

Importance Sampling Experiment

4/3/2022

Contents

Executive Summary	1
Experiment Overview	1
prior prior specification	1
Fully Bayesian Approach	3
Grid generation	4
Results	4
Potential Next Steps	5
Detailed Results	5

Executive Summary

In this experiment, I compare a fully Bayesian model averaging approach to grid search. I find the performance of the model averaging approach to be significantly inferior to the grid search. Specifically, the model averaging results in the independence graph (a precision matrix of 0's) far too often.

Next steps will be to consider weightings calculated via the total ELBO across all n weighted regressions, rather than individual specific sets of weights.

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.

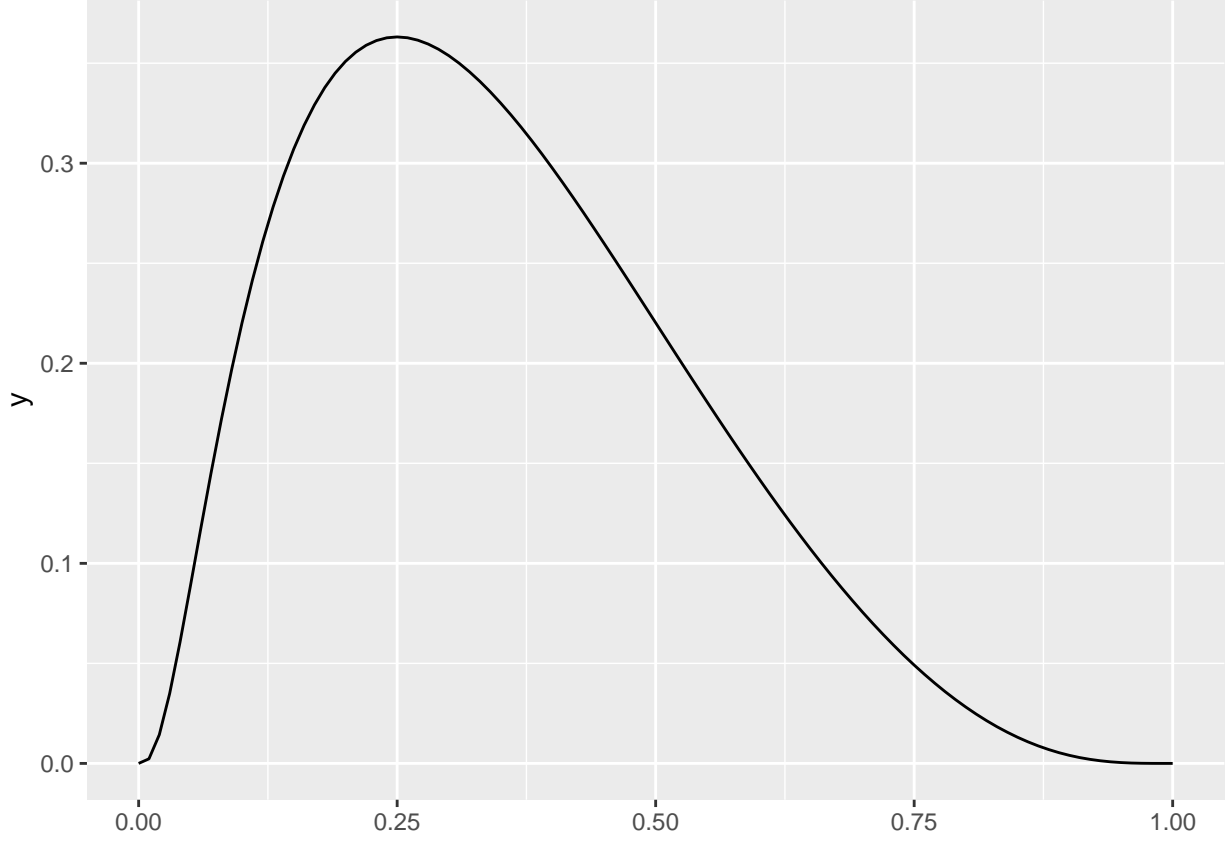
The first 2 methods (**prior** and **unif_prior**) employ a fully Bayesian approach of model averaging using importance weights, while the third method (**grid**) employs grid search.

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

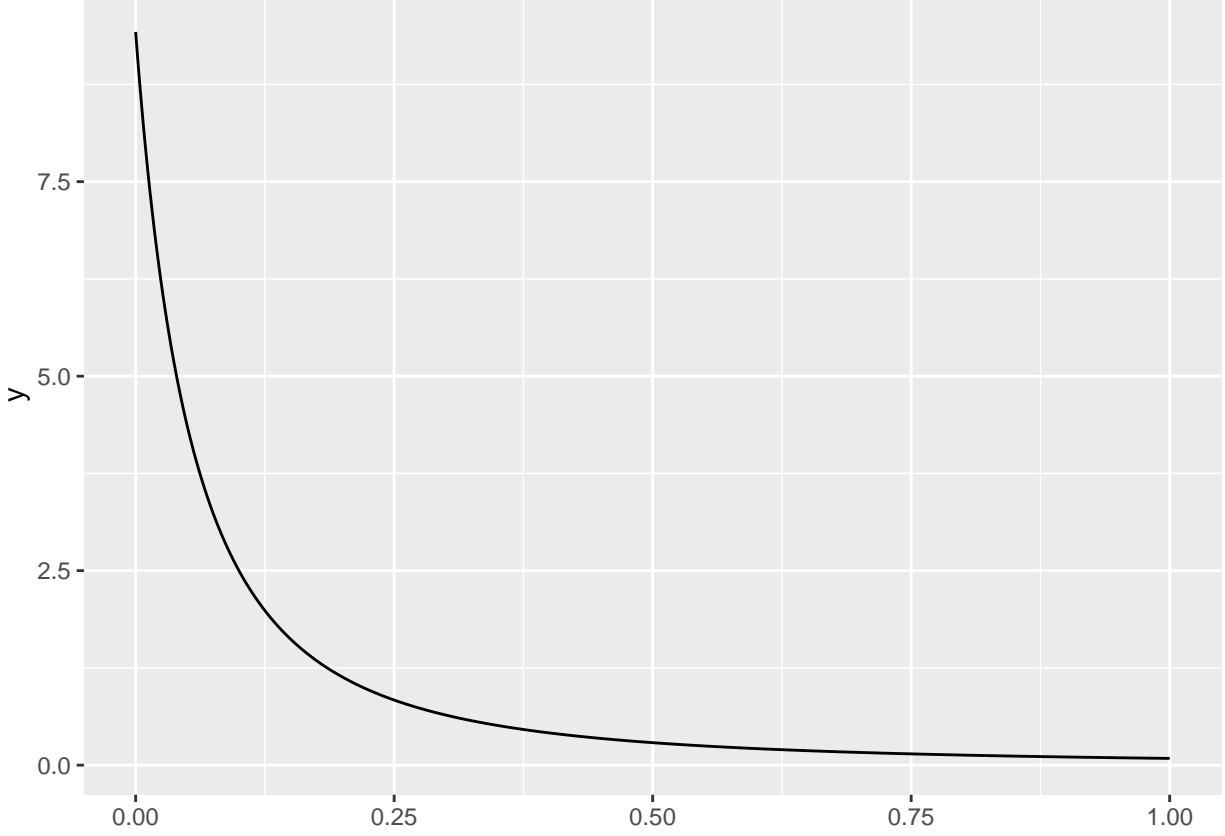
$$\hat{r} = \frac{\hat{s}_z}{1 + \hat{s}_z} \quad \hat{s}_z = \sigma_\beta^2 * \pi * T \quad T = \sum_{k=1}^p s_k^2 \quad (1)$$

Where s_k^2 is the sample variance of the data matrix with the k -th column removed.

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

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

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 $\tilde{w} = w(\theta_1), \dots, w(\theta_N)$, where $\theta_1, \dots, \theta_N$ 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 $\theta \in \tilde{\theta}$. The importance-weighted sum of the posterior quantities $p(\gamma_k = 1|X, y, \theta_i)$ will then be computed using:

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

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:

$$p(\theta_i) = p(\sigma_i^2) * p(\sigma_{\beta_i}^2) * p(\pi_i) \quad q(\theta_i) = q(\sigma_i^2) * q(\sigma_{\beta_i}^2) * q(\pi_i) \quad (4)$$

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

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

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

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

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

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

$$\sigma_{\beta, max}^2 = \frac{25}{\hat{\pi} \sum_{k=1}^p s_k^2} \quad (6)$$

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

Then the grid for the hyperparameter θ is generated by first creating a uniform grid on $[\log_{min} \theta, \log_{max} \theta]$ 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:

$$\mathcal{S} = \underset{\sim}{\sigma^2} \times \underset{\sim}{\sigma_\beta^2} \times \underset{\sim}{\pi} \quad (7)$$

Results

Detailed trial results are given below. It is readily apparent that a major shortcoming of the model averaging approach is that the sensitivity is too low. That is, `prior` and `unif_prior` produce graphs with no edges.

It is unclear why this is occurring. A potential explanation is some mistake in the code. However, this seems unlikely, since the average ELBO for each of the models is displayed. This is calculated by summing the weighted average of the ELBO across the individuals. If there were some mistake made in the calculation of the weights, then the average ELBO would presumably be different than the ELBO for grid search - however, the ELBO of the averaged models is relatively close to that of the grid search.

Potential Next Steps

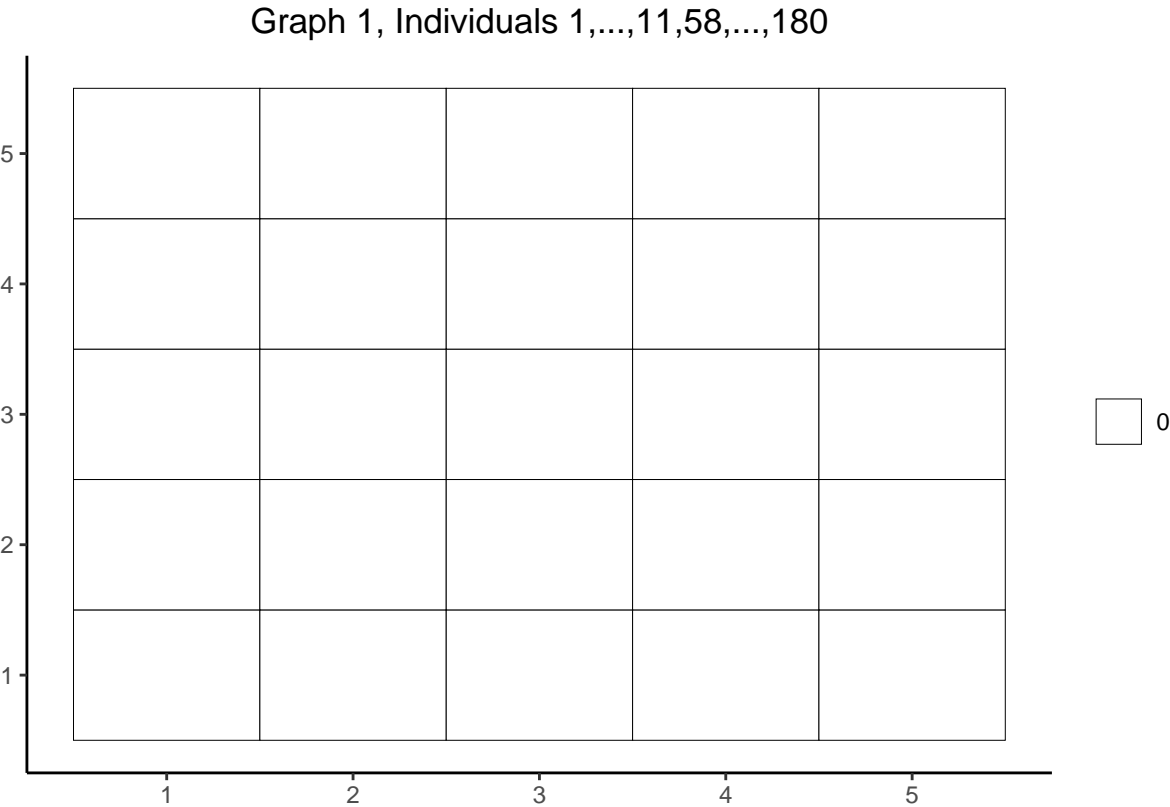
Next steps could be to calculate weights at the level of the variable fixed as the response instead of the individual level. Then, the weights would be calculated the same as before, except instead of approximating $p(y|X, \theta_i)$ by $\exp(ELBO_j^l(\theta_i))$, $p(y|X, \theta_i)$ could be approximated by $\exp(ELBO_j(\theta_i))$, where:

$$\exp(ELBO_j(\theta_i)) = \sum_{l=1}^n \exp(ELBO_j^l(\theta_i)) \tag{8}$$

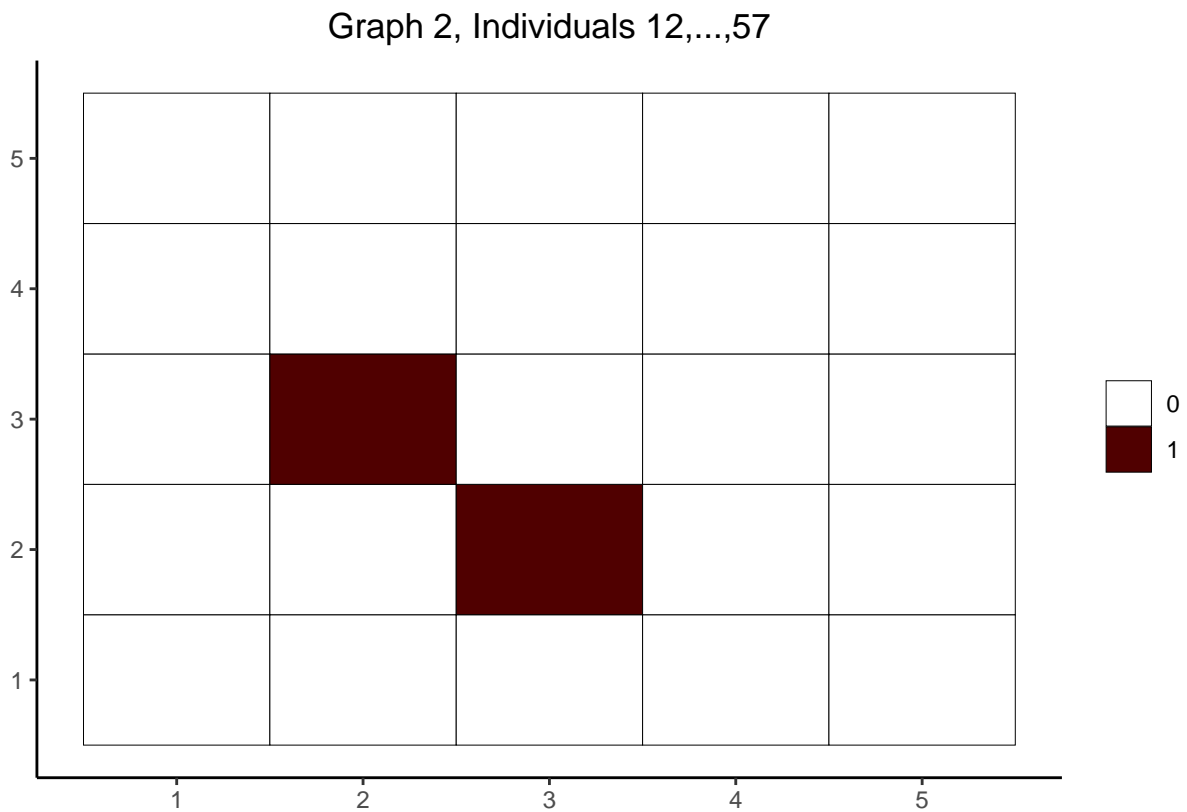
A potential issue with this is that it is unclear whether proposition 1 would still hold.

Detailed Results

```
## - - - - -
## TRIAL  1
##
## prior
##              Covariate Dependent Graphical Model
##
## Model ELBO: -9666.02           Unique conditional dependence structures: 2
## n: 180, variables: 5           Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 2.097 mins
## [[1]]
```

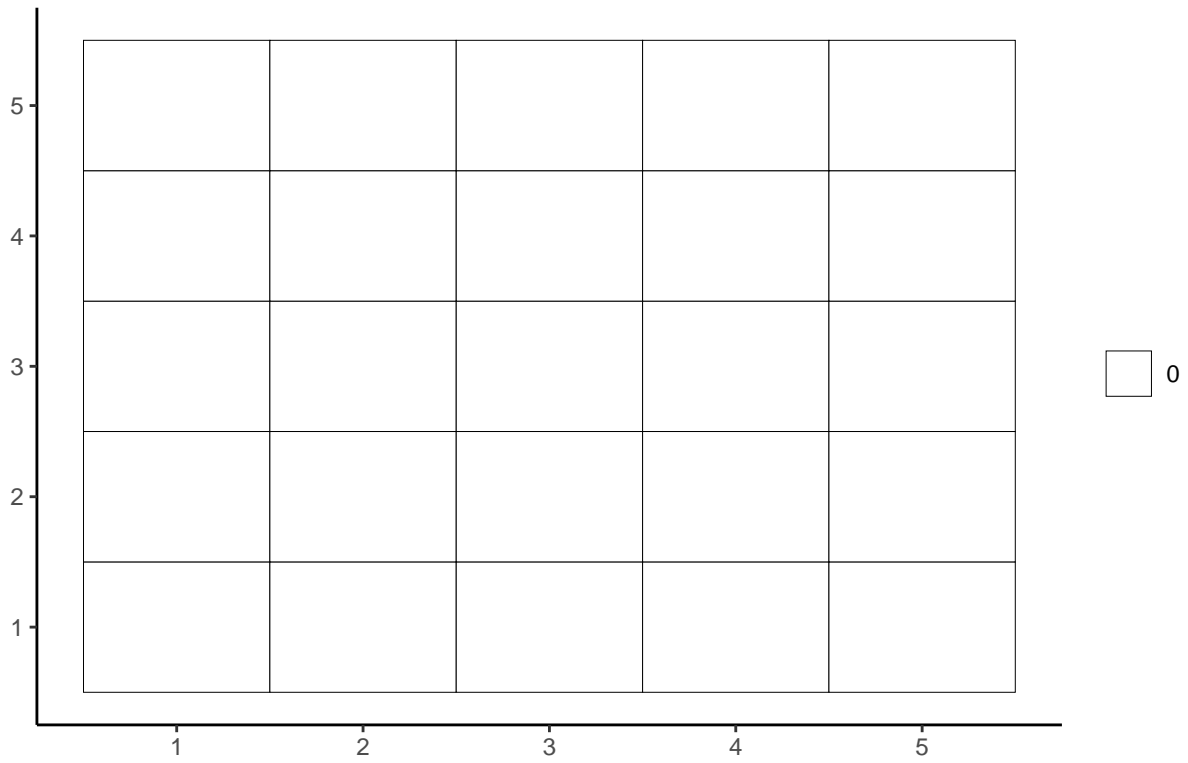


```
##
## [[2]]
```



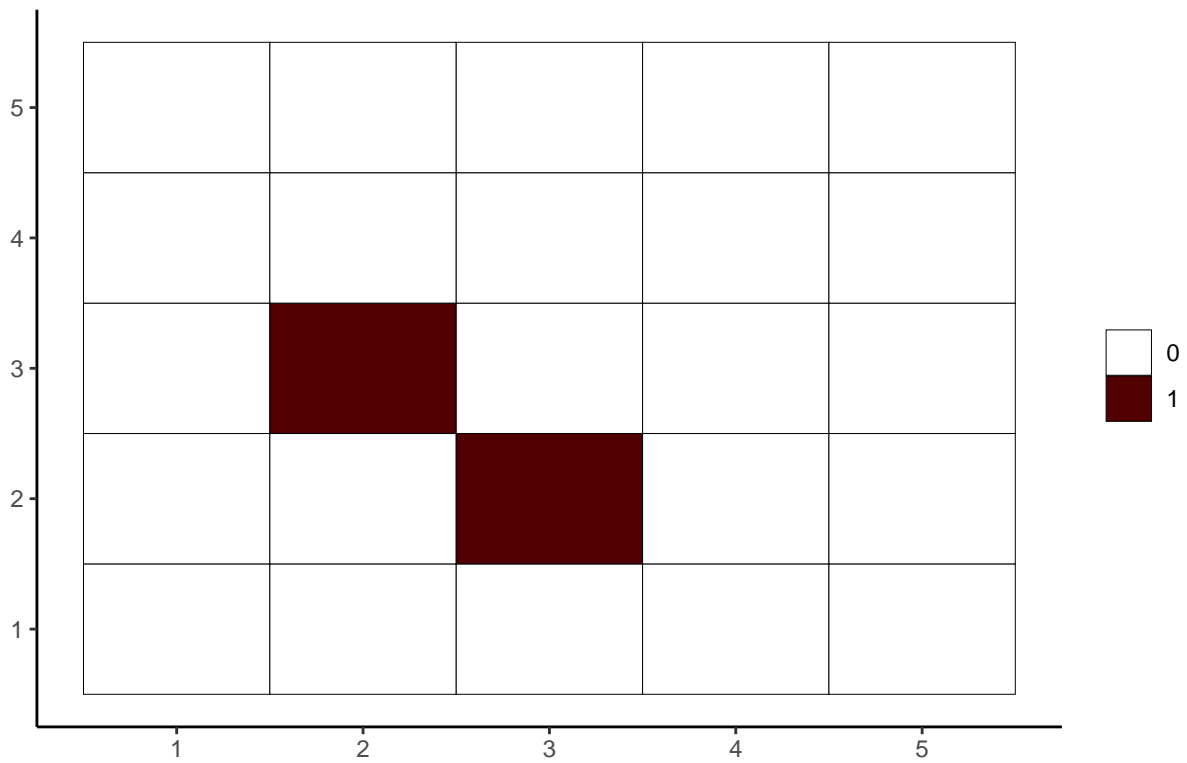
```
##
##
##
## TRIAL 1
##
## unif_prior
## Covariate Dependent Graphical Model
##
## Model ELBO: -9732.29 Unique conditional dependence structures: 2
## n: 180, variables: 5 Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 2.116 mins
## [[1]]
```

