

データ分析・統計解析
ソフトウェアの選び方

正しいソフトウェア選択がデータ分析成功のカギ

数多くのデータ分析ソフトウェアが存在する今、どのようなソフトを選択すべきでしょうか。

当ホワイトペーパーでは、数多くのソフトウェアの導入実績のあるスマート・アナリティクス社の視点よりデータ分析ソフトウェアの分類やデータ分析ソフトウェア選定の基準などを紹介します。

これからデータ分析を行うとお考えの皆様の一助になれば幸いです。

スマート・アナリティクス株式会社では、ソフトウェアの導入・購入のサポートから、データ分析の受託サービス、伴走型分析サポート、トレーニングコースの提供まで、データ分析をトータルにサポートいたします。

ビッグデータ・IoT・AIの世界へ

現代はまさにIoTをはじめとするデータが大量に発生し、ビッグデータの時代となりました。また、併せてAI（人工知能）の発達によりデータ分析が身近な時代となりつつあります。

ソーシャル



購買データ



IoT



データ分析対象とビジネスにおけるデータ活用例



「ヒトの分析」

新規顧客の獲得

併売・クロスセル

顧客の離反防止



「モノの分析」

歩留りの改善

故障の予知

予知保全



「カネの分析」

初期与信・途上与信

不正検知

コンプライアンス

そのような中、データ分析の活用範囲は多岐にわたります。
ビジネスにおいては、上記のように「ヒト」「モノ」「カネ」の各領域において利用が拡大しています。

データ分析で可能なこととは？

一般的にデータ分析でできることは、

「予測」 ・ **「パターン発見」** ・ **「分類」**

の3つに類型化できます。あらゆる活用例も分析の観点では上記の3つに当てはまります。



内容

具体例

	内容	具体例
予測	ある値を予測すること	優良顧客の購入確率の予測 将来の売上予測
パターン発見	何かと何かを一緒に購入されるなどの なんからのパターンを発見すること	解約パターンの特定 購入者パターンの特定
分類	データを何かの軸で分けること いわゆるセグメンテーションを行うこと	優良顧客のセグメンテーション 顧客クラスタリングの作成

データ分析と何をする事なのか？

そもそもデータ分析とは何を行うことなのでしょう？

ビジネス・研究に限らず、データ分析は、知りたい物事を判明・究明するためのアプローチ方法といえるでしょう。物事を究明するアプローチは必ずしもデータに基づかないものもありますが、データ分析は、「データ」を利用して知りたいことを究明するアプローチといえます。



