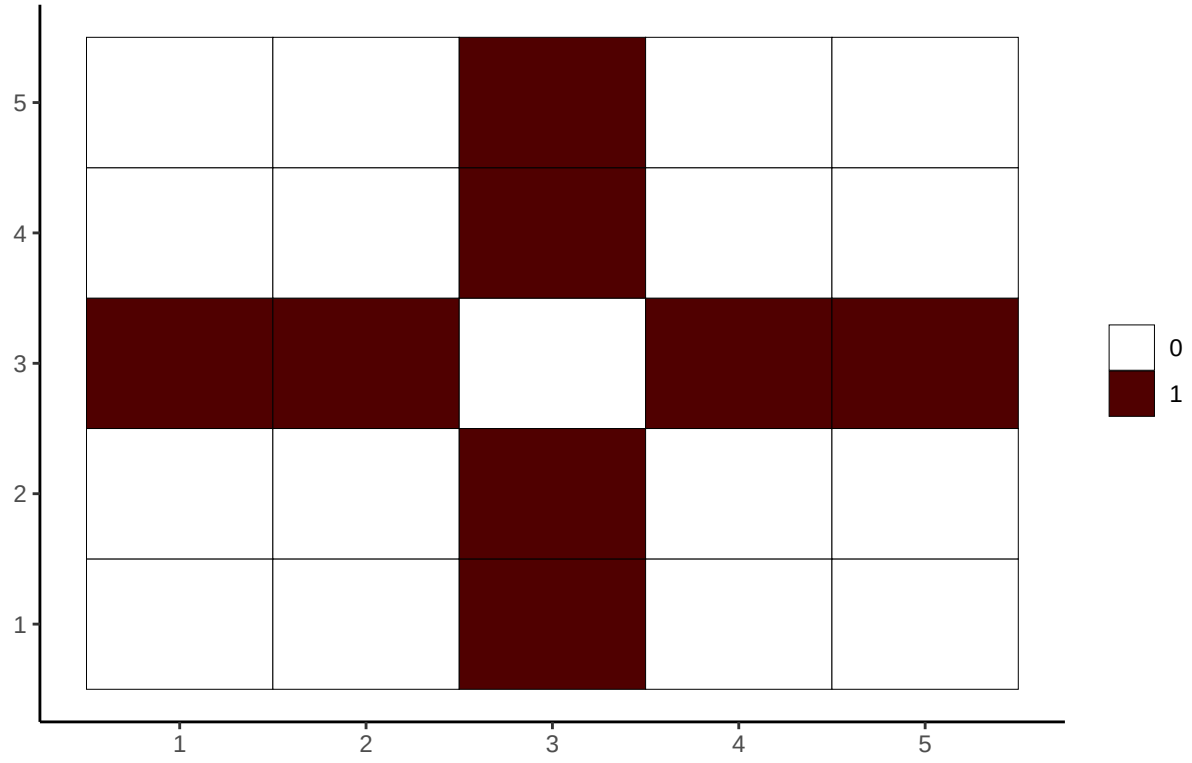
Graph 1, Individuals 1,...,180



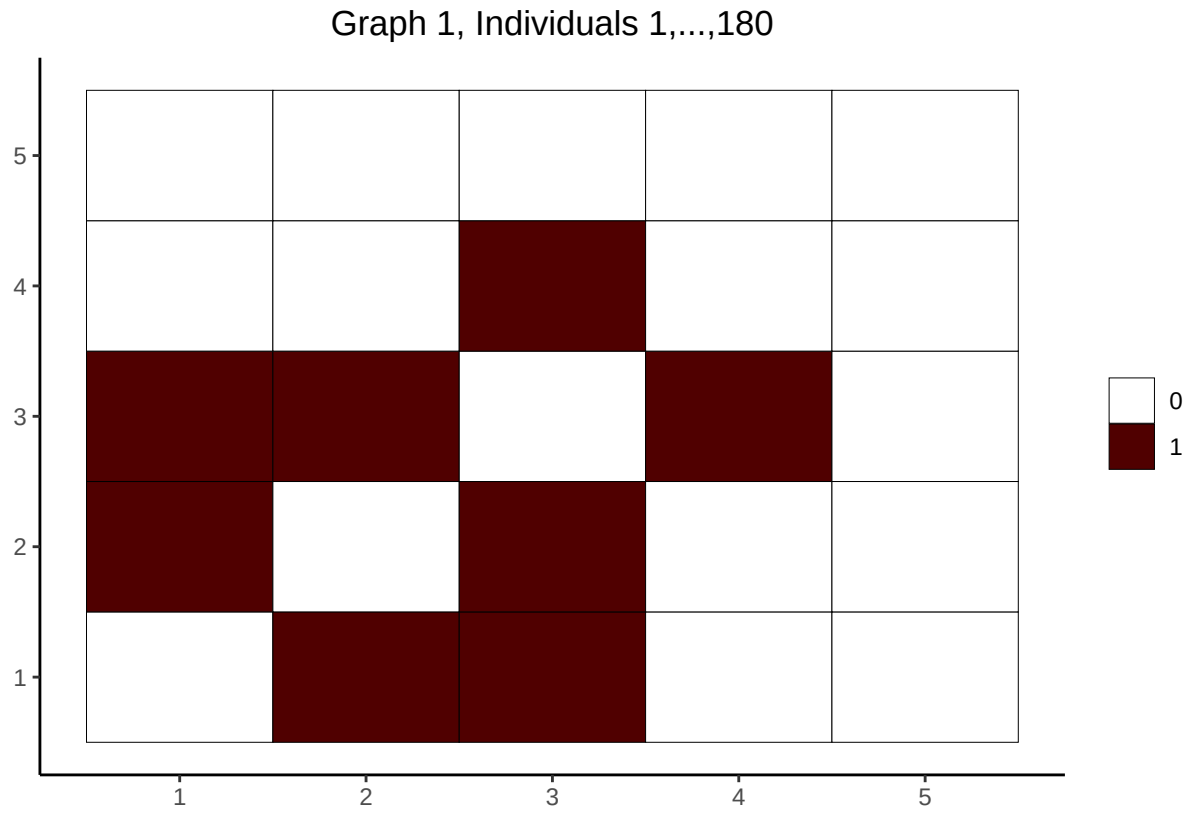

```
##
## - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
## TRIAL  3
##
## prior
##                       Covariate Dependent Graphical Model
##
## Model ELBO: -9243.52                Unique conditional dependence structures: 2
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 2.153 mins
## [[1]]
```




```
##
## [[2]]
```

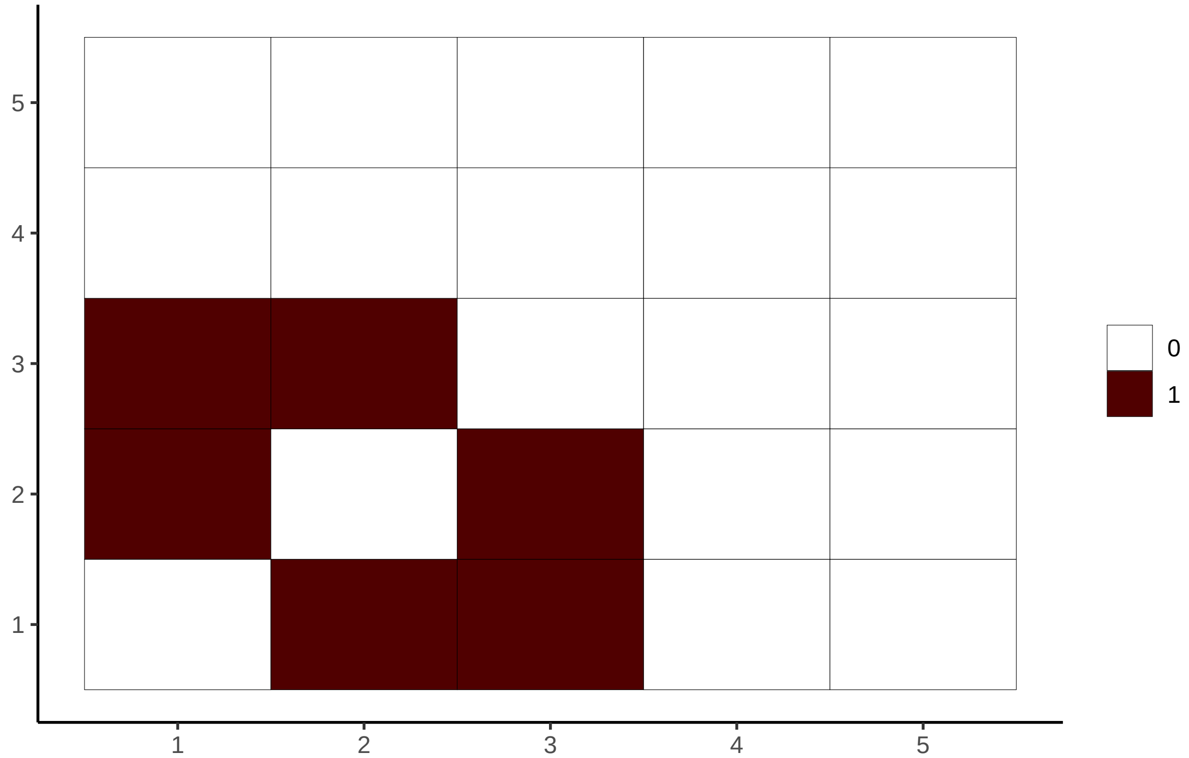



```
##
##
##
## TRIAL  3
##
## unif_prior
##                        Covariate Dependent Graphical Model
##
## Model ELBO: -9242.54                Unique conditional dependence structures: 1
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 2.097 mins
## [[1]]
```

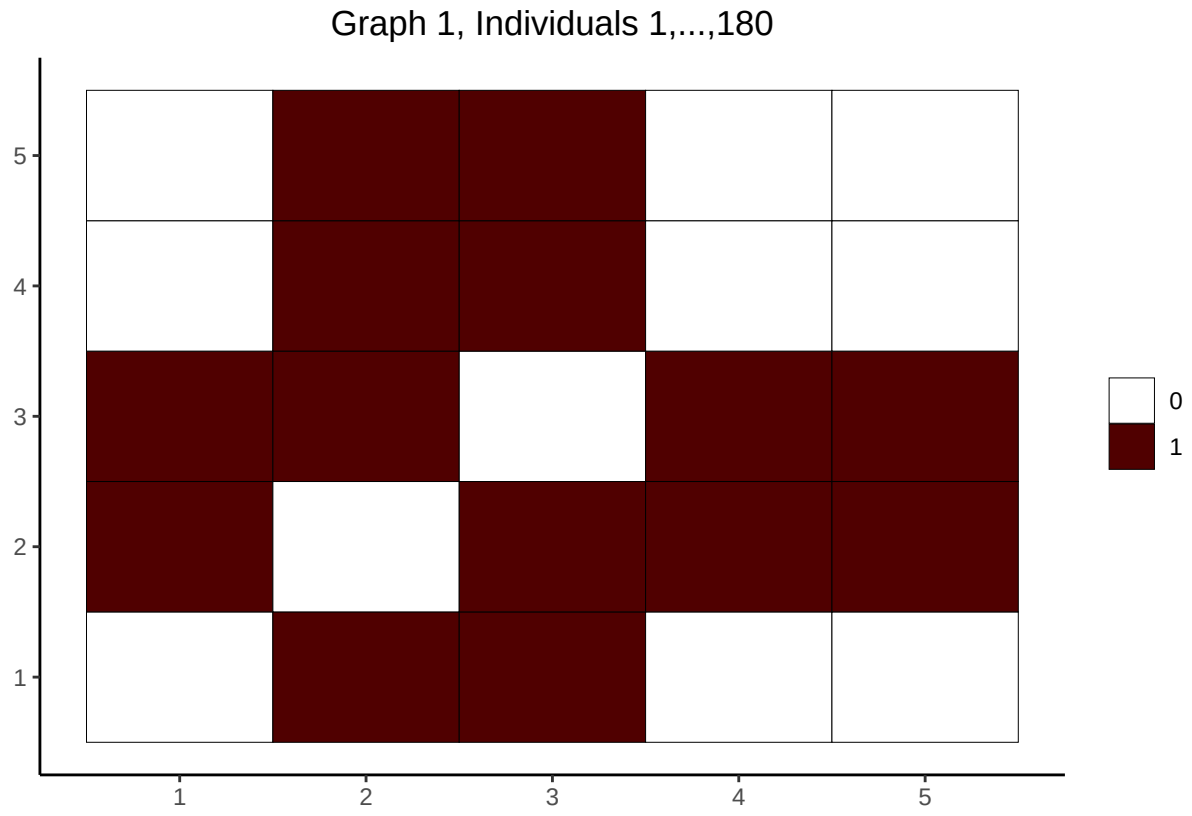



```
##
##
##
## TRIAL  3
##
## grid
##                        Covariate Dependent Graphical Model
##
## Model ELBO: -9241.55                Unique conditional dependence structures: 1
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 49.44 secs
## [[1]]
```