Graph 1, Individuals 1,...,29,48,...,180



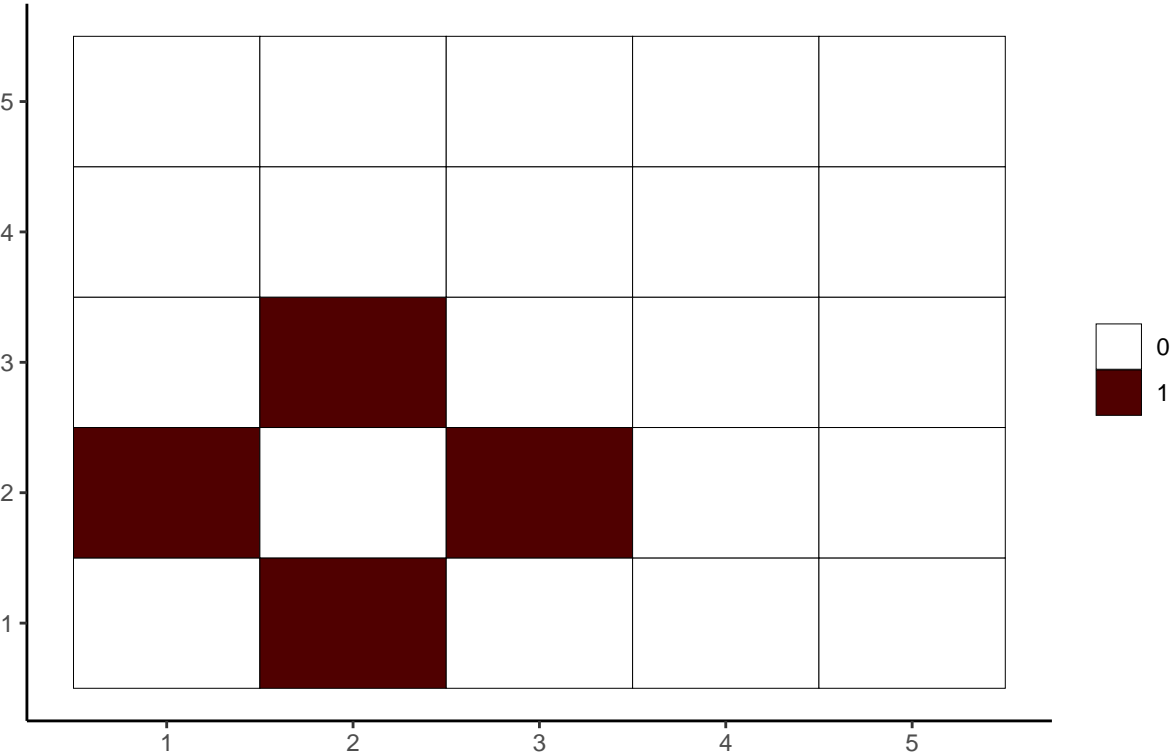
[[2]]

Graph 2, Individuals 30,...,47



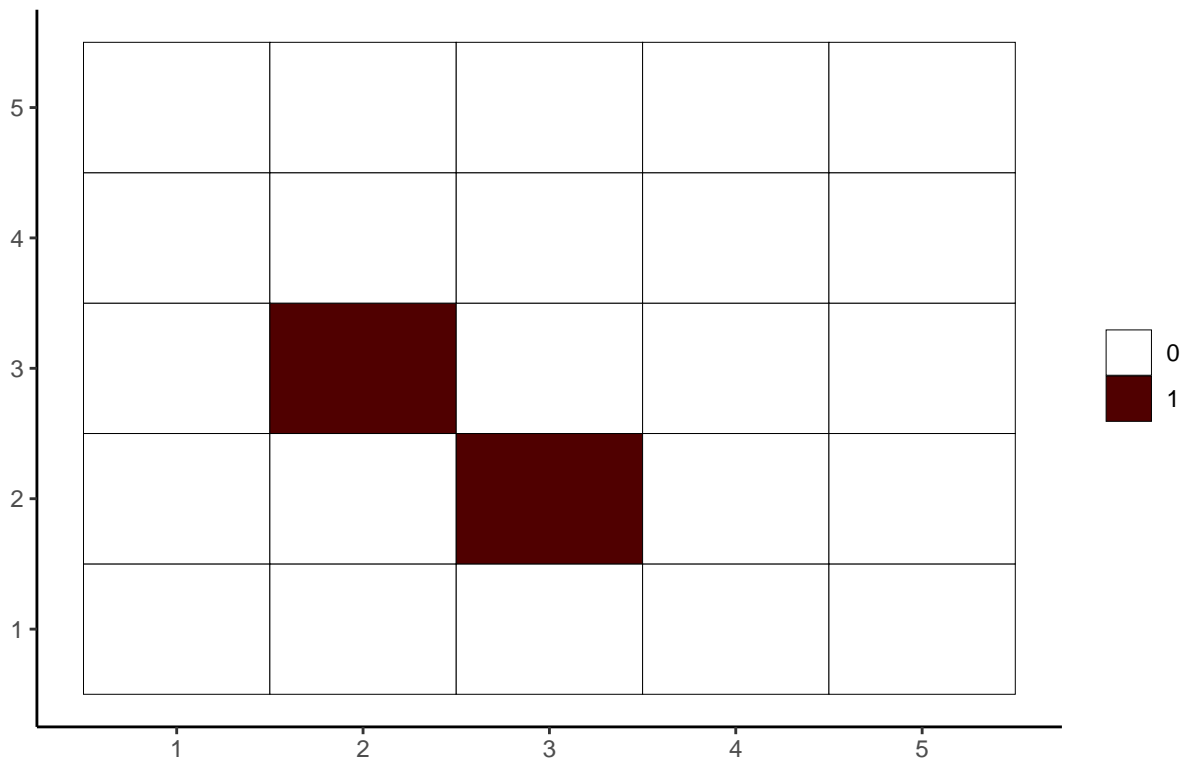
```
##
##
##
## TRIAL 1
##
## grid
##          Covariate Dependent Graphical Model
##
## Model ELBO: -9332.56          Unique conditional dependence structures: 2
## n: 180, variables: 5          Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 55.037 secs
## [[1]]
```


Graph 1, Individuals 1,...,152



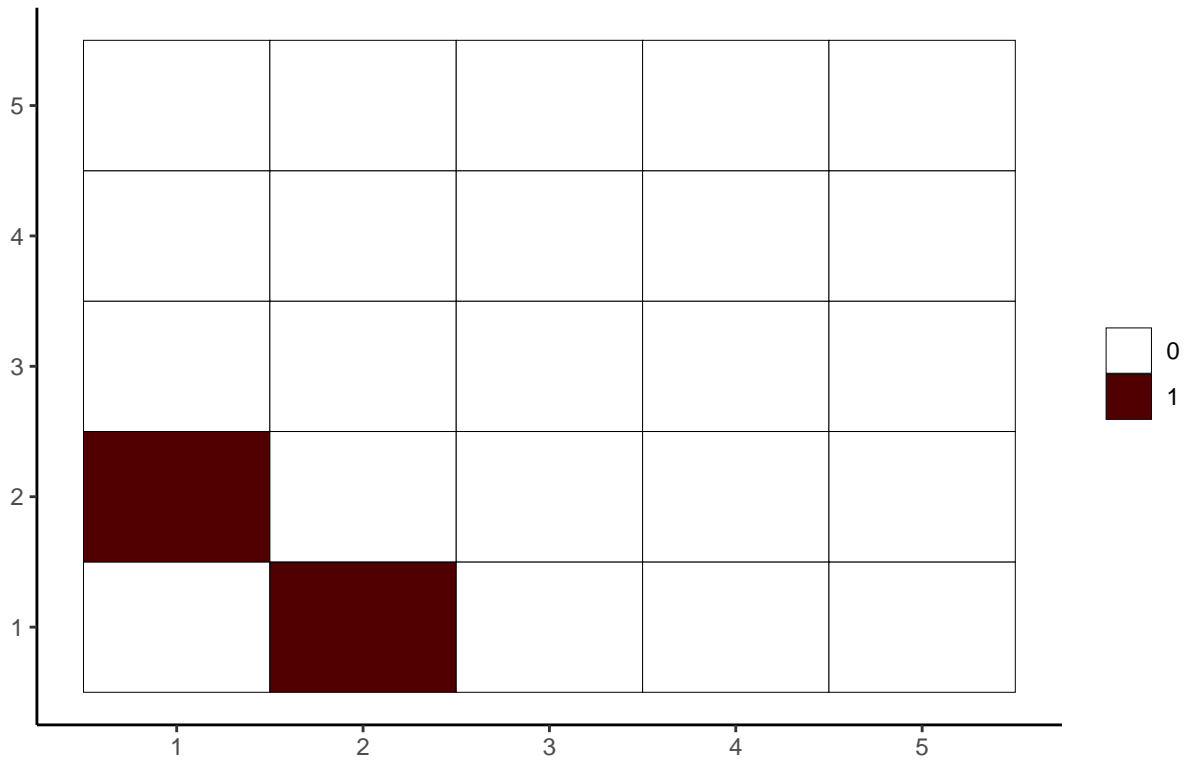
```
##  
## [[2]]
```

Graph 2, Individuals 153,...,180



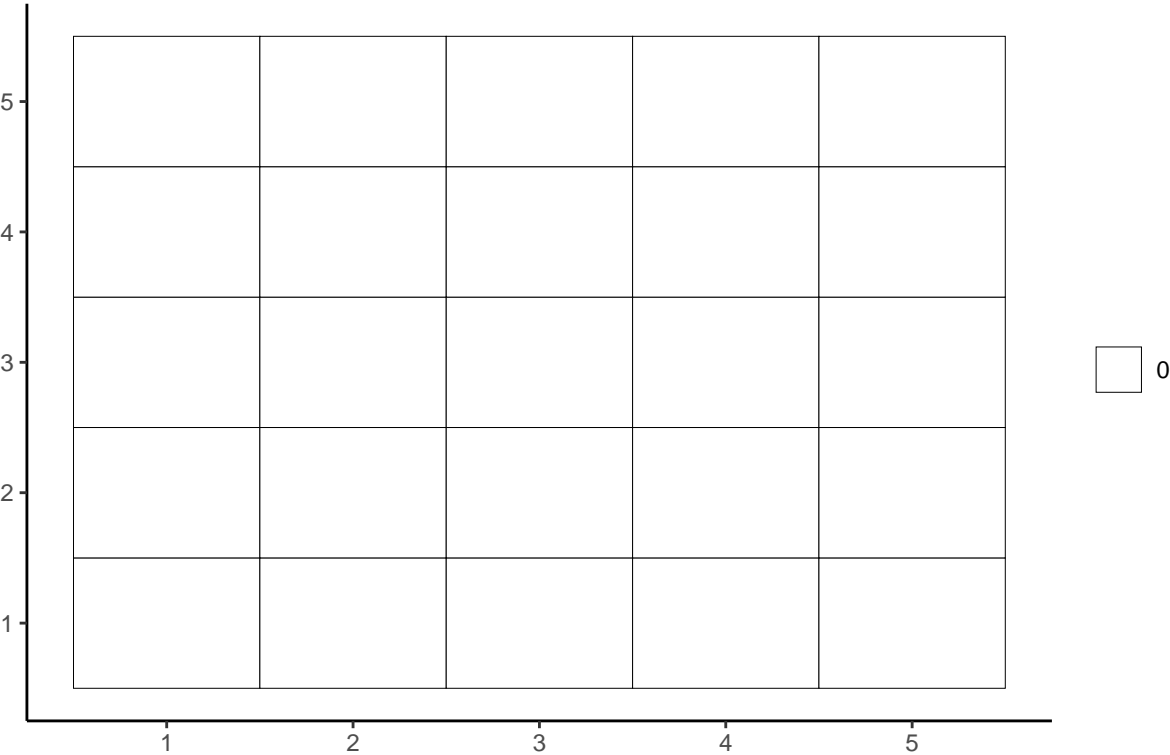
```
##
## -----
## TRIAL 2
##
## prior
##           Covariate Dependent Graphical Model
##
## Model ELBO: -9357.67           Unique conditional dependence structures: 3
## n: 180, variables: 5           Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 2.28 mins
## [[1]]
```

Graph 1, Individuals 1,...,25



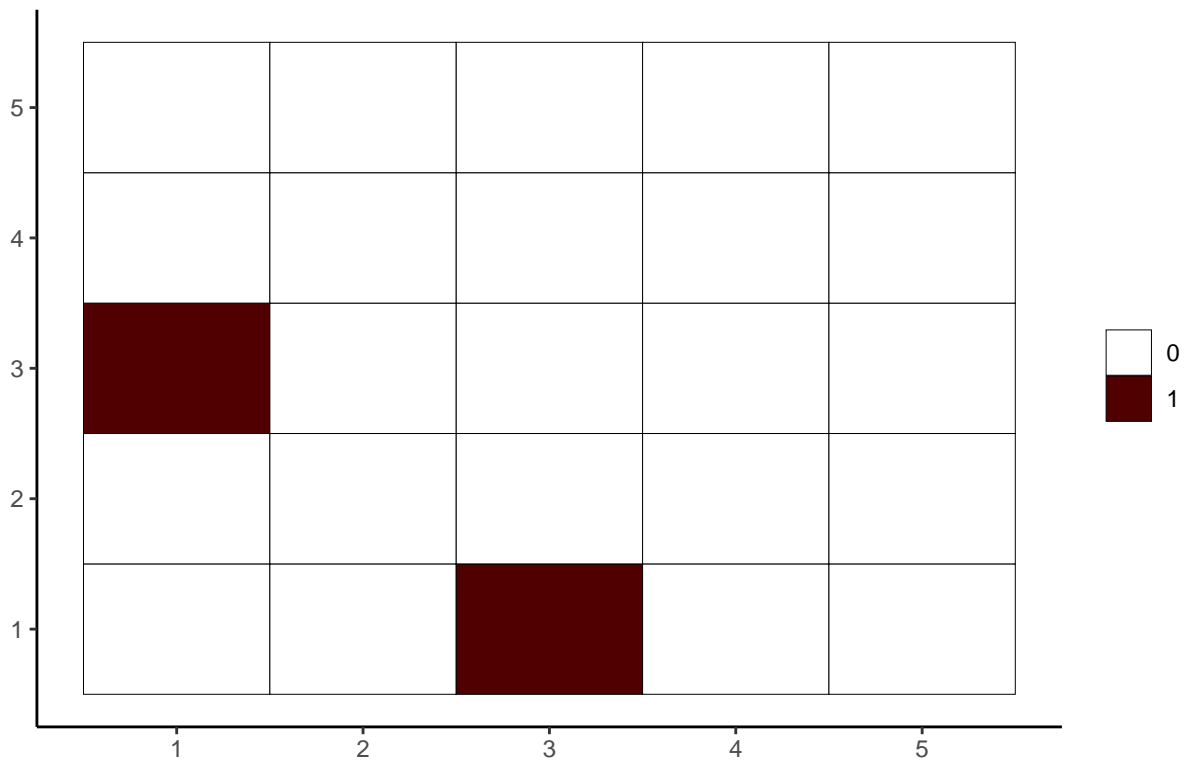
```
##  
## [[2]]
```

Graph 2, Individuals 26,...,153



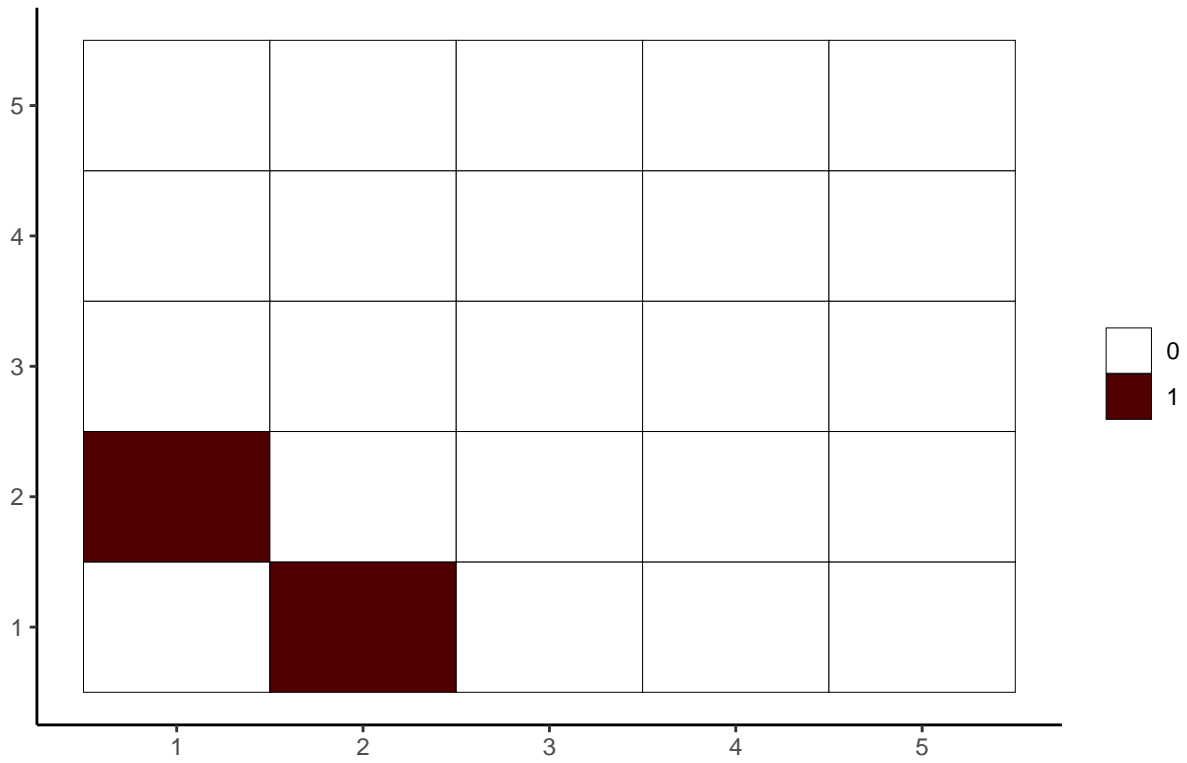
[[3]]

