

2023/12/11

IDRユーザフォーラム 2023
データ提供者セッション

リクルートのご紹介

～企業紹介とデータ利活用事例のご紹介

株式会社リクルート
プロダクト企画統括室 櫻井 一貴



Agenda

1. リクルートについて
2. IDRにご提供しているデータのこと
3. データ利活用事例
4. さいごに

リクルートについて

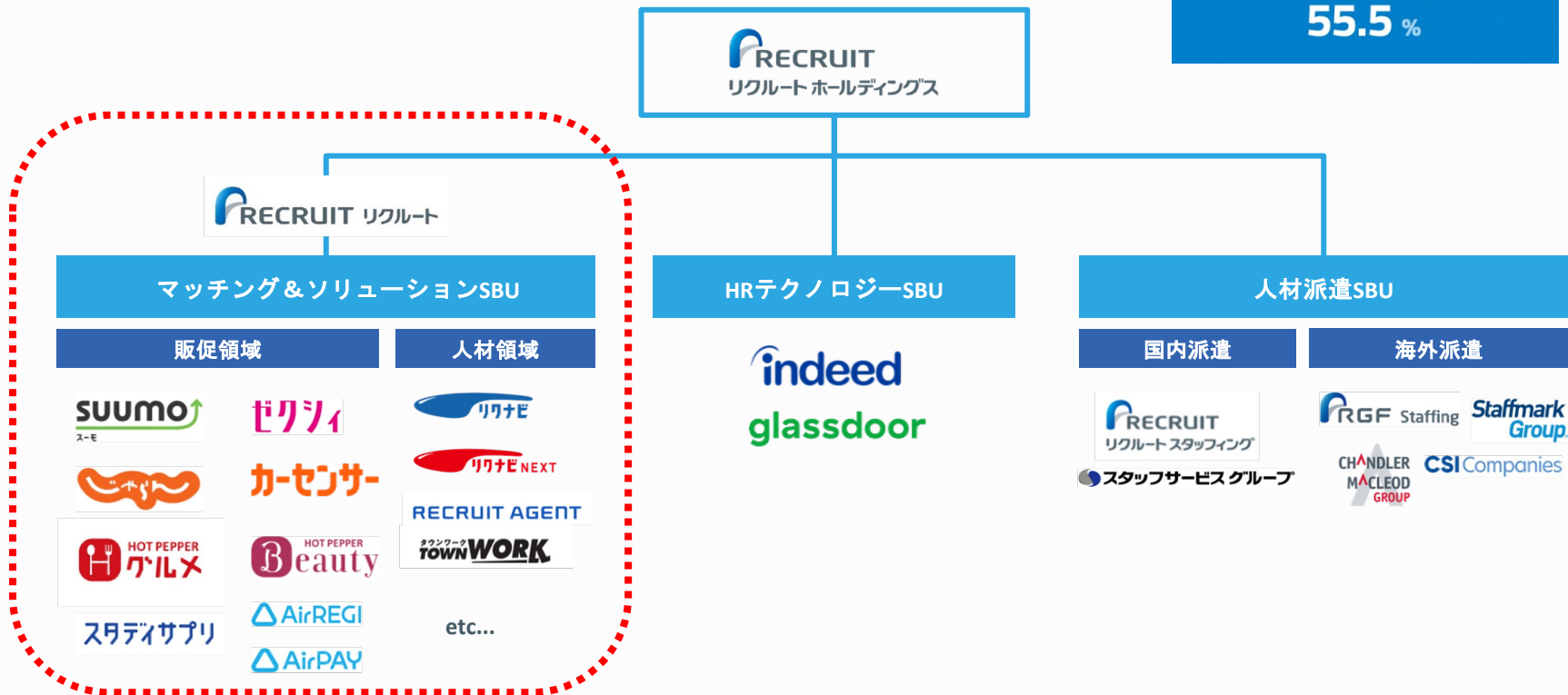
リクルートグループについて

リクルートは日本国内における事業会社

リクルートホールディングス全体の売上比率は「日本国内＜海外」となっています

2022年3月末時点

海外売上高比率
55.5%



事業領域は「人材」と「販促」

ミッション「まだ、ここにはない、出会い。」を「より早く、シンプルに、もっと近くに。」実現のために
人材マッチングビジネスと販促ビジネスで、事業者と個人のマッチング支援、事業者の業務・経営支援などを行っています

販促ビジネス

(主に国内)

suumo
スーム

ゼクシィ



カーセンサー

**HOT PEPPER
カルメ**

**HOT PEPPER
Beauty**

スマテイサブリ

AirREGI
AirPAY

人材マッチングビジネス

(国内外)

