

卒論で使える！ 効果的なデータの示し方

東京外国語大学附属図書館 学習相談デスク

永井哲平

自己紹介

名前

永井哲平（ながいてっぺい）

所属

大学院博士課程国際社会専攻

専門分野

経済学 -> 教育開発・計量経済学

専門地域

南・西アジア

受付時間

月曜日 15:00-19:00 金曜日 17:00-19:00

2017年卒論ガイダンス

この発表の狙いと流れ

1. 論文・レポートにおける主張の根拠（データ）の重要性
2. 論文・レポートを執筆する際の、主張の根拠（データ）の示し方を解説
3. 主張の根拠となる客観的事実の提示方法を紹介

論文執筆の大まかな流れ

集める



読む



先行研究を
まとめる

↑先行研究

自分の考え↓

自分の意見を
まとめる



分析する

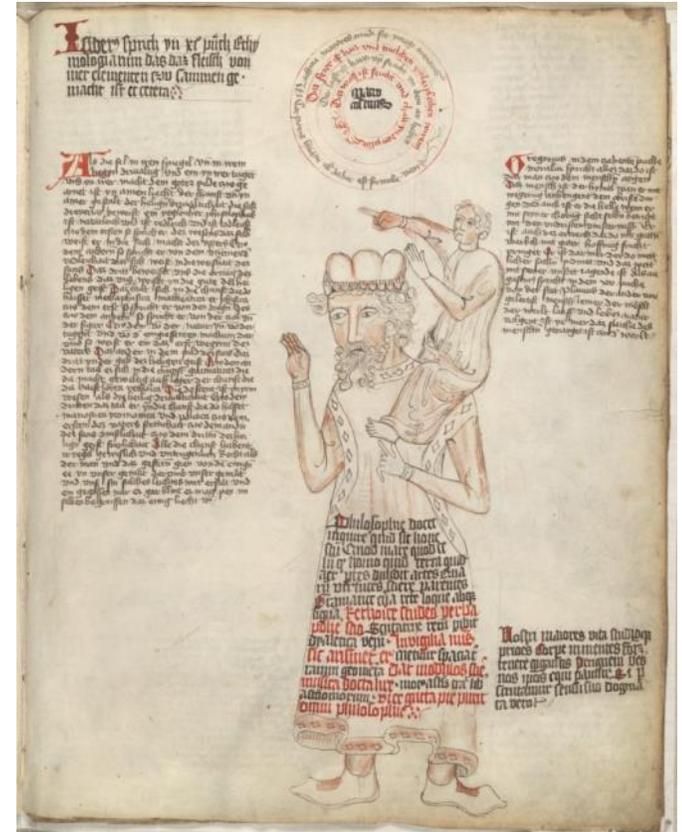


調べる

先行研究の重要性

- 既存の研究を精査してからの分析と実証
- 先行研究をいかに整理できるか
- 論文・レポート全体の完成度にかかわる

“Standing on the shoulder of giants”



先行研究のまとめ方

- 先行研究の情報の記載
執筆者、出版年、タイトルなど
- 興味のある（使えそうな）箇所を中心にまとめる
- 印刷した論文に線引きなどをして。
オレンジ：使える箇所
ピンク色：論点
黄色：用語の定義



先行研究をまとめる

先行研究の意見をまとめる

同意できる点

「（奨学金が）学生を逆に追い詰めている」

「格差が広がる事態は避けねばならない」

批判点

「給付型（奨学金）の検討を急ぐべき」

新たな視点

「幼児教育の段階的な無償化」

小中学生が対象の「就学援助」

先行研究をまとめる

先行研究を要約・批判する

朝日新聞(2016.4.15)によると、奨学金が学生を追い詰めており、貧富の差が広がる事態を避けなければならないと報じている。そのためには返済の不要な給付型奨学金制度の導入を急ぐべきであり、幼児教育の段階的な無償化や小中学生が対象とした就学援助も検討されるべきだが、奨学金の返済に追われている学生の実態を示してはいない。

主張の根拠を示す

例題

死刑制度は日本においても撤廃されるだろうか。近年では、死刑制度を根本的に見直す動きが世界で広がっている。しかし、国内世論では死刑制度が多くの支持を受けていることから、日本において死刑が即時撤廃される可能性は低い。

