

データーサイエンス

すいすい会 第36回

【東京ガス社ゲスト登壇】
事業会社におけるデータ利活用拡大
と分析組織強化のポイントをご紹介します



本日の話題

- 事業会社がデータサイエンス・サービス企業を上手く活用するテクニック
- GRIが提供しているデータサイエンス・サービス
- 地域に根差したビジネス展開する企業が行うべき
ロケーション・インテリジェンスのご紹介

事業会社がデータサイエンス専門企業を
上手く活用するテクニック

様々な内容に答えるための工夫

- TGグループのあらゆる課題に対して適切なプロジェクトを組成している

Tableauドクター
Pythonドクター

大まかな年間枠で人材を
抑える

R&DとPoC支援
(AI活用)

スポットで新規性のある
テーマのご依頼

データ利活用の研修

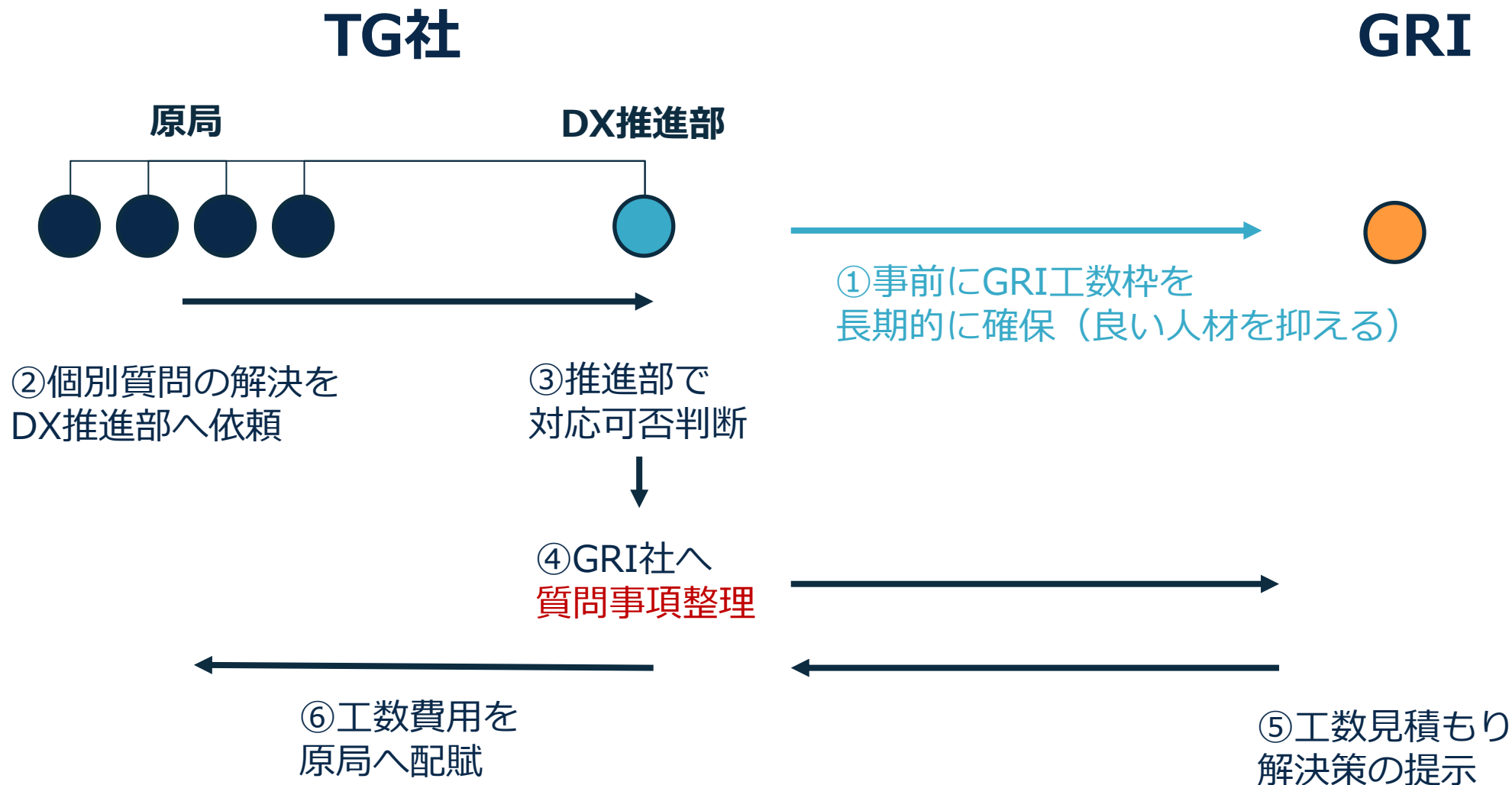
必要に応じてカスタム研修
(TG社のデータを用いて研
修内容を作る)

