# The University of British Columbia

## Master of Data Science

*Data Science 583*

Project

Due: 11 April 2019

Analyze the seismic timing data (located on the github site) using the `gam()` function.

The $z$ variable of this data set corresponds to seismic timings measured by geophones dropped into ditches dug along transects following the $(x, y)$ coordinates. The timings are related to depth of a particular substratum. The shape of this surface is of importance in oil exploration. This particular substratum represents an ancient riverbed (river bottom) in central Alberta.

Possible models to consider can be constructed using multiple linear regression, bivariate spline regression with equally spaced knots, thin-plate splines, and generalized additive models with normal and gamma families.

Write a short report (no more than 3 pages, including figures) which briefly summarizes each of the models you considered as well as your preferred model. For that model, construct a contour map that could be used by a geologist, and to give a short explanation as to why your method of modelling is better than the alternatives.