Graph 3, Individuals 154,...,180



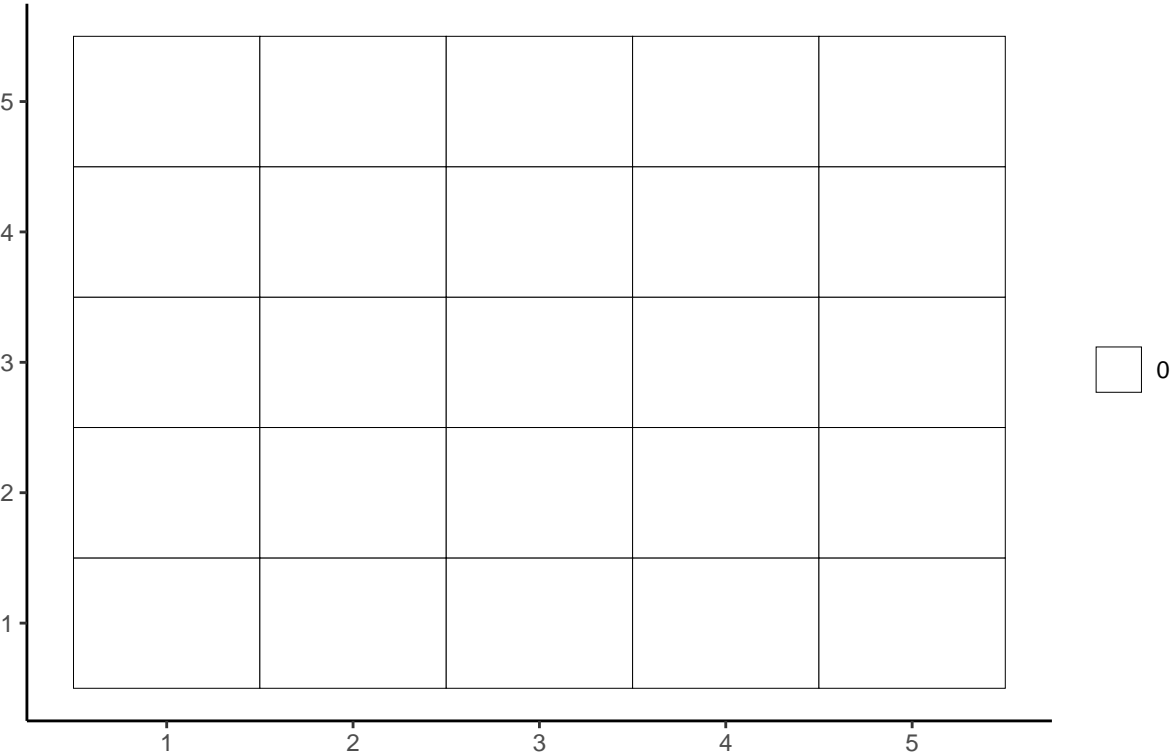
```
##
##
##
## TRIAL 2
##
## unif_prior
##          Covariate Dependent Graphical Model
##
## Model ELBO: -9417.18          Unique conditional dependence structures: 3
## n: 180, variables: 5          Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 2.258 mins
## [[1]]
```

Graph 1, Individuals 1,...,24



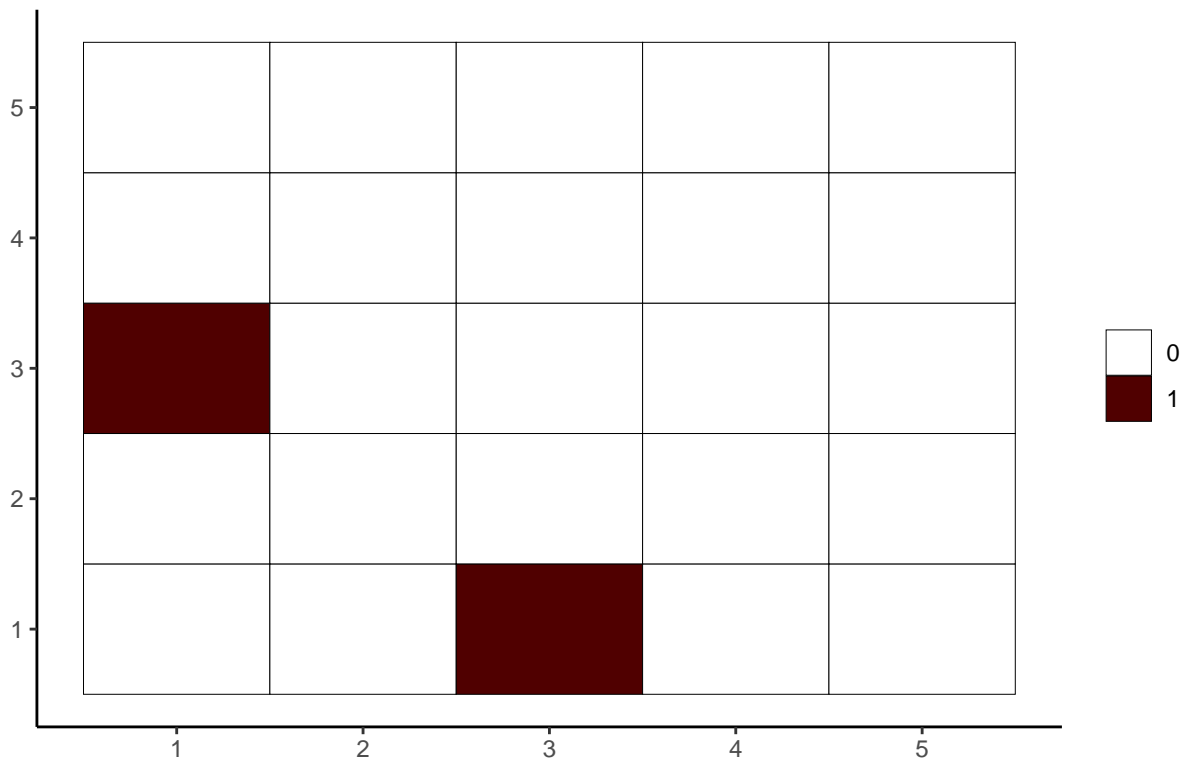
[[2]]

Graph 2, Individuals 25,...,154



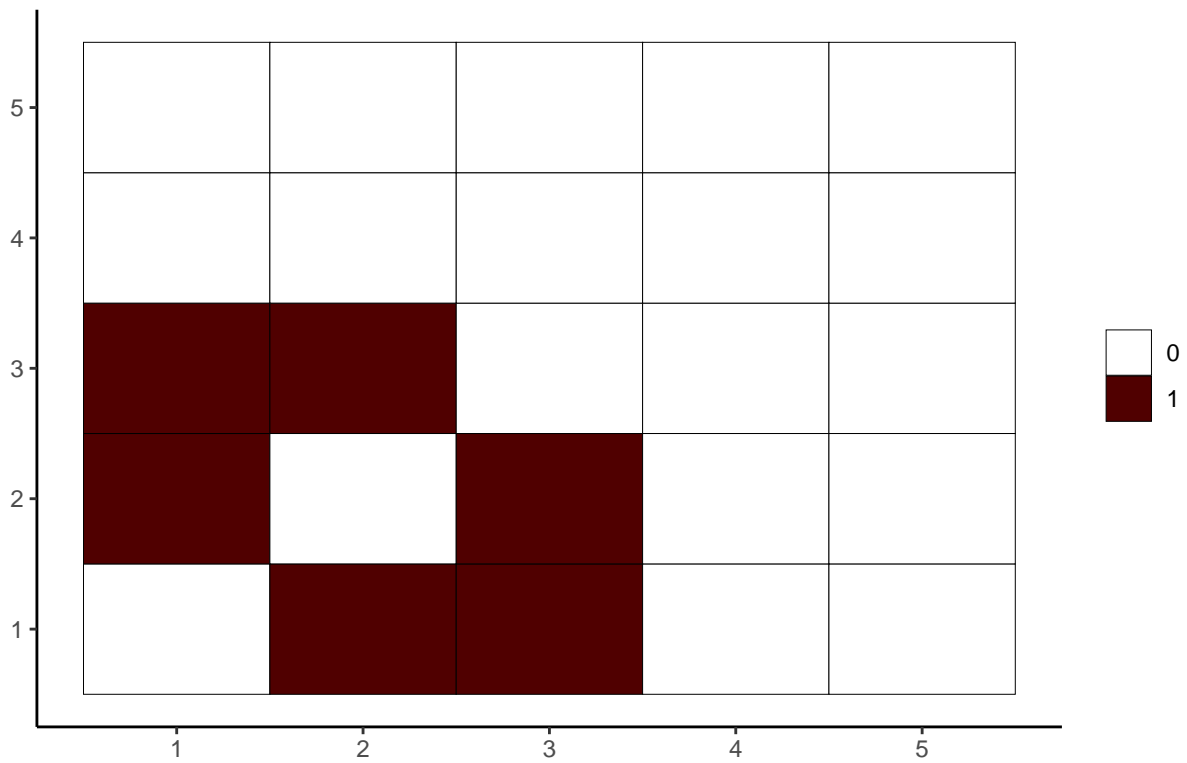
[[3]]

Graph 3, Individuals 155,...,180



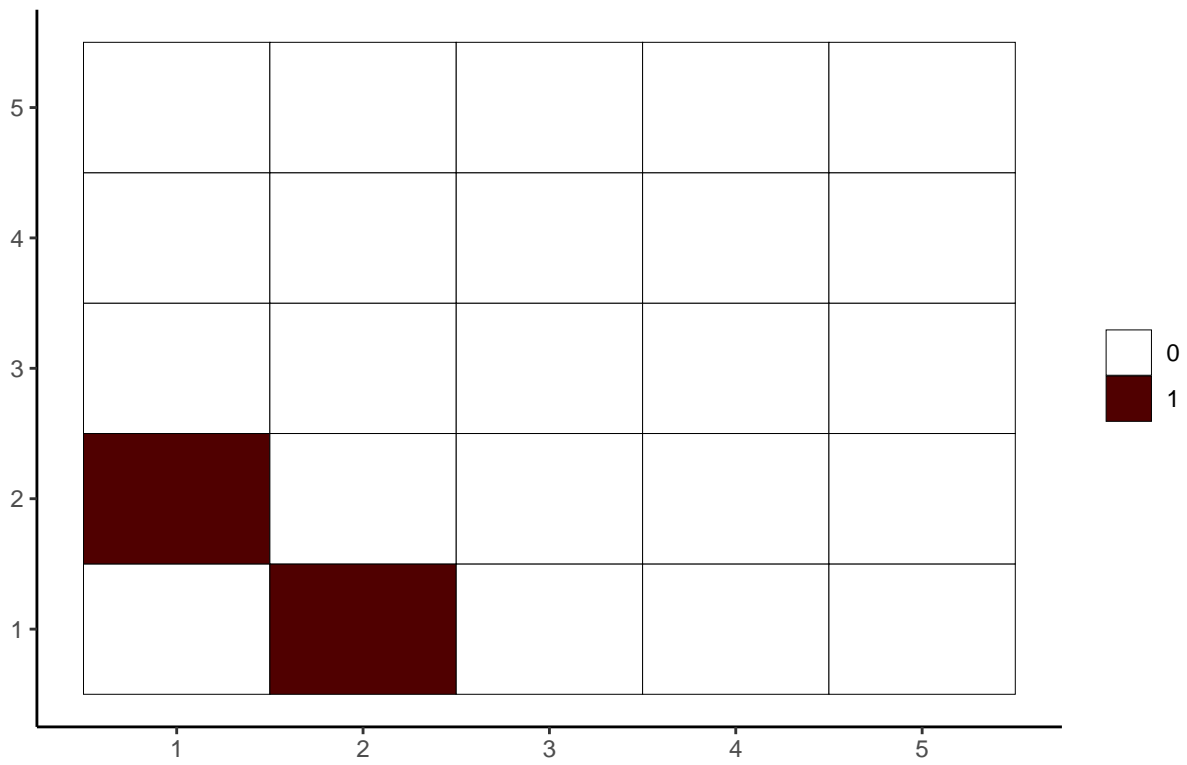
```
##
##
##
## TRIAL 2
##
## grid
##          Covariate Dependent Graphical Model
##
## Model ELBO: -9128.94          Unique conditional dependence structures: 1
## n: 180, variables: 5          Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 50.924 secs
## [[1]]
```


Graph 1, Individuals 1,...,180



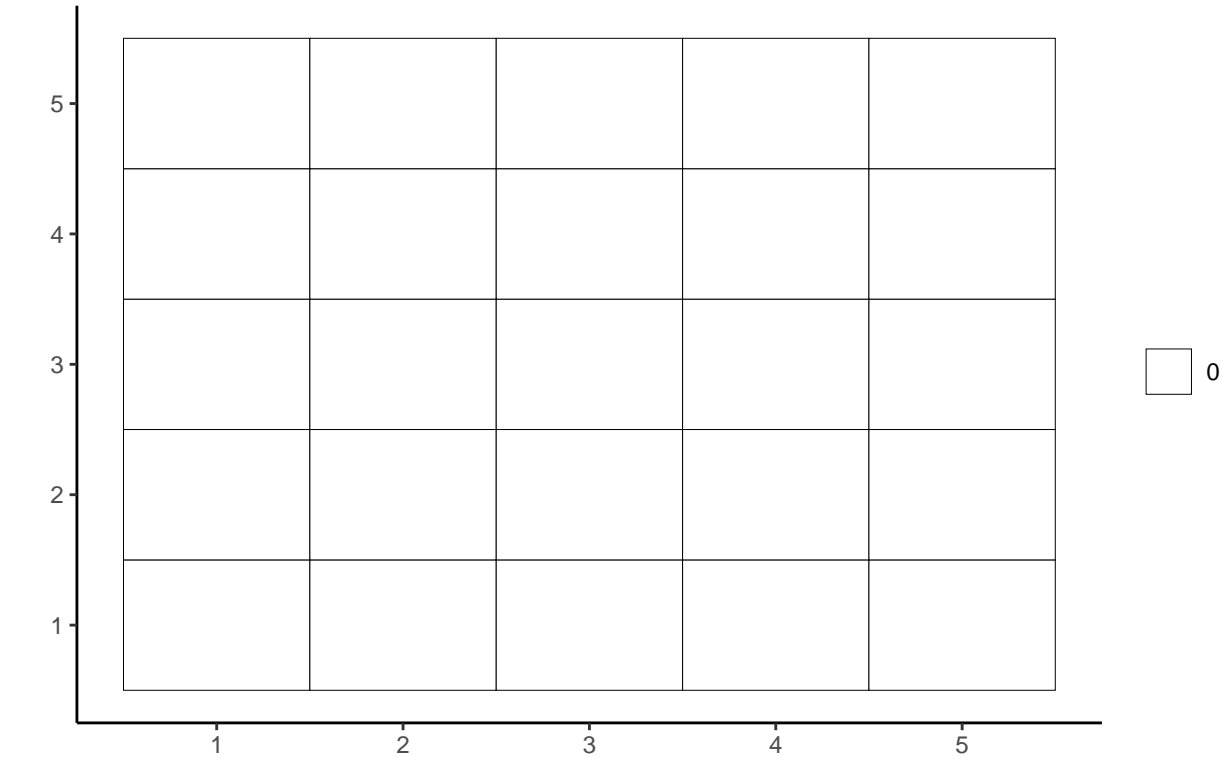
```
##
## -----
## TRIAL 3
##
## prior
##           Covariate Dependent Graphical Model
##
## Model ELBO: -9559.35           Unique conditional dependence structures: 3
## n: 180, variables: 5           Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 2.31 mins
## [[1]]
```

Graph 1, Individuals 1,...,28



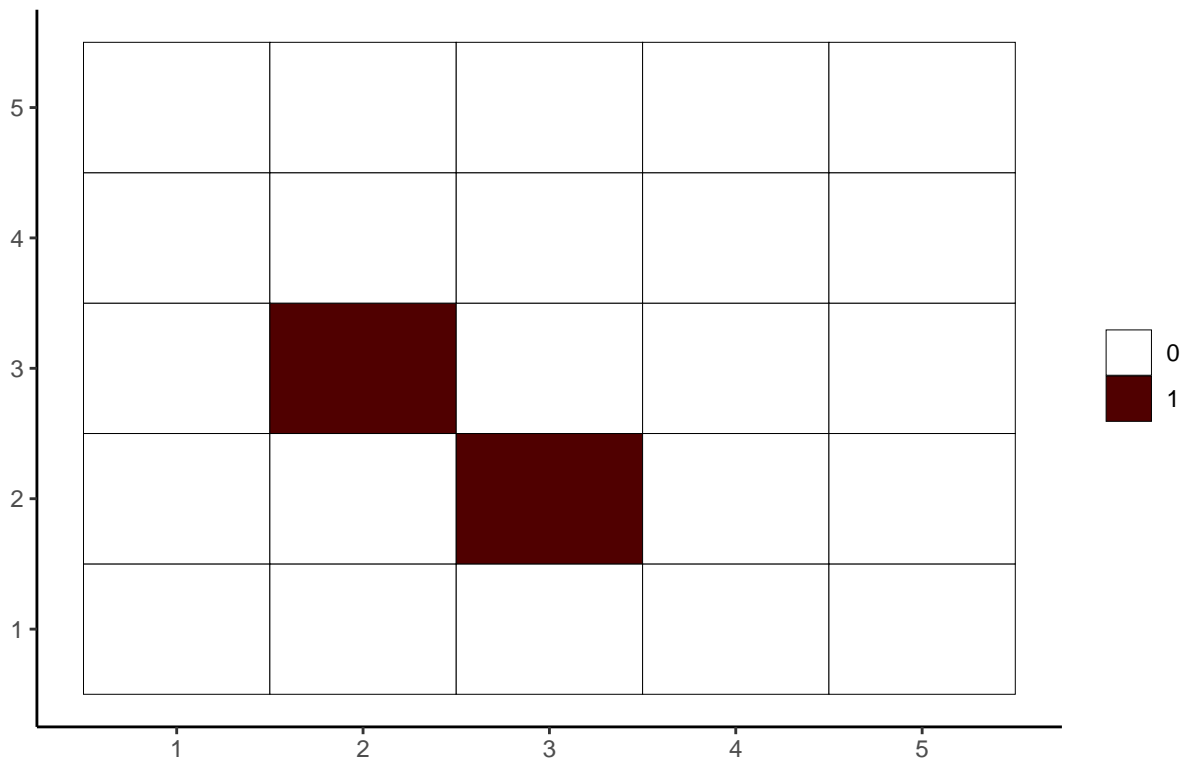
[[2]]

Graph 2, Individuals 29,...,123,153,...,180



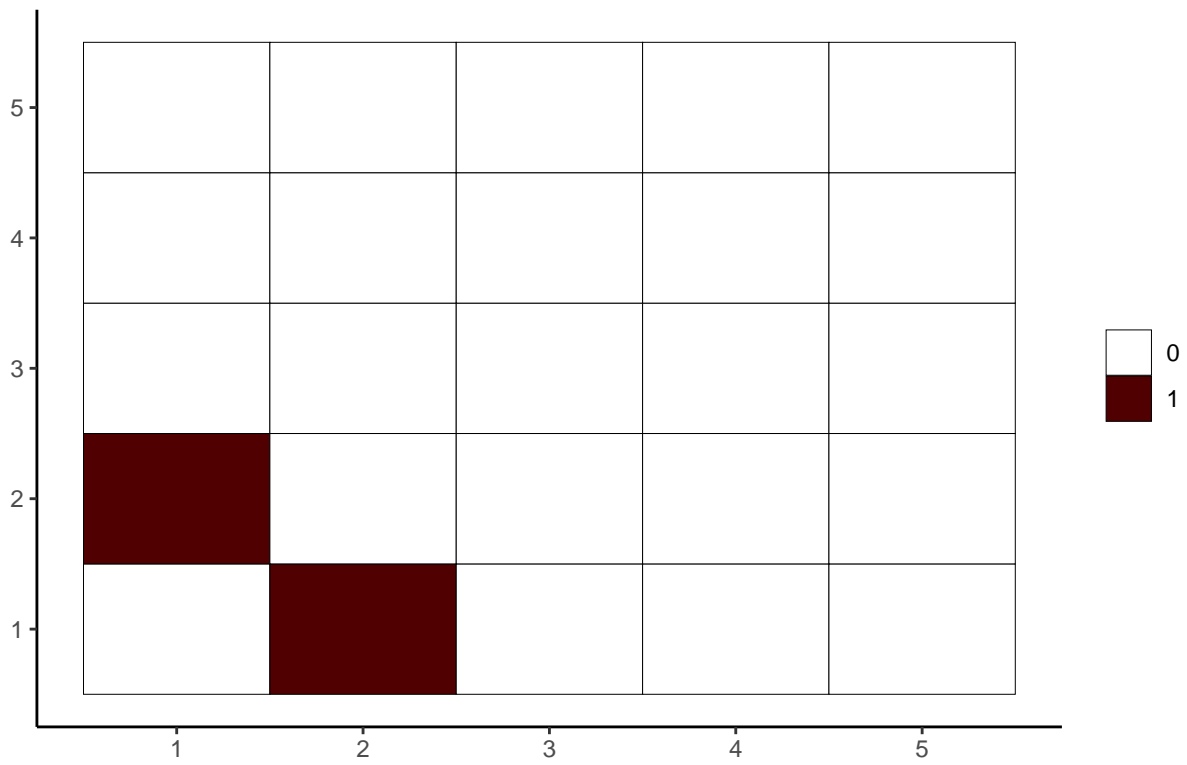
[[3]]

