

Day4

データを業務に活かす

技術部データ基盤チーム 鈴木頼子 / GMO PEPABO inc.

2022.07.15 データエンジニアリング研修 基礎編 Day4

GMOペパボ®

「データを業務に活かす」パートでお話すること

1. データエンジニアリングの前後で必要になる工程

ここまでで学んできたデータエンジニアリングを意味のあるものにするために、
その前後でどのような工程が必要になるか

2. 前後の工程の内容

前後の工程で具体的に何をすればよいか

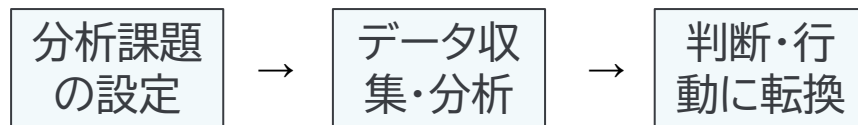
3. 前後の工程の注意点

前後の工程で陥りがちなポイントと気をつけること

データエンジニアリングを意味のある作業にするために

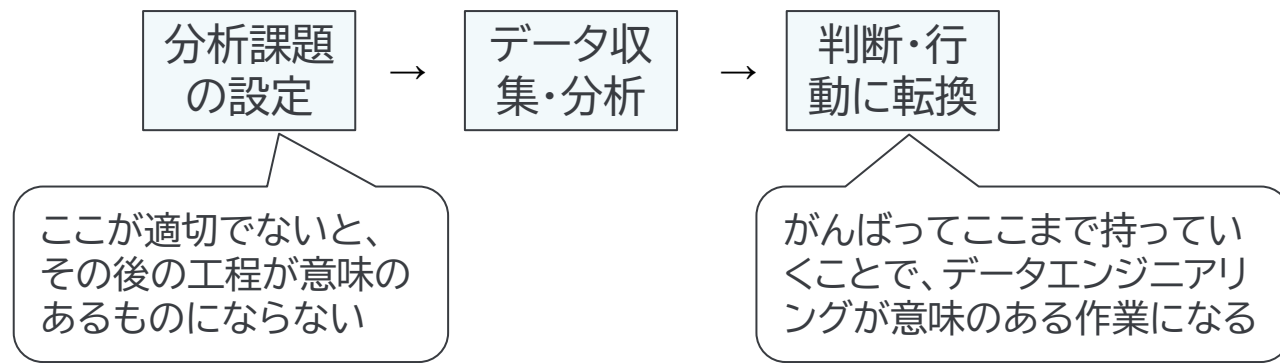
ここまでで学んできた、データエンジニアリングを意味のある作業にするためには、その前後も大切になります

- 前：分析課題の設定
- 後：判断して行動へ転換

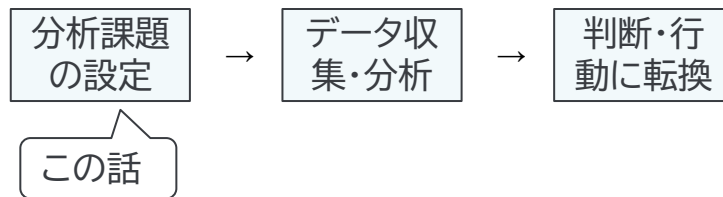


なぜ前後の動きが大切になるのか

- データエンジニアリングは、そのデータを何かに活用するために行います
- 事前にその目的を設定し、それに適した形で行う必要があるし、
事後に何かの判断に使わなければ、事業上は価値のある作業になりません



※ 分析により、その施策をやるべきでないと「わかった」ならそれも一つの成果で、無理やり行動に転換すべきという話ではないです 4



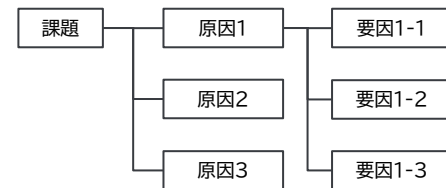
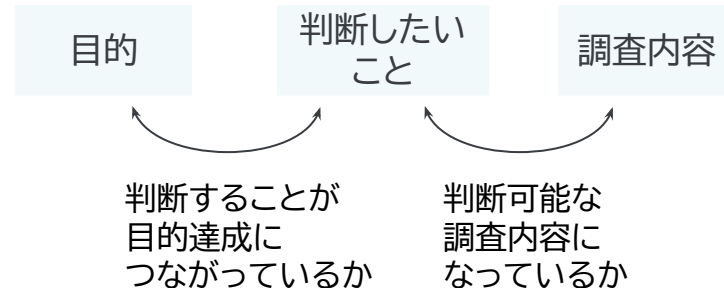
分析課題の設定

