

# Final Project

## INFO 5709

(Due: May 2<sup>nd</sup>)

### Dataset:

In this individual project, you will use the data from the U.S. Dept. of Education College scorecard. The dataset is shared with you termed as “[MERGED2018-19.xlsx](#)”. You can find the data dictionary in “[CollegeScorecardDataDictionary.xlsx](#)” file. The dataset can be found on “Canvas → Datasets”.

### Variables:

Investigate the following variables within the dataset.

	Variable name	Name of Data Element	Data type	Values and labels
1	ST_FIPS	FIPS code for state	Integer	Refer to “institution_data_dictionary” in CollegeScorecardDataDictionary.xlsx  For instance: 1 Alabama 2 Alaska ...  78 Virgin Island
2	CONTROL	Control of institution	integer	Refer to “institution_data_dictionary” in CollegeScorecardDataDictionary.xlsx  1- Public 2- Private nonprofit 3- Private for-profit
3	PREDDEG	Predominant undergraduate degree awarded	integer	Refer to “institution_data_dictionary” in CollegeScorecardDataDictionary.xlsx  0- Non-degree-granting 1- Certificate degree 2- Associate degree 3- Bachelor's degree 4- Entirely graduate degree granting
4	UGDS_WHITE	undergraduate degree-seeking students who are <b>white</b>	Character	-
5	UGDS_BLACK	undergraduate degree-seeking students who are <b>black</b>	Character	-

6	UGDS_HISP	undergraduate degree-seeking students who are <b>Hispanic</b>	Character	-
7	UGDS_ASIAN	undergraduate degree-seeking students who are <b>Asian</b>	Character	-
8	UGDS_AIAN	undergraduate degree-seeking students who are <b>American Indian</b>	Character	-
9	UGDS_NHPI	undergraduate degree-seeking students who are <b>Native Hawaiian</b>	Character	-
10	UGDS_2MOR	undergraduate degree-seeking students who are two or <b>more races</b>	Character	-
11	UGDS_NRA	undergraduate degree-seeking students who are non-resident <b>aliens</b>	Character	-
12	UGDS_UNKN	undergraduate degree-seeking students whose race is <b>unknown</b>	Character	-

## Requirements:

Perform the following tasks:

- 1) Create some initial charts and graphs in Tableau to provide an overview of the shape and structure of the dataset. Do not include such charts and graphs in the presentation.
- 2) Remove records containing missing or erroneous values from further analysis.
- 3) Develop at least **five interesting questions**. Create **different** charts and graphs to answer those questions.

## Project Presentation:

- 1) The presentation is open to you. Remember that the objective of your presentation is to effectively communicate the important information. At a minimum, your presentation should cover the following information:
  - A brief description of the dataset,
  - Number of records containing missing or erroneous values that were removed,
  - Developed questions, followed by appropriate charts and graphs providing compelling answers to those questions.

- 2) Create a power-point presentation and present it in a 3-minute movie on Flipgrid.
- 3) Watch at least one of your classmates' movies and record feedback.

**Grading Criteria:**

- 1) You can earn up to 60 points for this assignment (55 points for the presentation and 5 points for the feedback on your classmates' presentations).