リクナビ

indeed

RECRUIT
リクルートスタッフィング

usg people

**Staffmark
Group.**

リクナビNEXT

glassdoor

スタッフサービスグループ

**RGF Staffing
APEJ**

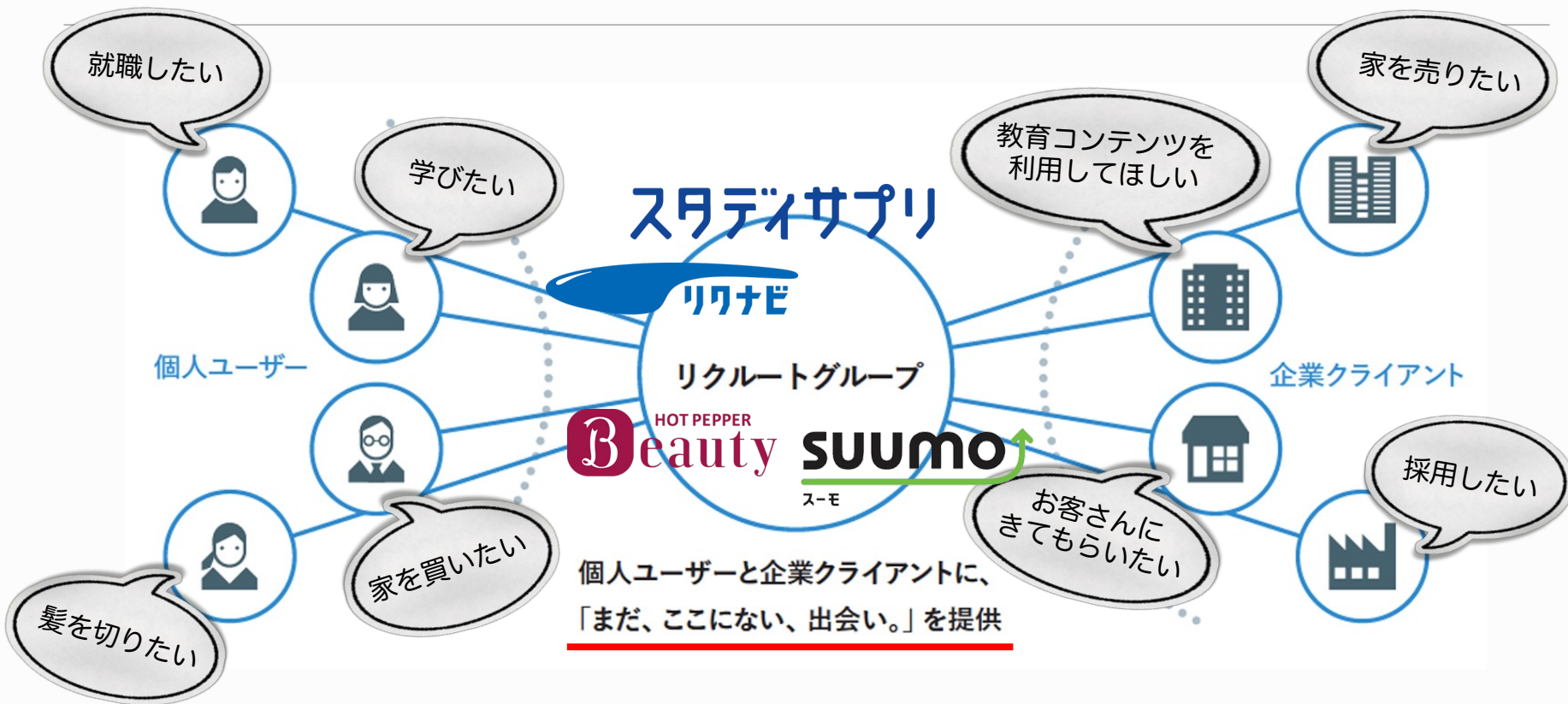
CSI Companies

RECRUIT AGENT

RGF Staffing UK

**タウンワーク
TOWNWORK**

基本となるビジネスモデル「リボンモデル」



基本となるビジネスモデル「リボンモデル」

スタディサプリ

リクナビ

リクルートグループ

HOT PEPPER
Beauty

SUUMO
スーモ

個人ユーザー

企業クライアント

就職したい

学びたい

教育コンテンツを
利用してほしい

家を売りたい

家を買いしたい

お客さんに
きてもらいたい

採用したい

髪を切りたい

個人ユーザーと企業クライアントに、
「まだ、ここにはない、出会い。」を提供

サービス多様化、ITシフトさらにSaaSソリューションへ

リクルートの60年は“変化対応”の歴史
紙→デジタル、グローバル経営の失敗と拡大、メディア事業に加えて
SaaSソリューションによる業務・経営支援への拡大



- 人材事業創業～
- 販促メディア事業への進出

1960-70年



- 販促メディア事業領域拡大
- 1990年代半ばにはネットにも挑戦

1980-90年



- 紙→ネット移行
- 中国での事業展開と撤退

2000年

CSI Companies

AirREGI AirPAY

indeed

glassdoor

RGF Staffing

スマディサプリ



- M&Aでのグローバル展開本格化
- デジタル事業拡大
- SaaS型業務経営支援

2010年

IDRにご提供しているデータについて

2014年からデータ公開をさせて頂いています

1. Yahoo!データセット
2. 楽天データセット
3. ニコニコデータセット
4. **リクルートデータセット**
5. クックパッドデータセット
6. LIFULL HOME'Sデータセット
7. 不満調査データセット
8. Sansanデータセット
9. インテージデータセット
10. オリコンデータセット
11. ダイエット口コミデータセット
12. 弁護士ドットコムデータセット
13. アットホームデータセット
14. . . .



The screenshot shows the NII IDR website interface. The main navigation bar includes 'HOME', 'データ一覧', '研究成果一覧', 'ユーザフォーラム', '組織', '関連リンク', and '問い合わせ'. The page title is '情報学研究データリポジトリ'. The left sidebar lists various data sets, with 'リクルートデータセット' highlighted in a red box. The main content area shows details for the '民間企業提供データ' (Private Company Provided Data) section, including 'Yahoo!データセット', '楽天データセット', and 'ニコニコデータセット'. The 'リクルートデータセット' section is also highlighted in a red box and lists the following data items:

