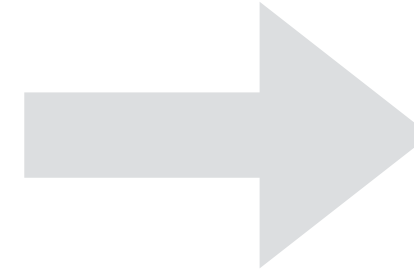
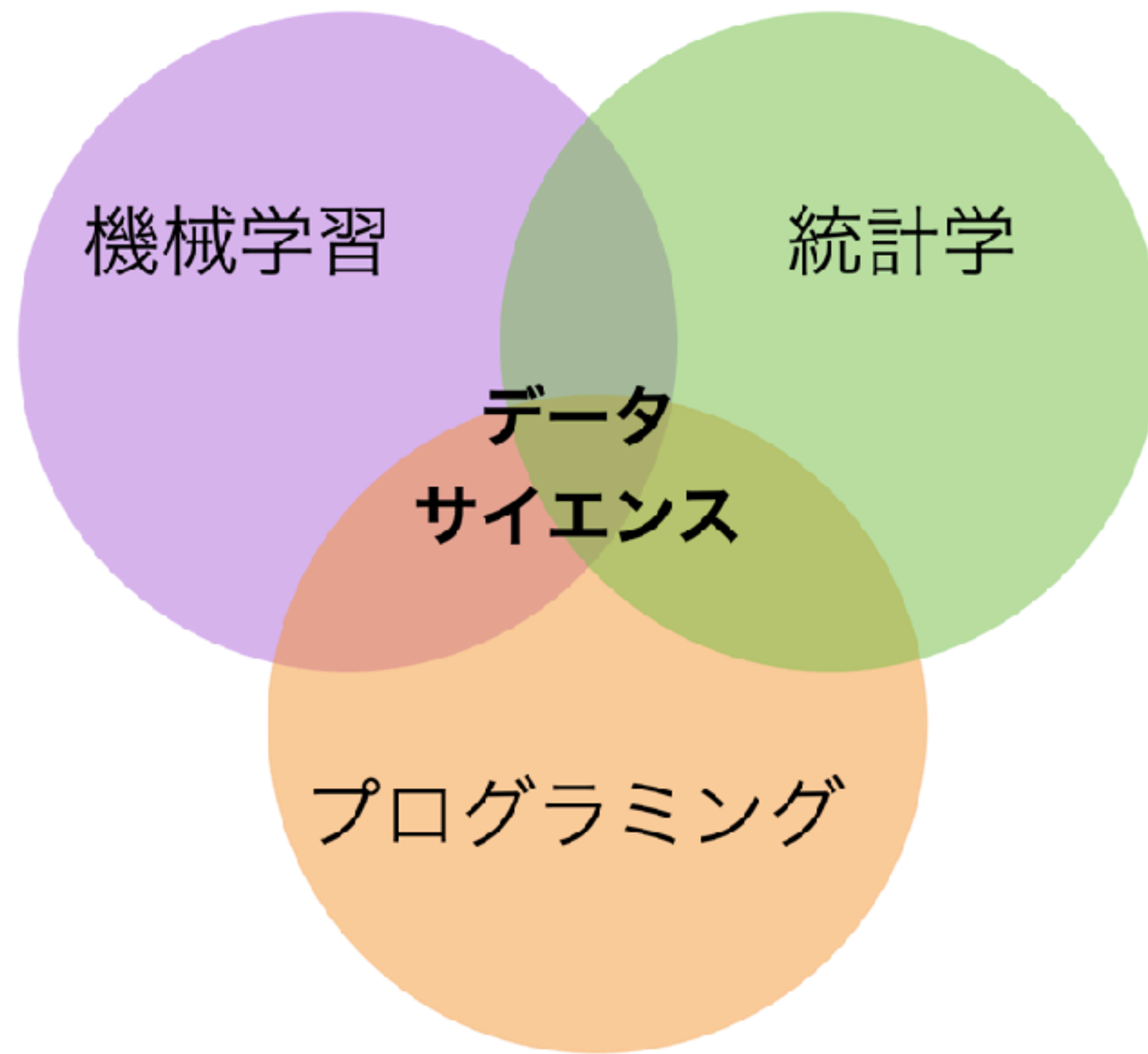


EXPLORATORY

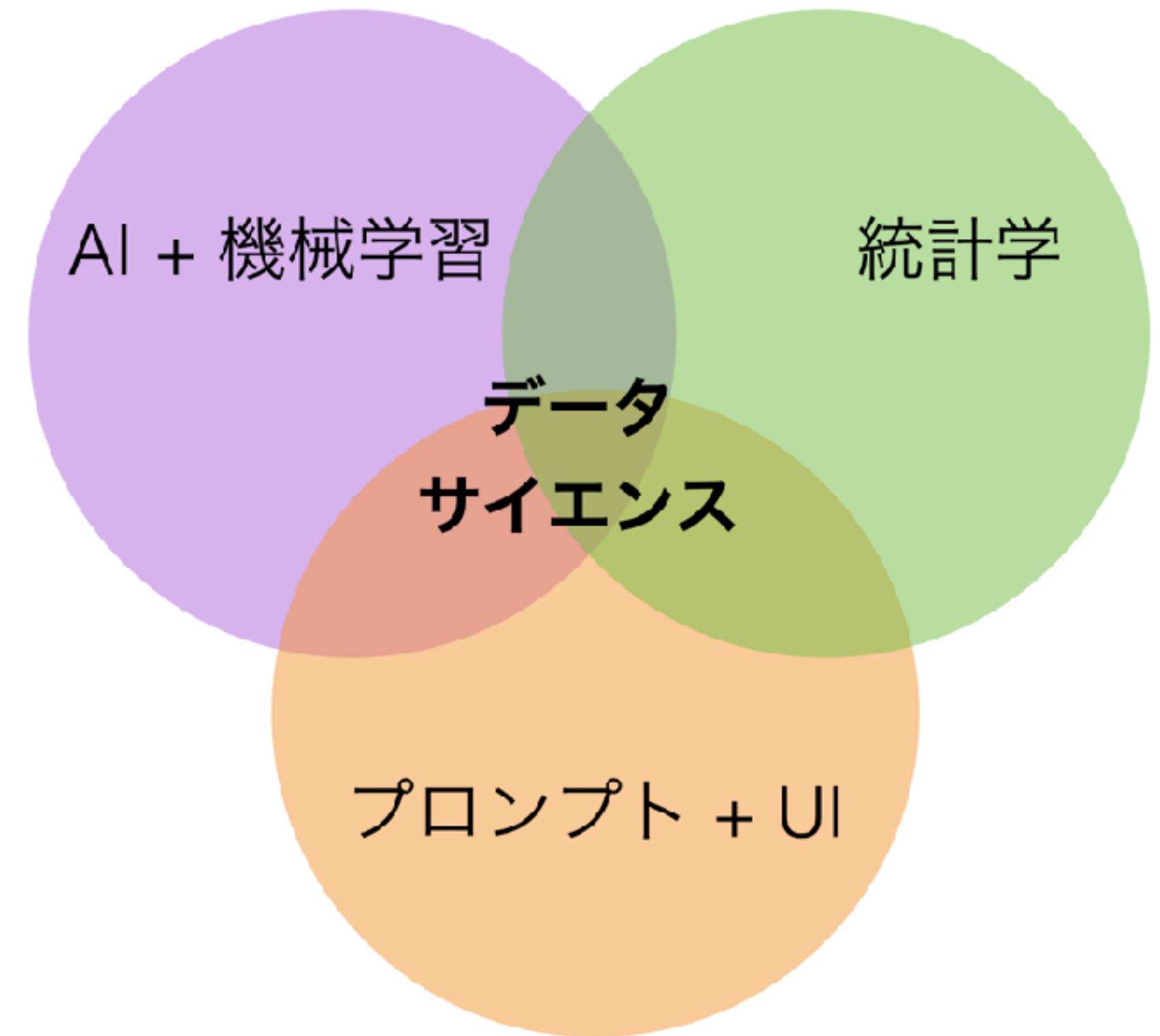
だれもがデータをもとに
よりよい意思決定ができる

データサイエンスの民主化

データサイエンス 1.0



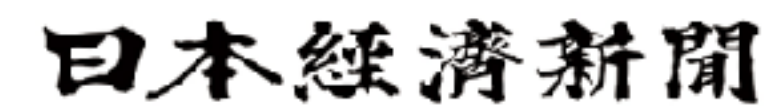
データサイエンス 2.0



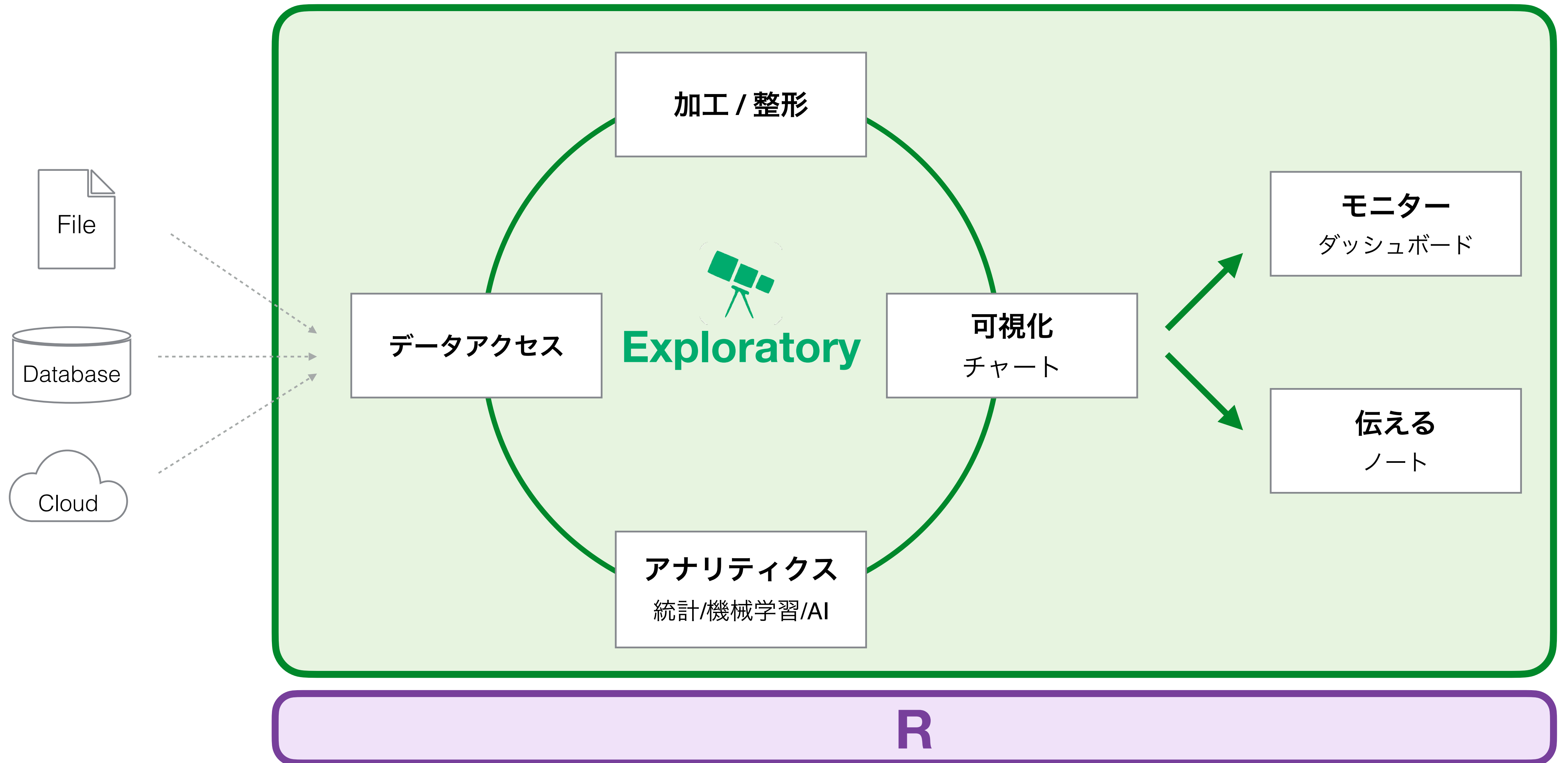
世界中の組織がデータを使って意思決定を行うための変革を支援しています。



日本の企業がデータを使って業務を改善していくための支援をしています。



分析する人のためのデータサイエンス・プラットフォーム



データの加工

- AIプロンプトまたはUIによる効率的なデータ加工
- 計算、結合、データ整形、名寄せ、テキスト変換、など
- データ加工の再現性と自動化

アナリティクス

- 統計的分析（仮説検定、多変量分析、相関分析、など）
- 予測モデルの構築と検証（統計学習、機械学習）
- クラスタリング、因子分析、マーケットバスケット分析、テキスト分析、時系列分析

データの可視化

- 集計データ、時系列、ばらつき、不確実性、地理情報の可視化
- 豊富なチャート機能
- インタラクティブな探索分析

ダッシュボード / レポート

- ビジネス指標のモニタリング
- インタラクティブなダッシュボードの作成
- 分析レポートの再現性と自動化
- AIノートエディター

データ・アクセス

データベース

Amazon Athena	Amazon Aurora	Amazon Redshift	Google BigQuery	MongoDB	Snowflake	MS SQL Azure SQL	MySQL	Postgre SQL
Oracle	Teradata	Vertica	Treasure Data	Salesforce	Presto	Hive	Access	ODBC

クラウド・ストレージ

Amazon S3	Google ドライブ	Google Cloud Storage	MS Azure Storage
--------------	----------------	-------------------------	---------------------

クラウド・アプリケーション

Google Analytics	Google Sheets	Salesforce	FRED 経済指標	株価
---------------------	------------------	------------	--------------	----

ファイル

CSV, 区切りあり	エクセル	Parquet	JSON	ログ	XML	Web Pages	PDF	統計 SPSS SAS SATA
---------------	------	---------	------	----	-----	-----------	-----	---------------------

データアクセス - AI データ

機能：

- 欲しいデータを日本語でリクエスト
- AIが取得用のRスクリプトを生成・実行
- 生成されたスクリプトは確認・修正が可能で再現性を担保
- 一度使った取得手順はスキルとして蓄積、2回目以降が速くなる

実現できることの例：

- 複数のExcelシート・CSVを1つのデータに結合
- PDFやWordの中の表を抽出
- SQLを書かずにデータベースから集計データを取得
- APIを介したデータの取得

AI データヘルプ

列名を対応表の設問文に変換し、値も回答ラベルに変換する

まず両ファイルの詳細な構造を確認します。

- ラベル対応表データの全内容を確認（設問コード・回答コード・ラベルの対応）
Rスクリプトの実行完了 (496ms)
- 回答データの列構成と質問タイプの種類を確認
Rスクリプトの実行完了 (336ms)
- Q11～Q14系（MA設問）の設問コードと設問文の対応を確認
Rスクリプトの実行完了 (255ms)

構造が把握できました。Q11～Q14系はMA（複数選択）で、各選択肢が「S1」、「S2...」という列名になっており、値は0/1 のダミー変数です。これらの列名は「設問文・回答ラベル」の形に変換するのが適切です。スクリプトを生成します。

