

データサイエンス演習 予習教材

1 データサイエンス，データ分析とは

データサイエンスを辞書で調べると以下のように定義されている。

データサイエンス

データの分析についての学問分野。統計学、数学、計算機科学などと関連し、主に大量のデータから、何らかの意味のある情報、法則、関連性などを導き出すこと、またはその処理の手法に関する研究を行う。これらの研究者および技術者はデータサイエンティストとよばれる。(デジタル大辞泉、下線は大西による)

機械学習等の技術が強調されがちであるが、ビジネス等の実用においては、

何らかの意味のある＝価値のある＝問題解決に資する

がすべてである。例えば、「売上減少を食い止める」というビジネス課題に対して、データ分析は、

統計解析してグラフを作って、「売り上げ減少の要因は来店頻度が低いためである」

と報告して完了というイメージを抱きがちであるが、これでは問題解決に資するということにはならない。データ分析の目的（ゴール）は統計量の算出ではなく、

戦略および打ち手の構築

である。この課題に対しては、

- 売り上げ減少の要因は来店頻度が低いためである
- そこで、月3回以上来店する顧客には四半期ごとにキャンペーン告知のダイレクトメールを送付し、来店頻度の低い顧客には商品を活用したお役立ち情報のメールマガジンを1か月に2回送付する
- なぜなら、月3回以上来店する顧客は、継続的に来店し、1回当たりの購入金額も高いからである

という報告を行うのが問題解決に資するデータ分析である。

2 データ分析のプロセス

2.1 データ分析のステップ

1. 目的・問題の特定

- 「10%下がった売上を元に戻す」といった目的を明確にする
- 「売上」といった指標を決める

2. 現状把握，課題の特定

- 「売上」が下がった原因がどこにあるかを特定する
「商品」「店舗」「顧客」…
- What を特定

3. 要因の特定

- 「なぜ特定の顧客層への売上が下がったのか」といった、課題が生じる原因を特定する
 - － 「リピーターが減ったのはフォローが足りないから」
 - － Why を特定

4. 方策（戦略，打ち手）構築

これらのプロセスについて次の事例を用いて説明する。

社長から「売上が落ちているので，売り上げ増加策を考えろ」と命じられた

2.2 目的・問題の特定

「問題とは，望ましい状態と現状との差異で，問題解決とはそれらの差異を解消すること」とよく説明される。目的は，望ましい状態，あるいは望ましい状態にすることである。誤った目的の下で突き進むと，必然的に望ましくない結果になる。

この事例では経営者より「売り上げを増加する」という明確な指示があるので，目的は「売り上げを増加する」ということで一応はよいことにするが，「売り上げをXX年の水準に戻す」のように具体的にになっていることが望ましい。指標も売上でよいであろう。

この事例では指示が明確であるが，指示が売上増加なのか利益増加なのか，はたまた利益率増加なのかははっきりしないこともあるかもしれない。もっとひどい場合には，ブームに煽られて，「データを使って何かやれ」，「AIを使って何かやれ」という指示（無茶ぶり）もあるらしい。指示があいまいな場合は，自ら目的や問題を定義し，指示を出した上司に確認をとる必要がある。

2.3 現状把握，課題の特定

仮説思考による課題の特定 売上の減少という問題では，すべての商品，すべての店舗，すべての客層等で様に売上が減少するということはあまりなく¹，特定の部分が大きな問題を起こしていることが多い。そこで，どこが問題を起こしているか，すなわち何が課題であるかを探る。その際，行き当たりばつりに調べるのではなく，

- 店舗別に売上増減に差があるのではないか
- 商品別に売上増減に差があるのではないか
- 顧客別に売上増減に差があるのではないか

といった課題の候補（仮説）に分解して，課題を探すのが効率的かつ効果的である。

課題について仮説を立てる際，図1のようなロジックツリーを用いると，問題の構造が可視化されて，自分の考えを整理したり，他者と共有するのに役立つ。ロジックツリーの枝分かれが仮説の細分化（分解）に相当するが，分解は抜けなくだぶりなく（MECE; Mutually Exclusive and Collectively Exhaustive）行うことが重要である。

