Evaluation script flow:

1. Take the original file (online\_retail\_test\_public.csv) and recommended file (vit\_recommended) as input
2. Check for No. of customers recommended against the original file
   1. If Original file has the customer ID , “A” , “B” , “C” , “D”
   2. Recommended file also should have the same list of customer ID , “A” , “B” , “C” , “D”
   3. Probably, getting length of the file would be a crude way of approaching it rather create an array for customerID from both the file and compare the array…not sure whether python/pandas have any better way to handle this
3. If the recommended file has lesser number of customerID, then penalty has to be included as listed below
4. For each customerID, create a group/array with the items
   1. Compare the items for each customerID from both the files and create a precision count.
   2. For example:
      1. CustomerID “A” [Items bought] = { 123,456,789,323,575,909}. (original file)
      2. customerID “A” [items recommended]={123,476, 999}. (vit\_recomended )
      3. precision = 1/1+2. = 0.33 (true positive / true positive + false positive) ( correctly recommended items / total recommended items )
      4. Recall = 1 / 1 = (correctly recommended items / total useful recommended items )
      5. F- meashure = 2 precision \* recall / precision + recall
      6. return 2 \* (1/3\* 1 ) / 1/3 + 1)
      7. F-measure 0.248
   3. Example 2:
      1. customerID”B”[items bought]={123, 789,117,333,888,223}
      2. customerID”B”[recommended items]={789}
      3. precision = 1/1
      4. recall = 1/1
      5. return 2 Precision \* recall / Precision + recall
      6. return 2 (1\* 1 / 1 + 1)
      7. 2 \* (0.5)
      8. F1 measure 1
   4. Example 3:
      1. customerID”C”[items bought]={123, 789,117,333,888,223}
      2. customerID”C”[recommended items]={111}
      3. precision = 0/1
      4. recall = 0
      5. return 2 Precision \* recall / Precision + recall
      6. return 2 (0)
      7. F1 measure - 0
   5. Taking average of all precision and return it as single precision score
   6. Penalty of 0 % in the consolidated score, if the recommended list has all the customer id
   7. Penaly of 10% in the score , if the recommended list has 90% of customer id
   8. Penalty of 20 % in the score , if the recommended list has 80% of customer id
   9. Penalty of 40% flat in the score, if the recommendation has less than 80% of customer
   10. Return the final score as the FINAL PRECISION
5. Other things to consider:
   1. Top “N” is 10 items is recommended. But no penalty for anything less..