Graph 3, Individuals 124,...,152



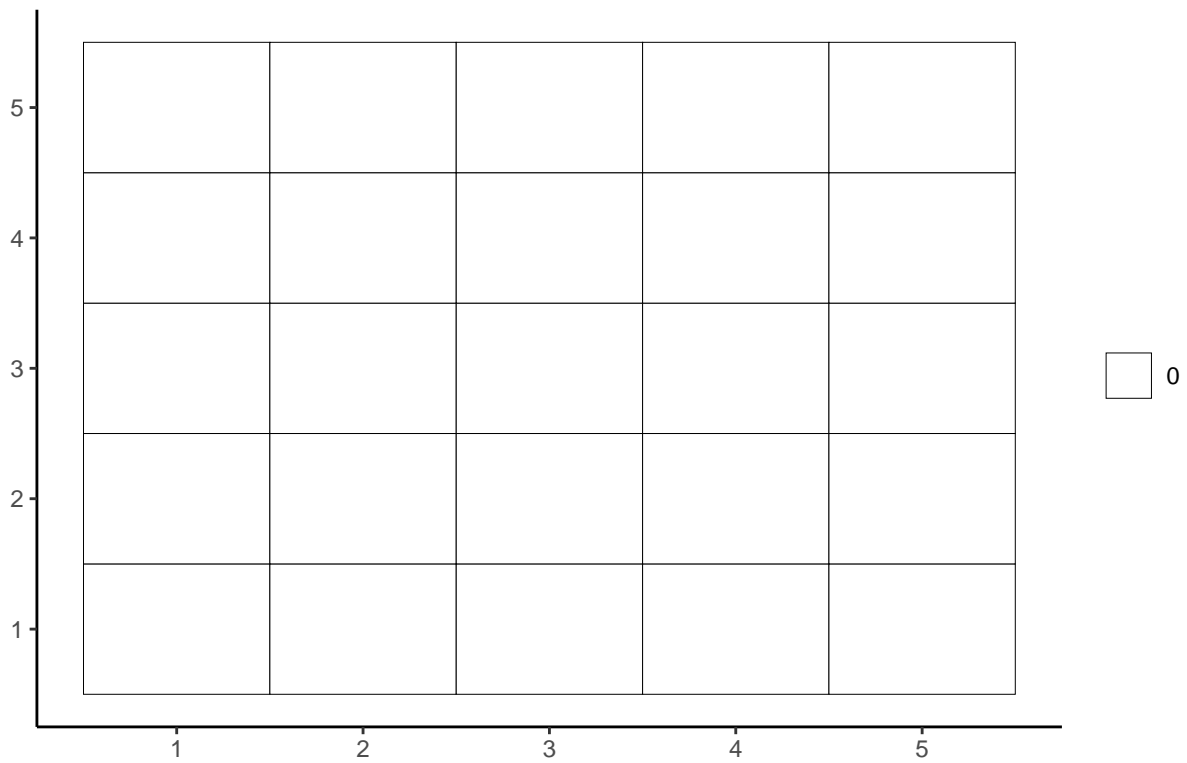
```
##
##
##
## TRIAL 3
##
## unif_prior
##           Covariate Dependent Graphical Model
##
## Model ELBO: -9623.85           Unique conditional dependence structures: 2
## n: 180, variables: 5           Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 2.27 mins
## [[1]]
```

Graph 1, Individuals 1,...,24



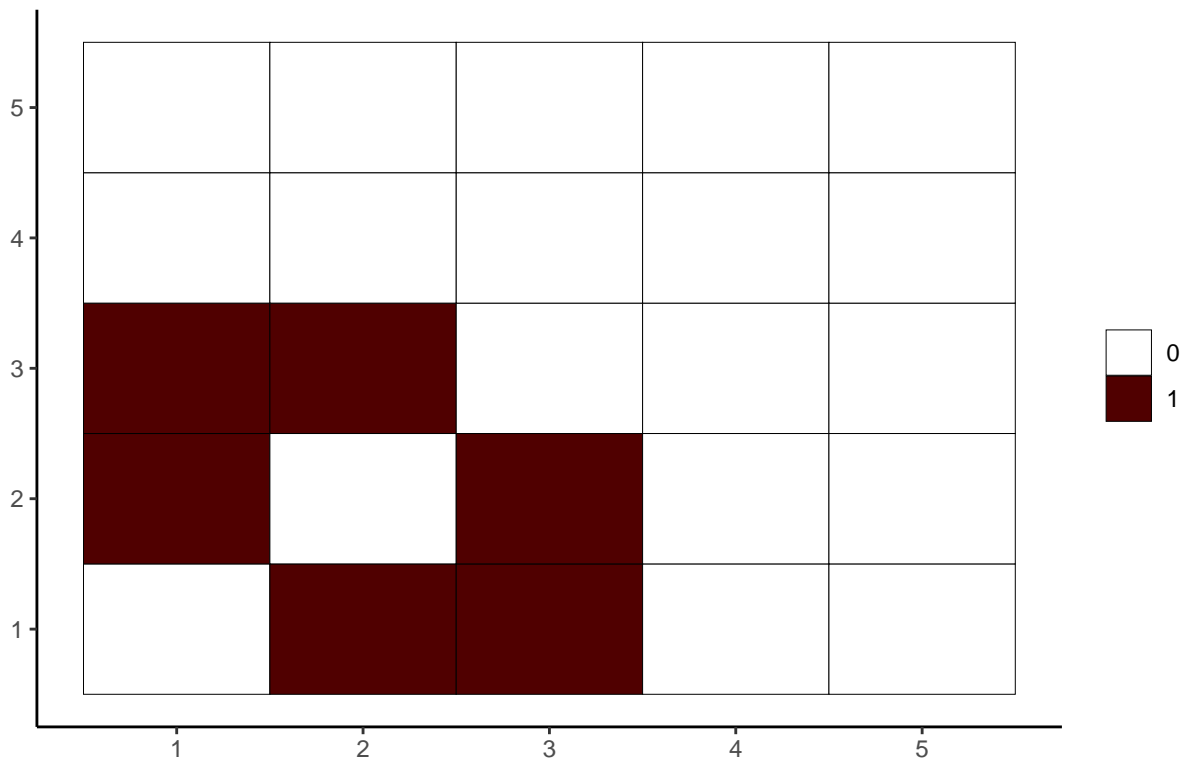
[[2]]

Graph 2, Individuals 25,...,180



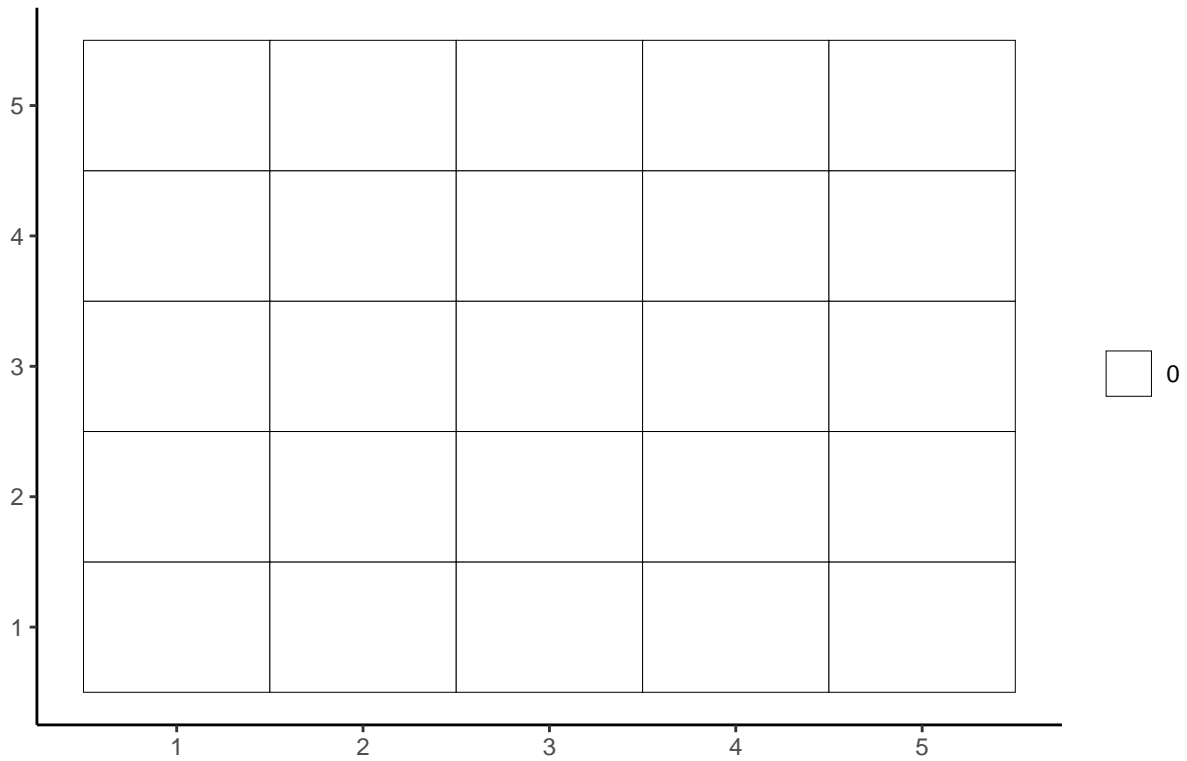
```
##
##
##
## TRIAL 3
##
## grid
##          Covariate Dependent Graphical Model
##
## Model ELBO: -9241.55          Unique conditional dependence structures: 1
## n: 180, variables: 5          Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 48.943 secs
## [[1]]
```

Graph 1, Individuals 1,...,180



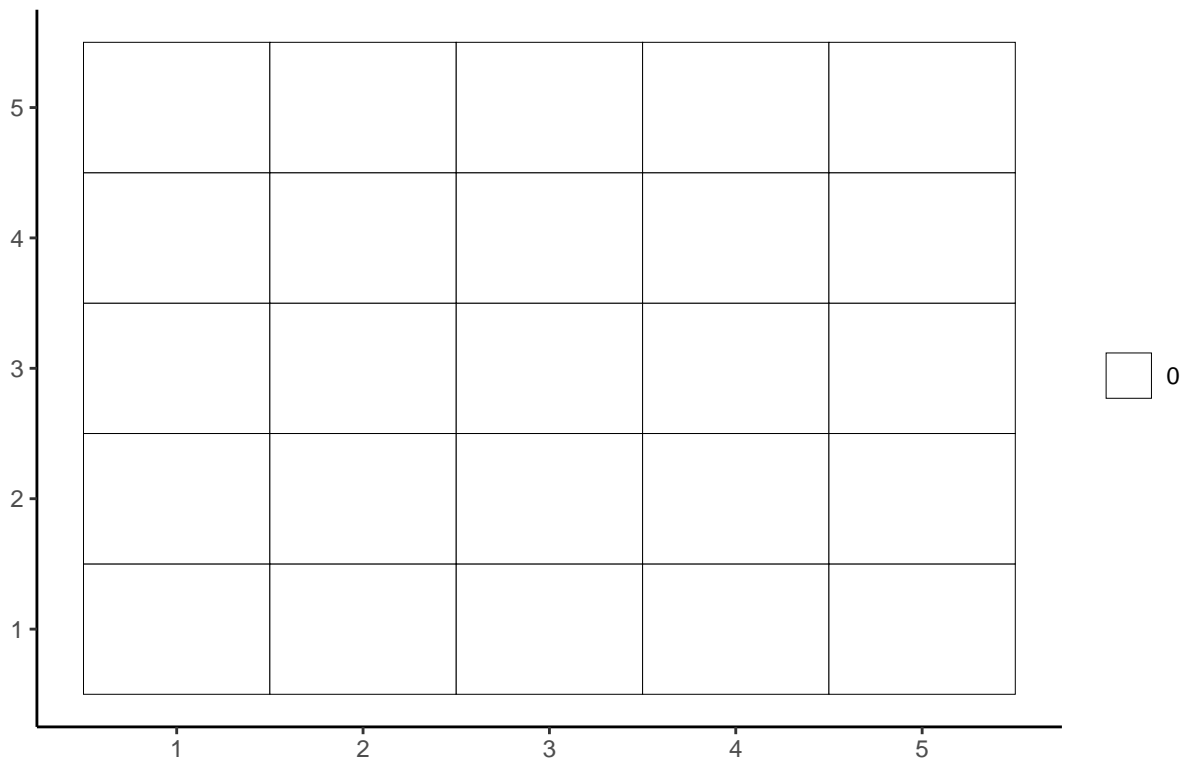
```
##
## -----
## TRIAL 4
##
## prior
##           Covariate Dependent Graphical Model
##
## Model ELBO: -9700.38           Unique conditional dependence structures: 1
## n: 180, variables: 5           Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 1.966 mins
## [[1]]
```

Graph 1, Individuals 1,...,180



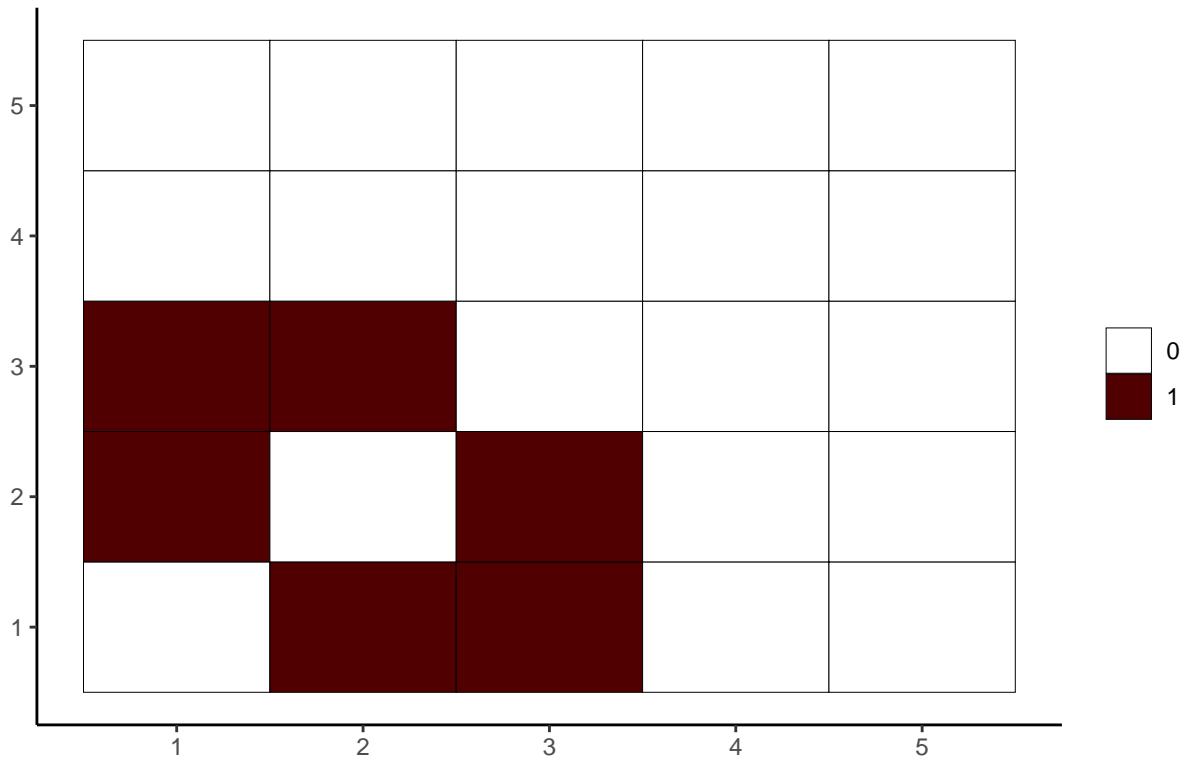
```
##
##
##
## TRIAL 4
##
## unif_prior
##          Covariate Dependent Graphical Model
##
## Model ELBO: -9758.63          Unique conditional dependence structures: 1
## n: 180, variables: 5          Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 1.922 mins
## [[1]]
```


Graph 1, Individuals 1,...,180



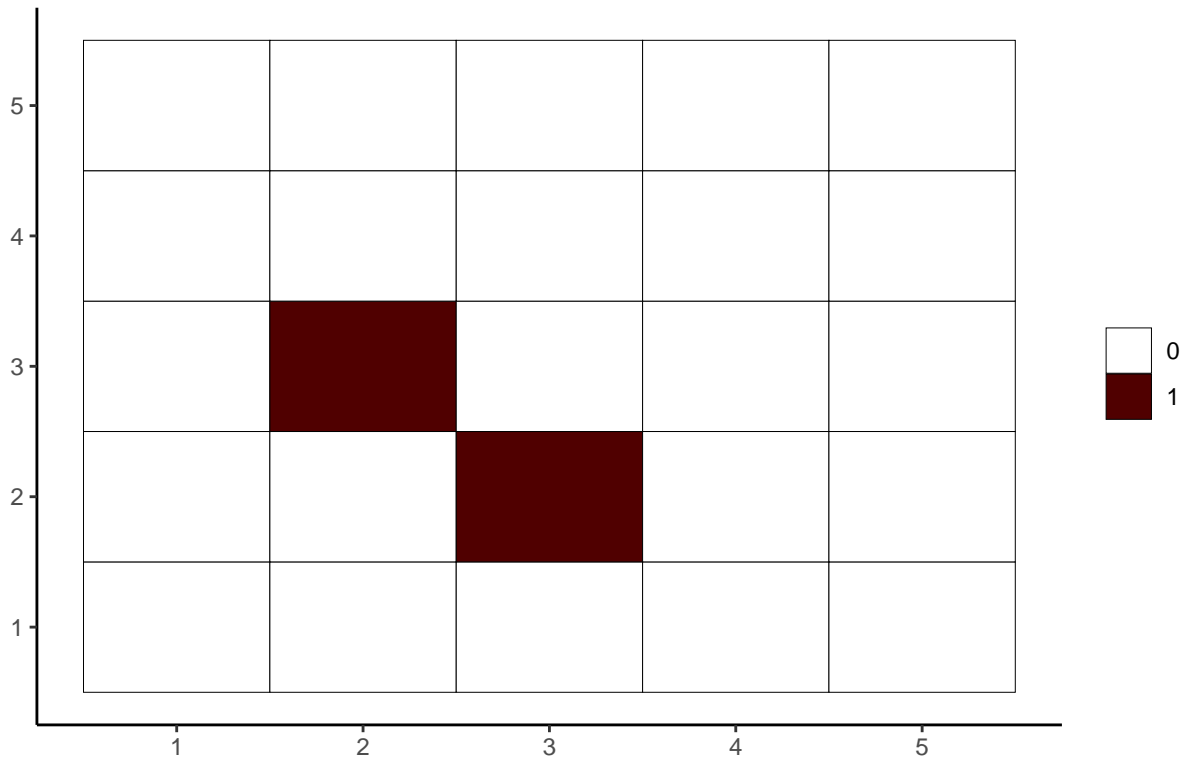
```
##
##
##
## TRIAL 4
##
## grid
##          Covariate Dependent Graphical Model
##
## Model ELBO: -9547.98          Unique conditional dependence structures: 1
## n: 180, variables: 5          Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 52.246 secs
## [[1]]
```