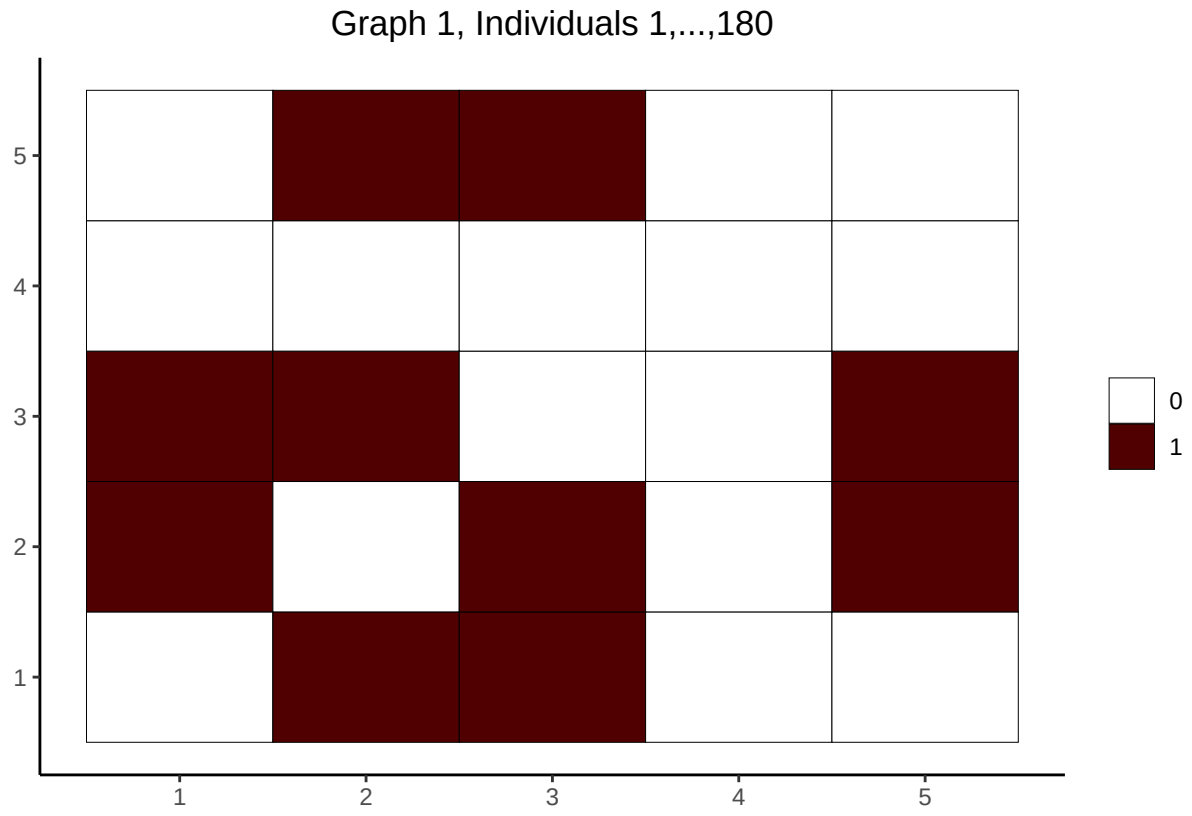
Graph 1, Individuals 1,...,180



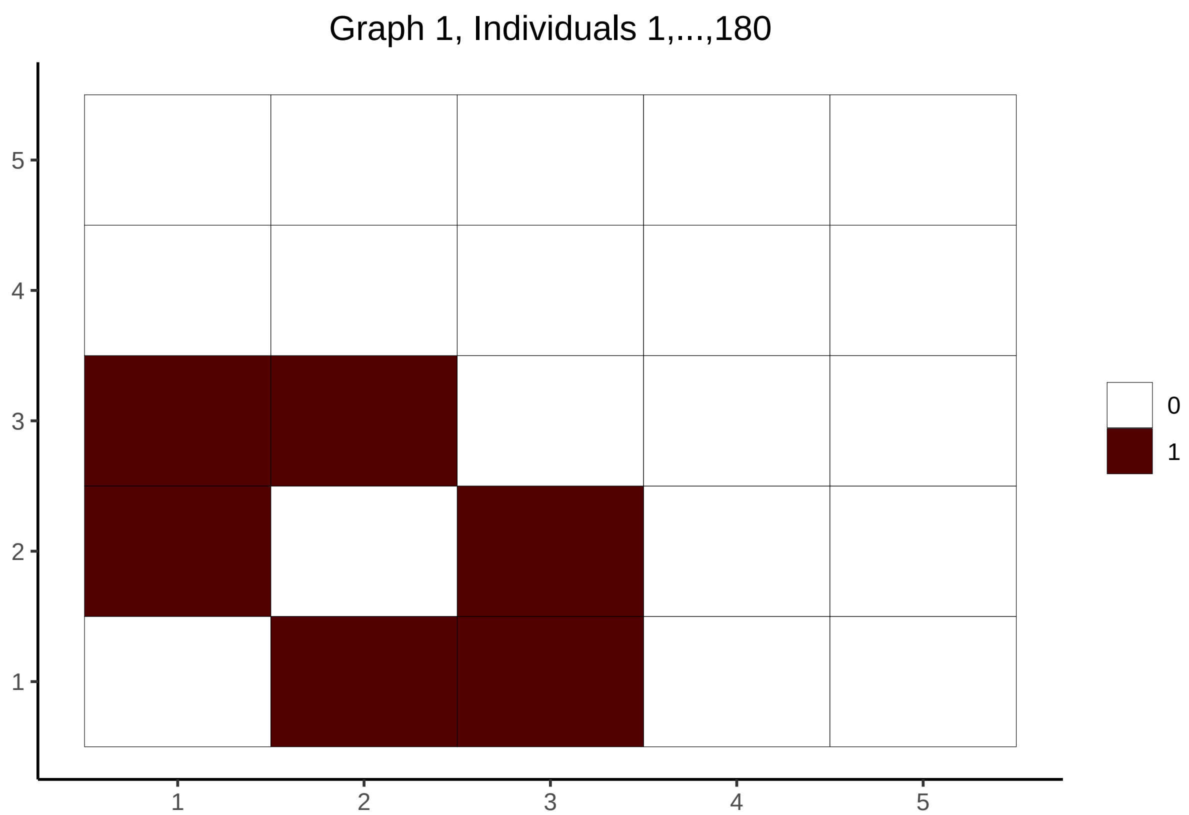

```
##
## - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
## TRIAL  4
##
## prior
##                       Covariate Dependent Graphical Model
##
## Model ELBO: -9549.51                Unique conditional dependence structures: 1
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 1.809 mins
## [[1]]
```



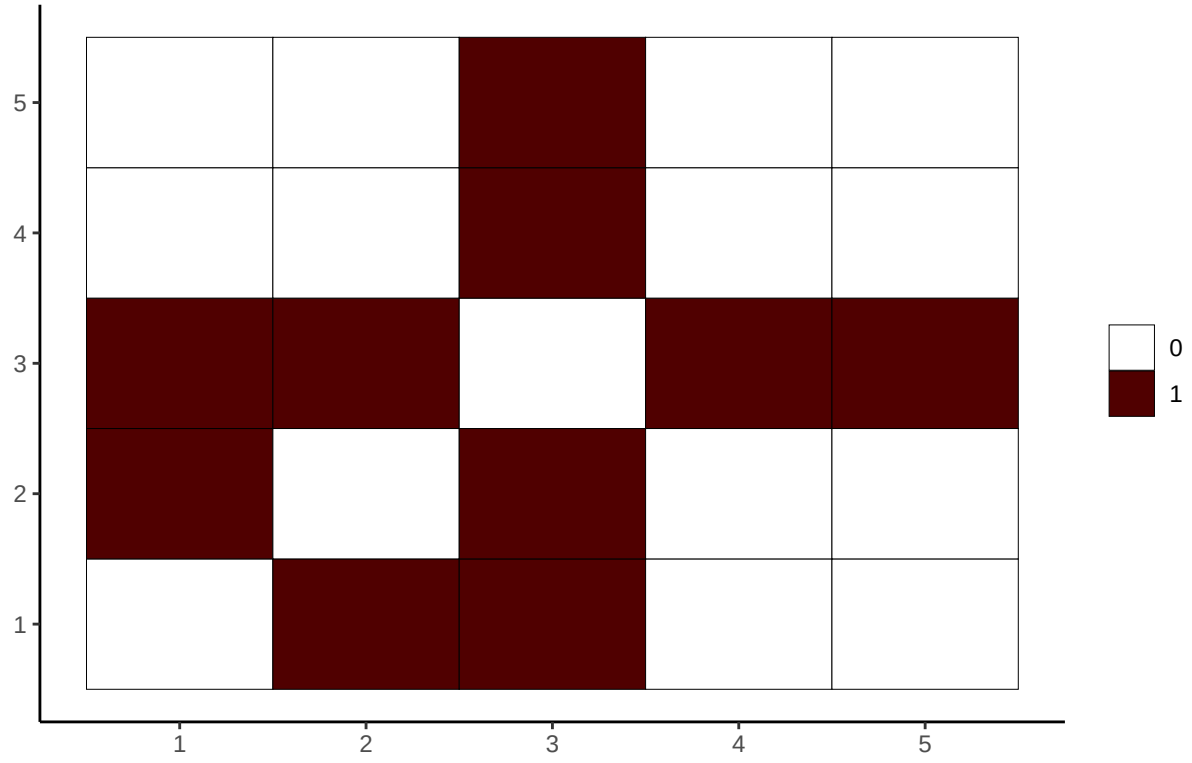

```
##
##
##
## TRIAL  4
##
## unif_prior
##                       Covariate Dependent Graphical Model
##
## Model ELBO: -9549.51                Unique conditional dependence structures: 1
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 1.755 mins
## [[1]]
```



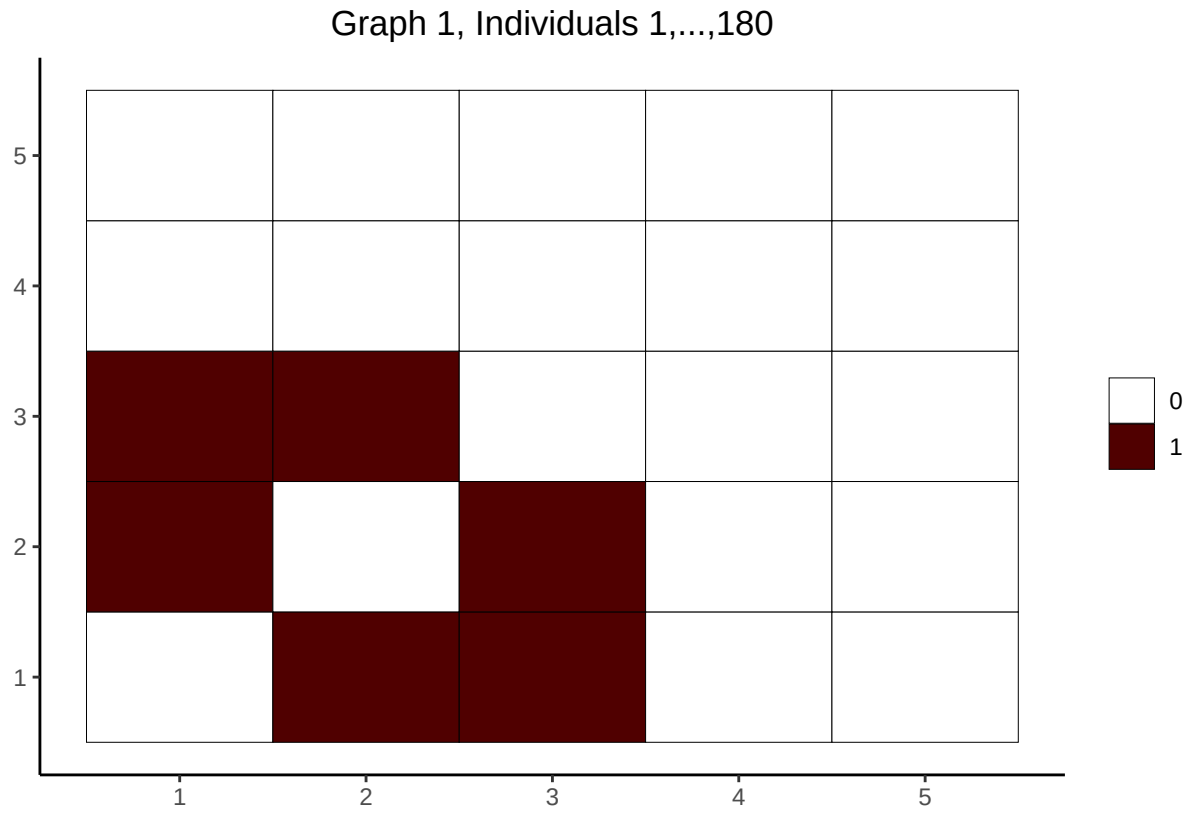

```
##
##
##
## TRIAL  4
##
## grid
##                        Covariate Dependent Graphical Model
##
## Model ELBO: -9548.53                Unique conditional dependence structures: 1
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 52.036 secs
## [[1]]
```



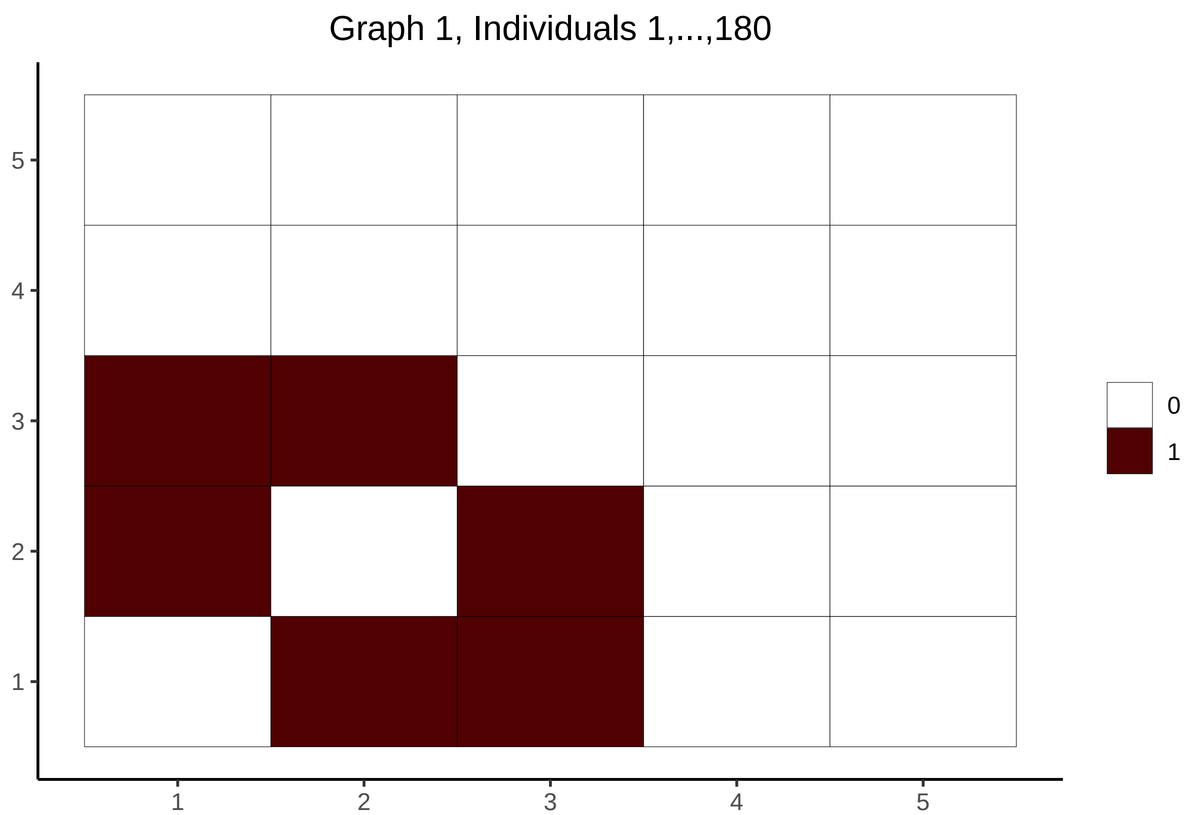

```
##
## - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
## TRIAL  5
##
## prior
##                       Covariate Dependent Graphical Model
##
## Model ELBO: -9290.02                Unique conditional dependence structures: 1
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 2.151 mins
## [[1]]
```



