EXPLORATORY

EXPLORATORY

データを「見る人」から
価値を「生み出す人」へ

データサイエンス・ブートキャンプ・トレーニング　紹介資料

2026年12月14日（月）・15日（火）・16日（水）

主催：Exploratory, Inc.

その問いに、答えを出す手法を持っていますか

それぞれの問いには最適な分析手法があります。手法を理解していなければ、データがあっても答えは出ません。

「この施策、効果があったと言える？」

▶ 信頼区間・仮説検定

「この落ち込みは異常？通常の変動？」

▶ ばらつきの評価・XmRチャート

「何を改善すれば、成果が一番動く？」

▶ 線形回帰・ロジスティック回帰による分析

「より精度の高い予測をしたい」

▶ ランダムフォレスト・XGBoost

「顧客をどう分ければ打ち手が変わる？」

▶ クラスタリング

「来月どれだけ売れる？何人必要？」

▶ 時系列予測・フォーキャスト

ブートキャンプの3日間は、この手法群そのものです

全24セッションの構成（各1時間）

基礎・加工 3	統計・推測統計 6	回帰・多変量 4	機械学習・AI 6	演習・総括 5

24セッションのうち**16セッションが、統計学・回帰分析・機械学習・AIを使った手法そのもの**に充てられています。ツールの操作説明ではなく、「どの問いに、どの手法で、どう答えるか」に時間を使います。

なぜ「集計と可視化」で止まってしまうのか

手法の引き出しがないこと、知識が断片的で応用が利かないこと、そしてその両方が悪循環になっていることが原因です。

STEP 1
データを準備する
結合・整形
欠損値の処理

▶

STEP 2
可視化して眺める
分布と傾向を
つかむ

▶

STEP 3
差や変化を確かめる
その差が偶然か
統計的に検証

▶

STEP 4
要因を特定する
何が結果に
効いているか

▶

STEP 5
予測して意思決定する
将来を見積もり
打ち手を決める

多くの現場はここまで

ここから先は、手法を持っていないと進めない

1. 手法の引き出しがない

- さまざまな問いに対応する手法があることを知らない
- 検定・回帰・機械学習は**「専門の知識が要る」**と感じて手が出ない
- 社内に、手法を教えられる人がいない

2. 知識が断片的で応用が利かない

- 手法の名前は知っていても、なぜその手法を使うのかが分からない
- 見よう見まねで試しては、**自分のデータに応用できず挫折**
- 「有意」「オッズ比」といった専門用語がわからず、結果を解釈できない

3. 学ぶ時間も、相談相手もいない

- 業務の合間の独学は、途中で頓挫しやすい
- 疑問が出たときに、その場で確かめられる相手がいない
- 結果、毎回同じ集計を繰り返すことになり、この悪循環から抜けられない

本ブートキャンプは、この3つを同時に解きます

統計・回帰・機械学習の手法を、実務の問いとの対応づけとともに体系的に扱います。

数式ではなく「なぜその手法か」「結果をどう解釈するか」から入り、演習で手を動かします。

3日間で一気通貫。講師とスタッフがその場で疑問を解消し、受講後もサポートが続きます。

ブートキャンプで学べること

扱うトピックと、それが実務のどんな場面で役立つか。次ページから1つずつご説明します。

トピック	内容	実務で役立つ場面
データの基礎・可視化	データタイプと性質、分布と傾向のつかみ方	受け取ったデータの、どこを最初に見るべきかが分かる
データラングリング	結合・集計・整形をステップとして積み上げる	分析の8割を占める前処理を、再実行できる形にする
統計の基礎・XmRチャート	ばらつきの指標、標準化、シグナルとノイズの分離	数字の上下が異常なのか通常の変動なのかを判断する
推測統計・信頼区間	確率分布、中心極限定理、区間推定	一部のデータから全体を語るときの確からしさを示す
仮説検定	t検定・比率検定によるグループ間の差の検証	「その差は誤差では？」と聞かれたときに根拠を示す
相関分析	カテゴリーと数値／数値と数値の相関、総当たりでの確認	何が結果に効いているかを特定し、効きそうな要因に当たりをつける
線形回帰	目的が数値のときの要因分析。係数の解釈	売上や満足度などの数値に影響する要因を特定し、改善の優先順位を決める
ロジスティック回帰	目的が二値（する／しない）のときの要因分析。オッズ比の解釈	解約・成約・離反が「起きるかどうか」の要因を特定する
機械学習・モデルの検証	決定木／ランダムフォレスト、変数重要度と予測精度	予測モデルを作り、その結果を評価する
クラスタリング	回答・行動パターンの似た対象をグループに分類	属性ではなく実態でセグメントを作る
時系列予測	売上・需要の予測	将来の数字を見積もり、要員や在庫の計画に使う
テキスト分析 with AI	AIモデルによる自由記述の分類・要約・スコア化	テキストデータを定量データと同じ土俵で扱う
エキササイズ（演習）	実データを使った4回の実践演習	学んだ手法を、業務に近い形で通しで使ってみる

データの基礎・可視化

データを受け取ったとき、最初にどこを見るべきかを身につけます。

どんなトピックか

分析の出発点になる数値・カテゴリ・日付・論理値といった**データタイプと、それぞれの性質と適切なデータ型への変換方法**を学びます。

よくある課題

読み込んだ時点でデータ型が正しく設定されていないことに気づかないまま進めてしまいます。数値が文字列のまま、日付が文字列のままでは、**集計も可視化も分析も正しく進みません。**

できるようになること

全列の分布とデータタイプを把握でき、どういう特徴をもったデータなのかを理解できるようになります。さらに適切なデータタイプへの変換方法を学べます。

データラングリング（データ加工）

分析作業の8割を占める前処理を、再実行できる形に変えます。

どんなトピックか

UI、あるいは、自然言語で指示するAIプロンプトを通して、結合・集計・条件に応じた計算などのデータ加工の基礎を学びます。さらに、ステップとして記録された加工の履歴を、あとから見直し・修正・再実行する方法を学びます。

