# Day 3 – Hybrid Machine Learning + Geostatistics

# Today's Goal

Integrate:

- Machine learning
- Spatial geostatistics
- Hybrid prediction modelling

## Load Data

```
library(readr)
hybrid_ml_geo <- read_csv("../data/hybrid_ml_geostatistical_data.csv")
head(hybrid_ml_geo)
```

```
# A tibble: 6 x 11
     id    utm_x    utm_y n_tested n_pos prevalence_true ml_signal_true
  <dbl>    <dbl>    <dbl>    <dbl> <dbl>           <dbl>          <dbl>
1     1  387215. 1485399.       59     0     0.000334            -7.06
2     2 1046370.  758144.      180     8     0.0496              -1.66
3     3  681525. 1102741.      164    22     0.167               -0.980
4     4  713355. 1127489.       62     0     0.00476             -4.17
5     5  276748. 1384709.       84     0     0.00000423         -11.6
6     6 -112576. 1315376.       60     0     0.000000669        -12.5
# i 4 more variables: S_residual_true <dbl>, elevation <dbl>, ndvi <dbl>,
```