EDUCATION PROGRAM FOR TOP SOFTWARE ENGINEERS

### レコメンデーションエンジン データ仕様

EDUCATION PROGRAM FOR TOP SOFTWARE ENGINEERS

購入ペアデータの仕様



- データはHDFS上の「/user/root/hadoop\_exercise/04/data/goods\_pair」に テキストデータとして配置されている
- 各行には、購入された商品のペアがカンマ区切りで記録されている

• • •

あずきキャラメル,チョコレートマシュマロ あんドーナツ,あずきキャラメル チョコレートマシュマロ,びわゼリー あずきキャラメル,フィナンシェ

• • •

# GINEERS EDUCATION OF THE NGINEERS OF THE NGINE

#### 関連度分母データの仕様

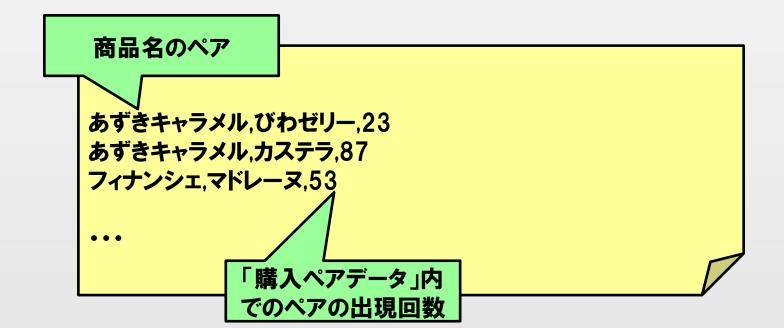
- データはHDFS上の「/user/root/hadoop\_exercise/04/data/」に「denomination」という名前のテキストデータとして配置する
- 各行には次の順にデータがカンマ区切りで記録されている
  - 1 商品名
  - 2 ①の商品が購入ペアデータのペア内に出現した回数



### 4 SINEERS EDUCATION PHONES

#### 関連度分子データの仕様

- データはHDFS上の「/user/root/hadoop\_exercise/04/data/」に「numerator」という名前のテキストデータとして配置する
- 各行には次の順にデータがカンマ区切りで記録されている
  - 1 カンマ区切りの商品名のペア(商品名は昇順)
  - 2 ①のペアが購入データペア内に出現した回数(ペアの順序は問わない)



## 5 CHEERS EDUCATION PRODUCTION OF THE PRODUCTION

### 関連商品データの仕様(参考)

- データはHDFS上の「/user/root/hadoop\_exercise/04/data/」に「related\_goods」という名前のテキストデータとして配置する
- 各行には、次の順でデータがカンマ区切りに記録されている
  - 1 商品名(その1)
  - 2 ①の商品と関連がある商品名
  - ③ ①の商品と②の商品の関連度(0.025~1)

• • •

あずきキャラメル,チョコレートマシュマロ,0.032 あずきキャラメル,草加せんべい,0.031 あんドーナツ,あずきキャラメル,0.04 チョコレートマシュマロ,びわゼリー,0.042 チョコレートマシュマロ,きんつば,0.026

• • •