中期経営計画に関連するテーマが多い

<https://www.tokyo-gas.co.jp/about/plan/index.html>

Tableau / Python Doctorを効率的に行う仕掛け

- コストパフォーマンスを最大化するための準備



質問事項の整理

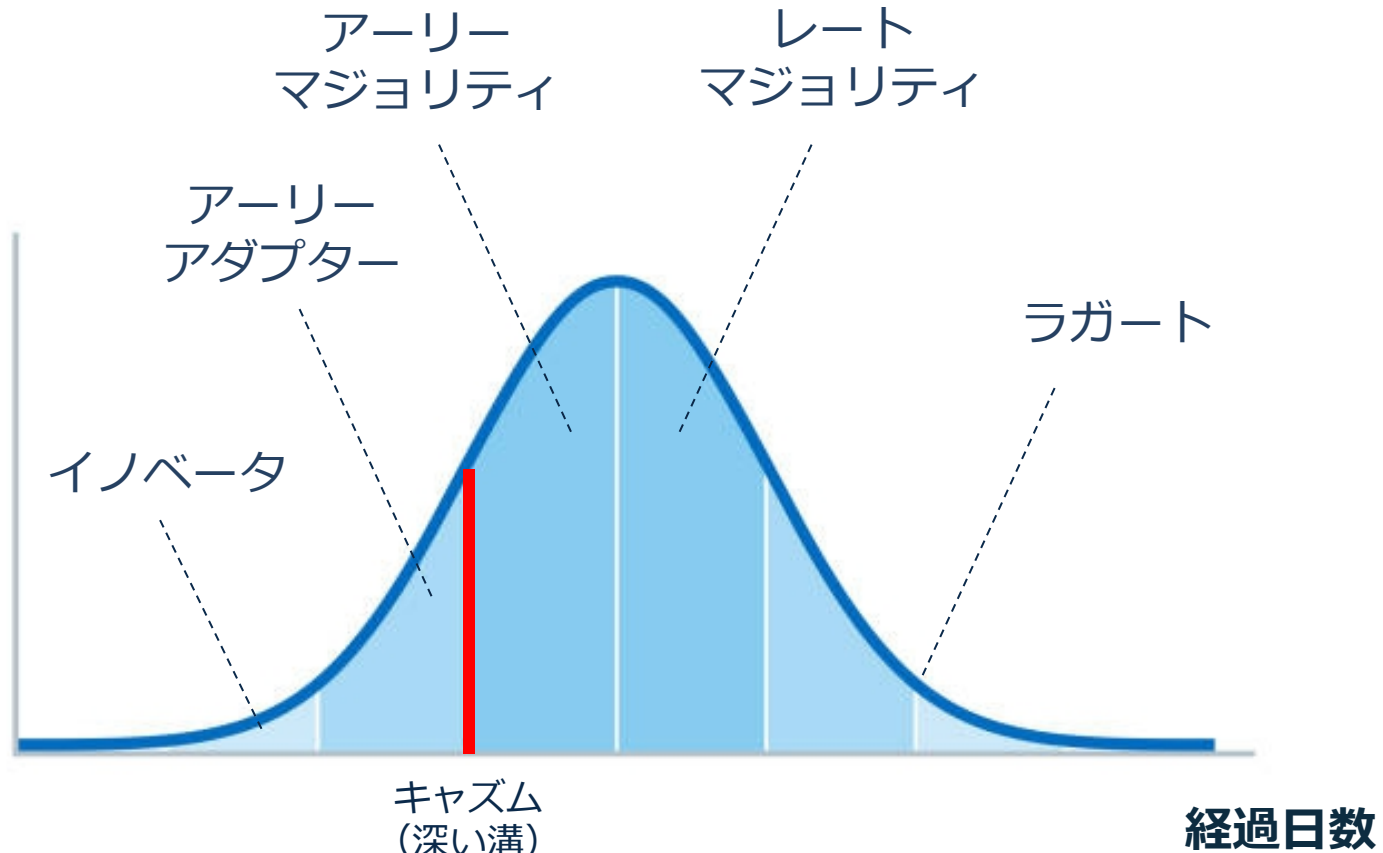
- 原局の組織の役割に照らし合わせ、解くべき分析課題か確認（期間と金額）
- 必要なデータを検討し、分析結果の仮説を立てる
- 自分たちで試してダメだったポイントをまとめる



- 分析結果の理解度が上がる
- 次回以降、類似の課題に対して、分析手順を理解できるためスキルが上がる

データ利活用の社内普及度合いに合わせて外部依頼のゴールを変える

デジタル人材の人数



パイロット部門
(キーパーソン、上長)



