**Thesis Summary (FINISH BY 15 SEP)**

**Contribution 1: Exploratory data analysis (Sem 1) – can probably omit – write this part last (tonight)**

* **Produce the following**
  + Score distributions (histograms) showing scores are quite skewed – why we need to discretise based on the population distribution for scores.
  + Distributions showing the number of problems completed
  + Outcome distributions
    - Most commonly occurring sequences (may have overlap with below)

For each student

* Worthwhile understanding how many times they made a submission before passing, and ran a problem before passing
* Make no attempt at slides but start attempting problem/pass problem
* Different journeys

Showed why you didn’t use slide views – too sporadic and problem views