よくある課題

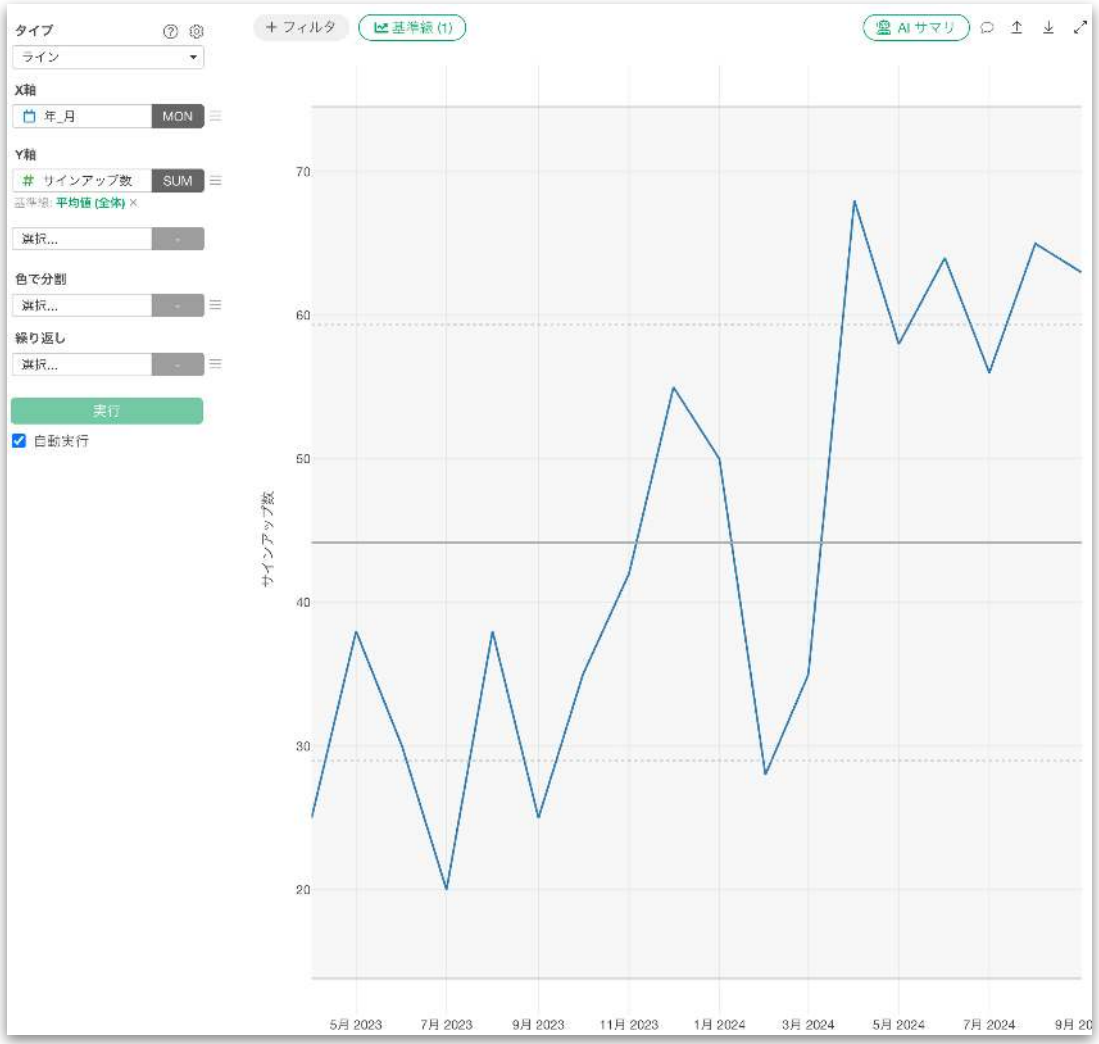
表計算ソフトの関数で組むのは手間がかかり、分析したい形にたどり着けません。手順も残らず、何をどう直したのか後から追えません。

できるようになること

関数では組みにくい加工も、日本語で伝えるだけで形にでき、生成された処理を確認したうえで適用できます。手順はステップとして残るため、次回はデータを差し替えるだけで済みます。**前処理に費やしていた時間を、そのまま分析と考察に回せます。**

