Package ‘SGPdata’

April 18, 2021

Type Package

Title Exemplar Data Sets for Student Growth Percentiles (SGP) Analyses

Version 25.0-0.0

Date 2021-4-15

Maintainer Damian W. Betebenner <dbetebenner@nciea.org>

Depends R (>= 4.0)

Suggests SGP (>= 1.9-0.0), knitr, rmarkdown

Imports crayon, data.table (>= 1.9.6)

Description Data sets utilized by the ‘SGP’ package as exemplars for users to conduct their own student growth percentiles (SGP) analyses.

URL https://CenterForAssessment.github.io/SGPdata,
   https://github.com/CenterForAssessment/SGPdata,
   https://cran.r-project.org/package=SGPdata

BugReports https://github.com/CenterForAssessment/SGPdata/issues

VignetteBuilder knitr

LazyLoad Yes

LazyData Yes

LazyDataCompression xz

License GPL-3

NeedsCompilation no

Author Damian W. Betebenner [aut, cre],
   Adam R. Van Iwaarden [aut],
   Ben Domingue [aut]

Repository CRAN

Date/Publication 2021-04-18 21:50:15 UTC
R topics documented:

- SGPdata-package
- sgpData
- sgpData_INSTRUCTOR_NUMBER
- sgpData_LONG
- sgpData_LONG_COVID
- sgptData_LONG

Index

<table>
<thead>
<tr>
<th>Package</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SGPdata-package</td>
<td>Exemplar Data Sets for Student Growth Percentile (SGP) Analyses</td>
</tr>
</tbody>
</table>

Description

The 'SGPdata' package contains five exemplar data sets utilized by the 'SGP' package for testing and examples.

Details

- Package: SGPdata
- Type: Package
- Version: 25.0-0.0
- Date: 2021-4-15
- License: GPL-3
- LazyLoad: Yes
- LazyData: Yes

<table>
<thead>
<tr>
<th>Package</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>sgpData</td>
<td>Longitudinal Student Assessment Data</td>
</tr>
</tbody>
</table>

Description

A dataset comprising a subset of five years of state assessment reading data suitable for student growth percentile and percentile growth projection/trajectory analyses. The dataset is used in examples provided in the documentation for the studentGrowthPercentiles and studentGrowthProjections functions, and serves as an exemplar for user construction of their own datasets for growth percentile analyses.
Format
A data frame of student level observations for five years on the following variables.

- **ID**  Student ID
- **GRADE_2017**  Student Grade Tested 2017, possibly missing
- **GRADE_2018**  Student Grade Tested 2018, possibly missing
- **GRADE_2019**  Student Grade Tested 2019, possibly missing
- **GRADE_2020**  Student Grade Tested 2020, possibly missing
- **GRADE_2021**  Student Grade Tested 2021, possibly missing
- **SS_2017**  Student Scale Score 2017, possibly missing
- **SS_2018**  Student Scale Score 2018, possibly missing
- **SS_2019**  Student Scale Score 2019, possibly missing
- **SS_2020**  Student Scale Score 2020, possibly missing
- **SS_2021**  Student Scale Score 2021, possibly missing

Note
All datasets used with studentGrowthPercentiles and studentGrowthProjections must be specifically formatted as wide format files. The first variable/column is the student ID variable. The next set of columns provide the grade of the student across all the years provided in the dataset (possibly missing). The last set of columns provide the scales scores of the student in the respective grades. Multi-year operational analyses benefit from putting data in long format. See sgpData_LONG and associated documentation for a comprehensive account.

Source
Anonymized student level state assessment data

See Also
sgpData_LONG

sgpData_INSTRUCTOR_NUMBER

*Lookup table associating INSTRUCTOR_NUMBER with student ID by YEAR and CONTENT_AREA*

Description
A lookup table comprising student IDs associated with INSTRUCTOR_NUMBERa subset of five years of state assessment reading and mathematics data suitable for student growth percentile and percentile growth projection/trajectory analyses. The dataset is in LONG format with each record representing a unique teacher by student by year by content area combination.
Format

A data frame of student level observations for five years and two content areas for the following variables.

- ID Unique Student Identification Number
- CONTENT_AREA Content area for student/teacher record (Reading or Mathematics)
- YEAR Year for student/teacher record
- INSTRUCTOR_NUMBER Unique instructor number identifier
- INSTRUCTOR_LAST_NAME Instructor last name
- INSTRUCTOR_FIRST_NAME Instructor first name
- INSTRUCTOR_WEIGHT Proportional weight associated with student/teacher record
- INSTRUCTOR_ENROLLMENT_STATUS Indicator of full enrollment status of student with teacher

Source

Anonymized student/teacher lookup table in LONG format

See Also

sgpData_LONG

Description

A dataset comprising a subset of five years of state assessment reading and mathematics data suitable for student growth percentile (SGP) and percentile growth projection/trajectory analyses. The dataset is in LONG format with each record representing a unique student, by year by content area combination. The example data is used in examples and illustration in the enclosed documentation to show how SGPs can be operationalized across years with the results being used to produce a variety of high quality visualizations both at the aggregate (e.g., school) and individual level.

Format

A data frame of student level observations for five years and two content areas for the following variables.

