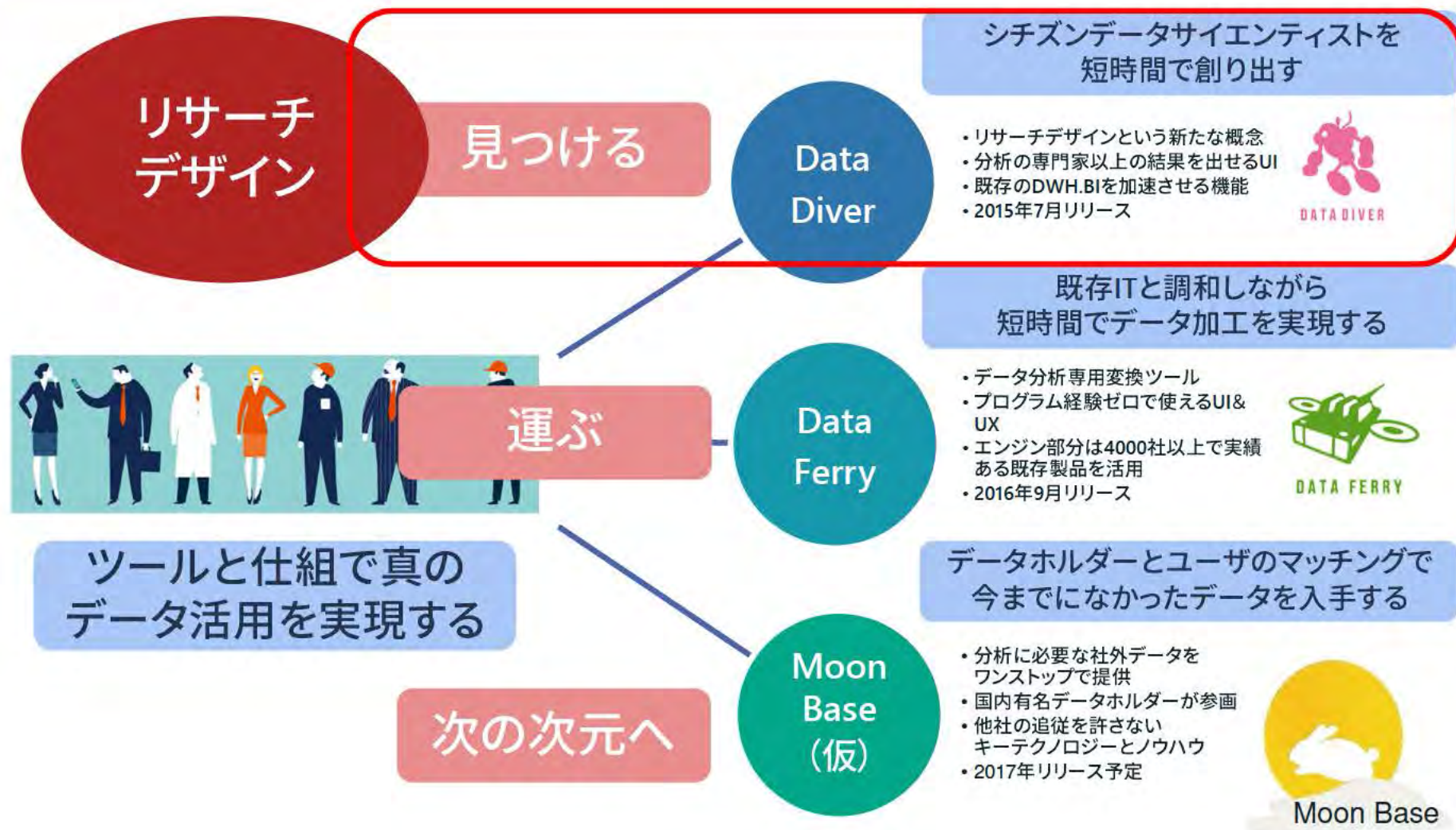




Data Diver 製品概要資料

株式会社データビークル
2017年9月

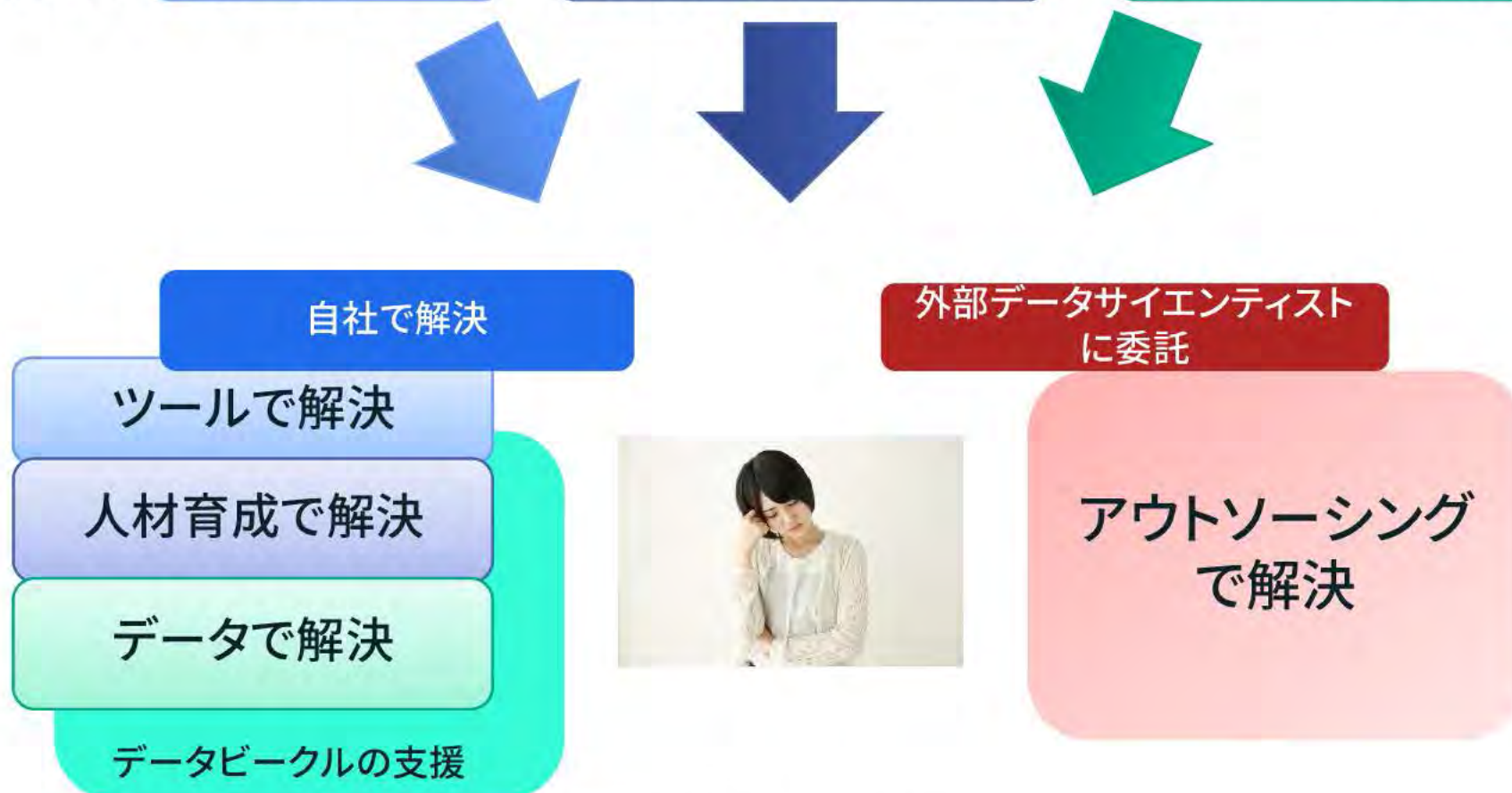
データビークルが提供する三つの武器



