

分散する生情報を使えるデータに変える

データ加工・ 集計サービス

grooveworks.co.jp

株式会社グループワークス
神奈川県川崎市高津区久本3-3-2
溝ノ口第一生命ビルディング 4F
Tel 044-271-6043 / fax 044-271-6063

はじめに

アンケートをしたけど、入力する時間がなくて・・・。

データはあるけど、どうやって集計したらよいのか・・・。

手元にあるリサーチ情報を活用するためには、まずはデータの入力や変換、クリーニングといった「下ごしらえ」が必要です。

本資料では、そうしたデータの前処理とデータを活用するための集計方法についてご紹介いたします。

<主な内容>

- データ活用に不可欠なデータの前処理と集計
- データの前処理～入力・変換・クリーニング
- データ集計
 - － 複雑なクロス集計
 - － ウェイトバック集計
 - － 有意差検定
- ご利用料金

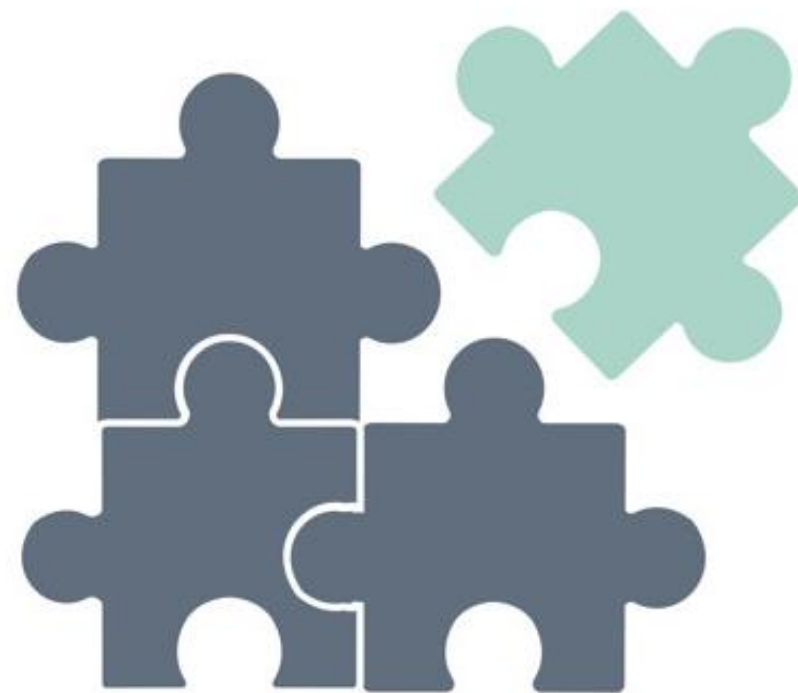
データ活用に不可欠なデータの前処理と集計

データ分析を行う際、元データをそのまま利用できるケースはまれで、多くの場合、集計・分析ツールに合わせてデータ形式を整える必要があります。

そうしたデータの前処理は、正しい分析のために不可欠な作業ですが、結構面倒で時間がかかります。

また、少し複雑な集計になると自社で対応するのは大変ですし、集計結果がでてきても、パワーポイント・チャート作成用にデータを加工したりする必要があります。

データ加工・集計サービスをご利用いただくことにより、**分析やプレゼン資料作成などのブレインワークに専念**していただくことができます。



データの前処理～入力・変換・クリーニング

調査の対象者による自記式アンケートでは、記入内容が判読しづらい回答や無回答など、機械的な処理では対応できないケースが少なからず発生します。

◆ あなたが外食したいと思うのはどのような時ですか。(いくつでも)

1. 誕生日など特別な日 4. 特別に食べたいものがある時 7. 新しいお店がオープンした時
2. 給料日・ボーナス日 5. 料理するのが面倒・時間がない時 8. その他
3. 家族や友人が集まる時 6. クーポンや割引がある時 9. どれもない

◆ あなたが外食するお店を選ぶ際に最も重視する点を教えてください。(1つだけ)

1. 料理の味 4. お店の雰囲気 7. 駐車場有 10. 座敷有 13. 人気店
2. 価格 5. 店員の対応 8. クレジットカード可 11. 個室有 14. その他
3. メニューの種類 6. 交通の便 9. 子連れ可 12. 営業時間 15. ほとんど外食しない

どちらの番号に○を付けているのか判別できない

単一回答のところ、複数回答されている

また、回答者の条件が設定されている質問に対して条件に該当しない人が回答していたり、回答データ間の矛盾があったりするなどの問題にも直面します。

誤回答や無回答の処理、論理矛盾のチェックおよびクリーニング、複数のデータの回答状況を組み合わせた**分析用変数の作成**など、生データからできるだけ多くの、できるだけ精度の高い情報を引き出し、データの活用の幅を広げることができるよう、データの前処理を行います。

