

データ科学2 レポート課題：相関ルール分析に関する計算機実験

---

- ライブラリ `arules` に付属しているデータ `Groceries` を対象に、頻出パターン・相関ルール分析を行い、結果を実験レポートとしてまとめなさい。なお、R を起動し

```
> library(arules)
> data(Groceries)
```

と入力することで、`Groceries` データが準備されます。

`arules` ライブラリがインストールされていることが前提です。インストールされていない場合は、資料を参考にインストールしてください。

- レポートでは、最小支持度や最小確信度を変えることで、結果がどの様に変化するか考察してください。
- また、以下の例を参考に、ある条件下での（飽和）パターンや相関ルールを抽出するなどしても良いでしょう。
  - 最小支持度を  $0.001 (=0.1\%)$  としたときの、飽和パターンの集合
  - 最小支持度を  $0.001 (=0.1\%)$  としたときの、“yogurt”を含むサイズ 3 以上の飽和パターンを支持度の高い方から 5 個
  - 最小支持度  $0.1\%$ 、最小確信度  $0.5$  としたときの、ルールの前提部に“whole milk”を持つ相関ルールのうち、Confidence の大きな方から 3 個
- 必要に応じて視覚化するのも良いかもしれません。
- レポートは、`Latex` を用いて作成し、A4 サイズの PDF ファイルに変換した上で提出すること。
- 提出ファイル名は、s54YyXxx-report3.pdf とすること（s54YyXxx の部分は、自身の学生証番号に合わせて読みかえてください）
- 分量は任意とするが、必ず実験レポートであることを踏まえた上でレポートを作成すること
- ※切 2020/07/26（レポート未提の場合、成績は E 判定とする）