

人材開発部門の データ活用

「社員意識調査」と「多面評価」を柱として

半蔵門オフィス 代表 南雲道朋

第7回 データ分析の基本 1 (Step1 ~ 3: 現状を可視化する)

今回から、データ分析の具体的な方法に入ります。前回紹介した7つのステップのうち、まずはStep1からStep3の「現状を可視化する」ステップを取り上げます。

Step1: 件数を数える

(1) まず1つの数表に整理

まず、回答データの全体を1つの数表として整理し、いつでも参照できるようにします。そのこと自体が、社員の意識改革にとって重要なステップとなります。データがない世界では、「みんなそう言っているよ!」という言葉が決め言葉になり、空気で意思決定されることになりがちです。しかし、データへのアクセスが可能になり、「何人がそう言っているのか」、「具体的にどのような発言があったのか」という問いを発することができるようになると、社員の意識は変わり始めます。

集計の前に、まず生データのままだと全体を眺めまわしてみることが重要です。その際、回答値の大小に応じて自動的に着色する

EXCELの「条件付書式」機能を用いて、セルに色をつけてみることをおすすめします。それだけで、まずはどこに山がありどこに谷があるのか、といった全体傾向がわかります。また、扱いに注意が必要な異常値（全問1点となっている等）や、エラー値が発見されるかもしれません。

(2) 「イチゼロ」形式の数表が基本

回答データは、ある回答選択肢を「選択したのか(1) / しなかったのか(0)」という、イチゼロの形式でそろえることが基本になります。そのイメージを図表1に示します。

イチゼロ形式の数表は、いかなる形式の設問にも対応でき、かつ、そこからいかなる集計・分析もできる、万能の土台となります。5段階の「そう思う」尺度設問であっても、「複数の選択肢から1つ以上を選択する設問」であっても、同じ形式の中で分析できますし、「自由記述設問」でさえ、「回答しているかどうか」、あるいは「ある単語が含まれているかどうか」といった見地から、同じ形式の中

で分析できます。

5段階の「そう思う」尺度設問において、5つの選択肢に1点から5点までの数値を割り当てて登録されるようシステム化しておけば、イチゼロ形式の数表にしなくても、「1~5点」の点数をそのまま数表化できます。しかし、いったんすべてイチゼロ形式にすることが基本です。それにより、「平均点だと2.1点と著しく低い」、その中で『まったくそう思わない』と回答した人は何人いるのか」といった具体的な掘り下げも、直ちに可能になります。

性別、部門といった属性についても、「男か女か」、「営業部門か製造部門か管理部門か……」といった選択肢別に、「そうであるか(1) / そうでないか(0)」のイチゼロ形式にしておきます。そうすることで、後のステップでの分析が容易になります。

(3) 回答者人数をそのまま語る意義

「どの回答選択肢を選択した人が何人いるか」という「人数の単純集計」は、最も初歩的で抽象度が低い集計方法です。しかし、そ

図表 1 設問別・選択肢別の集計（イチゼロ形式の数表から平均点化へ）

平均点化	平均点	【設問】 当社の社風はオープンである						
		5段階 回答値	5強く 思う	4そう 思う	3どちら でもない	2そう 思わない	1全く 思わない	ポジ ティブ 回答
件数カウント	全体	2.82	2%	24%	34%	34%	6%	26%
	機器部門	2.83	2%	24%	34%	35%	5%	26%
	設備部門	2.84	2%	27%	31%	32%	7%	29%
	装置部門	2.73	2%	19%	36%	35%	8%	21%
	本社部門	2.85	3%	25%	34%	33%	6%	27%
	件数	【設問】 当社の社風はオープンである						
		5段階 回答値	5強く 思う	4そう 思う	3どちら でもない	2そう 思わない	1全く 思わない	ポジ ティブ 回答
	全体	3772	82	907	1265	1284	234	989
	機器部門	1705	40	407	575	593	90	447
	設備部門	878	15	240	276	285	62	255
	装置部門	647	12	126	231	229	49	138
	本社部門	542	15	134	183	177	33	149

元データ（選択肢別：イチゼロ形式）

回答 連番	部門	性別	年齢	元データ 【設問】 当社の社風はオープンである						
				5段階 回答値	5強く 思う	4そう 思う	3どちら でもない	2そう 思わない	1全く 思わない	ポジ ティブ 回答
1	機器部門	男性	50～59歳	3	0	0	1	0	0	0
2	設備部門	男性	50～59歳	2	0	0	0	1	0	0
3	機器部門	女性	50～59歳	3	0	0	1	0	0	0
4	機器部門	男性	50～59歳	2	0	0	0	1	0	0
5	機器部門	男性	50～59歳	2	0	0	0	1	0	0
6	装置部門	男性	50～59歳	2	0	0	0	1	0	0
7	機器部門	男性	50～59歳	4	0	1	0	0	0	1
8	機器部門	女性	40～49歳	3	0	0	1	0	0	0
9	機器部門	男性	40～49歳	4	0	1	0	0	0	1
10	機器部門	男性	50～59歳	2	0	0	0	1	0	0

