GENDER CONVERSION

Input:

The input has three columns: **SentenceID**, **English** and **Arabic**.

These columns should not change throughout the annotation process.

Output:

The output has two columns:

SentenceType, and **SwitchTranslation**

The value of **SentenceType** should be a two-letter code, where the first code refers to the status of the 1st person reference in the sentence, and the second code refers to the status of the 2nd person reference in the sentence.

The status should be one of four values:

- M: Masculine
- F: Feminine
- B: Ambiguous
- D: does not exist

Here are all the possible combinations with examples:

Code	Meaning	Example
M-M	The sentence has a masculine 1st person reference and a masculine 2nd person reference	أنا صبور مثلك يا صديقي
M-F	The sentence has a masculine 1st person reference and a feminine 2nd person reference	أنا صبور مثلك يا صديقتي
F-M	The sentence has a feminine 1st person reference and a masculine 2nd person reference	أنا صبورة مثلك يا صديقي
F-F	The sentence has a feminine 1st person reference and a feminine 2nd person reference	أنا صبورة مثلك يا صديقتي
M-B	The sentence has a masculine 1st person reference and an ambiguous 2nd person reference	أنا صبور مثلك
F-B	The sentence has a feminine 1st person reference and an ambiguous 2nd person reference	مرحبا ، انا جارتك الجديدة
B-M	The sentence has an ambiguous 1st person reference and a masculine 2nd person reference	أنا من تنتظره
B-F	The sentence has an ambiguous 1st person reference and a feminine 2nd person reference	أحبك يا صديقتي
B-B	The sentence has ambiguous 1st person and 2nd person references	أحبك انت
M -D	The sentence has a masculine 1st person reference but NO 2nd person reference	أنا قوي
F- D	The sentence has a feminine 1st person reference but NO 2nd person reference	انا سعيدة جدا
D -M	The sentence has a masculine 2nd person reference but NO 1st person reference	قف !
D -F	The sentence has a feminine 2nd person reference but NO 1st person reference	انت تكتبين
B- D	The sentence has an ambiguous 1st person reference but NO 2nd person reference	حسناً. انا هنا
D -B	The sentence has an ambiguous 2nd person reference but NO 1st person reference	الجو ممتع في بلدك
D-D	The sentence has no reference for 1st nor 2nd person	الجو ممتع

The **SwitchTranslation** column should have the Arabic sentence after converting the masculine references to feminine and the feminine references to masculine. The only exceptions are M-F, F-M, M-M and F-F where only the gender of the 1st person reference should be changed (Note: we will have a second round later on to change the 2nd person in these four cases).

Here are the changes the sentences above should receive (the comment column does not exist in the annotations but it is just to let you understand the task).

Code	Arabic	SwitchTranslation	Comment
M-M	أنا صبور مثلك يا صديقي	أنا صبورة مثلك يا صديقي	Only the 1st person changes because it is an M-M case
M-F	أنا صبور مثلك يا صديقتي	أنا صبورة مثلك يا صديقتي	Only the 1st person changes because it is an M-F case
F-M	أنا صبورة مثلك يا صديقي	أنا صبور مثلك يا صديقي	Only the 1st person changes because it is an F-M case
F-F	أنا صبورة مثلك يا صديقتي	أنا صبور مثلك يا صديقتي	Only the 1st person changes because it is an F-F case
M-B	أنا صبور مثلك	أنا صبورة مثلك	
F-B	مرحبا ، انا جارتك الجديدة	مرحبا ، انا جارك الجديد	
B-M	أنا من تنتظره	أنا من تنتظريه	
B-F	أحبك يا صديقتي	أحبك يا صديقي	
B-B	أحبك انت		No change is needed. Leave **SwitchTranslation** empty
M -D	أنا قوي	أنا قوية	
F- D	انا سعيدة جدا	انا سعيد جدا	
D -M	قف !	قفي !	
D -F	انت تكتبين	انت تكتب	
B- D	حسناً. انا هنا		No change is needed. Leave **SwitchTranslation** empty
D -B	الجو ممتع في بلدك		No change is needed. Leave **SwitchTranslation** empty
D-D	الجو ممتع		No change is needed. Leave **SwitchTranslation** empty

VERY IMPORTANT

What should we do with 3rd person references?
Nothing. This task is only about 1st and 2nd person references.
For example: هذا هو الرجل الشجاع is an D-D case.

How to handle dual and plural cases?
We do not expect the dual and plural to happen much, but when they do add a "!" to the label (e.g., M!, F! and B!).
For example:
لقد سمعتكن تتكلمن عنها should be tagged as B-F!
لقد سمعتك تتكلمين عن أخلاقنا should be tagged as B!-F

How to handle proper nouns?
Handle proper nouns as ambiguous. Do not make any assumptions about proper nouns.
For example: مرحبا ، انا ماري should be tagged as B-D
However, مرحبا ، انا ماري جارتك الجديدة should be tagged as F-B because of the word الجديدة

What do to with writing mistakes?
Do NOT correct any mistakes of any kind.
Do NOT correct even hamzas, Ta' marboota, Alif maqsoura or misplaced punctuation.
Do not add any diacritization. Notice that أحبك انت is a B-B case as diacritization is absent.
If the Arabic sentence has major mistakes (ex: نعم ، أوه ، طفلرضيع ، أعرف), just make the sentence type B-S and leave **SwitchTranslation** empty.

What if the English column does not match the translation in the Arabic column?
Ignore the English column, do not worry about it. Also, do not use the English column to judge the gender or any other information. Just use the English column if you need to understand an Arabic sentence that is not clear for some reason. For example:
We know that the word ذهبت in ذهبت إلى هناك refers to 1S not 2MS because the English translation is "I went" not "You went".

Any hard cases we can encounter?
Yes, but this is infrequent.
For example: Sometimes you need to change the lemma itself for the gender conversion.
Examples:
انا رجل ➔ انا امرأة
أحبك يا أمي ➔ أحبك يا أبي
نعم أنا هو ➔ نعم أنا هي (note that although هي is 3rd person, but we need to change it in order to adjust the 1st person reference).
However, if you encounter any **HARD** case that you cannot figure out (e.g., انا حامل), just write the word **HARD** in the **SwitchTranslation** column.