
DATAEXPLORING

(データエクスプローリング)

業務案内
(2019年01月現在)

代表 出口慎二
shinji.deguchi@dataexploring.com

事業紹介

■ DATAEXPLORING(データエクスプローリング)について

- ◆ DATAEXPLORING(データエクスプローリング)は、個人事業の屋号です。代表者が、2001年より、データ分析およびその周辺業務を個人で受託していました。

■ 代表者について

- ◆ 出口慎二(でぐちしんじ):1972年生まれ。1997年、統計分析サービスを行う会社に入社。調査データの入力・集計から多変量解析による分析、関連するプログラミング業務などに携わる。2001年以降、個人でデータ分析事業を行なう。
- ◆ 2003～2004年、IRJ(インターネットリサーチ研究会)にて会員社合同の実験調査プロジェクトの統括ディレクターを務める。2004年、『自分でできるネットリサーチ』を3人共著で上梓。2005年～2007年、インフォプラント(現、マクロミル)にてRQI(Research Quality Institute)を設立し主任研究員を務める。2005年～2007年度、サーベイ・メソドロジー研究会(日本行動計量学会の研究部会のひとつ)に参加、インターネットリサーチを含むデータ収集法の研究に携わる。
- ◆ おもにデータの分析業務を行っていました(ALSCALやSMACOFなど2相3元データの分析、値が正規分布する表に対する対応分析、対応分析を用いたテキストマイニング、大規模データのクラスタリングなど)

おもな経歴・著書・発表論文等

1997年, 埼玉大学教養学部卒業.

1999年, 統計法第10条による統計官・統計主事になるための資格(現, (財)実務教育研究所認定統計士).

2000年, (財)社会通信教育協会認定生涯学習2級インストラクタ(統計).

2002年, (財)実務教育研究所認定データ解析士.

2005年, 立教大学社会学研究科応用社会学専攻(現, 経営学研究科)博士前期課程修了(修士(社会学)).

- ◆ 中山厚穂・出口慎二・岡太彬訓(2004). 百貨店における購買行動分析:選好度の写像分析による購買傾向の時間的変化. 日本オペレーションズ・リサーチ学会秋季研究発表会アブストラクト集(2004年度), 108-109.
- ◆ 平石郁生・出口慎二・森田英一(2004). 自分でできるネットリサーチ. ラトルズ.
- ◆ 中山厚穂・出口慎二・岡太彬訓(2005). 2相3元データの分析:INDSCALとPREFMAPを併用した方法とINDCLUSによる方法. 日本オペレーションズ・リサーチ学会秋季研究発表会アブストラクト集(2005年度), 206-207.
- ◆ 出口慎二(2005). ボランティアパネル型インターネットリサーチによる第44回衆院選直後調査結果報告. サーベイ・メソドロジー研究会:活動の記録, 研究会:2005年11月22日(abstract).
- ◆ 福田昌史・出口慎二(2006). インターネット調査の特性:面接ーネット並行調査. 第87回行動計量シンポジウム資料.
- ◆ 出口慎二(2008). インターネット調査の効用と課題. 行動計量学, 35(1). 47-57.
- ◆ 出口慎二・岡太彬訓・中山厚穂(2009). 2相3元データ分析に際してのINDSCAL, INDCLUS, 及びPREFMAP解併用の有効性ーケース研究. 行動計量学, 36(2). 105-114.
- ◆ 出口慎二. 東日本大震災ビッグデータワークショップの参加報告. (未刊行). (2012.10.29.公開)
- ◆ 出口慎二. 実データに見るO2O効果のいち事例. (未刊行). (2013.12.27.公開)
- ◆ 出口慎二・中山厚穂・高崎祐哉(2014)大規模データのクラスタリング:クラスタ中心を再計算しない非階層的クラスタリング:ドラッグストアの会員分類を例として. 日本行動計量学会第42回大会発表スライド.
- ◆ 出口慎二(2015)ビッグデータの要約としてのMDS. 日本行動計量学会第43回大会発表スライド.
- ◆ 出口慎二(2015)日本のネット調査が抱える問題点:あえて今鳴らず警告. 日本行動計量学会第43回大会発表スライド.
- ◆ 中山厚穂, 出口慎二, 烏谷正彦(2016)クラスタ中心を再計算しない大規模データのための非階層的クラスタリング. 日本行動計量学会第44回大会発表スライド.