XmRチャート

数字の上下が「異常」なのか「いつものばらつき」なのかを見分けます。



どんなトピックか

時系列の変動から**シグナル（意味のある変化）とノイズ（通常のばらつき）を切り分ける**XmRチャートを扱います。

よくある課題

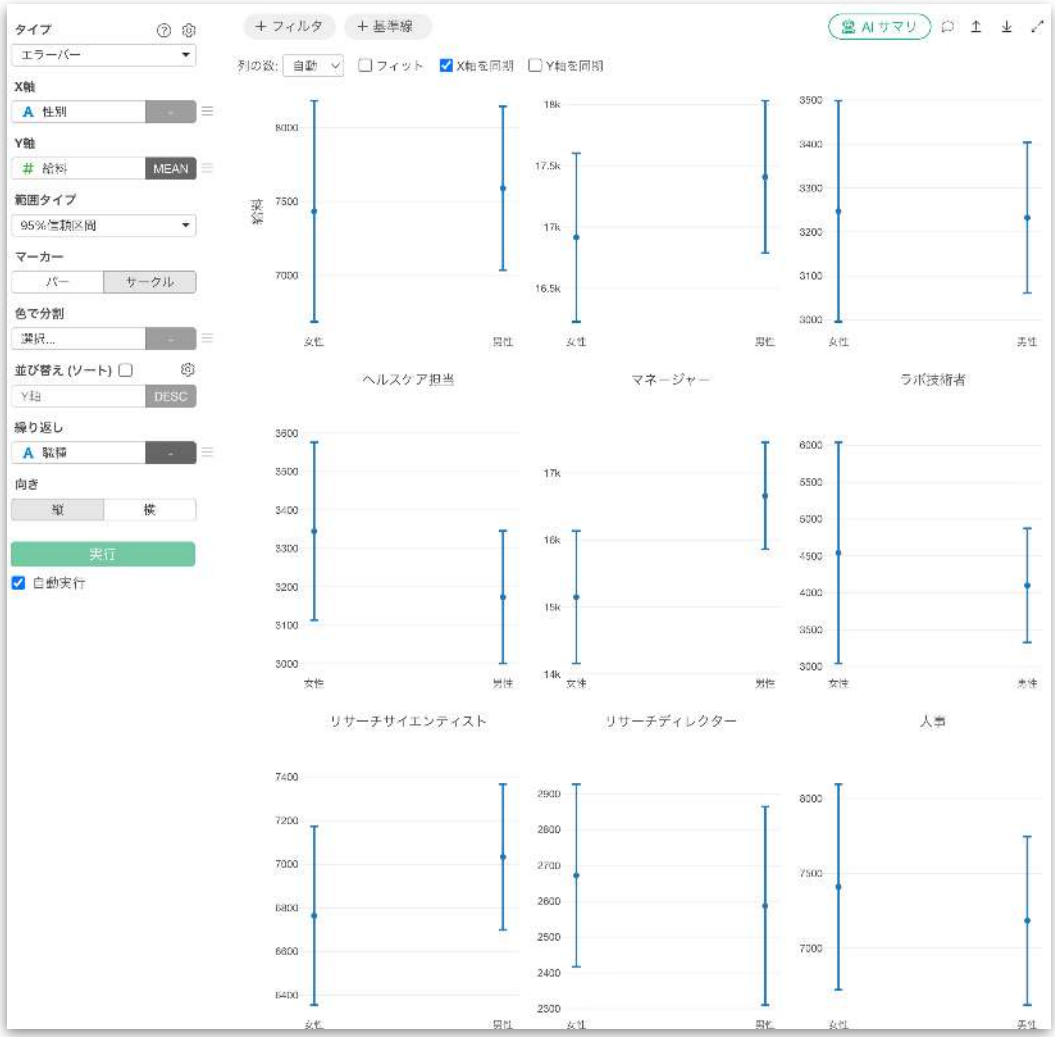
前月比・前年比の増減に一喜一憂し、**通常のばらつきの範囲でしかない変動に対して施策を打ってしまいます。**逆に、本当に対応すべき変化が「よくある上下」として見過ごされることもあります。

できるようになること

どこからが「対応すべき変化」なのかを**基準を持って線引きできる**ようになり、報告のたびに数字の上下を説明し直す必要がなくなります。

推測統計・信頼区間

手元の一部のデータから、全体について語れるようになります。



どんなトピックか

確率と確率分布、中心極限定理を土台に、標本から母集団を推定する考え方を扱います。推定値を1点ではなく**幅（信頼区間）として示す方法**を学びます。

よくある課題

サンプル数が少ないほど数字は揺れますが、その揺れの大きさが分からないため、細かい差にまで意味を読み取ってしまい、**誤った判断を下すことにつながります。**

できるようになること

「離職率は12%」ではなく「9〜15%の範囲にある」と、幅で報告できるようになり、**グループ間に統計的に有意な違いがあるかを判断できるようになります。**

仮説検定

グループ間の差に、統計的に意味があるのかを確かめます。

どんなトピックか

仮説を立て、データで検証し、受け入れるか棄却するかを判断するという**科学的手法のプロセス**を学びます。そのうえで t検定・比率検定・カイ2乗検定といった手法を、どの場面でどう使うかを扱います。

よくある課題

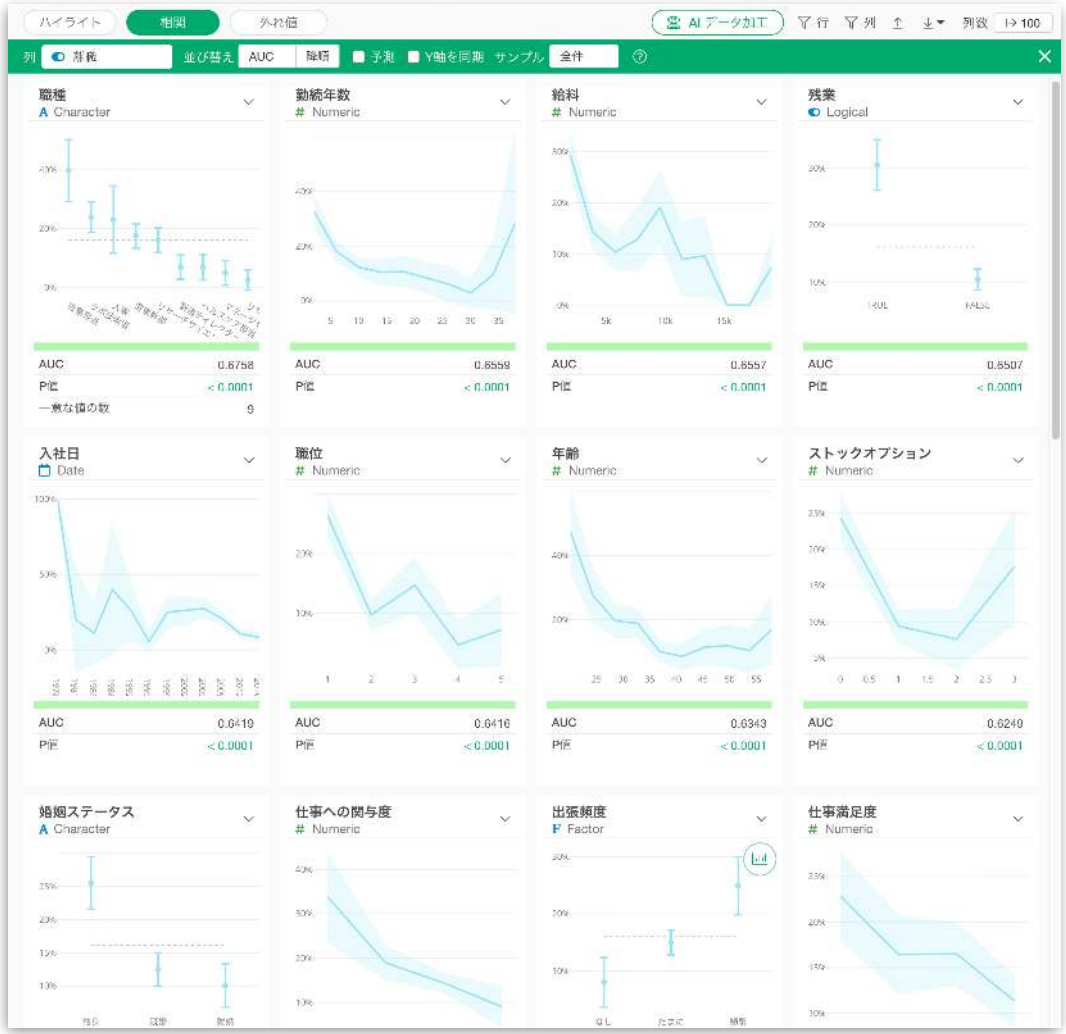
「A群とB群で5ポイント差」が本物の差なのか偶然なのか分からないまま施策を打ってしまいます。逆に「それ、誤差では？」と指摘されたときに**反論できず、せっかくの発見が意思決定に使われません。**