データ分析の手順と注意点

データ活用・分析のためのステップと課題

データ準備・加工

データ活用の目的を決定
必要なデータを定義
データを準備・加工

よくある課題

必要な情報が
すぐに手に入らない
データが分散されていてすぐデータを
入手できない

探索・分析

必要な情報を探索
知りたい事象を分析

よくある課題

難しすぎる分析ツール
前任者が作ったコードや分析プログラ
ムは解読不能

共有・レポートニング

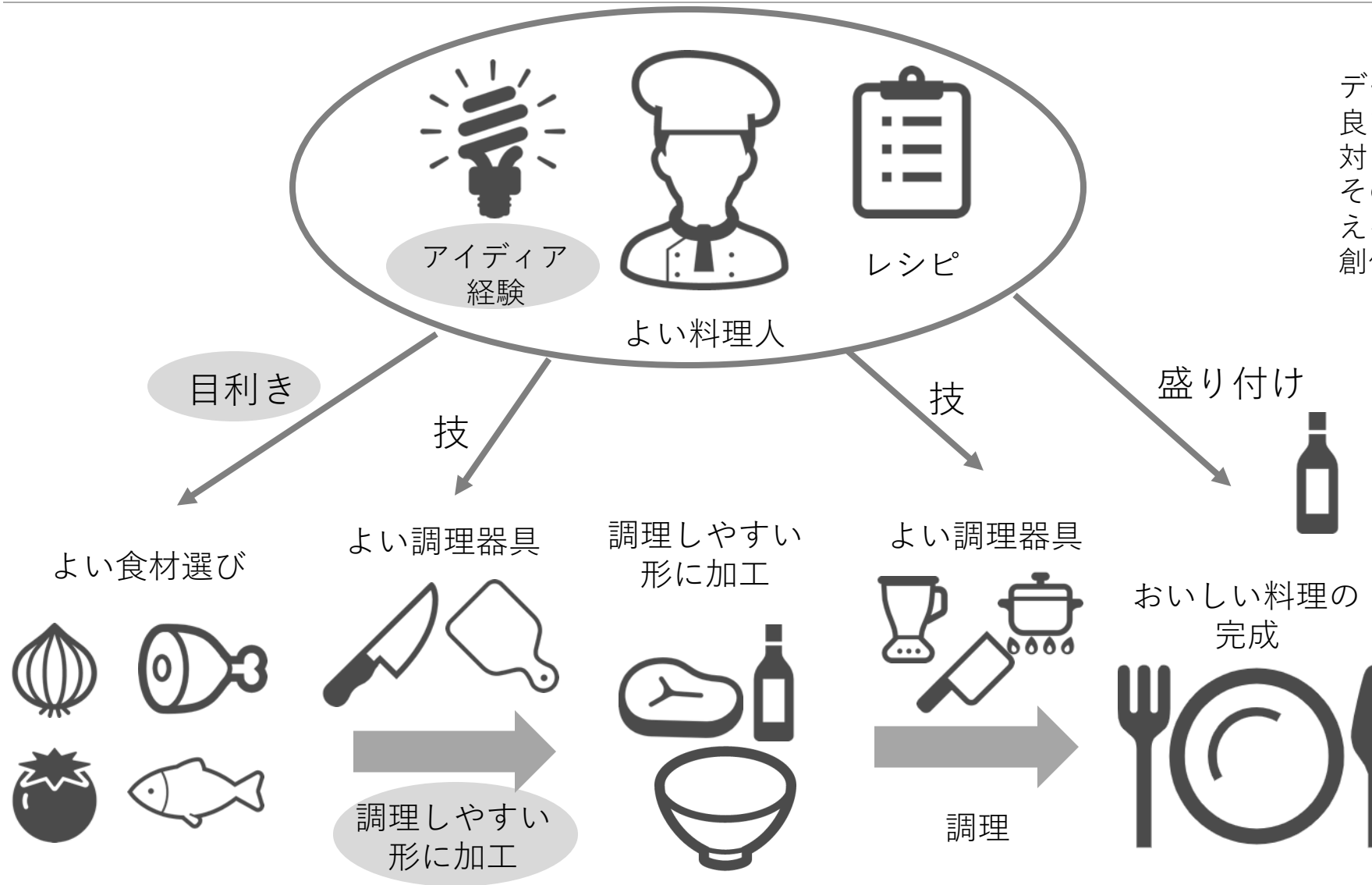
得られた知見をすばやく共有
知見を元にレポートニング

よくある課題

一部の専門家のみで共有
