

社会人のためのデータサイエンス演習

第5週:ビジネスでデータサイエンスを実現するために

第1回:各週のおさらい

講師名:菅 由紀子

講座内容

第1週

- データサイエンスとは

第2週

- 分析の概念と事例
ビジネス課題解決のためのデータ分析基礎(事例と手法)①

第3週

- 分析の具体的手法
ビジネス課題解決のためのデータ分析基礎(事例と手法)②

第4週

- ビジネスにおける予測と分析結果の報告
ビジネス課題解決のためのデータ分析基礎(事例と手法)③

第5週

- ビジネスでデータサイエンスを実現するために

第5週の内容紹介

第1回

- 各週のおさらい

第2回

- データ分析に基づく問題解決ケーススタディ①

第3回

- データ分析に基づく問題解決ケーススタディ②

第4回

- 様々な組織で活躍するデータサイエンティスト

第5回

- 組織におけるデータサイエンスの実現

第6回

- 講座のまとめ①

第7回

- 講座のまとめ②

第1週のまとめ

データの取得・処理性能とともに進化。新たなAIサービスの誕生

人々の生活のさまざまな局面で
データサイエンスが活用されるようになる



AI



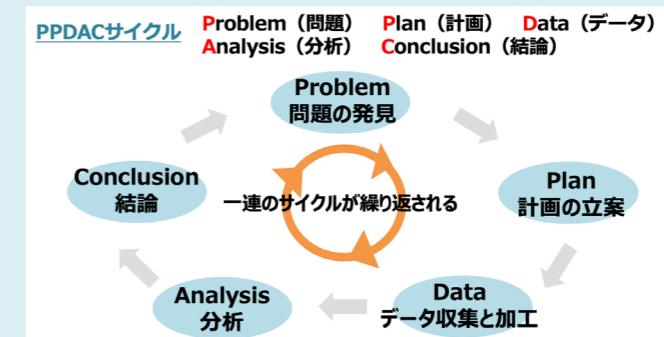
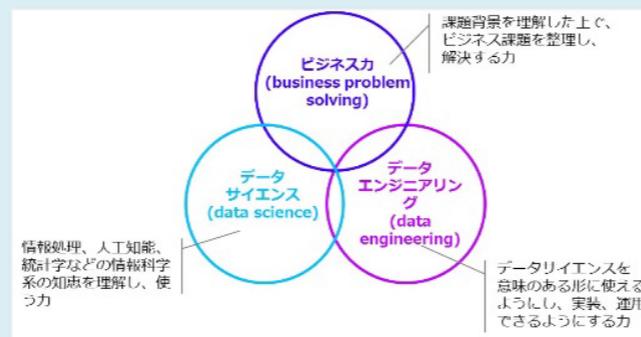
渋滞予測