できるようになること

「この差は統計的に有意です」と根拠を持って報告できるようになります。逆に「差とは言えない」と分かれば、無駄な打ち手を避けられます。

相関分析

数十列のデータから、結果と関係していそうな要因の当たりをつけます。



どんなトピックか

カテゴリーと数値、数値と数値それぞれの相関の見方を扱います。さらに相関モードを使って、**多数の列の総当たりの相関を一度に確認し、結果と関係の強い要因を絞り込む**方法を学びます。

よくある課題

気になる2列を選んではチャートを作る、という確認を1組ずつ繰り返すことになります。列が数十あると総当たりで見るのは現実的でなく、関係のある要因を見落としたまま先に進んでしまいます。

できるようになること

全列の相関を一度に見渡し、深掘りすべき要因を短時間で絞り込めるようになります。ここで当たりをつけてから回帰分析に進む、という分析の順序が身につきます。

線形回帰

興味の対象が数値のときに、複数の要因が結果にどれだけ影響を与えているかを評価します。

どんなトピックか

目的が数値（売上・給料など）のときの要因分析の手法を学びます。係数の意味と有意性、信頼区間、カテゴリー型の説明変数の扱い、R2乗によるモデル評価、多重共線性までを扱います。さらには相関と因果関係の違いまで扱います。

よくある課題

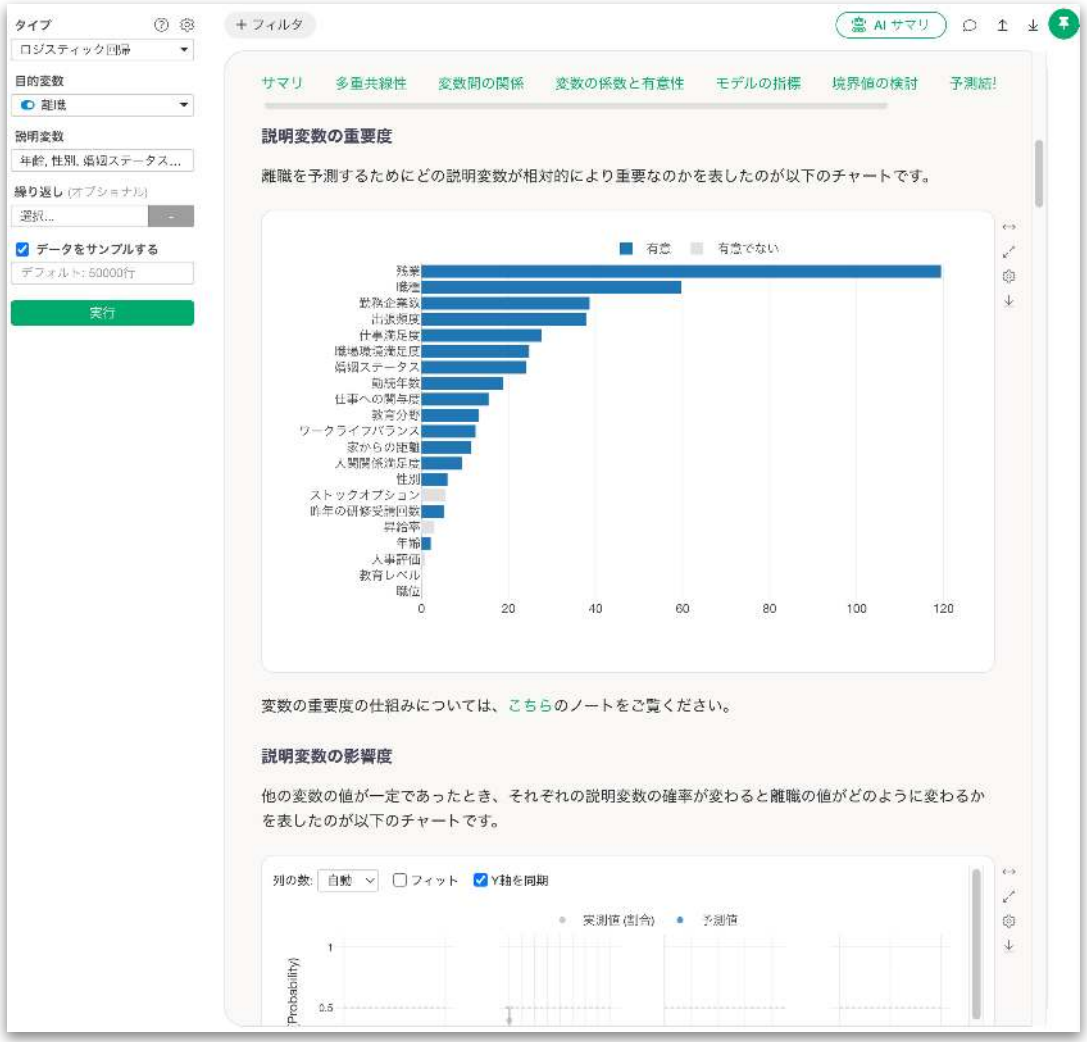
1対1の相関を見ただけでは、別の要因が両方を動かしているだけの見かけの関係を本物だと思い込んでしまいます。要因が複数あるとき、どれが本当に効いているのかを切り分けられません。

できるようになること

「職位や勤続年数が同じ人どうしで比べても、給料に差が残るのか」のように、他の要因の影響を取り除いたうえで、**それぞれの要因が独立して結果にどれだけ影響を与えているかを数字で示せるようになります。**

