

データ分析における問題定義や目的設定の有無によって生じる洞察結果の違いの調査

1532102 辻泰輝

指導教員：山崎治 准教授

1. はじめに

ビッグデータを意思決定の材料として活用するために、データ分析を活用したいと考えている企業は多く存在し、ビジネス現場での利用価値への期待は高まっている。しかし、目的通りに活用出来ていない企業が見られる。

データ分析を行うことで何かしらの結果は得られたものの、その結果が新しい発見に繋がるわけでも、ビジネスチャンスへ繋がるわけでもない。このような状況に陥ってしまう原因の1つとして考えられるのが、「データアプローチ」という考え方である(河村他, 2016)。「データアプローチ」とはデータありきで、その活用方法を導き出そうとする考え方のことである。本来データ分析というのは、現状抱えている課題の把握や、その課題を解決するために必要なデータや、分析の目的をまず明確に定める必要がある。その上で、データを収集し、分析を行うという流れが正しいデータ分析の考え方である。このような考え方を、河村他(2016)は「課題アプローチ」と呼んでいる。

2. 目的

本研究では、「データ分析の目的」に焦点を当て、目的を持ってデータ分析を行った場合と、そうでない場合での分析結果に差が現れるのかを調査し、それらの重要性を明らかにすることを目的とする。そこで実際のデータ分析を通じて考察を行う課題を利用した実験を行う。特に、目的設定に関して、実験者側から介入を行う参加者グループと、介入を行わないグループの違いに注目し、得られた洞察の内容について分析を行う。

3. 実験 目的の有無がデータ分析に及ぼす影響

実験では、目的のあり／なしが分析過程や結果の洞察に影響を及ぼすのかを調査した。

3.1 方法

実験参加者：本学情報科学部情報ネットワーク学科4年生10名と本学社会システム科学部経営情報科学科4年生2名が2人1組ごとに実験に参加した。計6組の参加者を、目的なし群(3組)、目的あり群(3組)に分けた。

実験計画：1要因2水準参加者間計画で行う。独立変数として目的の有無を取り上げ、「目的を与える」/「目的を与えない」の2水準を設ける。

材料：Kaggle社のWebページに掲載されている「Store Item Demand Forecasting Challenge (<https://www.kaggle.com/c/demand-forecasting-kernels-only/data>)」という商品売上のデータを用いた。同データを加工し、仮想的な実験用データとした。実験用データは「月」「日」「曜日」「店舗」「アイテム」「年齢」の6項目で構成され、それぞれに対して「売上(個数)」が示される。また資料として、1つの項目に対する売上を示すグラフ(単純集計)を6個、2つの項目を掛け合わせたものに対する売上を

示すグラフ(クロス集計)を15個Excel上で作成した。

手続き：実験は「データの観察・話し合い(25分間)・「記述(制限時間なし)」の二段階で構成される。目的あり群、目的なし群の共通目的として、実験用データから読み取れる「現状」と、それに対する「改善案」を提出することを求めた。この際、目的あり群のみに「既存商品・既存店舗についての弱みを知りたい」という分析の目的を伝えた。

「データの観察・話し合い」段階では、25分間2人でデータを観察しながら、気づいた点、考えた内容をメモ用紙として用意したA4用紙1枚に記入させた。また参加者の分析の様子をスクリーンキャプチャソフトで録画した。25分後、「記述」段階では、データを閲覧させず、メモ用紙に記した内容をまとめ「現状」と「改善案」の記述を行なわせた。参加者は、実験者がPowerPointで作成したフォーマットに「現状」および「改善案」を記入した。

3.2 結果

参加者が閲覧したグラフや、提出された現状と改善案に大きな差は見られなかった。しかし、閲覧したグラフが単純集計であったのか、それともクロス集計であったのかについての割合では、図1のような差が見られた。

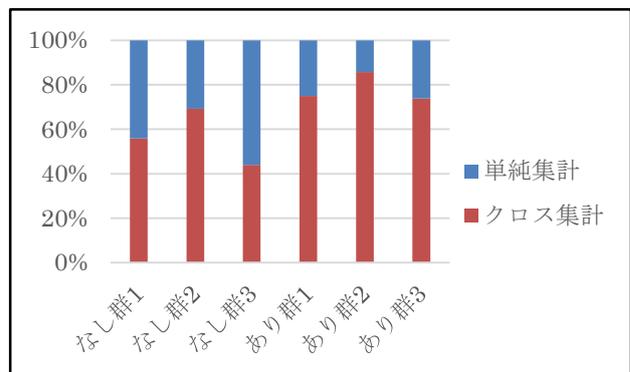


図1 単純集計とクロス集計の閲覧率

目的あり群は、実験時に提示された分析の目的に従い「アイテム」と「店舗」の項目に着目しやすく、それらの項目を他の項目と関連づけながら分析を進めたため、2項目に対する売上グラフ(クロス集計)をより多く閲覧したと考えられる。

4. まとめ

本実験は使用したデータや、実験者が提示した分析目的が単純であった為に、両群に明確な差が見られなかったと考えられる。今後の展開として、「分析で解決したい問題の定義」や、その問題定義に沿った「分析目的」をさらに詳細に設定し、データ分析過程の違いを検討していく必要がある。

参考文献

河村真一・日置孝一・野寺綾・西脇清行・山本華世(著)・日経情報ストラテジー(編)(2016). 本物のデータ分析力が身に付く本 日経BP社出版