Guidelines for creating a regression matrix

- List the covariate layers:

```
files <- list.files(path =
'./Data/covariates/covs1km/wgs84/stack1',
pattern="tif$", full.names=TRUE)
```

  Adjust file path to the location of the covariate stacks on your computer.

  Either use covariates in WGS84 or UTM (for UTM use 'utm' in the file path). If you would like to model in UTM then we suggest to use the covariates prepared in UTM and not re-project the WGS84 rasters in R.

  Choose stack folder:
  - stack1: stack of 59 covariates (use to try out random forest)
  - stack2: stack of 171 covariates
  - stack3: stack of 31 covariates (advised to use this for linear regression)

- We advise to first use the covariate layers in these stacks to make your code work. If you are able to fit a model and make spatial predictions. Then the next step can be to add your own data.

- Read stack into R:

```
r <- stack(files)
```

  Make sure the 'raster' package is loaded. This function works if all covariate layers have the same extent, resolution, projection. It should work on the stacks prepared for you. You can add your own layers to the stack.

- Load the prediction mask:

```
m <-
raster('./Data/covariates/covs1km/wgs84/mask/mask.tif')
```

Do not forget to adjust the file path (e.g. change to utm).

- Mask the stack:

```
r2 <- mask(r1, m)
```

- Convert masked stack to SpatialGridDataFrame

```
stack <- as(r2, "SpatialGridDataFrame")
```

- Check the categorical LCEE covariate (see data preparation
  tutorial) in your dataset. (Or remove the covariate from the stack
  for now). Check if the LCEE covariate is stored as a factor
  variable. If not, convert to factor.

- Convert sample data to a SpatialPointsDataFrame and overlay
  with the stack. Append extracted covariate values to dataset.

```
coordinates(d) <- ~x+y
dum <- over(x = d, y = stack)
d <- cbind(d, dum)
```

- Check the categorical LCEE covariate (see data preparation
  tutorial) in your dataset.