レポート作成に手間がかかる
バージョン管理が煩雑

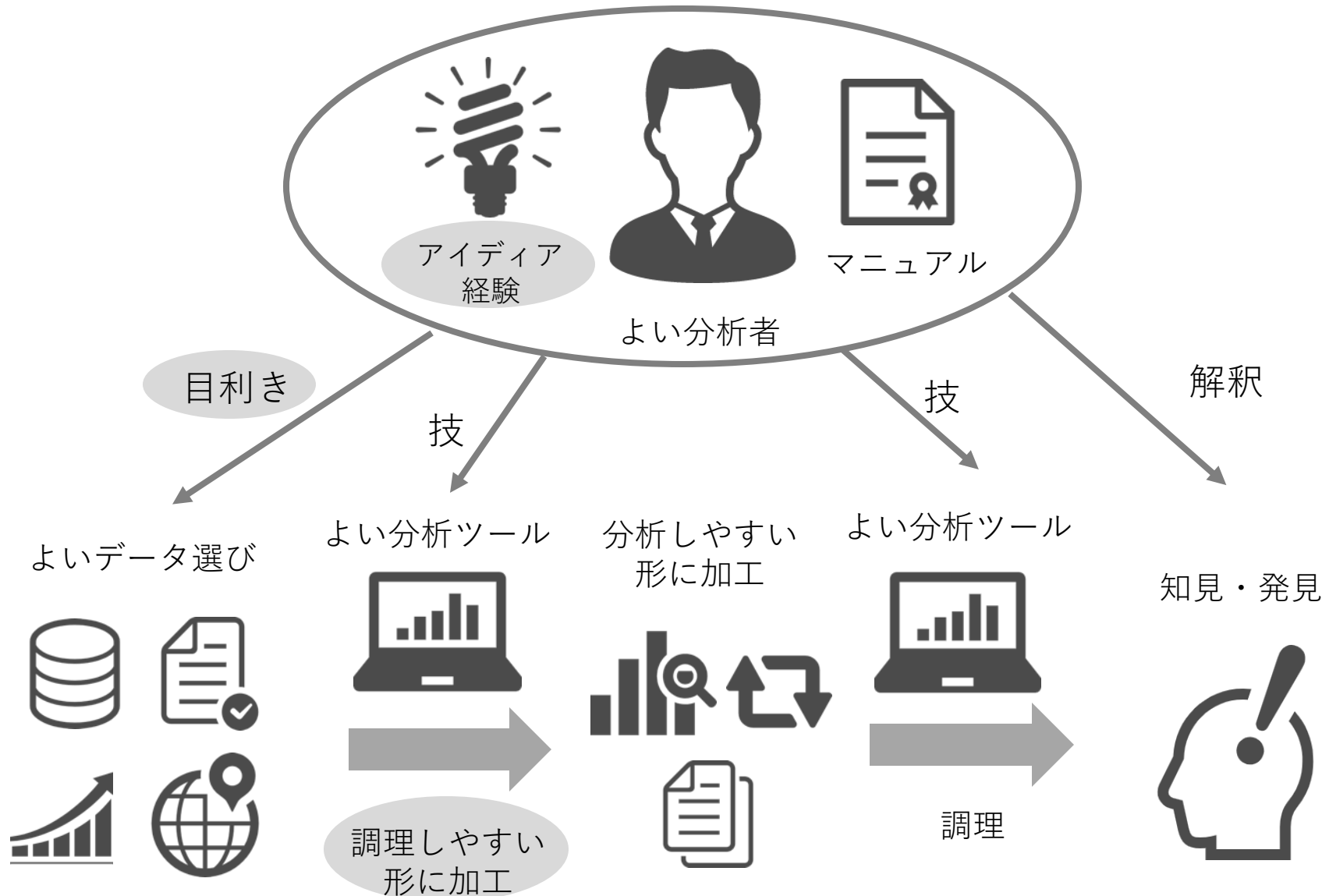
データ分析を行う際のプロセスは、データの取得から「データ準備・データ加工」に始まり、データの「探索、分析」そしてその結果を利用するための「共有・レポートニング」が一般的です。一方その実現には多くの課題も存在します。

データ分析は料理と似ている



データ分析は、よく料理に例えられます。良い調理人は、お客様からのオーダーに対してまず完成形の料理をイメージし、その料理に必要な素材を選び、下ごしらえをして、調理を行い、最終系の料理を創作します。

