Impact of High School Teachers on the Geographic Relative Population IQ

About

This particular study analyzes two varied datasets, through which a correlation is achieved based on several similar variables. The first dataset includes the count of teachers and high school students for each state in the United States, accounting each year between 2001 and 2019. The second dataset calculates the population and average Intelligence Quotient(IQ) of every state in the United States.

Data Creation Range: January 2001 - December 2019
Created by: World Population Review and the National Center for Education Statistics (NCES)
Content: List of Maps and Datasets
Sources:
World Population Review:
https://worldpopulationreview.com/state-rankings/smartest-states
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8954344/
NCES US Education Data:
https://www.kaggle.com/datasets/georgetryfiates/national-center-for-education-statistics?resource=download&select=pupils_fte_teachers_2001_2019.csv

Use Cases

Potential real-world application of the data set

- The correlation between the number of teachers and students in a state could provide insights into how well-funded a state's education system is, and whether there is a need for more teachers to support students. This information could be used to guide education policy decisions and funding allocation.
- By analyzing the IQ data, it may be possible to identify patterns in cognitive ability across different regions of the United States. This could be used to guide public health policy decisions related to education, healthcare, and social programs.
- By comparing teacher-student ratios across different states and years, it may be possible to identify which states have more effective teachers. This information could be used to guide teacher training programs and improve overall teaching quality.

Data Nutrition Label

Data Provenance:

- The data sources are World Population Review and the National Center for Education Statistics (NCES).
    - The data includes two varied datasets:
    - Count of teachers and high school students for each state in the United States, accounting for each year between 2001 and 2019.
    - Population and average Intelligence Quotient(IQ) of every state in the United States.

Data Granularity:

- The data covers the numbers of full time employed teachers in each state in the United States for each year between 2001 and 2019.
- The datasets provide yearly data for each state and do not provide data for specific regions.

Data Currency:

- The data creation range is from January 2001 to December 2019.
- The datasets must be updated and may not reflect any changes after 2019.

Data Accuracy:

- The data's accuracy depends on the sources' data collection methods and data validation procedures.
- The data sources, World Population Review and NCES, are reliable and authoritative.
    - World Population Review is an independent organization that provides global demographic data, including population, area, density, and rankings of countries, regions, and cities around the world. The organization sources data from reputable international organizations, including the United Nations, the World Health Organization, and the World Bank.
    - The National Center for Education Statistics (NCES) is a part of the United States Department of Education, and is the primary federal entity responsible for collecting and analyzing data related to education in the United States publishes statistics and reports on education in the United States, covering a wide range of topics including student achievement, school demographics, funding, and staffing. Their data is known to be widely used by researchers, policymakers, and educators to make informed decisions about education in the United States.
- The World Population Review has calculated the derived state IQ partly from NAEP Reading and Math scores (in 2015, 2017, and 2019) which is an assessment tested to measure a student's academic achievement.

Data Relevance:

- The datasets are relevant to the study's objective of analyzing the correlation between the number of teachers and students, the population, and the average IQ across different states in the United States.
- The data is useful in providing insights into education policy decisions, public health policy decisions related to education, healthcare, and social programs, and teacher training programs.

Data Accessibility:

- The NCES dataset is available on Kaggle, a data science community platform, and can be accessed after creating a Kaggle account.
- The World Population Review dataset is available on their website. Also, the IQ data can be accessed through academic article written by Bryan J. Pesta for the Journal of Inelligence published on 27 February 2022.

Objectives

What do the instances that comprise the dataset represent (e.g., documents, photos, people, countries)?

- The instances in this dataset represent different states in the United States. Each instance contains information about a particular state's population, density, education system, and intelligence quotient (IQ) levels, among other variables.

How many instances are there in total (of each type, if appropriate)?

- The total number of instances in the dataset is 50. Specifically, there are 50 instances, each corresponding to a different state in the United States. Therefore, the dataset is a sample of 50 states from the larger set of all 50 states in the United States.

Does the dataset contain all possible instances or is it a sample (not necessarily random) of instances from a larger set?

- The datasets are from World Population Review and the National Center for Education Statistics (NCES) and only cover 2001 and 2019. While it does cover all states in the United States, it is still a sample of the total number of years and could be missing data from other years or states. Therefore, it is important to remember that any insights or conclusions drawn from this dataset are based on this particular sample and may not necessarily generalize to other populations or periods.

What data does each instance consist of?

- Each instance consists of information about a particular state, including state code, state name, population, density, and education-related information such as the number of full-time equivalent (FTE) teachers, number of pupils, and pupil-per-teacher ratio IQ-related information such as the average IQ and rank.

Are relationships between individual instances made explicit (e.g., users' movie ratings, social network links)?

- Since the data is organized by state and year, it's possible to investigate relationships between variables over time within individual states and across states.

Are there recommended data splits (e.g., training, development/validation, testing)?

- Splitting the data into training and testing sets may be appropriate to develop and evaluate predictive models.

Are there any errors, sources of noise, or redundancies in the dataset?

- There aren't any errors or sources of noise that exists in this data set.

Variables Dictionary

state_code
- The abbreviation of the state in two letters

fips
- Federal Information Processing Standards

densityMi
- Density of the population per mile

pop2023
- Population of the state in 2023

pop2022
- Population of the state in 2022

pop2020
- Population of the state in 2020

pop2019
- Population of the state in 2019

pop2010
- Population of the state in 2010

growthRate
- Population Growth Rate from 2022 to 2023

growth
- Population change from 2022 to 2023

growthSince2010
- Change in population from 2010 to 2023

averageIQ
- Average IQ of the state

rank
- The Average IQ rank of the state among 50 states

year
- The year of data measurement

pupils
- Number of highschool students

pupils_per_fte_teachers
- Number of high school students per full time employed teachers

IQ_status
- New Categorical Variable
- The status of the state whether they are in the lower 25th percentile (low), the middle 50th percentile (mid), or the upper 25th percentile (high)

per_change_students
- New Continous Numerical Variable
- Percent change of the numer of high school students from the previous year (2018)