- ID Unique Student Identification Number
- LAST_NAME Student last name
- FIRST_NAME Student first name
- CONTENT_AREA Content area for student observation (Reading or Mathematics)
- YEAR Year for student observation
- GRADE Grade level of student observation
SCALE_SCORE  Student Scale Score
ACHIEVEMENT_LEVEL  Achievement level associated with student scale score
GENDER  Student gender
ETHNICITY  Student ethnicity
FREE_REduced_LUNCH_STATUS  Student free/reduced lunch status
ELL_STATUS  Student English Language Learner status
IEP_STATUS  Student Individual Education Plan status
GIFTED_AND_TALENTED_PROGRAM_STATUS  Student Gifted and Talented Program status
SCHOOL_NUMBER  School number associated with student record
SCHOOL_NAME  School name associated with school number and student record
EMH_LEVEL  Elementary, Middle, High school indicator for school attended
DISTRICT_NUMBER  District number associated with student record
DISTRICT_NAME  District name associated with school number and student record
SCHOOL_ENROLLMENT_STATUS  Indicator of full academic year student enrollment in school
DISTRICT_ENROLLMENT_STATUS  Indicator of full academic year student enrollment in school
STATE_ENROLLMENT_STATUS  Indicator of full academic year student enrollment in school
VALID_CASE  Indicator of whether the case is valid or invalid

Note

All datasets used with studentGrowthPercentiles and studentGrowthProjections must be specifically formatted as wide format files. The first variable/column is the student ID variable. The next set of columns provide the grade of the student across all the years provided in the dataset (possibly missing). The last set of columns provide the scales scores of the student in the respective grades. Operational users of this package are advised to use data formatted similar to the sgpData_LONG with the higher level functions prepareSGP, analyzeSGP, combineSGP, summarizeSGP, visualizeSGP, and outputSGP

Source

Anonymized student level state assessment data in LONG format

See Also

sgpData
**sgpData_LONG_COVID**  
*Longitudinal Student Assessment Data in LONG format with COVID-19 2020 interruption*

**Description**

A dataset comprising a subset of seven years of state assessment ELA and mathematics data across an 8 year span suitable for student growth percentile (SGP) and percentile growth projection/trajectory analyses. The dataset is missing 2020 data to help users model COVID related interruptions to student growth. The dataset is in LONG format with each record representing a unique student, by year by content area combination. The example data is used in examples and illustration in the enclosed documentation to show how SGPs can be operationalized across years with the results being used to produce a variety of high quality visualizations both at the aggregate (e.g., school) and individual level.

**Format**

A data frame of student level observations for seven years (across an 8 year span) and two content areas for the following variables.

- **VALID_CASE**  Indicator of whether the case is valid or invalid
- **CONTENT_AREA**  Content area for student observation (ELA or Mathematics)
- **YEAR**  Year for student observation
- **ID**  Unique Student Identification Number
- **IEP_STATUS**  Student Individual Education Plan status
- **GRADE**  Grade level of student observation
- **SCALE_SCORE**  Student Scale Score
- **ACHIEVEMENT_LEVEL**  Achievement level associated with student scale score
- **PRIOR_SCORE_1_YEAR**  Student scale score from 1 year prior (if available)
- **PRIOR_SCORE_2_YEAR**  Student scale score from 2 years prior (if available)
- **ETHNICITY**  Student ethnicity
- **FREE_REDUCED_LUNCH_STATUS**  Student free/reduced lunch status
- **ELL_STATUS**  Student English Language Learner status
- **GENDER**  Student gender
- **SCHOOL_NUMBER**  School number associated with student record
- **SCHOOL_NAME**  School name associated with school number and student record
- **DISTRICT_NUMBER**  District number associated with student record
- **DISTRICT_NAME**  District name associated with school number and student record
All datasets used with `studentGrowthPercentiles` and `studentGrowthProjections` must be specifically formatted as *wide* format files. The first variable/column is the student ID variable. The next set of columns provide the grade of the student across all the years provided in the dataset (possibly missing). The last set of columns provide the scales scores of the student in the respective grades. Operational users of this package are advised to use data formatted similar to the `sgpData_LONG_COVID` with the higher level functions `prepareSGP`, `analyzeSGP`, `combineSGP`, `summarizeSGP`, `visualizeSGP`, and `outputSGP`.

**Source**

Anonymized and altered student level state assessment data in LONG format

**See Also**

`sgpData_LONG`
SCALE_SCORE  Variable providing the scale score associated with the test record.
SCALE_SCORE_RASCH Variable providing the Rasch score associated with the test record.
SEM Variable indicating the standard error of measurement associated with the scale score.
ACHIEVEMENT_LEVEL Variable indicating the achievement level associated with the scale score.

Source
Anonymized student level interim assessment data in LONG format

See Also
sgpData_LONG
Index

* datasets
  sgpData, 2
  sgpData_INSTRUCTOR_NUMBER, 3
  sgpData_LONG, 4
  sgpData_LONG_COVID, 6
  sgptData_LONG, 7

* package
  SGPdata-package, 2
  SGPdata (SGPdata-package), 2
  sgpData, 2, 5
  SGPdata-package, 2
  sgpData_INSTRUCTOR_NUMBER, 3
  sgpData_LONG, 3, 4, 4, 7, 8
  sgpData_LONG_COVID, 6
  sgptData_LONG, 7