課題の特定は大きな傾向を見るところから始め，課題となる部分を深掘りするようにする。図1に示すロジックツリーは非常に小さいものだが，それでもすべて検証するのは大変である。目的に近い大きな傾向から検証し，課題がありそうな仮説だけさらに深掘りするようにするのが効率的である。

¹例えば，企業の不祥事等でブランド価値が落ちるとこのようなことが起こる。

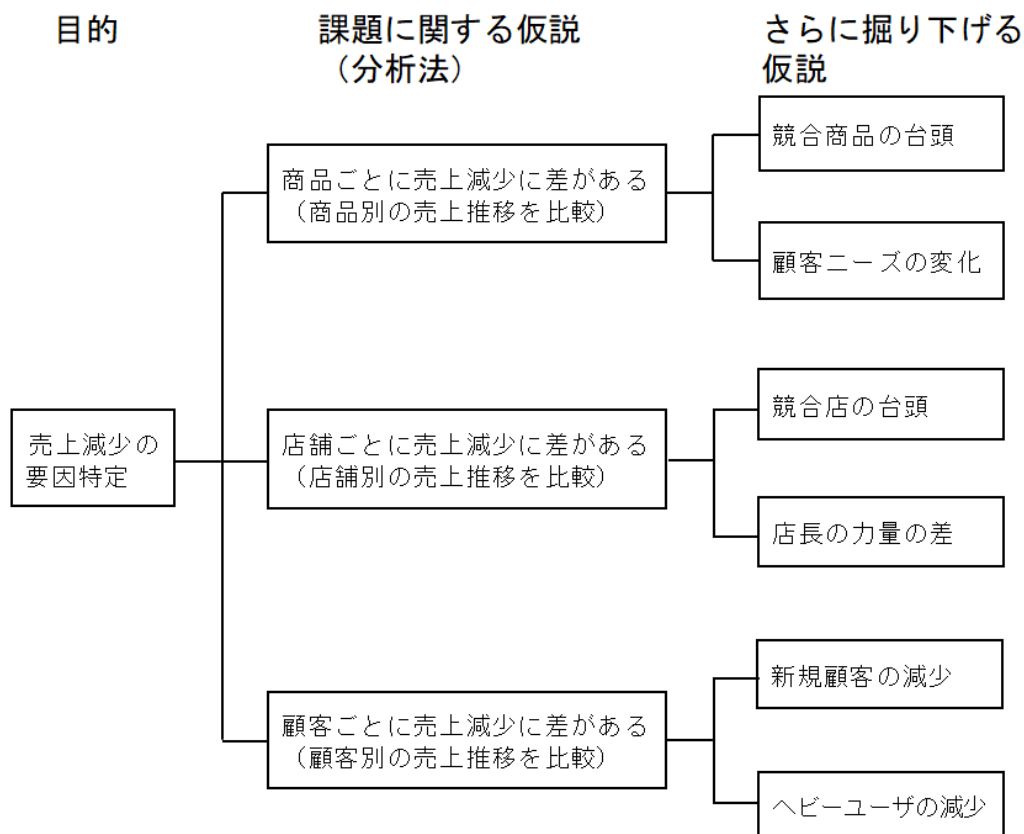


図 1: ロジックツリー

最も大きな傾向 表 1 に直近 5 年間の売上の推移を示す。年平均で 3.7%，5 年間で約 19% も売上が減少しており，確かに改善が必要そうである。

表 1: 売上推移

年	売上 (100 万円)
2014	1000.00
2015	948.14
2016	902.08
2017	854.12
2018	813.01

商品別の売上推移の比較 ロジックツリーに従い，商品別に売上の推移を比較する。商品別の売上の推移は表 2 に示す通りである。どの商品も 5 年間で売上が減少しており，減少率は 17%～20% 程度であり，商品により多少のばらつきがあるが，飛びぬけて減少率が大きい商品はないことから，売上の減少の要因として足を引っ張る商品はなさそうである。