- ラベル対応表を使って回答データの列名を設問文に、値を回答ラベルに変換してマージ
実行完了 (620ms)

正常に変換できました。右側パネルのプレビューを確認してください。

取得結果の概要は以下のとおりです。

- 行数: 1,000件（回答者1,000名分）
- 列数: 46列

列の変換内容：

変換前（設問コード）	変換後（設問文）
回答者ID	回答者ID（変換なし）
SC1	あなたの性別をお答えください
SC2	あなたの年齢をお答えください

ラベル対応表データ.csv ×

労働時間調査_回答データ.csv ×

+↑

R スクリプトv1: 労働時間調査_回答データ（列名・値を... 実行

```
library(dplyr)
library(tidyr)
library(readr)

# ファイル読み込み
resp <- read_csv("/Users/in/Dropbox/Seninar/AI-Data-Demo/survey_data.csv")
label_raw <- read_csv("/Users/in/Dropbox/Seninar/AI-Data-Demo/survey_label.csv")

# 設問コードごとに設問文・質問タイプを補完（fill）
label_filled <- label_raw |>
  fill(設問コード, 設問文, カテゴリー, 質問タイプ, .direction = "down")

# ---- 1. SA / NUM 列の値を回答ラベルに変換 ----
# SA列: 回答コード → 回答ラベル のマッピングを列ごとに適用
sa_cols <- label_filled |>
  filter(質問タイプ == "SA") |>
  pull(設問コード) |>
```

データ解説

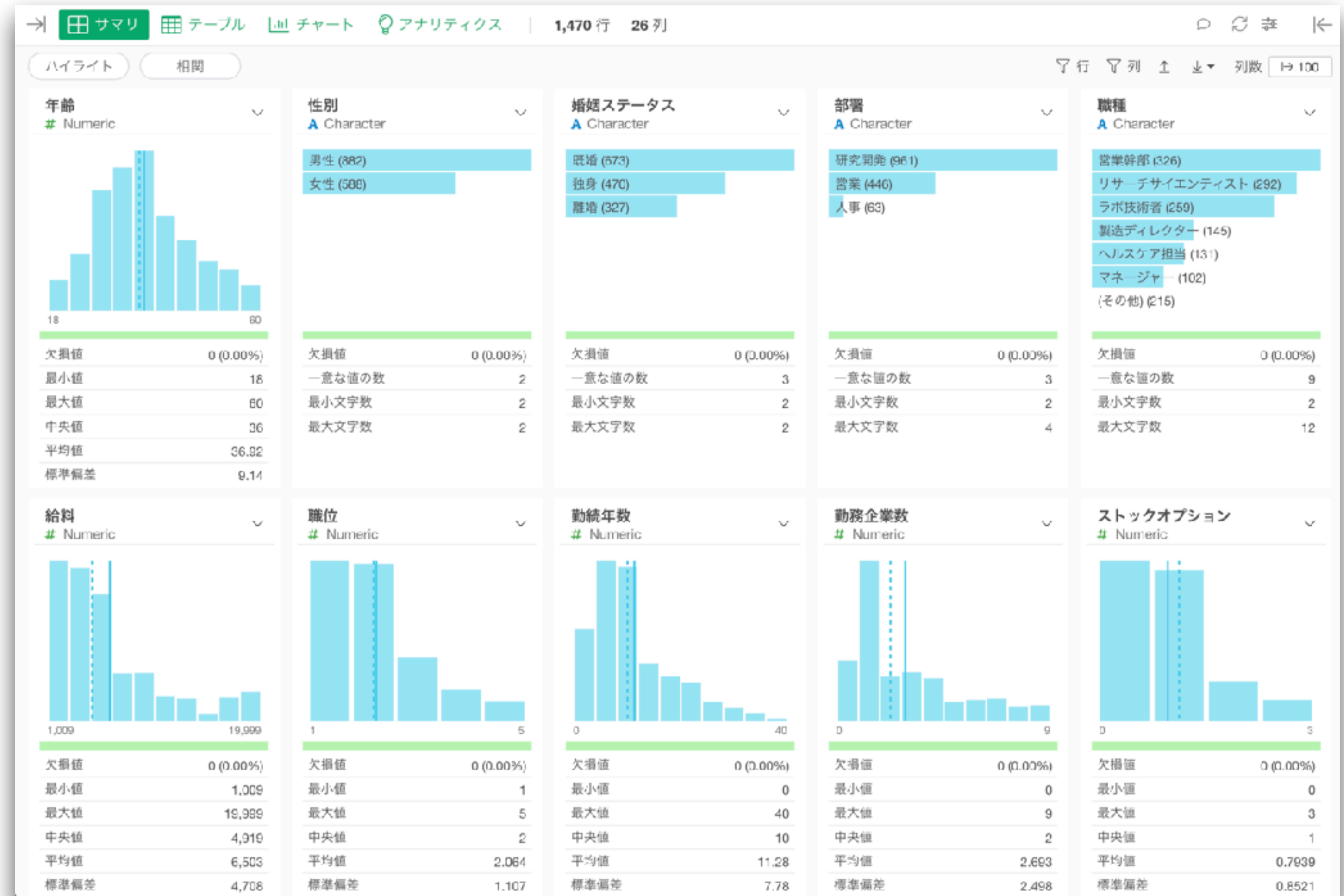
データ: 1,000 行 46 列 表示: 100 行 46 列

	回答者ID A Character	あなたの性別をお答えくだ... A Character	あなたの年齢をお答えくだ... Numeric	あなたの現在の役職 A Character
1	R0001	男性	34	一般社員
2	R0002	女性	30	課長
3	R0003	男性	50	課長
4	R0004	女性	48	部長以上
5	R0005	女性	25	一般社員
6	R0006	男性	39	主任・係長
7	R0007	男性	31	主任・係長
8	R0008	女性	22	一般社員
9	R0009	女性	42	部長以上
10	R0010	男性	39	課長

保存キャンセル

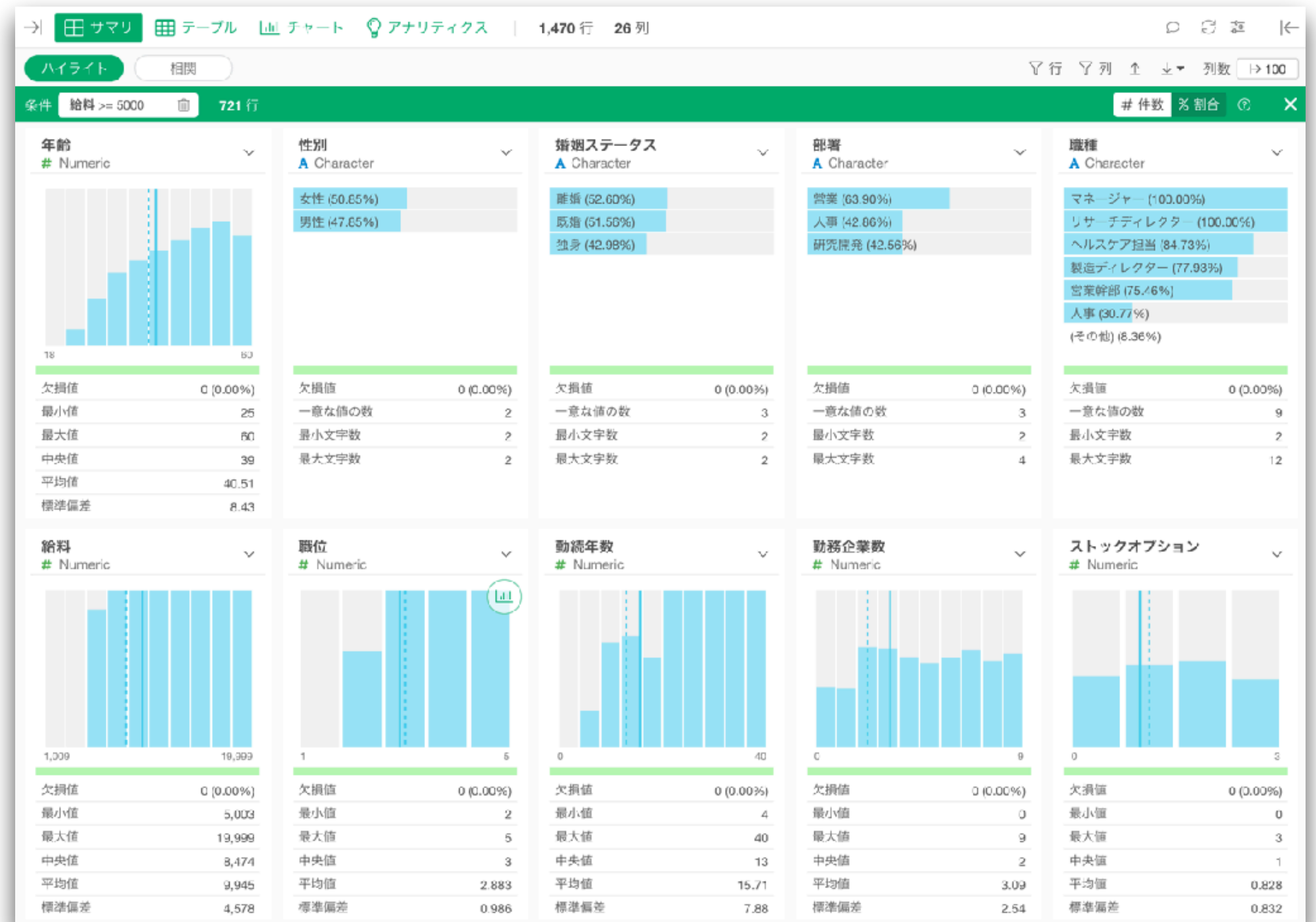
探索的分析 - サマリー・ビュー

- データのサマリー情報を高速で自動的に表示
- それぞれのデータ・タイプに適した可視化を自動的に選択
- 統計値、欠損値などを自動検出



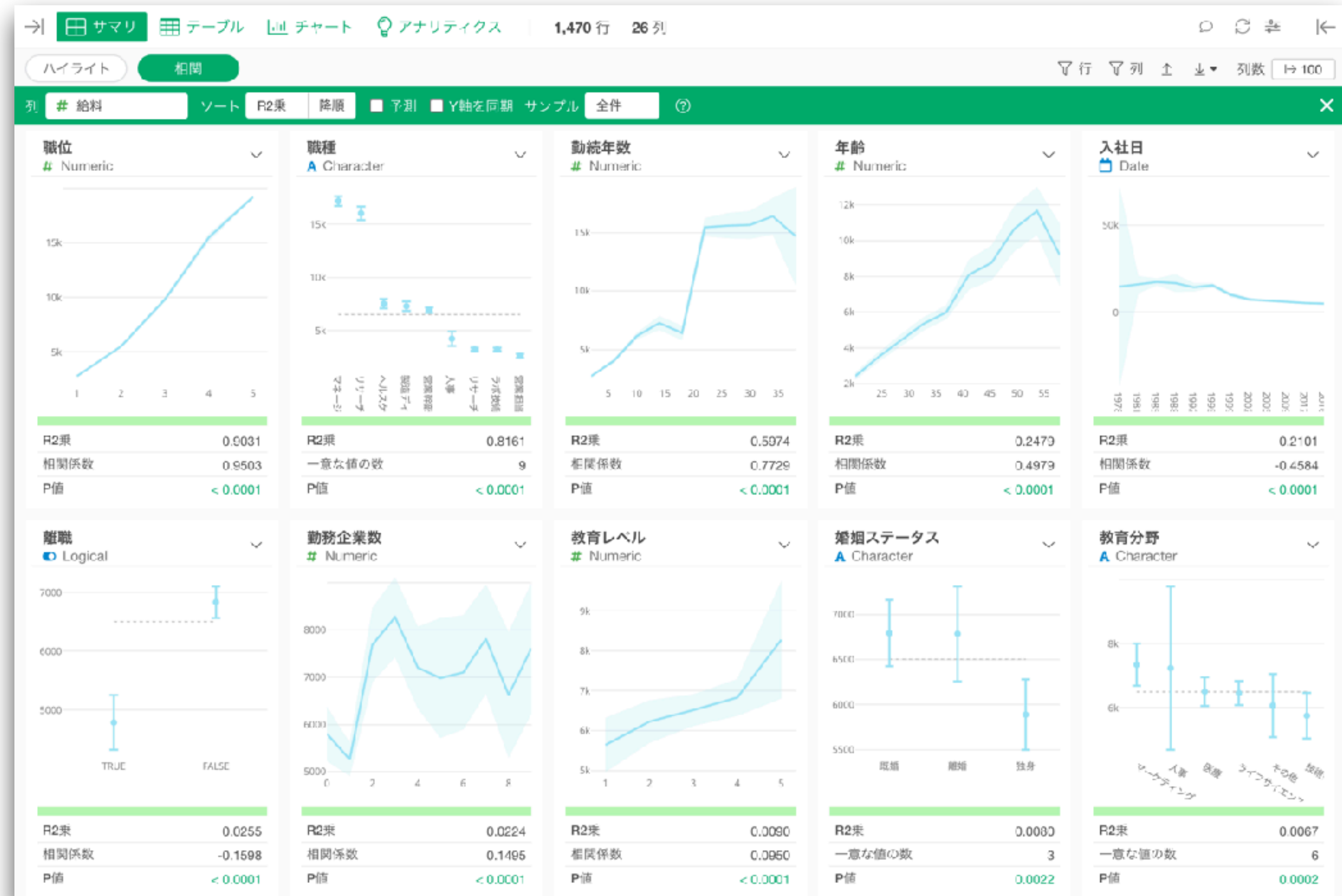
探索的分析 - ハイライト

- 設定した条件に合致するデータの件数、割合でハイライト
- データに隠れたパターンを素早く特定
- 外れ値や欠損値の特徴を素早くつかむことが可能



探索的分析 - 相関

- 興味のある対象列と全ての列の相関関係を自動で検証
- R2乗、AUC、P値、相関係数を自動計算
- 多変量の統計モデル、機械学習のモデルの構築が可能



データの加工- AI プロンプト

機能：

- 日本語でやりたいことをリクエスト
- リクエストを叶えるRスクリプトをAIが生成
- Rスクリプトの詳細な説明
- スクリプトの編集
- プレビュー

データラングリング（加工、整形） with AI プロンプト ヘルプ

データの加工に関するタスクを頼むことができます。（例：開始日の列を日付型に変換してください） サンプルサイズ

↑

Rコマンド

R スクリプト 編集

```
mutate(`産業名` = str_extract(`企業情報`, "[A-Za-z &-]+$"))
```

使用している関数の説明

- `mutate()` : 新しい列を作成または既存の列を変更する関数です。ここでは「産業名」という新しい列を作成しています。
- `str_extract()` : 文字列から特定のパターンに一致する部分を抽出する関数です。
 - 第1引数: 抽出元の列名（[企業情報](#)）
 - 第2引数: 抽出するパターンを指定する正規表現。 `[A-Za-z &-]+$` は「英字、スペース、アンパサンド、ハイフンの1つ以上の連続で、文字列の末尾にあるもの」を意味します。

期待される結果

このコードを実行すると、以下の処理が行われます：

- 「企業情報」列から産業名（文字列の末尾にある英字部分）を抽出し、「産業名」という新しい列に格納します。

具体例として：

リセット ステップとして実行 キャンセル

データの加工- AI 関数

企業名 A Character	正式な企業名 A Character
トヨタ自動車	トヨタ自動車株式会社
トヨタ自動車株式会社	トヨタ自動車株式会社
トヨタ自動車(株)	トヨタ自動車株式会社
トヨタ	トヨタ自動車株式会社
本田技研工業	本田技研工業株式会社
本田技研工業株式会社	本田技研工業株式会社
本田技研工業（株）	本田技研工業株式会社
ホンダ	本田技研工業株式会社
本田	本田技研工業株式会社
日産自動車	日産自動車株式会社
日産自動車株式会社	日産自動車株式会社
日産自動車(株)	日産自動車株式会社
ニッサン	日産自動車株式会社
日産	日産自動車株式会社
マツダ	株式会社マツダ
株式会社マツダ	株式会社マツダ
（株）マツダ	株式会社マツダ
MAZDA	株式会社マツダ
スズキ	スズキ株式会社
スズキ株式会社	スズキ株式会社
スズキ(株)	スズキ株式会社
SUZUKI	スズキ株式会社
SUBARU	株式会社SUBARU

AI モデル関数の作成

☐ 既存列を上書き ☒ 新しく列を作成

正式な企業名

関数に渡す列

AI プロンプト

☒ 新しいプロンプトを書く ☐ テンプレートを使用

コピー 新規テンプレートとして保存

各組織名について、正式な法人名称を返してください。

☒ データを分割して並列処理をする ⓘ

リセット 実行 閉じる

機能：

- 日本語で実現したい処理をリクエスト
- 各行にAIから出力された値を生成
- テンプレートを保存して関数と再利用
- 関数が存在しないタイプのデータ加工を実現

実現できることの例：

- 文字列の表記揺れの修正
- 文章からセンチメントスコアを算出
- 文章をラベル付け、など

データの加工 - UI

効率的な分析のためのデータの加工

- 複数のデータの結合
- テキスト・データの加工
- 条件分岐を使った計算
- 欠損値を予測手法により補完
- 条件つき集計
- IPアドレスから国・地域情報への変換
- 時系列情報の生成

他にもたくさん...!