データ分析は料理と似ている

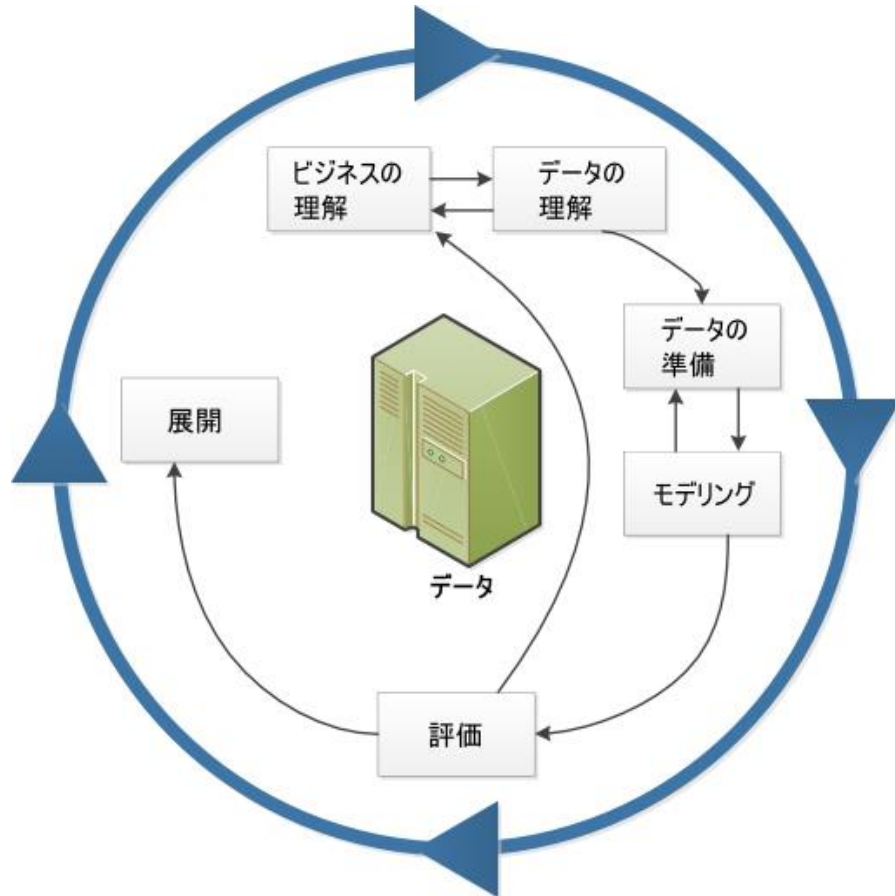


データ分析も料理と同じです。
最終的に何を知りたいのかをまず設定した上で、その究明したいもの明らかにするために必要なデータを選定します。

その上で、分析に必要なデータ加工を行い、最終的にアウトプットを行います。

データ分析メソドロジー「CRISP-DM」

業種横断、長年の経験に基づくデータ分析のメソドロジー



CRISP-DM

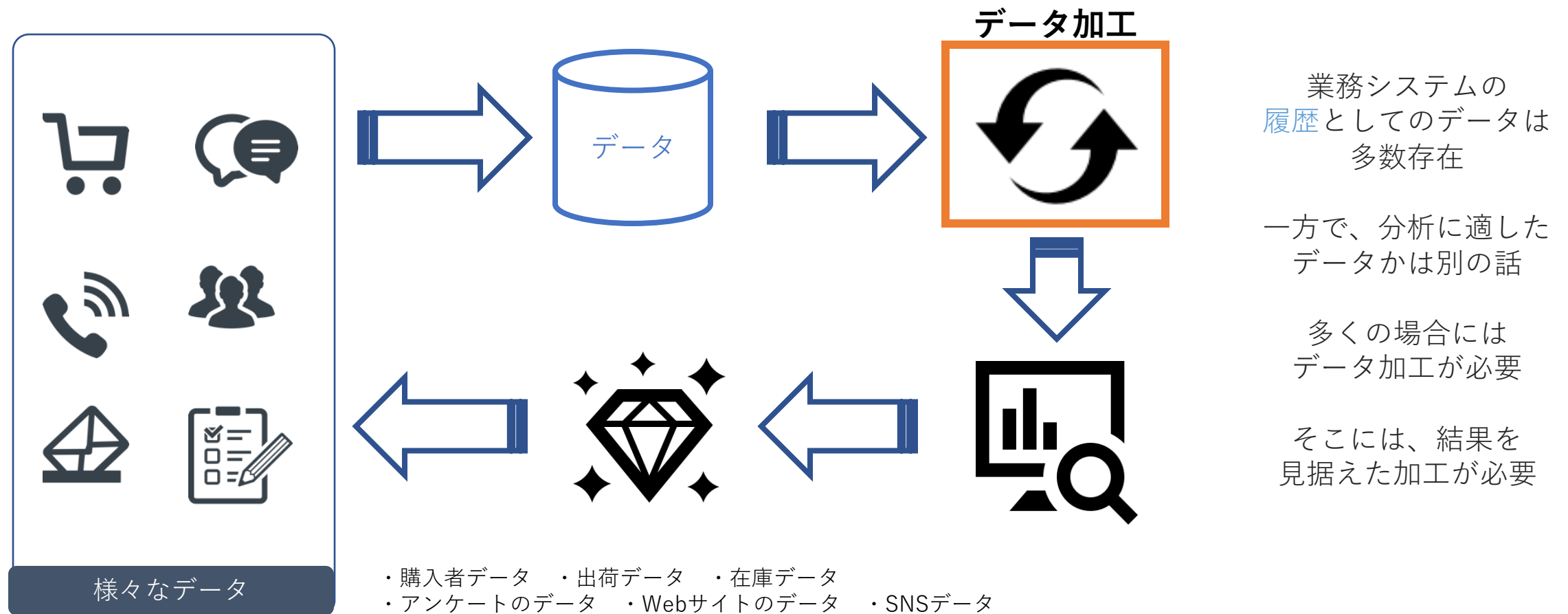
(Cross-Industry Standard Process for Data Mining)