れだけに集計上のブラックボックスやごまかしが入る余地がなく、その結果からは逃げも隠れもできません。その意味では、最も強力な集計方法であるともいえます。

たとえば、社員意識調査の結果を経営幹部で振り返り、ディスカッションしようという場合、あえてデータの加工は最小限にとどめ、設問ごとに「〇〇と答えた人が〇〇人いる」という生々しいデータを参照しながら議論を行うことが、効果的な場合があります。「当社の将来性はあるか」という設問への社員の回答結果を、「平

均点は3.17点と高くありません」と振り返るよりも、「全体の21%にあたる450人が、『まったくそう思わない』、あるいは『そう思わない』とネガティブに回答しています」と振り返るほうが、はるかに実感がわきやすく、その回答者の顔ぶれを想像しながら、「その450人の内訳はどうなのか」、「役職別に見るとどうなのか」、「年齢層別ではどうか」、「部門別ではどうか」と、ディスカッションは進むのです。

その場合、「平均3.17点」という1つの数字によってではなく、

「5点は120名、4点は710名……」という形で説明することになるため、報告資料の紙数が多くなってしまふデメリットはありません。資料化せずに、話し合いをしながら、その都度画面上で「その450人の内訳は、役職別に見るとこうです」、「年齢層別に見るとこうです」などと、細かく掘り下げて見る（ドリルダウンする）ことができるのであれば、あえて平均3.17点などと指標化する必要はないともいえるかもしれません。自在にデータを見ることができるITツールが発達することで、あえて高度で抽象的な集計・分析を行う必要性は薄れる、という逆説さえ成り立つのです。

Step2: 平均点化する

（1）回答者数から回答者率へ

選択肢ごとに回答者の人数を足し上げたあとに、それを回答者総数で割った平均を出すことにより、選択肢別の「回答人数」は選択肢別の「選択率」へと変わります。それによって、回答者母数が異なっても比較が可能となります。たとえば、「この設問に『まったくそう思わない』と回答した人は、男性が235人、女性が215人です」といったところで、男性と女性で母数が違えば比較する意味はありませんが、「男性は46%、女性は57%です」といえば、

図表2 設問別の集計（5点満点中の平均点から偏差値化へ）

偏差値化	偏差値	設問別集計						
		オープンな社風	革新的な社風	チャレンジの奨励	バリューの浸透	ビジョンの浸透	当社における意思決定スピード	部門間の円滑な情報流通
↑	全体	50.0	50.0	50.0	50.0	50.0	50.0	50.0
	機器部門	50.1	50.2	49.9	50.2	50.0	50.5	50.2
	設備部門	50.2	50.5	49.7	50.7	50.1	49.8	52.2
	装置部門	49.0	48.5	49.9	48.4	49.4	48.6	46.7
	本社部門	50.4	50.2	50.9	50.3	50.6	50.3	49.7
↑	全回答の標準偏差	0.94	0.80	0.83	0.89	0.88	0.88	0.90
	平均点	2.82	2.14	2.72	2.84	2.69	2.65	2.69
	機器部門	2.83	2.16	2.72	2.85	2.69	2.70	2.71
	設備部門	2.84	2.18	2.70	2.90	2.70	2.64	2.89
	装置部門	2.73	2.03	2.72	2.69	2.64	2.53	2.39
	本社部門	2.85	2.16	2.80	2.87	2.75	2.68	2.66

元データ（設問別：5択に点数割当）

回答番号	部門	性別	年齢	オープンな社風	革新的な社風	チャレンジの奨励	バリューの浸透	ビジョンの浸透	当社における意思決定スピード	部門間の円滑な情報流通
1	機器部門	男性	50～59歳	3	2	3	4	3	4	3
2	設備部門	男性	50～59歳	2	2	2	3	2	3	3
3	機器部門	女性	50～59歳	3	2	3	2	2	3	2
4	機器部門	男性	50～59歳	2	1	1	3	2	3	3
5	機器部門	男性	50～59歳	2	2	3	3	2	2	3
6	装置部門	男性	50～59歳	2	3	3	4	3	2	3
7	機器部門	男性	50～59歳	4	3	3	2	3	2	3
8	機器部門	女性	40～49歳	3	1	2	4	1	2	2
9	機器部門	男性	40～49歳	4	2	3	3	3	3	4
10	機器部門	男性	50～59歳	2	2	3	2	2	3	2

比較に意味が出てきます。

（2）設問別の平均点へ

選択肢別の集計では、5択の場合には、1つの設問に対して5つの値が対応することになり、設問間の比較が単純にはできません。そうした場合、設問ごとに「5点満点中の平均点」という1つの指標にまとめることによって、設問間の比較ができるようになります。そこから、どの設問項目が相対的に問題なのかを議論できます。そのイメージを図表2に示します。そして、一覧表にしたうえで、点数の高低に基づいて着色す

