

## ■コンテスト用データの見方(2020 年)

### 1. データの全体像

- データは大きく、1) アンケートデータ、2) 出稿データの2種類です。
- 1) アンケートデータはシングルソース形式で、様々な調査を行った結果が、分割して整理されています。
- アンケート調査は、関東1都6県を対象に、男女20～59歳に行った結果で、性・年代別で関東地方の人口の縮図になる形でデータを整理しています。
- 今年度のデータでは**2,500 サンプル**をご提供しています。
- 1) アンケートデータは、いずれも[SampleID]をキーとしてマージすることができます(異なるファイルであっても、[SampleID]が同じレコードは、同じ回答者が答えたレコードです)。
- 2) 出稿データは、テレビCM、雑誌広告、新聞広告について、どの番組(雑誌)に、どのような商品のCM(広告)があったのかを整理したものです。
- .

#### データの全体(お送りしたファイルの全体像)

大分類	ファイル名称	データ形式	データの内容
アンケートデータ	メインデータ	CSV/SPSS	個人属性、チャネル利用、商品別の購入実態・購入意向など
	テレビ番組別視聴状況	CSV/SPSS	2020年1月25日～4月4日のテレビ番組(地上波)の視聴の有無
	雑誌閱讀状況	CSV/SPSS	調査対象の雑誌(656誌)の閱讀の有無
	WEB利用状況	CSV/SPSS	利用しているWEBサービス、利用頻度、利用デバイスなど
	データ定義	EXCEL	上記の各変数の定義説明用
出稿データ	テレビ出稿	CSV	テレビ番組別のテレビCM出稿データ
	雑誌・新聞出稿	EXCEL	雑誌別の雑誌広告出稿データ 新聞別の新聞広告出稿データ

## 2. アンケートデータの見方

- SPSS 形式のファイルは変数ラベル、値ラベルが設定済みです。
- CSV、テキスト形式の場合、以下のようにデータが格納されています。

(1行目には変数名が入力されています)

SampleID	変数1	変数2	変数3	...
1	各サンプルの回答			
2				
3				

- なお、SampleID は1からの通し番号ではなく、欠落している番号もありますのでご注意ください。
- 変数名と回答に記載されている値の意味については、「データ定義」のファイルを参照してください。
- また、「データ定義」ファイルには、全ての変数およびその説明が記載されていますので、どのファイルにどのような変数が格納されているかを確認する場合にも利用することができます。

## 3. 「データ定義」ファイルの見方

- 各シートは以下の形式で記載されています。
- (各シート名は、アンケートデータのファイル名と同じです)。

項目	説明
変数	対応するファイル上の変数名(1番上の行)を示しています
ラベル	変数の内容が記載されています
値の定義	対応するファイル上の値(2番目以降の行に記載)が、それぞれどのような意味なのかが記載されています
備考	各変数に関する補足情報が記載されています

## 4. 出稿データの見方

### 1) 「テレビ出稿」ファイル

- ・ 番組 ID、広告主（企業名）、商品名（アイテム名）、CM の秒数のデータが格納されています。
- ・ 各番組（番組 ID）に、どのような会社・商品のテレビ CM の出稿があったのかがわかります。
- ・ 大きなデータになりますので、データ読み込みの際にはご注意ください。
- ・ 番組 ID は、アンケートデータのテレビ番組別視聴状況の変数名・番組 ID と共通のものを使っています。
- ・ したがって、ある商品 A のテレビ CM がどの番組で放送されたかについては「テレビ出稿」ファイルで確認でき、その番組の視聴状況は、アンケートデータの「テレビ番組別視聴状況」で確認できます。

### 2) 「雑誌・新聞出稿」ファイル

#### 【雑誌】

- ・ 雑誌コード、雑誌名、広告主（企業名）、商品名（アイテム名）、ページ（掲載しているページ番号）、色（カラー or 白黒）、大きさ（掲載ページ枚数）、広告の種類（純広・ペイパブ）のデータが格納されています。
- ・ 各雑誌に、どのような企業・商品の広告が、どのような形（掲載面、色、大きさ、純広・ペイパブの別）で載ったのかを整理したものです。

#### 【新聞】

- ・ 新聞コード、新聞名、刊行日、広告主（企業名）、商品（アイテム名）、ページ（テレビ欄かそれ以外か）、色、大きさ（段数）のデータが格納されています。
- ・ 各新聞に、どのような企業・商品の広告が、どのような形（掲載面、色、大きさ、純広・ペイパブの別）で載ったのかを整理したものです。

以上