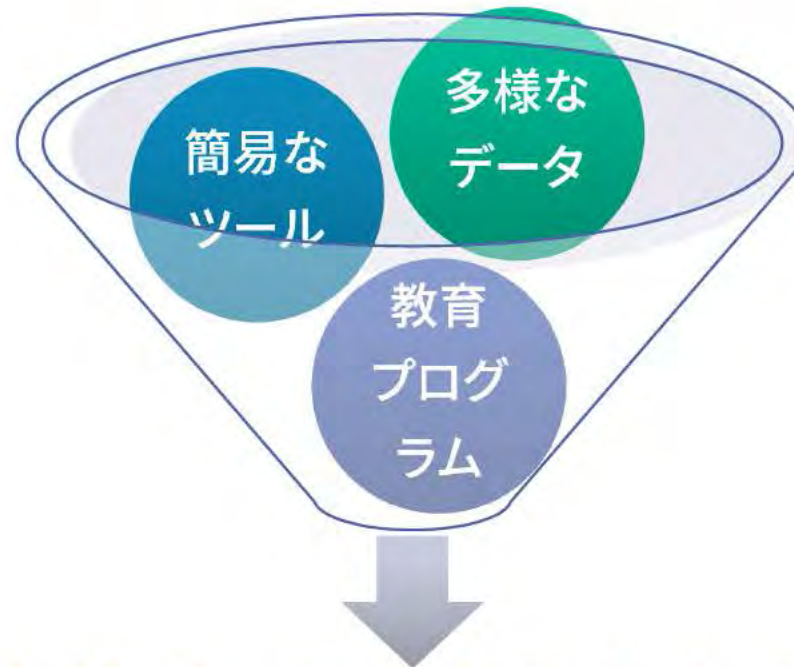
データサイエンスとデータビークル

課題：誰とデータサイエンスを実現するか？

構成要素：工数・時間・コスト ・ 分析スキル・データの質 ・ 分析データの多様性



データビークルが目指す未来

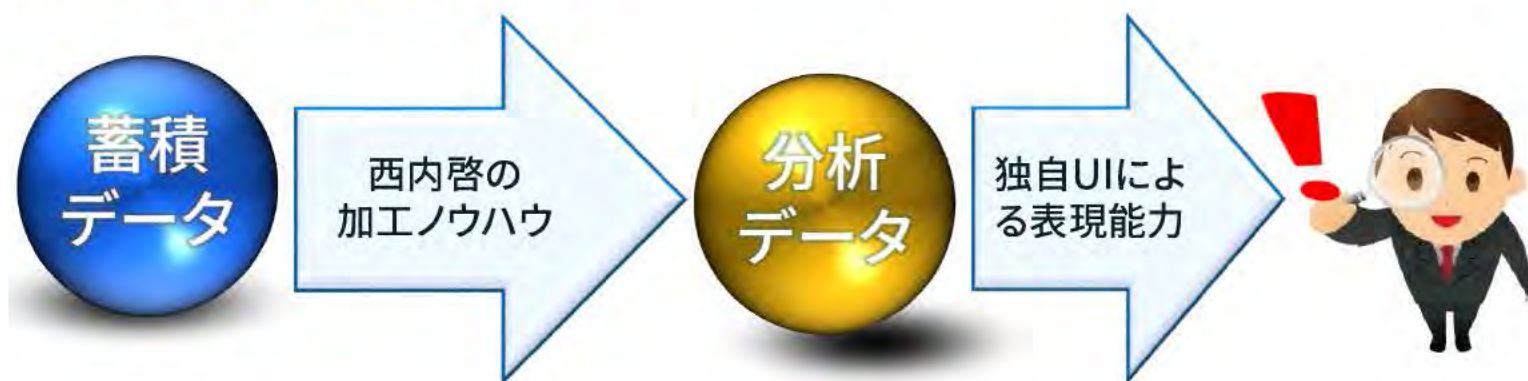


データサイエンスをみんなの手に



Data Diverとは

- ー従来のBIツールでもBAツールでもない「**リサーチデザイン**」ツール
- ープログラミング不要で高度な統計ノウハウを駆使できる
- ーリサーチデザインはトライ&エラーの連続
- ー現場がやるから仕事が進む



