Topic:

Impact of factors on high school graduation rate

Introduction/Motivation:

In 2004, there are only 71.7% students were graduated from high school successfully. However, in 2013, it increased to 81.4% with 1.8 million more people graduated successfully.

The government set up a goal that 90% of students will graduate from high school successfully in 2020.   To achieve this goal, it is necessary to find out root causes lead to high school graduation rate and predict dropout rate based on each factors.

This topic make me curious about the root factors about graduation rate and motivate to help more students finish their high school education.

Purpose:

To research for root causes affect high-school graduation rate and predict consequences dropout rate base on each factors.

Use result to give reasonable recommendations to help government achieve higher graduation rate.

Data:

http://datafordiplomas.devpost.com/details/resources

2010 census data merged 2013 graduation data

- 9898 records with missing value

- 565 variables including census/graduation factors

Detailed graduation data with graduation related variables

Mapping logic of censes tract overlap school districts

Process Dataset:

1. choose variables most related to topic (about 20-30)

2. clean data, eliminate missing values

3. Use programming, visualization, classification technics to analyze correlation between factors and results