1. ホットペッパービューティーデータ

データセットを利用した研究リスト

The image shows two overlapping browser windows of the DSC Reference Portal. The left window displays the homepage with a navigation menu and a list of datasets. A red callout box labeled "研究発表" (Research Publications) points to the "研究発表" link in the left sidebar and the "リクルートデータセット [9件]" entry in the dataset list. The right window shows the search results page for "研究発表", displaying a list of 9 items with titles, authors, and publication details.

研究発表

国立情報学研究所 (National Institute of Informatics) DSC Reference Portal

データセット共同利用研究開発センター (DSC) から提供するデータセットに関する文献情報のポータルです。以下の情報を公開しています。(SRCの音声コーパス及びNTCIRテストコレクション関係は順次掲載予定)

- 利用者から報告があった研究成果やデータセットを紹介した参考文献の文献情報
- DSC提供のデータセットのメタデータ, DOI, 謝辞, 引用方法

検索

インデックスリンク

- インデックスを選択して▼
- インデックスツリー
- 研究発表
- 音声コーパス
- 全般
- データセット

インデックスリスト

- Yahoo!データセット [226件]
- 楽天データセット [219件]
- ニコニコデータセット [33件]
- リクルートデータセット [9件]
- クックパッドデータセット [137件]
- LIFULL HOME'Sデータセット [25件]
- 不満足調査データセット [26件]

アイテムリスト

研究発表 > 情報学研究データリポジトリ (IDR) > リクルートデータセット

チェックしたアイテムをExport 実行 表示順 出版年 (降順) 表示数 20

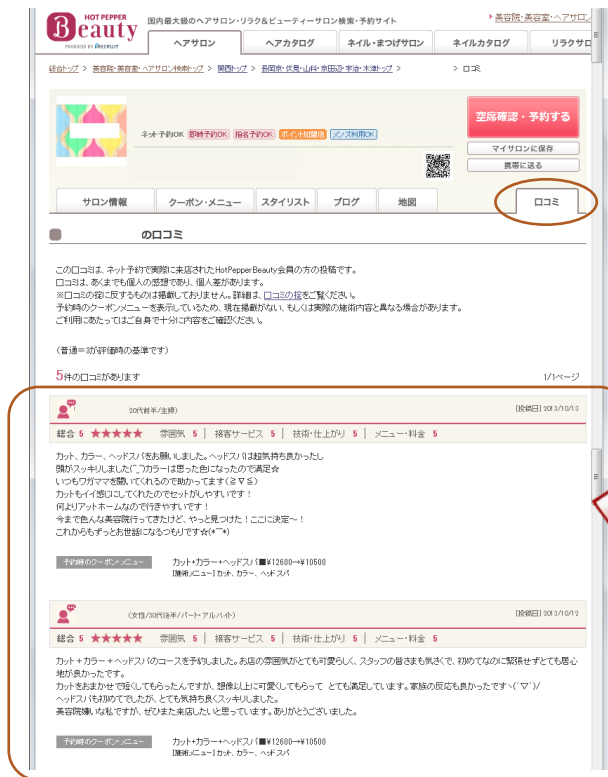
- ランダムフォレストを用いたカスタマーレビューの半自動分類
小倉 枝里子, 春藤 史郎, 石津 昌平
インテリジェントシステムシンポジウム2017予稿集, 154-157 (2017)
- 定数分析を用いた自作自演のロコミ判定モデルの構築と評価
谷 亮祐, 藤原 留也, 波多野 賢治
IDRユーザフォーラム2017, (2017)
- レビュー投稿者の日空度を考慮した評判情報の可視化と店舗検索への応用
日高 加菜, 豊田 哲也, 太原 剛三
IDRユーザフォーラム2017, (2017)
- 価値観モデルを用いた協調フィルタリングによるハイブリッド型推薦手法