需要予測

自動運転

買い物支援

データサイエンティストに求められるスキルレベルがより具体化した



第2週のまとめ

Analysisの具体的手法 分布 / 比較 / 傾向

KGIとそれに関連する要因の関係を分析する

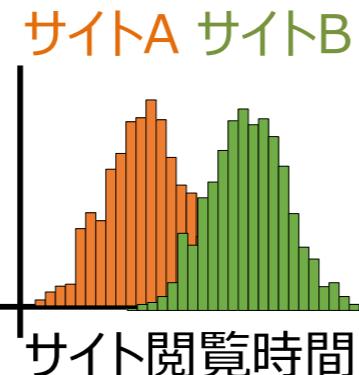
比較

名義 × 名義

来客数	男	女
サイトA	18	3
サイトB	4	16

クロス集計

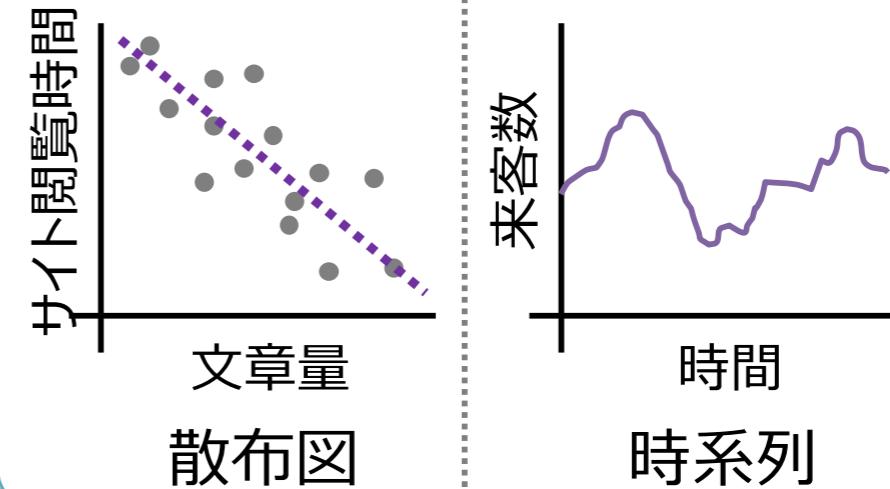
名義 × 比例



ヒストグラム

傾向

比例 × 比例



時系列

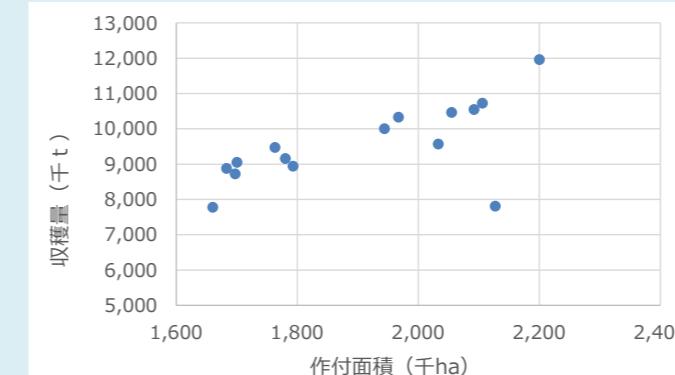
第3週のまとめ

比較と傾向の分析

クロス集計 / 散布図 / 相関 / 時系列

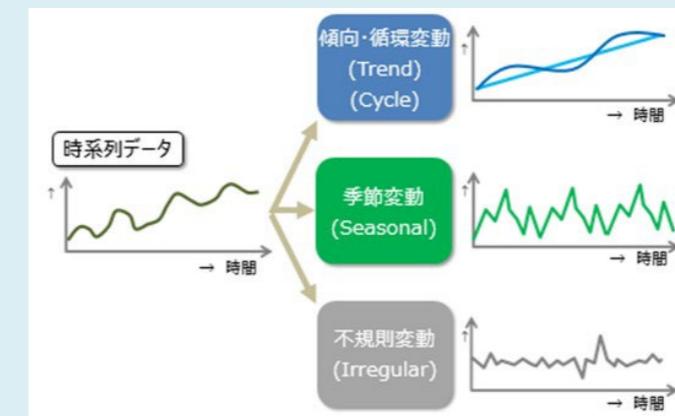
- ・可視化が重要

- ✓ 変数の関係性の把握
- ✓ 外れ値の除外



- ・時系列は分解して考える

- ✓ 解釈が容易に
- ✓ 将来予測が可能



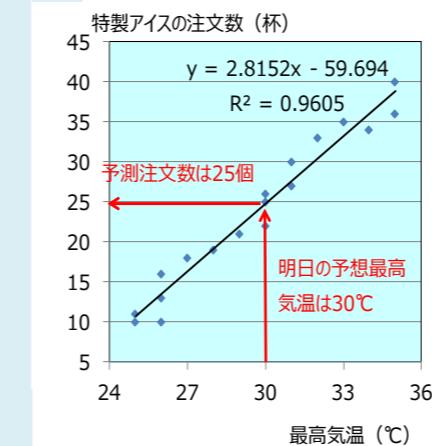
第4週のまとめ

Analysis応用編

予測 / 分析結果の報告 / 機械学習

- ・単回帰での将来予測

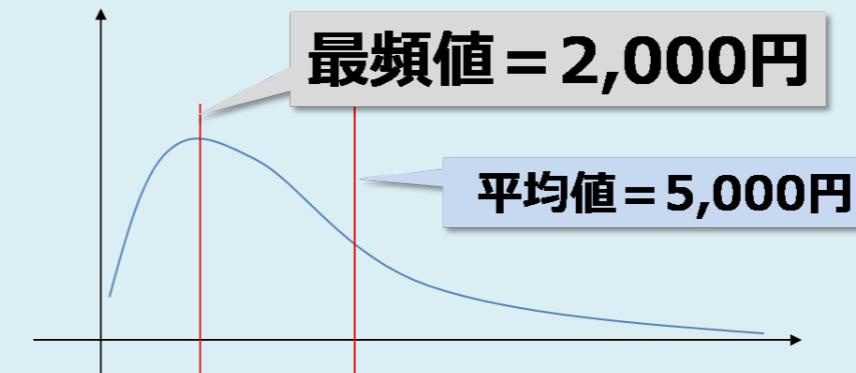
- ✓ 概念とエクセルでの実施



- ・分析報告の落とし穴

- ✓ 前提や指標を適切に提示

機械学習の応用先は幅広い



次回のテーマ

次回は

「データ分析に基づく問題解決ケーススタディ①」

お疲れ様でした！



総務省統計局

社会人のためのデータサイエンス演習

第5週:ビジネスでデータサイエンスを実現するために

第2回:データ分析に基づく問題解決ケーススタディ①

講師名:小野 陽子

第5週の内容紹介

第1回

- 各週のおさらい

第2回

- データ分析に基づく問題解決ケーススタディ①

第3回

- データ分析に基づく問題解決ケーススタディ②

第4回

- 様々な組織で活躍するデータサイエンティスト

第5回

- 組織におけるデータサイエンスの実現

第6回

- 講座のまとめ①

第7回

- 講座のまとめ②

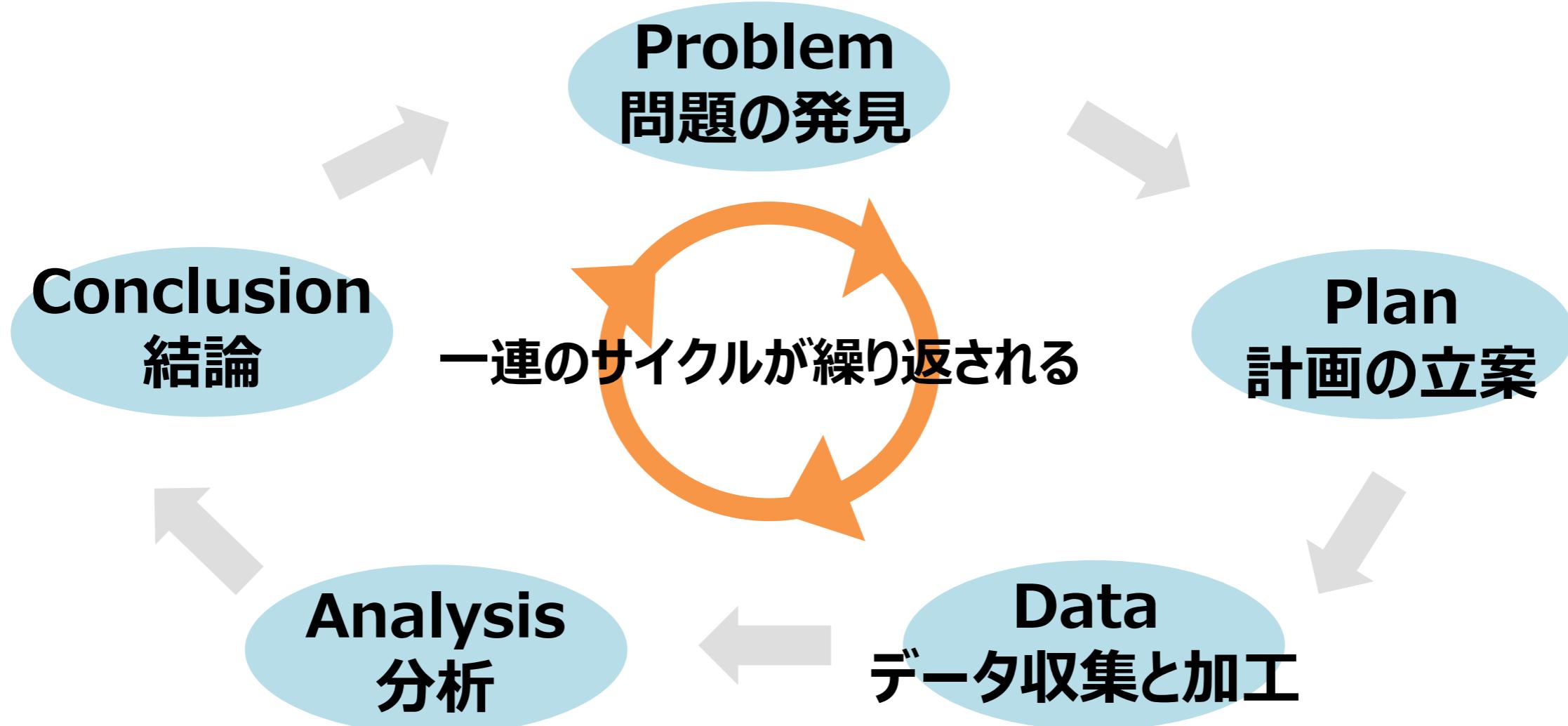
復習：課題解決のサイクル

PPDACサイクル

Problem (問題)
Analysis (分析)

Plan (計画)
Conclusion (結論)

Data (データ)



データ分析に基づく問題解決ケーススタディ

Problem(問題の発見)