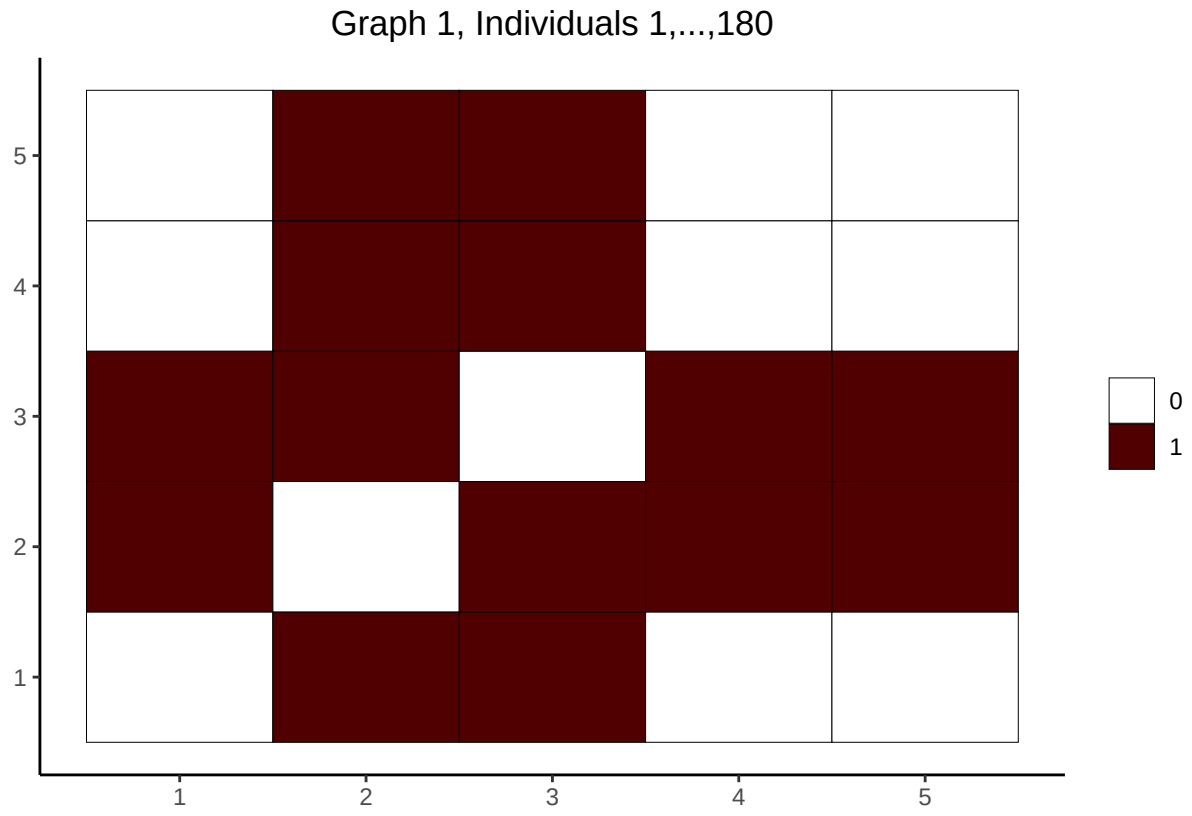

```
##
##
##
## TRIAL  5
##
## unif_prior
##                        Covariate Dependent Graphical Model
##
## Model ELBO: -9288.85                Unique conditional dependence structures: 1
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 2.105 mins
## [[1]]
```



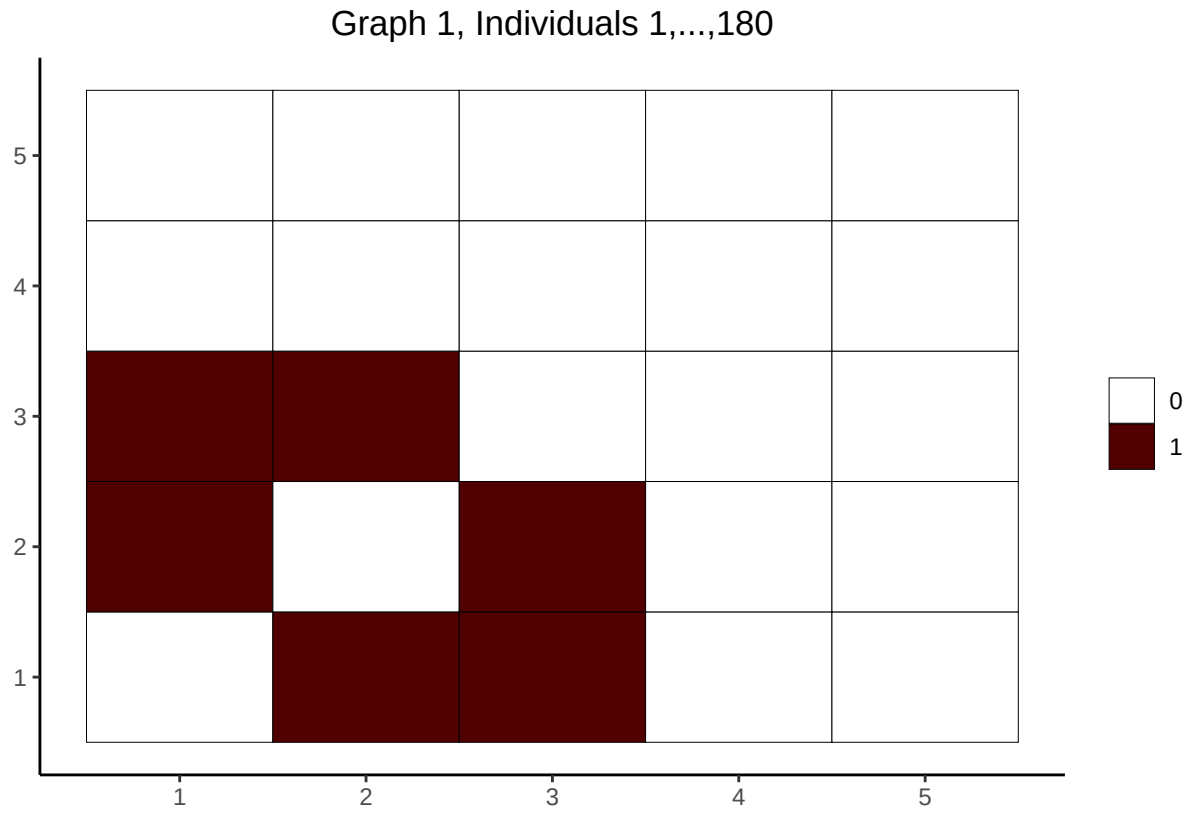

```
##
##
##
## TRIAL  5
##
## grid
##                        Covariate Dependent Graphical Model
##
## Model ELBO: -9288.12                Unique conditional dependence structures: 1
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 50.482 secs
## [[1]]
```



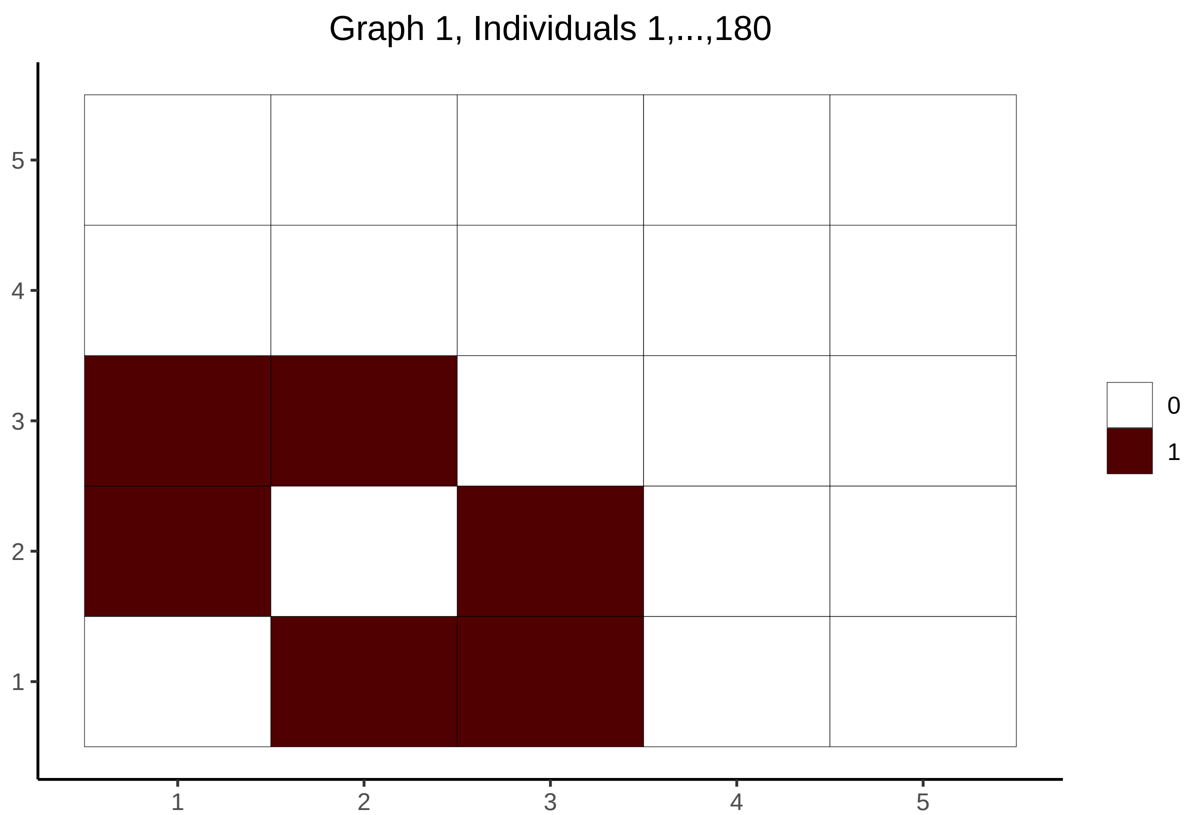

```
##
## - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
## TRIAL  6
##
## prior
##                      Covariate Dependent Graphical Model
##
## Model ELBO: -9571.61                Unique conditional dependence structures: 1
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 2.093 mins
## [[1]]
```



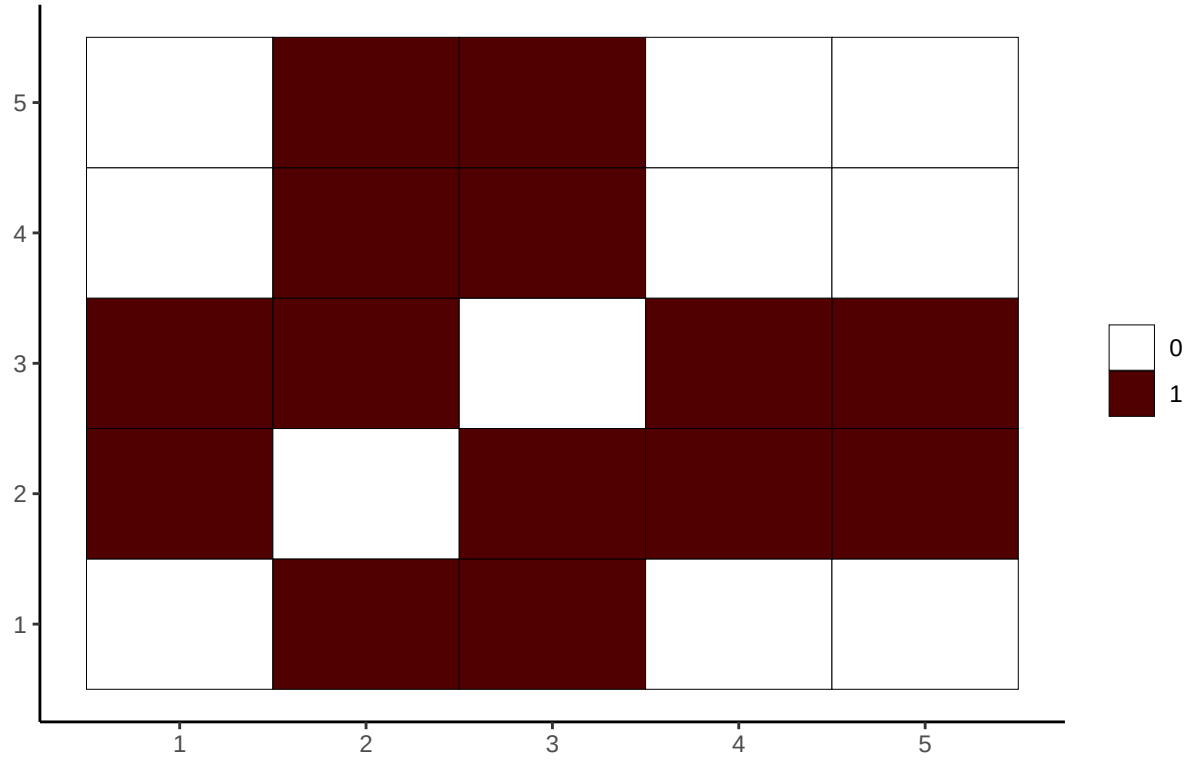

```
##
##
##
## TRIAL  6
##
## unif_prior
##                       Covariate Dependent Graphical Model
##
## Model ELBO: -9568.74                Unique conditional dependence structures: 1
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 2.025 mins
## [[1]]
```