Graph 1, Individuals 1,...,180



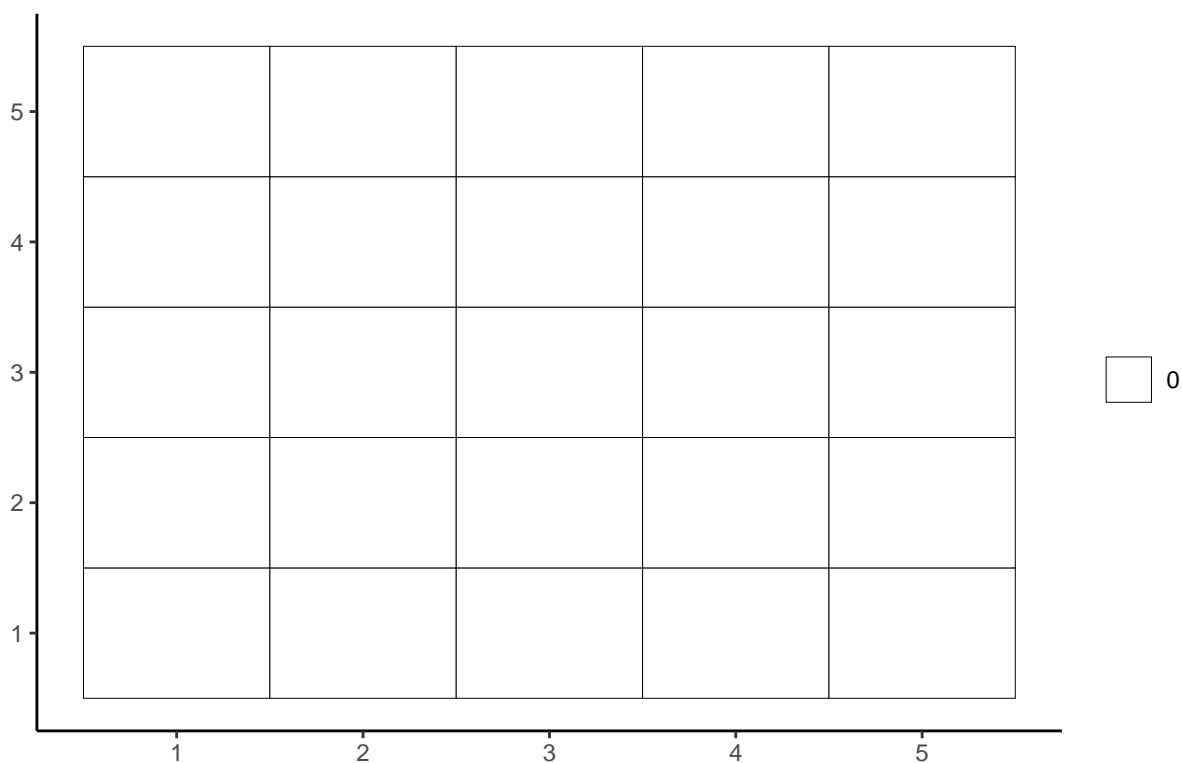
```
##
## -----
## TRIAL 5
##
## prior
##           Covariate Dependent Graphical Model
##
## Model ELBO: -9571.91           Unique conditional dependence structures: 2
## n: 180, variables: 5           Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 2.321 mins
## [[1]]
```

Graph 1, Individuals 1,...,26



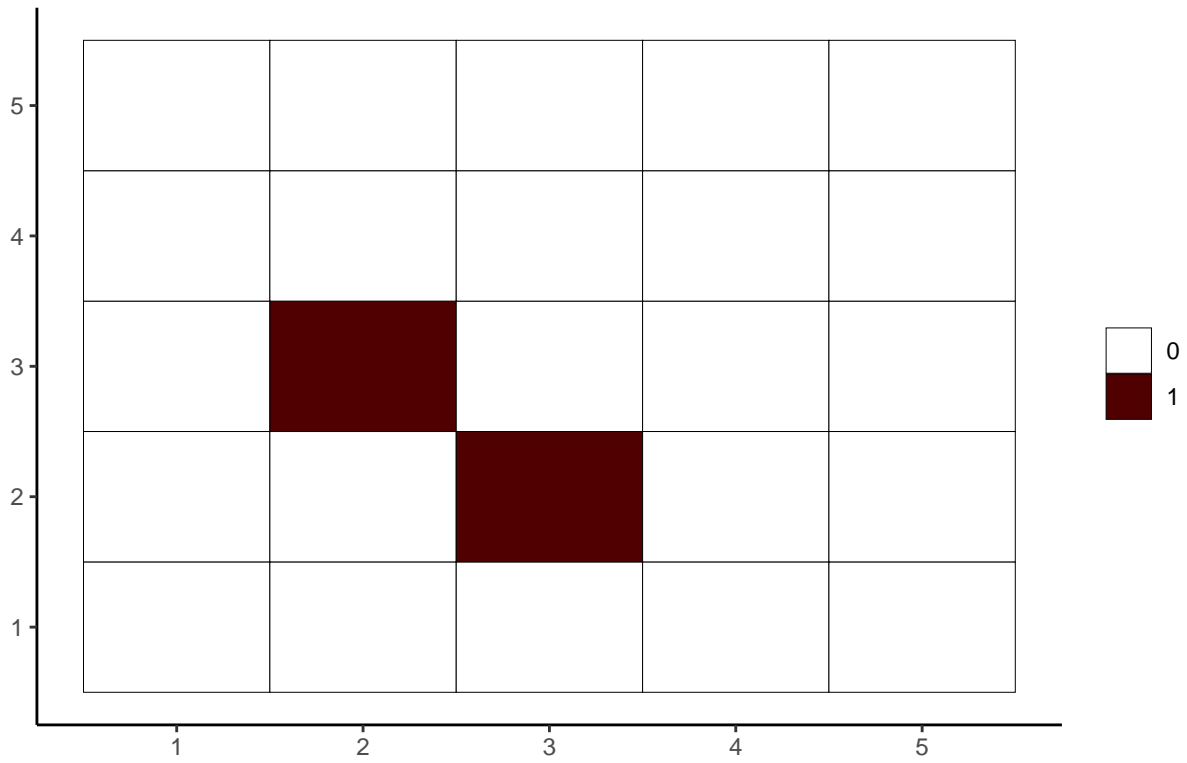
```
##  
## [[2]]
```

Graph 2, Individuals 27,...,180



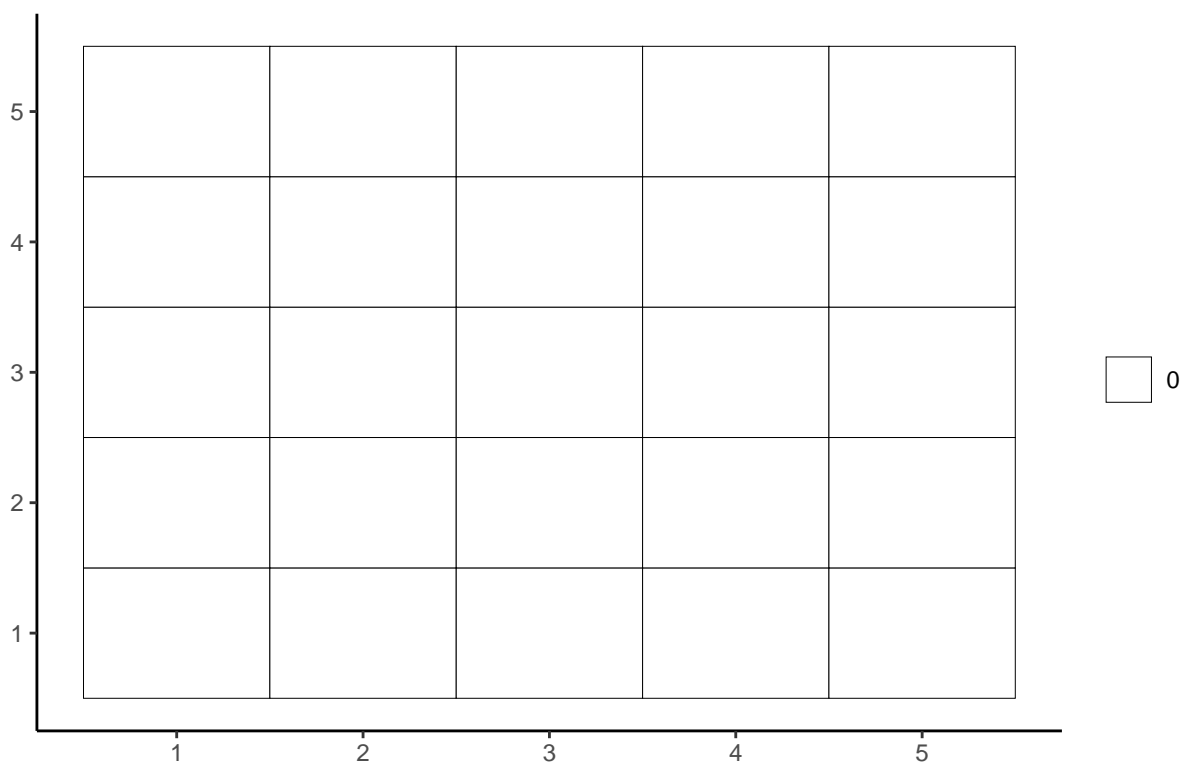
```
##
##
##
## TRIAL 5
##
## unif_prior
##           Covariate Dependent Graphical Model
##
## Model ELBO: -9639.22           Unique conditional dependence structures: 2
## n: 180, variables: 5           Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 2.345 mins
## [[1]]
```

Graph 1, Individuals 1,...,21



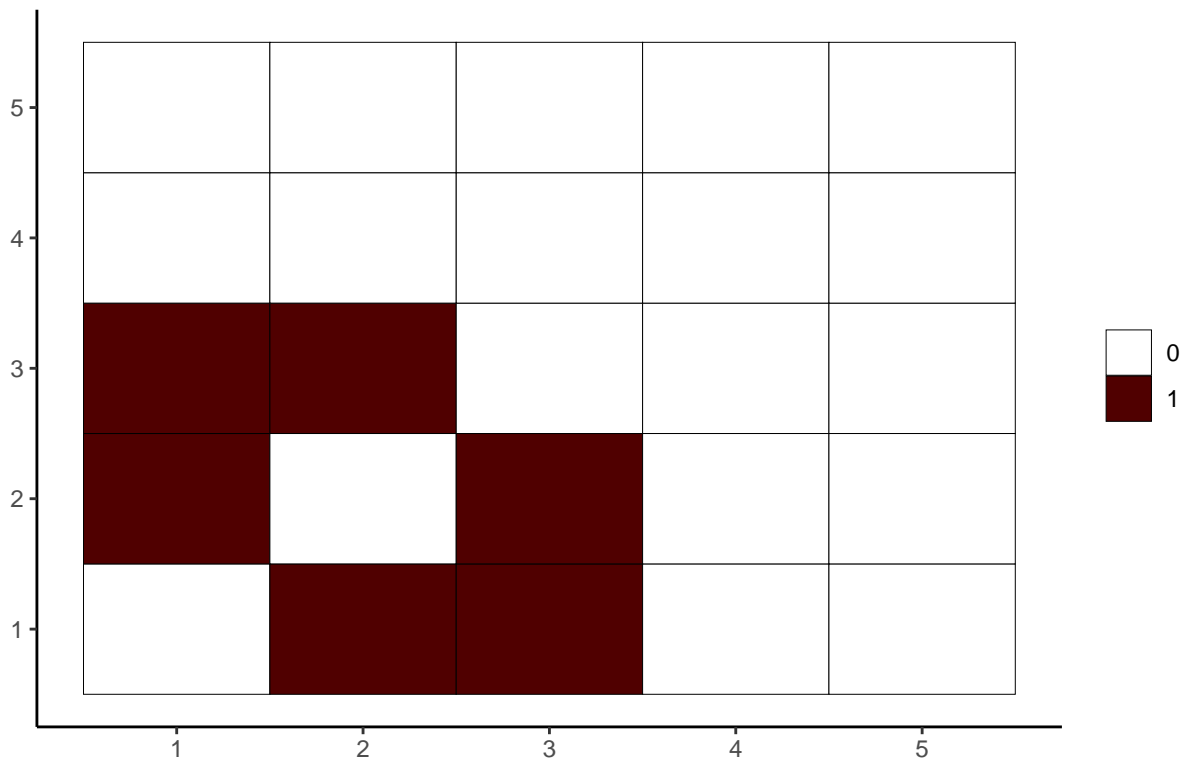
```
##  
## [[2]]
```

Graph 2, Individuals 22,...,180



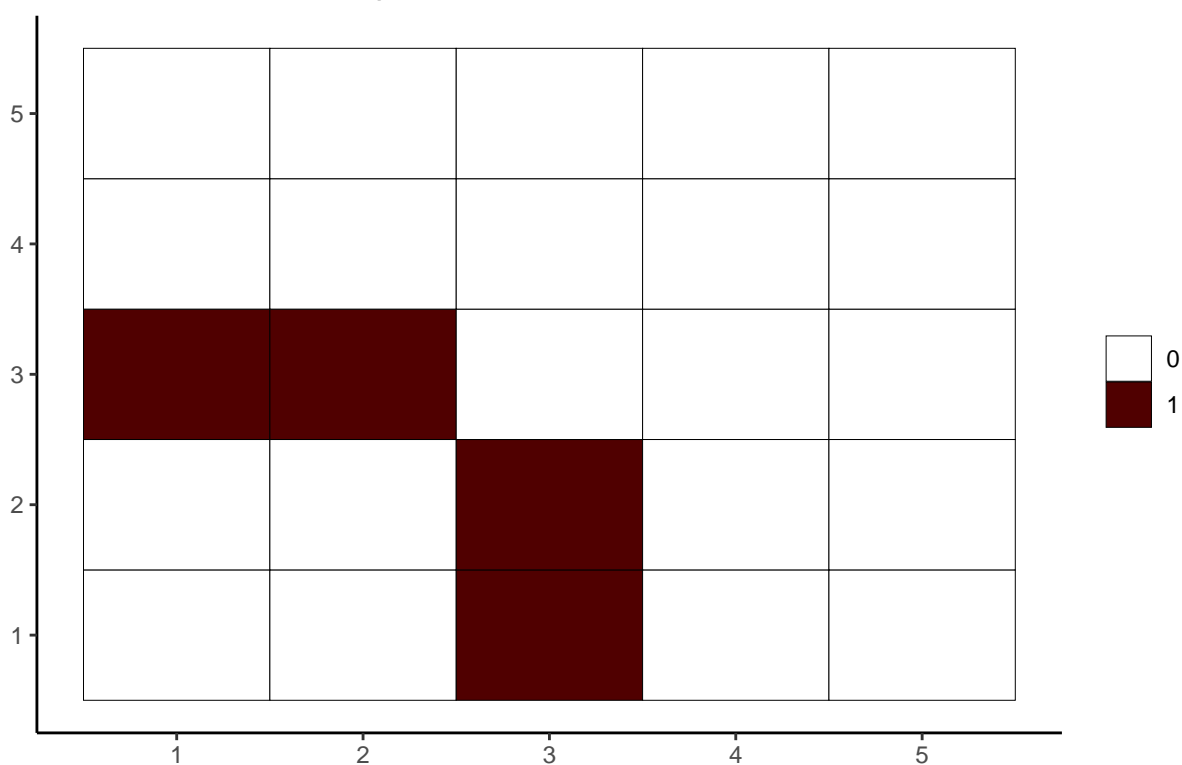
```
##
##
##
## TRIAL 5
##
## grid
##          Covariate Dependent Graphical Model
##
## Model ELBO: -9286.99          Unique conditional dependence structures: 2
## n: 180, variables: 5          Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 51.207 secs
## [[1]]
```

Graph 1, Individuals 1,...,63



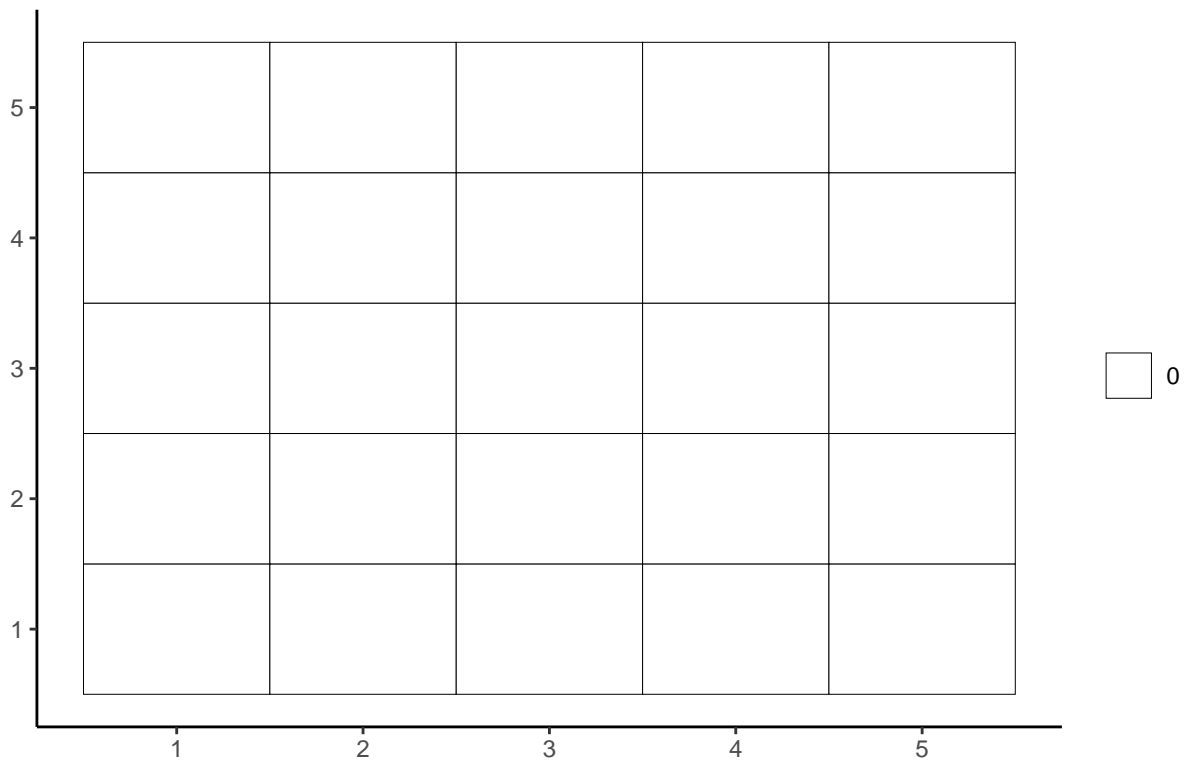
```
##  
## [[2]]
```

Graph 2, Individuals 64,...,180



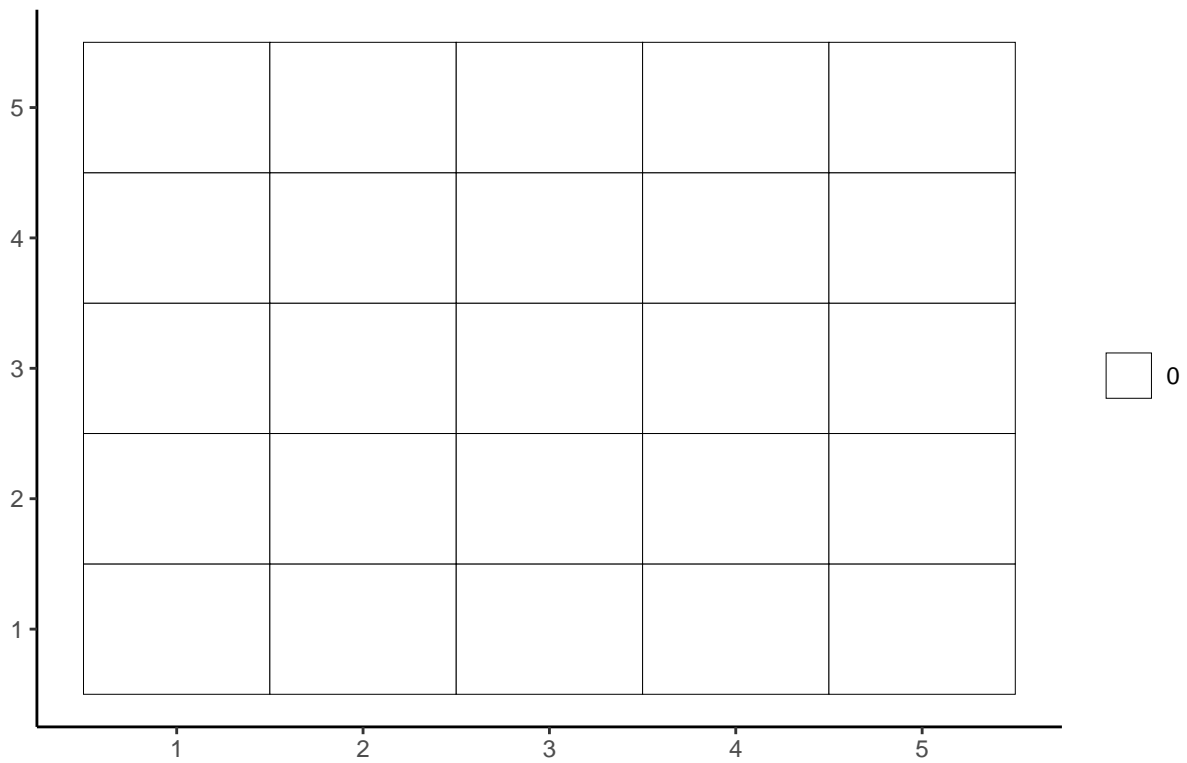
```
##
## -----
## TRIAL 6
##
## prior
##           Covariate Dependent Graphical Model
##
## Model ELBO: -9806.11           Unique conditional dependence structures: 1
## n: 180, variables: 5           Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 2.237 mins
## [[1]]
```