客観的なデータ（主張の根拠）の重要性

死刑制度は日本においても撤廃されるだろうか。近年では、死刑制度を根本的に見直す動きが世界で広がっている。しかし、国内世論では死刑制度が多くの支持を受けていることから、日本において死刑が即時撤廃される可能性は低い。

論文の重要な構成要素

問題提起

自分の主張

主張の根拠

客観的な事実

客観的なデータ（主張の根拠）の重要性

死刑制度は日本においても撤廃されるだろうか。近年では、死刑制度を根本的に見直す動きが世界で広がっている。しかし、国内世論では死刑制度が多くの支持を受けていることから、日本において死刑が即時撤廃される可能性は低い。



データセクションの書き方

- どのような調査・分析を行ったかを紹介する
使ったデータをどのように収集したかを示す。使用したデータが一次データなのか、二次データなのかを明記する。一次データの場合、方法を示す。
 - データの収集方法：横断面的、遡及的（時系列的）、予測的
 - 被験者を、どのように治療群と対照群に割り当てたか
 - データの出典：質問表、インタビュー、各種調査
 - データに与える潜在的な偏り
 - 調査内容と方法のプラス面とマイナス面

データの妥当性・適合性

- 妥当性

妥当性を評価するためには、扱うトピックとデータをよく理解しなければならぬ。関係性の薄い調査結果を持ち出すことは避けなければいけない。例えば、高齢者をサンプルとした意見や嗜好は、若年層や子供のそれとは異なる。

- 適合性

対比する数値を選択する前に、その分野で用いられる慣習的な標準、比較値について調べる

- 比較可能性

背景となる例では類似の文脈から得たデータを示す。分析したい次元だけを変化させ、2017年卒論ガイダンス他の次元はそのままにする。

根拠を表す悪い例

1. アメリカの発電量のランクは1である。
2. 1990年、南部と中西部の人口比は1.43であった。
3. 2000年のアメリカ連邦政府の財政支出は61%が人的資源、12%が利息、16%が国家防衛、6%がその他の活動、5%が物的資源である。

4つの「W」

When: いつ

Where: どこ

Who: だれ

What: なにか

4つの「W」

When: いつのデータか

データを1つの時点に関連付けるか、幾つかの時点にまたがって比較するか、一定期間に渡って追跡するか明記し、適切な日付と時間の単位を示す

- 調査期間は2017年2月17日から2月28日までの12日間で、配布数は216票であった。（脇坂ほか 2017）
- 国土交通省河川局が作成している年間水害被害額と死者数の推移から、近代以降の全国的な水害発生動向を把握した。ここで、水害被害額は1875年から直近の2010年まで、死者数は1902年から2010年までの期間を対象とし（谷淵 2017）

4つの「W」

Where: どのデータか

データをどこから採取したのか示す。地理的単位(市、国、大陸など)または組織(学校、研究機関、企業)など。地理に馴染みのない読者のために地図を用意すると良い。

- A社（本社：東京都）の従業員を対象に、アンケート調査を実施した。A社は、情報システムの開発、製造、販売を主たる事業内容とし、従業員数5,000名を超える企業である。(古田 2017)

4つの「W」

Who: 誰から得られたデータか

A) 全事例かサンプルか？

特定の時間と場所におけるすべての事例ではなく、一部のみを切り出して対象とすることもあります。限定された分析サンプルかどうかを明記し、調査内容が研究に適しているかを説明する。

- 調査協力者は日本語を専攻する一年生59人である。うち31人は後述するように事前テストで実験用動詞の知識を要することが確認されたため（中略）大学入学後に日本語学習を開始しており、調査時点での学習期間は6ヶ月であった。(姚, 菅谷 2017)

4つの「W」

Who: 誰から得られたデータか

B)回答率

調査対象となるすべての事例からデータが得られるとは限らない。アンケート調査などでは回答率を示す。

- 保護者は82名中、41名から回答が得られ、回収率は50.0%である。また、寄宿舍指導員は52名中、37名から回答が得られた。そのうち未記入の項目があった回答を除いた結果、有効回答数35名（67.3%）の回答を使用して集計した。回収率は71.2%である。（細谷、須佐 2017）