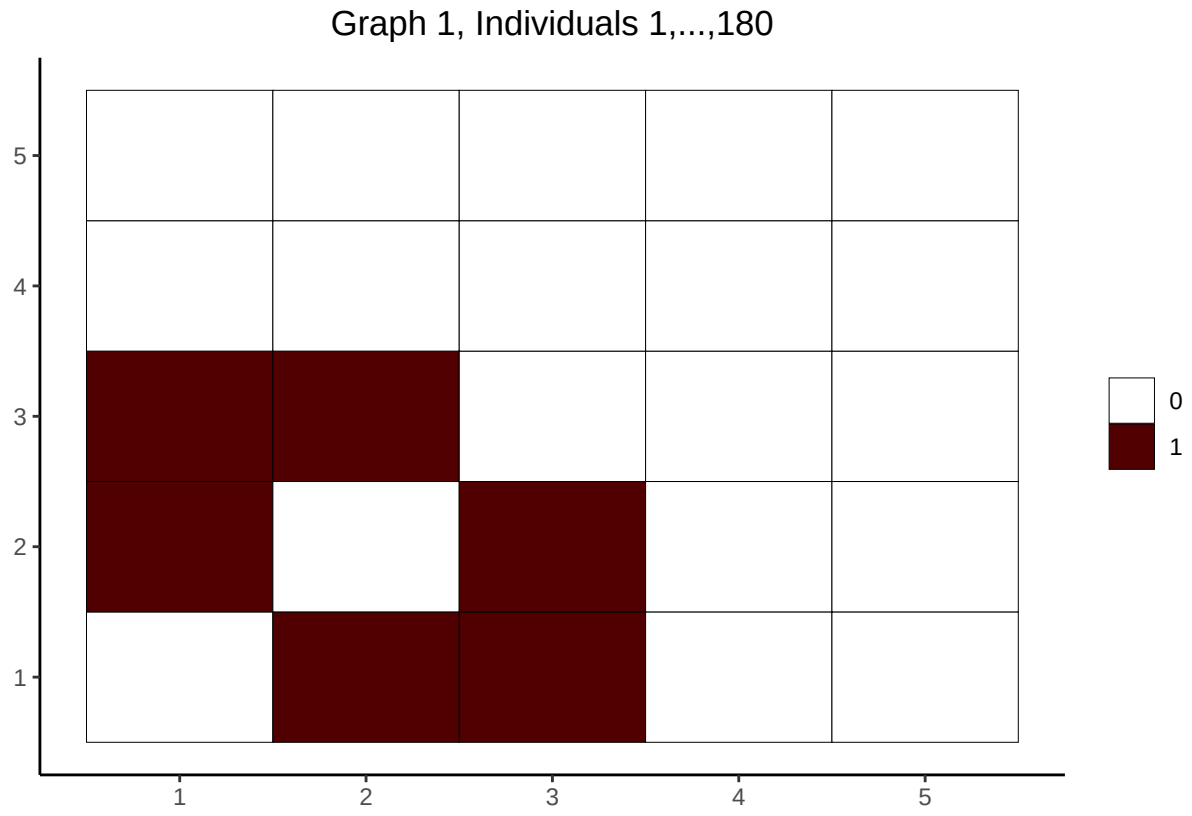

```
##
##
##
## TRIAL  6
##
## grid
##                        Covariate Dependent Graphical Model
##
## Model ELBO: -9568.07                Unique conditional dependence structures: 1
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 49.113 secs
## [[1]]
```



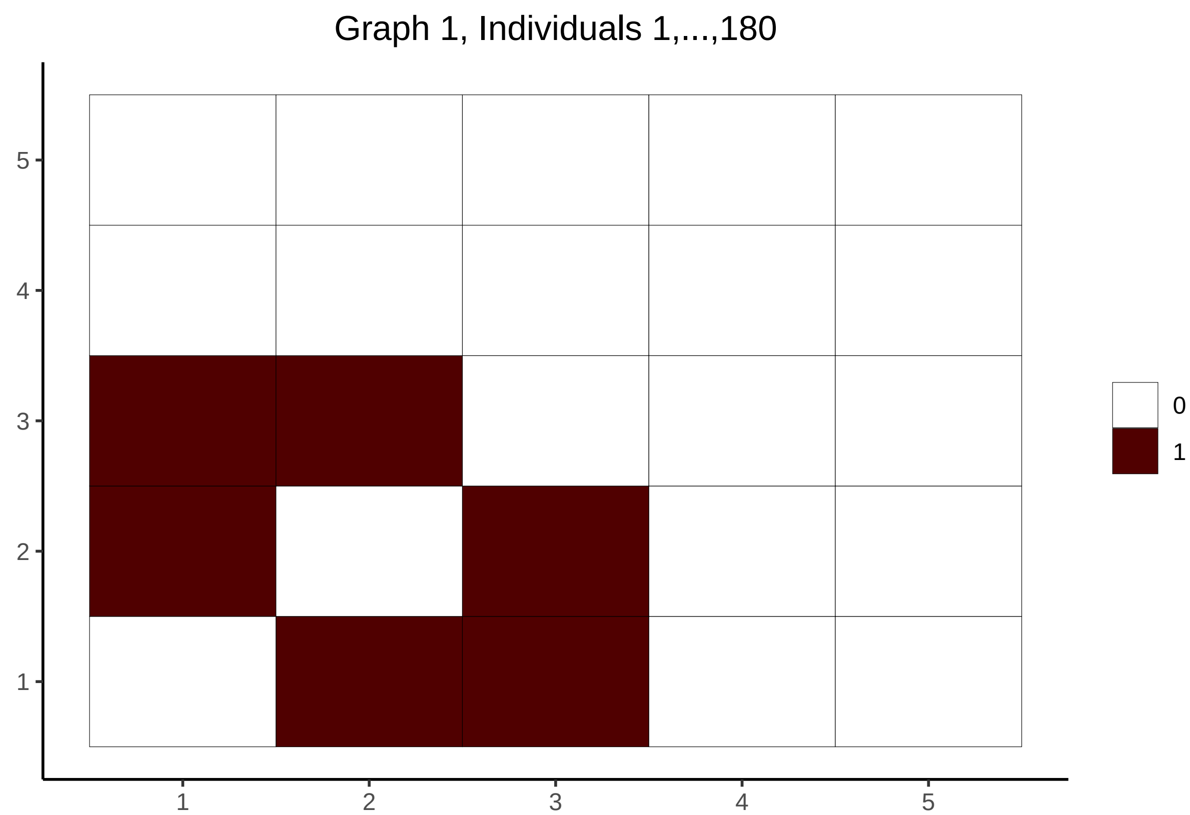

```
##
## - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
## TRIAL  7
##
## prior
##                       Covariate Dependent Graphical Model
##
## Model ELBO: -9634.16                Unique conditional dependence structures: 1
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 1.976 mins
## [[1]]
```



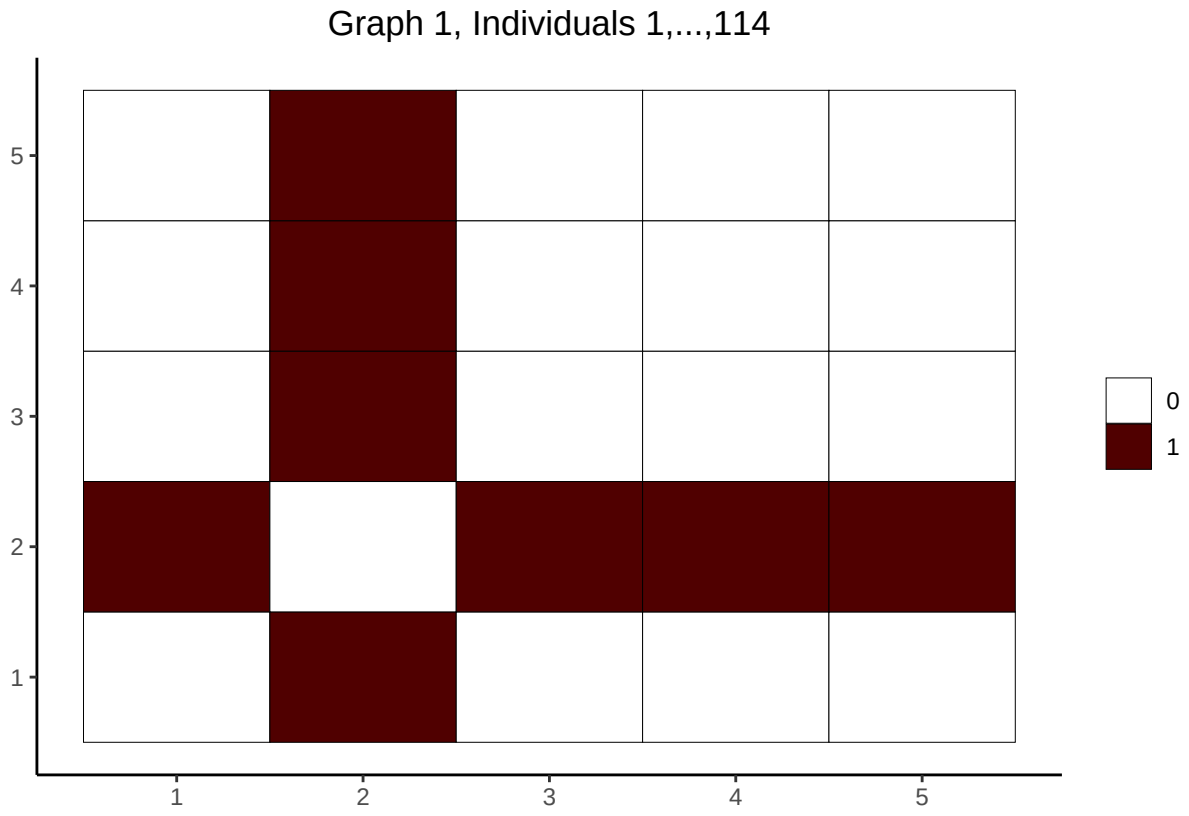

```
##
##
##
## TRIAL  7
##
## unif_prior
##                        Covariate Dependent Graphical Model
##
## Model ELBO: -9632.38                Unique conditional dependence structures: 1
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 1.903 mins
## [[1]]
```



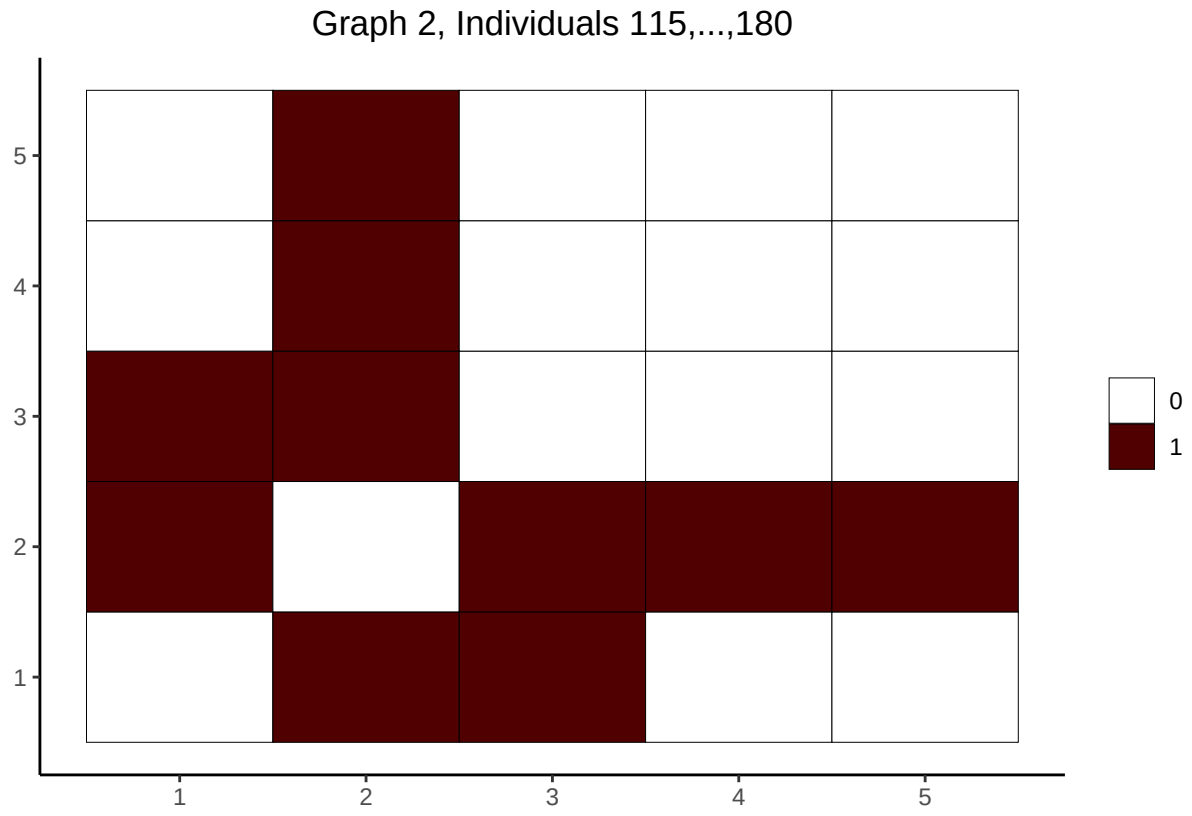

```
##
##
##
## TRIAL  7
##
## grid
##                       Covariate Dependent Graphical Model
##
## Model ELBO: -9631.28                Unique conditional dependence structures: 1
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 44.307 secs
## [[1]]
```