The screenshot displays the Tableau Desktop interface with a data table. The table has the following columns: 注文日 (Order Date), 出荷日 (Shipping Date), 配送タイプ (Shipping Type), 顧客ID (Customer ID), 顧客名 (Customer Name), 顧客セグメント (Customer Segment), 郵便番号 (Postal Code), 市 (City), 州 (State), and 国 (Country). A context menu is open over the '顧客ID' column, listing various data manipulation actions. The menu options include: 列名を変更 (Change column name), 列を選択 / 削除 (Select / Delete column), 列を並び替える (Reorder column), フィルタ (Filter), グループ化 (Group), 集計 (Summarize) (with a sub-menu for Sum, Average, Count, etc.), 計算を作成 (Create calculation), 表計算を作成 (Create table calculation), 並び替え (ソート) (Sort), 結合 (列を追加する / Join) (Join), マージ (行を追加する / Union) (Merge), ワイド型からロング型へ (Convert wide to long), ロング型からワイド型へ (Convert long to wide), ピボット (Pivot), 分割 (Split) (with sub-menus for splitting by column, row, and location), 複数の列をつなげる (Unite) (Concatenate), データタイプを変換 (Convert data type), 値を置き換える (Replace values), テキストデータの加工 (Text data manipulation), 欠損値 (NA) の加工 (Handle missing values), 'その他'グループを作る (Create 'Other' group), 値の順序をセット (Set value order), アナリティクス (Analytics), 予測 (アナリティクス・ビュー) (Predict (Analytics view)), and 仮置 (Placeholder).

	注文日 Date	出荷日 Date	配送タイプ Character	顧客ID Character	顧客名	顧客セグメント	郵便番号 Numeric	市 Character	州 Character	国 Character
1	2019-11-12	2019-11-14	ファーストクラス	AB-100151		ユーザー	73,120	Oklahoma City	Oklahoma	United States
2	2019-02-06	2019-02-08	セカンドクラス	JR-162107			<NA>	Wollongong	New South Wales	Australia
3	2019-10-18	2019-10-19	ファーストクラス	CR-127307		ユーザー	<NA>	Brisbane	Queensland	Australia
4	2019-01-29	2019-01-31	ファーストクラス	KM-163754		業主	<NA>	Berlin	Berlin	Germany
5	2019-11-06	2019-11-07	当日配送	RH-949511		ユーザー	<NA>	Dakar	Dakar	Senegal
6	2019-06-29	2019-07-02	セカンドクラス	JM-156557			<NA>	Sydney	New South Wales	Australia
7	2017-11-07	2017-11-09	ファーストクラス	TS-213409		ユーザー	<NA>	Porirua	Wellington	New Zealand
8	2018-04-15	2018-04-19	標準クラス	MB-180355		ユーザー	<NA>	Hamilton	Waikato	New Zealand
9	2019-11-12	2019-11-14	ファーストクラス	AB-100151		ユーザー	73,120	Oklahoma City	Oklahoma	United States
10	2017-03-07	2017-03-08	ファーストクラス	AB-100151		ユーザー	98,103	Seattle	Washington	United States
11	2017-03-07	2017-03-08	ファーストクラス	AB-100151		ユーザー	98,103	Seattle	Washington	United States
12	2018-04-20	2018-04-23	ファーストクラス	AJ-107801			<NA>	Kabul	Kabul	Afghanistan
13	2017-12-27	2017-12-29	セカンドクラス	MM-72601					Jizan	Saudi Arabia
14	2018-11-14	2018-11-14	当日配送	VF-217151					Parana	Brazil
15	2019-06-07	2019-06-09	セカンドクラス	PF-191202			<NA>		Heilongjiang	China
16	2020-07-31	2020-08-03	セカンドクラス	BP-111854			<NA>		Ile-de-France	France
17	2017-02-19	2017-02-25	標準クラス	AB-100151		ユーザー	76,01		Texas	United States
18	2020-09-08	2020-09-14	標準クラス	PJ-188356			<NA>		Tuscany	Italy
19	2020-02-01	2020-02-02	ファーストクラス	JS-156857			<NA>		Queensland	Australia
20	2020-12-05	2020-12-07	セカンドクラス	RH-955512		ユーザー	<NA>		Kigoma	Tanzania
21	2018-08-09	2018-08-11	ファーストクラス	AB-600103			<NA>		Silesia	Poland
22	2017-03-07	2017-03-08	ファーストクラス	AB-100151		ユーザー	98,10		Washington	United States
23	2017-05-02	2017-05-03	ファーストクラス	JK-1532527	Jason Klamczynski	法人	<NA>		Anhui	China

再現性

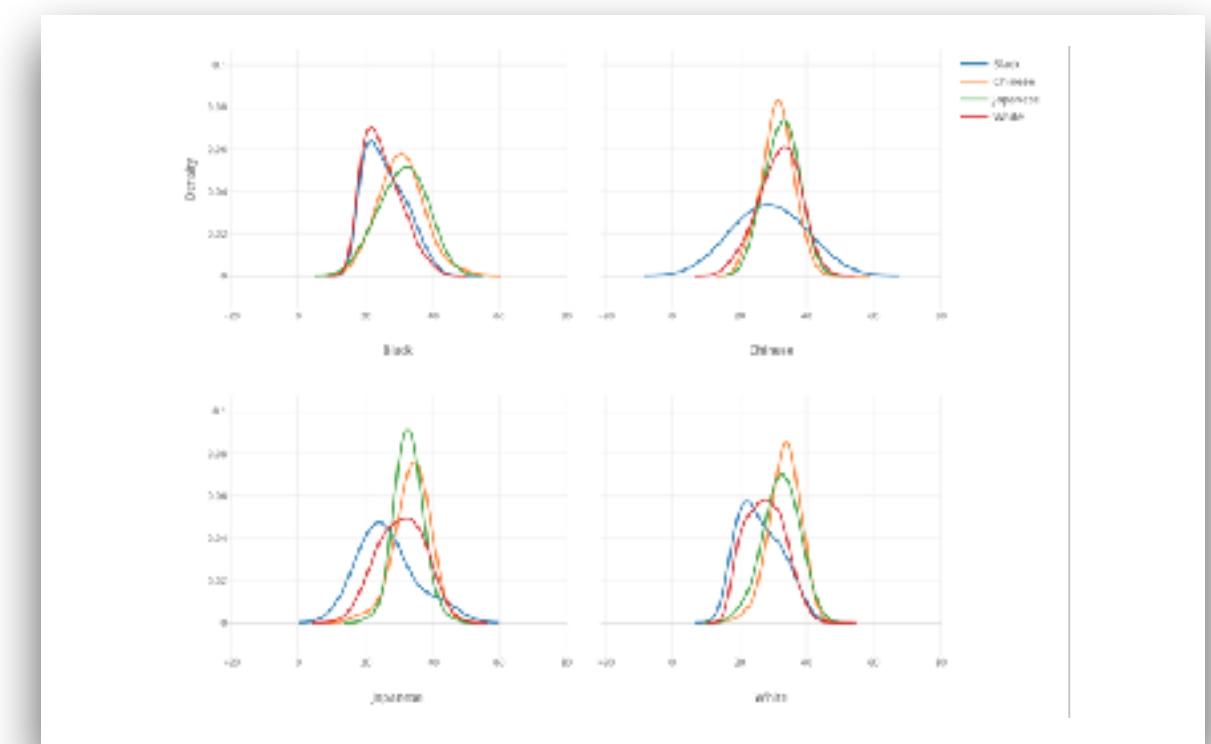
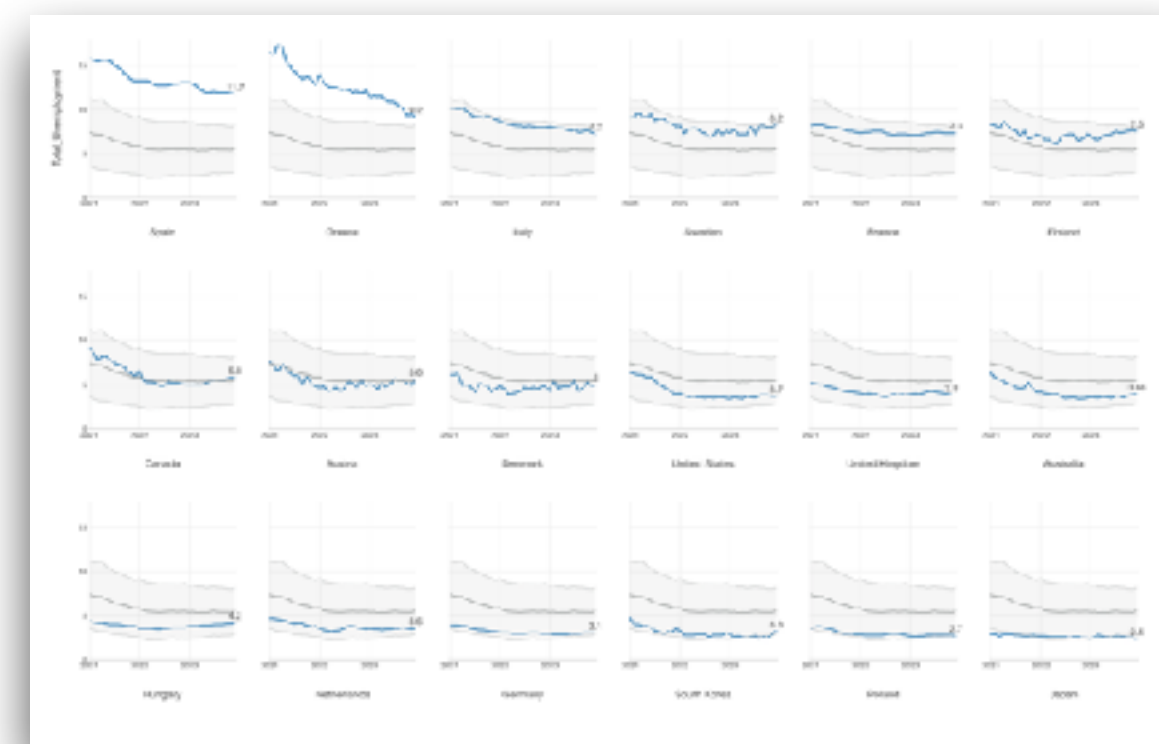
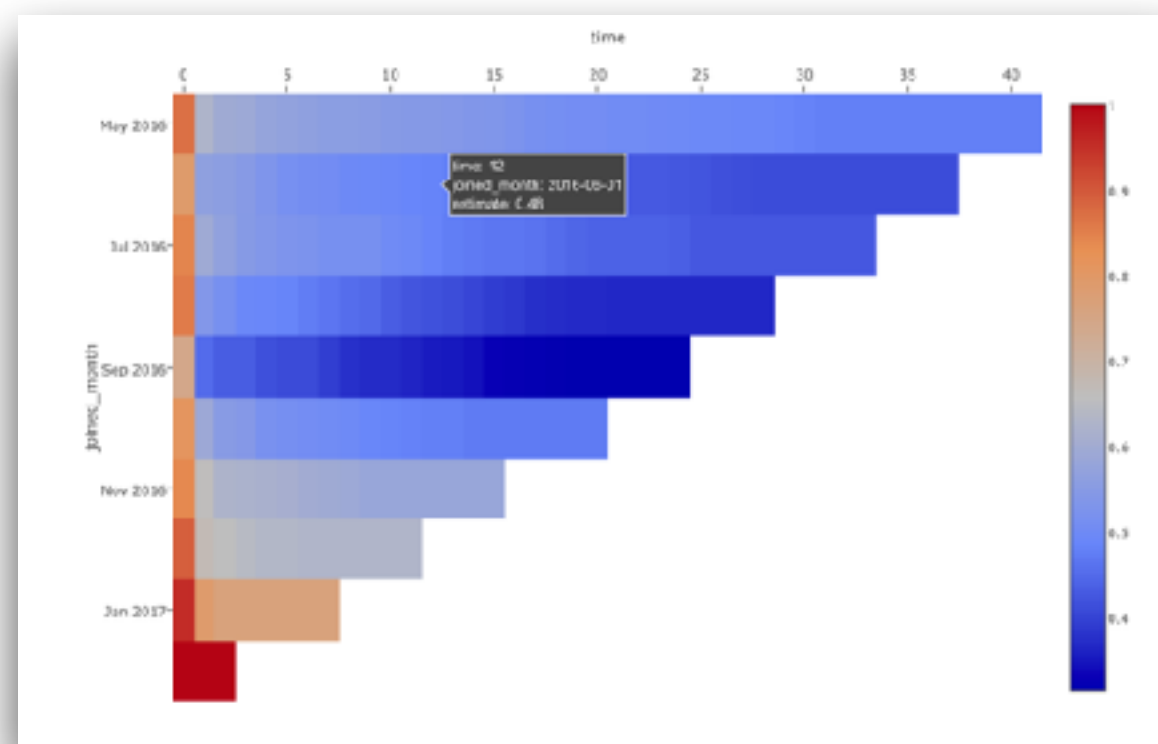
- データを更新した際は関連付けされている**全てのデータが自動的に更新**

The screenshot displays a data analysis tool interface. At the top, a navigation bar includes tabs for 'サマリ' (Summary), 'テーブル' (Table), 'チャート' (Chart), and 'アナリティクス' (Analytics). The 'テーブル' tab is active, showing a table with 51,290 rows and 17 columns. A modal window titled 'Rスクリプトを生成' (Generate R Script) is open, displaying R code for data processing. The code includes library imports, token setting, data retrieval from Google Drive, cleaning, and output generation. A green button 'クリップボードにコピー' (Copy to Clipboard) and a '閉じる' (Close) button are at the bottom of the modal. The right sidebar contains a 'ステップ' (Steps) panel with 14 numbered steps for data manipulation, including sorting, calculation, and summarization. A table of data is visible at the bottom of the interface.

