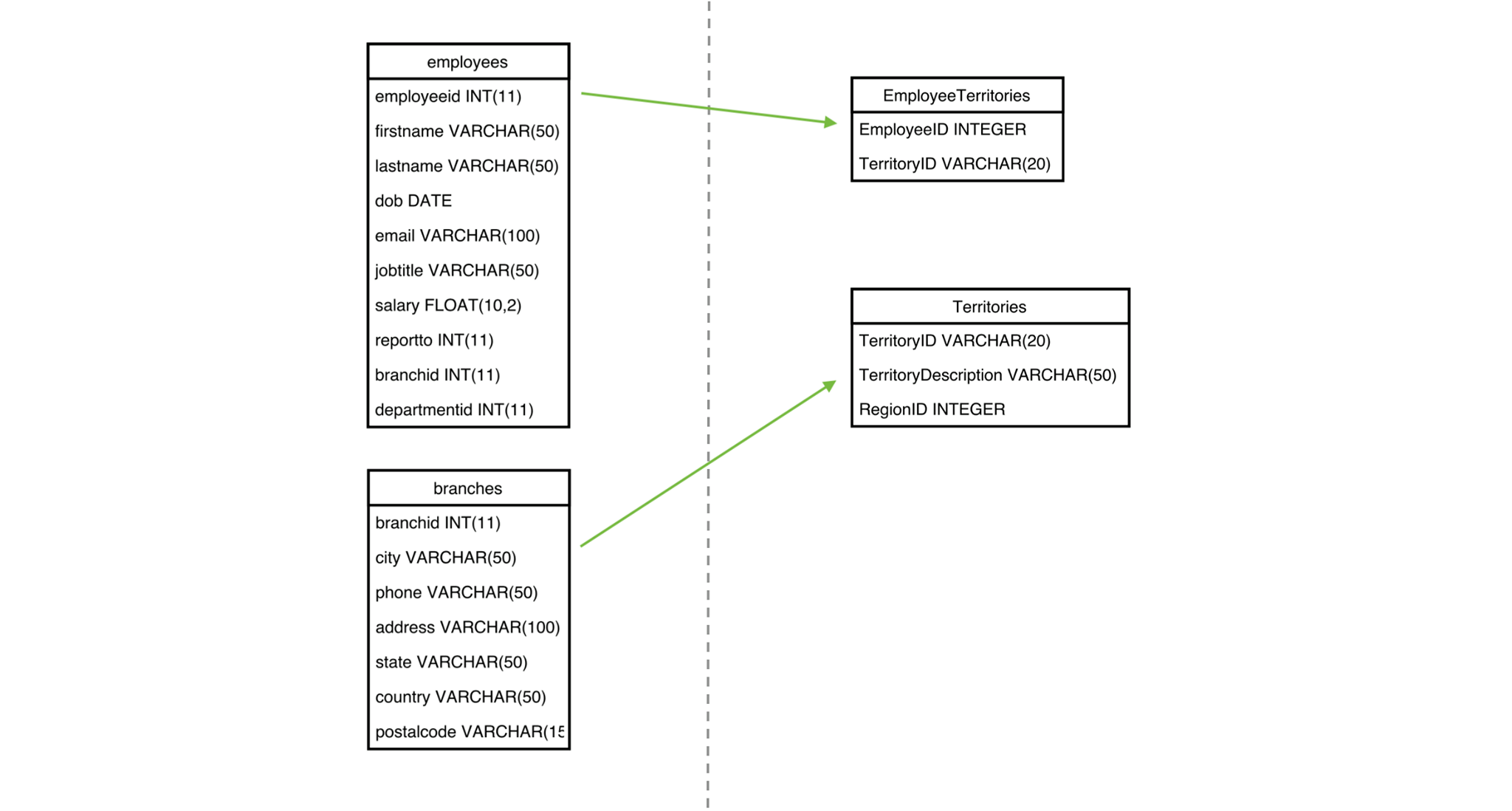


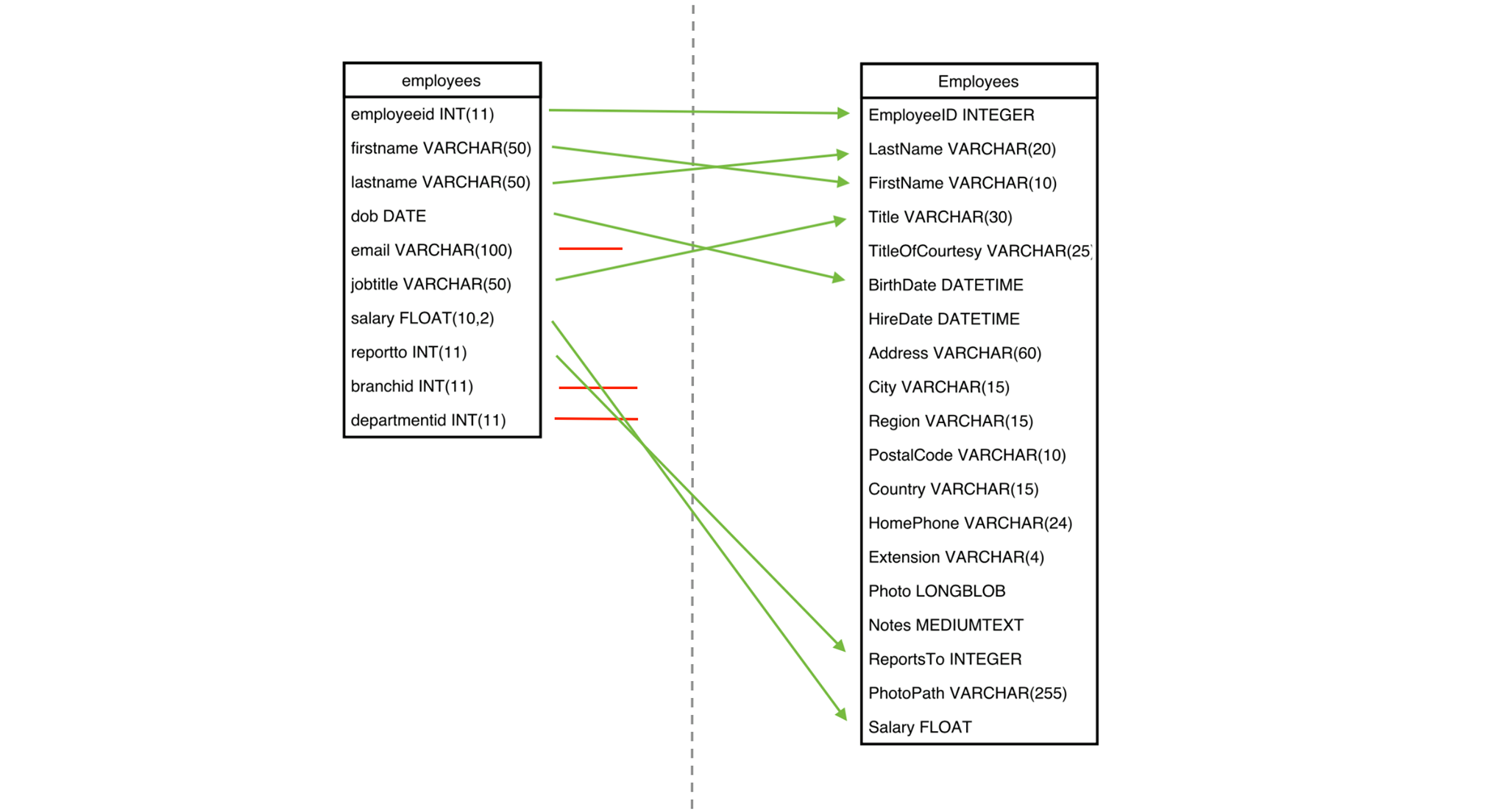
Análise e Integração de dados

Ana Leitão 81365

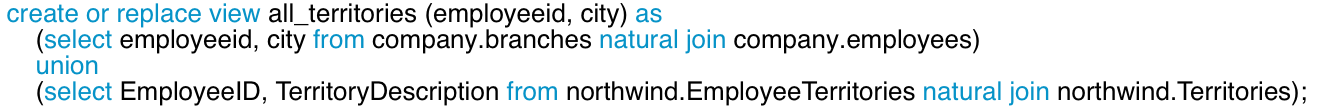
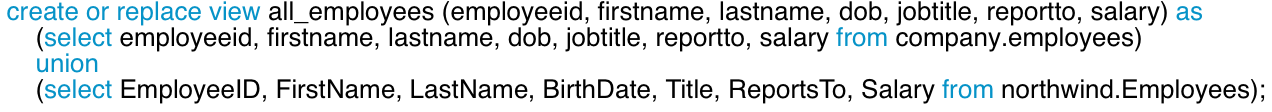
João Ramos 80915