あなたはある飲食店のオーナーです。
お店の利益の低下が悩みです。

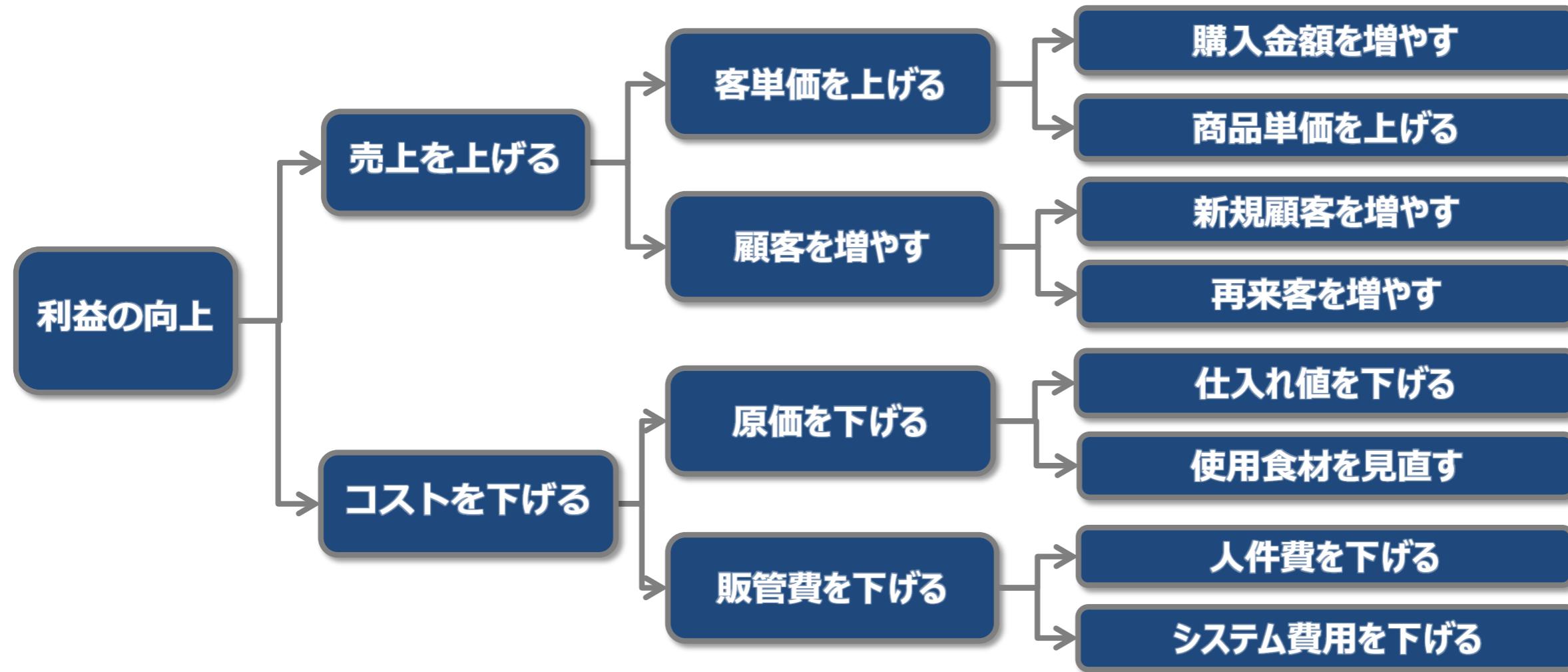
今後の継続的な店舗運営のために
お店の利益を増やしたいと考えています。



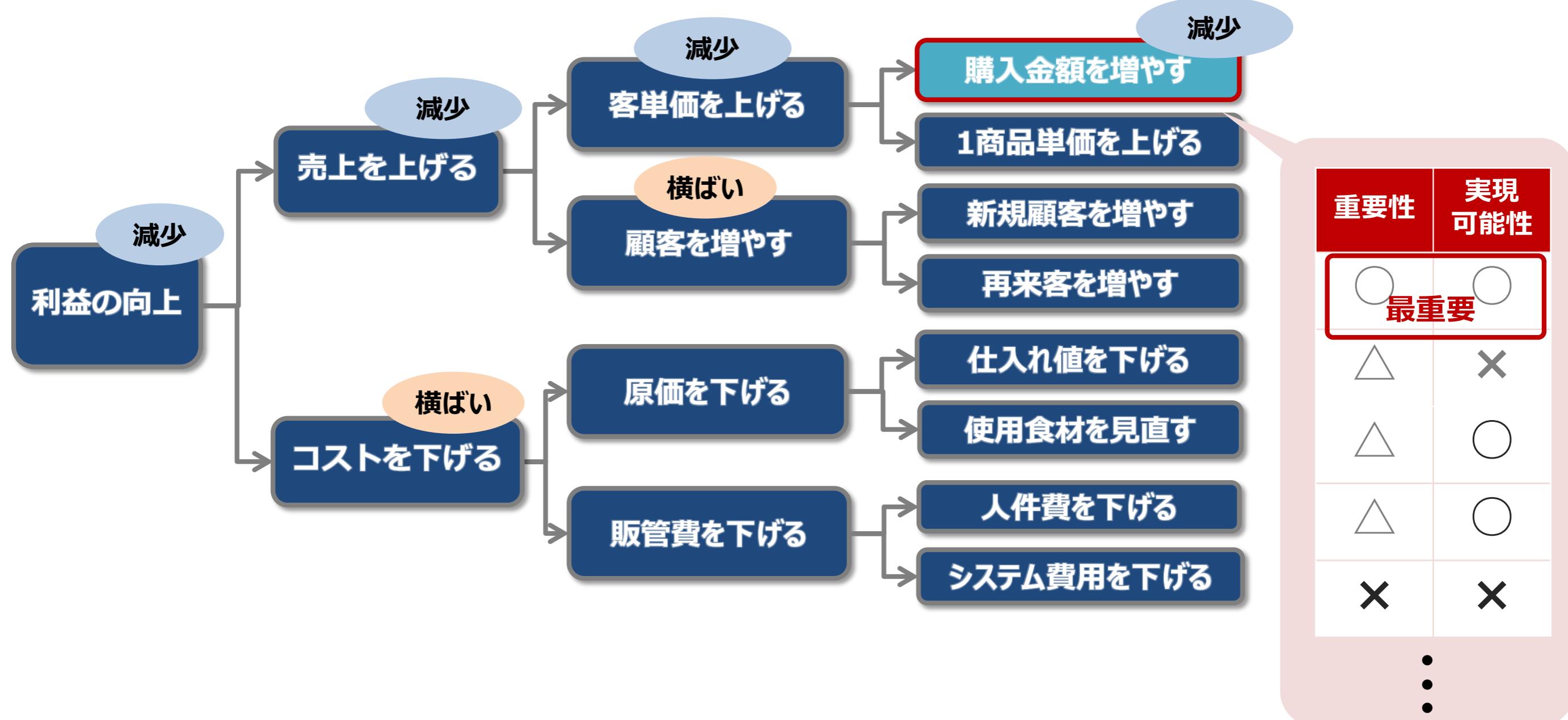
課題： 利益の向上

課題解決のためのKGI：店舗の利益

データ分析に基づく問題解決ケーススタディ Plan(課題の要素分解と課題の優先順位付け)



データ分析に基づく問題解決ケーススタディ Plan(課題の要素分解と課題の優先順位付け)



データ分析に基づく問題解決ケーススタディ Plan(KGI達成のための因果の仮定の設定)

購入金額を増やす

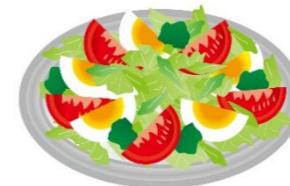
「クーポン」を使ったA/Bテスト実施

仮説①

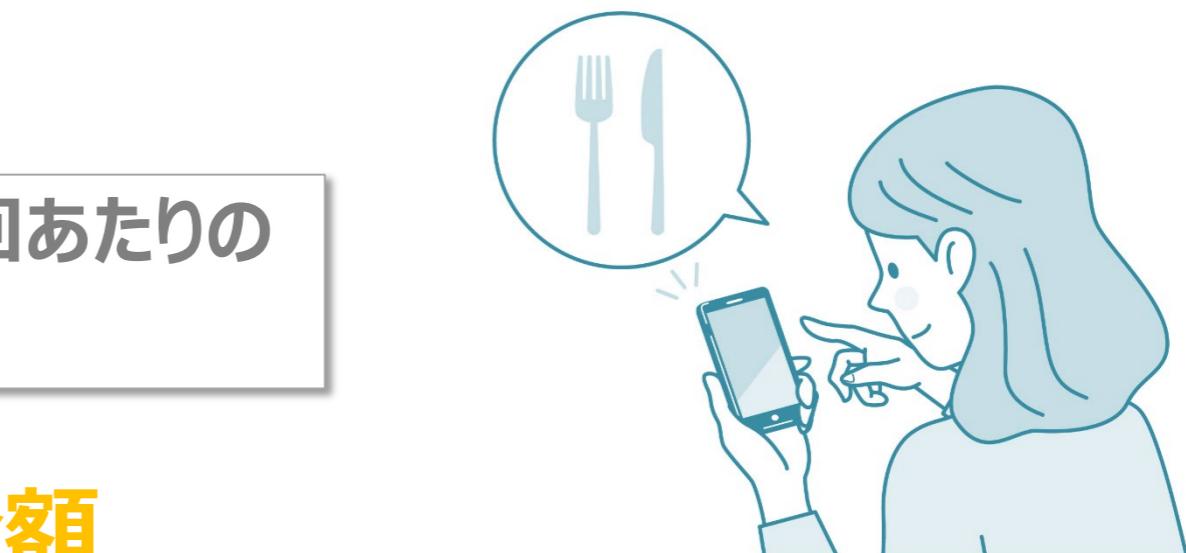
クーポンの発行有無によって、1回あたりの平均注文金額が異なる

特別クーポンのご案内

以下のサイドメニューがお得！



KPI：1回あたりの平均購入金額



データ分析に基づく問題解決ケーススタディ Plan(計画まとめ)