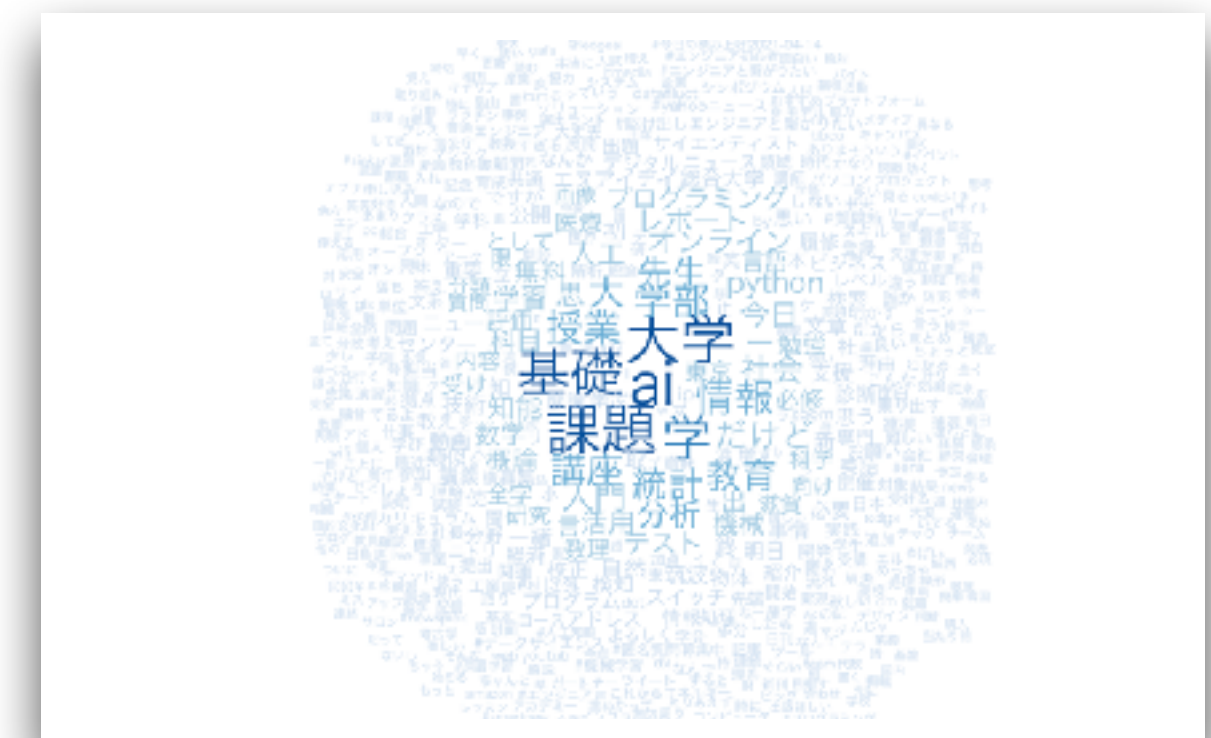
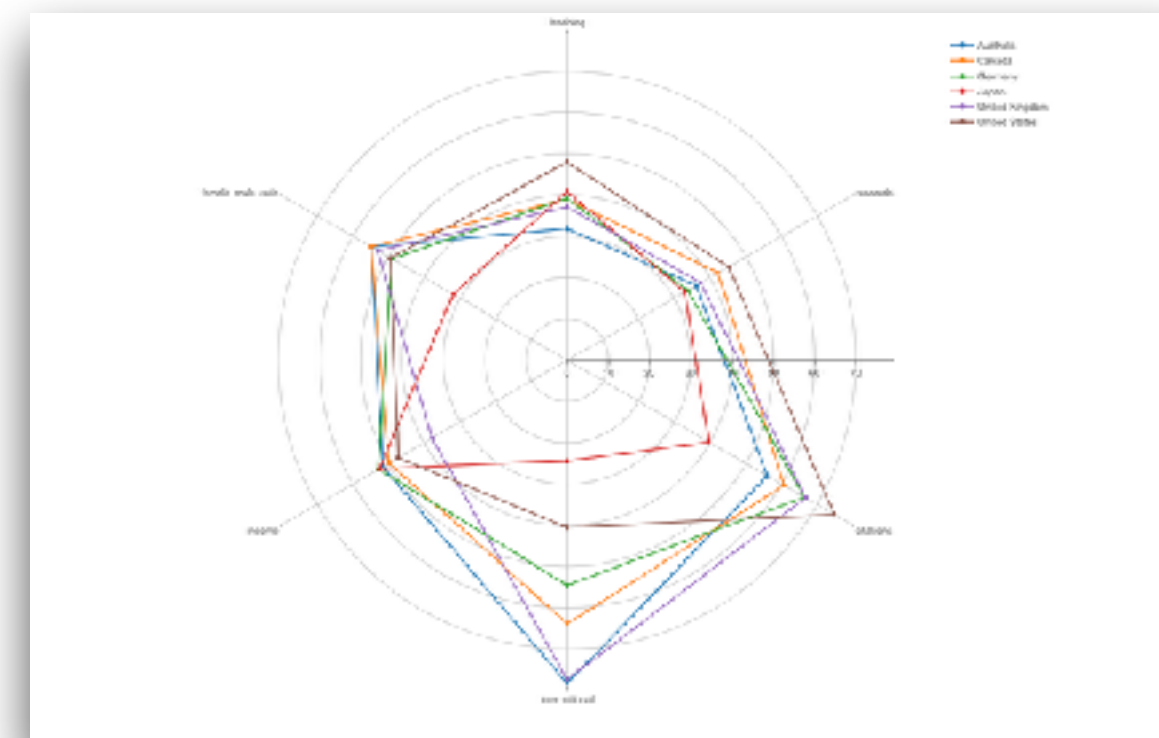
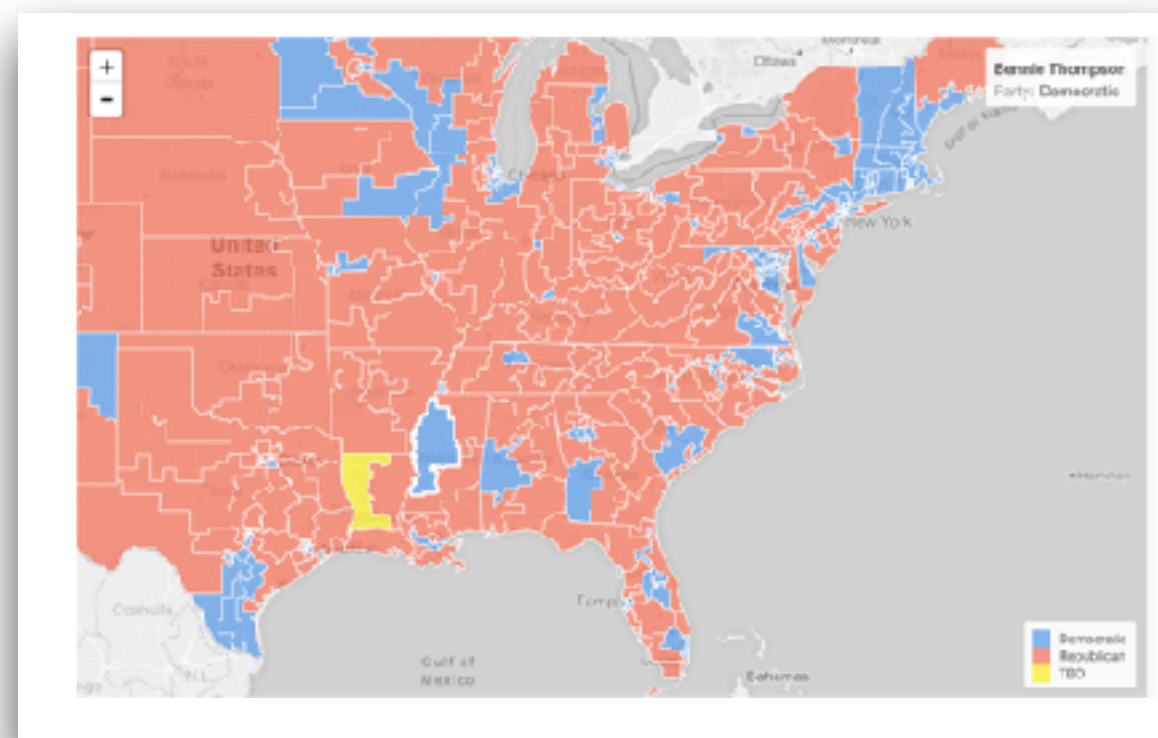
行	列	行	列	行	列	行	列
21	2019-08-09	2019-08-11	ファーストクラス	AB-800103	Ann Blume	法人	FUR-1A-4644
22	2018-03-07	2018-03-08	ファーストクラス	AB-100151404	Aaron Bergman	コンシューマー	OFF-ST-3744