schema matching

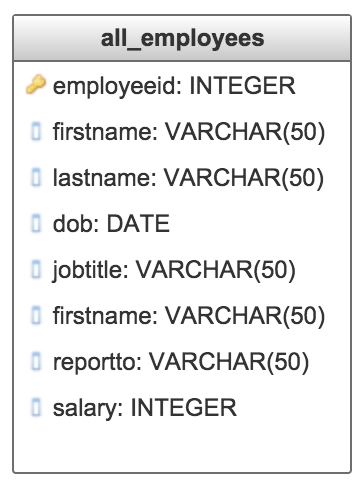


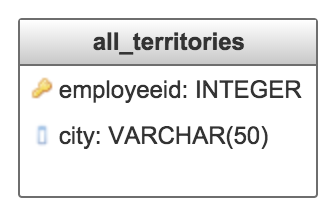


SQL views in text

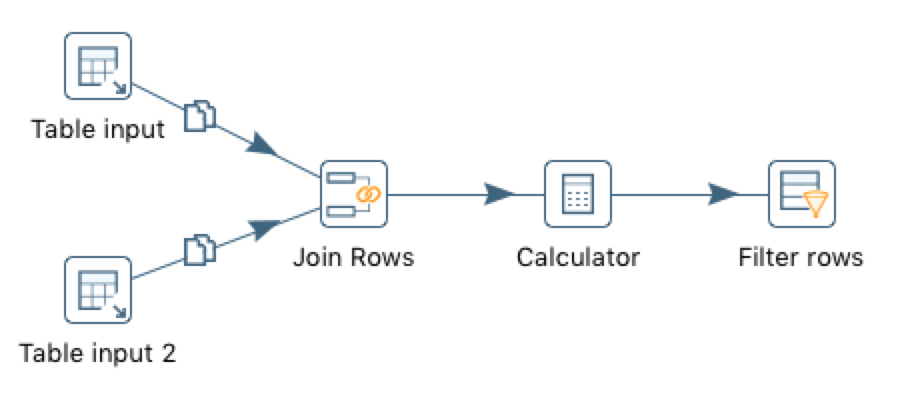


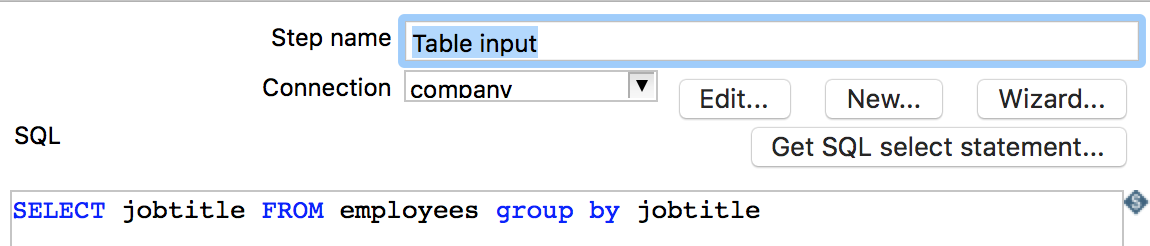
mediated schema

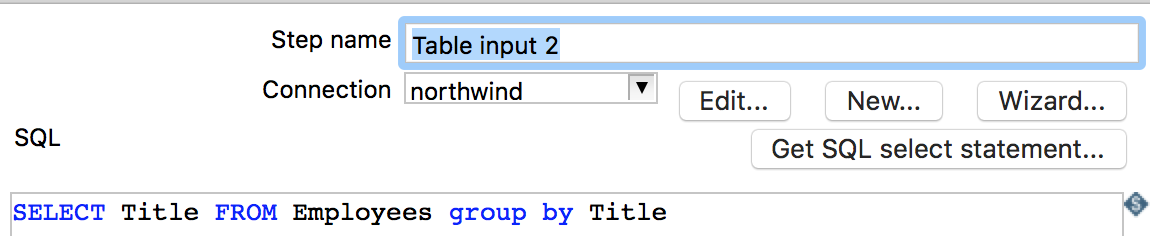