4つの「W」

What: 何を示しているか

複数の情報源を用いる場合、どの変数がどの情報源から来ているのかを示す。また、連続型変数には単位を示し、順序値には対応する数値を基準を示す。

- 人口属性と個人意識については質問表に基づいており、健康指標については医療記録に基づいている。
- 『高所得』『低所得』は所得分布の最上位/最下位の五分位数に相当し、『中所得』は中間の3つからなる。

データの表現

基準値の選択

- 標準と慣習の使用

例) 物価、インフレ/デフレを示す場合は消費者物価指数が用いられる。日本では2015年を、米国では1982-'84年の平均を基準としている。

例) 個々人の点数や身長を評価する際は、標準分布が用いられる。

例) 貧困境界値は、世帯所得を評価する際によく使われる。

データの表現

ランクを表す語法

数値、基準となるグループ、グループ内でのランクを理解しやすい用語で伝える

アメリカの発電量のランクは1である

アメリカの再生可能エネルギーの発電量のランクは1である

- **アメリカの再生可能エネルギーの発電量は14万7千トンであり、OECD加盟国で最も高い**

データの表現

相対差を表す語法

何を比較しているのか、どちらがより大きい/小さいのか、
をわかりやすく記述する

1990年、南部の分子は中西部の分母の1.43倍であった

1990年、南部と中西部の人口比は1.43であった

➤ **1990年、南部の人口は中西部の1.43倍であった**

データの表現

最頻値を表す書き方

最頻カテゴリ、順位、相対的割合を使用して、数値どうしを比較しやすく示す。

2000年のアメリカ連邦政府の財政支出は61%が人的資源、12%が利息、16%が国家防衛、6%がその他の活動、5%が物的資源である。

- 2000年のアメリカ連邦政府の財政支出の**60%以上**は人的資源に使われており、大きく離れて**第2位**は国家防衛(16%)、**第3位**は利息(12%)となっている。**残り**は物的資源とその他で**ほぼ均等**に分けられている。

データをどのように入手するか

- アンケート
- 公開されているデータベース
 - コーパスデータベース
 - ひまわり (国立国語研究所)
 - Japanese EFL Learner Corpus
 - BYU British National Corpus
 - 社会・経済統計データベース
 - e-Stat (総務省統計局)
 - OECD Statistics (経済協力開発機構)



