Will's notes

Forward Feature Selection

Given $F=\{f_1, f_2 \dots f_k\}$ features
Want: Reduce #F s.t. $F=\{f_1, f_2 \dots f_n\}$ where $n<k$

(A) Sol. ① take f_1 s.t. $Y=\beta_0+\beta_1 f_1$ - model is ran.
② Next establish a training in sample dataset and a test outsample data set
③ Next Collect performance score $a_1 \sim f_1$

Repeat to f_2 s.t. $A=\{a_1, a_2\}$ ... to f_n s.t. $A=\{a_1, a_2 \dots a_n\}$

(B)
① Take $\max\{A\}=a_i \sim f_i$ s.t. $Y=\beta_0+\beta_1 f_i$
② take f_i s.t. $Y=\beta_0+\beta_1 f_i+\beta_2 f_1$ model is ran.
③ Collect performance metric, similar to part A s.t.

$B\{b_1, b_2 \dots b_k\}$, where b_j is performance score associated with $Y=\beta_0+\beta_1 f_i+\beta_2 f_j$

(C) We repeat the process till times s.t. $Y=\beta_0+\beta_1 f_i+\beta_2 f_j \dots \beta_n f_n$ / doe

Will take a long time to run /