Graph 1, Individuals 1,...,180



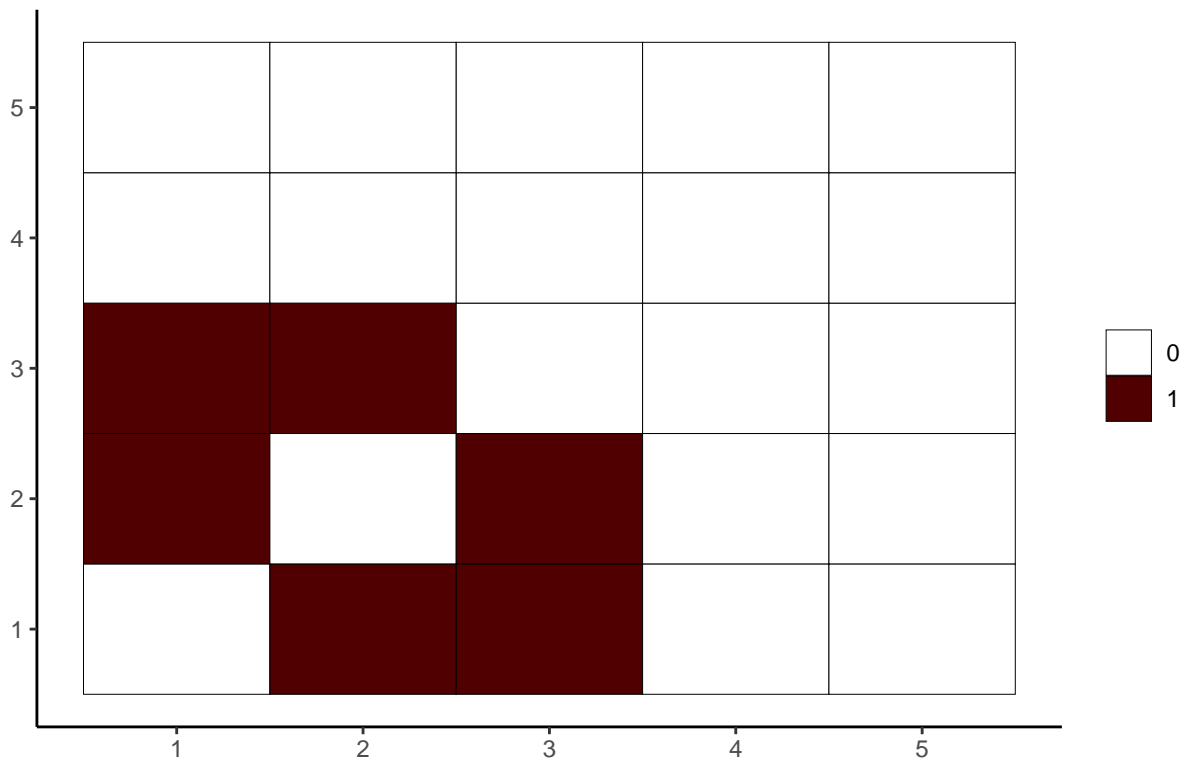
```
##
##
##
## TRIAL 6
##
## unif_prior
##          Covariate Dependent Graphical Model
##
## Model ELBO: -9870.69          Unique conditional dependence structures: 1
## n: 180, variables: 5          Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 2.218 mins
## [[1]]
```

Graph 1, Individuals 1,...,180



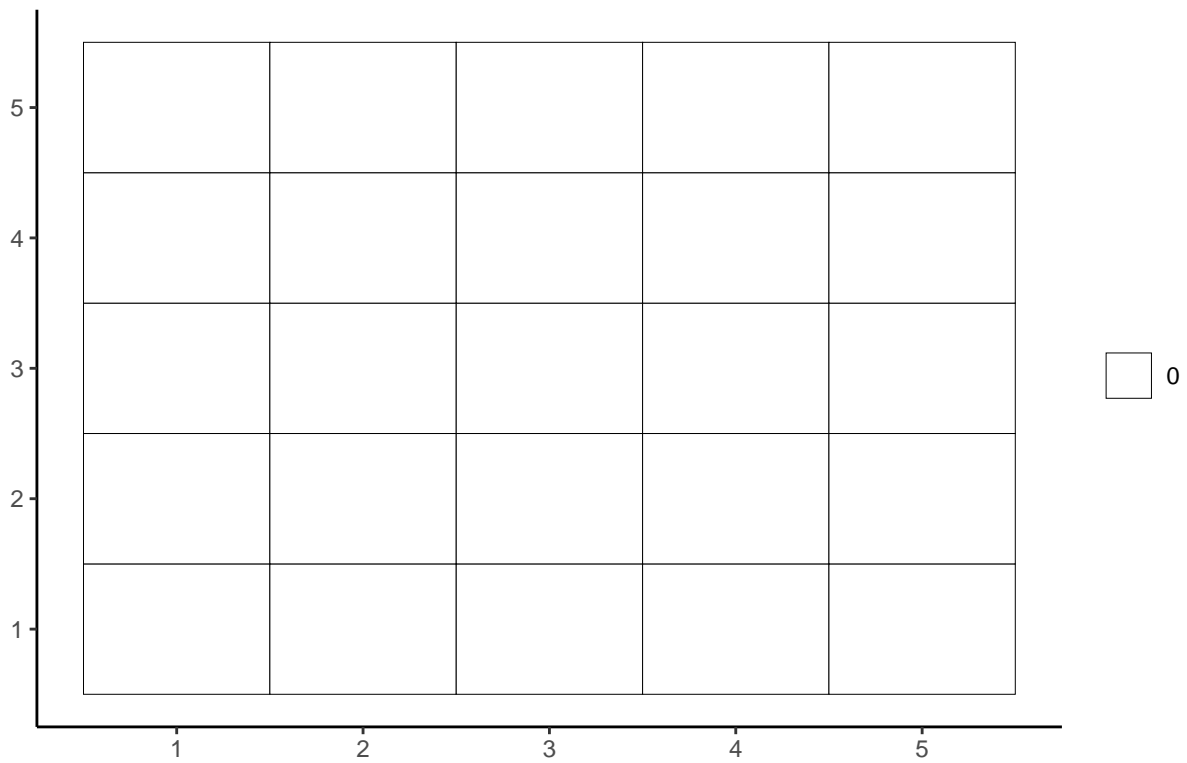
```
##
##
##
## TRIAL 6
##
## grid
##          Covariate Dependent Graphical Model
##
## Model ELBO: -9568.07          Unique conditional dependence structures: 1
## n: 180, variables: 5          Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 48.941 secs
## [[1]]
```

Graph 1, Individuals 1,...,180



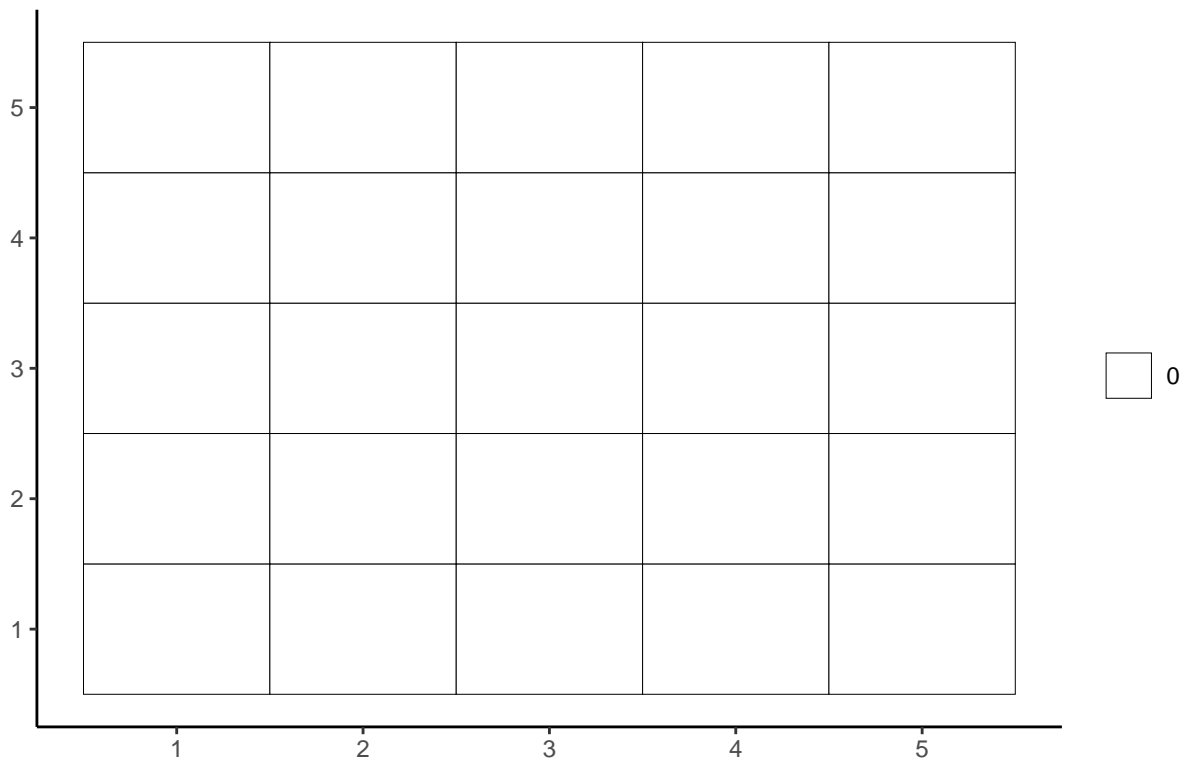
```
##
## -----
## TRIAL 7
##
## prior
##           Covariate Dependent Graphical Model
##
## Model ELBO: -9779.71           Unique conditional dependence structures: 1
## n: 180, variables: 5           Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 2.15 mins
## [[1]]
```

Graph 1, Individuals 1,...,180



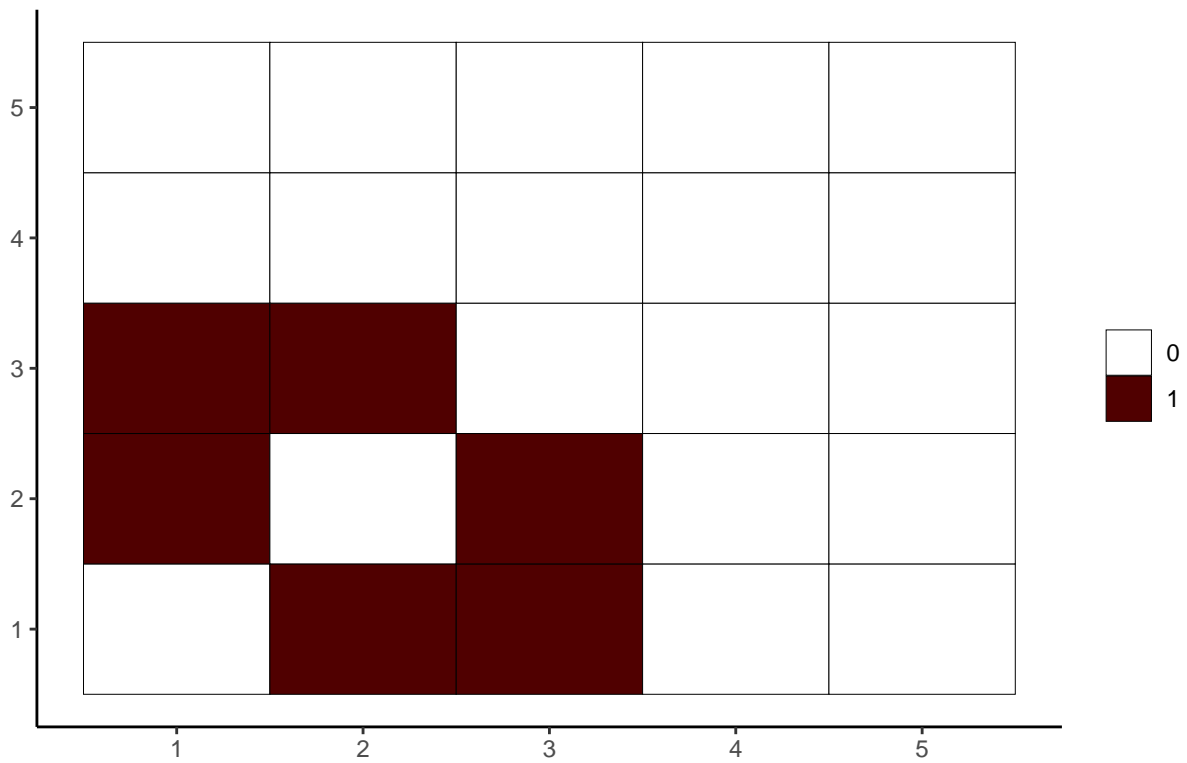
```
##
##
##
## TRIAL 7
##
## unif_prior
##          Covariate Dependent Graphical Model
##
## Model ELBO: -9840.91          Unique conditional dependence structures: 1
## n: 180, variables: 5          Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 2.104 mins
## [[1]]
```

Graph 1, Individuals 1,...,180



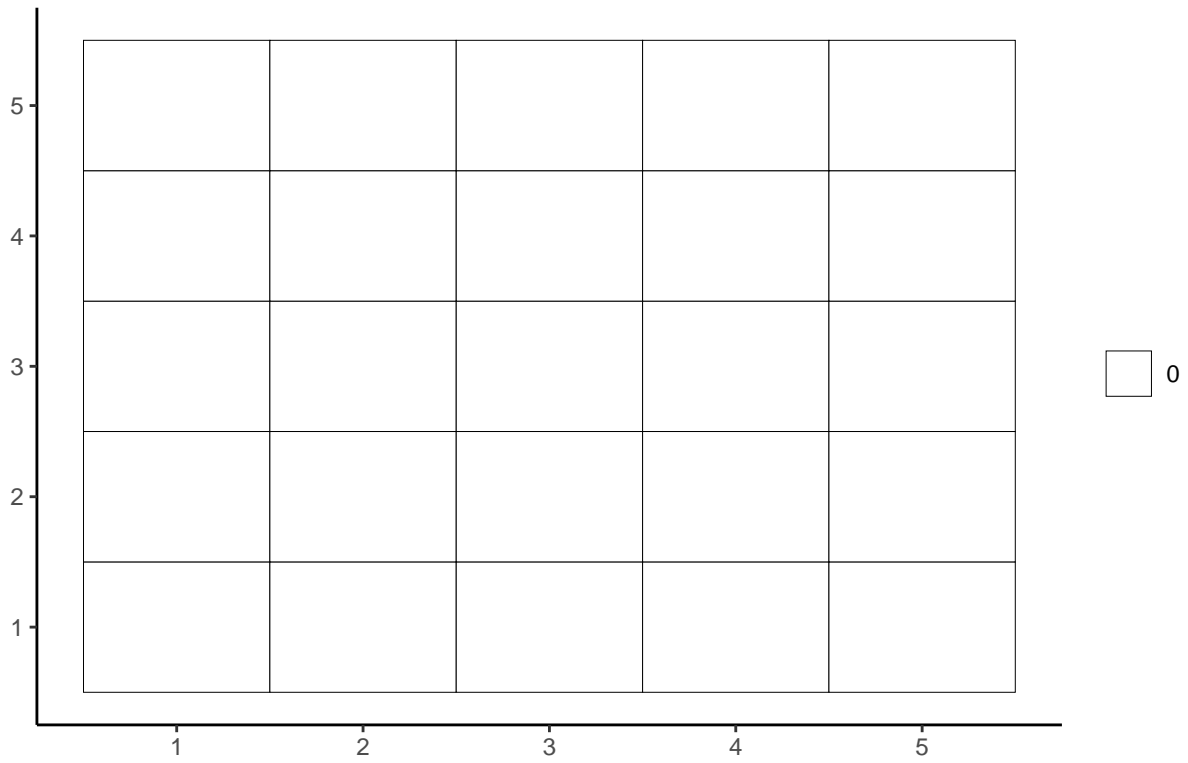
```
##
##
##
## TRIAL 7
##
## grid
##          Covariate Dependent Graphical Model
##
## Model ELBO: -9631.19          Unique conditional dependence structures: 1
## n: 180, variables: 5          Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 45.51 secs
## [[1]]
```

Graph 1, Individuals 1,...,180



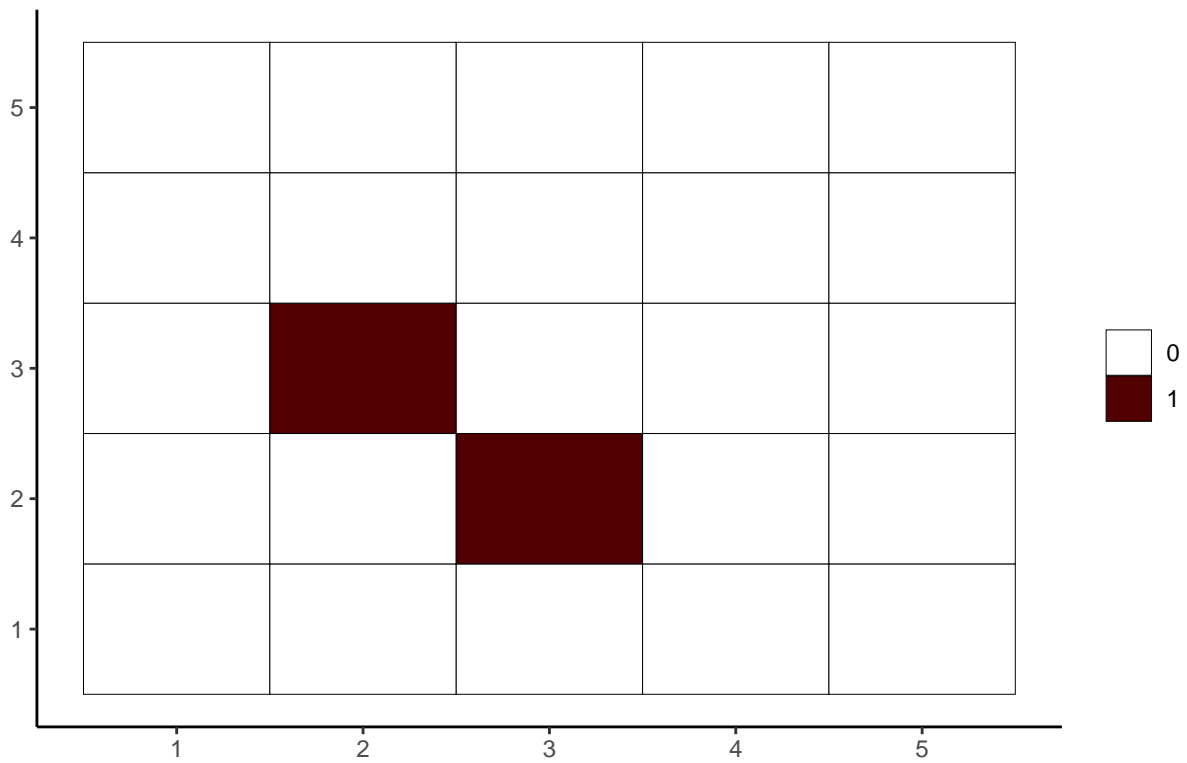
```
##
## -----
## TRIAL 8
##
## prior
##           Covariate Dependent Graphical Model
##
## Model ELBO: -9642.58           Unique conditional dependence structures: 2
## n: 180, variables: 5           Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 2.179 mins
## [[1]]
```

