Methods 3, Week 4:

Multilevel Modelling

Maxime Sainte-Marie

Theoretical foundations of Bayesian hierarchical modelling:

1. Bayesian
2. Regression
3. Hierarchy/Structure
4. Modelling

Models

"All models are wrong, but some are useful" (George Box)

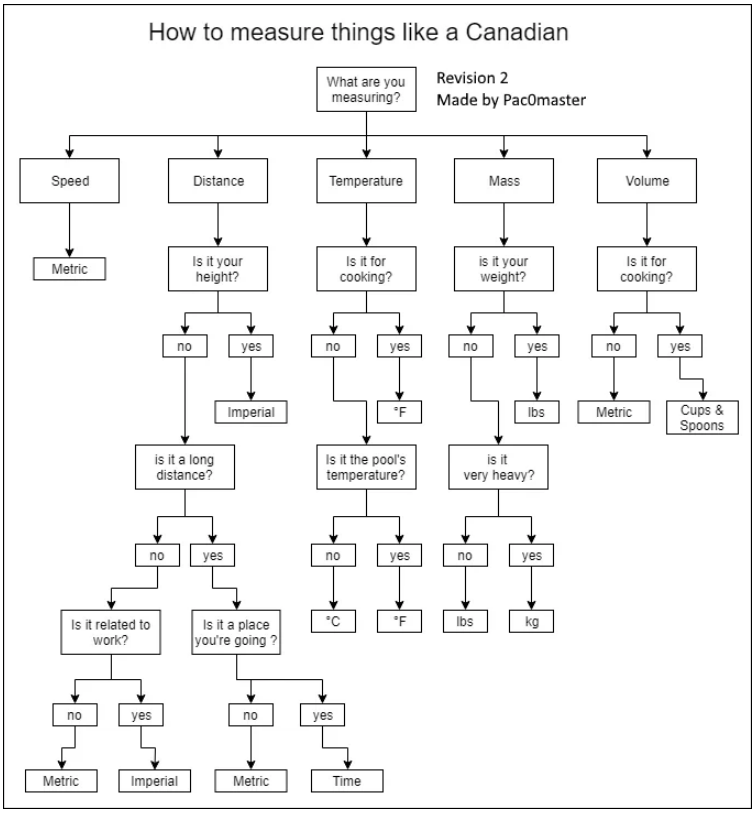
What is Modelling?

```
## PhantomJS not found. You can install it with webshot::install_phantomjs(). If it is installed, pleas
```

Modelling as reduction (Lévi-Strauss, *La pensée sauvage*)

What virtue is there in reduction of scale or properties? - A model "compensates for the loss of sensory dimensions by the acquisition of intelligible dimensions" - "Reversal of the process of knowing" - Reverse epistemic engineering - "With a scale model the knowledge of the whole precedes that of its parts" - Artificiality, "human-made-ness" - "Actual experiment" on what is being modeled. - Brings a supplementary dimension to our understanding of what it represents - Acquisition of procedural knowledge

Modelling the Canadian Measurement System



Bayesian Inference

- Probabilistic modelling framework that updates beliefs or knowledge about parameters based on observed data.
  - Uses Bayes' theorem to estimate the posterior distribution of parameters, given prior beliefs and the likelihood of the observed data under the different values of each parameter

$posterior = \frac{prior \cdot likelihood}{normalizing\ constant} \propto prior \cdot likelihood$

- The apparent increase in complexity and subjectivity is simply an effect of you becoming aware of the modelling choices and their implications

Regression

- *Regression towards mediocrity in hereditary stature* (Francis Galton, 1886)
  - Some biological types were more stable than others and hence were resistant to evolutionary change
  - Phenomenon of bivariate distributions discovered through his studies of heritability.
  - The use of regression in Galton's sense does survive in the phrase regression to the mean
    - On average, extremes do not survive
    - The statistical explanation attributed to Galton appeared during the biometrician-mutationist debate in the early 1900s
- Statistical technique used to model the relationship between one or more independent variables (predictors) and a dependent variable (outcome)
  - Bayesian regression estimates the posterior distribution of regression coefficients.

Hierarchy/Structure

- Data is organized into multiple levels or groups
- The goal is to model how variables at different levels influence the outcome variable
  - Bayesian context: hierarchy of prior distributions for the parameters at each level of the model
- 3 approaches to group structured data
  1. Complete pooled models
  2. No pooled models
  3. Partially pooled models

Complete pooling

- Lumps all observations into one population or one "pool."
- Assumptions
  1. Each observation is independent from the others
  2. Information about groups is irrelevant, as a universal model is appropriate for all groups
- However:
  1. Though observations on one group may be independent of those on another, observations within a group are correlated.
  2. With respective to given pair of variables, groups can be inherently different
- Can produce misleading conclusions about the relationship of interest and its significance

No pooling

- Build a separate model for each group
- Assumptions
  - Groups do not contain relevant information bout each other
- Drawbacks
  1. No generalization or application to groups outside sample.
  2. Underutilizes the data and thus ignores potentially valuable information
     - This is especially consequential when we have a small number of observations per group.

Partial pooling

- Provides a middle ground
  - Though each group might have its own model, one group can provide valuable information about another
- Information sharing across levels
  - Useful when you have limited data at certain levels
- Allows for the modelling of two important features
  1. Within group variability
     - Variability among multiple observations within each group can be interesting on its own
  2. Between group variability
     - Allows us to examine the variability from group to group.

Random effects

- Allows to estimate group-specific parameters while borrowing information from the overall dataset to improve the stability and reliability of those estimates
  - Especially for groups with limited data.
- Shrinkage
  - We adjust each group level data with what we learn from the other participants, pulling them towards the mean

– The more outlying the group, the more the shrinkage
– The fewer datapoints in the group, the more the shrinkage