コンサルの時間の8割はこの機械的作業

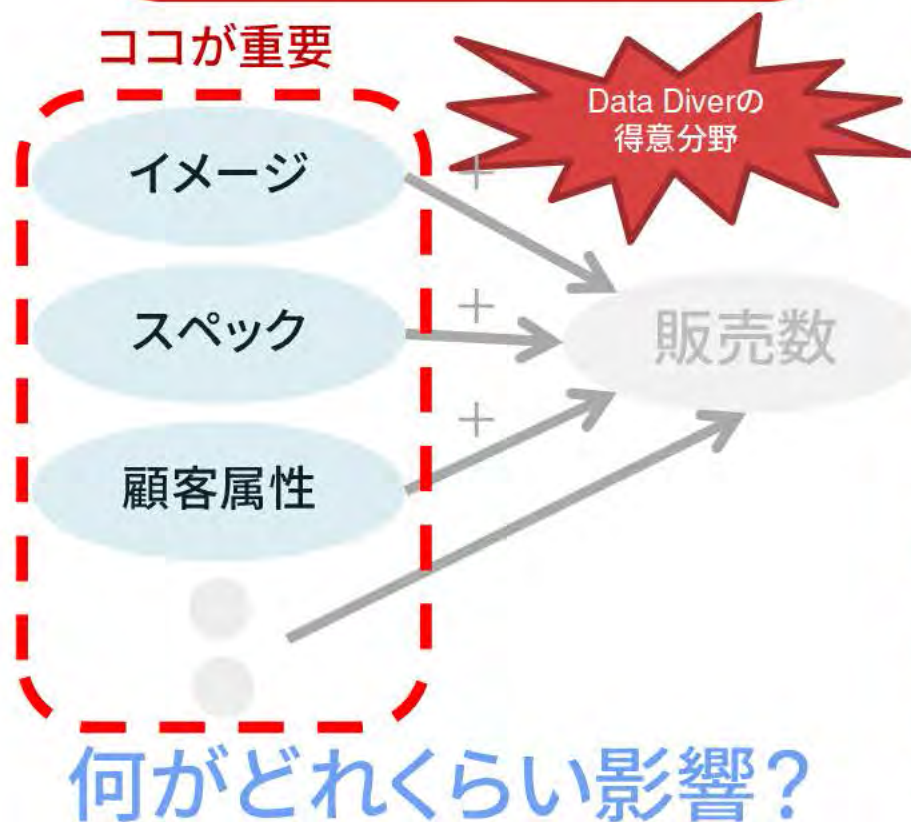
現場の苦勞の8割は結果の理解

統計学の知識、データ分析スキル不要で
世界トップレベルのデータサイエンティストと同等の
「気付き」

を得ることができます

こういう人におすすめ

洞察



予測



Data Diver 二大特長

1.蓄積データの分析データ化

2.二種の分析に絞った機能提供

「蓄積データの分析データ化」とは？

・蓄積データ(業務データ)

顧客ID	性別	年齢	..
ID001	男	25	..
ID002	女	28	..
ID003	男	35	..
ID004	女	32	..
ID005	女	23	..
.			
.			

顧客マスター



購入日	顧客ID	商品ID	数量	金額	..
2017年4月1日	ID001	A	2	480	..
2017年4月2日	ID001	B	1	150	..
2017年4月2日	ID001	A	2	480	..
2017年4月1日	ID002	A	1	240	..
2017年4月1日	ID002	C	5	1750	..
2017年4月1日	ID004	B	2	300	..
2017年4月10日	ID004	D	4	396	..
2017年4月15日	ID004	E	3	630	..
.					
.					

売上トランザクション



商品ID	商品名	単価	..
A	食品	240円	..
B	飲料	150円	..
C	食品	350円	..
D	その他	99円	..
E	飲料	210円	..
.			
.			

商品マスター

Data Diverがありとあらゆる項目を必要な形式・単位で自動集計し分析データ化

・分析データ(解析単位をキーに様々なデータを一つの表に集計)

解析単位(顧客ID)	性別	年齢	金額合計	購入回数	平均額	...	食品割合	飲料割合	...	月曜日割合	火曜日割合	...
ID001	男	25才	30000	5	6000	...	20%	60%	...	0%	40%	...
ID002	女	28才	48000	8	6000	...	10%	30%	...	20%	...	...
ID003	男	35才		2			50%	50%	...			
ID004	女	32才		10			20%	30%	...			
ID005	女	23才	4000	4	1000	...	25%	50%	...	0%	0%	...

顧客毎の購入総額、
回数などを計算

顧客毎の購入品
の割合を計算

顧客毎の購入曜日の
割合を計算など

二種の分析に絞った機能提供とは？

1.定量的な分析(重回帰分析)