マネージャ陣の巻き込み
(マネージャ、育成サポート)



全社の巻き込み
(経営層、初心者)



浸透ステップの踏み方 (1/2)

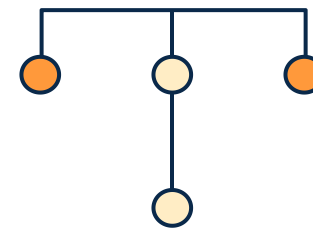
パイロット
部門



パイロット
部門上長



マネジメン
会議



- 上手くいく部門を選ぶ

- クイックウィン
- データリテラシーが高い
- 手を動かす時間の確保
- データの存在

- 部門上長の理解度

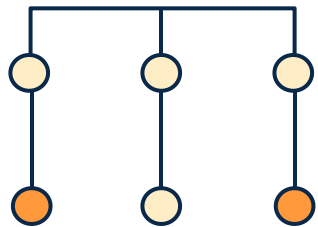
- 分析事例を説明可能なレベルまで理解
- 上長自らデモを見せられるようになる

- 良さが伝わる

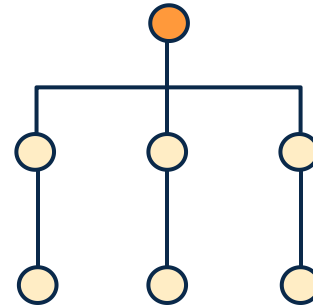
- マネジメント層の目に触れる場で共有
- 横の部門長が真似しなくなる

浸透ステップの踏み方 (2/2)

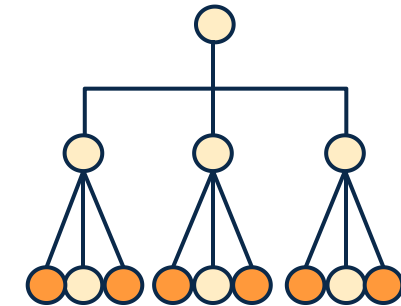
コア人材の拡充



経営戦略



全社浸透



- 横の部門で分析が進む

- 分析でつまづいた際の助け
- 分析事例の共有

- 経営の方向性に示す

- 分析結果と経営方針の一致
- 今後の方向性に従った分析

- 経営の方向性が伝わる

- 初心者講習
- 分析結果を使うのが当たり前
- KPIや評価と連動
- データ定義書

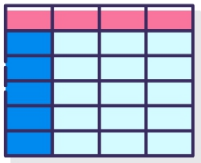
どのようなサポート人材を配置しているか？

- 基本知識
 - データ構造の把握テクニック
 - データ構造ごとの分析初手の理解
 - 分かりやすい情報設計と分析結果の共有
 - データ整形のテクニック
- マインドセット
 - 型にはまったソリューション提供ではなく、自分事として課題解決する
 - 受動的に分析作業を提供せず、データと会話をして能動的に提案をする
 - 新しさを追求し続ける