出典: <https://dsc.repo.nii.ac.jp/>

リクルートデータを利用した研究リスト (一部)

No	研究タイトル	発表先
1	ネットワーク特徴量を用いた経験財の評価手法の提案	サービス学会 第8回国内大会, (2020)
2	Hot Pepper Beautyによる顧客獲得	NII-IDRユーザフォーラム2019, (2019)
3	カスタマーレビューに基づいた美容院業界における地域特性の可視化	NII-IDRユーザフォーラム2019, (2019)
4	Service Switching in Case-based Decisions following Bad Experiences: Online reviews data of Japanese hairdressing salons	The 27th International Conference on Case-Based Reasoning (ICCBR 2019) Workshop Proceedings,74-84 (2019)
5	Graph-based semi-supervised classification for online customer reviews using consensus clustering	IEEE International Conference on Industrial Engineering and Engineering Management (IEEM 2019),1-5 (2019)
6	価値観モデリングを利用した協調フィルタリングにおけるモデル関係学習の提案	ARG 第15回Webインテリジェンスとインタラクショナル研究会 (WI2) 予稿集 [優秀研究賞受賞],15-20 (2019)
7	Introduction of Model Relationship Learning for Matrix-based Collaborative Filtering Employing Personal Values	Proceedings of the 6th International Workshop on Advanced Computational Intelligence and Intelligent Informatics (IWACIII 2019), (2019)
8	美容室のレビューデータを利用した口コミ分析	NII-IDRユーザフォーラム2019, (2019)
9	ランダムフォレストによる“顧客の声”の半自動判別モデル	NII-IDRユーザフォーラム2018, (2018)
10	コンセンサスクラスタリングに基づくグラフベース半教師あり学習	日本品質管理学会 第48回年次大会研究発表会要旨集,79-82 (2018)
11	ランダムフォレストの半教師あり学習による“顧客の声”の分類	感性工学会論文誌,17(5),537-545 (2018)
12	ランダムフォレストを用いたカスタマーレビューの半自動分類	インテリジェントシステムシンポジウム2017予稿集,154-157 (2017)
13	定量分析を用いた自作自演の口コミ判定モデルの構築と評価	NII-IDRユーザフォーラム2017, (2017)
14	レビュー投稿者の甘辛度を考慮した評判情報の可視化と店舗検索への応用	NII-IDRユーザフォーラム2017, (2017)
15	価値観モデルを適用した協調フィルタリングによるハイブリッド型推薦手法	知能と情報 (日本知能情報ファジィ学会誌),29(4),628-636 (2017)
16	レビュー投稿者の甘辛度を考慮した美容院評判情報と利用者情報の可視化	情報処理学会 全国大会講演論文集,2017(1),521-522 (2017)
17	レビューに対する分散表現を用いた評判分析	言語処理学会 年次大会発表論文集,23,198-201 (2017)
18	価値観モデルベース協調フィルタリングのハイブリッド推薦システムへの拡張	ARG 第8回Webインテリジェンスとインタラクショナル研究会 (WI2) 予稿集,21-26 (2016)
19	Proposal of Hybrid Recommender Systems Based on Personal Values-based Collaborative Filtering	Proceedings of the 6th International Symposium on Computational Intelligence and Industrial Applications (ISCIIA2016), (2016)
20	定量分析による自作自演の口コミ判定モデルの構築とその評価	電子情報通信学会 総合大会講演論文集,情報・システム(1),43-43 (2016)