CRISP-DMは、IBM社やダイムラー社らによって提案されたデータ分析プロジェクトの標準フレームワークです。ビジネスにおけるデータ活用において、分析手法やアルゴリズムの選定のみならず業務課題の設定からデータの理解、データの準備、モデリングさらには作成したモデルを評価し、ビジネスプロセスに展開することまでをカバーしています。

データ分析の作業は、
「ビジネスの理解」「データの理解」「データの準備」に
約80%の工数が取られるといわれています。

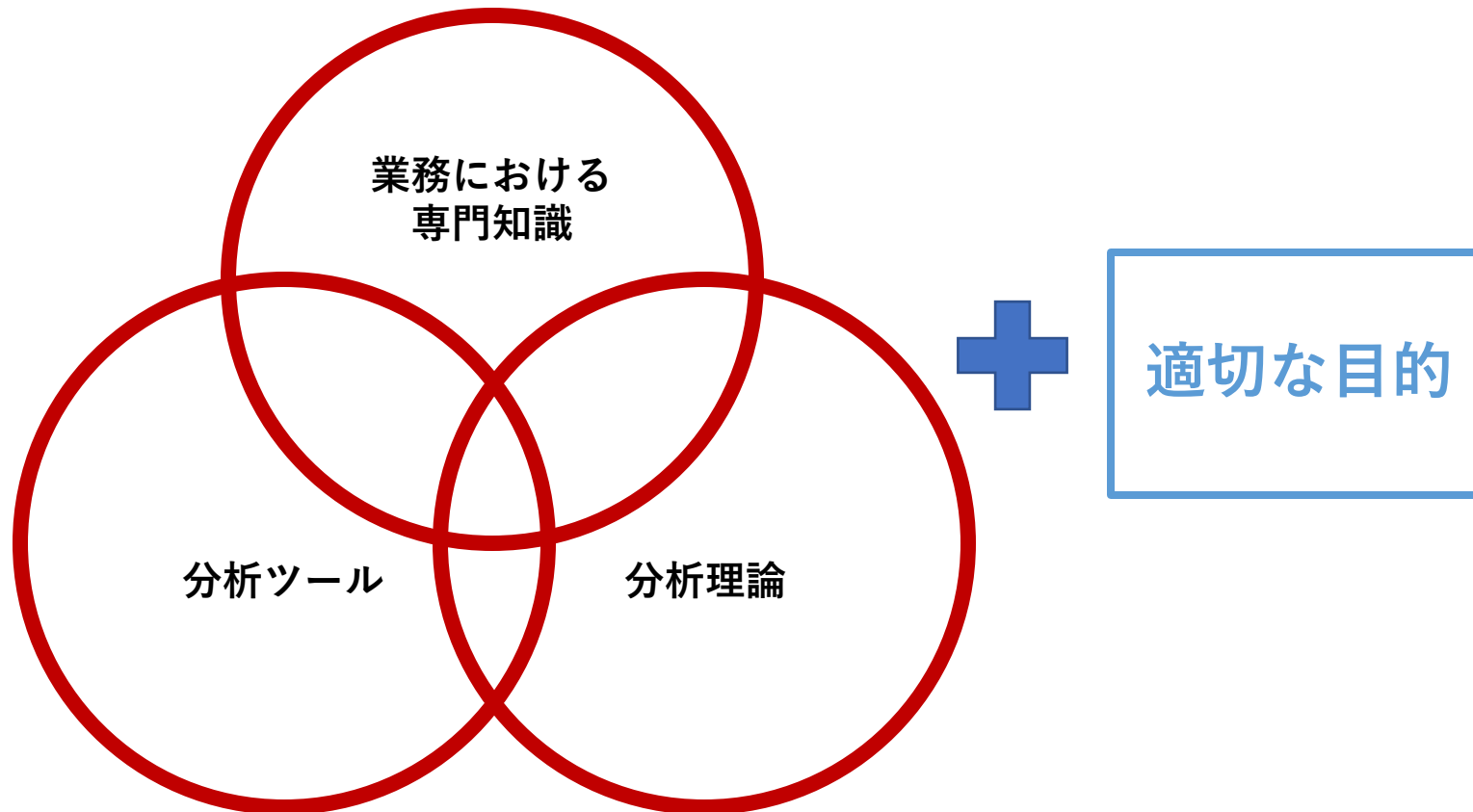
分析に利用するデータについて

データ分析の際に抑えておかなければならないことの一つに「データ加工」があります。特にビジネスのデータ分析において、知りたい事象について自動的もしくは手動で取得するデータでは、データ加工がなければ分析できない場合があります。分析ツール選定においてもデータ加工の機能を考える必要があります。