GRIデータサイエンス事業部が提供するサービス

あらゆるデータを初見で対応

職種x業種ごとに異なる分析テーマを
初見で成果に繋げられる理由の一つは
データ構造と分析手法の類似性



業務システム データ

- ID-POSシステム
- ERPシステム
- CRMシステム
- SCMシステム
- WMSシステム
- HRシステム
- ポイントシステム



時系列 ログデータ

- Webログ
- SNSデータ
- センサーログ
(IoT)
- TV視聴ログ
- 生体信号
- 金融経済データ



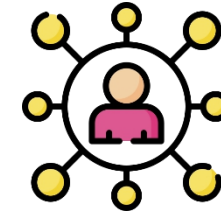
社会調査 市場調査

- アンケート
- 世論調査
- カルテ
- レビュー



非構造化データ

- 自然言語
- 音声
- 画像
- 動画



ネットワーク データ

- 交通データ
- SNSデータ
- 会話・文字データ
- 組織図
- コミュニティ
- 人間関係データ



オープンデータ

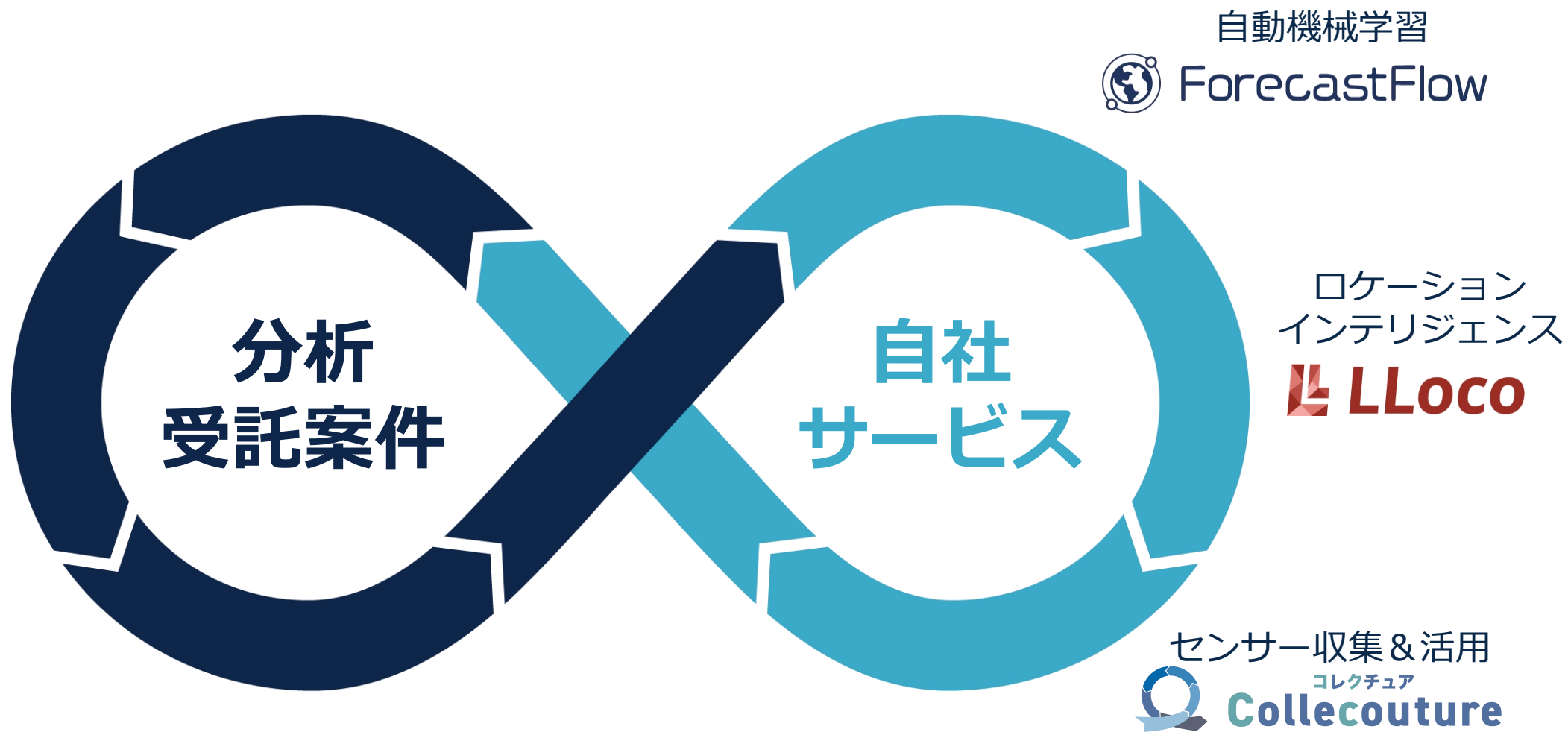
- 政府統計
- 天気
- 人口統計
- 地域
- カレンダー
- コーパス
- スポーツ