ることで、グラフにするまでもなく、問題の把握が容易になります。なお、そのように着色した数表は「ヒートマップ」と呼ばれます。

設問別の結果を1つの数値で表す指標として、「平均点」以外に「中央値」や「最頻値」といった指標もあります。「平均点／中央値／最頻値／最大値／最小値」を組み合わせることで、回答の分布を正しく把握できる、という考え方もあります。しかし、項目と項目の比較や、どの項目がどの項目に影響しているのかといった項目間に関連づけた分析は、平均点に基づ

くことが妥当であり、「平均点および平均点を中心とするバラツキ」を基礎にして、豊富な統計分析手法が組み上げられています。よって、平均点を用いることをまず考えるべきです。

（3）平均点としてのポジティブ回答率

平均点の一種として、5択を5点から1点に割り振ったうえでの単純平均点ではなく、「ポジティブ回答率」、すなわち「肯定的な回答をした人の割合」を用いることもできます。5択の上位2つを「1」に、下位3つを「0」に置き換えたうえで平均するのです。

ポジティブ回答率には、意味合いがわかりやすいというメリットがあります。たとえば、調査結果を社内に広く、あるいはCSRレポートを通じて社外に公開する場合において、「『当社で長く働きたいと思っている』という設問の平均点は4.1点でした」というよりも、「肯定的な回答をした社員の割合は79%でした」というほうが、意味合いが伝わりやすいといえます。

また、設問の内容によっては、ポジティブ回答率のほうが意味の解釈のために適しています。「経験を積むため他社に出向してみたい」という設問であれば、「平均点は3.4点」というよりも、「肯定的／前向きな回答をした社員の割合は63%」というほうがわかりやすいといえます。1か0かを

平均する割合指標には、誤差の範囲をすぐ導ける統計分析上のメリットもあります。

こうしたメリットがある反面、ポジティブ回答率には、5択を2値に変換する過程でデータの情報量が落ち、また、「5段階で回答したのに、結局はポジティブ回答率を算出するためにしか用いられないのだ」という認識が、回答者の間に広まってしまった場合には、率直な回答が得られなくなるリスクもあります。したがって、原則として単純平均点を用いることが妥当であるというのが本稿の考えです。

(4) 属性別集計とそのクロス集計

集計単位としては、全体集計、属性別集計（年齢層別等）に加え、複数の属性を組み合わせたクロス集計（年齢別×職種別、年齢別×役職別等）を行います。3つ以上の属性のクロス集計（年齢別×職種別×男女別等）まで行うこともあります。

属性のすべての組み合わせで集計すると、何百という数表やグラフができ上がってしまうこともあります。そのすべてに意味があるとはかぎらないため、「その属性での集計は重要か」、「属性を組み合わせても十分な母数が確保できるかどうか」の2つの観点から、どこまでクロス集計を行うかを判断します。属性を組み合わせた結果、母数が3人を割り込む場合に

は集計は行わないか、属性区分を統合（「65歳以上」を「60歳以上」に統合する等）するべきでしょう。

「どのような属性の組み合わせで集計することを優先すべきか」は、「どのような視点で組織をみることが重要か」ということを反映します。たとえば、「男女での役割分担色が強い」、かつ「終身雇用・年功序列色が強い」組織であれば、男女別、年齢別、そして、男女別×年齢別の枠組みで集計・分析することが優先されます。

あるいは、「職種別に求める人材像が異なり」、かつ「事業部門別に顧客業界に沿ったカルチャーをもつ」組織であれば、職種別、事業部門別、そして、職種別×事業部門別の枠組みで集計・分析することが優先されます。

Step3: 偏差値化する

平均点は、最も有用な指標ですが、1つだけ難点があります。それは、単なる平均点の高低によっては、強み／弱みを判断することができないということです。

強み／弱みを判断するには、母集団全体の中でどのあたりに位置するかということを見る必要があります。そのための方法が、「標準化」、または「偏差値化」という方法です。統計における標準化は、業務プロセスにおける標準化とは意味が違い、「特定の値が、

平均点を中心としたバラツキの幅の中で、どこに位置するか」を指標化することを指します。

偏差値を出すためには、平均点とともに、回答のバラツキの幅を示す「標準偏差」という値を、あらかじめ出しておかなければなりません。そして、「個別の点数と平均点との差」を「標準偏差」で割ることで、回答母集団の中における位置を示す点数である「標準得点」を導き、それに（わかりやすい点数にするために）10を掛けて50を足してやることで、50を平均点とする二桁の数値、いわゆる「偏差値」が導かれます。

平均点を主に用いながらも、強み／弱みを議論する際には偏差値を用いることがより妥当です。たとえば、社員意識調査において部門別の特徴や強み／弱みを明らかにしたい場合、図表2に示すように、単純に平均点だけでみると、装置部門の弱みは「革新的な社風」となります。しかし、全体の回答分布の中での偏差値に着目するならば、弱みは「部門間の円滑な情報流通」となります。これこそが、組織内における装置部門の特徴といえるのです。

強み／弱みは偏差値によって語ることが妥当であり、とりわけ個人に焦点を当てた多面評価のフィードバックにおいては、偏差値を用いた強み／弱みの提示が効果的になります。