

IR担当者のデータ分析プロセスに関する試行調査

2020年11月5日
IR実務担当者セッション (Zoom)

田中 秀典 (宮崎大学)、山本 鉦 (九州工業大学)、岡部 康成 (帯広畜産大学)、
鈴木 達哉 (山形大学)、山本 幸一 (明治大学)、藤原 宏司 (山形大学)

背景：IRの役割

IRとは、

情報提供を目的とした学内外のデータ収集・分析
を行う活動（機能）

目的を持ってデータ収集をしないと、何を分析したら良いのかわからない



結局、単純な集計だけで報告を終わる



そんなんじゃないよ！
もっと役立つ分析を！
IRの意味がない！



背景：データ分析におけるプロセスの重要性

どうすれば良い？

何か良いツールを導入すれば解決するのか？



いいえ

ツールよりも **プロセスが重要**



戦術



ツール



どのようなプロセスを経て分析を行っているのか？
使用するツールによってプロセスは異なるのか？



戦術に相当するテンプレートは存在する？

目的と手法

目的： データ分析プロセスにおける共通点の把握

手法： 明確なデータリクエストが与えられなかったというシナリオの元、架空のデータを用意して、複数のIR担当者における**データ分析のプロセス**を調査。

データマネジメント



分析



報告

■ 方法：データセット

当日投影のみで
ご説明いたします

方法：調査対象者（被験者）



担当者A

データ分析担当歴：約3年



担当者B

データ分析担当歴：約10年



担当者C

データ分析担当歴：約5年



担当者D

データ分析担当歴：約10年



担当者E

データ分析担当歴：約20年



担当者F

データ分析担当歴：約20年



使用するツールは自由に決めてもらい、実際に分析を行いながら
“どのように分析するのか”というステップをあえて書き出してもらった。
データ以外の情報は一切与えていない。

結果

今回、使用したツールは様々であった。

 担当者A
R→Rmd(xaringan)

 担当者B
RDB + R

 担当者C
Tableau

 担当者D
Tableau

 担当者E
Excel

 担当者F
R→PowerBI

目的に応じて好みのツールを使い分けていた

結果



担当者A

1. 調査目的の把握

- 今回は明確なリクエストがないため、設問の構成から推定した。

2. データ構造の確認

- 個人を紐付けられるキーの確認
- 個人の属性を表す項目の確認
- 設問形式の確認

3. データ全体の確認

- エラー値や欠損値の確認と、処理方法の決定
- 総数の確認（サイズによってツールを使い分ける）
- 分布などを確認（極端な偏りの有無）

結果



担当者A

4. キーとなる設問の分析

- 重要だと思われる設問について分析
(本来はリクエストに基づき対象の設問を決定)

5. 残りの設問の分析

- キーとなる設問と関連性の高い設問は、回答の関係性を確認
(クロス集計などで回答の傾向を把握する)

6. 説明資料の作成

- バラバラの分析結果を統合して、報告用のストーリーを構成
- 今回はRで分析を行ったので、R markdown+Xaringanで資料作成

結果



担当者F

1. データの理解

- アンケートの選択肢を見ながら、ポイントとなる部分を考える。

2. データ構造の確認

- 件数確認
- ID重複確認
- データのラベル処理
(1や2となっているものを、ラベルに変換)

3. データ全体の確認

- エラー値や欠損値の確認
- 総数や分布などを確認 (極端な偏りの有無)

