

データエンジニアリング研修 基礎編

GMOペパボ®

Day1 導入 (このスライド)

Day1 データの集合の形

Day2 データベースとデータの参照

Day3 複数テーブルの参照

Day4 データの要約・可視化

Day4 データを業務に活かす

Day1 導入

技術部データ基盤チーム 鈴木頼子 / GMO PEPABO inc.

2022.07.12 データエンジニアリング研修 基礎編 Day1

GMOペパボ®

「導入」パートでお話すること

1. この研修のゴール

研修により、何をできる状態になることを目指しているか

2. 研修の背景

目指しているデータ活用の状態と、今回の研修内容の関係性

3. 研修によりできるようになる具体例

研修で学んだことを活かすと、どのようなアウトプットができるか

4. 研修で学んでいく内容

ゴールに向けて、これから学んでいく内容を確認しておく

「データから知識を抽出して判断できるようになる」こと

もう少し噛み砕くと、判断したい対象に対して

1. どのような知識が得られれば判断できるか考えて
2. (必要なデータを取得してためておき)
3. そこからデータを取り出して集計し
4. その結果から知識を得て
5. それを元に何らかの判断を行う

そのために、データを扱う技術が必要になります

1. どのような知識が得られれば判断できるか考えて
2. (必要なデータを取得してためておき)
3. そこからデータを取り出して集計し
4. その結果から知識を得て
5. それを元に何らかの判断を行う

データを収集・蓄積する
技術

データを取得・集計する
技術

データを要約・可視化・
分析する技術

マーカ一部分が研修で扱う項目です

なぜ、上記の動きができるようになってほしいのか… 背景を次にお話します

データ活用で目指したい姿を俯瞰的に確認する

データに基づいた判断

a. 数値を処理した結果を元に、**人**が判断する

例: 集計した結果を見て、次に行う施策を決める

← 今回の研修はこちら

b. 数値を処理した結果を元に、**機械が自動**で判断する

例: 機械学習で各ユーザーに最適な表示を行う

← 最終的に目指す状態

この2つの段階を達成するために、どんな技術が必要になるかを次に話します

データに基づいた判断を行う段階

1 顧客接点のデジタル化

2 事業活動データの収集

3 データ蓄積・分析基盤

4 データ処理パイプライン

5 データ可視化とリテラシー

6 機械学習プロジェクト管理

7 マーケティング自動化

8 自動的な意思決定

1: 収集

データが出力され、取りまとめられている段階

2: 分析

取りまとめたデータを可視化、一元的に分析できる段階

3: 活用

データにより継続的なサービス改善を行える段階

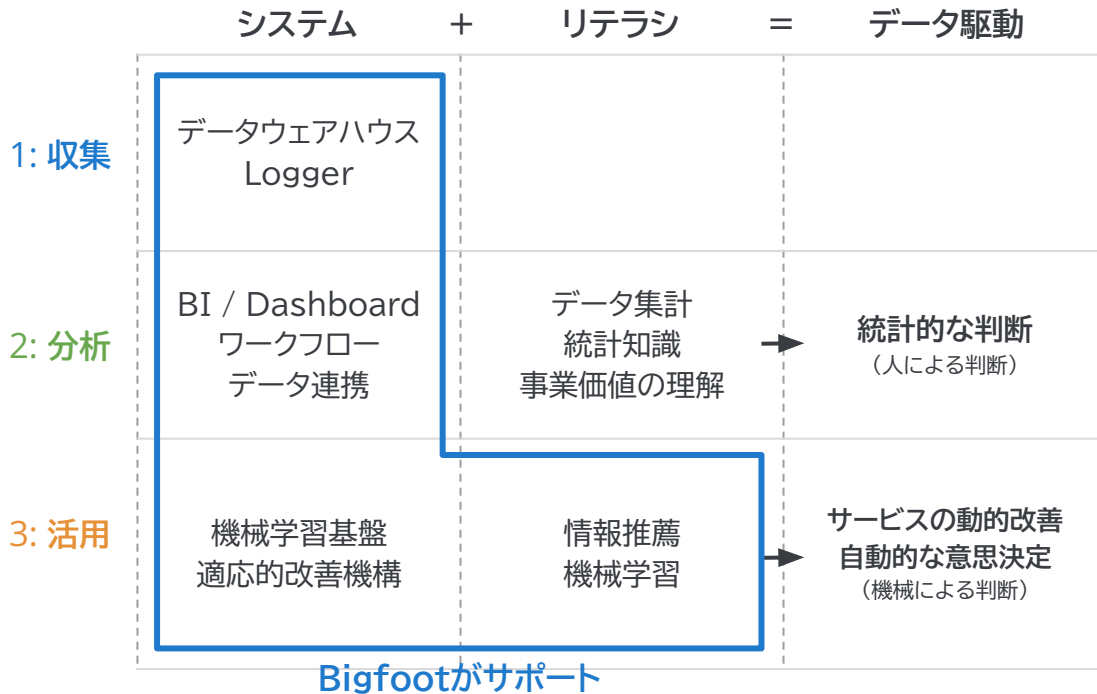
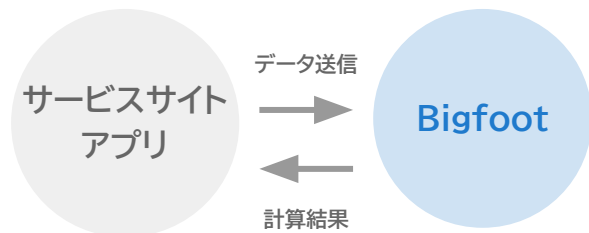
各段階で必要になる知識・技術