ロジスティック回帰

解約・成約・離反といった「起きる／起きない」の要因を突き止めます。



どんなトピックか

結果が二値（Yes / No）の場合の分析方法を扱います。**オッズとオッズ比**の考え方、係数の解釈、信頼区間と仮説検定、そして予測確率の読み方を学びます。

よくある課題

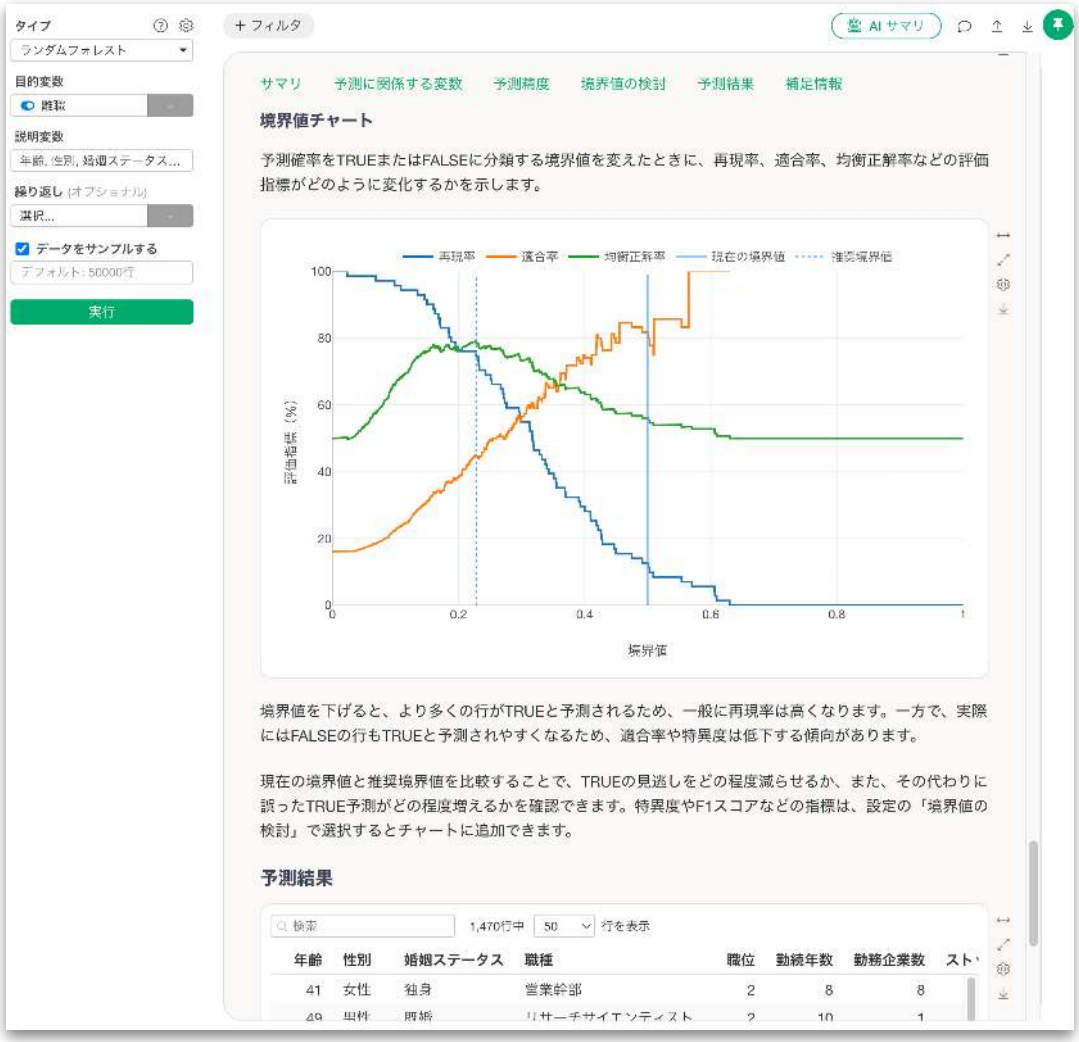
離職率や成約率をセグメント別に集計するだけでは、**複数の要因が重なったときに何が効いているのか分かりません。**また「オッズ比1.8」と言われても、それが何を意味するのかを説明できないまま報告してしまいます。

できるようになること

「辞める／辞めない」「買う／買わない」といった二値の結果について、それが起きる確率を左右している要因を、影響の大きさの順に示せるようになります。

機械学習・モデルの検証

決定木・ランダムフォレストで予測モデルを作り、その精度を検証します。



どんなトピックか

決定木とランダムフォレストの仕組みを理解し、予測モデルを構築します。あわせて**トレーニングデータとテストデータを分けてモデルを評価する**方法、予測精度の指標の読み方、パラメーター・チューニング、統計モデルとの使い分けを扱います。