ホットペッパービューティーのクチコミデータ



データ概要

国立情報学研究所が株式会社リクルートテクノロジーズから提供を受けて研究者に提供しているデータセットです。

1. ホットペッパービューティーデータ 2014-09-01 新規

ホットペッパービューティーに2012年1月～2014年1月に掲載されたデータです。各種IDはユニーク番号に変換済みで、各データ間はユニーク番号により紐付けが可能です。

- 店舗データ (約8万件)
店舗名, 住所, 店舗メッセージ, ホームページなどのデータです。TSV形式のファイルで約45MB (展開後) です。
- 店舗ブログデータ (約361万件)
タイトル, 本文, 登録日などのデータです。TSV形式のファイルで約2.4GB (展開後) です。
- スタ일리ストデータ (約18万件)
キャリア年数, 得意技術, 自己PR, 性別などのデータです。TSV形式のファイルで約64MB (展開後) です。
- クーポンデータ (約15万件)
クーポン名, 利用条件, 有効期限, 料金などのデータです。TSV形式のファイルで約56MB (展開後) です。
- メニューデータ (約52万件)
メニュー名, 料金, 施術時間などのデータです。TSV形式のファイルで約53MB (展開後) です。
- セットメニューデータ (約7万件)
メニュー名, 料金, 施術時間などのデータです。TSV形式のファイルで約9MB (展開後) です。
- 口コミデータ (約73万件)
ニックネーム, 性別, 世代, レビュー内容, 個別および総合ポイント, 投稿日などのデータです。TSV形式のファイルで約465MB (展開後) です。

例：カスタマーレビューに基づいた美容院業界の地域特性

カスタマーレビューに基づいた美容院業界における地域特性の可視化

岩崎友秋, 永井一平, 三浦和也, 齊藤史哲 (千葉工業大学)

背景と目的

厚生労働省が発表した平成 29 年度衛生行政報告例によると、ヘアサロンの店舗数は年々増え続け 2019 年には約 25 万件に達している [1]。これはコンビニエンスストアの店舗数の約 4 倍と比べて競争が激しいことがうかがえる。こうした**オーバーストア状態**は、客の奪い合いに起因する利益低下の要因となっている [2]。しかしオーバーストア状態なのにも関わらず店舗が一定の地域に偏っていたり、近年高い技術を求めて都内にわざわざ遠方から訪れる人がいること [3] から美容院業界内で地域格差が生じている。

本研究では**地理的要因**と**カスタマーレビュー**から**顧客満足度**につながる総合評価の関係性に着目する。そしてカスタマーレビューを解析し美容院業界ごとの**地域や店舗のニーズ**を可視化することで**地域特性を明らかにする**。

ホットペッパービューティーデータセット

2012 年 1 月～2014 年 1 月の期間に掲載されたもの

本研究で用いたデータセット

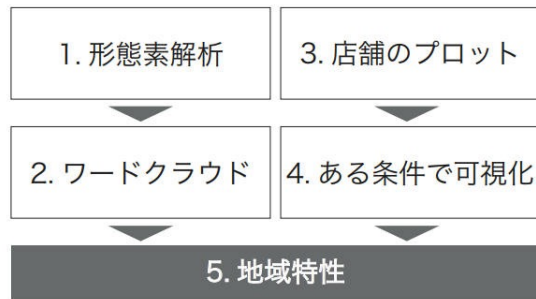
- ・店舗データ：約 10,000 件
 - キャッチコピー
 - 位置情報
- ・クチコミデータ：約 360,000 件
 - 5 段階評価（総合評価・テクニックなど）
 - カスタマーレビュー
 - 世代・性別などのデモグラフィクス



例：カスタマーレビューに基づいた美容院業界の地域特性

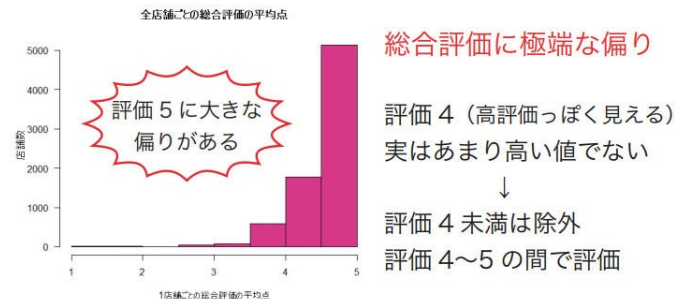
提供データに対する分析の事前調査

分析方法



1. キャッチコピーとカスタマーレビューを形態素解析
2. 地域ごとにワードクラウドで可視化
3. 都内の店舗を地形図にプロット
総合評価4未満は除外
4. レビュー数に比例して半径を大きく
総合評価5に近いほど明るく
5. 地域特性の可視化