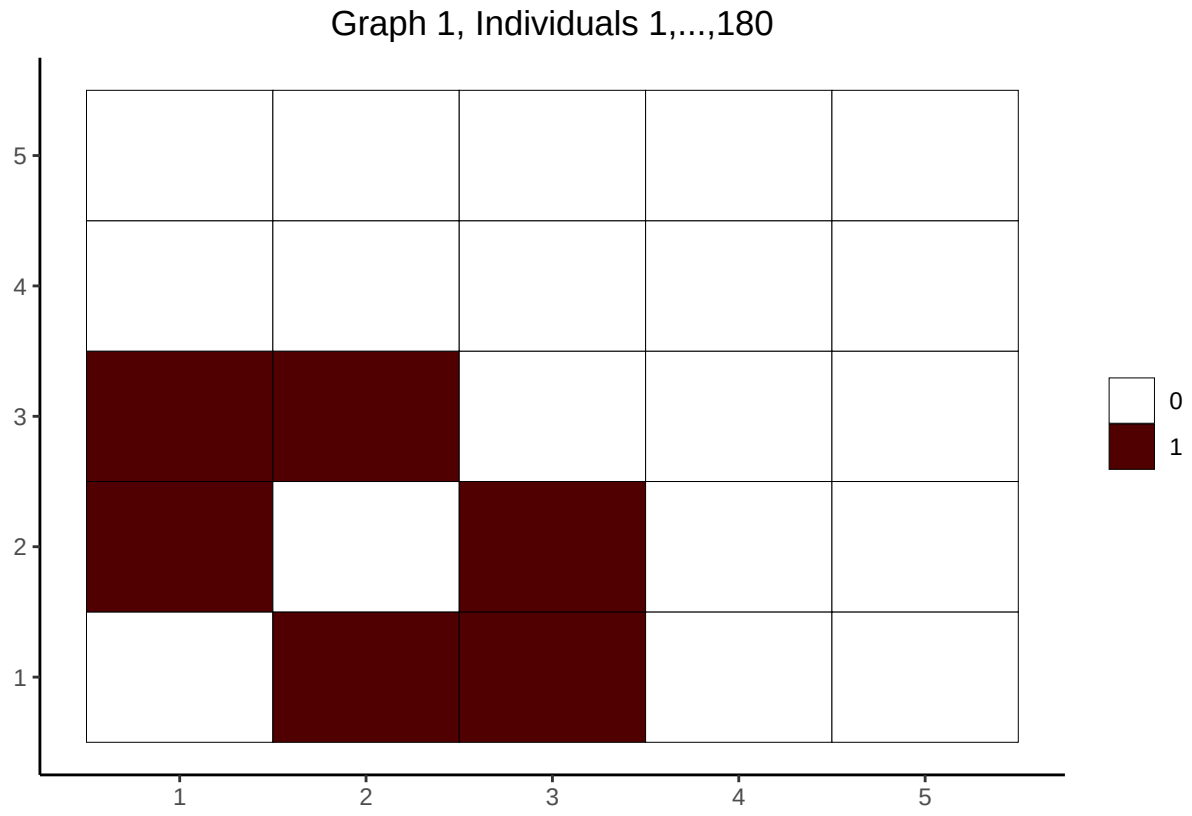

```
##
## - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
## TRIAL  8
##
## prior
##                       Covariate Dependent Graphical Model
##
## Model ELBO: -9300                   Unique conditional dependence structures: 2
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 1.977 mins
## [[1]]
```



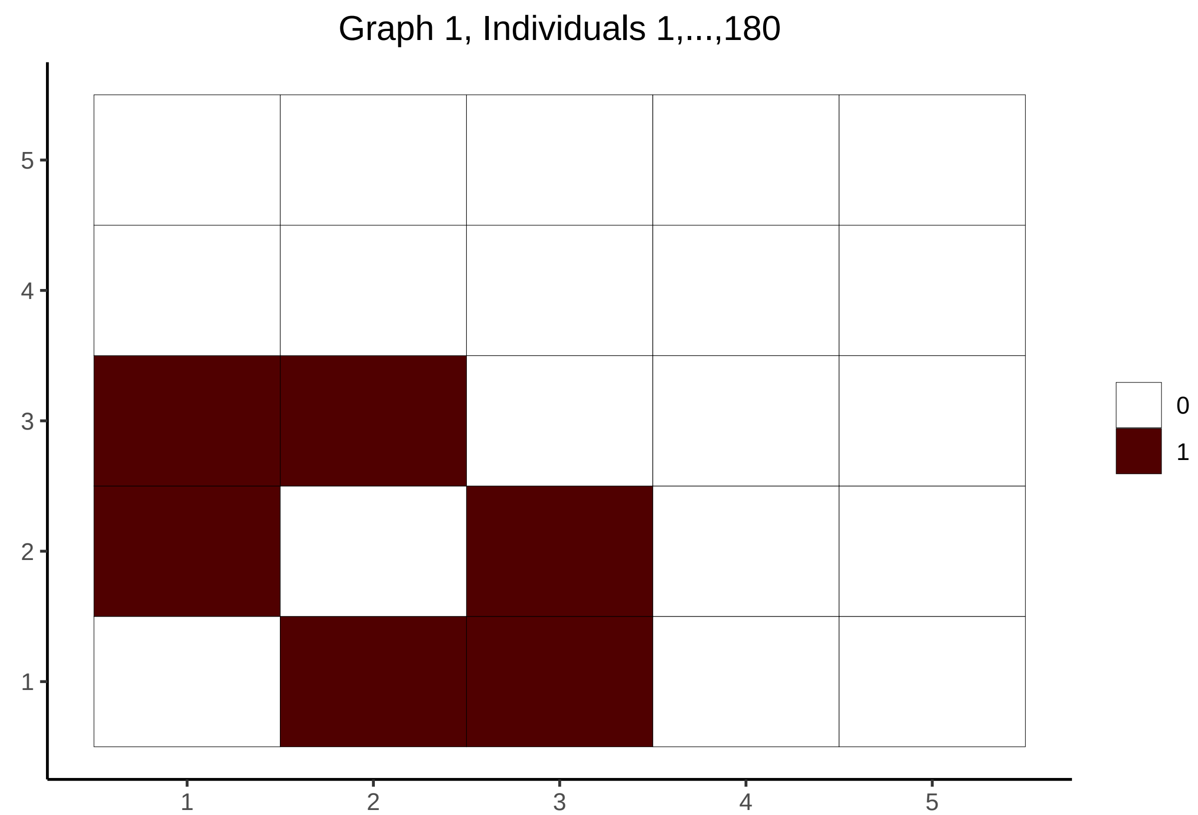

```
##
## [[2]]
```



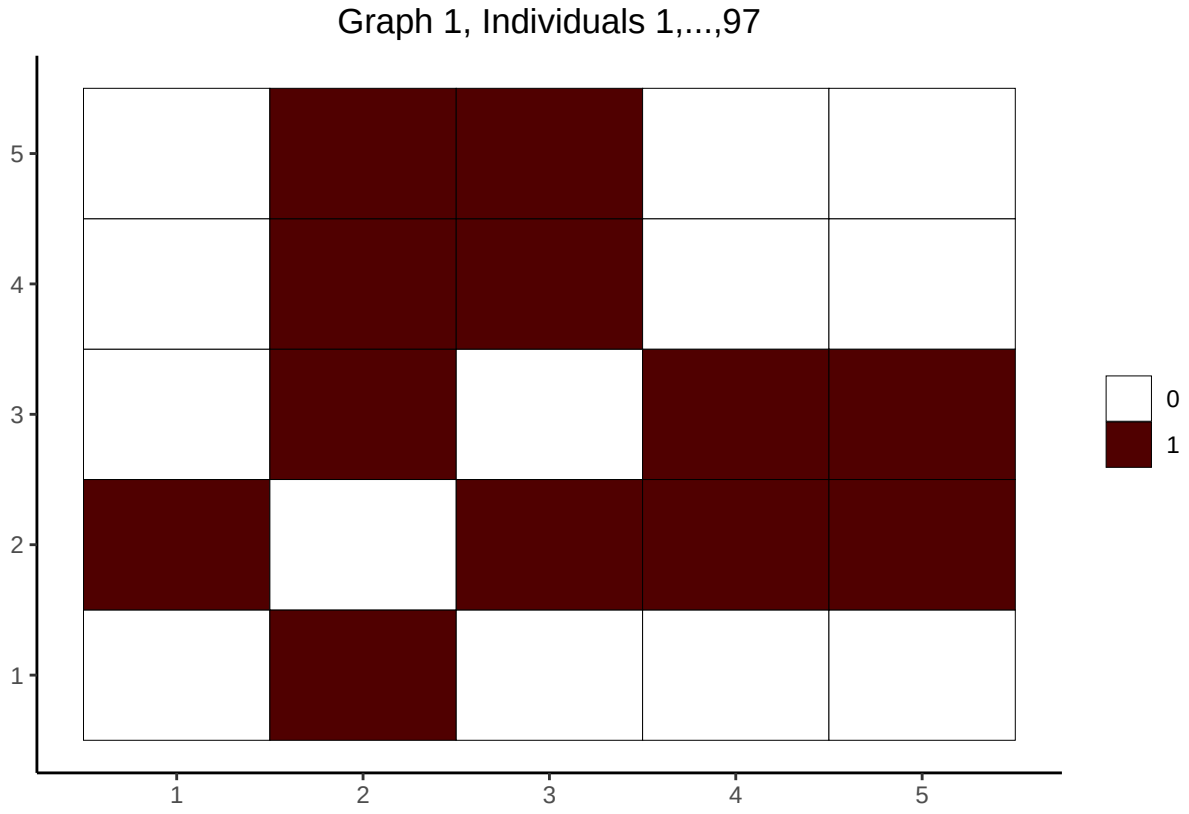

```
##
##
##
## TRIAL  8
##
## unif_prior
##                        Covariate Dependent Graphical Model
##
## Model ELBO: -9297.35                Unique conditional dependence structures: 1
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 1.971 mins
## [[1]]
```



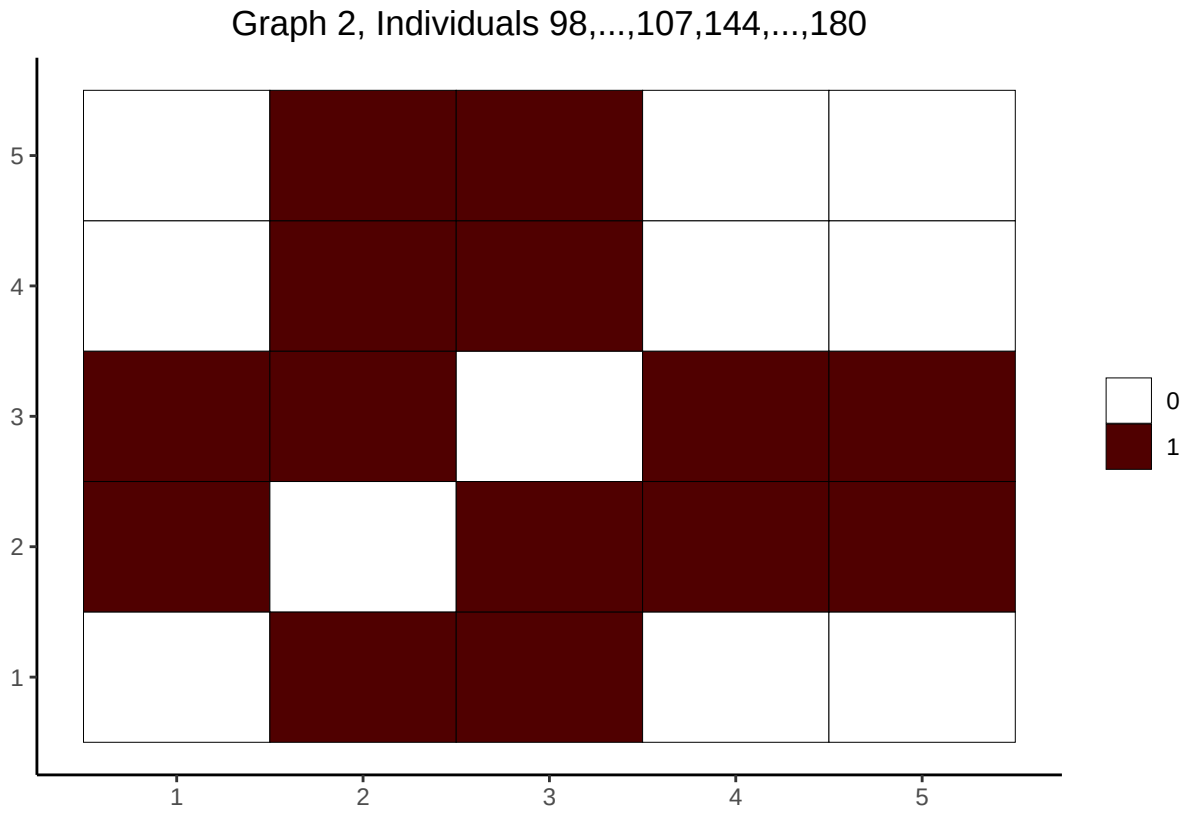

```
##
##
##
## TRIAL  8
##
## grid
##                        Covariate Dependent Graphical Model
##
## Model ELBO: -9296.41                Unique conditional dependence structures: 1
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 1.114 mins
## [[1]]
```