よくある課題

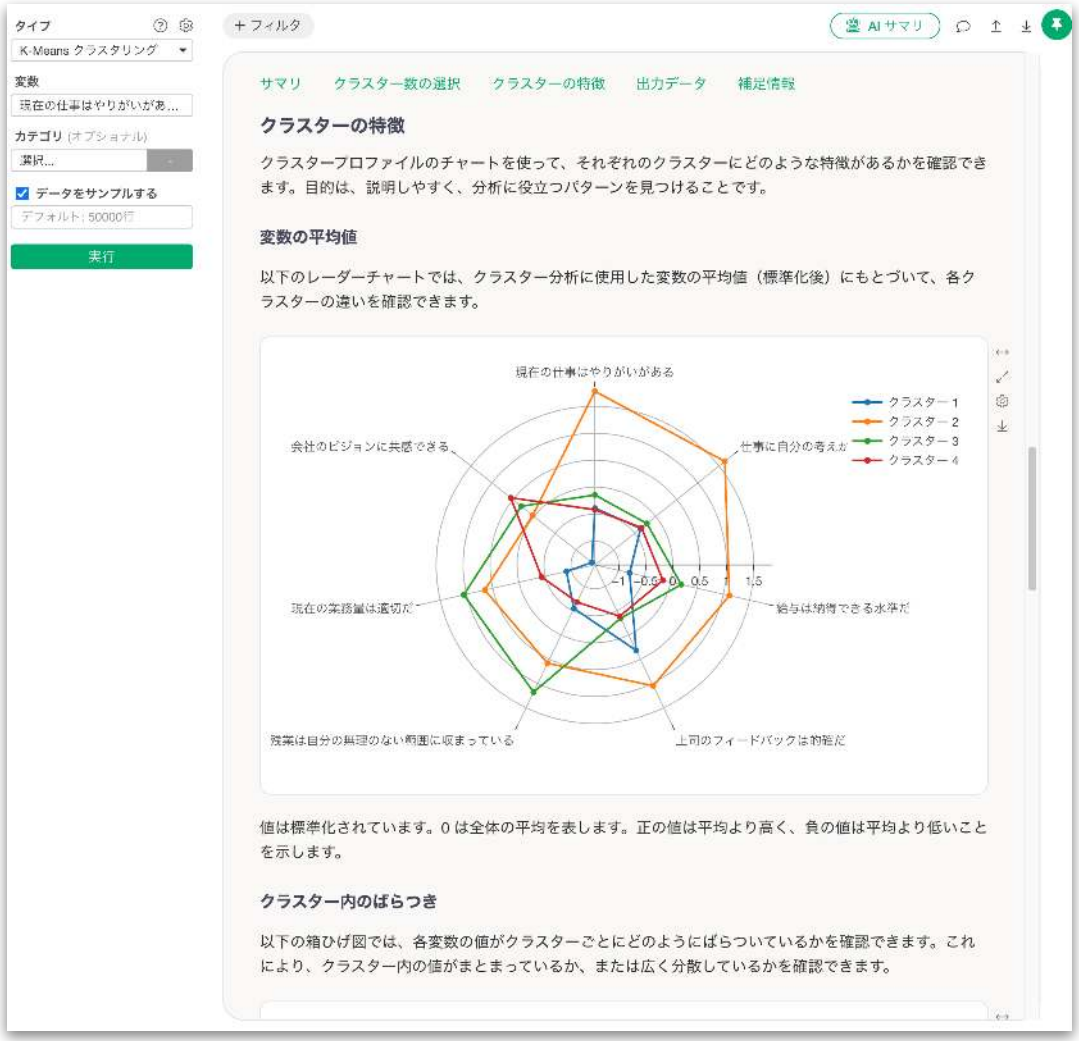
手元のデータにはよく当てはまるのに、新しいデータではまるで当たらないモデルを作ってしまいます。**予測精度が何を意味するのかを検証しないまま**、モデルの結果を意思決定に使ってしまう危険があります。

できるようになること

過去のデータから予測モデルを作り、まだ**結果が出ていないデータに予測結果を追加できるようになります。**さらに、テストデータで精度を測り、実務に載せてよいモデルかも判断できるようになります。

クラスタリング

属性ではなく、実際のデータで顧客や対象をグループに分けます。



どんなトピックか

似ているデータどうしを**グループ（クラスター）にまとめる**手法を扱います。またクラスターごとの特徴をどう解釈するか、そして最適なクラスター数をどう探すかまで含めて学びます。

よくある課題

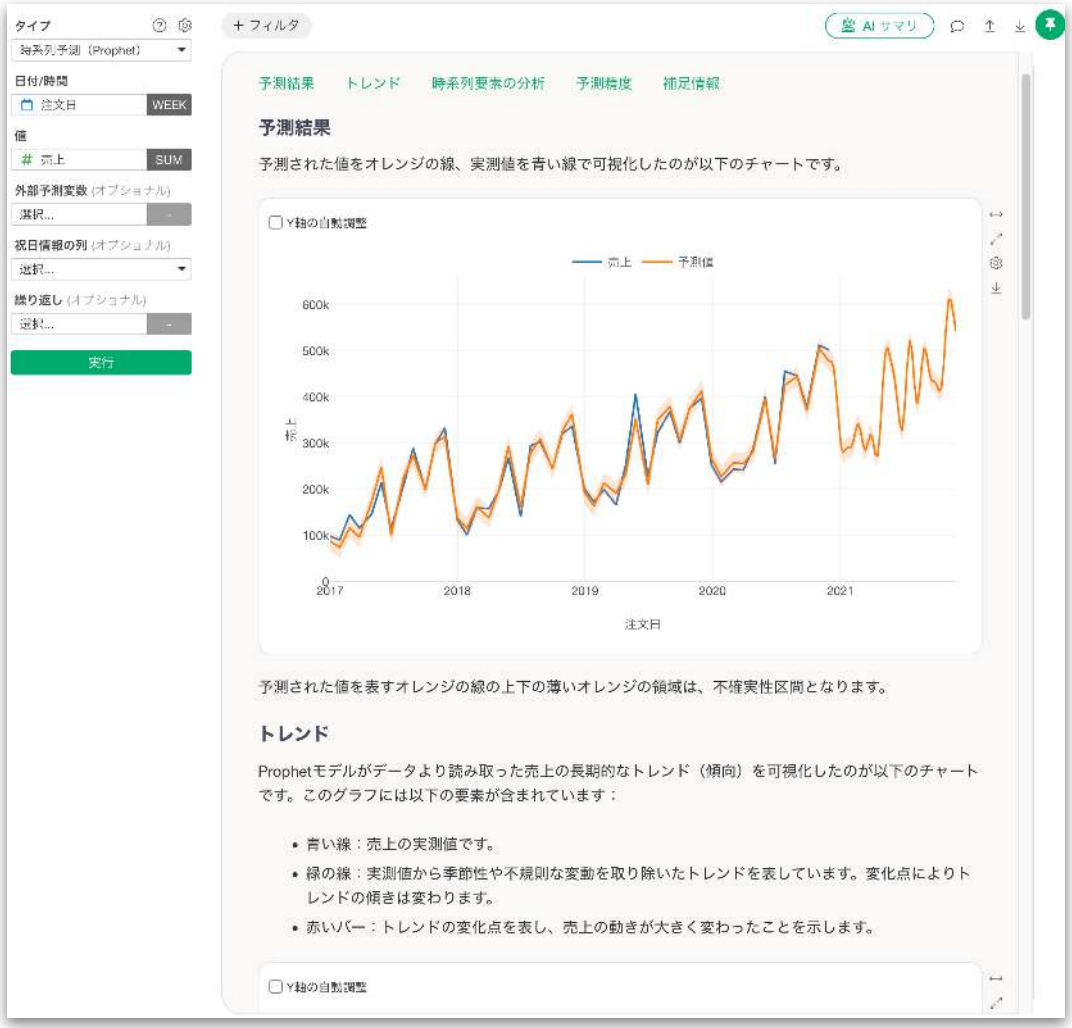
セグメントを「20代・首都圏」のような**属性でしか切れません。**しかし実際に行動を分けているのは利用実態や価値観であることが多く、属性別に集計しても差が出ないまま終わります。