データ集計～複雑なクロス集計

複雑なクロス集計

性別、年代別や地域別などのシンプルなクロス集計でしたら、Excelでも、ピボットテーブルを使って簡単に集計することができます。

しかしながら、地域×性×年代といった多重クロスや、集計ベースを複雑な条件で絞り込む必要があったり、集計項目数が多かったりすると、**一般的な表計算ソフトで対応するのは大変**で、分析の入り口にすぎない集計段階で、手間や時間がかかってしまい、なかなか先に進むことができません。

データ加工・集計サービスでは、

専用の集計プログラムにより、
複雑な集計のご要望にお応えします。

分析者の視点に立った集計軸のご提案や
表側設定など、タテ・ヨコ・ナナメ自在の集計
により、**リサーチデータの可能性を
最大限に引き出します。**

地域別⇒	東京			
性別⇒	男性			
年代別⇒	20～29歳	30～39歳	40～49歳	50～59歳
回答ベース：	(50) %	(50) %	(50) %	(50) %
設問1				
○○○○	xx	xx	xx	xx
△△△△	xx	xx	xx	xx
□□□□	xx	xx	xx	xx

データ集計～ウェイトバック集計

ウェイトバック集計

アンケートで回収したサンプルの構成比が実際の市場(母集団)と異なる場合、**サンプルに重みづけをして母集団と構成比を合わせた形で集計**することがあります。
(ウェイトバック集計)

年代区分	母集団	母集団(%)	回収 サンプル数	ウェイトバック後 のサンプル数	ウェイト値
20～29歳	123,777	15.5	100	77	0.77
30～39歳	156,070	19.5	100	98	0.98
40～49歳	183,950	23.0	100	115	1.15
50～59歳	154,455	19.3	100	97	0.97
60～69歳	180,989	22.6	100	113	1.13
合計	799,241	100	500	500	

※ 例えば、調査対象の20～69歳について
母集団の年代別構成比がわかっている場合、
左表のようにウェイト値を算出して
ウェイトバック集計を行うことができます。

一般的な表計算ソフトでウェイトバック集計を行おうとする場合、集計のたびに忘れずにウェイトをかける必要があったりして結構大変です。

データ加工・集計サービスでは、

ウェイトバック集計はもちろん、**面倒なウェイトの算出**も行います。

データ集計～有意差検定

有意差検定

ある調査で、20代女性における自社商品の評価が競合他社より10ポイント高かったとしても、当該セグメントのサンプル数が少なかったりすれば、必ずしも統計的に有意な差があるとはいえません。

特に回答ベースが小さい場合などは、見た目の評価差だけで判断するのではなく、**有意差検定による確認**が求められるところです。

しかしながら、検定の対象が多くなると、一般的な表計算ソフトで有意差検定を行うのは非常に大変な作業になります。

データ加工・集計サービスでは、

専門の統計ツールを使って**複数のセルにまたがる検定**を含め、多様な集計が可能です。

<検定つき集計のアウトプット例>

	X社	Y社	Z社
	(190)	(410)	(273)
回答ベース:	A	B	C
商品の総合評価			
(5) 非常によい	23.2% BC	6.5%	7.1%
(4) ややよい	35.3% BC	21.2%	25.7%
(3) ふつう	31.1%	48.6% A	46.1% A
(2) あまりよくない	6.8%	10.8%	12.3%
(1) まったくよくない	3.7%	12.9% A	8.9% A
平均値	3.67 BC	2.98	3.10

ご利用料金

データ加工・集計に関する各サービスの料金は以下の通りです。

	料金
クロス集計	3万円～
【追加オプション】	
ウェイトバック集計	2万円～
有意差検定	3万円～
データ加工	別途お見積り
その他のご要望	応相談

※ クロス集計は分析軸の数や内容によっても料金が変わります。

データの入力や分析のご相談もお受けしています。

グルーブワークスについて

アンケート／市場調査ならグルーブワークスにおまかせください。

アンケートやデータ分析に関するあらゆる業務に精通したプロのリサーチャーが、調査票の作成からデータの集計・分析までワンストップでサポートします。



アンケート作成



データ入力



集計



分析



自由回答



レポート

- この聞き方で大丈夫？
- アンケートがうまく作れない

- 人手も時間も足りない！
- 回答データの入力が面倒

- 性別などのクロス集計がしたい。でも難しい

- もっと高度な分析もできるはず。でもやり方がわからない

- たくさんありすぎて読み切れない。誰かまとめてほしい

- パッと見てわかるようなグラフにしたい。でも時間がない

こちらからお問い合わせください。



044-271-6043



info@grooveworks.co.jp

営業時間 9:30 - 17:30 [土日・祝日除く]

株式会社グルーブワークス <https://grooveworks.co.jp>

〒213-0011 神奈川県川崎市高津区久本3-3-2 溝ノ口第一生命ビルディング4階

プライバシーマーク登録番号：第10824102(05)号