分析 アプローチ 方法

- ・売上データから設定したKPIの現状を知り、どのようなA/Bテストを実施するか決定する
- ・A/Bテストの効果検証を行い、次に打つべき対策を検討する

データ

- ・注文実績データ（顧客ID / 注文ID / 注文日時 / 商品ID / 単価 / 注文数）
- ・商品マスター（商品ID / 商品名 / カテゴリ / 単価）
- ・会員属性データ
- ・天候データなど

データ分析に基づく問題解決ケーススタディ Data(データの収集・加工)

データ収集方法の検討



- 自社で取得が可能かどうか
(データの所在、依頼すべき部門の確認)
- 外部で手配するデータの確認
公的統計などのオープンデータや
購入が必要な商用データ

データの粒度、期間

- 分析に必要なデータ量が
十分に揃っているかどうか、
期限までに取得が可能か
どうか

分析に必要なデータが現時点で揃わない場合は、
新たに取得することも検討する

データ分析に基づく問題解決ケーススタディ Data(分析軸の作成①)



注文実績データ

注文日時	顧客ID	商品ID			
2020/4/10 11:25	00003	A0005			
2020/4/10 12:42	00003	B0001	620	1	
2020/4/10 12:43	02451	D00010	1,280	1	
2020/4/11 11:05	00101	A0003	1,008	2	
2020/4/11 11:06	00218	XX00001	1,560	1	
2020/4/11 11:15	00218	A0001	796	2	
2020/4/11 11:23	00714	XX0005	1,350	1	

共通のIDを使って各データセットの紐づけが可能

注文実績に、商品名や顧客の詳細データを掛け合わせることで、年代や性別ごとの注文傾向など、より詳細に、データを理解することができる。

会員属性データ

顧客ID	登録日	性別	生年月日
00001	2016/5/1	男	1978/5/10		
00002	2016/5/2	女	1982/9/15		
00003	2016/5/2	女	1964/2/24		
00004	2016/5/2	男	1999/10/29		
00005	2016/5/4	女	2001/2/8		
...					

商品マスター

商品ID	商品名	カテゴリ	単価
A0005	ガトーショコラ	デザート	604
B0001	オムライス	メイン	535
D00010	抹茶スフレ	デザート	610
A0003	トマトソースパスタ	パスタ	1,280
XX00001	カフェオレ	ドリンク	504
...			

データ分析に基づく問題解決ケーススタディ Data(分析軸の作成②)

注文実績データ

注文日時	顧客ID	商品ID	単価	注文数	・・・
2020/4/10 11:25	00003	A0005	504	1	
2020/4/10 12:42	00003	B0001	620	1	
2020/4/10 12:43	02451	D00010	1,280	1	
2020/4/11 11:05	00101	A0003	1,008	2	
2020/4/11 11:06	00218	XX00001	1,560	1	
2020/4/11 11:15	00218	A0001	796	2	
2020/4/11 11:23	00714	XX0005	1,350	1	

KPI項目に影響を与える要素について、
あらかじめ集計・追加を行う

会員属性データ

顧客ID	登録日	性別	生年月日
00001	2016/5/1	男	1978/5/10
00002	2016/5/2	女	1982/9/15
00003	2016/5/2	女	1964/2/24
00004	2016/5/2	男	1999/10/29
00005	2016/5/4	女	2001/2/8

来店回数	合計購入数	合計売上	平均購入金額	平均購入数
3	13	¥3,279	¥1,093	4
2	16	¥6,512	¥3,256	8
5	13	¥10,523	¥2,105	3
4	8	¥7,328	¥1,832	2
1	3	¥2,690	¥2,690	3

次回のテーマ

次回は

「データ分析に基づく問題解決ケーススタディ②」

お疲れ様でした！



総務省統計局

社会人のためのデータサイエンス演習

第5週:ビジネスでデータサイエンスを実現するために

第3回:データ分析に基づく問題解決ケーススタディ②

講師名:小野 陽子

第5週の内容紹介

第1回

- 各週のおさらい

第2回

- データ分析に基づく問題解決ケーススタディ①

第3回

- **データ分析に基づく問題解決ケーススタディ②**

第4回

- 様々な組織で活躍するデータサイエンティスト

第5回

- 組織におけるデータサイエンスの実現

第6回

- 講座のまとめ①

第7回

- 講座のまとめ②

データ分析に基づく問題解決ケーススタディ

Problem(問題の発見)

あなたはある飲食店のオーナーです。
お店の利益の低下が悩みです。

今後の継続的な店舗運営のために
お店の利益を増やしたいと考えています。



課題解決のためのKGI : 店舗の利益

データ分析に基づく問題解決ケーススタディ Plan(KGI達成のための因果の仮定の設定)

購入金額を増やす

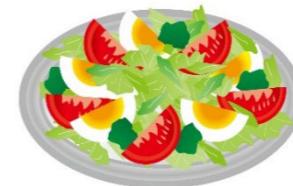
「クーポン」を使ったA/Bテスト実施

仮説①

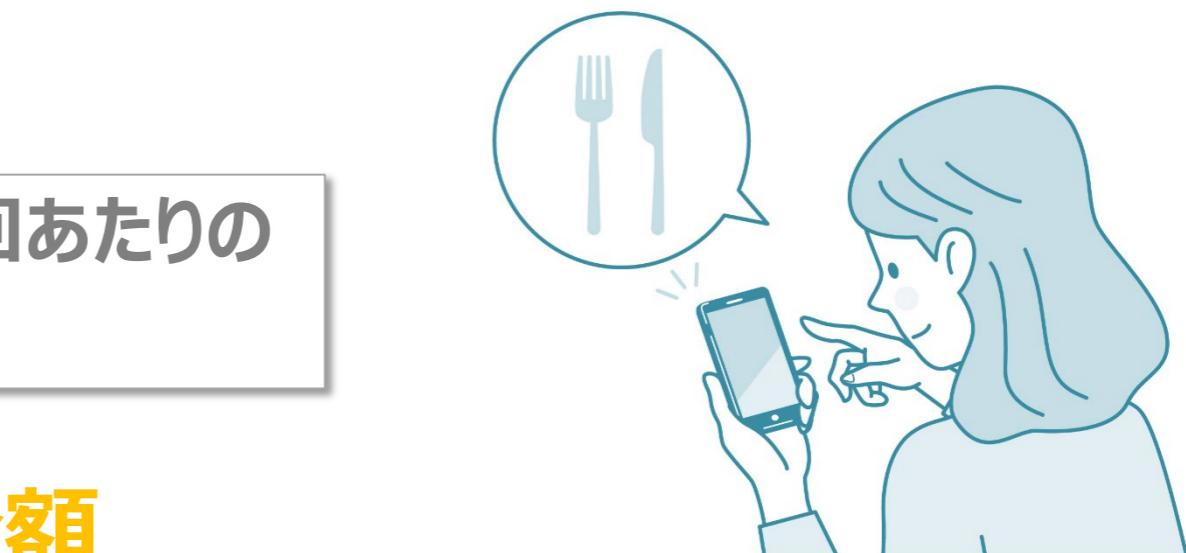
クーポンの発行有無によって、1回あたりの平均注文金額が異なる

特別クーポンのご案内

以下のサイドメニューがお得！



KPI：1回あたりの平均購入金額



データ分析に基づく問題解決ケーススタディ Plan(計画まとめ)