総合評価をプロットして気づいたこと

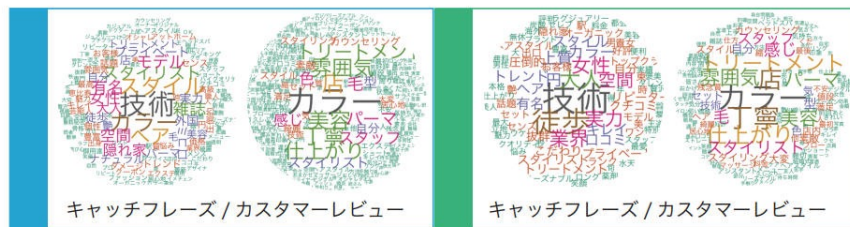


分析結果



渋谷区

中央区



平均年齢	25歳前後	35歳前後
ターゲット	若者	大人の女性
評価	甘口	辛口
店舗の売り	人気モデルに合わせたカット	徒歩圏内

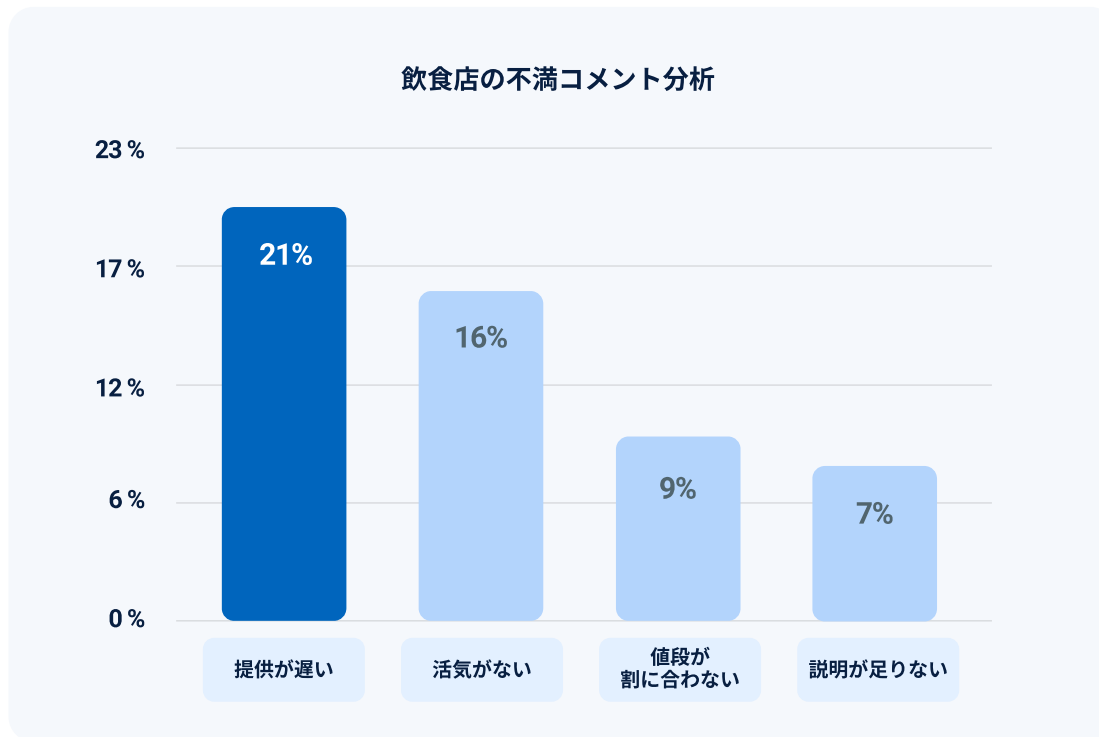
データ活用事例

事例：データサイエンティスト × 飲食領域

- 飲食店の調理順序を最適化（生産技術を応用）
 - 料理の提供遅れを防ぐ（クレームを減らす）