データの可視化

データ・サイエンス、統計で必要になる豊富な可視化のオプション

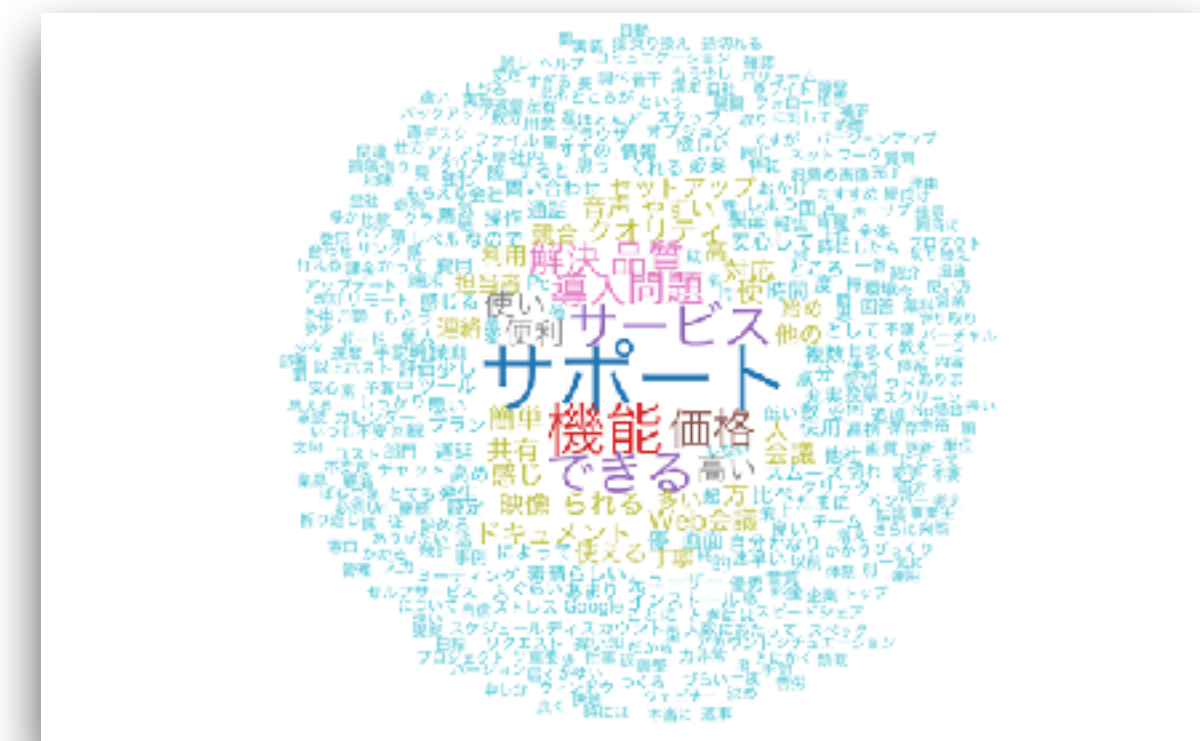
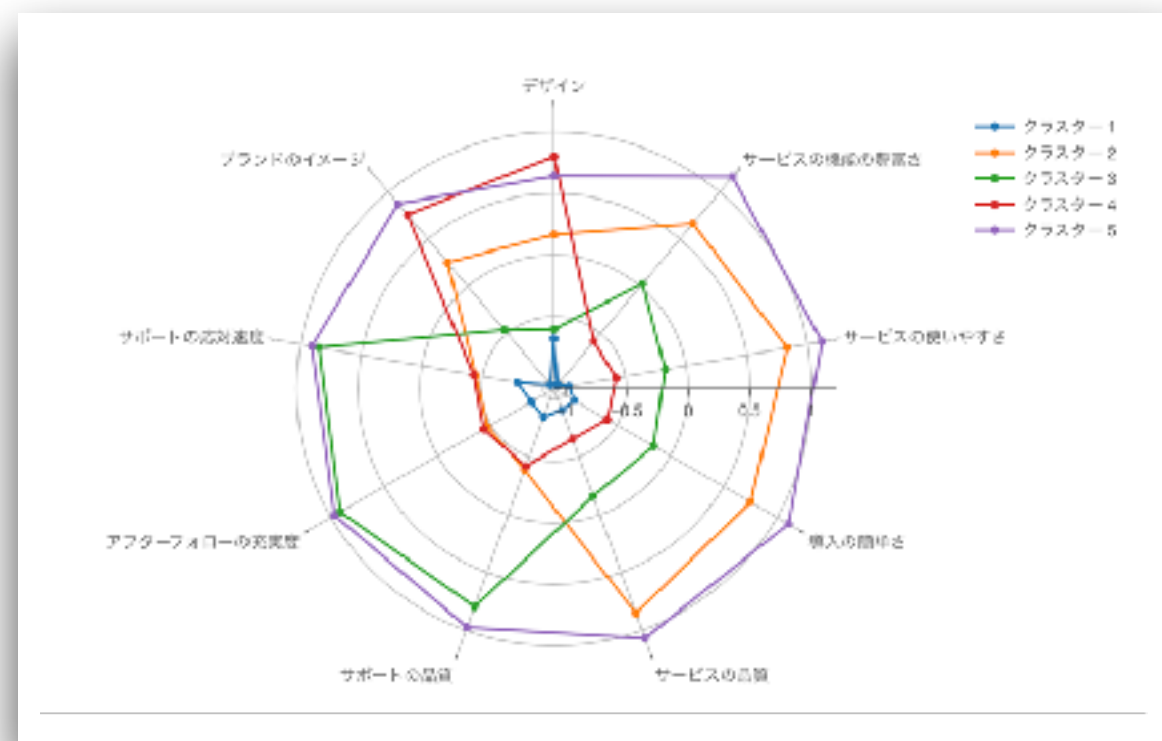
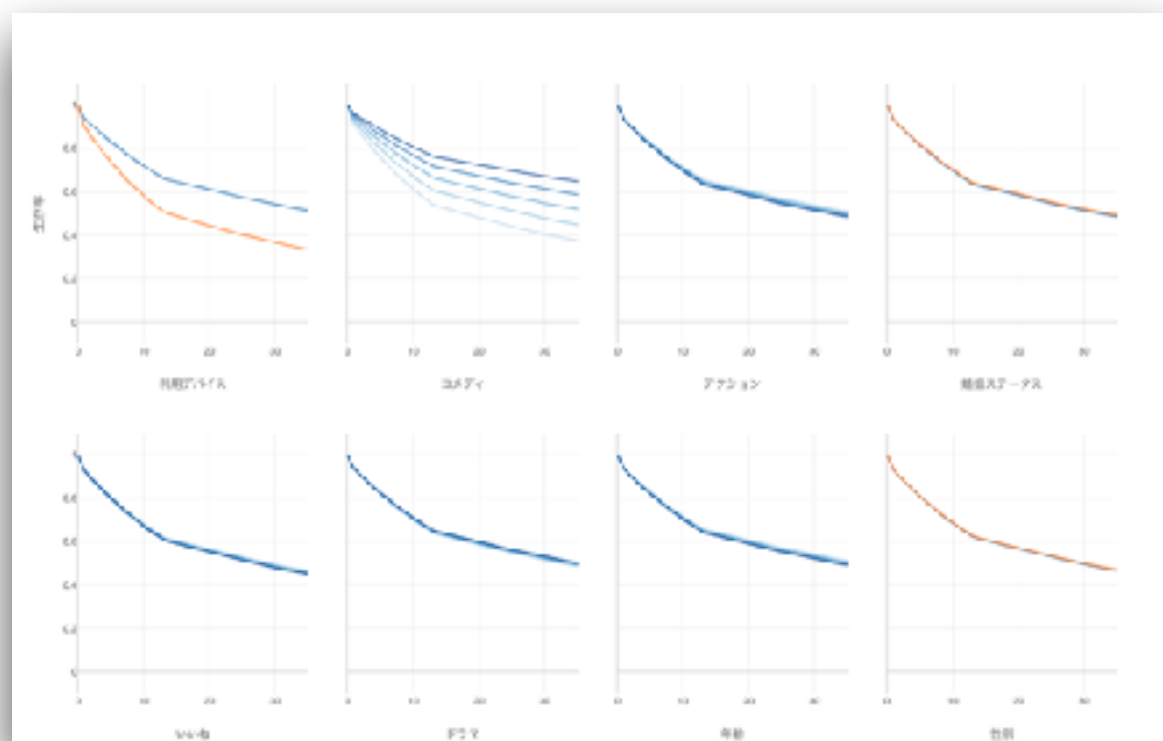
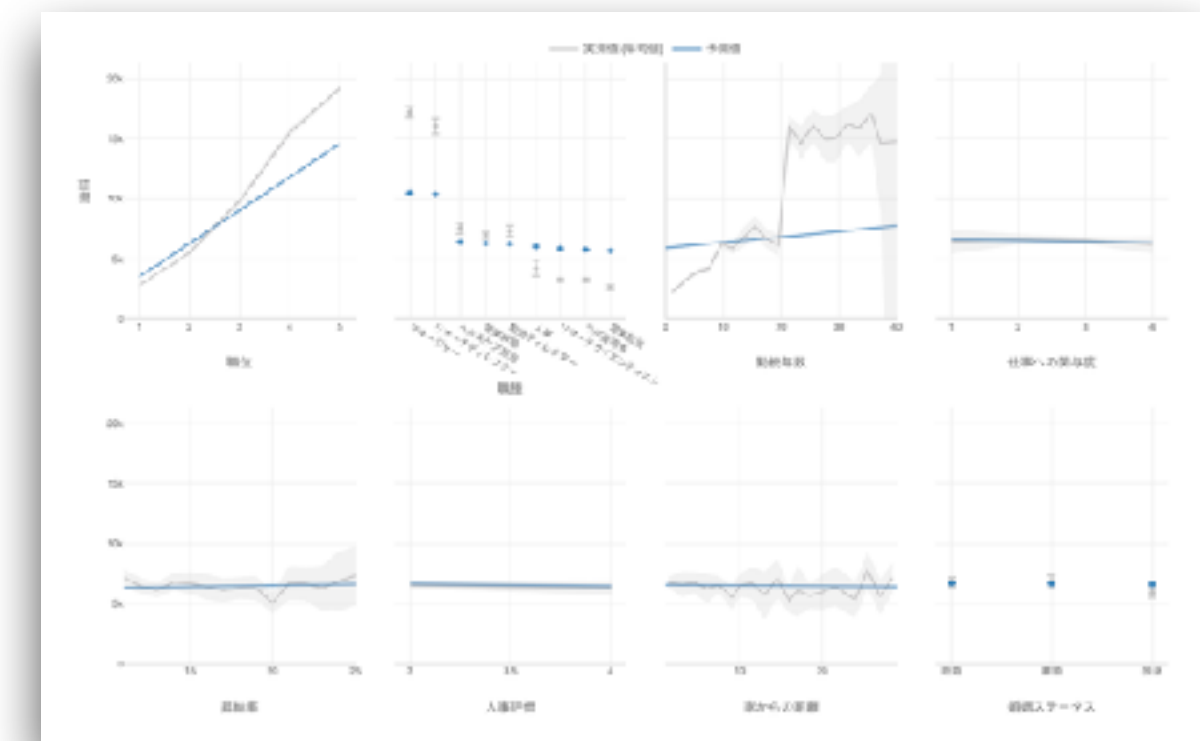
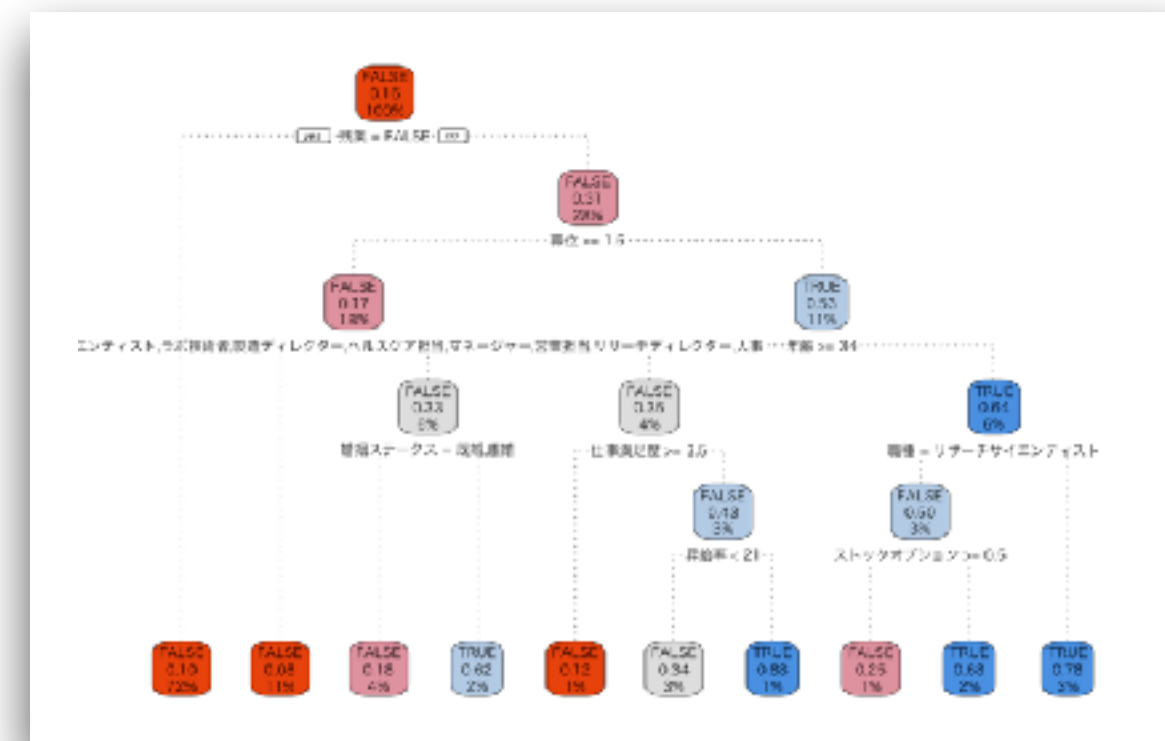
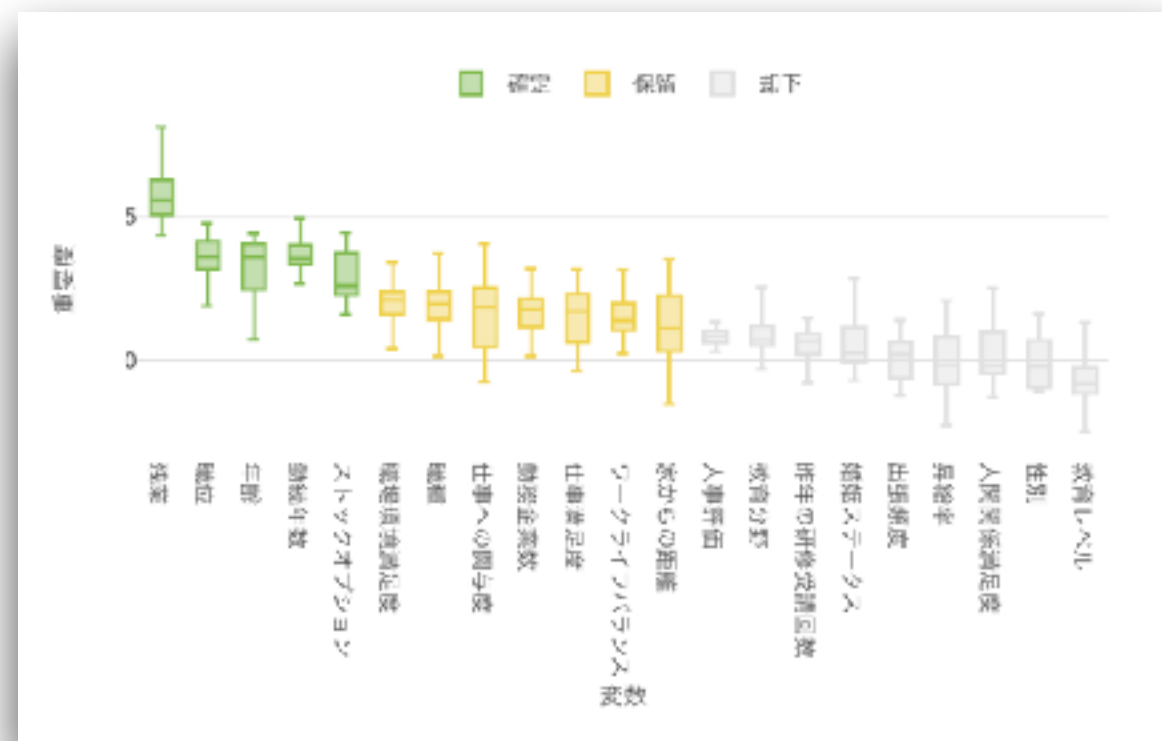
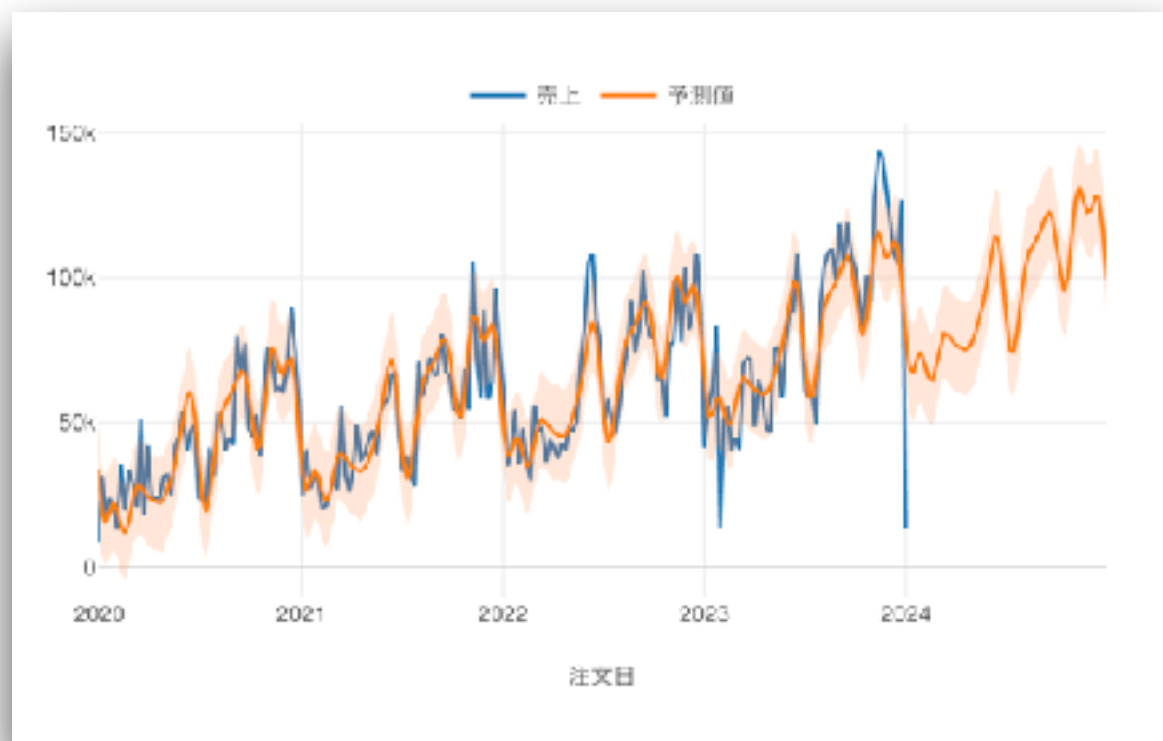


Market	Category	Sales (\$)	# of Orders	Marketing Spent	Sales Comp
Africa	Technology	\$322,367.04	911	\$74,693.54	39,649.49
	Office Supplies	\$269,755.53	3,045	\$28,748.41	31,755.94
	Furniture	\$194,890.84	631	\$44,404.41	20,434.04
Asia Pacific	Technology	\$1,537,454.7	3,034	\$299,291.3	169,364.42
	Furniture	\$1,461,651.3	2,323	\$281,368.54	158,542.3
	Office Supplies	\$1,043,652.27	8,345	\$187,163.06	110,831.44
Europe	Technology	\$1,233,509.06	2,234	\$234,861.06	130,636.37
	Office Supplies	\$1,163,880.58	7,718	\$229,611.07	125,392.83
	Furniture	\$860,106.56	1,777	\$170,314.68	92,870.83
LATAM	Furniture	\$611,670.62	2,382	\$127,285.21	97,900.57
	Technology	\$788,713.82	2,050	\$160,226.42	85,470.36
	Office Supplies	\$563,990.75	5,882	\$109,940.02	69,254.69
USCA	Technology	\$682,452.84	1,312	\$191,261.48	88,656.25
	Furniture	\$752,172.79	2,147	\$41,854.86	81,438.62
	Office Supplies	\$749,503.4	6,319	\$167,209.36	79,662.55
Total		\$12,642,601.91	51,290	\$2,368,103.46	1,358,085.7