営業について

■ 基本情報

- ◆ 営業日: 月～金(祝祭日を除く)
- ◆ 営業時間: 9:00.～18:00.
- ◆ 夏季休業: 6～8月中に1週間
- ◆ 年末年始休業: 12/30～1/5

- その他: しばしば不在にしていることがあります。メールを送っていただけると確実です。外出時でもできるかぎりメールは確認します。

■ 費用について

- ◆ 以下を基本とします。
 - 分析I(対応分析, 多次元尺度構成法, クラスタ分析, 因子分析)
 - ▶ 分析1件につき, 4万円+税を基本とします(訪問しての打ち合わせは含みません。ご来所いただいたの打ち合わせは含みます。電話やメール等による簡単なやりとりを含みます)。分析出力一式(数表およびグラフ)のほか, 納品時に出力についてのコメントを文書にて添えますが, 別途報告書は作成しません。
 - 分析II
 - ▶ その他あるいは組み合わせでの分析は1人日4万円+税を目安に算定します。
 - 複数の分析を同じデータに対して行う場合
 - ▶ 同データに対して行う2件目以降の分析はこの75%を基本とします(訪問しての打ち合わせは含みません)。同時に行う分析でも, データが異なれば別の分析として扱います。
 - その他
 - ▶ その他の分析等については1人日4万円+税を目安に算定します。

分析の例

■ 対応分析

- ◆ 対象者を分ける場合は数量化III類を行います。
- ◆ 多重対応分析では数量化III類を勧める場合もあります。
- ◆ 布置の解釈に角度と距離の散布図を利用する場合があります。
- ◆ 行要素と列要素の同時布置も行います。（図の解釈に少々の注意があります）
- ◆ 基本的に、第1軸と第2軸の2次元を重視します。

■ 多次元尺度構成法

- ◆ 2相3元データに対して, ALSCAL (SPSS)あるいはSMACOF (R)を行います.
- ◆ RでSMACOFを行う場合, Package smacof を使います (重み布置の値に癖があります).
- ◆ 1相2元データに対するALSCALなども必要に応じ行います.
- ◆ 必要に応じ, 古典的方法も行います.

■ クラスタ分析

- ◆ 「分析I」(定額)では, 階層的方法(Ward法)を行います. 変数の数にもよりますが, 5,000SS程度までは対応可能です.
- ◆ より大きなデータの場合(億以上のデータ, 200列を超えるデータ), 複数のサンプリングデータセットに対する階層的方法の繰り返しから, クラスタ数とクラスタ平均を決めた後, これに基づいて非階層的なクラスタリングを行います. この場合, 「分析II」扱い(作業手数に応じた費用)となります.
- ◆ 非階層的なクラスタリングを行う場合, データ件数により, クラスタ中心の再計算を行う場合(k-means法)と行わない場合があります. クラスタ中心の再計算を行わない場合は, より大きなデータの処理が可能です.

■ 因子分析

- ◆ 「分析I」(定額)の場合, 直交回転のみとなります.
- ◆ 直交回転の場合でも, 変数選択の検討などを行う場合, 「分析II」(作業手数に応じた費用)扱いとなります.
- ◆ 斜交回転は「分析II」(作業手数に応じた費用)となります.
- ◆ 因子得点を用いたクラスタリングは, あまり行っていません. あるいはSEM(共分散構造分析)を勧める場合もあります.

■ 決定木 (Decision Tree)

- ◆ 交互作用はみません. 詳しくは下記ご覧ください.

http://dataexploring.com/info/docs/report_nri2013.pdf

- ◆ 木の構造そのものを探索したい場合不向きです. 影響を持つであろう変数を探す場合は, 通常のAIDよりも見やすい結果を与えることが期待できます.

■ テキストマイニング

- ◆ 茶釜を利用した形態素解析＋対応分析を用いた布置によるテキストマイニングを行います(「分析II」).
 - 形態素解析では, 標準辞書を使いますが, 「未知語」も見するため, 標準辞書にない語も多くの場合, 拾うことができます.
 - 対応分析では, 絶対寄与/相対寄与を見ることで, 次元の構成に大きく影響を果たした語を見付けやすくなっています.
 - 品詞情報は使いません(簡易的な分析になります).

Web について / 事業所案内

- この案内に記載の内容は、すべてホームページ上に記載があります。また、出力サンプルなど、この案内資料よりも多くの情報があります。あわせてご覧ください。

<http://dataexploring.com/>

■ 事業所案内

DATAEXPLORING

代表 出口慎二

データエクスプローリング
<http://dataexploring.com/>
〒180-0006 東京都武蔵野市中町 1-12-10-701
TEL 090-1110-5444
E-mail shinii.deguchi@dataexploring.com