Graph 1, Individuals 1,...,10,22,...,120



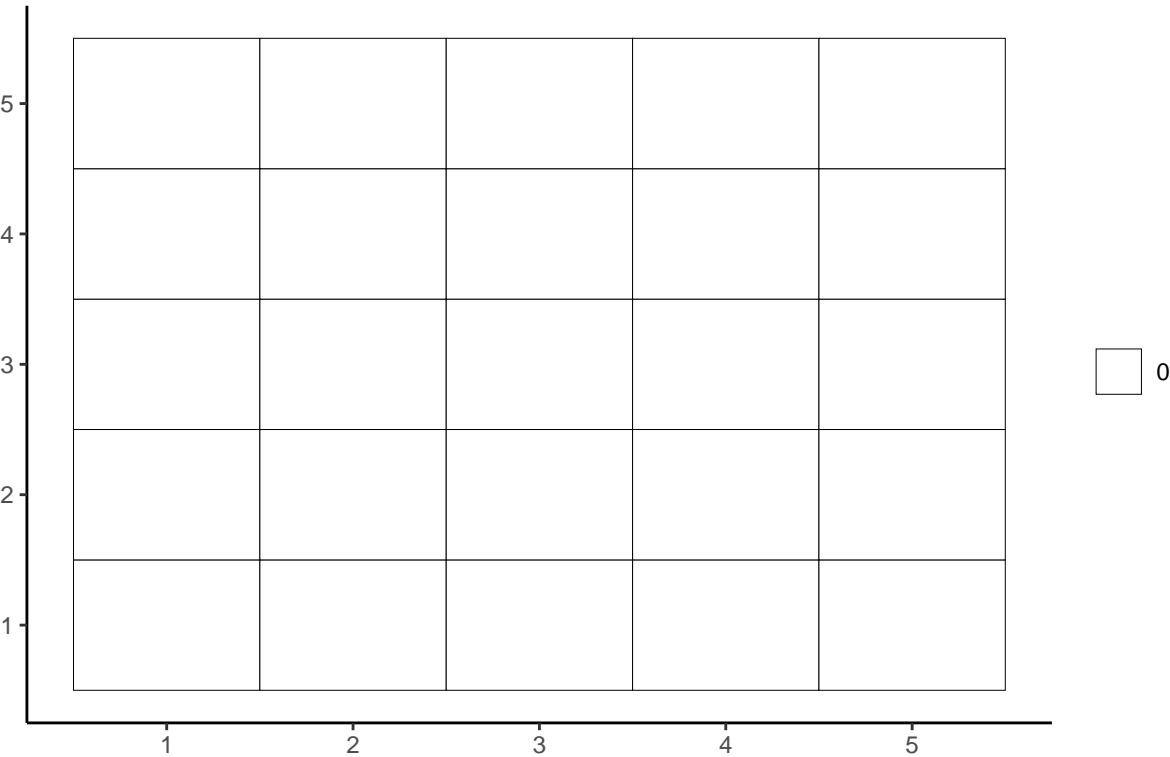
```
##  
## [[2]]
```

Graph 2, Individuals 11,...,21,121,...,180



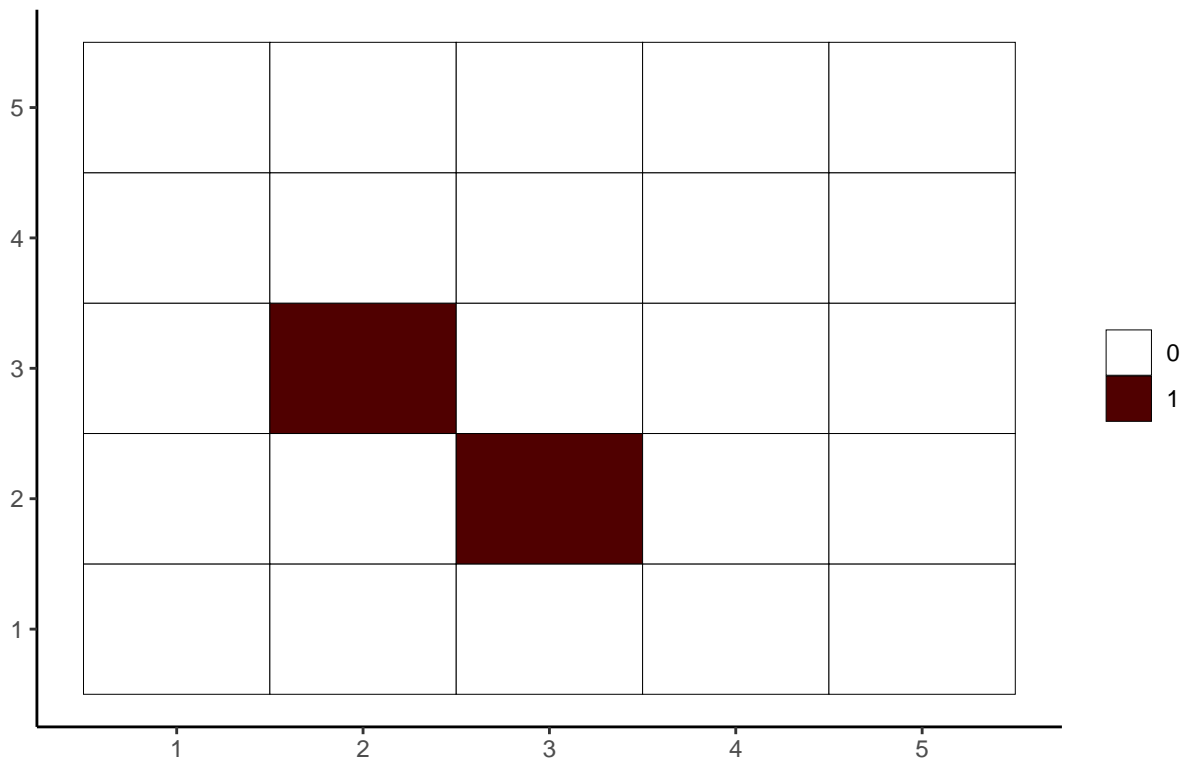
```
##
##
##
## TRIAL 8
##
## unif_prior
##           Covariate Dependent Graphical Model
##
## Model ELBO: -9714.68           Unique conditional dependence structures: 2
## n: 180, variables: 5           Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 2.177 mins
## [[1]]
```


Graph 1, Individuals 1,...,122



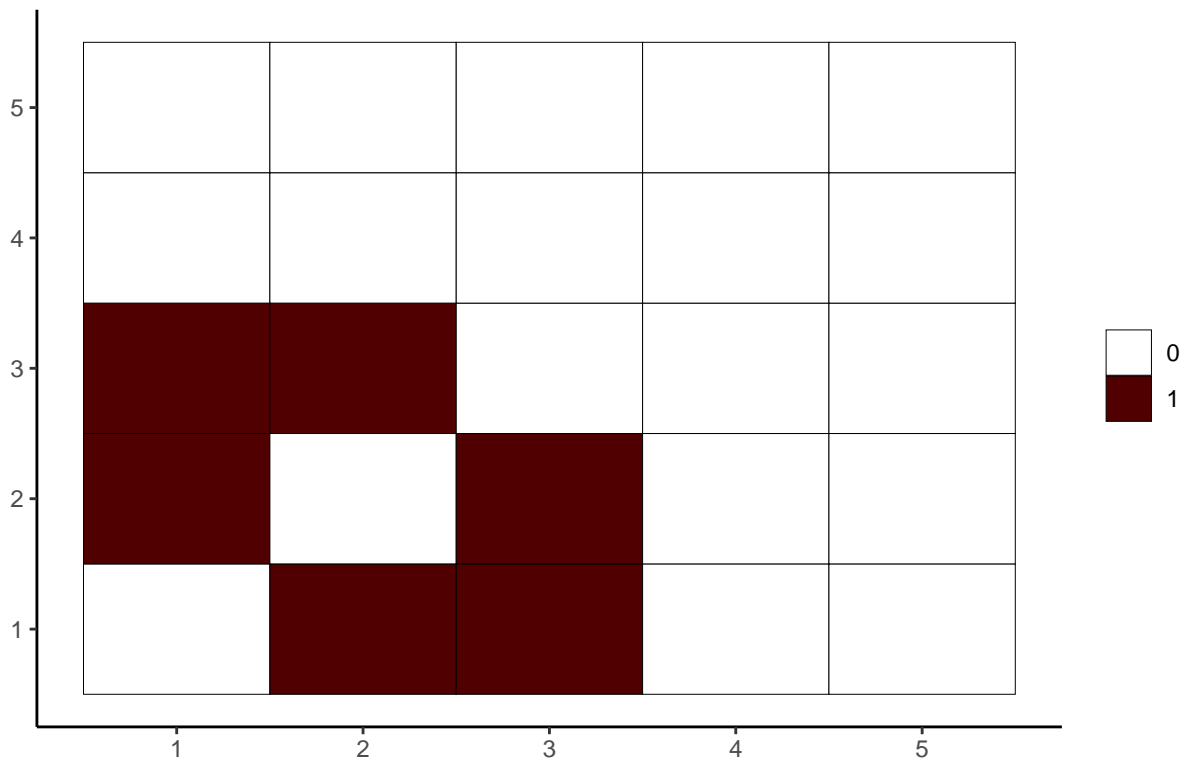
[[2]]

Graph 2, Individuals 123,...,180



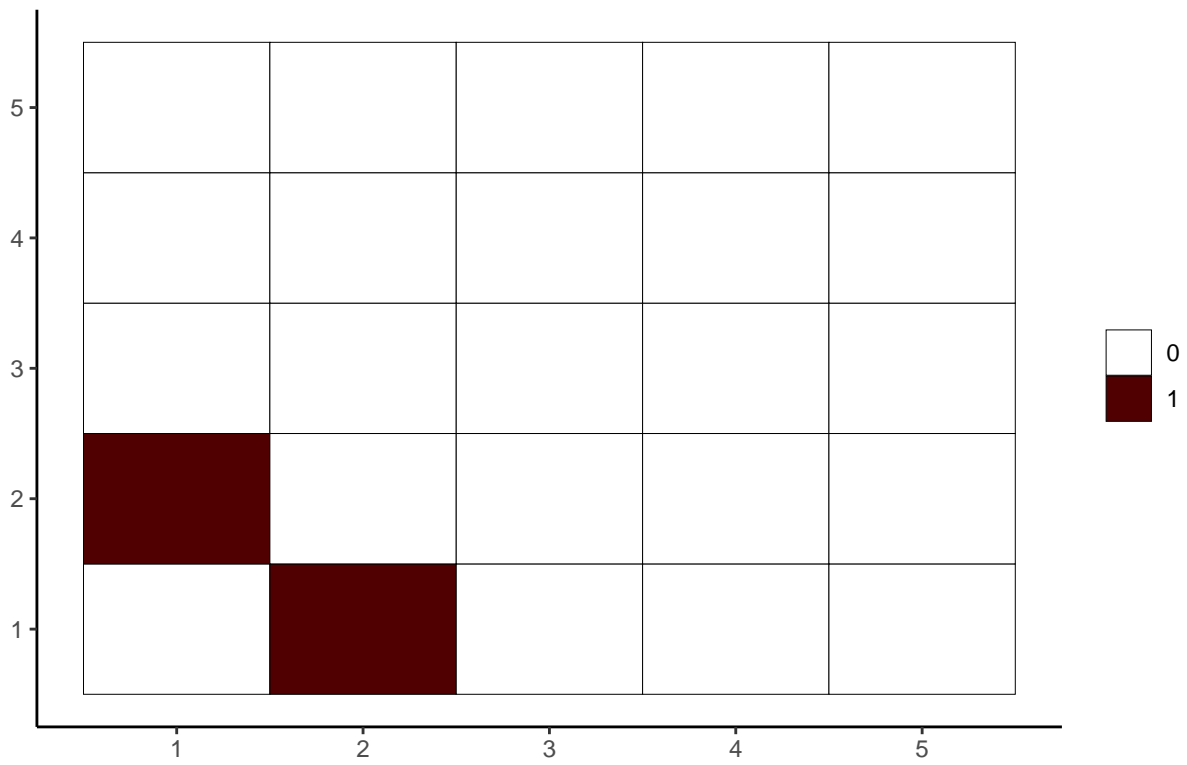
```
##
##
##
## TRIAL 8
##
## grid
##          Covariate Dependent Graphical Model
##
## Model ELBO: -9296.41          Unique conditional dependence structures: 1
## n: 180, variables: 5          Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 1.096 mins
## [[1]]
```

Graph 1, Individuals 1,...,180



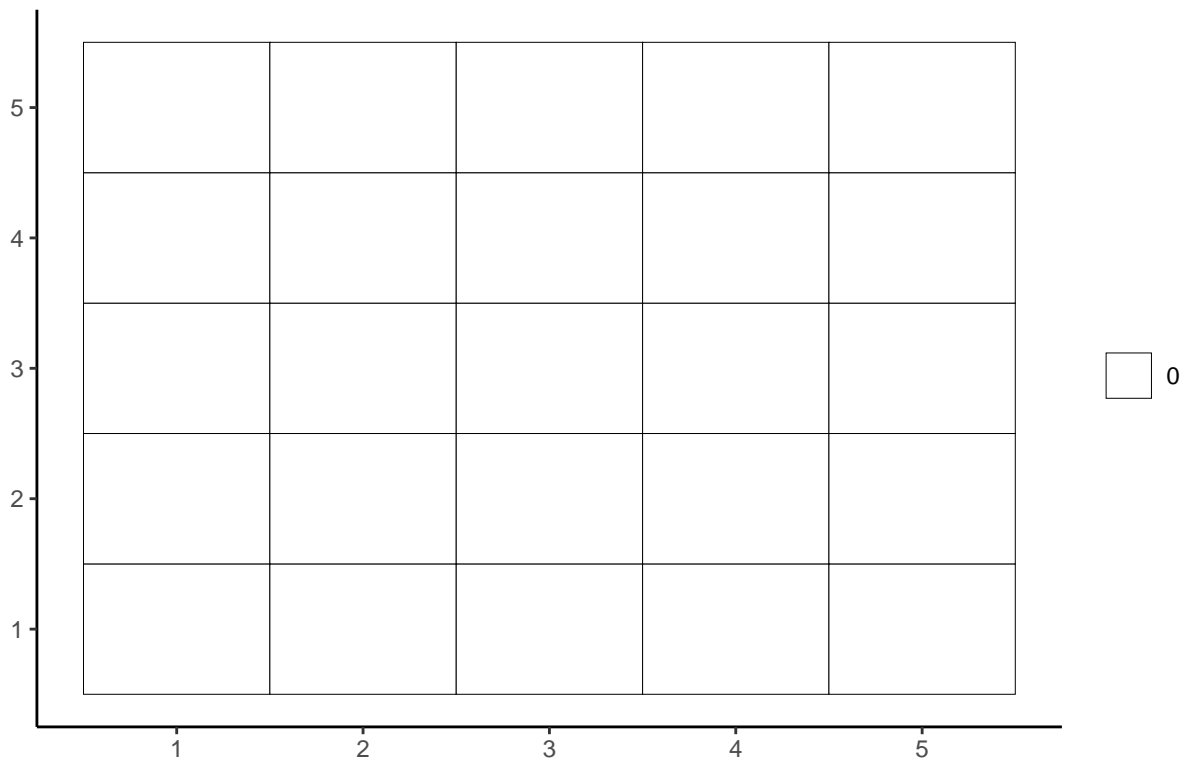
```
##
## -----
## TRIAL 9
##
## prior
##           Covariate Dependent Graphical Model
##
## Model ELBO: -9675.22           Unique conditional dependence structures: 2
## n: 180, variables: 5           Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 1.902 mins
## [[1]]
```

Graph 1, Individuals 1,...,48



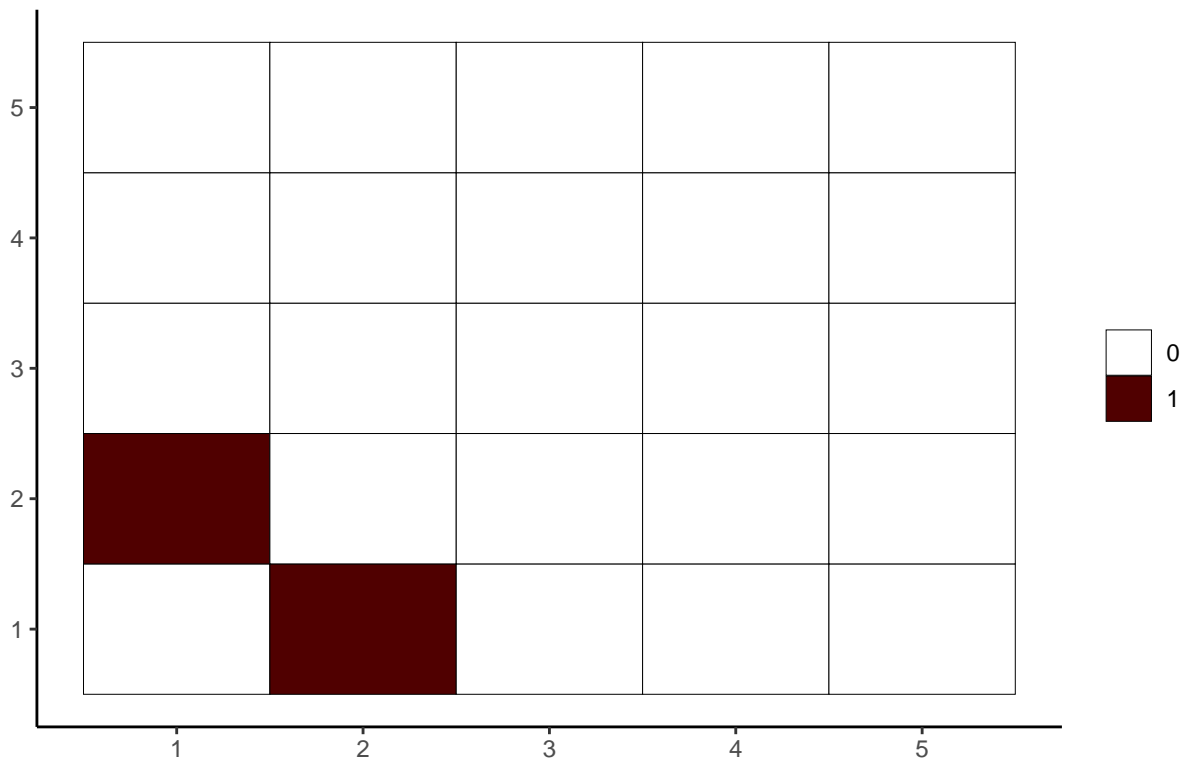
[[2]]

Graph 2, Individuals 49,...,180



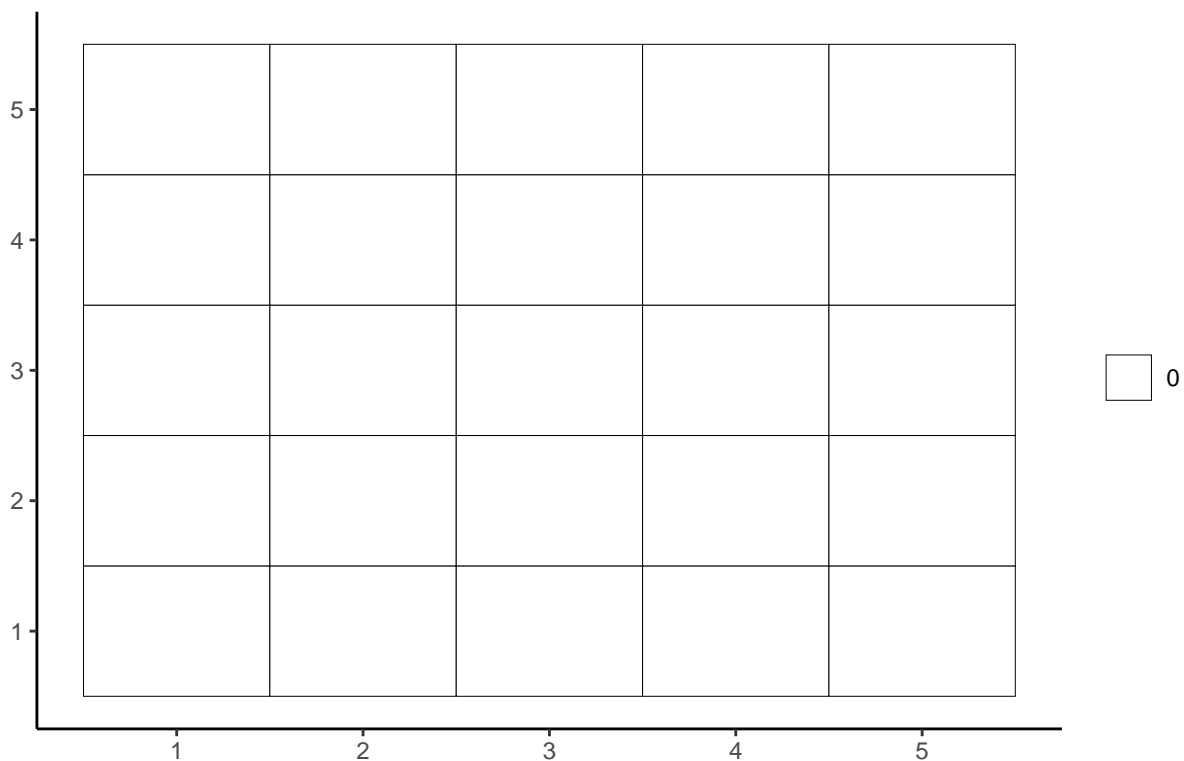
```
##
##
##
## TRIAL 9
##
## unif_prior
##           Covariate Dependent Graphical Model
##
## Model ELBO: -9733.31           Unique conditional dependence structures: 2
## n: 180, variables: 5           Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 1.879 mins
## [[1]]
```