データ分析を成功させるために必要な要素

「ツール」「理論」「専門知識」の3つのノウハウがデータ分析プロジェクトの成功には不可欠です。



データ分析を成功させるためには、適切な「分析ツール」の選択、適切な「分析の理論」の選択が当然必要です。

さらには現場の「専門知識」が不可欠です。なぜならば、「分析ツールはあくまでも道具です。最後の解釈は分析ツールではなく、ビジネスユーザーや研究者が行う必要があります。

当然、それを支える目的の設定、組織設定も必須の要素といえるでしょう。

分析ツールの種類と代表的なソフトウェア



分析と言っても目的によって使い分けるべき

分析ツール群ごとの特徴と機能

ツール	得意な機能	データ収集 データ加工	可視化	分析 (統計)	共有 展開
BIツール	基礎集計と可視化・定点観測 (ダッシュボード)	○	◎	×	○
セルフBI	多くのデータを自由に結合して 簡単な分析を行う	○	◎	△	○
統計解析	統計的な処理	○	△	◎	×
データマイニング	仮説発見 仮説検定・機械学習	◎	×	◎	△
Auto ML (AIツール)	機械学習を自動的に実行	△	△	○	△
最適化ツール	最適資源配分を実行	×	×	◎ (最適化問題)	△

分析ツール選定の基準は？

分析ツールの選定においては、分析ツールの価格だけで決定するべきではありません。

まず、前項のように「分析の目的」、誰が分析を行うのか、導入までのコスト（習得コストや学習時間）を含めて検討する必要があります。

特に目的の決定は分析ツール選定において非常に重要な要素となります。

例えば、BIツールは集計や可視化が得意。でも統計解析やデータマイニングツールは苦手。統計解析やデータマイニングツールは、当然高度なデータ分析や統計は得意。でも可視化はとても苦手のように、ツールの向き、不向きがあります。

導入する際には、分析ツールの習得のしやすさなども考慮しましょう。例えばプログラミングツールはオープンソースで無料で利用できるものもありますが、プログラミング経験がある方でなければ操作を習得するには時間がかかります。その習得までの日数も考慮して導入のコストを算出してみましょう。

また、Auto ML（自動AI）ツール群は非常に人気の高い製品群です。

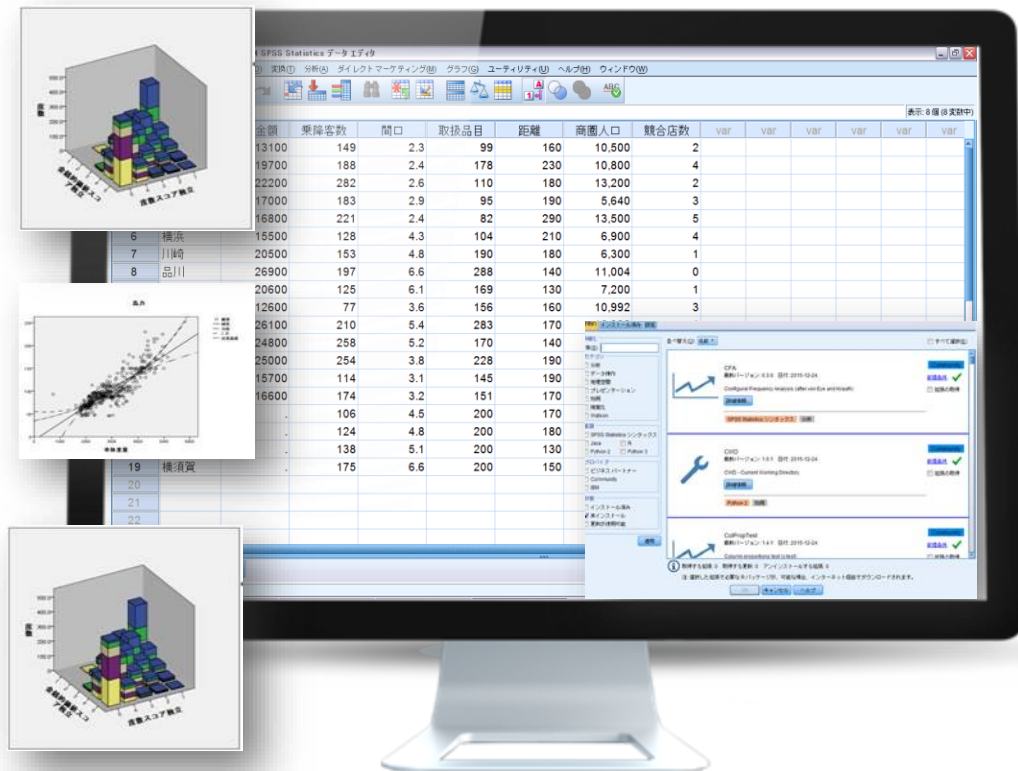
データを投入すれば結果を出してくれるというキャッチフレーズの製品が多いのですが、分析の世界では、「Gavage In Gavage Out（ゴミからはゴミしかでない）」という格言がある通り、データ次第でよい結果が出ることも出ないこともあります。