データサイエンス事業の双対循環ビジネスモデル

- データサイエンティストによる受託案件のノウハウを自社サービスに集約
- 自社サービスで得られた幅広い知見を受託案件の付加価値化



データサイエンス事業の双対循環ビジネスモデル



重点的なR&DとPoCテーマ

時系列予測

コンテクスチュアル
ターゲティング

地理空間
トークンエコノミー

画像処理

NLP & LLM

Marketing
Mix Model

データ交換

巡回セールスマン問題
最適化

大規模な
調査分析基盤

地域に根差したビジネス展開する企業が行うべき
ロケーション・インテリジェンスの紹介



ロケーション・インテリジェンス

地域に根差した企業が地理空間データをAIやBIと組み合わせ高度な意思決定に利用

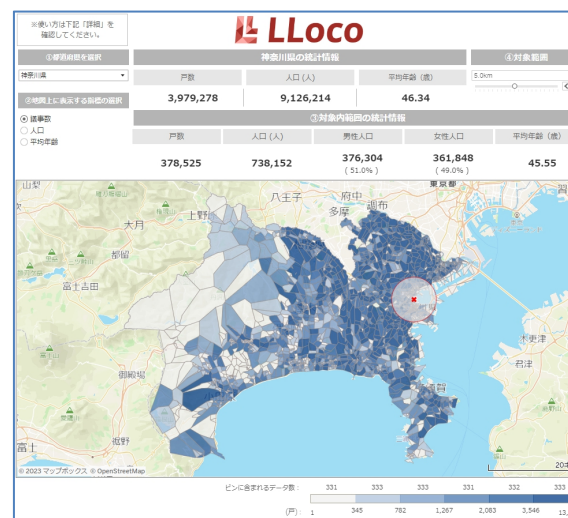
地域の
状況



データ
収集 &
加工



美しい表現で分析



意思
決定

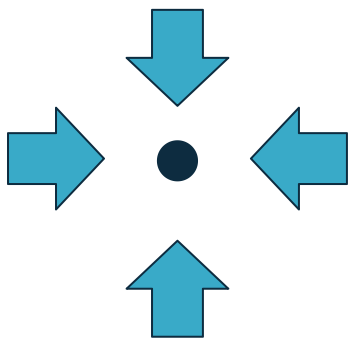
Artificial Intelligence
Business Intelligence

□ケーション・インテリジェンスのパターン

□ケーションデータは地域を示す軸に紐付く特徴

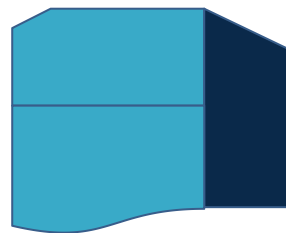
- 住所、緯度経度
- 郵便番号界、町丁目界
- 順序

地点メイン



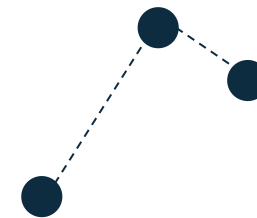
- 店舗集客
- 出店戦略
- 商圈分析

領域メイン



- エリア住民の特性推定
- エリア販促分析
- テリトリ分割

Pathメイン



- 訪問巡回分析
- 来店経路分析
- 回遊分析

ロケーション・インテリジェンスの具体例

■ 代理店営業の訪問戦略

代理店営業での週次レベル・月次レベルで訪問すべき代理店を戦略的に策定

■ 販促キャンペーンの地域反応分析

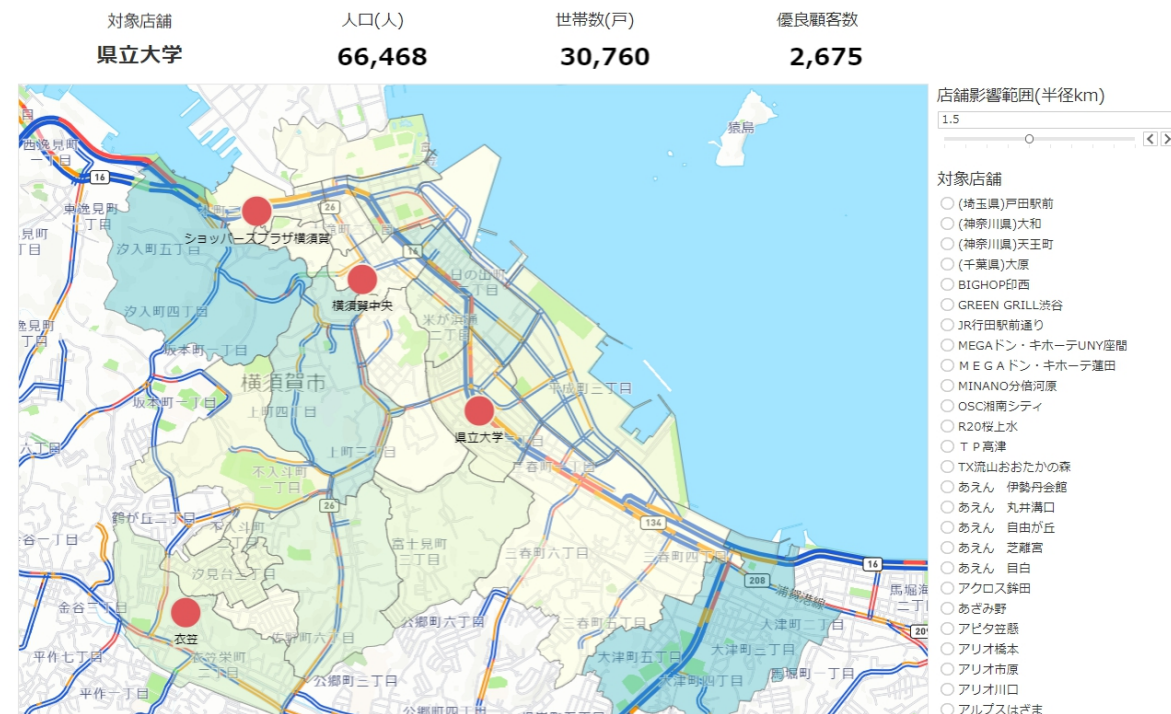
ファーストフード・チェーンでのTVCM、ポスティング、アプリを用いた重要地域の反応分析