データを読み解く

- 記述統計 Descriptive statistics
- 作図

学習相談デスクガイダンス『データの処理と表現方法』



付属図書館
TOPページ



学習相談デスク
TOPページ

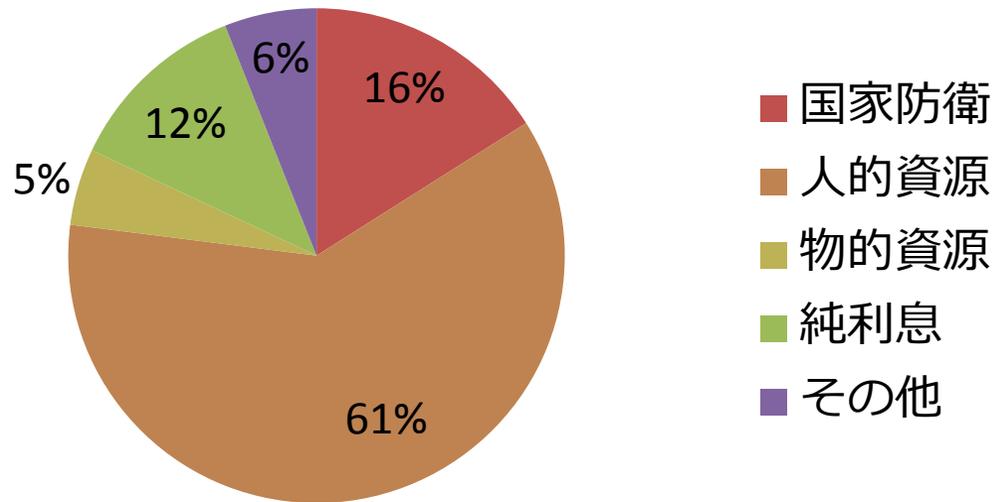


ガイダンスと
配りもの

図で表現する

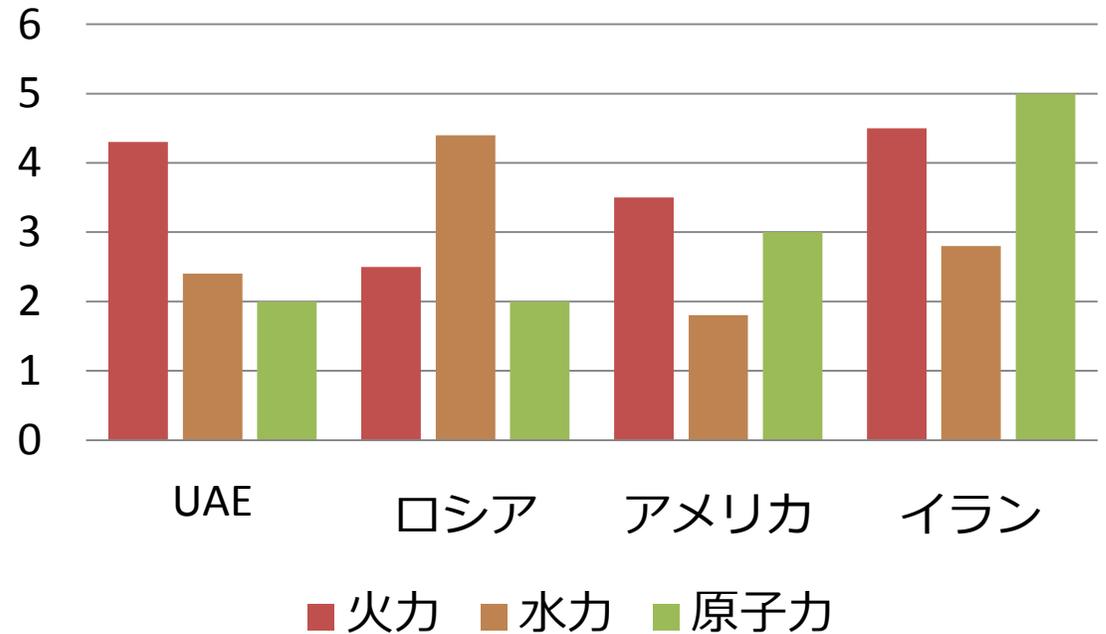
円グラフ

アメリカ連邦政府
合計支出：1.8兆ドル



棒グラフ

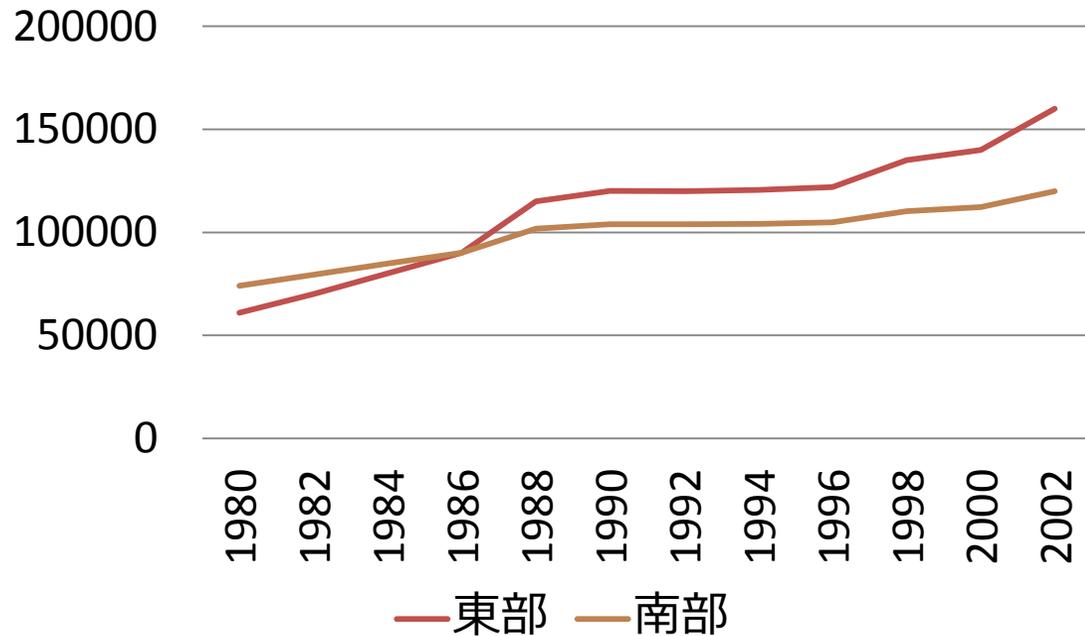
主要国のエネルギー源



図で表現する

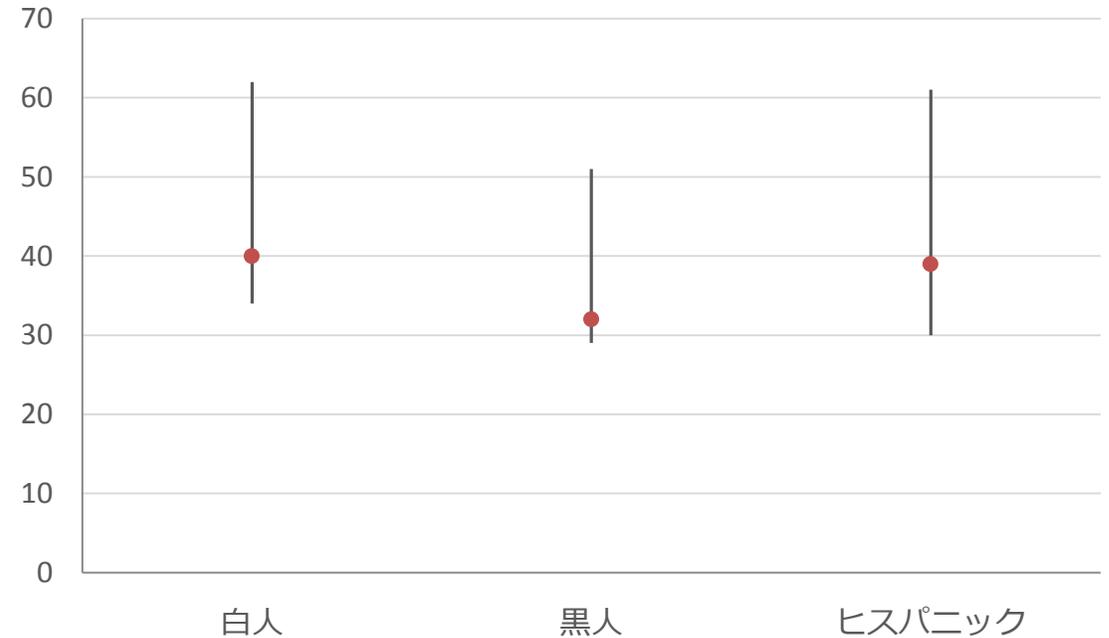
折れ線グラフ