数値の大小で比較できる貴社課題に最適

例：顧客の購入額を増やしたい



DATA DIVER

相関の有る仮説を出力

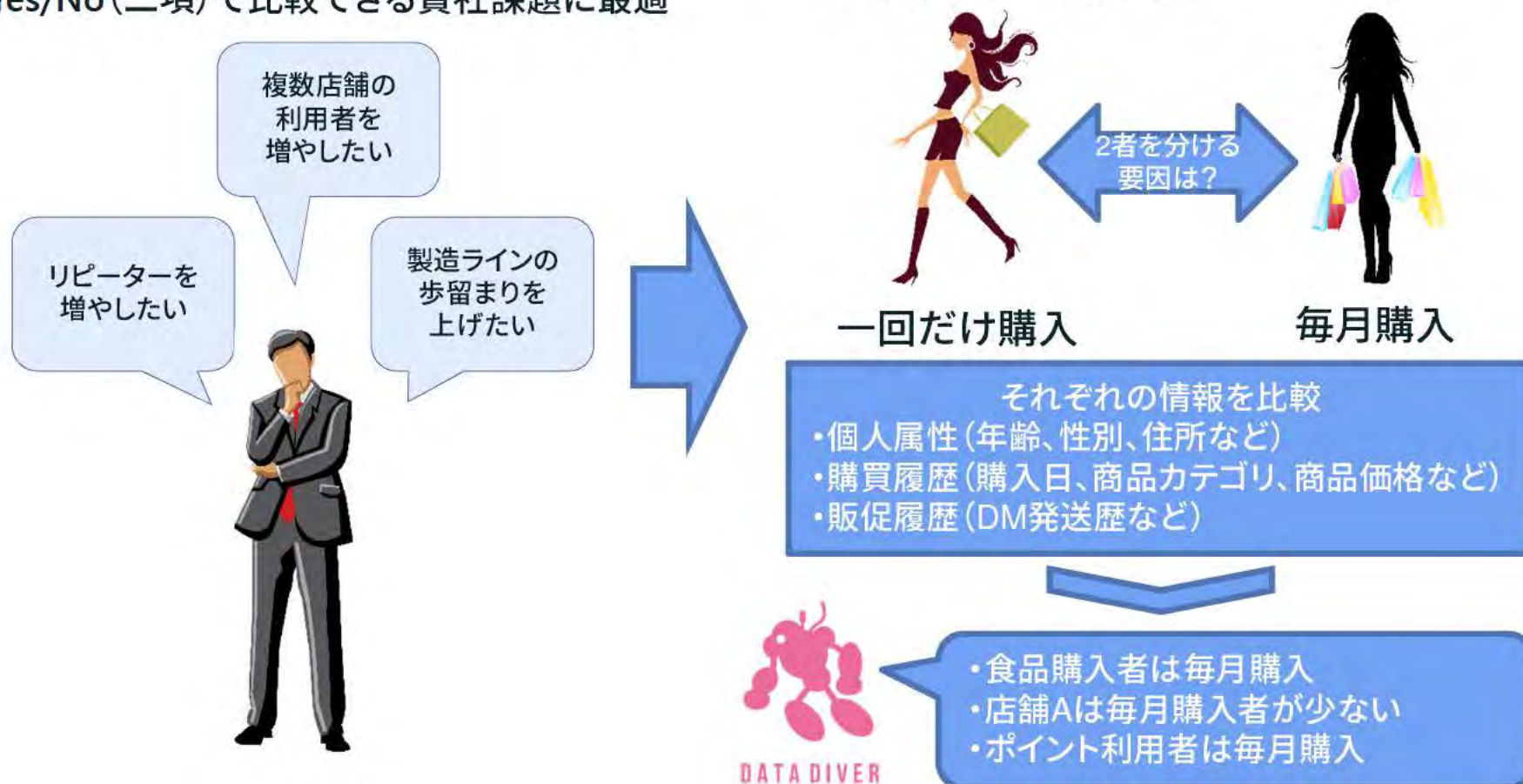
Copyright © 2017 DATAVEHICLE Inc.

二種の分析に絞った機能提供とは？

2.定性的な分析(ロジスティック回帰分析)

Yes/No(二項)で比較できる貴社課題に最適

例:リピーターを増やしたい



関連の有る仮説を出力

Data Diverがもたらす変化

・これまで



・これから



最新事例：ヤマハ発動機様・NTTドコモ様

報道資料発表



2017年8月7日

株式会社データビークル

ヤマハ発動機、データビークル「Data Diver」「Data Ferry」を
グローバルビジネスにおける分析ツールとして
テスト利用開始。

～現場が使えるデータサイエンスの実現に向けて～

株式会社データビークル（本社：東京都港区、代表取締役社長：油野達也）は、ヤマハ発動機株式会社がデータ分析ツールとして、「Data Diver（データダイバー）」「Data Ferry（データフェリー）」のトライアルを開始したことをお知らせします。

両製品は43万部発行のビジネス書のベストセラー「統計学は最強の学問である」の著者であり、データビークル副社長である西内啓がそのノウハウを詰め込んだ「現場が使える」データサイエンスツールです。従来、分析に使われてきたBI（ビジネスインテリジェンス）と呼ばれるグラフ描画ツールとは違い、データサイエンスに基づいた仮説（例：売上向上に影響する要素はなにか？）を短時間で探し、打ち手立案への貢献を実現します。

ヤマハ発動機ではグローバルビジネスにおいてデータサイエンスの普及と実践に着目しています。そこで短期間かつ低コストで習熟可能なデータサイエンスツール「Data Diver」、「Data Ferry」のテスト利用をすることとしました。



報道資料発表



2017年8月22日

株式会社データビークル

データビークルが提供する
「Data Diver」「Data Ferry」を
NTTドコモにてテスト利用開始

～現場が使えるデータサイエンスの実現に向けて～

株式会社データビークル（本社：東京都港区、代表取締役社長：油野達也）は、株式会社NTTドコモがインターネットサービス向けデータ分析ツールとして、「Data Diver（データダイバー）」「Data Ferry（データフェリー）」有償トライアルを開始したことをお知らせします。