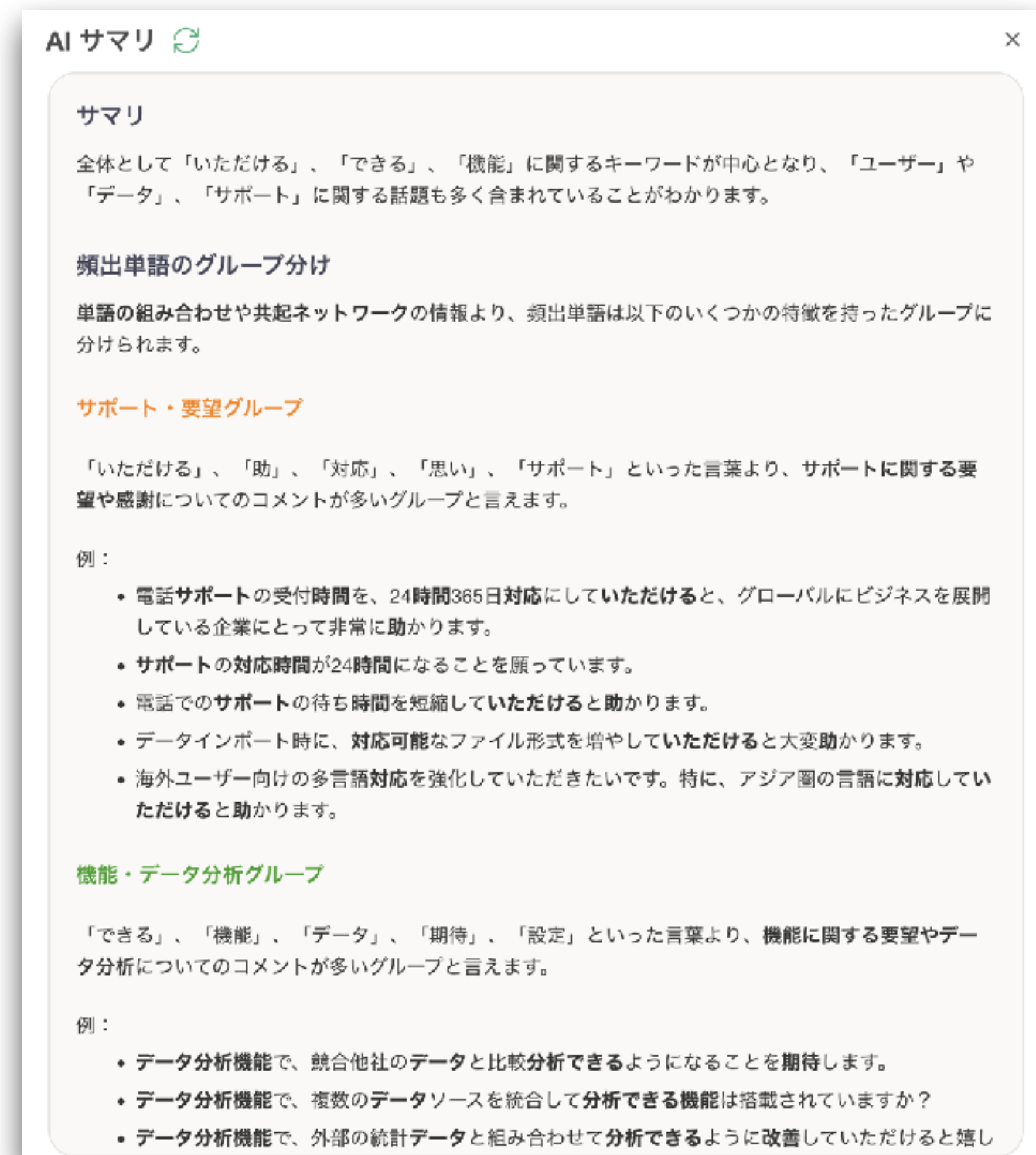
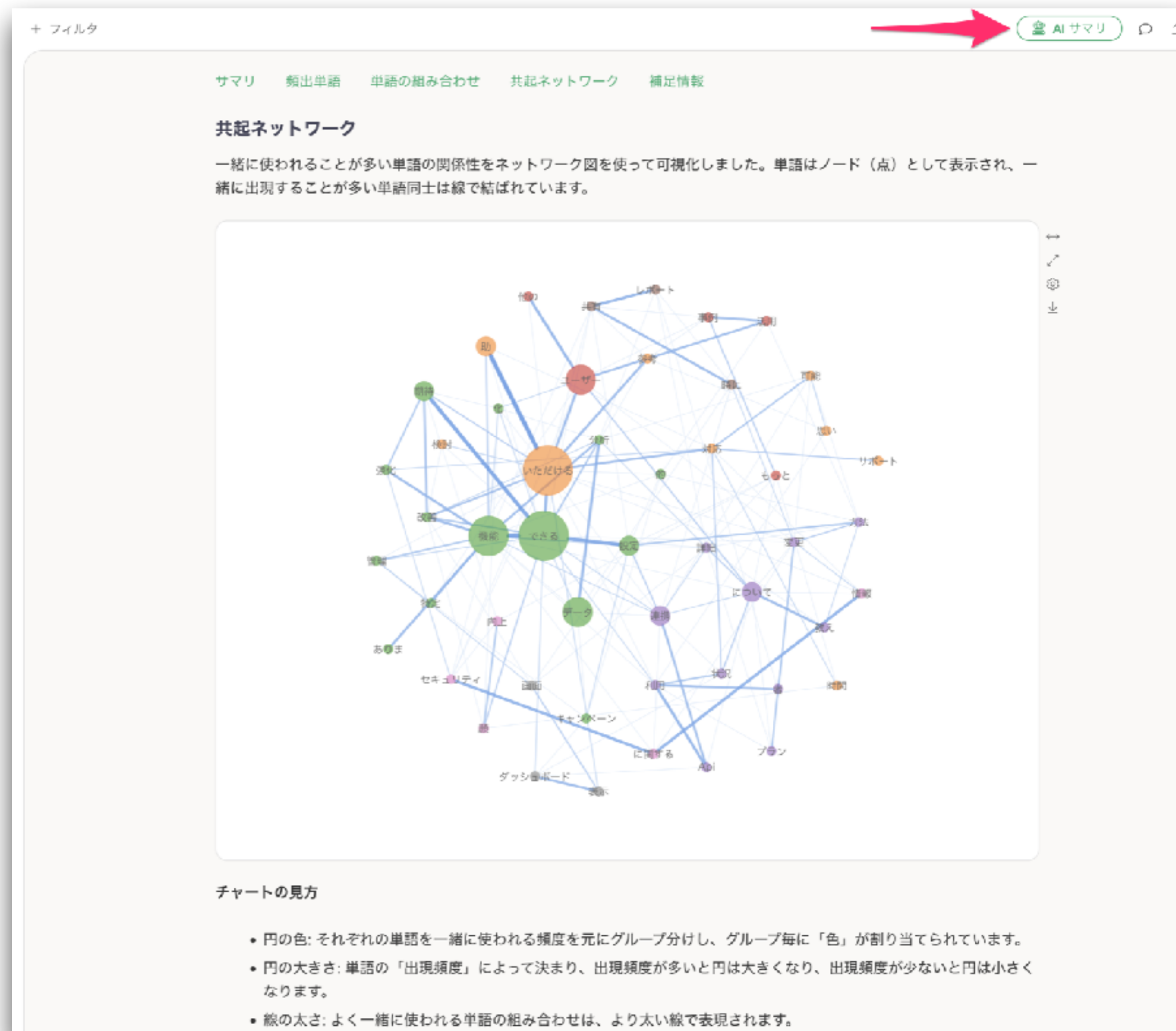
アナリティクス

