*Step 3- Develop Overall Strategy:*

NOTE: in green are additional actions needed based on the initial actions.

Based on the questions and what's in the dataframes, I will do the following:

* Remove all columns except those identified in Observations. ✅
  + Dete\_survey
* To answer questions, Role Start and Cessation dates for employment provided
* To answer questions, Age provided
* To answer questions, Job dissatisfaction and dissatisfaction with department provided. No missing data in these columns, boolean data
* Tafe\_survey
  + To answer questions, gives Length of Service at workplace and at institute.
  + To answer questions, Age provided
  + To answer questions, column indicating Contributing Factors. Job Dissatisfaction and Contributing Factors. Dissatisfaction provided. Not boolean, value missing if Dissatisfaction not a contributing factor.
* Convert the Ages and Length of Service ranges to numeric values. Use the midpoint of the range. ✅
* Make a new column indicating the survey name and populate with the name. ✅
* Make a new column in each survey df where any True row(index) in either dissatisfaction column registers as True. Name the new column Combined\_Dissatisfaction.✅
* For the dete\_survey df , convert the date to numeric (reference the DateTime module from a while back). Use the start and terminates dates and arithmetic to create a new column length of service (in years).✅
* Combine the tafe\_survey Length of Service columns into one, using same column name as dete\_survey Length of Service.✅
* Convert the tafe\_survey Length of Service to numeric. ✅
* Change column names so that Age and Length of Service have the same name (so they make a single column at merge).
* Merge the two surveys.✅
* Convert the Age columns to numeric (did earlier).✅
* Make a Histogram plot of Age, Length of Service. ✅
* Based on Histogram, divide survey into "older/younger" and "long/short" term of service.✅
* Make a Bar plot with basic statistics to see if Combined\_Dissatisfaction differences exist older/younger and long/short. ✅
* NOTE: if I knew how to do logistical regression in Python I'd do that after the merge.