住宅価格の変化



ハイ・ロー・クローズチャート

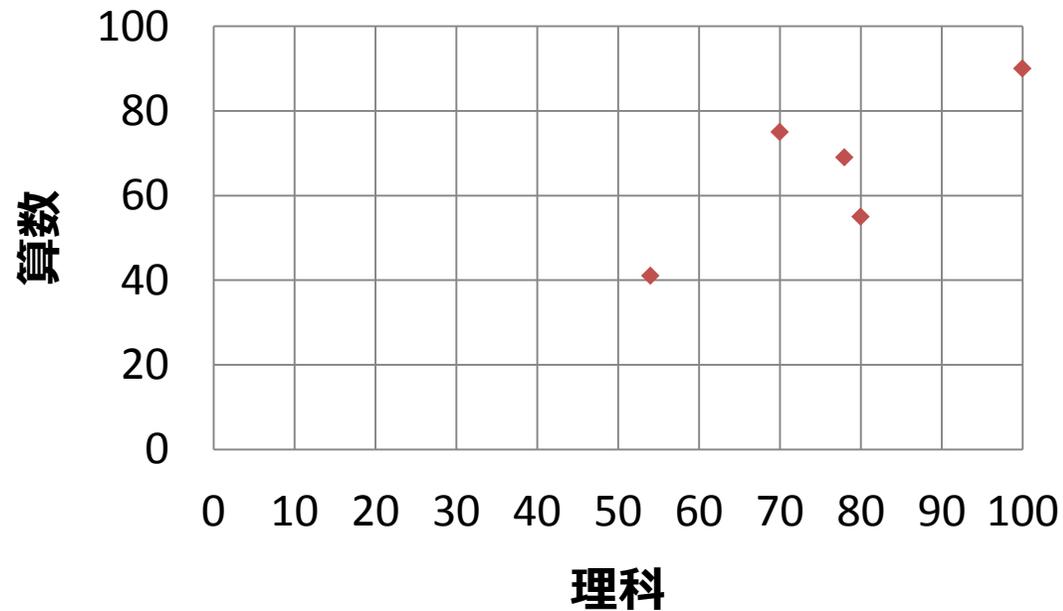
人種/民族別の社会経済指標



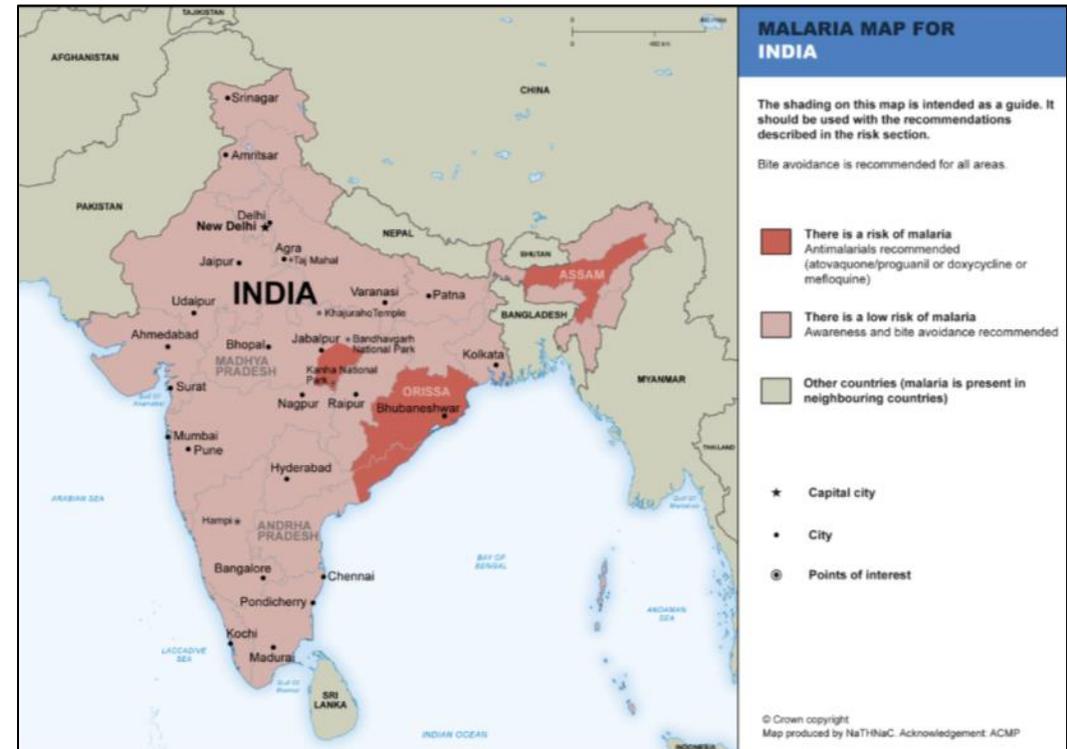
図で表現する

散布図

2科目のテストの関連

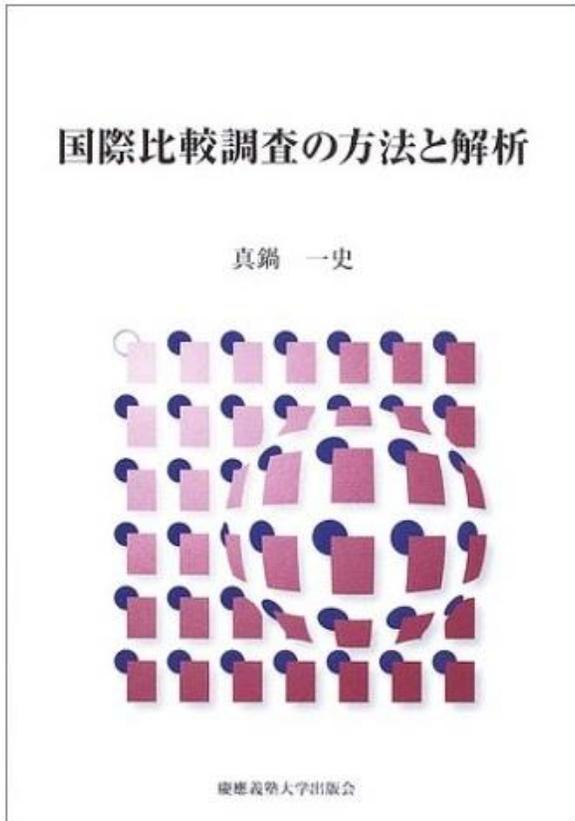


数値データの地図



出典：TRAVEL HEALTH PRO

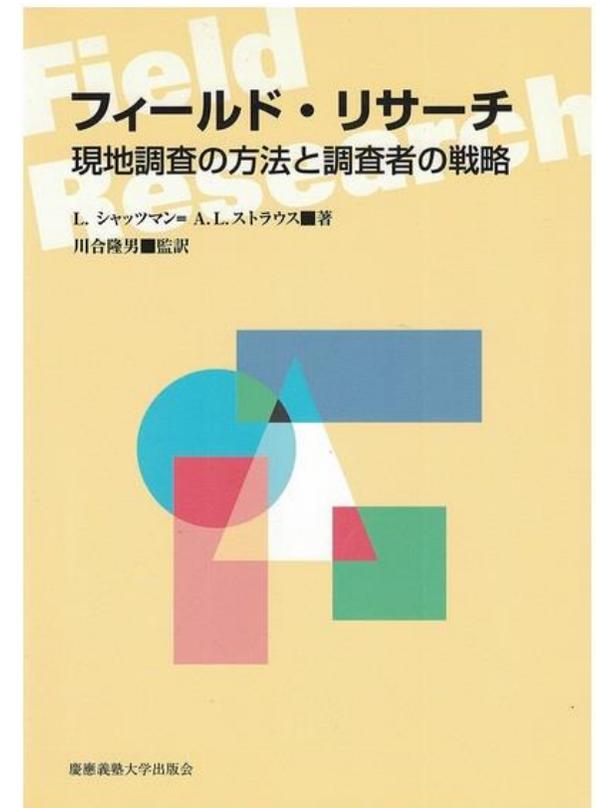
さらに、



3階閲覧室
A/361/529264



3階閲覧室
A/361/715174
2017年卒論ガイダンス



3階閲覧室
A/361/10397

参考文献

