Context: Healthcare Industry

Problem: Identify patient treatments based on receipt history

Details: Imagine we have a source table (see attachment) containing, per patient, all the drugs he took from the pharmacy in the past 5 years.

> In the first picture (*input*)we see the structure of the source table
> PatientID = unique identifier per patient
> DrugName = just the drug name
> StartDate = the date when the patient took the drug from the pharmacy
> EndDate = until which date the prescribed quantity is sufficient

> Based on this source data we want to compute the monotherapies or combined therapies the patient is following and for what periods
> In the second picture (*output*) we see the expected result.

> Patient = PatientID
> Treatment = combination of drugs that are taken in the same time. For example, if Patient 1234 takes both drugs A and B in the last week of March 2021, then he will be on the treatment A+B during that week
> StartDate = start date of the treatment
> EndDate = end date of the treatment
> **Note: Please compact the treatment periods to have as few rows as possible.
>   Please ensure the treatment periods do not superpose

Input

PatientID	DrugName	StartDate	EndDate
1234	A	01-Jan-21	01-Feb-21
1234	A	01-Feb-21	01-Mar-21
1234	A	01-Mar-21	01-Apr-21
1234	B	15-Mar-21	15-Apr-21
1234	A	01-Apr-21	01-Jun-21
1234	B	15-Apr-21	01-Jun-21
1234	C	01-Jun-21	01-Jul-21
1234	C	01-Jul-21	01-Aug-21
1235	D	01-Jan-21	01-Feb-21
1235	D	01-Feb-21	01-Mar-21

Output

Patient	Treatment	StartDate	EndDate
1234	A	01-Jan-21	15-Mar-21
1234	A+B	15-Mar-21	01-Jun-21
1234	C	01-Jun-21	01-Aug-21
1235	D	01-Jan-21	...