

Package ‘IRexamples’

January 20, 2025

Title Collection of Practical Institutional Research Examples and Tutorials

Version 0.0.4

Description Provides examples of code for analyzing data or accomplishing tasks that may be useful to institutional or educational researchers.

Depends R (>= 3.4.0)

Imports dplyr (>= 0.8.5), twang, survey, ggplot2, ggthemes, rstanarm, RSelenium, rvest, DisImpact, forcats, tidygeocoder, maps, geodist, sf, tigris, tidycensus

License GPL-3

URL <https://github.com/vinhdizzo/IRexamples>

BugReports <https://github.com/vinhdizzo/IRexamples/issues>

Encoding UTF-8

LazyData true

RoxygenNote 7.1.2

Suggests knitr, rmarkdown, markdown, prettydoc

VignetteBuilder knitr

NeedsCompilation no

Author Vinh Nguyen [aut, cre]

Maintainer Vinh Nguyen <nguyenvq714@gmail.com>

Repository CRAN

Date/Publication 2023-10-06 06:40:02 UTC

Contents

| | |
|---------------------------|---|
| ccc_list | 2 |
| survey_response | 2 |
| survey_results | 4 |
| uc_list | 5 |
| writing_center | 5 |

Index**8**

| | |
|----------|---|
| ccc_list | <i>California Community Colleges address data set</i> |
|----------|---|

Description

List of California Community Colleges (CCC) and their addresses, obtained from [this](#) site on August 31, 2021.

Usage

```
data(ccc_list)
```

Format

A data frame with the following variables:

College College name.

District District name.

Address Address of the college.

Phone Phone number of the college.

Website Web address of the college.

Examples

```
data(ccc_list)
```

| | |
|-----------------|---------------------------------|
| survey_response | <i>Survey response data set</i> |
|-----------------|---------------------------------|

Description

Synthetic survey data from a community college used to illustrate survey non-response/response bias and methods to reduce this bias. This data set (`survey_response`) contains a list of respondents and non-respondents in a student satisfaction survey. It also contains demographic information of all students that were invited to participate in the survey. Actual survey results for the respondents are stored in the [survey_results](#) data set.

Usage

```
data(survey_response)
```

Format

A data frame with student level data attributes:

Email_Masked student's masked email address, also serving as the student's unique identifier.

Responded 1 (yes) or 0 (no) indicating whether or not a student responded to the survey.

Age the age of the student.

Female 1 (yes) or 0 (no) indicating whether or not the student identifies as female.

Ethnicity ethnicity (one of: American Indian or Alaska Native, Asian, Black or African American, Hispanic / Latino, Native American or Pacific Islander, Two or More Races, Unreported or Unknown, or White).

K12_Student 1 (yes) or 0 (no) indicating whether or not a student is K-12 special admit student taking college courses.

First_Time_College_Student 1 (yes) or 0 (no) indicating whether or not a student is a first time college student (not a university student taking courses at a community college, not a continuing student, not a returning student, etc.).

Veteran 1 (yes) or 0 (no) indicating whether or not a student is a veteran.

DSPS 1 (yes) or 0 (no) indicating whether or not a student has ever leveraged the Disabled Student Programs and Services (eg, has a learning disability and requires extended test time).

EOPS 1 (yes) or 0 (no) indicating whether or not a student has ever been a part of the Extended Opportunity Programs and Services, a program focused on helping low income students succeed.

FinAidEver 1 (yes) or 0 (no) indicating whether or not a student has ever received federal or state financial aid at the district.

Honors 1 (yes) or 0 (no) indicating whether or not a student is part of the college's honors program.

International 1 (yes) or 0 (no) indicating whether or not a student is on an international student visa.

ESL 1 (yes) or 0 (no) indicating whether or not a student is considered an English language learner.

Athlete 1 (yes) or 0 (no) indicating whether or not a student participate in the college's athletics program.

GPA_Beg_Of_Term the student's cumulative grade point average (GPA) at the district prior to the start of the term in which the survey was administered. Students that have not attempted any courses previously will have a GPA of 0.

Units_Attempted_Beg_Of_Term the total number of units the student attempted at the district prior to the start of the term in which the survey was administered.

Units the student's total number of units enrolled in the term in which the survey was administered.

Noncredit_Student 1 (yes) or 0 (no) indicating whether or not a student is non-credit student in the term in which the survey was administered.

Part_Time_Student 1 (yes) or 0 (no) indicating whether or not a student is a part-time student (>0 and <12 units) in the term in which the survey was administered.

Full_Time_Student 1 (yes) or 0 (no) indicating whether or not a student is a full-time student (>=12 units) in the term in which the survey was administered.

Any_Online 1 (yes) or 0 (no) indicating whether or not a student was enrolled in at least 1 online course in the term in which the survey was administered.

All_Online 1 (yes) or 0 (no) indicating whether or not a student was enrolled in at all online course in the term in which the survey was administered.

Examples

```
data(survey_response)
```

| | |
|----------------|--------------------------------|
| survey_results | <i>Survey results data set</i> |
|----------------|--------------------------------|

Description

Synthetic survey data from a community college used to illustrate survey non-response/response bias and methods to reduce this bias. This data set (`survey_results`) contains responses to some of the student satisfaction survey questions. For a list of students invited to participate in the survey, as well information on which students responded to the survey, see the [survey_response](#) data set.

Usage

```
data(survey_results)
```

Format

A data frame with responses to a student satisfaction survey. For questions with Likert scale responses and asking about satisfaction, the options are Very Dissatisfied, Somewhat Dissatisfied, Somewhat Satisfied, or Very Satisfied.

Email_Masked student's masked email address, also serving as the student's unique identifier.