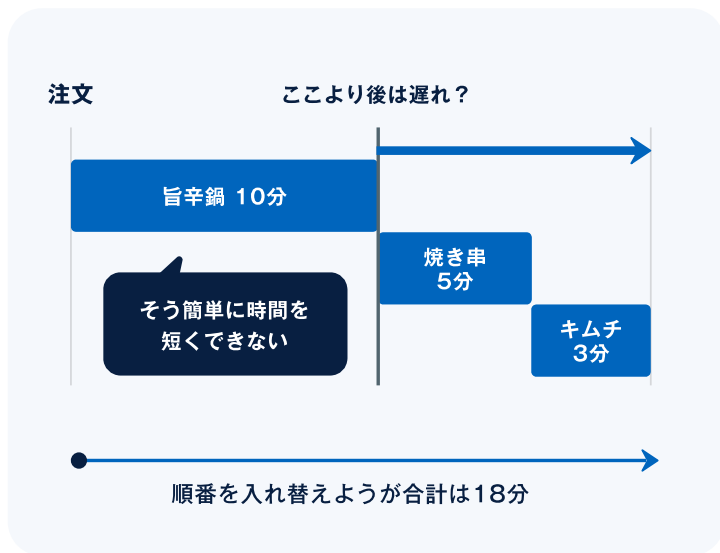


■ 飲食店クレーム要因のトップは“提供遅れ”

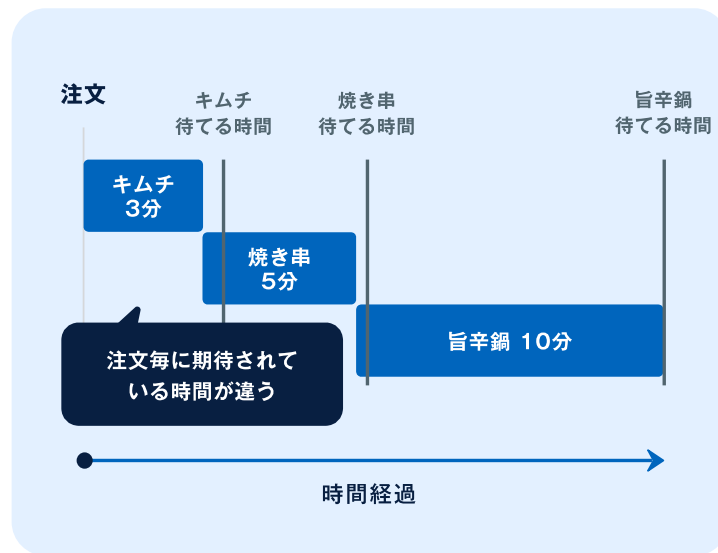


事例：データサイエンティスト × 飲食領域

- 「調理にかかる時間の総和を短くする」問題ではなく、
- 「調理遅れ時間の総和が最小化されるような調理順を求める」形に



合計時間が変わらないのならば
調理遅れは減らせない?



注文毎の待てる時間に合わせて
注文順を決めることで遅れを減らせる!