分析 アプローチ 方法

- ・売上データから設定したKPIの現状を知り、どのようなA/Bテストを実施するか決定する
- ・A/Bテストの効果検証を行い、次に打つべき対策を検討する

データ

- ・注文実績データ（顧客ID / 注文ID / 注文日時 / 商品ID / 単価 / 注文数）
- ・商品マスター（商品ID / 商品名 / カテゴリ / 単価）
- ・会員属性データ
- ・天候データなど

データ分析に基づく問題解決ケーススタディ Plan(計画まとめ)

売上実績データから設定したKPIの現状を知り、どのようなA/Bテストを実施するか決定する

現状把握

- ① 売上、平均購入金額の推移
- ② 平均購入金額が減少した顧客グループの特定
→ クーポン配布対象の決定

A/Bテスト 設計

- ① クーポンの利用条件の設定
- ② クーポンの内容の設定

A/Bテストの効果検証を行い、次に打つべき対策を検討する

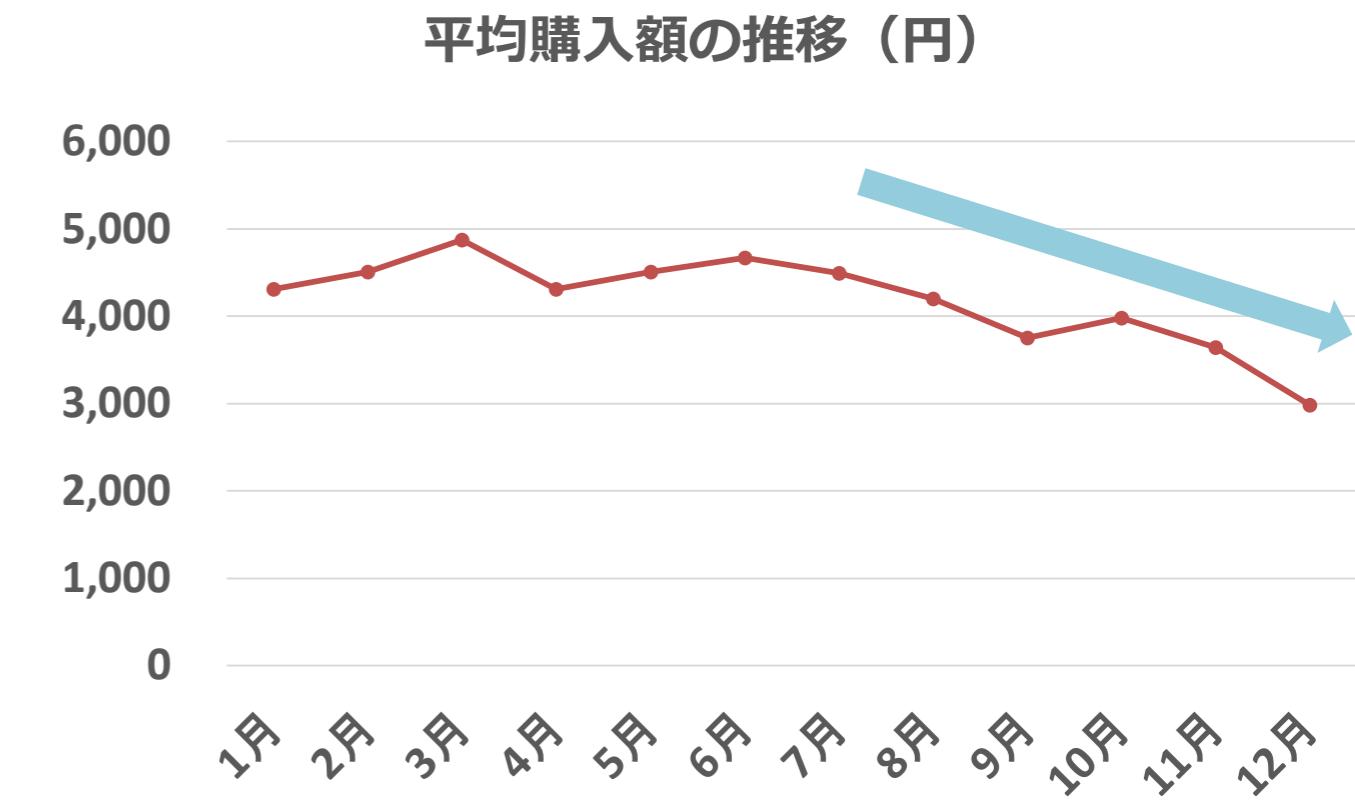
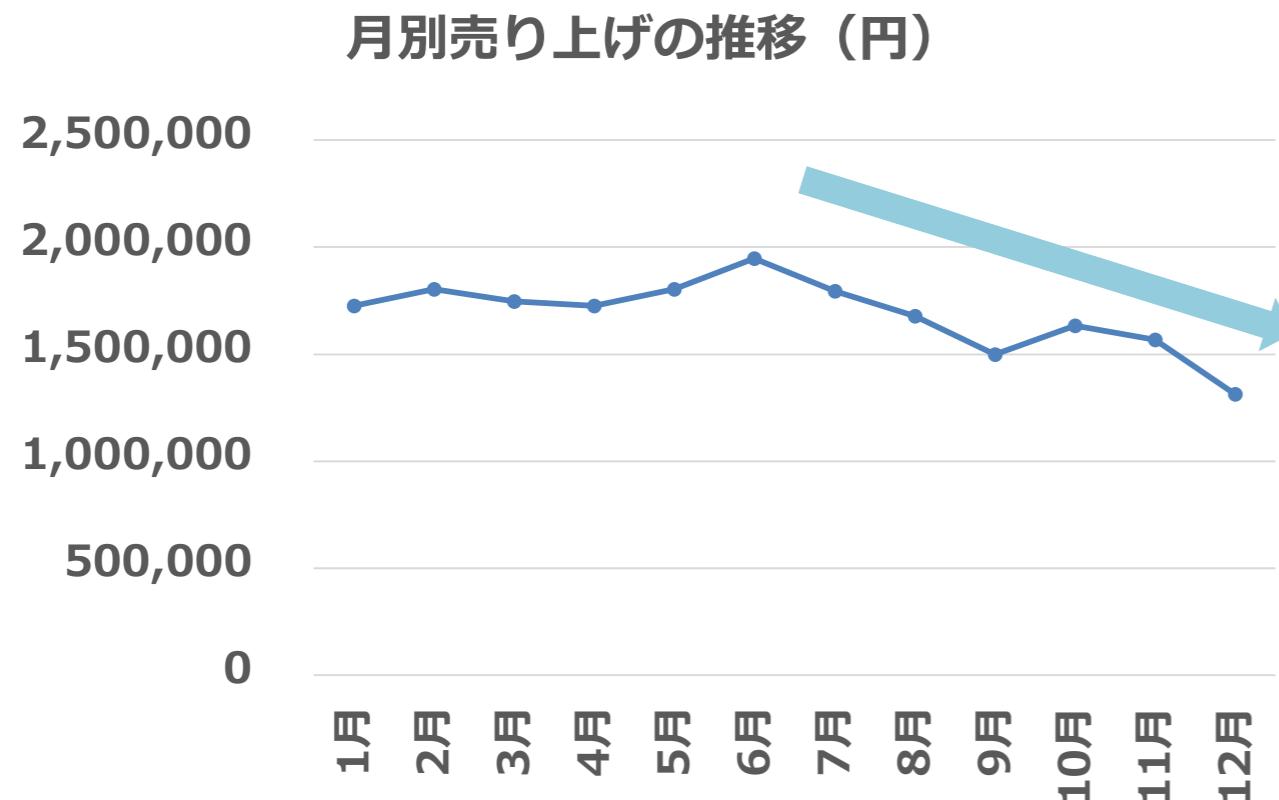
クーポン配布するグループとしないグループに分け