Overall Satisfaction (Likert scale) Student's overall satisfaction at the college.

Classroom learning experience (Likert scale) Student's satisfaction with classroom learning.

Your ability to register for classes you want/need (Likert scale) Student's satisfaction with the ability to register for classes.

Variety of courses offered (Likert scale) Student's satisfaction with variety of courses offered.

Class size (number of students in a class) (Likert scale) Student's satisfaction with class size.

Campus safety/security (Likert scale) Student's satisfaction with campus safety.

Financial Aid Office (Likert scale) Student's satisfaction with the Financial Aid Office.

Tutoring Services (Likert scale) Student's satisfaction with Tutoring Services.

Selecting a major (Likert scale) Student's reported difficulty in selecting a major at the college. The options are Very Difficult, Somewhat Difficult, Somewhat Easy, and Very Easy

Canvas (Likert scale) Student's satisfaction with Canvas, the college's online learning management system.

Have you heard about the new California Initiative called Guided Pathways? (Binary) 1 (Yes) or 0 (No)

Examples

```
data(survey_results)
```

| | |
|---------|--|
| uc_list | <i>University of California address data set</i> |
|---------|--|

Description

List of Universities of California (UC) and their addresses, gathered on September 1, 2021.

Usage

```
data(uc_list)
```

Format

A data frame with the following variables:

University University name.

Address Address of the university.

Examples

```
data(uc_list)
```

| | |
|----------------|--------------------------------|
| writing_center | <i>Writing center data set</i> |
|----------------|--------------------------------|

Description

Sample data from a community college used to illustrate program evaluation methods. The data set contains students' first enrollment/attempt in ENG 1, a first year college level writing course, over several years. Students may choose to receive additional support outside of the classroom at the college's writing center by enrolling in a 0.5-unit writing center/conference support course. Students may use the writing center as a study center where they have access to computers, dictionaries, thesauri, etc. At the center, students may also schedule 1:1 conferences with the instructor on duty, to get assistance and/or feedback on their writing by bringing a writing prompt or draft of their paper. Students registered for this support course are required to spend at least 12 hours in the center, and to have at least 2 formal conferences.

Usage

```
data(writing_center)
```

Format

A data frame with student level data attributes:

Student_ID a unique identifier of the student.

Term Describes the year and term (fall, spring, summer) in which the student was enrolled in ENG 1.

Section_ID an identifier that describes the ENG 1 section for which the student was enrolled in. Students with the same Section_ID are considered to be in the same class.

Instructor_ID an identifier that describes the instructor that taught the section. Students with the same Instructor_ID were taught by the same instructor.

Main_CourseID The course for which the student was enrolled in. For this sample data set, this field should be ENG 1 for all rows.

Gender gender (one of: Male, Female, Decline).

Age the age of the student when they enrolled in ENG 1.

Ethnicity ethnicity (one of: African American, Asian, Decline to State, Hispanic / Latino, Mixed Ethnicity, Native American, Pacific Islander, or White, Non-Hispanic).

FirstGen 1 (yes) or 0 (no) indicating whether or not a student is a first generation college student.

Military 1 (yes) or 0 (no) indicating whether or not a student is using military veteran benefits (ie, a veteran).

FosterYouth 1 (yes) or 0 (no) indicating whether or not a student was a foster youth.

DSPS 1 (yes) or 0 (no) indicating whether or not a student has ever leveraged the Disabled Student Programs and Services (eg, has a learning disability and requires extended test time).

FinAid 1 (yes) or 0 (no) indicating whether or not a student has ever received federal or state financial aid at the district.

Units_Attempted_Beg_Of_Term the total number of units the student attempted at the district prior to the start of the term for which they enrolled in the current course.

GPA_Beg_Of_Term the student's cumulative grade point average (GPA) at the district prior to the start of the term for which they enrolled in the current course. Students that have not attempted any courses previously will have a GPA of 0.

TermUnitsAttempted the total number of units attempted in the current term.

K12_Student 1 (yes) or 0 (no) indicating whether or not a student is K-12 special admit student taking college courses.

First_Time_College_Student 1 (yes) or 0 (no) indicating whether or not a student is a first time college student (not a university student taking courses at a community college, not a continuing student, not a returning student, etc.).

Nonresident_Tuition_Exempt 1 (yes) or 0 (no) indicating whether or not a student is exempt from nonresident tuition based on AB 540.

International 1 (yes) or 0 (no) indicating whether or not a student is on an international student visa.

Nonresident 1 (yes) or 0 (no) indicating whether or not a student is not a resident of the college's state.

WR_Center 1 (yes) or 0 (no) indicating whether or not a student is enrolled in the writing center support section.

Main_Course_SuccessFlag 1 (yes) or 0 (no) indicating whether or not a student was successful (grades of A, B, C, or P) in the main writing course, Main_CourseID.

Main_Course_GradePoints a value of 4 (A), 3 (B), 2 (C), 1 (D), or 0 (F) indicating the student's grade in Main_CourseID.

HS_GPA student's high school GPA, if available.

Online 1 (yes) or 0 (no) indicating whether or not the student's Main_CourseID section was an online or in-person.

N_Center_Visits the student's number of writing center visits in the term.

Center_Attendance_Hours the total amount of time the student spent at the writing center in the term, measured in hours.

N_Conf the number of formal conferences the student had with an on-duty instructor at the writing center.

Examples

```
data(writing_center)
```

Index

* datasets

ccc_list, 2

survey_response, 2

survey_results, 4

uc_list, 5

writing_center, 5

ccc_list, 2

survey_response, 2, 4

survey_results, 2, 4

uc_list, 5

writing_center, 5