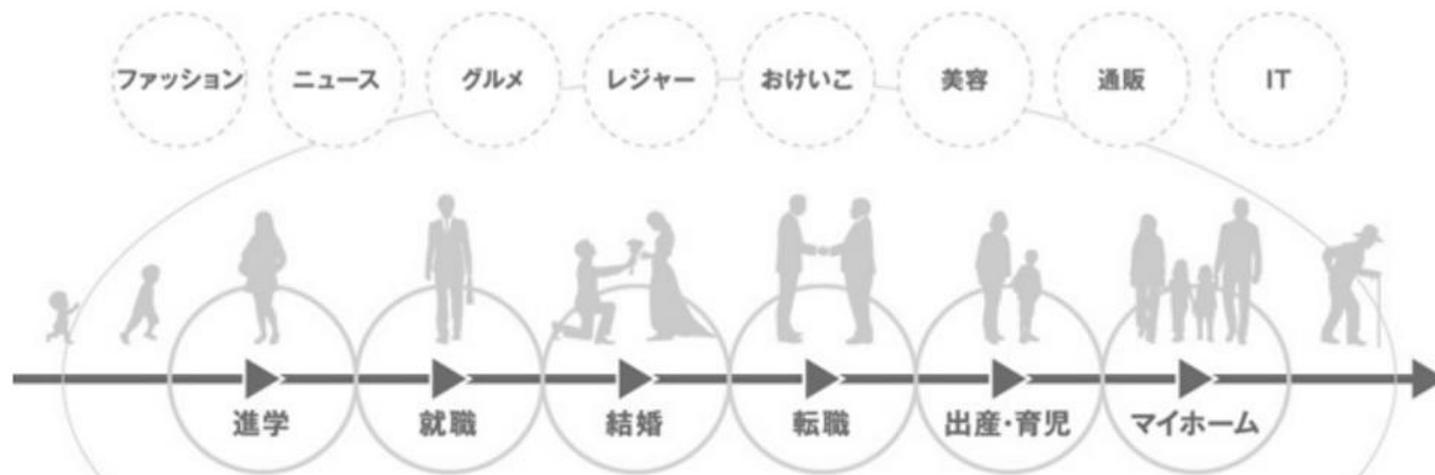


リクルート提供データのご紹介 @DBSJ Data Challenge

リクルートテクノロジーズ 櫻井 一貴



今回リクルートからは
「検索キーワード」および
周辺のデータをご提供します



サービスはたくさんありますが...



and more ...

ジャンルから商品を探す

すべて

+ 詳細条件

検索する

カート

お気に入り

購入履歴

クーポン

マイページ

- ファッション >
- 靴・バッグ >
- インテリア・家具 >
- 日用雑貨・文房具・キッチン用品 >
- 食品・スイーツ・お取寄せ >
- 水・ドリンク・お酒 >
- 美容・コスメ・健康 >
- 家電・カメラ・オーディオ >
- パソコン・スマホ・周辺機器 >
- 本・音楽・DVD >
- ゲーム・おもちゃ・ホビー >
- スポーツ・アウトドア >
- 花・ガーデニング・DIY >
- ペット・ペット用品 >
- キッズ・ベビー・マタニティ >
- 車・バイク >
- サービス他 >

注目ワード: 物置 クーポン 吸汗速乾Tシャツ 米 水着 コンタクトレンズ



最低価格保証

送料 無料 **3,980円**

シニアアソムリエ厳選
すべて金賞ボルドー5本

毎日更新!
最低価格保証商品

今回はコチラからご提供します

あなたへのおすすめ

ポンパレモール

【メーカー正規品】新型オカカ
蟹節削り器

5,100円

毎日更新!
お買い物クーポン

食品

店内全品対象 **100円OFF**
美味しさは元気の源...

店内全品対象 **3%OFF**
【全商品対象】京橋...