■ コンビニ払いの未払い退会者の削減

コンビニ払いで未払いによる強制退会を少なくするため、機械学習で退会予測をする際、地域特性の特徴量を入れ、退会抑止策の立案に貢献

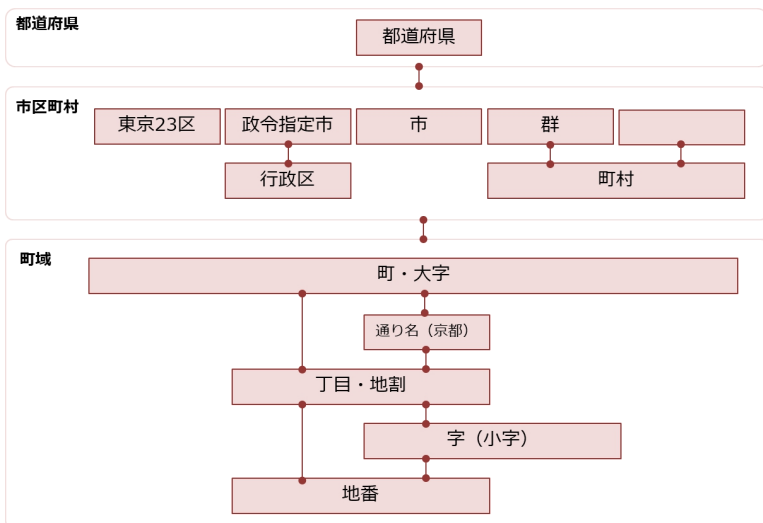
■ イベント会場の来場者分析

イベント会場の開催期間に来訪した人の地域、地域属性、行動属性を分析し、スポンサーにレポート



日本の住所データの扱いの困難さ（例外だらけの領域表示）

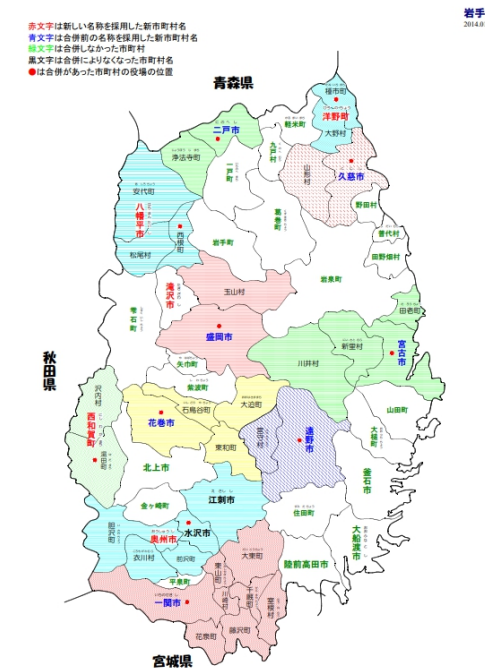
例外だらけの領域表示



表記ゆれ

阿佐谷
 阿佐ヶ谷 小さい「ヶ」をつける
 阿佐ヶ谷 大きい「ヶ」をつける

繰り返される 市区町村合併



ロケーション・インテリジェンスの実践時の典型的な課題

- 企業の持つ住所データの名寄せに工数
- 国勢調査データ（e-Stat）の収集が手間
- 地域ポリゴンのレンダリングが重い
- GISツールが高額で社内共有できない



ロケーションインテリジェンス用の
共通コンポーネントをパッケージ化



LLoco (エルロコ) の解決策

e-Stat

政府統計の総合窓口

- 地域境界ポリゴン (町丁目)
- 地域統計データ

JP 日本郵便
POST

- 郵便番号データ



LLoco

- 地域境界ポリゴン
 - 郵便番号界 / 市区町村界
 - 詳細版 / 高速版
- 地域統計データ (各集約)