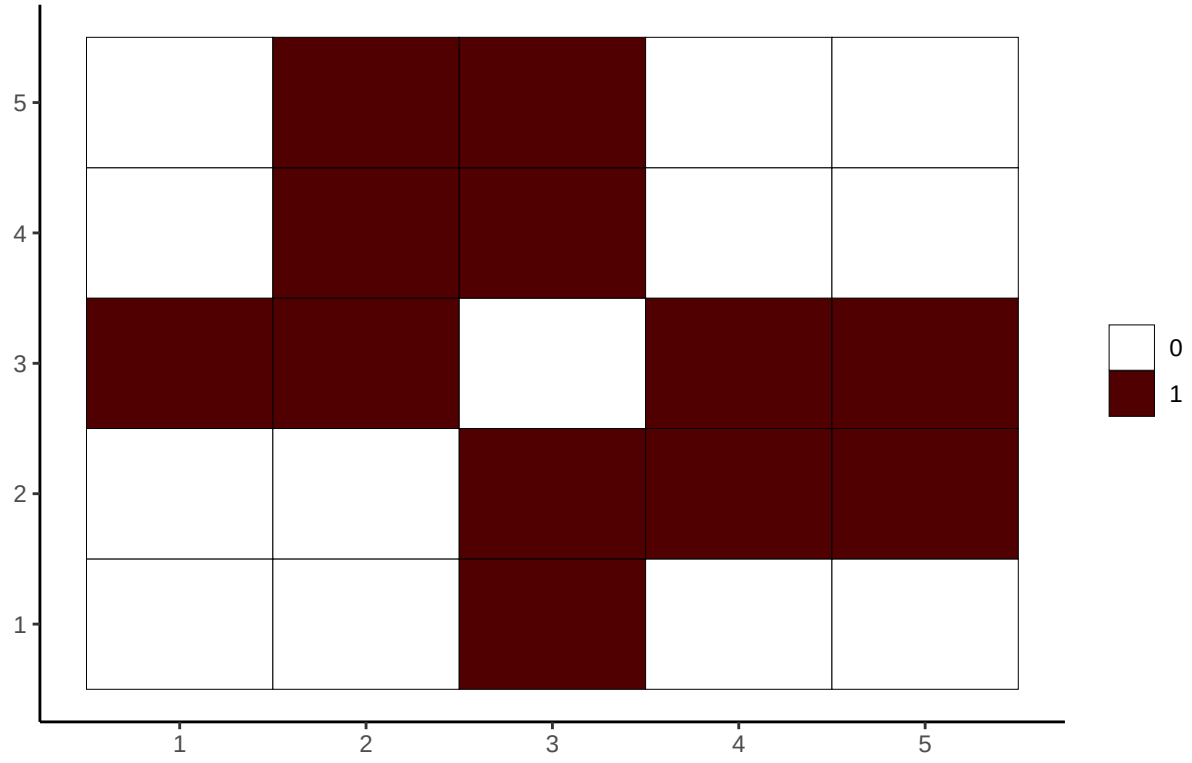

```
##
## ------------------------------------------------------------
## TRIAL  9
##
## prior
##                       Covariate Dependent Graphical Model
##
## Model ELBO: -9542.46                Unique conditional dependence structures: 3
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 1.758 mins
## [[1]]
```



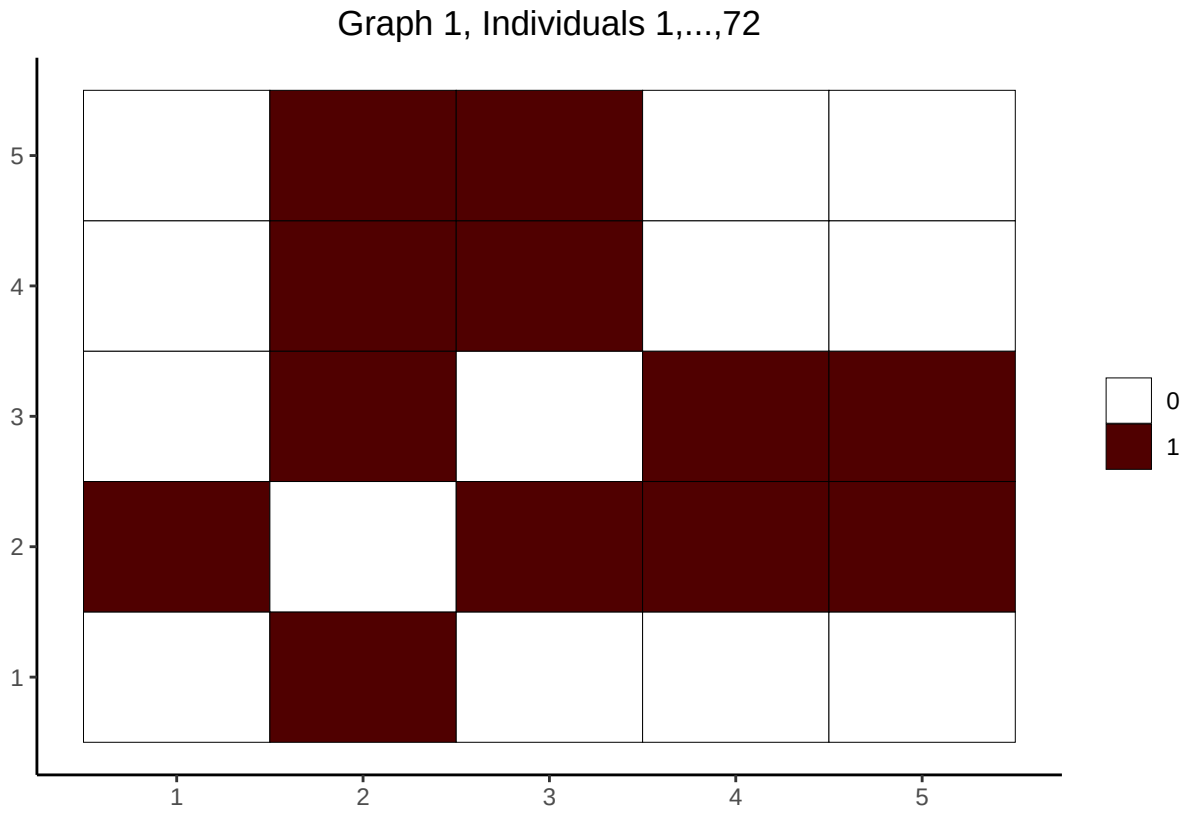

```
##
## [[2]]
```



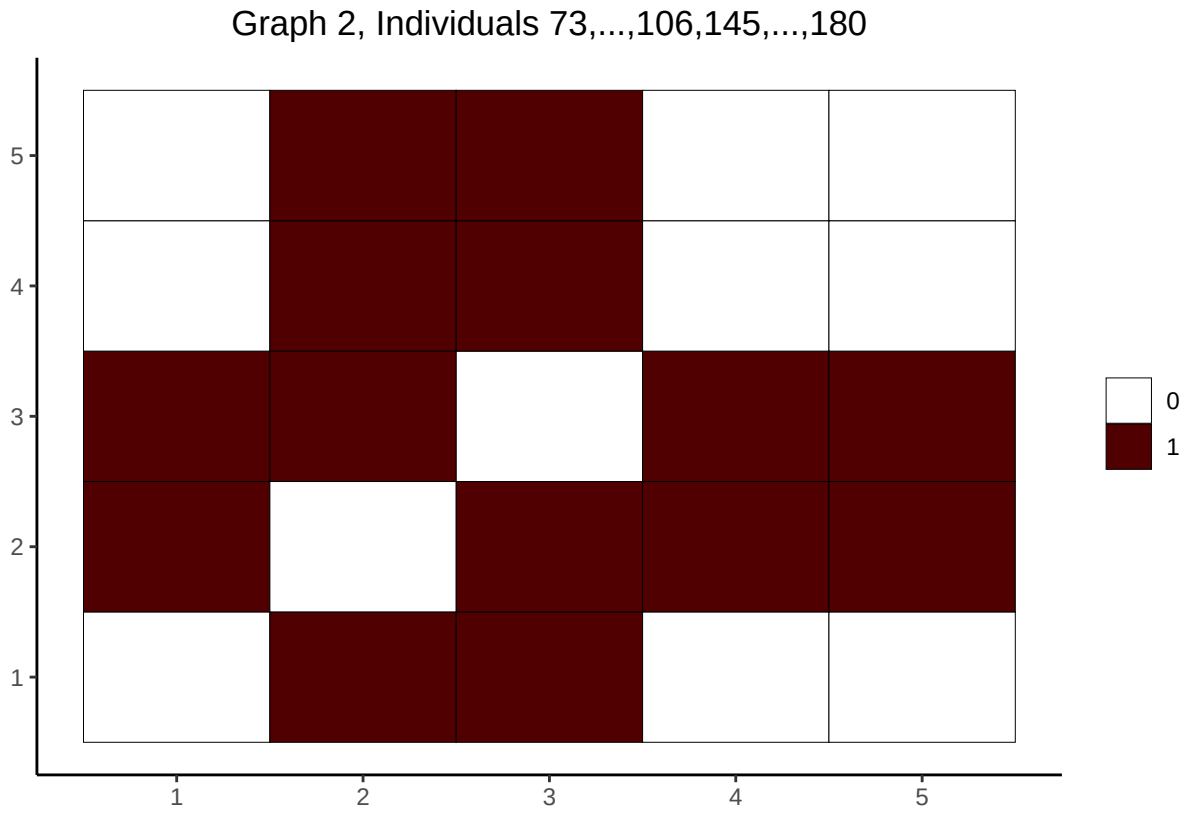

```
##
## [[3]]
```



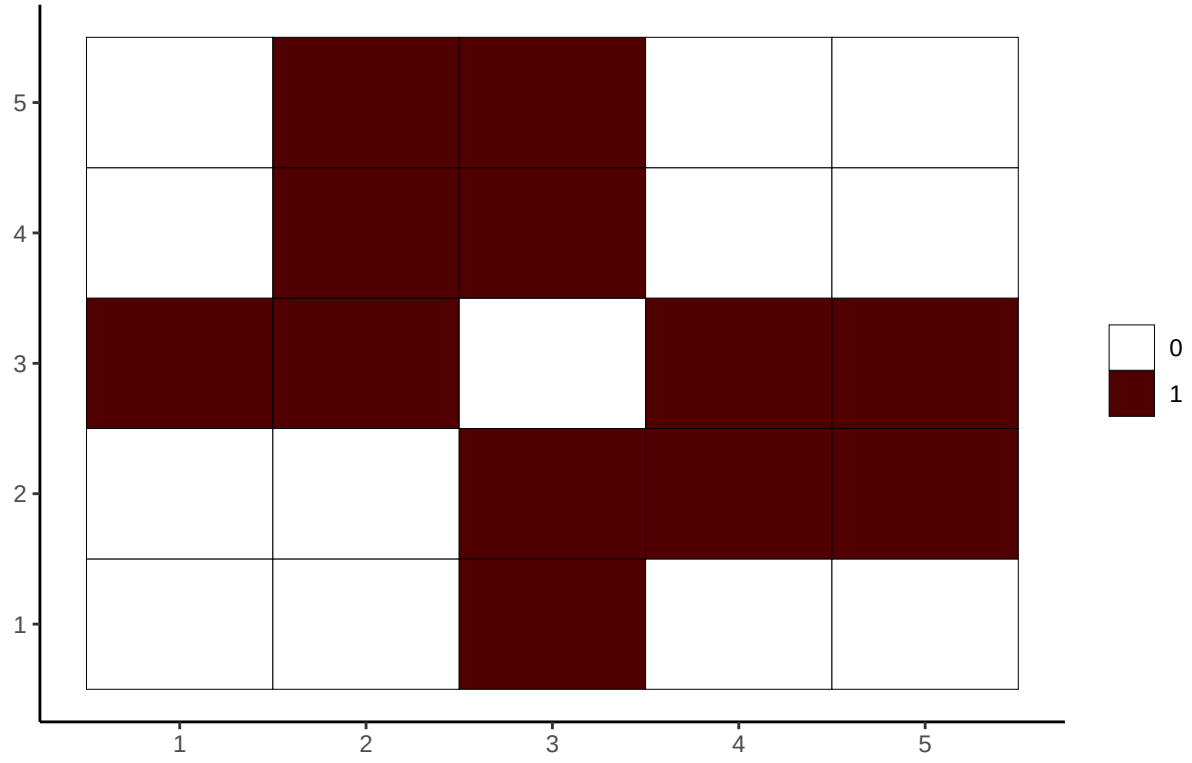

```
##
##
##
## TRIAL  9
##
## unif_prior
##                        Covariate Dependent Graphical Model
##
## Model ELBO: -9541.52                Unique conditional dependence structures: 3
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 1.758 mins
## [[1]]
```