店内全品対象 **100円OFF**

スイーツ・お菓子・ケーキランキング 2017/06/19更新

ランキングをもっと見る

1位 無添加 大森焼きアーモンド 1,580円 3% 43ポイント	2位 無添加 無糖・無油 2,780円 5% 128ポイント	3位 無添加 ミックスナッツ 2,990円 3% 83ポイント	4位 無添加 アーモンド小倉 777円 3% 21ポイント	5位 1kg 生チョコ 1,720円 3% 47ポイント	6位 新発売! ネット通販限定 ミルクカタラーナ 1,799円 3% 49ポイント	7位 無添加 ミックスナッツ 1,000円 5% 46ポイント	8位 厚手の濃厚抹茶あん 宇治抹茶だいふく 1,360円 3% 37ポイント	9位 無添加 大森焼きアーモンド 880円 3% 24ポイント	10位 1kg 送料無料 1,980円 3% 54ポイント	11位 最高級 ミックスナッツ 2,780円 5% 128ポイント	12位 Happy Mix 1,000円 3% 27ポイント
---	--	---	---	--	---	---	--	---	---	---	---

スイーツ・お菓子・ケーキ対象のクーポン

クーポンをもっと見る

店内全品対象 100円OFF 美味しさは元気の源【目	Coupon 200円OFF しゃぶまるの全商品対	オーケーフルーツ 2,000円以上で使える! 100円OFF	100円OFF 全品対象! 【オーケーフ	店内全品対象 2,000円以上で使える! 100円OFF 【ジェイエイトんどうフ	訳あり干し芋 100円OFF みりのりの里で使える100
--------------------------------------	-------------------------------------	---	-------------------------	--	---

抽選で当たる!
無料プレゼント!

本日の
ピックアップ!

メルサボン
インバス
ボディリトリートメント
ハーバルグリーン
500ml

リクルートカード

最大 **6,000** 円分
ポイントプレゼント!

×ポンパレモールで最大4.2%還元!

提供データ概要

1. 検索キーワードリスト

- ポンパレモールで検索されたキーワード
 1. ユーザーID
 2. 検索キーワード
 3. タイムスタンプ

2. アイテムクリックキーワードリスト

- キーワード検索でアイテムクリックされた際の対象データ
 1. ユーザーID
 2. 検索キーワード
 3. タイムスタンプ
 4. アイテムジャンル名

3. 商品情報

1. アイテム名
2. アイテムジャンル名
3. 登録日時

期間：

2016年6月～2017年12月

※全体規模の推定ができないよう
一定基準でリサンプリングした
数千万行のデータセット

データ公開に対する考え方

- 公開できる情報は極力公開したい

- 公開できないもの

- サービス全体の売上や規模が推定できてしまうもの
 - 上場企業では行ってはならない
- 利用者が特定できてしまうもの
 - 利用者の属性情報だけでなく、特定の利用者だけが検索したキーワードも危険
- 特定の商品やショップの人気の推定できてしまうもの
 - 特定のショップの売上や規模が推定できてしまうため

ご了承下さい



ヒントというかエールというか

データはいたってシンプルです

- 与えられた制約の中でいかに結果を出すか。分析屋さんの腕の見せどころ

分析は提供するデータで閉じる必要はありません

- 世の中に公開されているデータと絡めてみる？
- 競合サービスと比較してみる？

企業である以上、直近のモチベーションは利益が増えることです

- 利益が増える… 売上up, コストdown, マッチング精度up …

もちろんクライアントもカスタマーも嬉しくなってほしいです

- より商品が売れる, より便利に使える, …etc.

(参考) 当該データを用いた分析の事例について

- ・ 購買検索ログを使用した**注釈付きカテゴリ**の提案
- ・ クリック履歴を考慮した**商品ランキング**方式の提案
- ・ ネットショッピングにおける**探索行動のモデル化**手法の提案
- ・ 検索キーワード**表記ゆれ改善**のための分析

(分析された方の主な感想)

- ・ 自身も普段「検索」をするので仮説を立てやすかった
- ・ 想定以上にデータがキレイじゃなかった…
- ・ 仮説検証をしすぎてタスク確定までに時間がかかってしまった

たくさんのご応募を
お待ちしております！

dbsj_datahack@googlegroups.com