クーポン配布前後の平均購入金額の変化を比較する

データ分析に基づく問題解決ケーススタディ

現状把握（全体）

①売上、平均購入金額の推移



両者とも6月から減少傾向

データ分析に基づく問題解決ケーススタディ 現状把握（属性別）

②平均購入金額が減少した顧客グループの特定

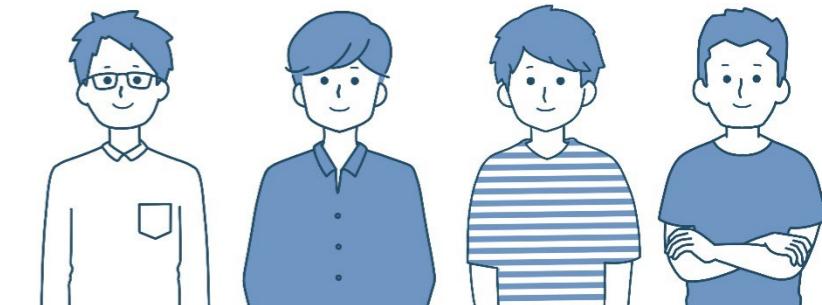
顧客の属性データ
(顧客ID,性別,年代)

注文実績データ
(顧客ID,注文日時,金額)

平均購入金額変化率（6月と12月）

	男性	女性
~10代	-10%	-5%
20代	-40%	-20%
30代	-10%	-8%
40代	-8%	-9%
50代	-20%	-10%
60代	6%	5%
70代～	3%	-2%

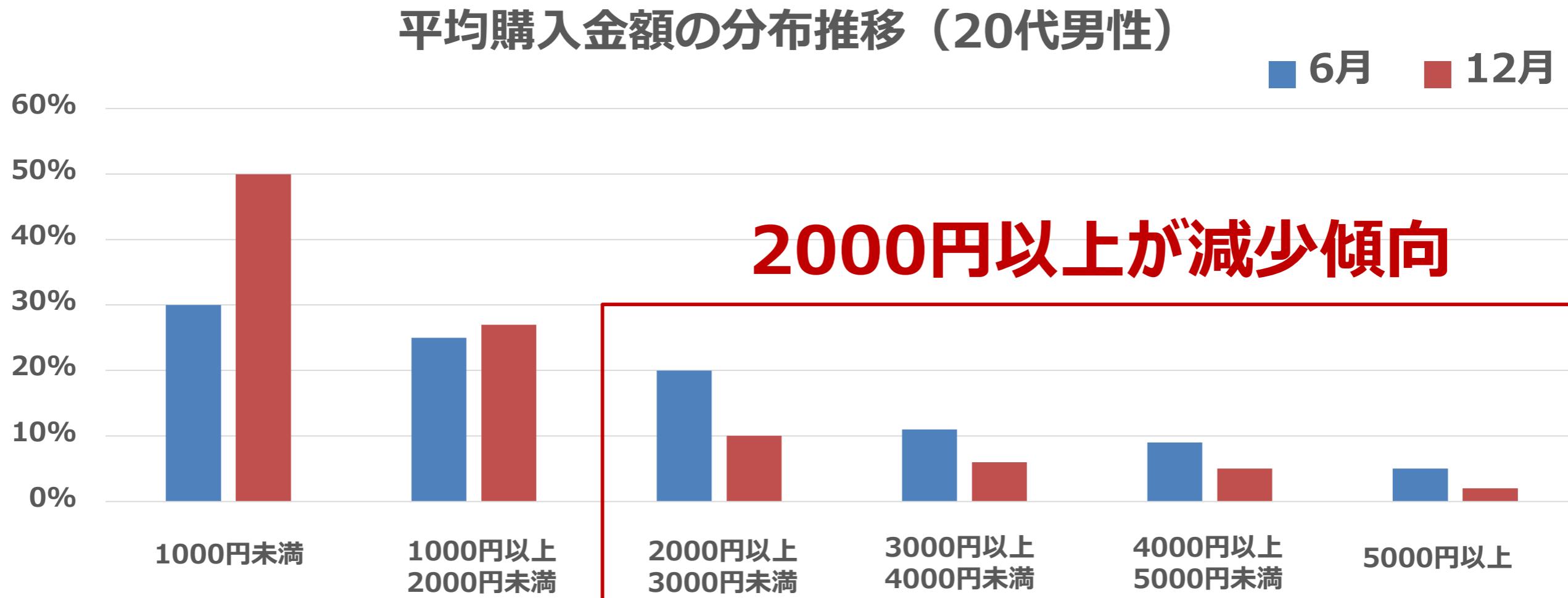
20代男性の
減少が最も大きい



データ分析に基づく問題解決ケーススタディ A/Bテスト設計

①クーポン利用条件の設定

20代・男性の平均購入金額の内訳を把握



データ分析に基づく問題解決ケーススタディ A/Bテスト設計

②クーポンの内容設定

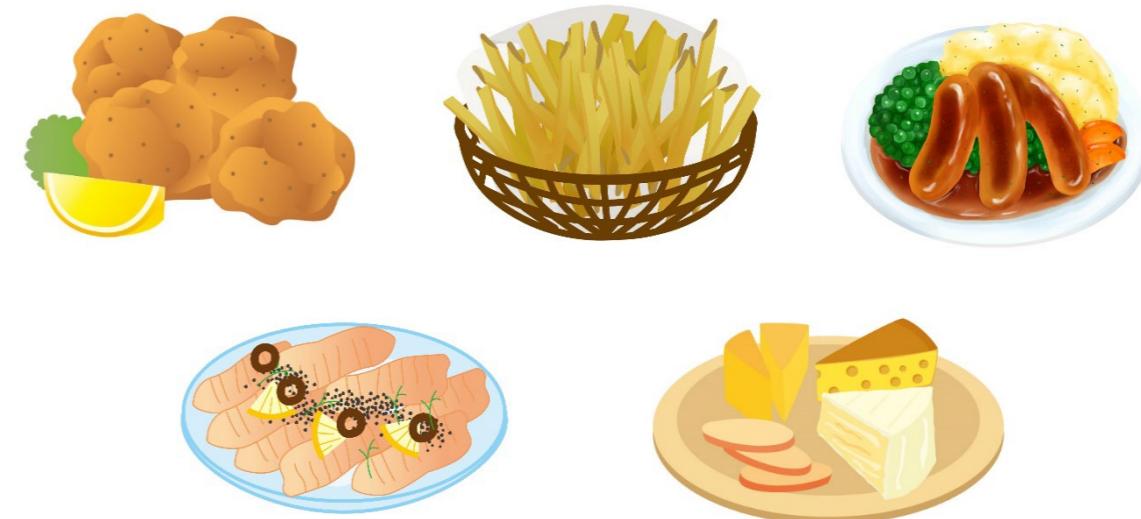
20代・男性に人気のサイドメニューを特定する

メニュー別注文数ランキング[月間]

売上順	メニュー名	価格	平均注文数(月間)
1	唐揚げ	498	160
2	フライドポテト	420	153
3	グリルソーセージ	528	136
4	サーモンのカルパッチョ	1280	107
5	チーズ盛り合わせ	870	101
6	ポテトサラダ	680	94
7	自家製ピクルス	459	88
8	スペイシーチキン	680	180
9	生ハムサラダ	580	126
10	カブレーゼ	688	108