表 2: 商品別の売上推移

年	商品 1	商品 2	商品 3	商品 4	商品 5	商品 6
2014	66.90	262.41	62.30	276.84	147.19	184.36
2015	63.46	250.47	59.18	261.91	138.54	174.57
2016	60.13	241.45	55.88	247.70	131.97	164.94
2017	57.70	228.01	53.74	233.99	123.62	157.05
2018	54.31	217.72	51.01	222.02	118.00	149.95
増減	-18.8%	-17.0%	-18.1%	-19.8%	-19.8%	-18.7%

店舗別の売上推移の比較 次に、店舗別の売上推移の比較を行う。店舗別の売上の推移は表 3 に示す通りである。どの店舗も 5 年間で売上が減少しており、減少率は 17%~21%程度であり、店舗により多少のばらつきがあるが、飛びぬけて減少率が高い商品はないことから、売上の減少の要因として足を引っ張る商品はなさそうである。

表 3: 店舗別の売上推移

年	店舗 A	店舗 B	店舗 C	店舗 D	店舗 E	店舗 F	店舗 G	店舗 H
2014	240.27	162.63	93.11	58.89	25.33	231.66	138.44	49.66
2015	229.80	153.91	88.81	55.49	23.99	218.08	130.96	47.09
2016	219.40	145.10	85.91	52.29	22.63	207.76	124.31	44.69
2017	208.71	137.09	81.26	49.47	21.34	195.62	118.16	42.46
2018	197.74	130.59	77.24	46.72	20.33	187.65	112.08	40.66
増減	-17.7%	-19.7%	-17.0%	-20.7%	-19.7%	-19.0%	-19.0%	-18.1%

顧客別の売上推移の比較 続いて、顧客別の売上推移の比較を行う。ここでは、新規顧客と既存顧客に分け、既存顧客は年間購買回数で分けた。顧客別の売上の推移は表 4 に示す通りである。すべての顧客に関して売上減少傾向にあるが、既存顧客が微減なのに対して（年間 3 回購買の顧客 7.9%減は微減とはいえないが）、新規顧客が 5 年間で 54.6%減少と顕著である。新規顧客は、2014 年には全売上の 30%、2018 年でも全売り上げの 16.7%を占めており、新規顧客の減少が売上減の最大の課題と言える。

表 4: 顧客別の売上推移

年	既存顧客（購買回数）				新規顧客
	≥ 4	3	2	1	
2014	100.00	150.00	200.00	250.00	300.00
2015	98.81	148.46	196.13	246.97	257.77
2016	98.18	143.84	196.19	247.80	216.07
2017	97.80	141.65	194.92	246.41	173.35
2018	97.92	138.14	194.18	246.61	136.15
増減	-2.1%	-7.9%	-2.9%	-1.4%	-54.6%

2.3.1 要因の特定

先の分析で, 課題は新規顧客の売上減少ということが分かった. 次のステップでは, 新規顧客の売り上げ減少の要因を特定することである. 新規顧客の減少ということなので, 広報に問題がありそうだという仮説が立つ. そこで, 表5に示す損益計画書 (P/L) を見ると, やはり広報費が大幅に削減されていることが判明した.

表 5: 損益計算書 (P/L)

	売上	原価	粗利	人件費	物流	広報	その他	販管費計	営業利益
2014	1000.00	479.39	520.61	149.50	89.32	198.72	48.12	286.94	233.67
2015	948.14	451.28	496.86	148.21	78.83	148.78	49.47	276.50	220.36
2016	902.08	431.95	470.13	151.62	71.92	144.28	48.92	272.46	197.67
2017	854.12	415.68	438.44	160.74	65.48	98.30	44.57	270.79	167.65
2018	813.01	388.43	424.58	158.79	56.88	79.01	45.88	261.56	163.02
増減	-18.7%	-19.0%	-18.4%	6.2%	-36.3%	-60.2%	-4.7%	-8.8%	-30.2%

2.3.2 方策 (戦略, 打ち手) 構築

ここまでの分析から, 売り上げ減少は新規顧客への売上減少によるところが大きく, 広報費の大幅な削減による広報不足が減として考えられることが分かった. 次のステップは方策を構築することである. 方策は要因に対しての方策である必要がある. すなわち, 新規顧客となる人々を来店させるための広報活動となる.