事例：データサイエンティスト × 飲食領域

■Airレジ オーダー『キッチンモニター』画面

注文を自動で最適な順番に並び変え



新 A-1 2名	
メイン	1
前菜	1
デザート	1
生ビール	2

0分経過
青空 太郎

すべて調理済

新 A-1 | 2名

生ビール	2
前菜	1
メイン	1
デザート	1

0分経過
青空 太郎

すべて調理済

目標提供時間を設定することで、調理しやすい順に自動で並ぶので、伝票を並べ替える手間もなく、提供遅れ防止に役立ちます。

提供が遅れそうな注文を色でお知らせ



新 A-1 4名	
シーザーサラダ	1
アスバラ串	2
ねぎま(塩)	2
チキン南蛮(ハーフ)	2

11分経過
青空 太郎

すべて調理済

商品ごとに設定した注意時間をすぎると、提供アラートが、色でお知らせするので提供が遅れる前に気づくことができます。

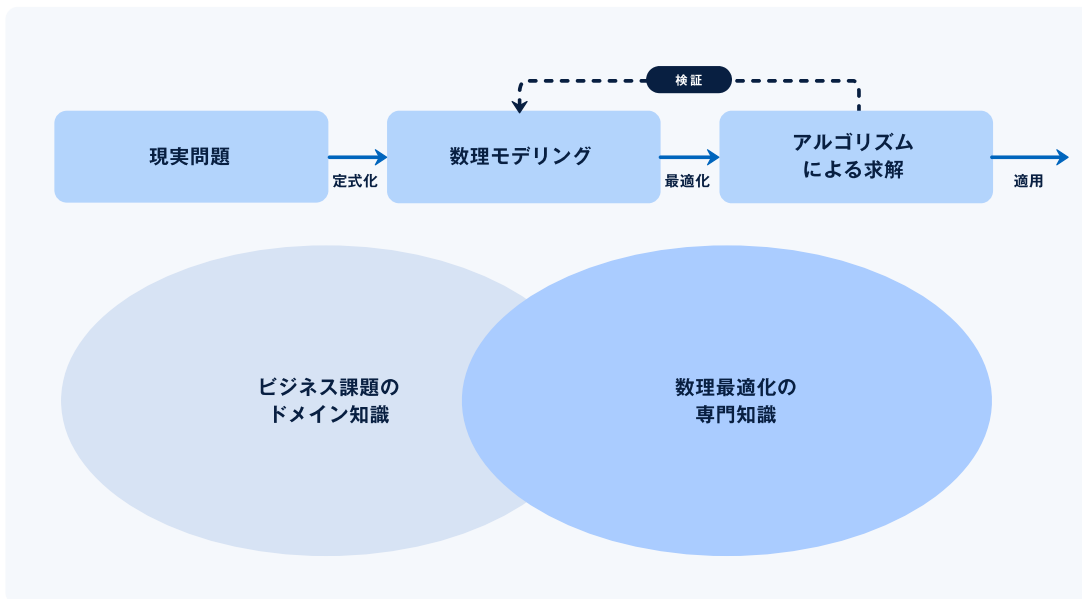
事例：データサイエンティスト × 飲食領域

- 「1機械スケジューリング問題」全てのジョブを1台の機械で処理する場合に、ジョブの処理順序を決定する問題を応用して課題解決



事例：データサイエンティスト × 飲食領域

- 最適化の入力データ「調理所要時間」「お客さまが待てる料理提供時間」は、直接計測することは難しいため、過去に取得したデータと機械学習や統計モデルを利用した手法を用いて、これらの値を推定、ロジックの磨きこみを実施
- 協力していただける店舗でのテストやヒアリングなどを繰り返し実施して開発



さいごに

より詳細についてはぜひ各組織のブログにて…！

The screenshot shows the homepage of the RECRUIT Recruit Data Blog. The header includes the site name and navigation links for categories, tags, and a search function. The main content area features a grid of article thumbnails with titles and dates. A right sidebar contains a tag cloud with various technical terms like 'BigQuery', 'AWS', and 'GCP'.

データ

The screenshot shows the homepage of the RECRUIT 1996+ blog. The header features a navigation menu with 'ホーム', '記事', 'マガジン', '採用サイト', and '月別'. The main content area includes a featured article with a large blue graphic and a search bar. A right sidebar contains a list of recent posts.

プロダクト
デザイン

The screenshot shows an article from Hatena Blog titled 'SLOの監視に Burn Rate Monitoring を導入してみた話'. The article is dated 2022-12-01 and is categorized under 'Engineering' and 'SRE'. The content discusses the implementation of Burn Rate Monitoring for SLO monitoring.

エンジニア

スデテサマリ

The screenshot shows the homepage of the Advanced Technology Lab BLOG. The header includes navigation links for 'HOME', 'BLOG', 'ABOUT', 'PROJECTS', 'BLOG', 'LIBRARY', 'RECRUIT', 'JP / EN'. The main content area features a large 'BLOG' heading and a grid of article thumbnails. A right sidebar contains a list of tags.

R&D

リクルート エンジニア ブログ

検索

ポスターブースでお待ちしています！

Case02

飲食店の調理順序を最適化し

料理の提供遅れ防止を実現

データサイエンティスト × 業務・経営支援



Case01

機械学習モデルによる

予測値を利用し、カスタマーの

新ジャンル利用数を最大化

データサイエンティスト × ビューティ



Case03

対話型チャットボットで実現する

旅行代理店のような検索体験

機械学習エンジニア × 旅行



ポスターブースでお待ちしています！

リクルート統合後初の大型技術カンファレンス

「RECRUIT TECH CONFERENCE 2024」開催

2024年2月21日（水）、オンライン開催、参加費無料



2024/2
大規模な
技術カンファレンス
開催します！

就活生対象！市場調査 兼 キャリア支援面談のご案内

市場調査に
協力頂ける方を
募集中！

概要

- ・ IT人材の育成やキャリアイメージ醸成機会の検討のため、市場調査に協力いただける方を募集します
 - ・ 話を聞かせていただいた上で、キャリア選択についてのアドバイスや相談を個別にお受けします
- ※ご協力いただいた方には先着で弊社オリジナルのノベルティグッズをプレゼント！

実施形式

12/12(火)~12/22(金) ※先着20名・オンライン形式・30分~1時間

こんな方におすすめ

- ・ 研究経験や研究を通して学んだことを活かしたいが、キャリアの選び方に悩んでいる方
- ・ 自分が携わってきた研究分野や技術がビジネスの場でどのように活かせるかを知りたい方

引き続きIDRユーザフォーラムを
楽しんでいきましょう！