両製品は43万部発行のビジネス書のベストセラー「統計学は最強の学問である」の著者であり、データビークル副社長である西内啓がそのノウハウを詰め込んだ「現場が使える」データサイエンスツールです。従来、分析に使われてきたBI（ビジネスインテリジェンス）と呼ばれるグラフ描画ツールとは違い、データサイエンスに基づいた仮説（例：売上向上に影響する要素はなにか？）を短時間で探し、打ち手立案への貢献を実現します。

NTTドコモでは、多種多様な分析ニーズに対応するため、短期間かつ低コストで習熟可能なデータサイエンスツール「Data Diver」、「Data Ferry」の1年間のテスト利用を開始しました。

公開事例：日本版DMCデータ分析基盤提供（既存サービスのパッケージ化）

モデルケースの作成

モバイル端末接続データ、目的地検索データ、
ID-POSデータ、食品成分マスタ、



TRUE DATA
Customer Communications, Ltd.

Data Loves Truth.



DATA DIVER



DATA FERRY

データ解析基盤

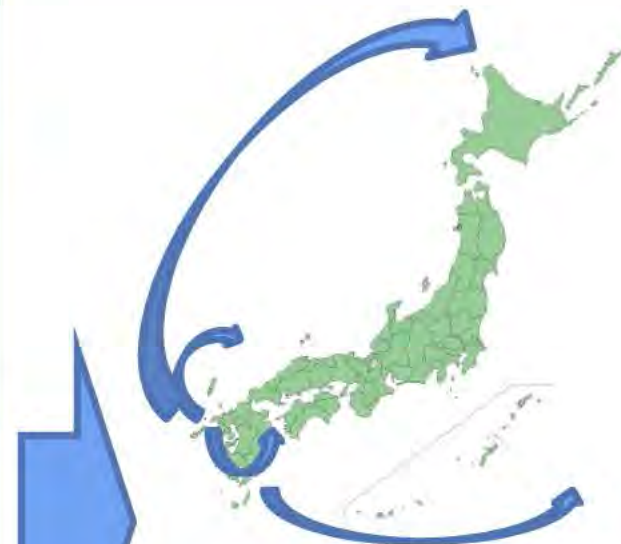


Microsoft

クラウド技術支援



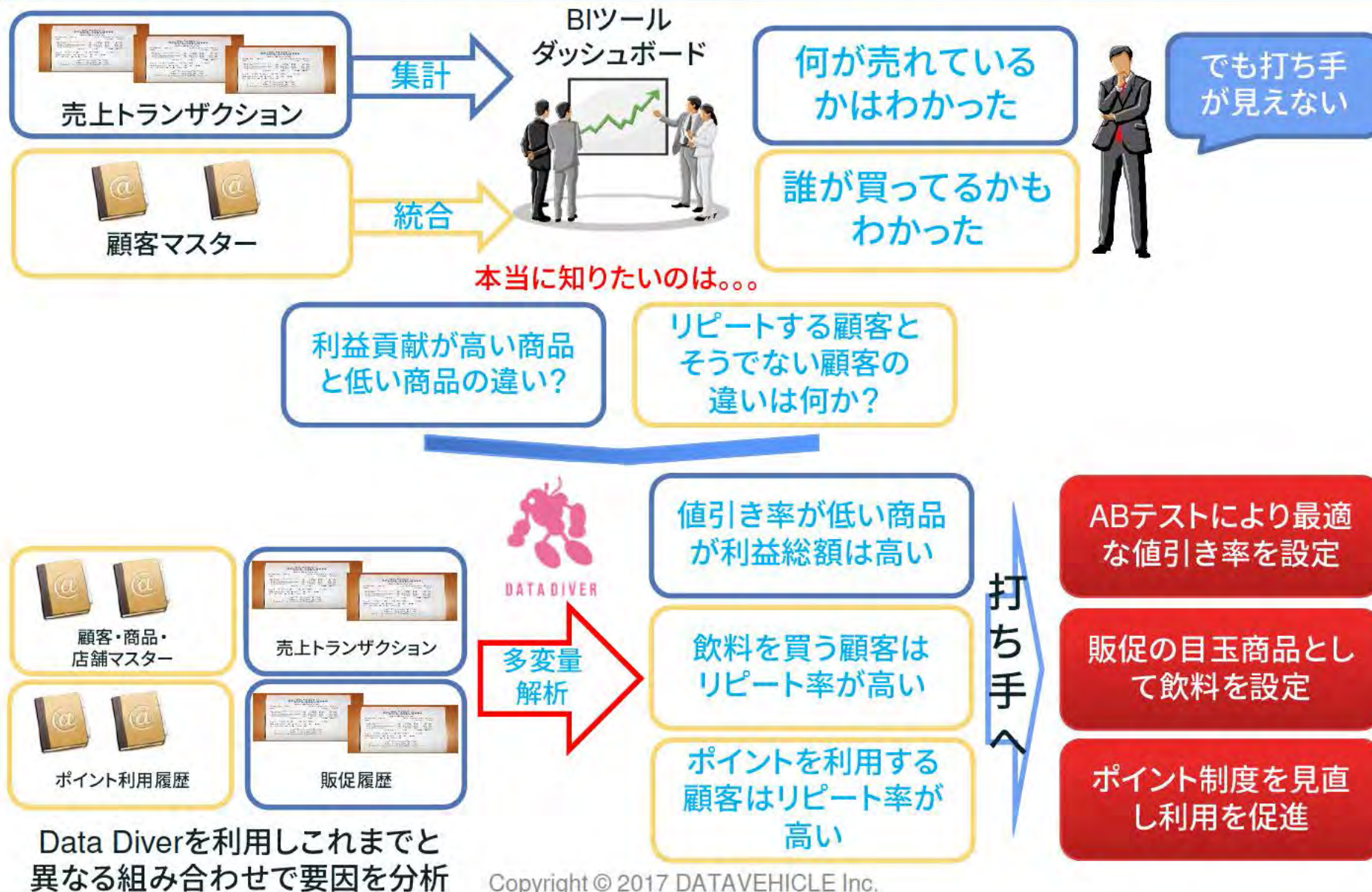
DMC=目的地型
観光振興会社



2020年に全国100 DMOへ
横展開を目指す

※2017年1月末現在120以上存在
参考：日本観光振興協会・会員数675

流通企業様事例(売上拡大・利益貢献)



製造業事例（不良率の削減）



Data Diverの分析時ポイント



Data Diverで分析に必要な3つの要素

御社や顧客について「何がどう変わるとうれしいか」ということを考えましょう。

例)

アウトカム:顧客の月間購買平均単価を上げたい。

解析単位:顧客/販促施策(DM、メルマガなど)/商品など

説明変数:年齢が高い顧客ほど一回の購買金額が高い/DMを月4回以上出すと売上が下がる/A商品はプレミアム版が売れているなど

1. アウトカム

- 「望ましさ」を具体的に定義
(売上、在庫コスト...)

2. 解析単位

- 「望ましさ」を比べる単位
(顧客、従業員、商品...)

3. 説明変数

- 「望ましさ」を左右しうる特徴
(属性、行動、心理特性、...)

画面イメージ1

- Data Diverで分析するデータを登録します。

The screenshot displays the Data Diver application interface, which is used for data analysis. The main window is titled 'Data Diver' and shows a list of registered data sources on the left and a central table of data. The table has columns for 'id', 'regist', 'customer_id', 'regist_date', 'gender', 'age', and 'pref'. The data is sorted by 'regist_date' in descending order.

Annotations highlight key features:

- カラム名やデータ型も変更できます** (Column names and data types can also be changed): This annotation points to the 'データ名' (Data Name) section on the left, which lists fields like 'customer_id', 'regist_date', 'gender', 'age', 'pref', '商品情報' (Product Information), '購買情報' (Purchase Information), and '登録情報' (Registration Information).
- 登録データのプレビュー機能があります** (There is a preview function for registered data): This annotation points to the central table of data, which shows a list of registered data points.
- 登録データ間の関係もD&Dで簡単設定** (Relationships between registered data can also be easily set with D&D): This annotation points to the 'Data Driver' section on the right, which shows a diagram of data relationships between '商品情報' (Product Information), '購買情報' (Purchase Information), and '登録情報' (Registration Information).

The bottom right corner of the interface shows the copyright notice: © Data Vehicle Inc. 2015-2016.