Graph 1, Individuals 1,...,46



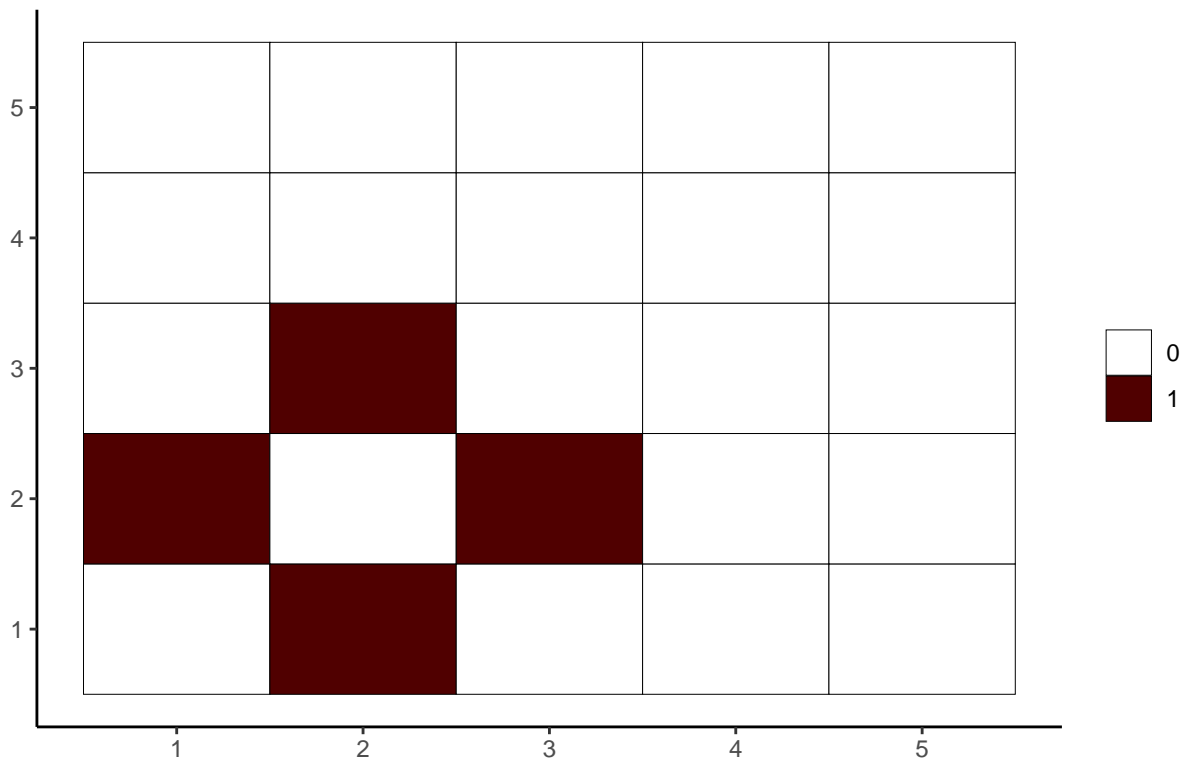
```
##  
## [[2]]
```

Graph 2, Individuals 47,...,180



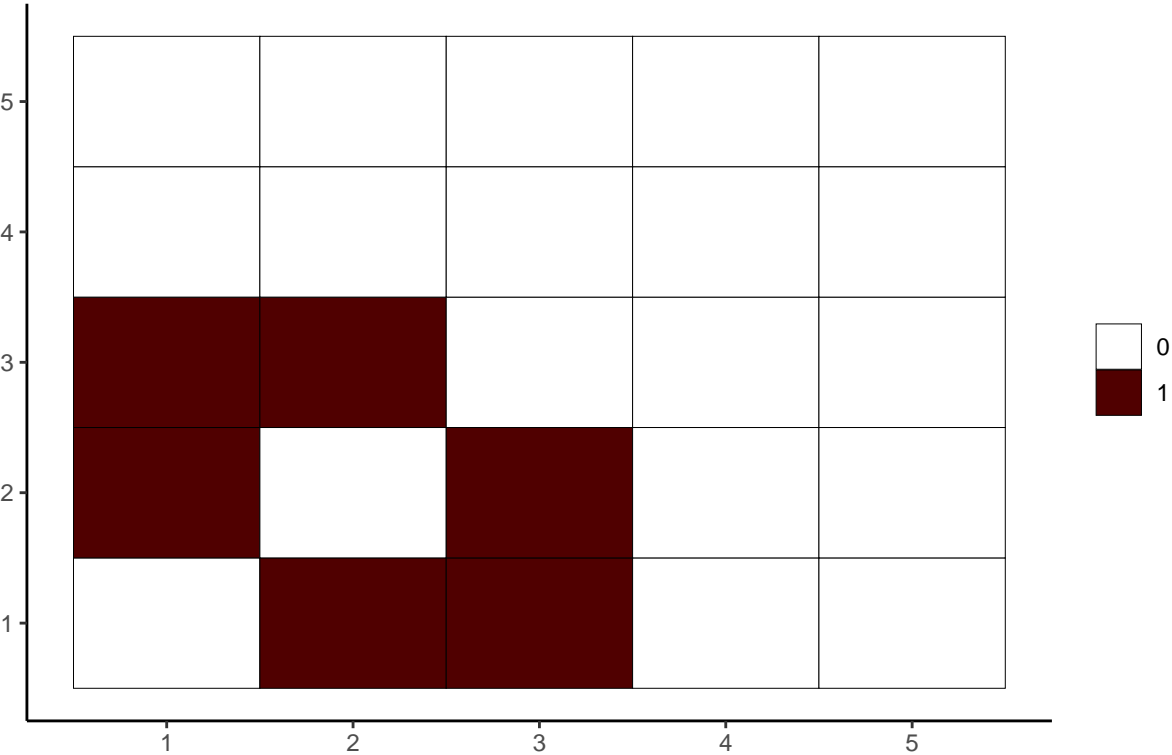
```
##
##
##
## TRIAL 9
##
## grid
##          Covariate Dependent Graphical Model
##
## Model ELBO: -9540.6          Unique conditional dependence structures: 3
## n: 180, variables: 5          Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 1.088 mins
## [[1]]
```

Graph 1, Individuals 1,...,71



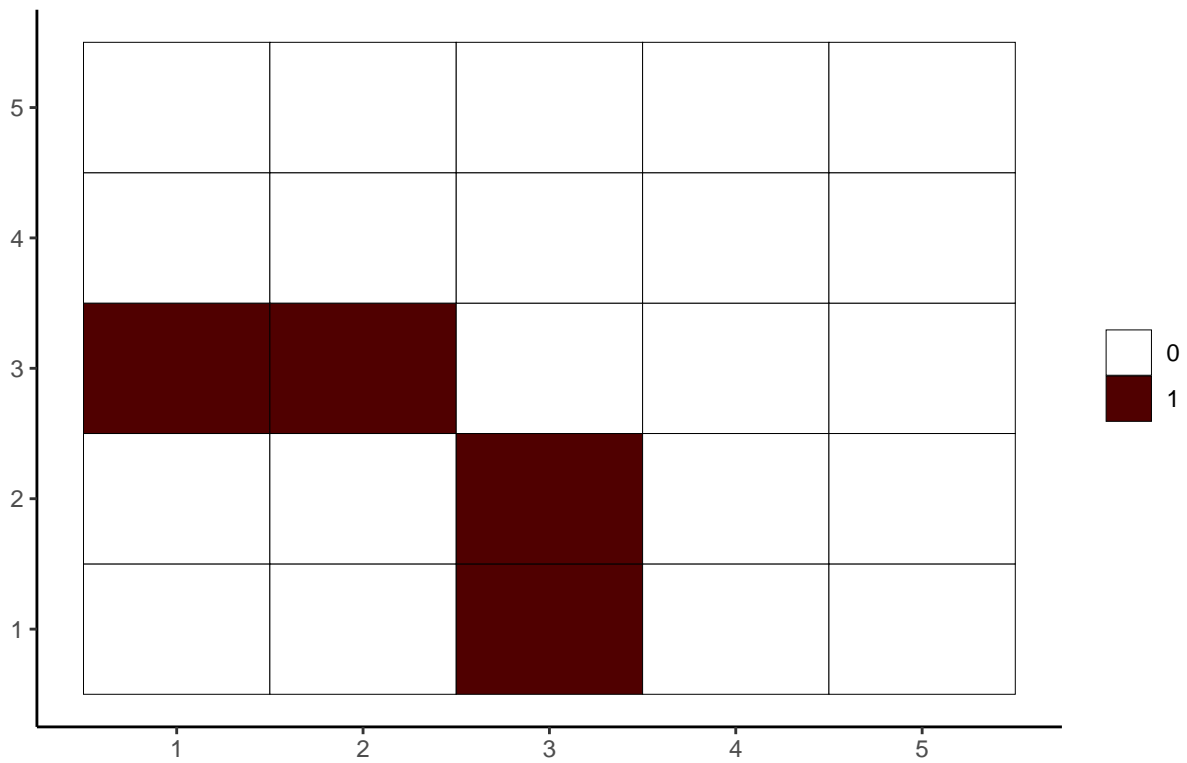
[[2]]

Graph 2, Individuals 72,...,110,141,...,180



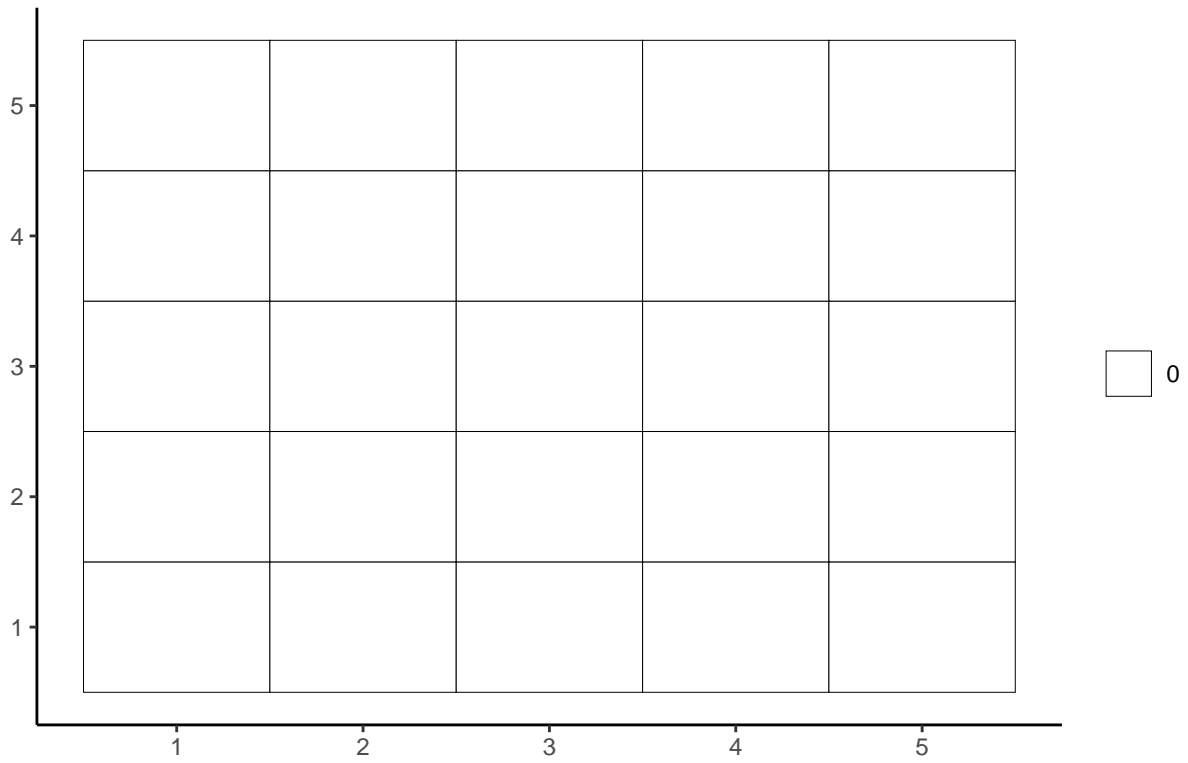
```
##  
## [[3]]
```

Graph 3, Individuals 111,...,140



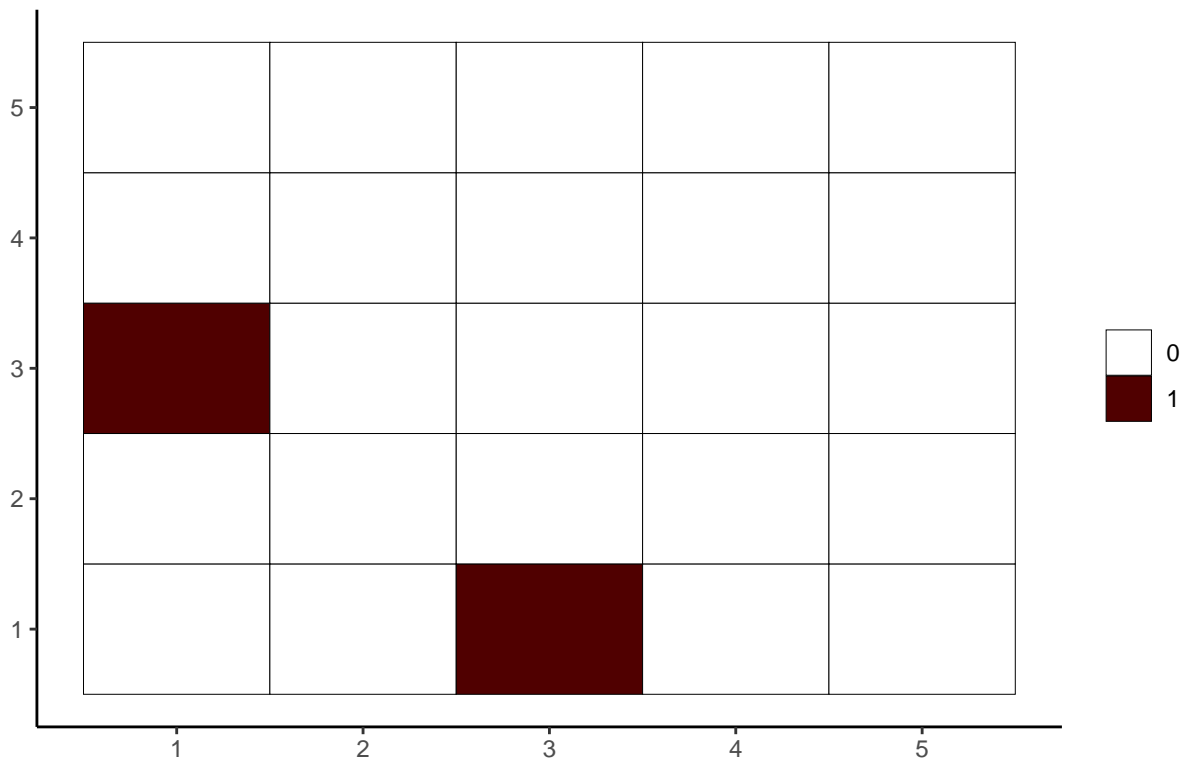
```
##
## -----
## TRIAL 10
##
## prior
##           Covariate Dependent Graphical Model
##
## Model ELBO: -9725.71           Unique conditional dependence structures: 2
## n: 180, variables: 5           Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 1.963 mins
## [[1]]
```

Graph 1, Individuals 1,...,165,179,180



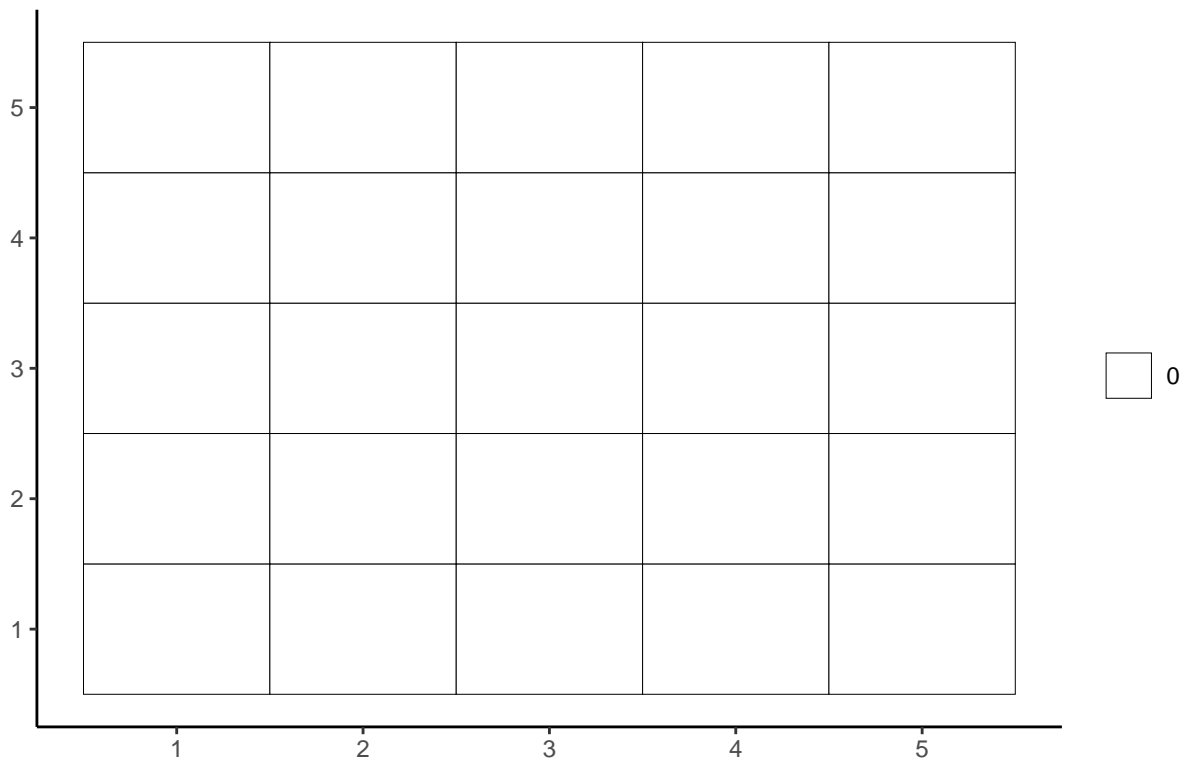
[[2]]

Graph 2, Individuals 166,...,178



```
##
##
##
## TRIAL 10
##
## unif_prior
##          Covariate Dependent Graphical Model
##
## Model ELBO: -9794.72          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]]
```

Graph 1, Individuals 1,...,180



```
##
##
##
## TRIAL 10
##
## grid
##          Covariate Dependent Graphical Model
##
## Model ELBO: -9386.55          Unique conditional dependence structures: 1
## n: 180, variables: 5          Hyperparameter grid size: 1000 points
## CAVI converged for 5/5 variables
##
## Model fit completed in 45.4 secs
## [[1]]
```

Graph 1, Individuals 1,...,180