2,000円以上注文で利用可能

以下のメニューがどれでも30%オフ



データ分析に基づく問題解決ケーススタディ A/Bテストの効果検証

A/Bテストの効果を検証する

20代男性のクーポン配布有無による変化を見る

		クーポン非配布 グループ	クーポン配布 グループ
A	平均購入額 (購入額÷合計来店回数)	1040	1248
B	平均購入点数 (購入点数÷合計来店回数)	2.60	3.38

1.2倍

1.3倍

クーポン配布グループの方が、非配布グループに
比べて平均購入額が1.2倍高い結果となった

データ分析に基づく問題解決ケーススタディ 結果・結論

分析結果

現状把握

- 直近半年間の売上、平均購入金額は減少している。
- 特に、20代男性の減少率が最も大きい。

A/Bテスト 設計

- 20代男性で特に減少率が高い購入金額2,000円を対象に
対象者に人気のメニュー上位5品を割引提供。
- 顧客グループを2つに分けて平均購入金額の変化を見る

A/Bテストの効果検証

- クーポン配布グループの方が平均購入額が高い

(クーポン配布効果1.2倍)

今後のアクションプラン

- クーポン配布ならびに効果検証の継続

次回のテーマ

次回は

「様々な組織で活躍するデータサイエンティスト」

お疲れ様でした！



総務省統計局

社会人のためのデータサイエンス演習

第5週:ビジネスでデータサイエンスを実現するために

第4回:様々な組織で活躍するデータサイエンティスト

講師名:石川 恵理香

第5週の内容紹介

第1回

- 各週のおさらい

第2回

- データ分析に基づく問題解決ケーススタディ①

第3回

- データ分析に基づく問題解決ケーススタディ②

第4回

- **様々な組織で活躍するデータサイエンティスト**

第5回

- 組織におけるデータサイエンスの実現

第6回

- 講座のまとめ①

第7回

- 講座のまとめ②

データサイエンティストの活躍の場



医療・製薬

罹患予測、画像診断結果などから、治療や診察の精度を向上させる。



製造業

工場の機械の稼働データを把握し、分析することで、機械トラブルの予防保全や物流・人員の最適化につなげる



金融業

市場調査や分析を行い、金融商品の評価や予測をデータに基づき実施。投資家向けにアドバイスを行う。



エンターテイメント

イベント施設の入場者データを活用し、運営の最適化や混雑緩和施策を実施する。



スポーツ

試合映像や結果データを用いて、選手のパフォーマンス分析をしたり、試合における有効な戦略を立案する。

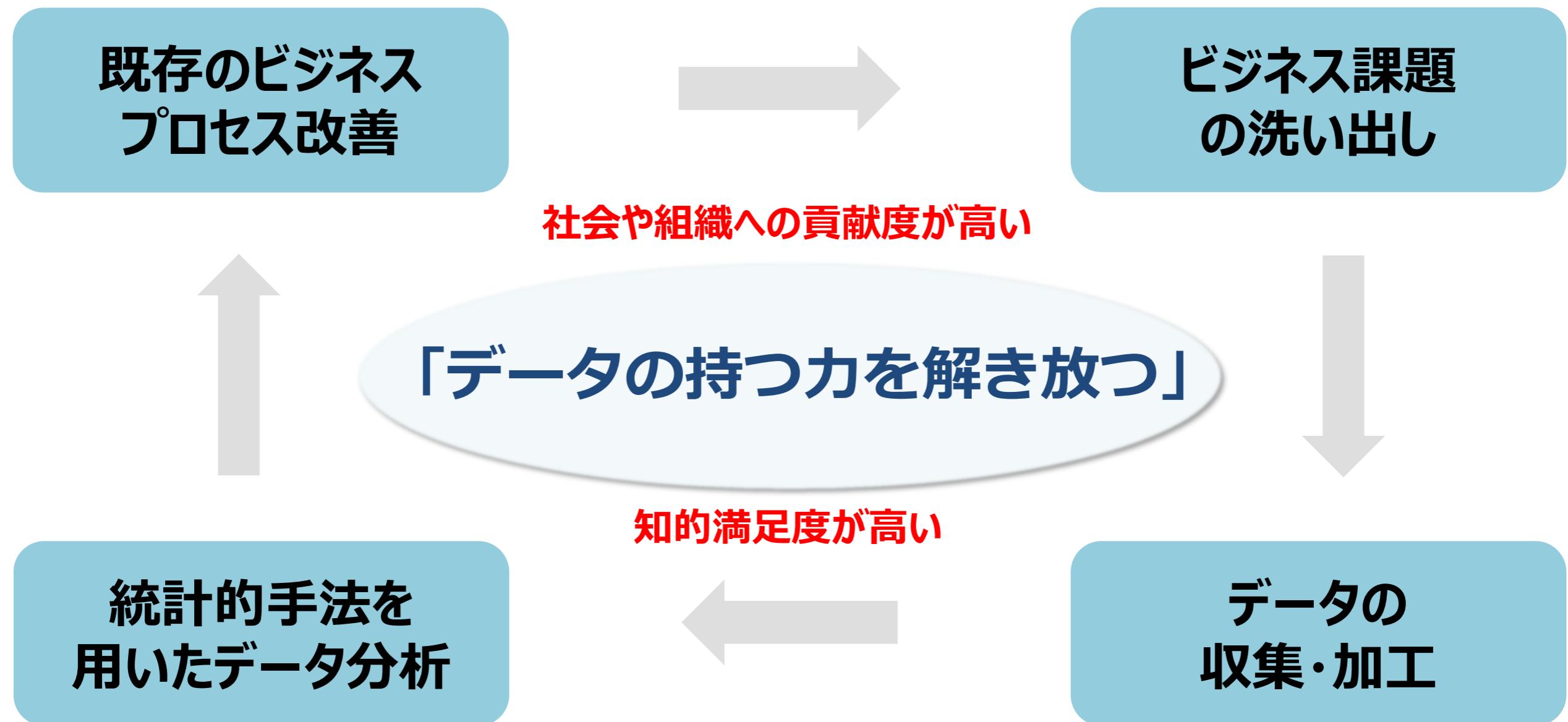


コンサルティング

クライアント企業の課題に対し、活用すべきデータの見極めやデータ解析を行い、解決のための有効な施策を提案する。

さまざまな業界において、データから価値を見出し
意思決定をすることができるデータサイエンティストが求められている

データサイエンティストの仕事の魅力



タイプ別① マーケティング担当

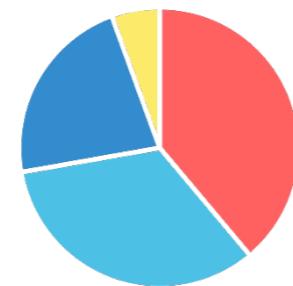


タイプ①

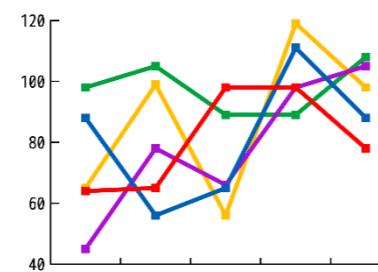
事業会社におけるマーケティング担当

データ活用業務の一例

データに基づくマーケティング施策の立案

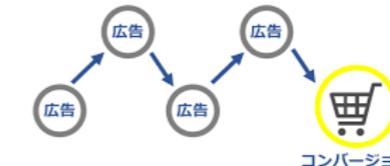
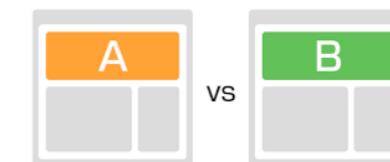


