金融大數據 2023/4/14 筆記

* 最好的分組方式

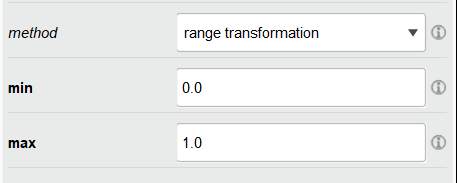
群內的相似性越高越好

群間的相異程度越高越好

* R/F/M
* R: Recency 最近什麼時候來買  
  →距離今天為止，購買的時間越長，Recency越低。

〔作業Hint〕Aggregate 聚合

* + 顧客ID → 計算Recency
* F: Frequency 購買頻率
* M: Money 金額
* 當用RFM分開後，觀察並解釋其他特徵。  
  → 因為理論上用RFM分開後，RFM會分得很開、彼此間最相像。
* 未知的屬性 拿來做集群分群
* 已知的屬性 拿來做解釋
* 正規化



Range transformation → 將屬性分到0~1之間