できるようになること

「ヘビーユーザー層」「価格重視層」といった、**意味のあるセグメントをデータから作れるようになります。**各クラスターの特徴を比較することで、誰に何を訴求するかの根拠も得られるようになります。

時系列データの分析と予測

売上・需要・アクセス数の将来を、確からしさとともに見積もります。



どんなトピックか

トレンドと季節性を分解して捉えたうえで、**時系列の予測モデル**を構築します。予測精度の検証、外部の予測変数を加えた分析まで扱います。

よくある課題

需要や売上の見通しが担当者の勘に頼り、担当が代わると再現できません。また、前年同月比や移動平均だけで将来を語った場合、予測精度が低くなってしまうことがあります。

できるようになること

季節性やトレンドを織り込んだ予測を、予測区間（幅）つきで示せるようになります。要員配置・在庫・予算といった計画に、根拠のある数字を使えます。

テキスト分析 with AI

自由記述やレビューを「抜粋」ではなく「全件」として扱えるようにします。

どんなトピックか

AIを使ってテキストを**分類・要約・スコア化**します。ポジティブ／ネガティブの判定や、どのトピックがどれだけ出現しているかの傾向把握を行います。

よくある課題

数百〜数千件のテキストを読み切れず、「代表的な意見」を数件抜粋するだけになります。**抜粋には選ぶ人の主観が入り、全体の傾向を表しません。**せっかく集めた定性データが、報告書の付録で終わってしまいます。

できるようになること

テキストデータを定量データと同じ土俵で比較でき、「不満の38%は初期設定に関するもの」といった形で、テキストデータを数字で語れるようになります。

3日間のカリキュラム

各日 9:00〜18:00。講義と、実データを使った演習（エキササイズ）を組み合わせて進めます。

DAY 1 — 12月14日（月）

統計学と推測統計の基礎

09:00-10:00	はじめに、データ分析とは
10:00-11:00	データの基礎
11:00-12:00	統計の基礎 - 標準偏差、標準化
12:00-13:00	ランチ休憩
13:00-14:00	XmRチャート
14:00-15:00	データラングリング
15:00-16:00	エキササイズ Part 1
16:00-17:00	確率論と推測統計
17:00-18:00	信頼区間

DAY 2 — 12月15日（火）

推測統計と統計学習モデルを使った分析

09:00-10:00	仮説検定の考え方
10:00-11:00	仮説検定手法 - t検定・比率検定
11:00-12:00	相関分析
12:00-13:00	ランチ休憩
13:00-14:00	エキササイズ Part 2
14:00-15:00	線形回帰の基礎
15:00-16:00	予測モデルの解釈
16:00-17:00	ロジスティック回帰
17:00-18:00	エキササイズ Part 3

DAY 3 — 12月16日（水）

AI・機械学習の手法を使った分析

09:00-10:00	機械学習 - 決定木・ランダムフォレスト
10:00-11:00	モデルの検証／統計学習 vs. 機械学習
11:00-12:00	パラメーター・チューニング
12:00-13:00	ランチ休憩
13:00-14:00	クラスタリング
14:00-15:00	時系列データの分析と予測
15:00-16:00	テキスト分析 with AI
16:00-17:00	エキササイズ Part 4
17:00-18:00	総括

講師

西田 勘一郎

CEO
EXPLORATORY

@KanAugust

略歴

- 米オラクル本社で、16年にわたりデータサイエンスの開発チームを率い、機械学習、ビッグ・データ、ビジネス・インテリジェンス、データベースに関する数多くの製品を世に送り出した。
- 2016年春、**データサイエンスの民主化**のため、Exploratory, Inc を立ち上げる。
- Exploratory, Inc.でCEOを務めるかたわら、データサイエンス・ブートキャンプ・トレーニングなどを通してデータサイエンスの技術と手法の普及と教育に取り組む。
- 2025年、「データに触れながら学ぶ統計学」を執筆、出版。

実績

これまでの受講企業・団体（一部）

データサイエンス・ブートキャンプ・トレーニングは通算50回以上、これまでに1,000名以上の方に受講いただいています。

受講者の声

実際に受講された方からいただいたコメントです。

> 「参加して最も良かったことは、データ分析に対する解像度が格段に上がったことです。『なぜこの手法が使われるようになったのか』が解説され、断片的な知識だったものが全体のまとまりをもって理解できました。」

山口 郷彬 様 ／ Gcomホールディングス株式会社

> 「統計学の概要から具体的なデータのクレンジング、分析手法まで幅広く、時に奥深く学ぶことができ、濃密な3日間を過ごすことができました。プログラミングの詳しい知識が無くても統計・分析・モデリングまで行えるので、非常に実用的だと感じています。」

安部 香織 様 ／ 株式会社 FIELD MANAGEMENT EXPAND

> 「学術的なバックボーンは押さえつつ、実務で成果を出すことに主眼を置いた内容で、実務担当者の日々の業務に直結したトレーニングです。Excelでは対処しきれないデータを扱う方、すべてにお勧めします。」

マーケティング担当 ／ レジャーサービス業界

> 「文系人間でも、プログラミング言語に関する学習を省略することで『データと意思決定』だけに注力できたのは得難い経験でした。濃厚な3日間は適度に脳を酷使し、他業界の方々との交流も良い刺激になりました。」