様々な機械学習、統計アルゴリズムに簡単にアクセス



アナリティクス - 分析レポートの自動生成

分析結果をレポートとして自動生成、さらにAIによる要約機能



ノート

分析より得たインサイトを効果的に伝えるためのノート

- インタラクティブなチャート
- 数式
- Rスクリプトによる拡張
- イメージ
- 柔軟なテキスト入力
- パラメーターによるデータの動的更新



ノート - AI エディター



機能：

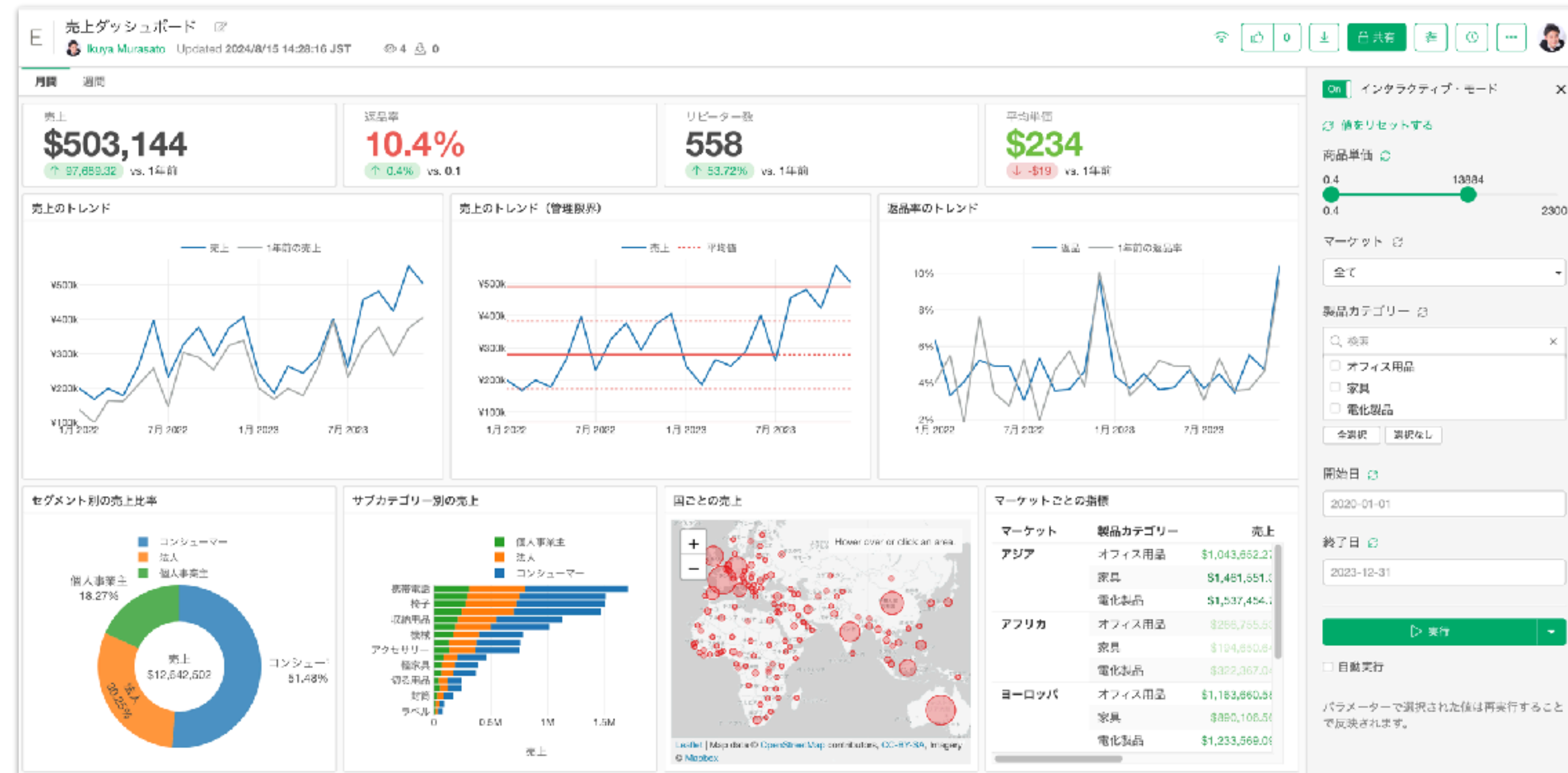
AIを使ったノートの自動生成と編集

- 分析レポートの自動生成
- チャートの自動解釈
- カスタムプロンプトの登録
- 文章の要約
- 文章の改善
- 他言語への翻訳

ダッシュボード

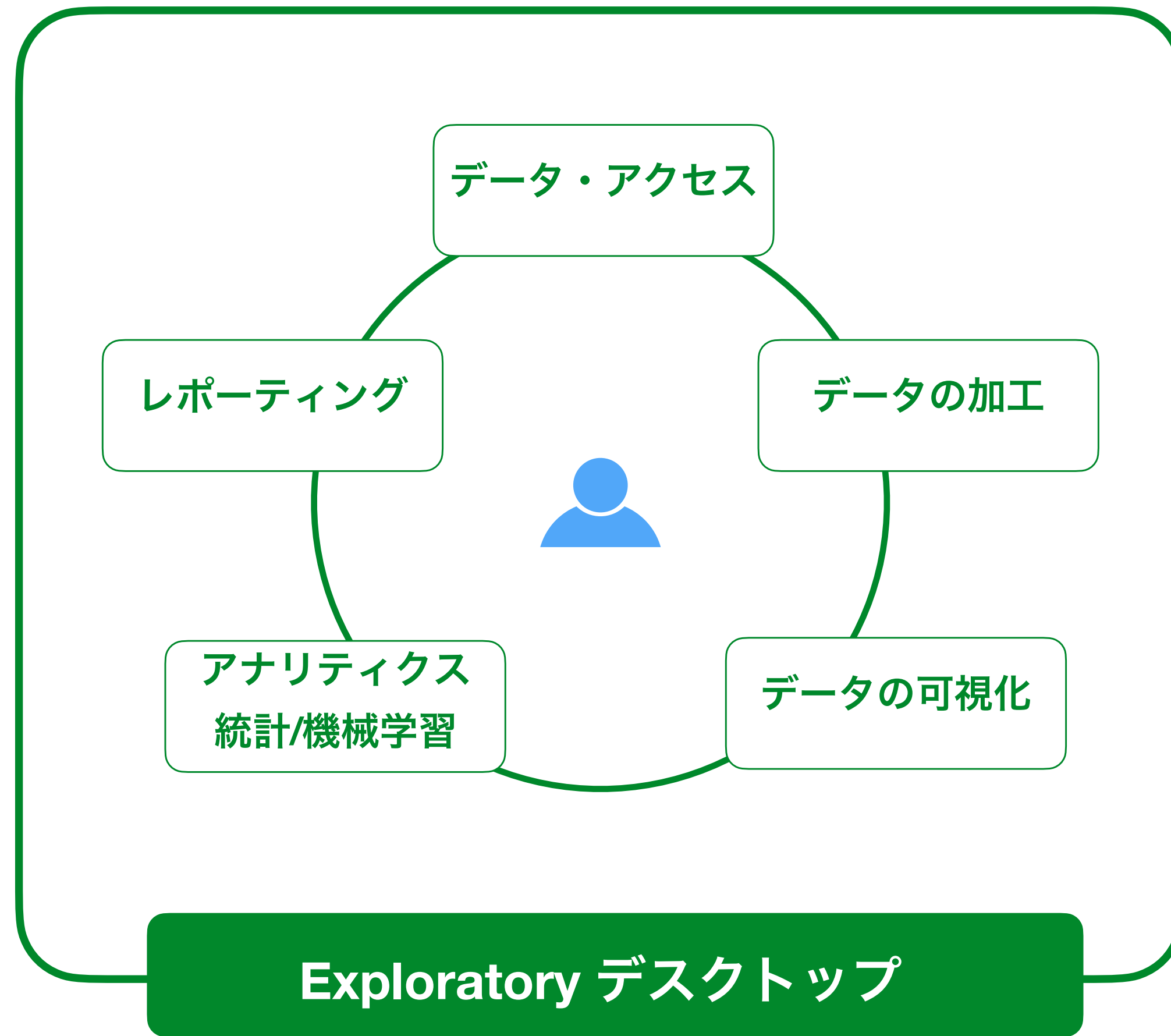
ビジネス改善に必要な指標を
モニター

- 指標・KPI
- インタラクティブ・チャート
- アナリティクスのチャート
- パラメーターによるデータの動的更新
- 自動化のためのスケジュール



Exploratory

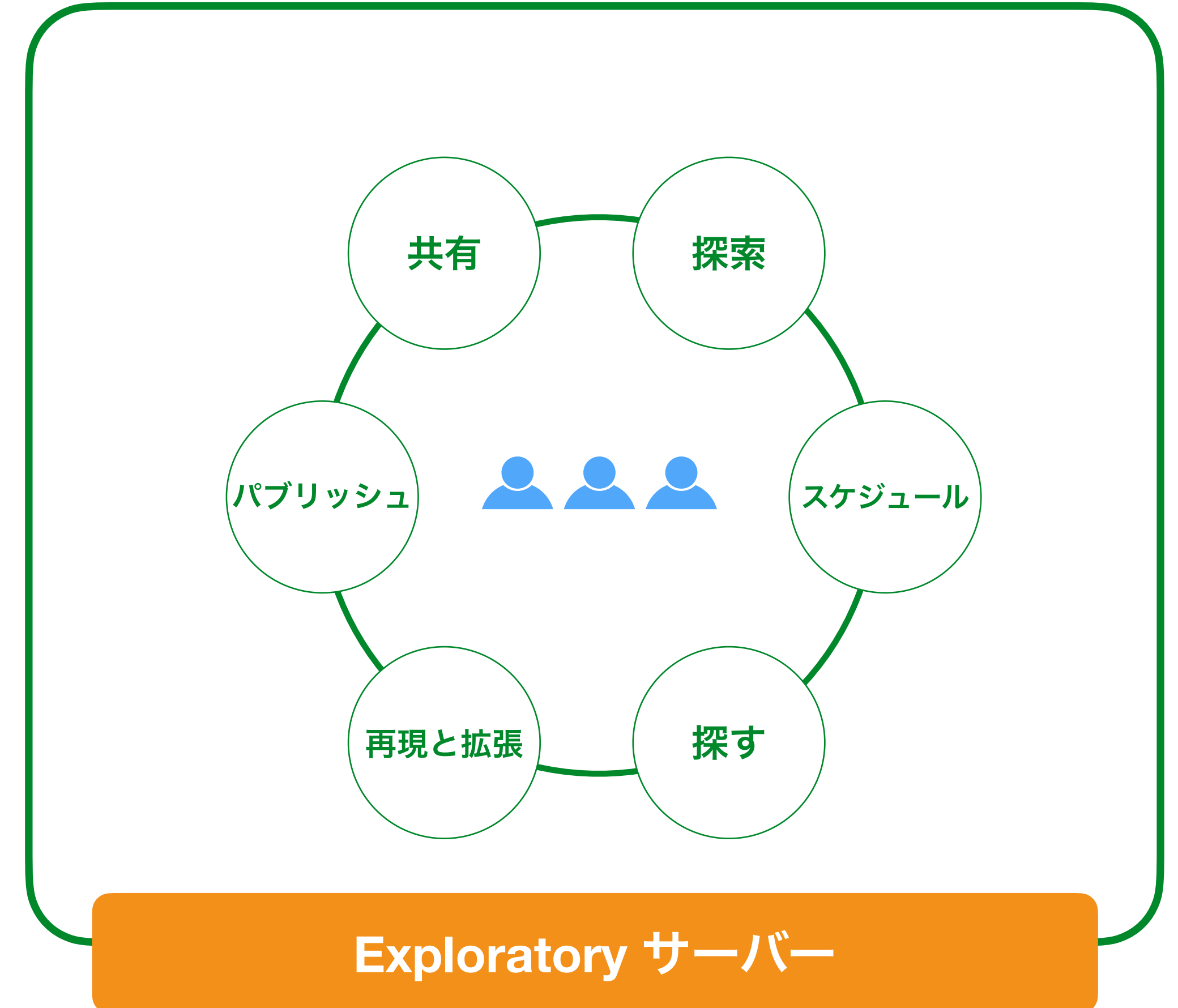
データからインサイトへ



PC (Mac / Windows)

➡
パブリッシュ

インサイト & データの一元化



Exploratory サーバー

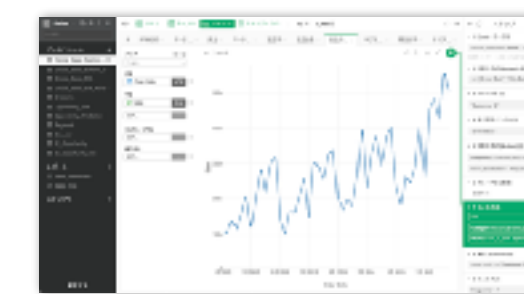
Exploratory サーバー

インサイト・カタログ



データ・カタログ

プロジェクト



検索・閲覧

共有の管理

インタラクティブ

ダウンロード・API

スケジュール

バージョン管理

Exploratory
デスクトップ

Exploratory
デスクトップ

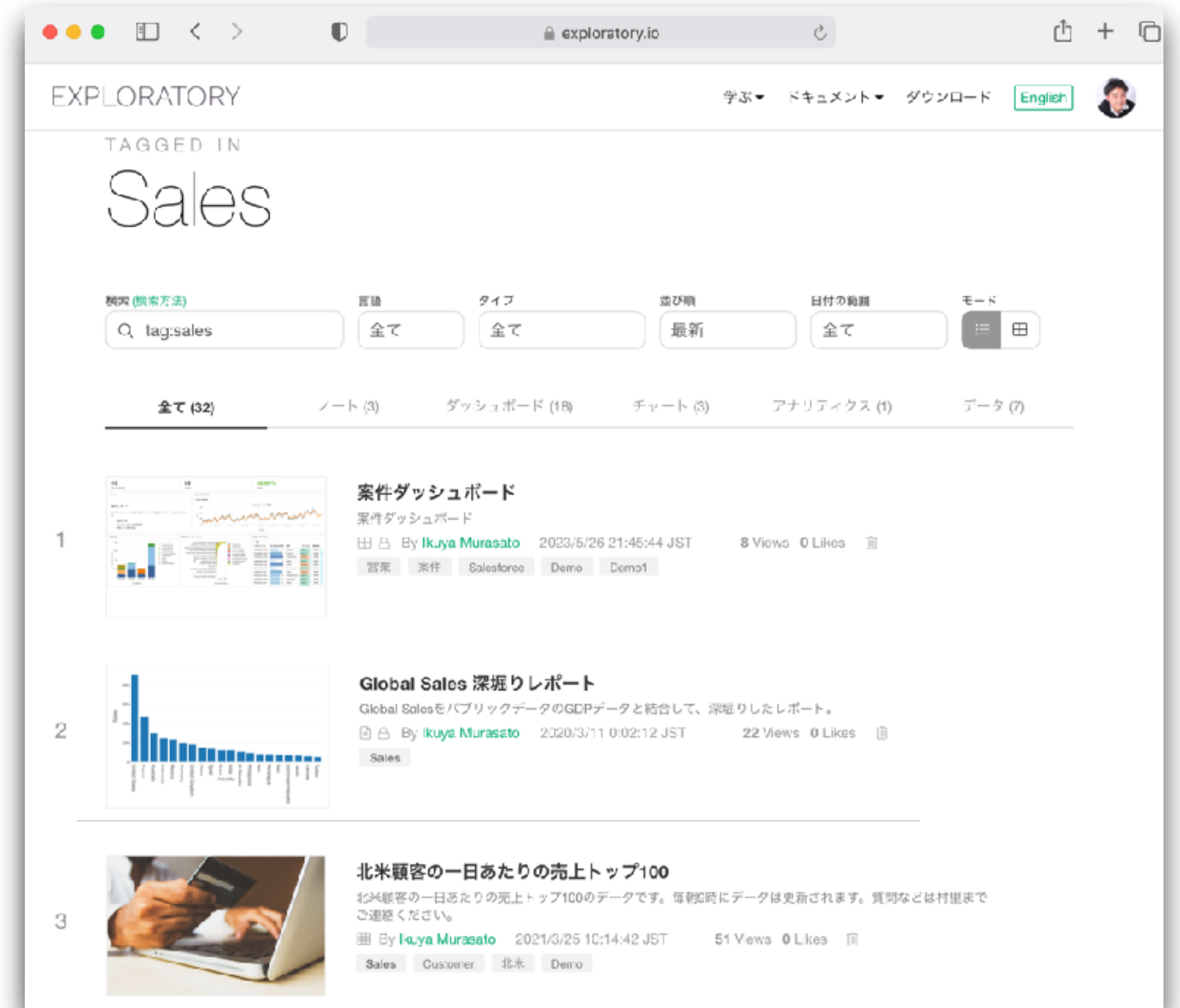
Exploratory
デスクトップ

閲覧者 (Viewer)

作成者 (Author)

共有したインサイトの一元管理

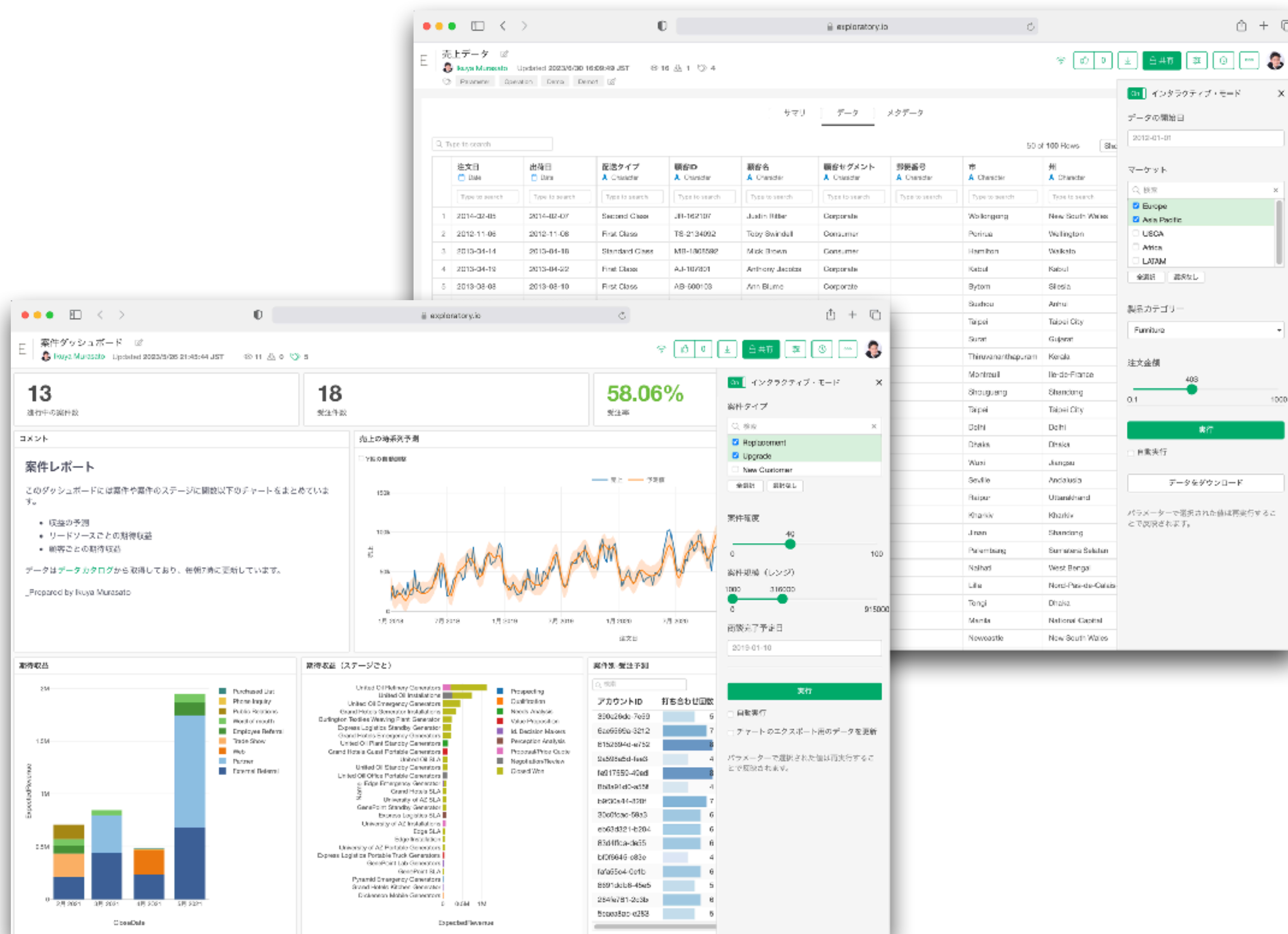
- ダッシュボード、ノート、チャート、データなどを一元管理
- 意味合い検索、キーワード検索、タグ検索
- PCからでもモバイルからでも、いつでも簡単にアクセス



