**Assignment No. 13.2**

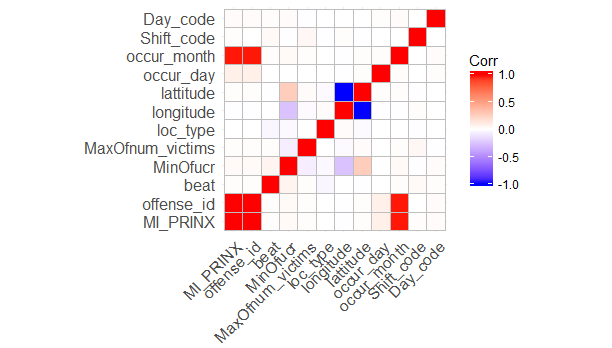
1. Visualize the correlation between all variables in a meaningful and clear way of representing. Find out top 3 reasons for having more crime in a city.

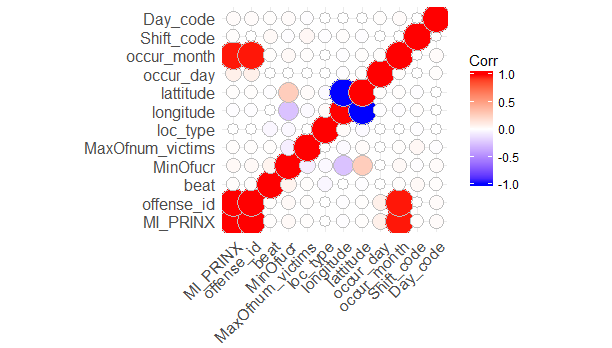
> library(ggcorrplot)

> crimecor <- round(cor(numcrime, use= "complete.obs"),3)

> ggcorrplot(crimecor)

> ggcorrplot(crimecor, method = "circle")





1. What is the difference between co-variance and correlation? Take an example from this dataset and show the differences if any?

Covariance is a measure of co-variability of two variables, and correlation is a measure of degree of linear association and strength of relatedness, and standardised co-movement.

Covariance is measured in a squared term, and usually has the units of measurement, that is same as that of the variables’ whose covariance is measured. The value of covariance can lie anything between –infinity and + infinity.

Correlation does not have a unit of measurement, since it is a standardised measure of covariance itself, and hence it has a value between +1 and -1

When the way in which the variable is measured changes, the covariance is automatically affected. But the correlation is not.

> colnames(numcrime)

[1] "offense\_id" "longitude" "lattitude"

> cov(numcrime)

offense\_id longitude lattitude

offense\_id 8.213043e+19 -3.497858e+08 1.321459e+08

longitude -3.497858e+08 5.966877e+01 -2.386342e+01

lattitude 1.321459e+08 -2.386342e+01 9.546876e+00

> cor(numcrime)

offense\_id longitude lattitude

offense\_id 1.000000000 -0.004996624 0.004719228

longitude -0.004996624 1.000000000 -0.999835504

lattitude 0.004719228 -0.999835504 1.000000000