- Jane E. Miller著、長塚隆監訳、2006年「数を表現する技術 伝わるレポート・論文・プレゼンテーション」オーム社
- 望月源、2006年「コーパス言語学実践」授業資料
- 古田克利、2017年『職業観、仕事の意味深さおよび組織適応感の関係-組織で働くプロフェッショナルを対象にした定量分析-』関西外国語大学研究論集、vol.106
- 姚新宇、菅谷奈津恵、2017年『中国人日本語学習者による動詞コロケーションの習得-明示的帰納法と暗示的帰納法の比較を中心に-』国際文化研究、vol.23
- 細谷一博、須佐絵里加、2017年『北海道内の肢体不自由特別支援学校に併設されている寄宿舎における教育的役割: 寄宿舎指導員と保護者のアンケート調査から』北海道教育大学紀要（教育科学編）、vol. 68(1)
- 朝日新聞『社説』2016年4月15日朝刊、14
- 脇坂幸子、森詩恵、藤澤宏樹、橋本理、2017年『シルバー人材センターにおける「地域班」活動の現状と課題: 「地域班」の世話人に対するアンケート調査から』大阪経大論集、vol. 68(1): 49-58.
- 谷淵郷、2017年『昭和戦前期における水害の地域的特徴: 道府県別の水害に関する統計の分析』2017年度日本地理学会春季学術大会、筑波大学.
- National Travel Health Network and Centre “Malaria Map for India”
<<https://travelhealthpro.org.uk/admin/web/uploads/India-malaria-map.png>> 2017.10.17 accessed.