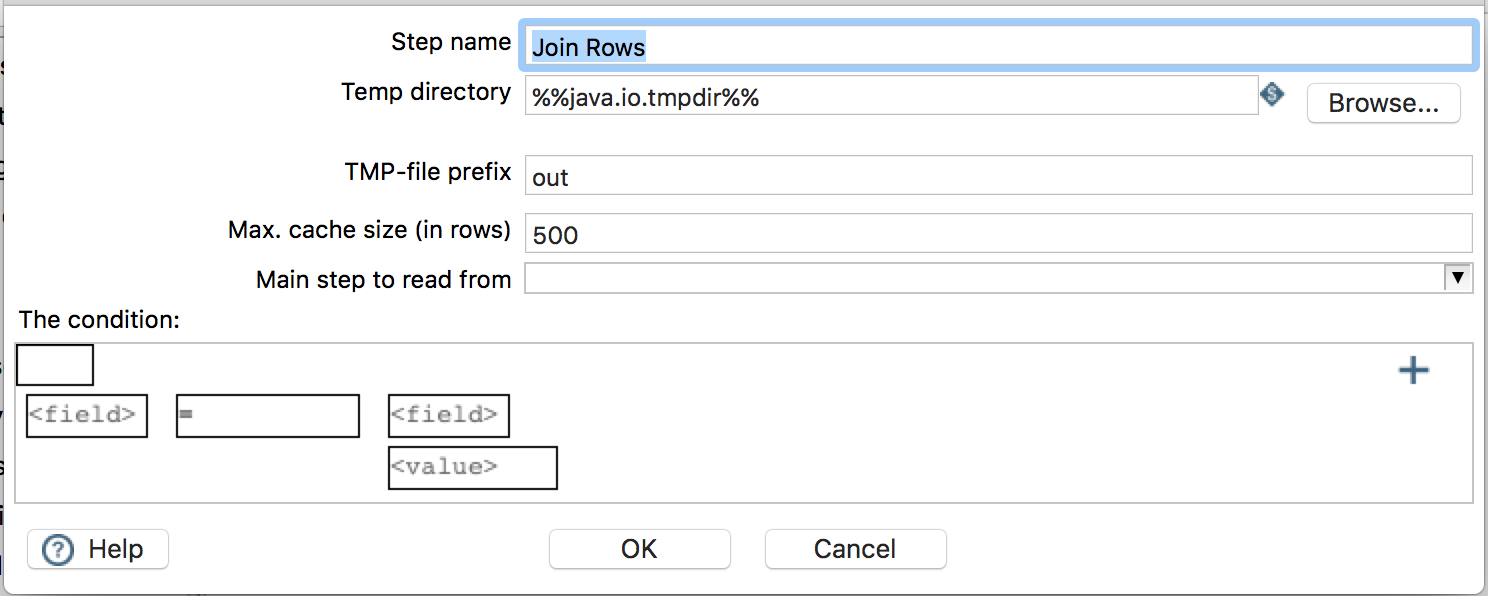


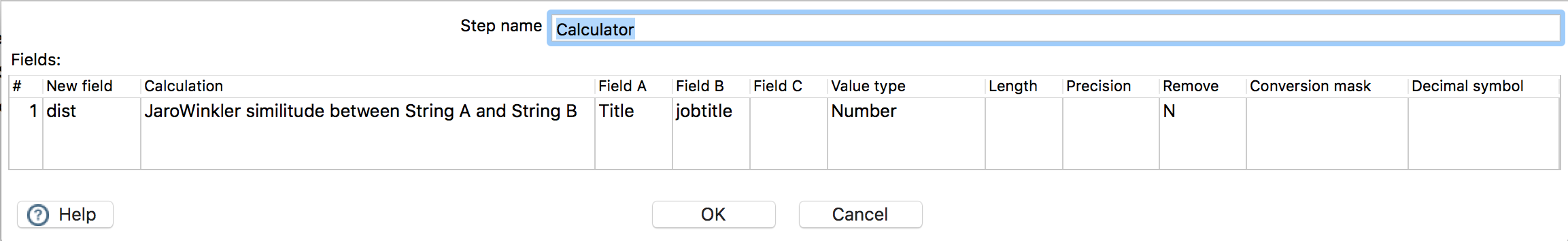
Duplicated job titles

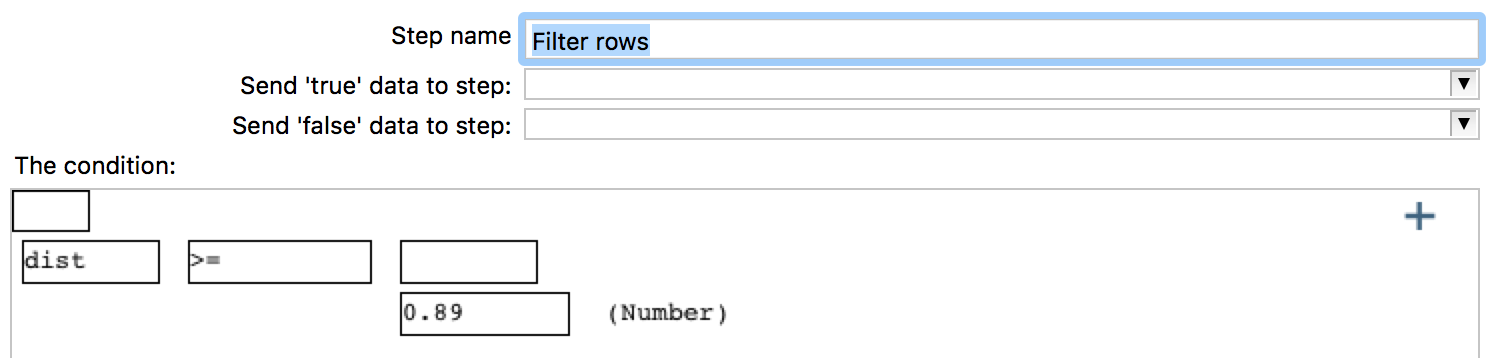


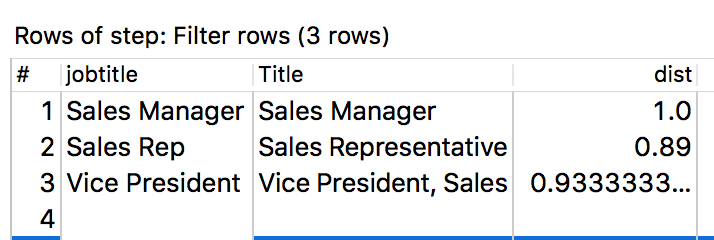












Comparing the job titles in both databases we concluded that a similarity measure would be more appropriate than a distance based one. For example we have the title “sales rep” and “sales representative”, they have the same meaning but a distance based string matching technique would have a high result which is the opposite of what is desired.

Northwind data warehouse