パラメーター

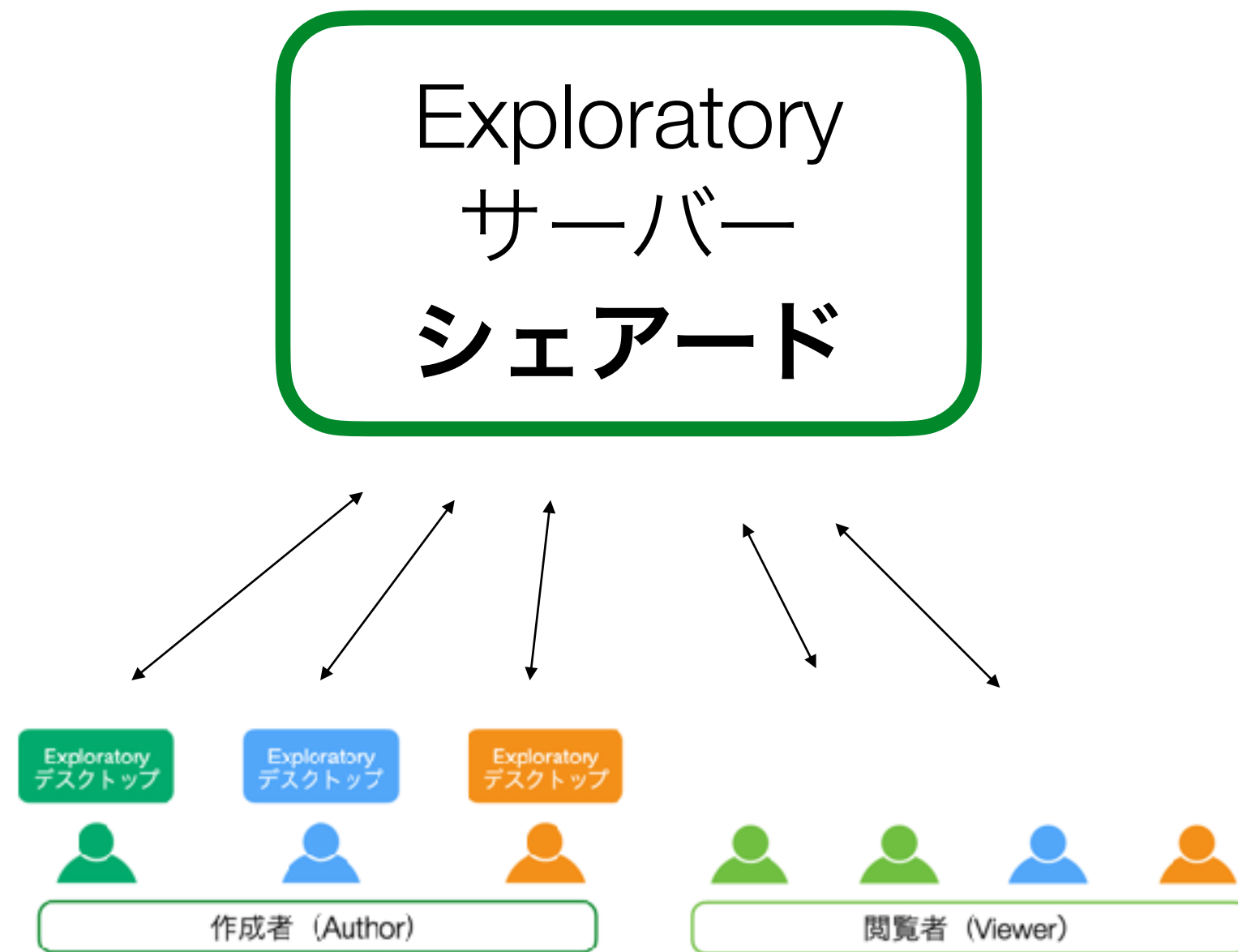
データやダッシュボードのフィルター、計算などをインタラクティブに変更するパラメーターでデータ関連のタスクをセルフサービス化

- **SQL**をパラメーター化して、**データの抽出**をセルフサービスで提供
- **ダッシュボード**をパラメーター化して、興味がある対象への**フィルタ**をセルフサービスで提供
- **計算**をパラメーター化して、**シミュレーション**をセルフサービスで提供



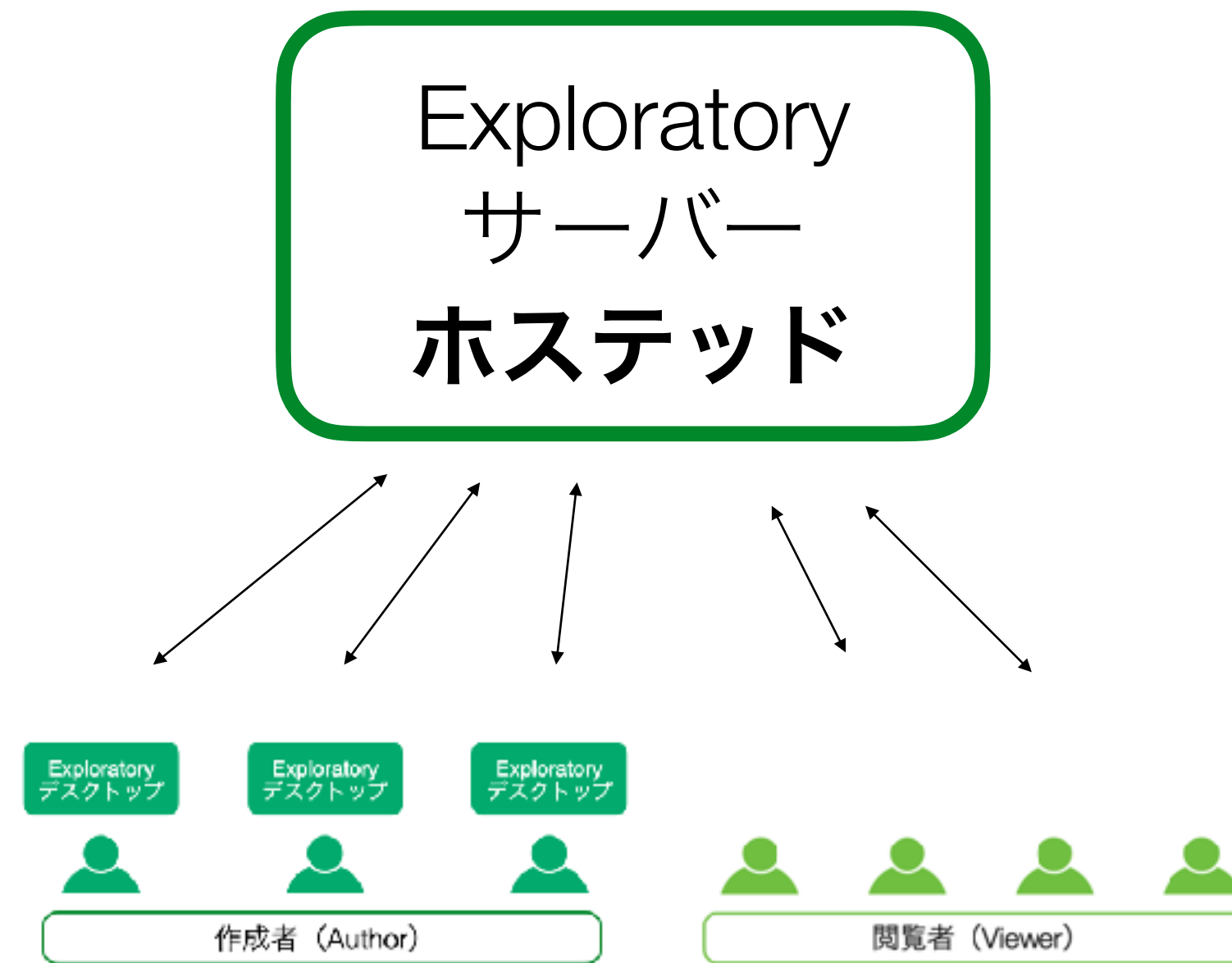
運用オプション

シェアード・クラウド



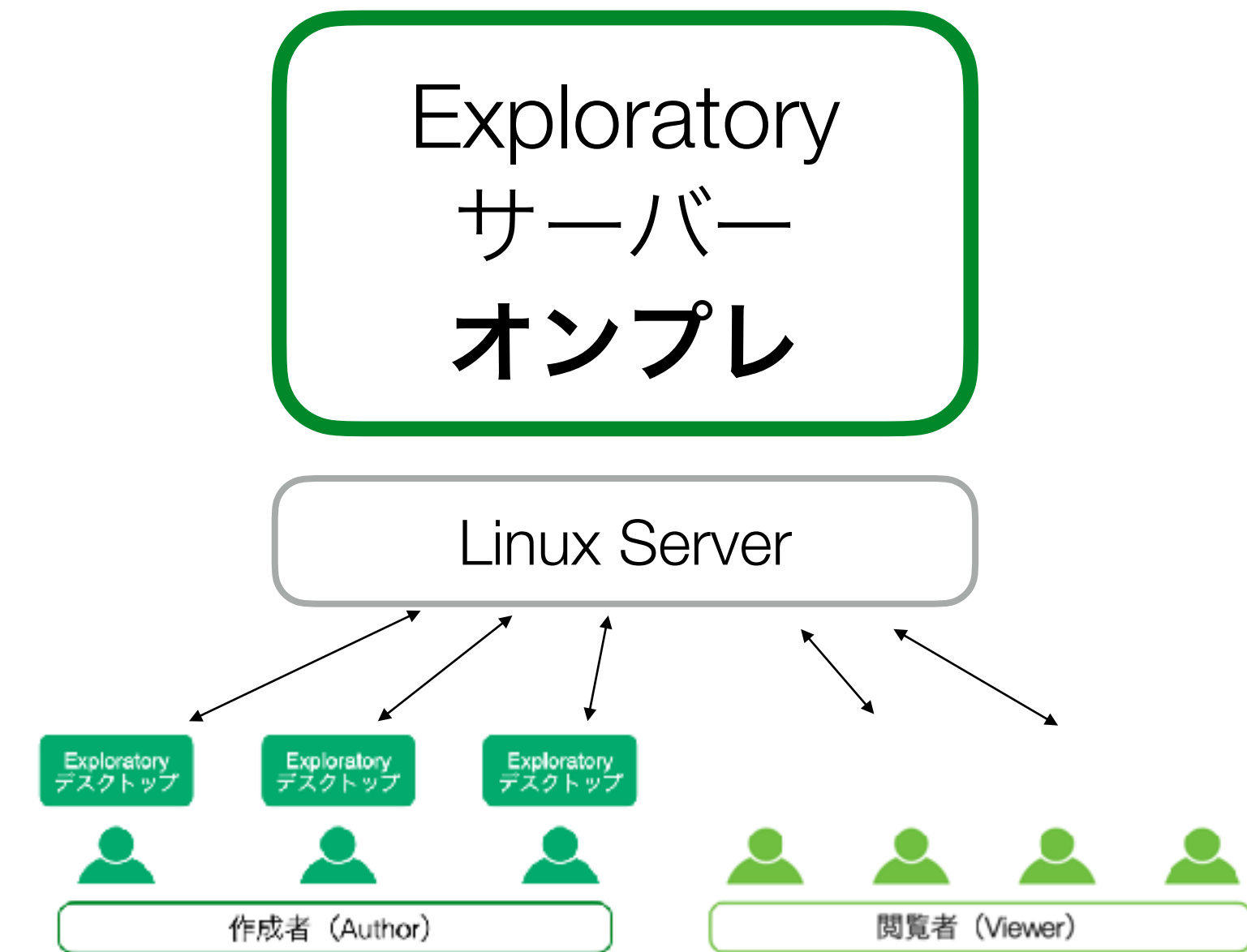
- 無料で開始
- 使用量によって柔軟なアップグレード

プライベート・クラウド



- プライベートの専用環境
- セットアップは不要
- メンテナンスはExploratory
- ユーザー・ライセンスは自社で管理

自社ネットワーク内



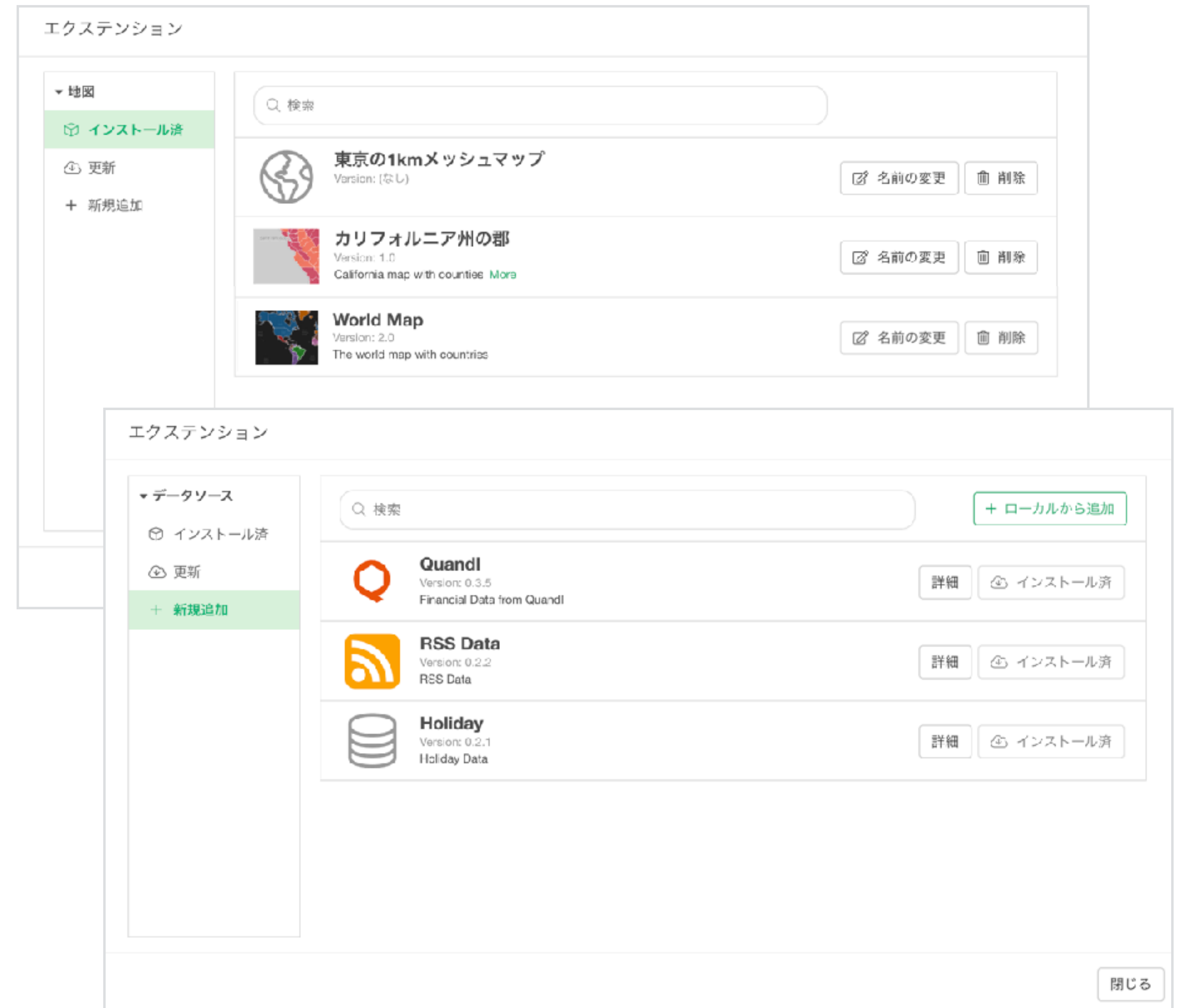
- 社内ファイアウォールに守られた環境
- 完全なコントロール
- サーバー環境・ユーザー・ライセンスは自社で管理

価格と機能

	パーソナル	ビジネス	ビジネス・プラス
価格（月額/年間契約の場合）	月額: \$49 年額: \$588	月額: \$99 年額: \$1188	月額: \$118 年額: \$1,416
チーム機能	なし	✓	✓
プロジェクトの共有	なし	✓	✓
AI機能	1ヶ月に100回まで利用可能	1ヶ月に200回まで利用可能	1ヶ月に400回まで利用可能
サーバー・ストレージ	5GB	10GB	100GB
サーバーでの 月間スケジュール可能行数	1000万行	5000万行	5億行
サポート	- メール - チャット	- メール（24時間以内） - ライブチャット - Web会議（Zoomなど）	- メール（24時間以内） - ライブチャット - Web会議（Zoomなど）
オンボーディング プログラム	なし	✓	✓

拡張フレームワーク

- ワンクリックで必要なRパッケージを1万以上のコレクションの中から簡単にインストール
- ユーザー独自のRスクリプト/関数の作成
- データ・ソース・エクステンション
- 分析の自動化・テンプレート化
- 機械学習モデル・エクステンション
- 地図・エクステンション
- ユーザー独自のカラーパレット



EXPLORATORY