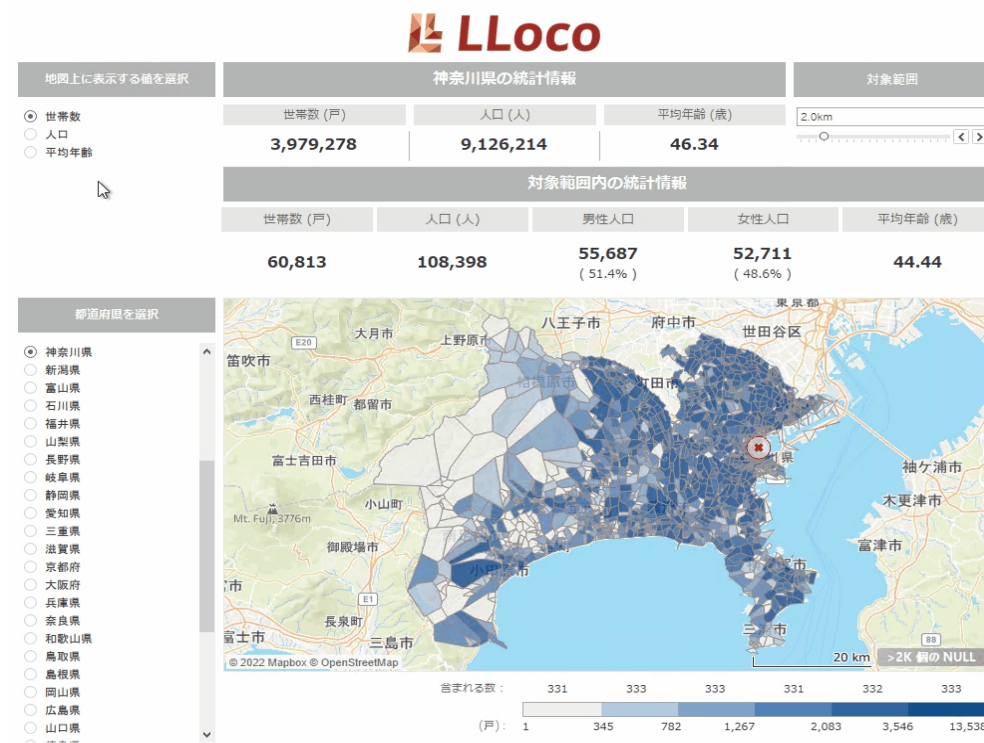
お使いのBIツールや
Pythonと一緒に
利用

- 自然言語処理と人力による町丁目と郵便番号のマッチング
- トポロジカル処理でポリゴン集約 & 高速レンダリング用のポリゴン生成 (e-Statポリゴンと比較して10~10,000倍高速)

LLocoで標準提供されるデータと価格

LLoco提供データ（標準パッケージ）

データの種類	データ分類（粒度）	データソース
地域境界線データ（ポリゴン）	郵便番号	eStatの小地域shapefile
	市区町村	eStatの市区町村shapefile
地域粒度の変換マスタ	郵便番号 → 郵便番号	日本郵便の郵便番号
	町域名 → 郵便番号	国勢調査
	町丁目 → 郵便番号	国勢調査
	郵便番号 → 市区町村	国勢調査
地域人口統計情報	基礎情報	国勢調査
	性別×年齢	国勢調査
	世帯類型	国勢調査
	性別×産業種類別	国勢調査
	性別×職業大分類	国勢調査

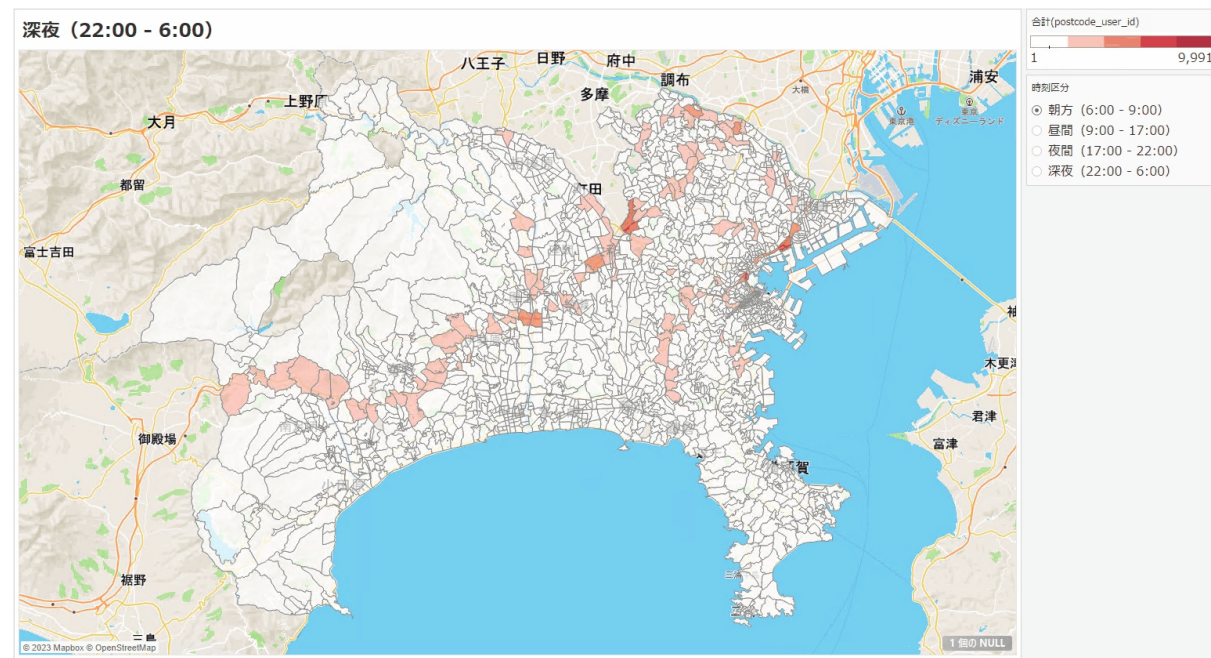


1年間、自由にダウンロードできます。同一企業内で無制限に利用可能です。
 (但し、許可なく他社へ販売すること不可)