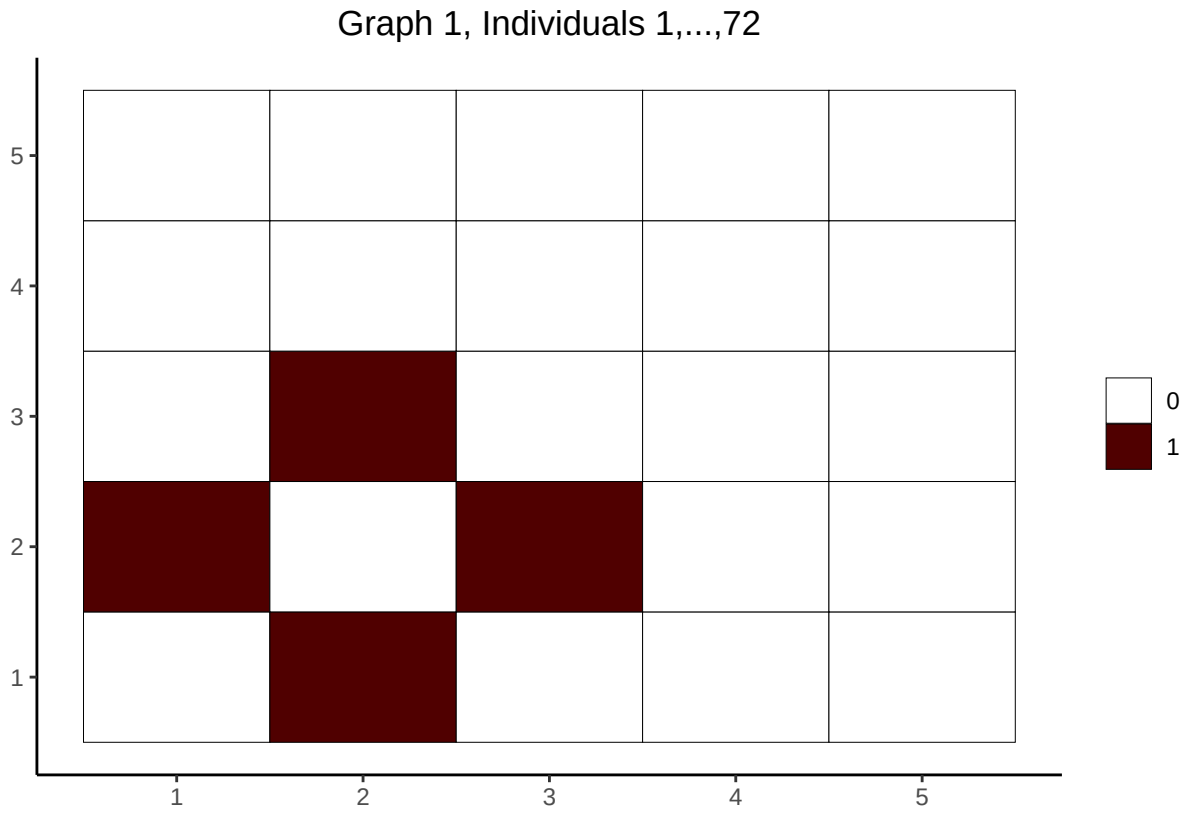

```
##
## [[2]]
```



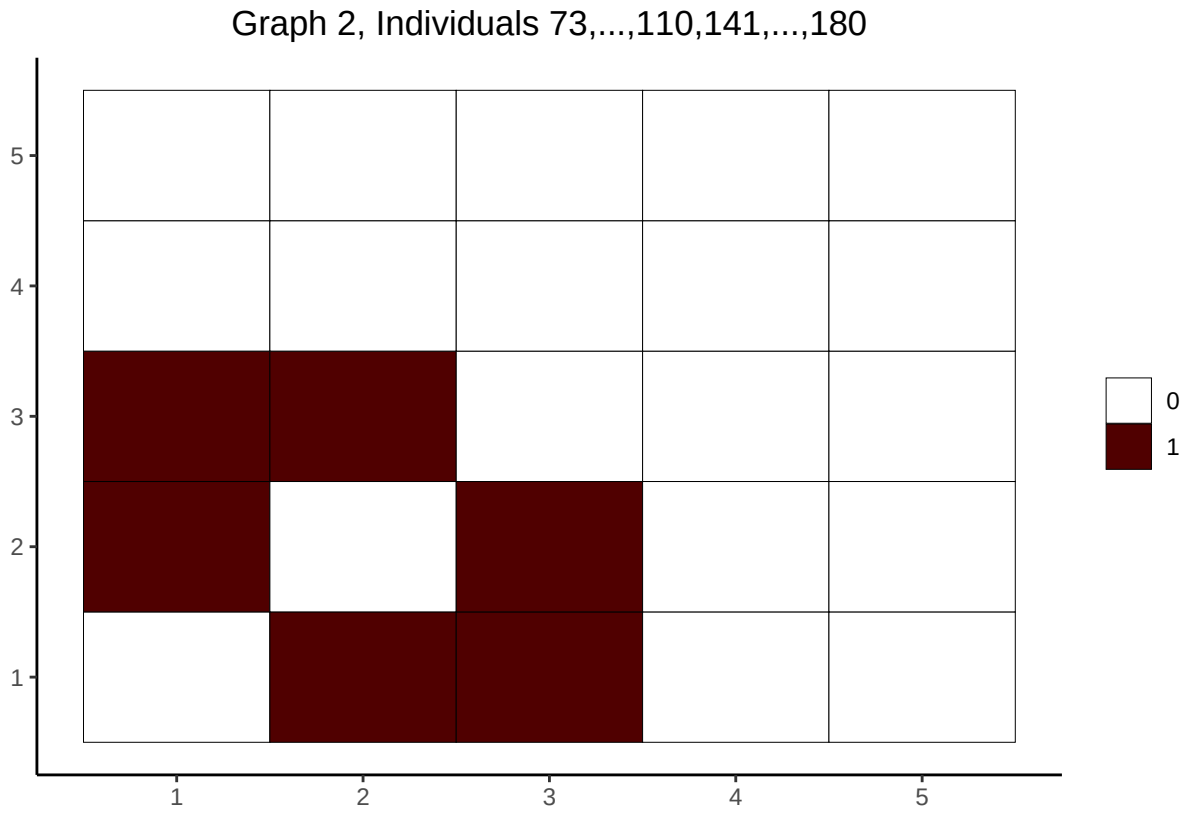

```
##
## [[3]]
```



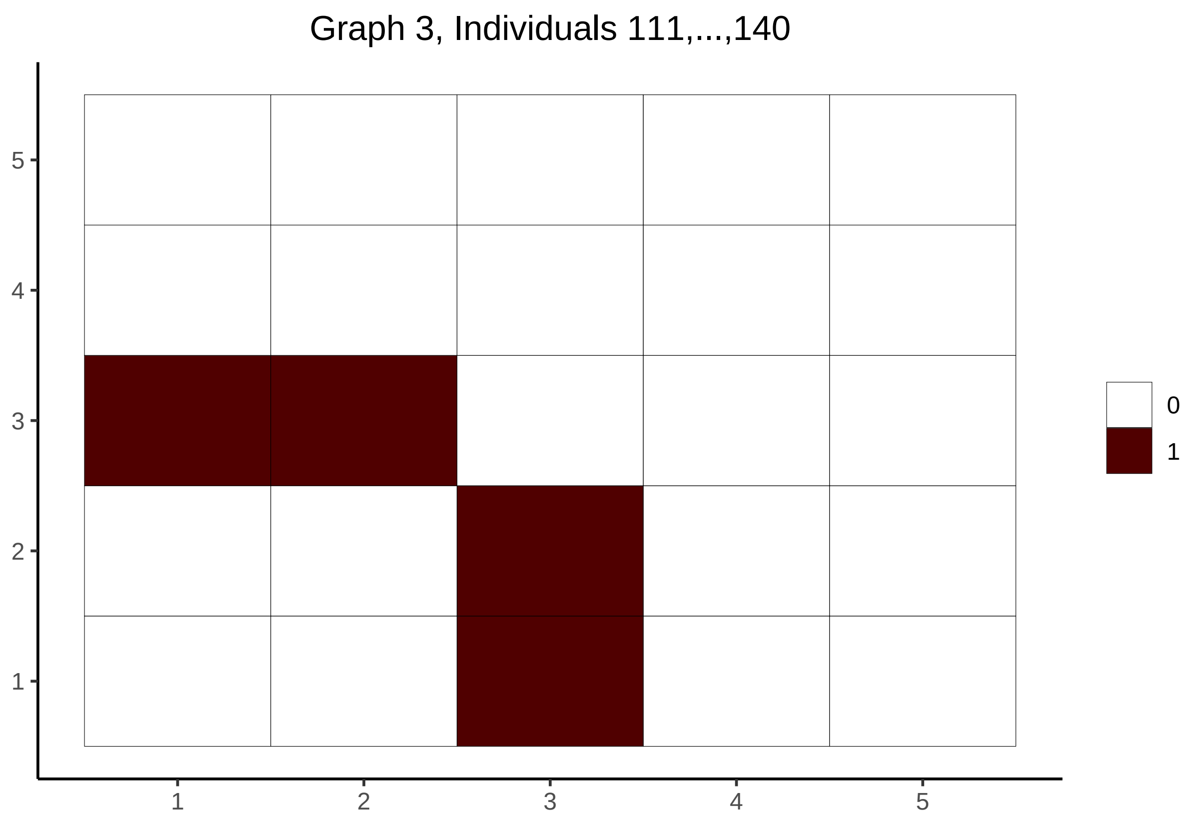

```
##
##
##
## TRIAL  9
##
## grid
##                        Covariate Dependent Graphical Model
##
## Model ELBO: -9540.94                Unique conditional dependence structures: 3
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 1.089 mins
## [[1]]
```