画面イメージ2

- アウトカム、解析単位を設定します。

Data Diver

分析 データ

新規分析

解析目的

customer_id

解析単位を選択

price

の 解析単位を選択して下さい

合計

件数

平均

アウトカムを定義

次へ >

簡単にアウトカムと解析単位の設定が行えます。

設定を終えると、あとは説明変数に向けた解析をData Diverが行います。

© Data Vehicle Inc., 2015-2016

画面イメージ3

- 解析～説明変数表示

自動解析

2

説明変数が一覧表示

分析結果	クリア	評価
ダイレクトメール_idの件数が1増えるごとに 購買情報.priceの合計は299.00低い傾向にあります。	△とてもクリア	👎
ダイレクトメール_send_dateがダイレクトメール_10月の件数が1増えるごとに 購買情報.priceの合計は5943.50低い傾向にあります。	△とてもクリア	👎
ダイレクトメール_send_dateがダイレクトメール_10月の割合が1ポイント増えるごとに 購買情報.priceの合計は51.73高い傾向にあります。	△とてもクリア	👎
ダイレクトメール_send_dateがダイレクトメール_1月の件数が1増えるごとに 購買情報.priceの合計は4409.51低い傾向にあります。	△とてもクリア	👎
ダイレクトメール_send_dateがダイレクトメール_3月の割合が1ポイント増えるごとに 購買情報.priceの合計は28.68高い傾向にあります。	△クリア	👎
ダイレクトメール_send_dateがダイレクトメール_4月の件数が1増えるごとに 購買情報.priceの合計は1712.83低い傾向にあります。	△とてもクリア	👎
ダイレクトメール_send_dateがダイレクトメール_土曜日の割合が1ポイント増えるごとに 購買情報.priceの合計は11.59低い傾向にあります。	△クリア	👎
ダイレクトメール_send_dateがダイレクトメール_深夜(24-4時台)の割合が1ポイント増えるごとに 購買情報.priceの合計は53.86高い傾向にあります。	△とてもクリア	👎
ダイレクトメール_send_dateがダイレクトメール_営業日の件数が1増えるごとに 購買情報.priceの合計は1977.34低い傾向にあります。	△とてもクリア	👎
商品ジャンルがおもちゃの件数が1つ増えるごとに 購買情報.priceの合計は2914.08低い傾向にあります。	△とてもクリア	👎
商品ジャンルがおもちゃの割合が1ポイント増えるごとに 購買情報.priceの合計は56.45高い傾向にあります。	△とてもクリア	👎

設定したアウトカムに影響を及ぼす説明変数を列挙
順番の入れ替え、削除を行い、より良い検索結果へ。

© Data Vehicle Inc., 2015-2016

画面イメージ4

- 説明変数の詳細をグラフで確認します。

統計的に補正されたデータを確認できるためこれまで見えなかった情報を見ることができます。



元データ



補正データ

画面イメージ5

- 有効と思われる分析結果を保存します。



主要スペック

入出力フォーマット	• CSV テキストフォーマット (ZIP形式も可) • SFTPによるバッチ投入 • その他フォーマット・連携方式は弊社別ツール「Data Ferry (データフェリー)」にて対応予定
管理データサイズ	• DataDiverとしてデータサイズの制限はありません。 ※利用可能な範囲は契約毎に設定されるストレージサイズ範囲内となります。 • 初期契約では50GBのストレージが提供されます。 ※有償にて追加可能
分析処理時間	• データ分析時には、サンプリングされたデータが 用いられ、数秒から数分で分析が終わります。

サービス仕様

- **提供方式**
 - クラウドサービスとして提供します。
※オンプレミスについては要相談
- **価格**

月額60万円(税別)最低契約期間6ヶ月
- **セキュリティ**
 - 契約毎に個別のシステムが割り当てられ、他社データへのアクセスは出来ません。
 - サーバアクセスは全てSSLによる暗号化が施されます。またご希望に応じてVPNの設定も可能です。
- **提供環境**
 - 1契約＝1アカウントでの提供となります(複数アカウントについては要相談)。
 - ストレージ容量は初期容量50GB、要望に応じて追加可能です(別途有償)。
※契約ストレージ容量内でアップロード、加工、分析、結果の保持を行います。
分析元データの最大サイズはストレージ容量の1/2から1/3が目安となります。

他製品比較

IoT・クラウド時代はデータ利用者が自分でデータ分析することを求められます

	データ前処理	分析操作	分析機能	利用者
一般的なBIツール	なし	手作業でドリルダウン	クロス集計で見える化	業務担当者 リサーチャー SE
SAS/SPSS/Rなどの専門ツール	コードを記述しデータを加工	分析手法やオプションを手作業で設定	記述統計 多変量解析 データマイニング	リサーチャー データサイエンティスト
Data Diver	独自ノウハウで自動化	3ステップですぐに結果へ	多変量解析	業務担当者

まとめ

Data Diverとは

- 従来のBIツールでもBAツールでもない「**リサーチデザイン**」ツール
- 統計学の知識、データ分析スキル不要で世界トップレベルのデータサイエンティストと同等の「**気付き**」を得ることができます
- 自社体制で**短サイクル**の分析~施策までを実現します

外部データサイエンティストの場合



シチズンデータサイエンティスト+Data Diver(Data Ferry)



