

データ分析力向上のための学習指導への示唆

— 統計データ分析コンペティション2021応募論文からの課題抽出 —

東京学芸大学(元 独立行政法人統計センター)
独立行政法人統計センター

山下 雅代
高田 聖治

SSDSE (教育用標準データセット)

データサイエンス教育のための汎用素材をエクセルとCSV形式で提供
欠測のない完備なデータセットで、**無料ですぐに分析可能**
分かりやすい解説書も付設

★2022年度から、SSDSEに日本語名称が追加

名称	内容
SSDSE-市区町村 (SSDSE-A) *1	1741市区町村 × 多分野125項目
SSDSE-県別推移 (SSDSE-B)	47都道府県 × 多分野107項目 × 12年次
SSDSE-家計消費 (SSDSE-C)	47都道府県 × 家計消費226項目
SSDSE-社会生活 (SSDSE-D) *2	47都道府県 × 社会生活119項目 × 男女別

※1 SSDSE-市区町村 (SSDSE-A) の2022年度版は6月末公開予定
※2 SSDSE-社会生活 (SSDSE-D) の次回更新は2023年の予定

★ファイルサイズが比較的小さい、使いやすいデータセットを新たに開発

名称	内容
SSDSE-基本素材 (SSDSE-E)	47都道府県 × 多分野90項目

★データセットの詳細、解説書の閲覧、ダウンロードは
右の二次元バーコードからアクセス



統計データ分析コンペティション

☆趣旨

SSDSEを用いた統計分析の論文を募集し、アイデアと解析力を競うことで、高校生、大学生等の統計リテラシーの向上を図る

☆賞及び副賞等

優秀作品には、**総務大臣賞、優秀賞、統計数理賞、統計活用奨励賞**と**各副賞**を授与、**受賞論文が月刊誌「統計」に掲載**

2022年度 実施予定概要

主催：総務省統計局・独立行政法人 統計センター

大学共同利用機関法人 情報・システム研究機構 統計数理研究所・一般財団法人 日本統計協会

後援：文部科学省・国立研究開発法人 科学技術振興機構 (JST)・全国高等学校長協会

一般社団法人 日本統計学会・全国統計教育研究協議会

募集部門：高校生の部………高校、高専 (1~3年次) の生徒

大学生・一般の部………短大、高専(4,5年次 専攻科)、大学生、大学院生、一般の府

主な日程：4月15日 (金) 開催発表

5月10日 (火) エントリー及び論文募集開始

8月10日 (水) エントリー 締切り

9月1日 (木) 論文 締切り (大学生・一般の部)

9月9日 (金) 論文 締切り (高校生の部)

10月18日 (火) 受賞論文 発表 (統計の日)

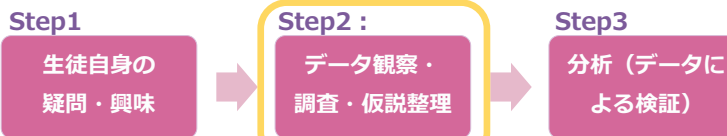
詳細は、
2次元コードに
アクセス



高校生のデータ分析力向上への課題

統計データ分析コンペティション2021 高校生の部への応募論文全55件から、データ分析力向上への課題を抽出

① 仮説生成にあたっての問題意識の持ち方



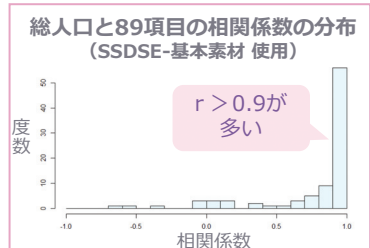
分析すべき仮説を探究すること

<仮説整理のためのツールの例>

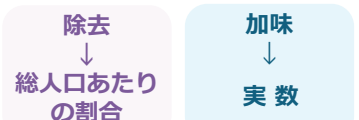


Step2が弱い論文が目立った

② 観察対象の比較では規模の大きさに注意



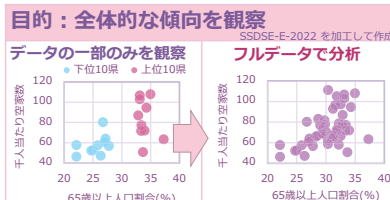
規模 (総人口) の影響を



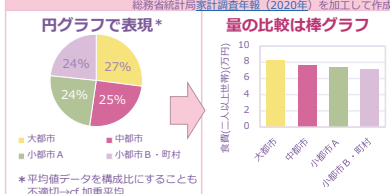
割合の解釈・活用が重要

③ グラフ描写の基本作法

<不適切なグラフの使用例>



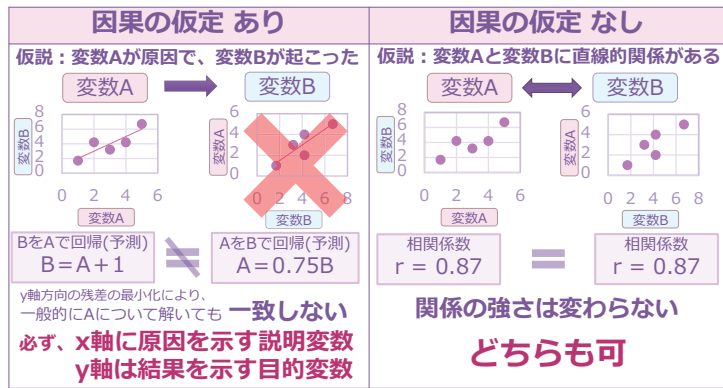
目的：都市の規模別の食費の差を比較



目的にあった分析方法・グラフを用いること

④ 散布図・相関分析・回帰分析と因果の関係

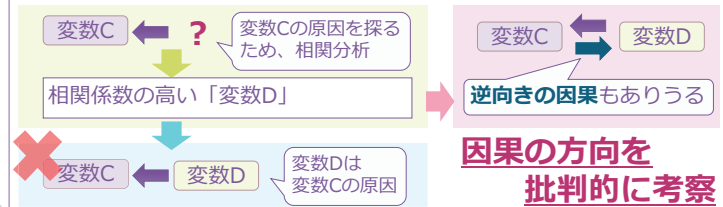
相関関係：一方が増加するとき、他方が増加または減少する傾向
因果関係：原因と結果の関係 (ex.人口→ゴミの量) (ex.空家数⇔ゴミの量)



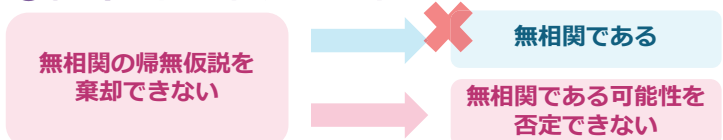
回帰分析 (回帰直線) : 因果の仮定が必須

相関分析 : 因果の仮定はどちらも可

相関分析と因果関係に関する考察



⑤ 統計的仮説検定の解釈：無相関検定の例



【指導の例】統計的仮説検定は、「アリバイ証明」
アリバイ成立 → 帰無仮説を棄却 = 犯人ではない
アリバイ不成立 → 必ずしも棄却できず = 犯人ではないとは言いきれない
これだけで「犯人だ」と言い切ることはできますか？

スパイラル学習を推奨