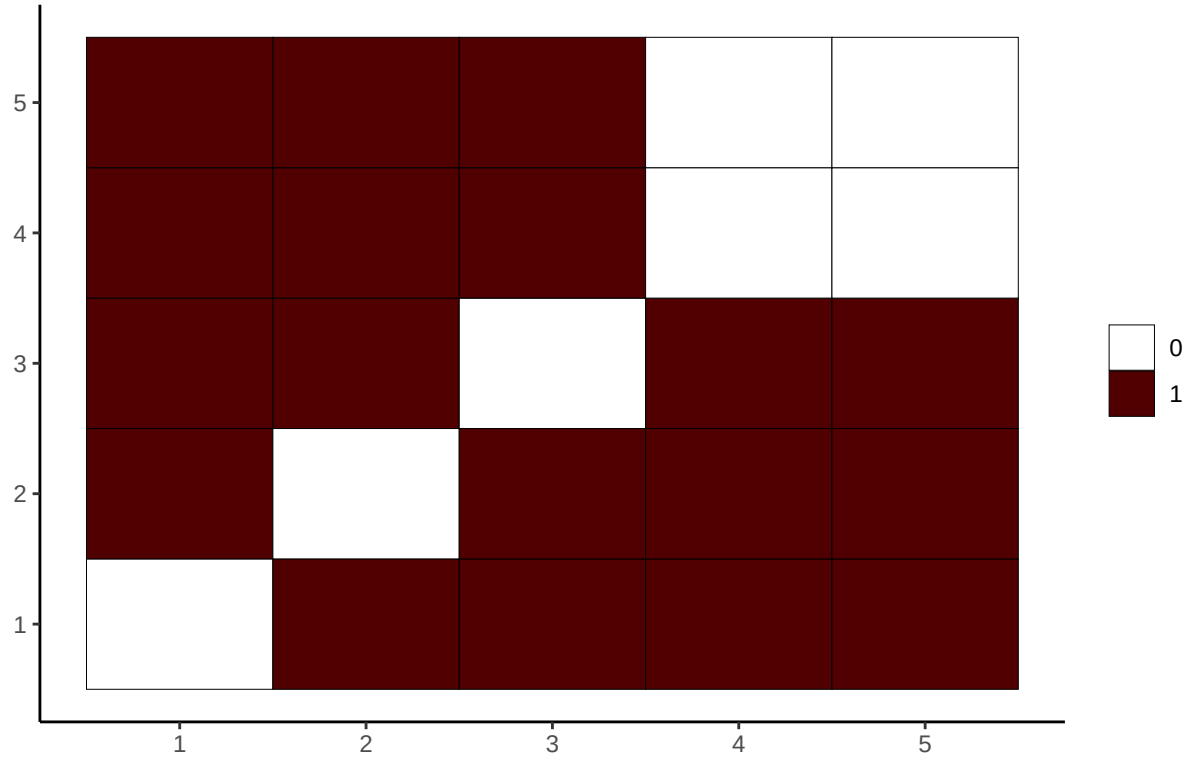

```
##
## [[2]]
```



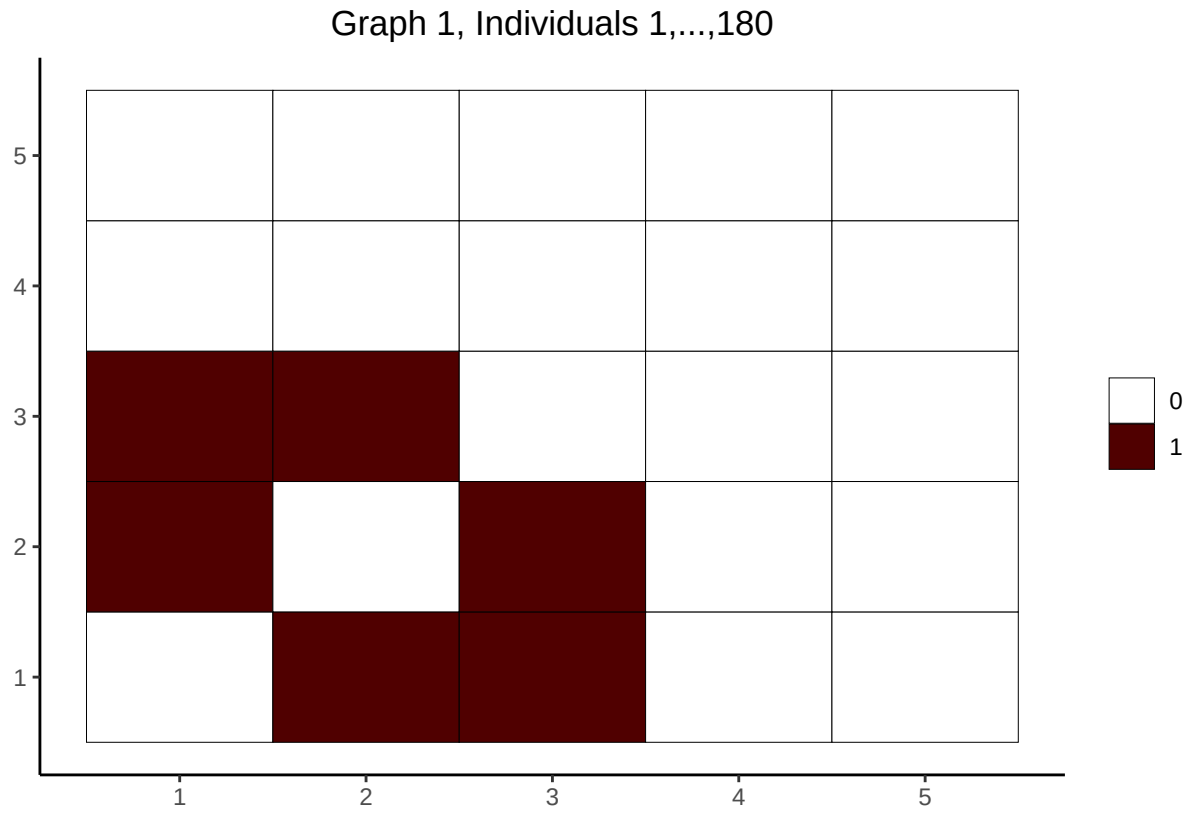

```
##
## [[3]]
```



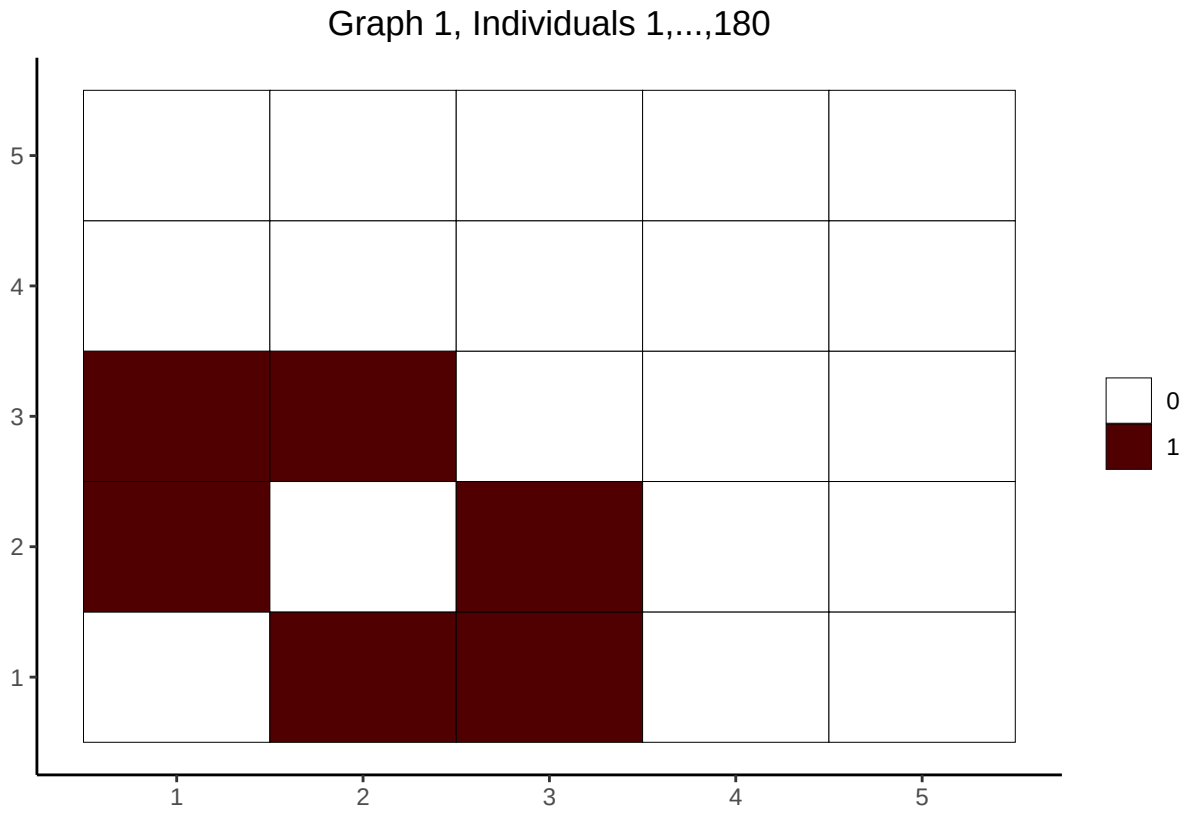

```
##
## - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
## TRIAL  10
##
## prior
##                       Covariate Dependent Graphical Model
##
## Model ELBO: -9402.02                Unique conditional dependence structures: 1
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 1.824 mins
## [[1]]
```



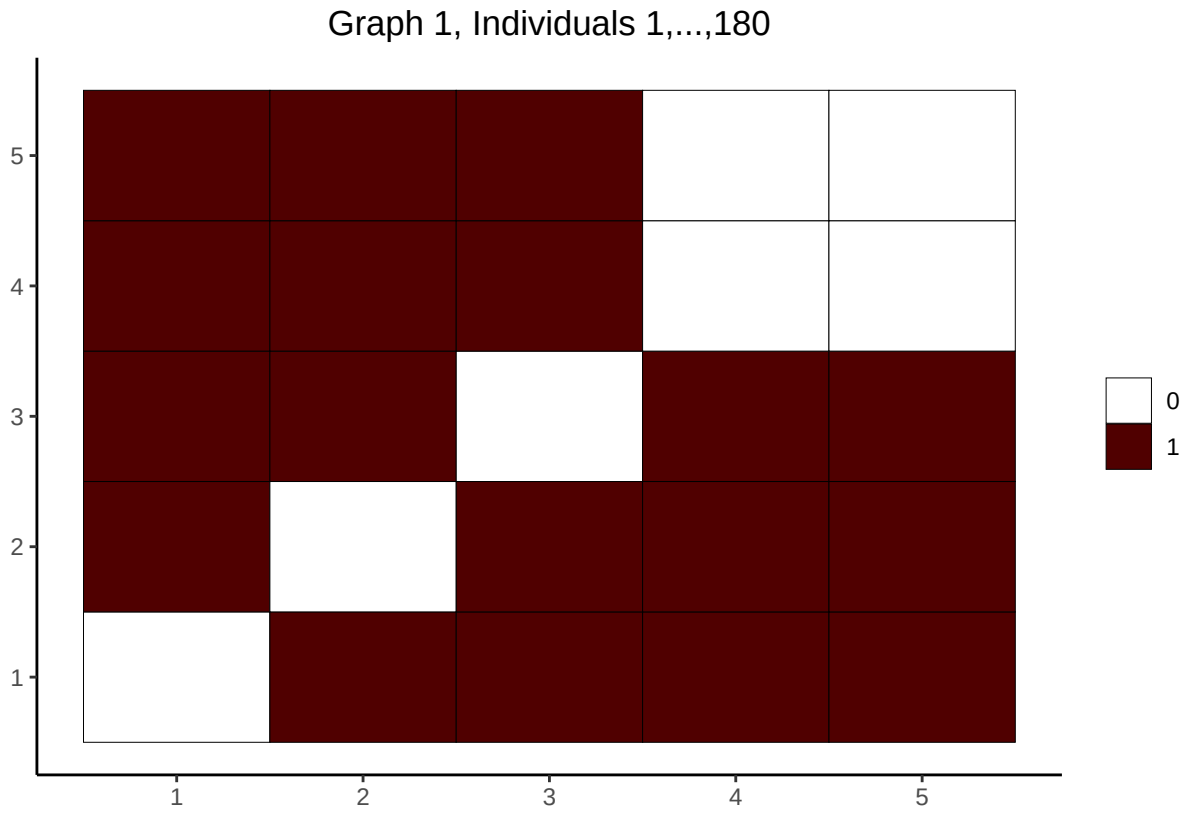

```
##
##
##
## TRIAL  10
##
## unif_prior
##                        Covariate Dependent Graphical Model
##
## Model ELBO: -9401.74                Unique conditional dependence structures: 1
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 1.737 mins
## [[1]]
```



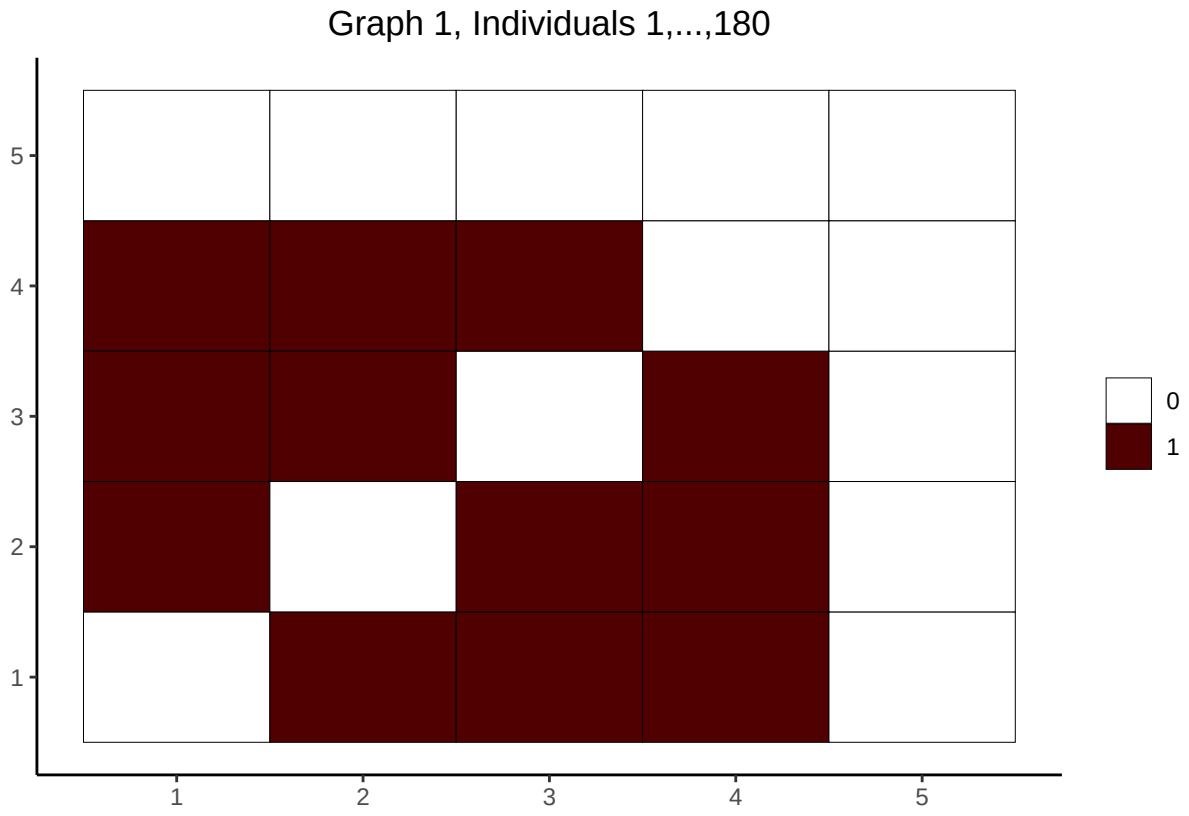

```
##
##
##
## TRIAL  10
##
## grid
##                        Covariate Dependent Graphical Model
##
## Model ELBO: -9401.2                 Unique conditional dependence structures: 1
## n: 180, variables: 5                      Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 44.453 secs
## [[1]]
```



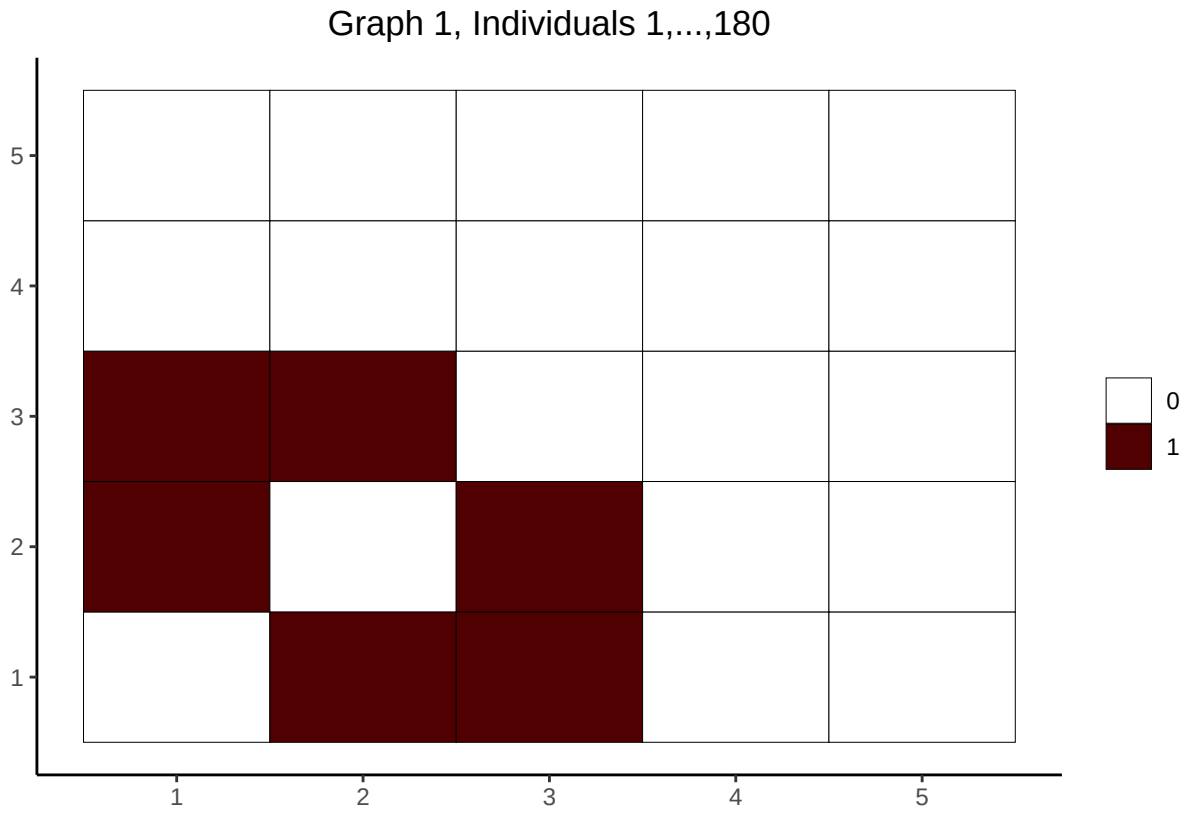

```
## [[1]]
```




```
## [[1]]
```




```
## [[1]]
```


Graph 1, Individuals 1,...,180
5
4
3
2
1
1
2
3
4
5
0
1