データ加工などをあまり得意としていないツールも多いことからその点も含めて検討をすすめましょう。

おすすめのデータ分析ソフトウェア

<おすすめ製品-1> IBM SPSS Statistics ファミリー

世界中28万、国内3万ユーザーが利用する統計解析のスタンダードツール



- ✓ 分類：統計解析ソフトウェア
- ✓ アンケートデータや調査研究に最適
- ✓ 1968年から続く伝統の統計解析ソフトウェア
- ✓ ユーザフレンドリーな操作性
- ✓ すぐに利用できる分析手法を30種以上標準搭載
- ✓ データ入力から出力までカバー

SPSS Statisticsの特徴：統計分析に使える機能が満載

社会調査から生まれたアンケート分析・研究調査に強いデータ分析ソフトウェア

多くのデータ・フォーマットに対応



豊富なデータ加工・準備機能

- 強力な変数変換機能
- データ結合（変数結合・ケース結合）
- 数量データのカテゴリ化
- 文字列データも連続数に変換

すぐに使える分析手法を多く搭載

記述統計・多重回答の集計・相関分析、因子分析
クロス集計（カイ2乗検定）、t検定、ANOVA、反復測定、ノンパラメトリック検定、ロジスティック回帰など各種研究用の解析手法を多く搭載

