

【大学IR業務、教学DXにおけるデータ一元化の考慮点について】

出典元：令和6年9月9日 文部科学省 全国学生調査(本格実施)の実施方針より



文部科学省

参考

令和7(2025)年度より本格実施を開始します

全国学生調査

全国学生調査とは？
- 調査結果はどのように活用できます -

在学生の学修実態の把握
全国共通の質問項目により全国の大学との比較分析や各大学の教育改善に活用できます

好事例の収集・分析と自大学のアピール
質問項目ごとの結果上位校一覧と教育方法・教育改善のグッドプラクティス事例を公表します

教学IRにおける活用
調査結果をまとめた教学IRレポート等参加大学の教育改善に役立つ資料を提供します

調査の目的

急速な少子化の進展等、高等教育を取り巻く環境が大きく変動する中において、各大学は社会が期待する役割や求める人材像を自ら把握し意識しながら、各自の強み・特色を生かした教育研究活動について積極的に発信することにより、組織や立地、知名度等による入学者選抜の選抜性の高低によらず、教育研究の質の高度化に向けた取組について国際社会を含む外部から適切な評価を得ていくことがこれまで以上に求められています。このような状況を踏まえ、学修者本位の教育への転換を目指す取組の一環として、学修の主体である学生目線からの大学教育や学びの実態把握を通じて、以下①～④への活用を目的とする「全国学生調査」を実施します。

- ① 各大学**
在学生の实態や重層、他大学との比較分析を踏まえた教育改善に活用すること
- ② 学生**
これまでの学びを振り返り、今後の学修や大学生生活をより充実させ、卒業後の道路を考える上での一つの契機とすること
- ③ 大学進学希望者や地域社会等**
各大学における学生の学修成果や大学全体の教育成果に関心を持ってもらい、大学に対する理解を深めてもらうこと
- ④ 国**
今後の政策立案に際しての基礎資料として活用すること

IR業務用データ一元化について 基本的な考え方 ①

- データは分析における全ての業務のインプットであるため、その整備状況に応じて分析の精度等への影響を及ぼします。
- また、データは日々発生・蓄積されるものでもあり、データ整備にあたってのルール策定によって、分析業務の継続的な品質確保が期待されるものと考えます。
- なお、貴学リソースにて分析業務を運用するためにも、ルール策定にあたっては、貴学としての判断・決定が必要と認識します。

分析までの概略フロー

データ準備

・CSVファイルの準備/格納

整備状況が不十分
の場合…

分析にあたっての
データ結合

・BI分析用ソフト(Tableau、
Power BIなど)を利用した、
分析のためのデータ結合(分析
シナリオに基づくデータ結合)

結合漏れ等による
データの欠損

分析

・BI分析用ソフト(Tableau、Power
BIなど)を利用した、分析の実施(分
析シナリオに基づくダッシュボード、
グラフ等による可視化)

誤った意思決定の恐れ

IR業務用データ一元化について 基本的な考え方 ②

- データ整備にあたっては、「整然データ(Tidy data)」をルール策定のための基本的な考え方とします。

「整然データ」とは

「データ分析においてどのようなデータが扱いやすいのか？」という問いへの解答のひとつ（データ分析のすべての状況において扱いやすい唯一の解答は存在しない）であり、ニュージーランド出身の統計学者ハドリー・ウィッカム(Hadley Wickham)氏^{*1}が提唱。

*1) <https://ja.wikipedia.org/wiki/%E3%83%8F%E3%83%89%E3%83%AA%E3%83%BC%E3%83%BB%E3%82%A6%E3%82%A3%E3%83%83%E3%82%AB%E3%83%A0>

「整然データ」のルール

- ① ひとつの列はひとつの変数である
- ② ひとつの行はひとつの観測である
- ③ 観測の集まりが表(テーブル)となる
- ④ ひとつのセルにはひとつの値だけ

※整然でないデータのことを雑然データ(messy data)と呼ぶ。

① ひとつの列はひとつの変数である

整然データ

変数		
年月	地区	天気
2023/1/1	東京	晴れ
2023/1/1	名古屋	曇
2023/1/1	大阪	雨
2023/2/1	東京	晴れ
2023/2/1	名古屋	曇
2023/2/1	大阪	晴れ
2023/3/1	東京	晴れ
2023/3/1	名古屋	雨
2023/3/1	大阪	曇

雑然データ

地区	2023/1/1	2023/2/1	2023/3/1
東京	晴れ	晴れ	晴れ
名古屋	曇	曇	雨
大阪	雨	晴れ	曇

見やすいのでレポートとしては良いが、ひとつの列がひとつの変数になっていない

② ひとつの行はひとつの観測である

整然データ

年月	地区	天気
2023/1/1	東京	晴れ
2023/1/1	名古屋	曇
2023/1/1	大阪	雨
2023/2/1	東京	晴れ
2023/2/1	名古屋	曇
2023/2/1	大阪	晴れ
2023/3/1	東京	晴れ
2023/3/1	名古屋	雨
2023/3/1	大阪	曇

雑然データ

地区	2023/1/1	2023/2/1	2023/3/1
東京	晴れ	晴れ	晴れ
名古屋	曇	曇	雨
大阪	雨	晴れ	曇

ひとつの行がひとつの観測になっていない

③ 観測の集まりが表(テーブル)となる

整然データ

年月	地区	天気
2023/1/1	東京	晴れ
2023/1/1	名古屋	曇
2023/1/1	大阪	雨
2023/2/1	東京	晴れ
2023/2/1	名古屋	曇
2023/2/1	大阪	晴れ
2023/3/1	東京	晴れ
2023/3/1	名古屋	雨
2023/3/1	大阪	曇

- ひとつの表には種類の違う観測が入ってはいはならない
- 同じ種類の観測が複数の表にまたがってはいはならない

④ ひとつのセルにはひとつの値だけ

雑然データ

年月	地区	天気
2023/1/1	東京、大阪	晴れ
2023/1/1	名古屋	曇

ひとつのセルに2つの値が入っている

(ご参考) 「Tidy data」
https://ja.wikipedia.org/wiki/Tidy_data