顧客アンケートデータからニーズを把握し、商品開発に活用。



売上実績の時系列の傾向から需要予測を実施。

広告配信・運用の最適化



広告の配信実績データ等を分析し、より効果的な広告を見極めて配信の最適化を行う。

タイプ別② データ分析コンサルタント

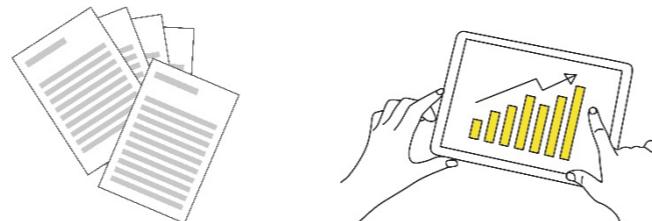


タイプ②

データ分析コンサルタント

データ活用業務の一例

データ分析プロジェクト マネジメント



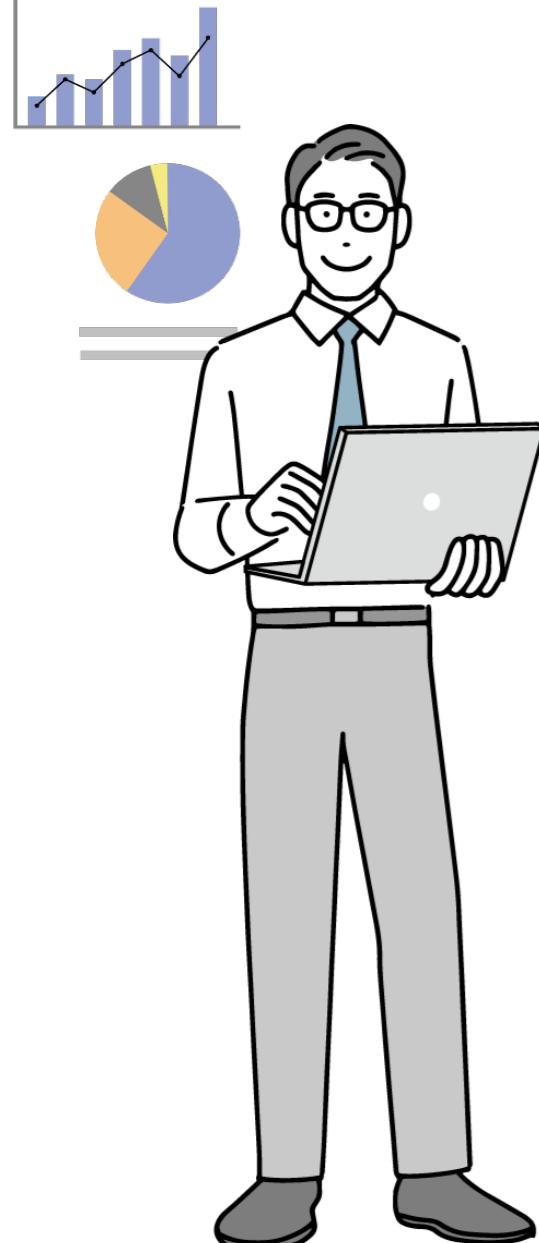
クライアントに対して、ビジネス課題の整理やプロジェクトを遂行するための全体設計やアドバイザリー。また、分析結果に基づく課題解決策の提示などを行う。

AI導入コンサルティング



組織にAIシステムの導入を検討する際の、初期設計や運用プランを考案。継続的なデータ利活用のためのアドバイスを行う。

タイプ別③ データ利活用推進アーキテクト



タイプ③

組織内のデータ利活用を推進するデータアーキテクト

データ活用業務の一例

データビジョンの策定

データの利用要件を理解し、技術の要件へと変換、成果を出すためのビジョン策定を行う。システム連携やデータ戦略、方針だけでなく、データの保管ポリシーやセキュリティにも配慮する。

データアーキテクチャの設計

ビジョンに基づき、データの収集・保管方法と使い道を設計する。ここで設計したユーザビリティが経営チャンスに直結する、重要度の高い仕事。

データフローの定義

収集データの保管形態、更新頻度、変換処理などを図であらわす。データの利用ユーザーに対して、アクセス権や利用の方法を定める。

高まるデータサイエンス人材ニーズ

データ分析関連人材規模とは？

以下の4人材を対象に、その合計を人数ベースで算出。

①分析コンサルタント

(データ活用戦略の策定など)

②データサイエンティスト

(データ収集、プログラミング言語を用いた分析、モデルの開発など)

③分析アーキテクト

(開発モデルをシステムに実装するフェーズに関与)

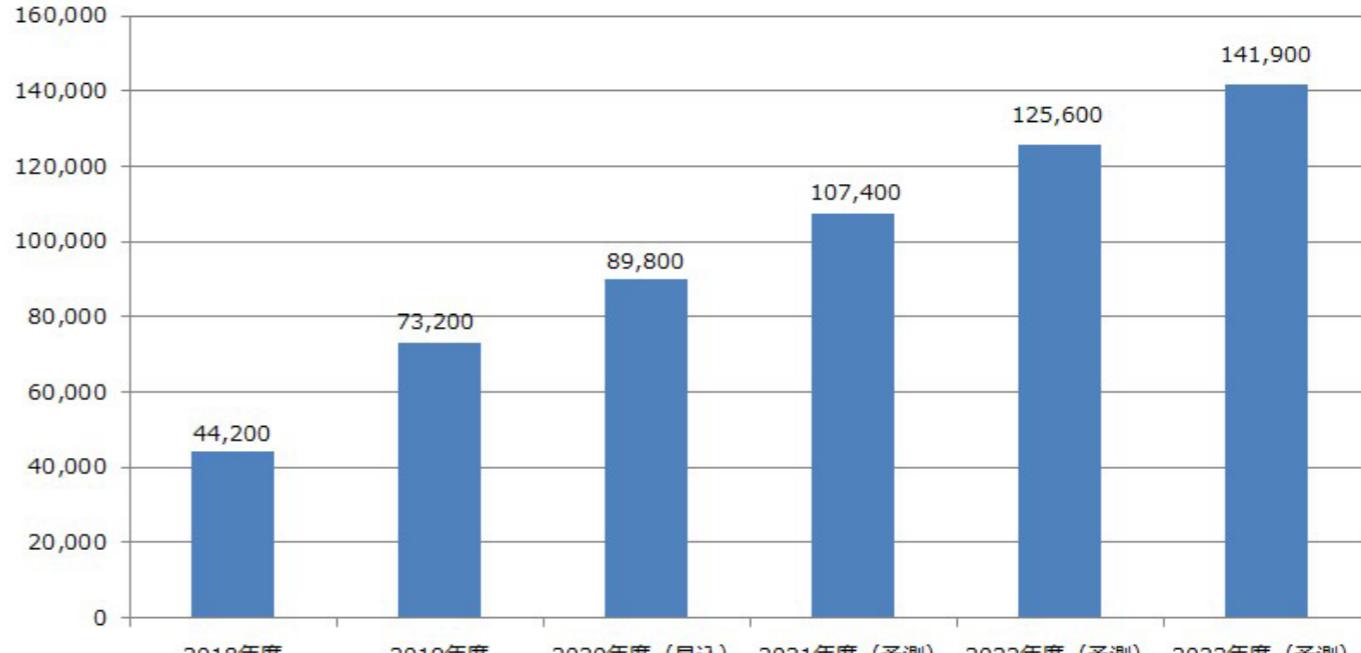
④プロジェクトマネージャー

(データ分析案件の統轄)

出典

https://www.yano.co.jp/press-release/show/press_id/2639

(人)



矢野経済研究所調べ

注1. 人数ベース

注2. データ分析関連人材規模は①分析コンサルタント、②データサイエンティスト、③分析アーキテクト、④プロジェクトマネージャーの4人材の合算値

注3. 2020年度見込値、2021年度以降予測値

2020年度の国内データ分析関連人材規模は89,800人の見込み。
2023年度には141,900人に達すると予測されている。

次回のテーマ

次回は

「組織におけるデータサイエンスの実現」

お疲れ様でした！