		実行のために 仕組みとして必要なもの	データを解釈するために 必要なもの	
段階		システム	リテラシ	データ駆動
1 顧客接点のデジタル化	1: 収集 データが出力され、 取りまとめられている段階	データウェアハウス Logger		この基礎ができることが 研修のゴール
2 事業活動データの収集				
3 データ蓄積・分析基盤				
4 データ処理パイプライン	2: 分析 取りまとめたデータを可視化、 一元的に分析できる段階	BI / Dashboard ワークフロー データ連携	データ集計 統計知識 事業価値の理解	→ 統計的な判断 (人による判断)
5 データ可視化とリテラシー				
6 機械学習プロジェクト管理	3: 活用 データにより継続的な サービス改善を行える段階	機械学習基盤 適応的改善機構	情報推薦 機械学習	→ 最終的なゴール サービスの動的改善 自動的な意思決定 (機械による判断)
7 マーケティング自動化				
8 自動的な意思決定				

全部できないといけないの？ …Bigfootがサポートします

Bigfootでできること

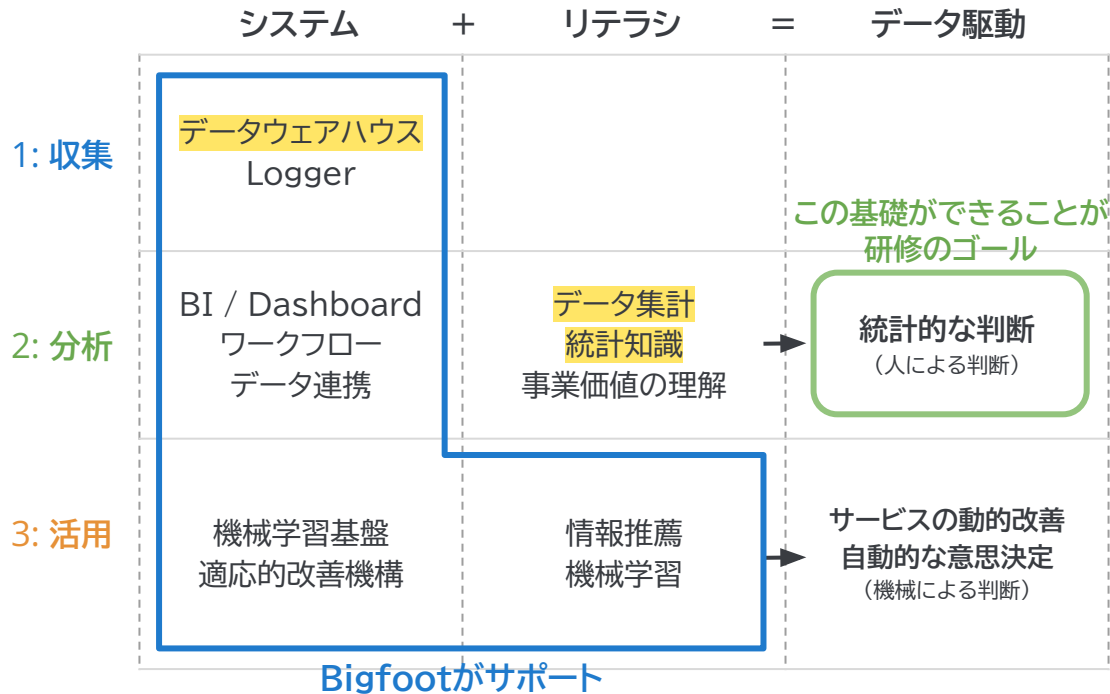
- データを受け取り、蓄積する
- データを元に計算する



この研修で学ぶこととの関係性

マーカーを引いた部分が研修で扱う項目です

- 人による判断ができるようになるために必要なデータ集計・統計知識の基礎を学びます
- システム側の基本的な仕組みも知ってもらうため、データウェアハウスにも少し触れます



全体像の話が続いたので、研修でできるようになる具体例も確認しておきます

ここでは具体的な社内での分析事例を紹介しました
公開資料からは削除しています

データに基づいた判断を行えるようになるために必要な力をつけていこう💪

1. Day1: データの集合の形
2. Day2: データベースとデータの参照
3. Day3: 複数テーブルの参照
4. Day4: データの要約・可視化
5. Day4: データを業務に活かす

データを取得し、それを要約・
可視化して情報を得る方法

データから得た情報を実際に
業務で活用していく方法