- 日本全国： 50万円/年
- 各県： 10万円/年
- 神奈川県： 無料

LLocoのオプションデータ

- 各企業の保有するデータをオプションとして出品が可能です
- オプションの例（詳細はお問合せください）
 - ① 気象データ、天気予報データ
 - ② 地域アンケート
 - ③ GPS人流データ
 - ④ 折込チラシデータ
 - ⑤ 新聞購読データ
 - ⑥ 結線テレビを用いたTV視聴データ



まとめと今後の予定

本日のポイント

- データサイエンス・サービス会社に頼むポイント
 - 質問事項の整理
 - 事前の工数確保と原局への費用配賦
- GRIデータサイエンス事業部では双循環モデルでサービス提供
 - データサイエンス受託案件
 - 自社プロダクト
- LLocoを使うとロケーション・インテリジェンスを手軽に始められる

人材募集

- データサイエンティストの発想による面白い未来を創りたい人！
- 大企業の歯車にならず、自分の個性と能力でワクワクできる仕事を作り出したい人！
- 古臭いことに囚われず、新しくて心動かされる社会変革にチャレンジしたい人！
- 募集職種例
 - フルスタック・データサイエンティスト
 - AI開発マネジャー
 - BIスペシャリスト
 - データエンジニア
- お問い合わせ先
info@gri.jp

ハイブリッド開催

GRI | MATILLION

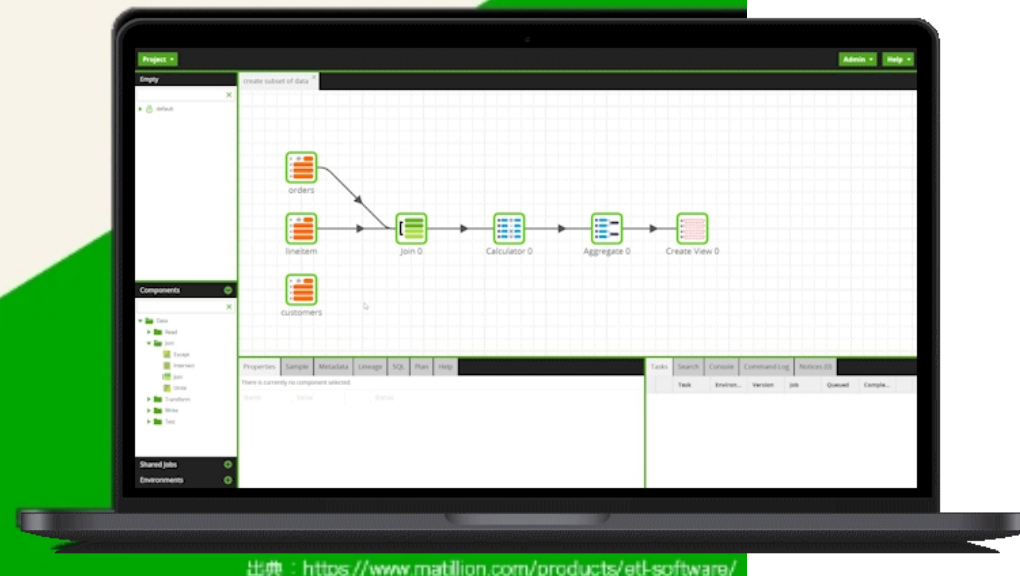
あの有名企業がBigQueryやSnowflakeを 使いこなす秘訣とは？

クラウドDWH時代の大規模パイプライン自動化ツール
Matillionユーザ会発足記念ミートアップ

2023/7/26 (水)

OPEN 18:15 START 18:30

📍 東京都港区



出典: <https://www.matillion.com/products/etl-software/>

こんな方におすすめ

- Matillionユーザや導入検討の方
- 大規模データ基盤構築、保守運用管理に課題感をお持ちの方（社内SEをデータエンジニアに育成）
- Matillionユーザ企業のデータ基盤構築成功事例にご興味のある方
- データエンジニア、データサイエンティスト、ビジネスアナリストの方

お申込み: <https://techplay.jp/event/905463>

Q&A

- ForecastFlowのSlackコミュニティとYoutubeのご質問にお答えします