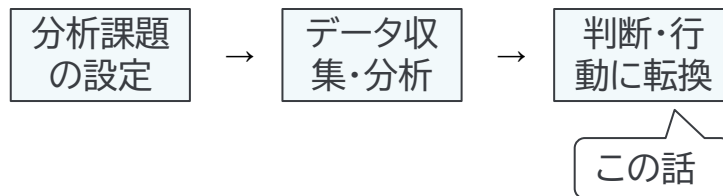
- 何を判断するために何を調べるのか決めます

- 目的と判断したいことと調査内容が
つながっているか確認しよう

- つながりを考えるために

- 頭の中で考えるだけでなく、文字に起こしてみる
- ロジックツリー等のフレームワークを利用して
分解してみる





判断・行動へ転換しやすくするために

- 分析課題の設定の時点で、施策への落とし込みを意識して設定しておこう
- 調査結果とそこからの判断内容を初見の人にもわかるよう伝えよう

陥りがちなポイント：目的と手段が逆転する

新しいデータが手に入った！
せっかくだから、これを使った施策をしたいなあ。
このデータからどんなことがわかるか、分析してみよう…

「新しいデータが手に入ったので、何か見出せないか分析してみよう」から、
今まで思いもしなかったすごい施策につながるような分析結果はまず得られません
(時間ばかりかかって、特にウィズダム*が得られなかった…となりがち)



データが手に入ったので分析するのではなく、
先に知りたいこと・仮説があって、それを明らかにするために分析するようにしよう

陥りがちなポイント：効果測定の方法を検討せずに施策を実施する

施策が終了したので効果を測りたいけど、どうやって測れば良いか、考えていなかったな…どうしたらいいだろう？

測定方法を決めずに施策を実施してしまうと、実施後に施策効果の測定ができずに困ることになります



施策実施前に、効果測定をどのように行うか決めておこう

- 効果測定に必要なデータを収集・蓄積しておく
- そのデータを使ってどのように効果を測るか決めておく

陥りがちなポイント：施策に転換できない

分析から～ということがわかりました！

へえ、おもしろいね！ それでそれって何に生かせるの？

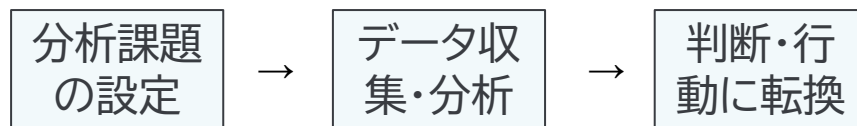
漠然と分析作業に着手してしまうと、何かしらの情報は得られたけど施策にはつながらない
…という結論に行き着く場合があります



着手前に、何のためにその分析を行い、その結果で何を判断するのかを明確にしておこう
仮説を立てておき、「仮説が正しいことを確認できたら～をする、正しくないことが
確認できたら～をする」と事前に決めておくぐらいでもよいです

データエンジニアリングの作業を意味のあるものにするために

- 前後の過程も大切になります



- 特に着手前の段階では、判断や行動への転換を行えるような調査内容になっているかをよく確認しておこう

研修内容は以上で終了です

研修全体を少しだけ振り返ります

この研修のゴールは「データから知識を抽出して判断できるようになる」でした

Day1 導入より

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

ゴールに向けて、以下の内容を学びました

Day1 導入

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

データを取得・集計する

データを要約・可視化する

データエンジニアリングを業務につなげる

この研修では、統計的な判断の基礎を学びました

	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	+			
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	→	この基礎ができることが 研修のゴール <div>統計的な判断 (人による判断)</div>
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 <div>サービスの動的改善 自動的な意思決定 (機械による判断)</div>
7 マーケティング自動化						
8 自動的な意思決定						

今後どうやって学んでいけばよい？

- クエリの書き方をもう少し練習したい
 - [SQLZOO](#) や [Progate](#) など、問題を解けるサイトがあります
- データエンジニアリングや統計の基礎を学びたい
 - データ基盤チームおすすめの本がまとまっているページ^(内部資料) があります
- 機械による自動的な意思決定の実践例を知りたい
 - 多腕バンディットで始めるデータ駆動^(内部資料) がわかりやすいです