Notes and Objective:

- The dataset has been divided (by the organisers) into two different excel sheets
  - **Training dataset**
  - **Testing dataset**
    - Note: Some samples in the training dataset may contain missing values for one or two traits.
- The **Testing dataset** contains 69 spectra.
- Your task is to quantify for each of the spectra in the **"Testing Dataset"** the following;

  - **Kappa casein**
  - **Casein micelle size**
  - **pH.**

Sample Collection:

622 milk samples from 622 cows were collected between August 2013 and August 2014 from 7 different Irish research herds. The samples originated from Holstein-Friesian, Jersey and Norwegian Red cows, as well as their crosses; all cows were fed a predominantly grass-based diet with occasional concentrate and grass silage supplementation. The samples were collected during morning and evening milking and represented different stages of lactation and different parities. All samples were analyzed by the same MilkoScan FT6000 (Foss Electronic A/S, Hillerød, Denmark) producing 1,060 transmittance data points in the mid-infrared light region. Milk pH, casein micelle size (CMS) and kappa casein are provided for the analyses. Milk pH of all samples was assessed with a SevenCompact pH meter S220 (Mettler Toledo AG, Greifensee, Switzerland). The casein micelle hydrodynamic diameter was determined using a Zetasizer Nano system (Malvern Instruments Inc., Worcester, UK). Milk proteins were determined using reverse-phase high performance liquid chromatography (HPLC) using an adaptation of the method of Visser et al. (1991) and are expressed as grams per liter of milk. The data contains 1,060 transmittance values and no records belonging to the same animal are present.

If you have any queries please contact either of the following;

alessandro.casa@ucd.ie **or** maria.frizzarin@ucdconnect.ie