Multilevel Modelling Analysis Process with R:: **CHEAT SHEET**

Author: Dr. Kaiyi Wang

Step 1 : Research Question

- Identify multilevel propositions
  - o Lower level factors (e.g. Patients)
  - o Higher level factors (e.g. Hospitals)
  - o Interactions (e.g. Patients in Hospitals)
- Research question establishment

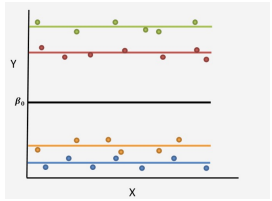
o Is the risk of developing Alzheimer's Disease really higher in rural areas?
o How does ethnicity and occupation Influence residential segregation in Sydney?

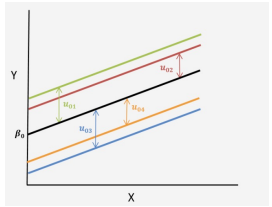
Step 2 : Exploratory Data Analysis

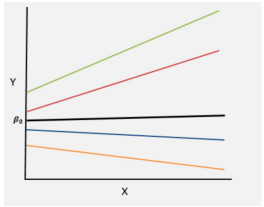
- Data structures
  - `summary(data)`
  - `dplyr::glimpse(data)`
- Distribution of the variables
  - `plot(y ~ x)`
  - `boxplot(y ~ x)`
  - `hist(x)`
- Categorical data relationship
  - `mosaic(~x + y, data)`
- Basic linear model
  - `lm(y ~ x, data)`
  - `summary(model)`
  - `predict(model)`

Step 3.1 : Multilevel Model Fitting Process

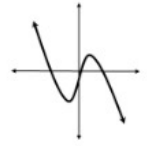
- **L**inear **M**ixed **E**ffects basic formula (**lme4** package)
  - `lmer(y~x1 + x2 + (1|level-2-grouping-var, data)`
- Model characteristics:
  - `summary(model), confint(model), fixef(model) VarCorr(model)`





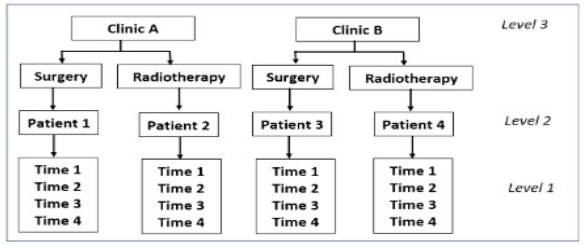


- **Model 0: non-hierarchical intercept-only model** $y_i = \beta_0 + e_i$
  - `lm(y ~ 1, data)`
- **Model 1: Empty random intercept model** $y_{ij} = \beta_0 + u_{0j} + e_{ij}$
  - `lmer(y ~ 1 + (1|groupvar), data)`
- **Model 2: Random intercept model (Fixed effects)** $y_{ij} = \beta_0 + \beta_1 x_{ij} + u_{0j} + e_{ij}$
  - `lmer(y ~ x + (1|groupvar), data)`
- **Model 3: Random slope and random intercept model (Random effects)** $y_{ij} = \beta_0 + \beta_1 x_{ij} + u_{1j} + u_{0j} + e_{ij}$
  - `lmer(y ~ x + (1 + x|groupvar), data)`


Cubic Function

- **Model 4: Growth curve model**
  - Quadratic: `lmer(y ~ x + I(x^2) + (1 + x|groupvar), data)`
  - Cubit: `lmer(y ~ x + I(x^2) + I(x^3) + (1 + x|groupvar), data)`

Step 3.2 : Longitudinal Multilevel Model Fitting Process



- **N**on **L**inear **M**ixed **E**ffects (**nlme** package)
- **Model 0: non-hierarchical intercept-only model**
  - `gls(y ~ 1, data, method = "ML", na.action = na.exclude)`
- **Model 1: Intercept only multilevel linear model**
  - `lme(y ~ 1, data, random = ~1|Person, method = "ML", na.action = na.exclude)`
- **Model 2: Longitudinal Multilevel Linear Model**
  - `lme(y ~ Time, data, random = ~Time|Person, method = "ML", na.action = na.exclude, control = list(opt="optim"), correlation = corAR1(0, form = ~Time|Person) ))`
- **Model 3: Longitudinal Multilevel Linear Model + Growth Curve**
  - `Model 2 + I(Time^2) in random intercept`
- **Model 4: Longitudinal Multilevel Linear Model + Growth Curve + Addition of a fixed covariate**
  - `Model 3 + covariate, correlation = corAR1(0, form = ~1|Person)`
- **Model 5: Model 4 + Level 3 Random (e.g. clinic)**
  - `lme(Y~ Time + I(Time^2), data, random = ~1|Clinic/Person, method ="ML", na.action = na.exclude, control = list(opt="optim"), correlation = corAR1(0, form =~1|Clinic/Person))`