荒田 孝幸 様 ／ 日本証券金融株式会社

開催概要・受講料

2026年12月回の詳細です。

開催概要	
日程	2026年12月14日（月）・15日（火）・16日（水）
時間	各日 9:00〜18:00
形式・会場	東京・丸の内（対面）
定員	25名（最小催行予定数 10名）
受付締切	2026年12月2日（水）　※定員に達し次第終了
受講資格	前提知識・経験は不要 （統計・プログラミングの知識も不要）
必要な環境	Mac（macOS 11 Big Sur以降）または Windows（Windows 8以降・64bit版）のノートPC
録画	全講義を録画。受講者は後から視聴可能
支払方法	クレジットカード（Visa・Master・Amex） または銀行振込（請求書払い可）

受講料

お一人あたり

247,000円（税別）

税込 271,700円

※2026年10月31日まで20%オフの早割を実施

含まれるもの

- 3日間の研修（全24セッション）・教材費
- 特典：1年間の Exploratory Business版ライセンス
  （年額$1,188相当／約19万円相当）
- 全講義の録画（受講者は後から視聴可能）
- 終了後の個別相談／チャットサポート／無料オンボーディング

割引制度

- **グループ割引**：3名以上まとめてのお申し込みで 10%OFF
- **学生割引**：50%OFF（いずれも詳細はお問い合わせください）

費用対効果の考え方

社内でご説明いただく際の整理としてお使いください。

外部に委託した場合と比べると

データの集計・分析を外部に委託した場合、1案件あたり数十万円規模になることが一般的です。**本研修の費用は、その分析1本分と同等かそれ以下**でありながら、支出の性質が異なります。

委託：その案件の**成果物**が残る

研修：**以降すべての分析に使えるスキル**が社内に残る

分析が終わるまでの時間も変わります

- 委託では、依頼・見積・データの受け渡し・初稿・修正依頼と、やり取りのたびに待ち時間が発生します。
- 意図が伝わらず、出てきた分析が想定と違えば、やり直しが必要になります。
- **内製できれば、この往復がなくなり、**見たいと思ったその日に確かめられます。

投資の回収イメージ

- **分析の内製化**：定常的な分析を社内で回せれば、年間の外注費を継続的に圧縮できます。
- **AI活用の質が上がる**：AIが出した分析結果を検証・解釈できるようになり、そのまま意思決定に使えます。
- **作業時間の削減**：データ加工とレポートが再利用できる形になり、更新のたびの手作業が減ります。
- **意思決定の速度**：外部とのやり取りや手戻りがなくなり、確かめたいその日に確かめられます。
- **属人化の解消**：分析の型と手順が共有され、担当者交代時の引き継ぎが容易になります。
- **ツール込み**：1年分の Business版ライセンス（約19万円相当）が含まれるため、受講直後から実務に適用できます。

上申用サマリ　— 稟議・社内申請用

このページの内容は、そのまま社内申請書に転記してお使いいただけます。

項目	内容
件名	データサイエンス・ブートキャンプ・トレーニング 受講
目的	データ分析を外部委託・属人対応から内製化し、AIや分析ツールが出した結果を自ら検証・解釈して意思決定につなげられる人材を育成する
対象者	業務でデータを扱い、分析や意思決定に関わる者（マーケティング／事業企画／研究開発／情報システム／人事など）　※統計・プログラミングの前提知識は不要
実施概要	2026年12月14日（月）・15日（火）・16日（水）　各日9:00〜18:00／東京・丸の内（対面）／定員25名の少人数制 提供：Exploratory, Inc.（講師：CEO 西田勘一郎）　※全講義を録画し、受講者は後から視聴可能
費用	247,000円（税別）／1名　※税込 271,700円（2026年10月31日まで20%オフの早割を実施） 特典として Exploratory Business版ライセンス1年分（年額$1,188相当／約19万円相当）を含む 3名以上の同時申込でグループ割引（10%OFF）あり。請求書払い（銀行振込）対応可
習得内容	データの基礎・可視化／データラングリング／統計の基礎・XmRチャート／推測統計・信頼区間／仮説検定／相関分析・線形回帰／ロジスティック回帰／機械学習（決定木・ランダムフォレスト）／モデルの検証・パラメーターチューニング／クラスタリング／時系列予測／AIによるテキスト分析／実データ演習（全4回）
期待効果	①AIや分析ツールが出した結果の妥当性を自分で検証できる　②数字の変動が異常か通常のばらつきかを判断できる　③何が結果に効いているかを特定し、施策の優先順位を決められる　④データ加工・レポートを再利用できる形にし、工数を削減　⑤分析の外部委託費の削減

よくいただくご質問

社内でのご検討にあたって想定される論点をまとめました。

ご質問	考え方
AIがある今、 なぜ必要なのですか	AIは「答え」を返してくれますが、その答えが正しいかを判断するのは人です。その差に意味があるのか、データ量は十分か、他の要因の影響ではないか——こうした問いは、こちら側に視点がなければ浮かびません。AI時代だからこそ、統計リテラシーと手法の理解が意思決定の質を分けます。
統計やプログラミングの 知識がなくても大丈夫ですか	前提知識は不要です。加工も分析もUI操作とAIプロンプトで行うため、言語の習得に時間を使いません。毎回多くの未経験の方が参加され、受講を通じて実務で使えるスキルを習得されています。
書籍や動画での 独学では不十分ですか	独学では、疑問が出たときに確かめる相手がいません。本研修は定員25名の少人数制で、講義中の疑問はその場で解消できます。演習時間には講師とスタッフが会場を回り、手が止まった箇所を個別に支援します。
他のデータ分析研修とは 何が違いますか	多くの研修はプログラミング言語やツール操作の習得に時間を使います。本研修は順番を逆にし、「なぜその手法か」「結果をどう解釈するか」という本質に受講時間のほぼすべてを充てます。通算50回以上・1,000名以上への実施実績を持つ講師が直接教えます。
費用は妥当ですか	分析の外部委託1案件と同等以下の費用です。加えて Exploratory Business版ライセンス1年分（年額$1,188相当／約19万円相当）と無料オンボーディングが含まれており、受講後すぐに実務で使い始められます。
受講者に定着しますか	終了後に個別相談ミーティング・チャットサポート・無料オンボーディングを提供しており、実務で使い始める段階を継続的に支援します。録画を見返しながら復習することもできます。

連絡先

メール

support@exploratory.io

ウェブサイト

https://ja.exploratory.io

X (Twitter)

@ExploratoryJP

データ・アカデミー

https://exploratory.io/**data-academy-jp**