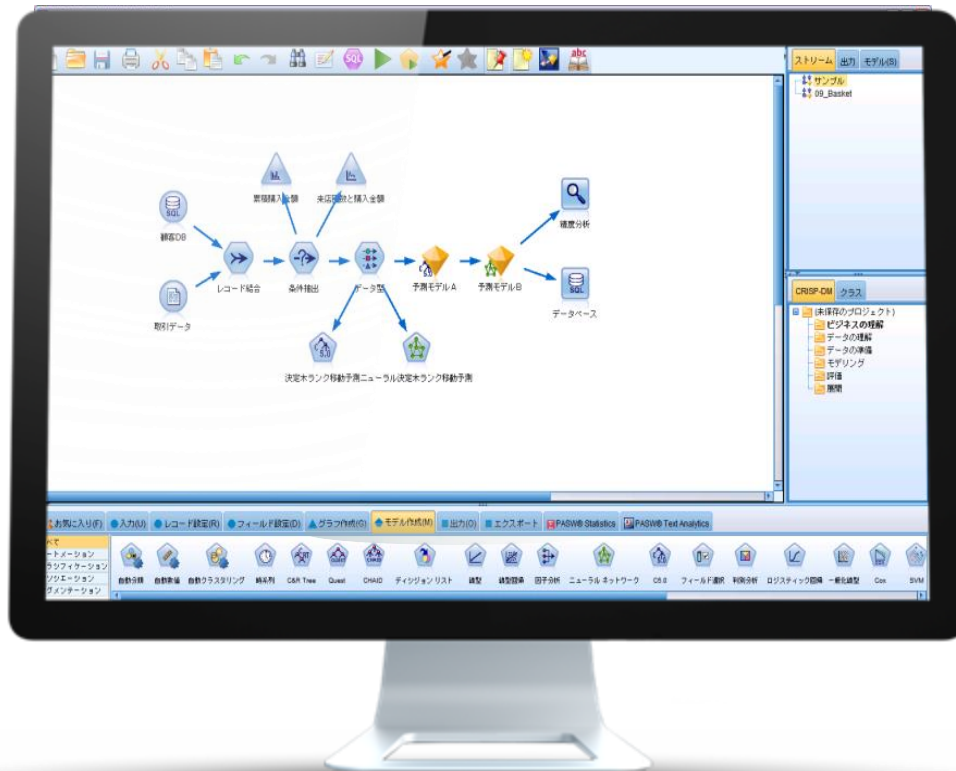
グラフ作成からOffice連携まで

豊富なグラフ機能を搭載。

Ver.25からOffice連携機能により
レポート作成がより効率的に

<おすすめ製品-2> IBM SPSS Modeler

人工知能モデルも搭載のデータ分析ワークベンチ



- ✓ データ入力からデータ加工、モデリング、展開までをワンプラットフォームで
- ✓ 人工知能を含め最新アルゴリズムを多数搭載
- ✓ 既存システムとの連携
- ✓ 国内シェアNo1のデータ分析プラットフォーム*

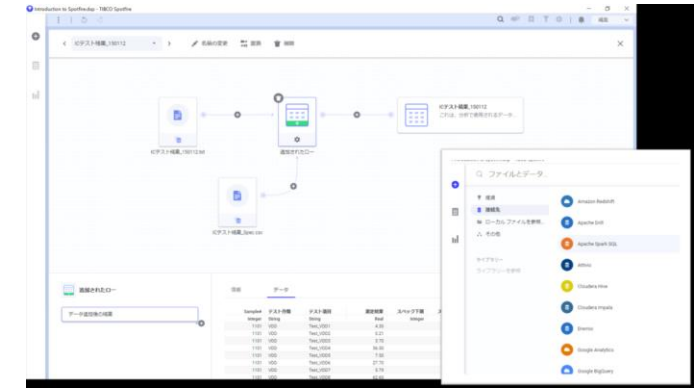
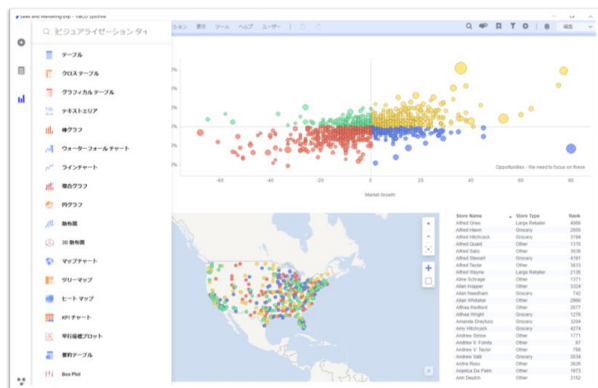
*出典：ITR「ITR Market View：DBMS / BI市場2016」

<おすすめ製品-3>TIBCO Spotfire

ビジュアルデータ分析&データ分析機能を搭載する「TIBCO Spotfire®」



- ✓ 分類：可視化ソフトウェア
- ✓ 各種DBやCSV、Excelなど社内外のデータにアクセス
- ✓ 可視化したい軸で自由にデータを操作しビジュアル分析を実現。
- ✓ ちょうどいい、現場で使えるデータ分析ソフトウェア。



データ分析のことならスマート・アナリティクスまで

スマート・アナリティクス株式会社は、あらゆる方にデータ活用のパワーを提供します。

分析ソフトウェア

IBM SPSS製品を始め、データ分析の各種ソフトウェアを提供中



分析サービス 教育・研修

CRISP-DMメソドロジーを利用し、かつ産学連携による独自プログラムの提供。また製品操作から理論まで各種教育・研修を提供

- ・ CRISP-DMコンサルティング
- ・ データ分析受託サービス
- ・ 製品アドバイザー
- ・ トレーニング・教育研修



産学ブリッジ

アカデミアの輝きのある知見（エンジン）をビジネスに活用するためのブリッジ事業

- ・ 最適化理論プログラムHOPE
- ・ マーケティング協働プログラム
- ・ 実証実験
- ・ 大学・ゼミとの協働活動



Thanks!

お問い合わせ先

スマート・アナリティクス株式会社

〒107-0062 東京都港区南青山3-8-2 サンプリッジ青山3F 302

TEL : 03-5786-4580 / FAX 03-5784-4577

E-mail : info@smart-analytics.jp

URL : <https://smart-analytics.jp/>