結果



担当者F

4. ポイントに設定した設問の分析

5. 全体の分析

- 全ての設問の組み合わせでグラフを作成
(Rで一気に実行)
- 面白そうな傾向を示したものは属性でスライスして深掘り
- 報告候補を決める

6. 説明資料の作成

- 報告項目が多い場合PowerBIを使用してまとめる

結果



担当者C

1. データの理解

- ・ 項目や選択肢を見ながら、作成方針を考える

2. データ構造の確認

- ・ データの型を確認
- ・ キー・インデックスを作成
- ・ データをロング型に変更
- ・ データのラベル処理（リレーションで解決）

3. データ全体の確認

- ・ 総数の確認

結果

担当者C

4. 全体の分析

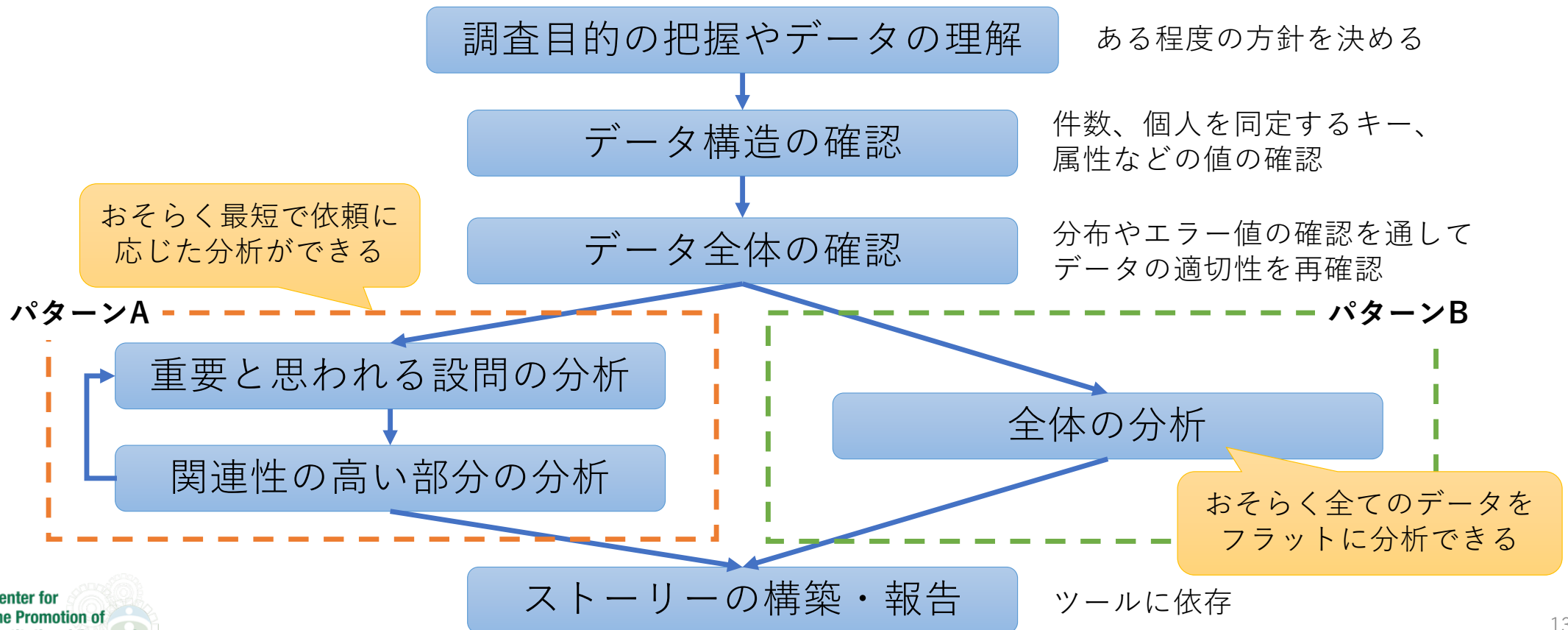
- グラフの標準パターンを作成して、それに各データを割り当てる
(実際はTableauのシートを複製し、フィルタでデータを変更)
- クロス集計シートの作成
- 探索的にシートを変更しながら、面白そうな傾向を示したグラフは複製して残す

5. プレゼンテーションに向けた準備

- 面白そうな傾向のものを中心に説明しながら、オンデマンド型で行う

結果

共通する部分もありながら、概ね2つのパターンが見られた



結果

対象とした6名は、概ね以下のパターンに分かれた

パターンA



担当者A

R→Rmd(xaringan)



担当者E

Excel



担当者F

R→PowerBI

架空のテーマを掲げて仮分析し、そこから関連性の高い分析を行いながら、分析を全体に展開したグループ

パターンB



担当者B

RDB + R



担当者C

Tableau



担当者D

Tableau

まずは全体を分析して、傾向を把握したグループ

結果



担当者D：専門員

Tableau

データ分析担当歴：10年

全体を分析して傾向を把握する方針

IR担当教員に相談しながら構築
必要に応じてさらに分析



相談しやすいように全体を分析して
スナップショットを準備

調査目的の把握やデータの理解

データ構造の確認

データ全体の確認

全体の分析

ストーリーの構築・報告

まとめ

データマネージメントに相当する部分は、概ね同じステップを踏んでいた。

分析は、架空のテーマを掲げて分析を開始するグループと、全体をまんべんなく分析するグループに分かれた。

➡ 分析経験年数よりも、**報告ツールに依存**している？

調査目的の把握やデータの理解

データ構造の確認

データ全体の確認

分析：2パターン

ストーリーの構築・報告

概ね同じ

まとめ

- データマネージメントの重要性が再確認された。
- 架空のテーマ設定を含め、分析方法には多様性が見られた。
- 分析方法は、使用するツールにも依存しているかもしれない。
 - ✓ 近年のBIツールを使えば、全体を網羅的に分析して、オンデマンドで切り口を変えながら報告できるから？
- 分析プロセスの全ては困難だが、データマネージメントを中心にテンプレート化することは可能だと思われる。
- 分析方法の多様性は、事例集のようなまとめ方が可能かもしれない。