


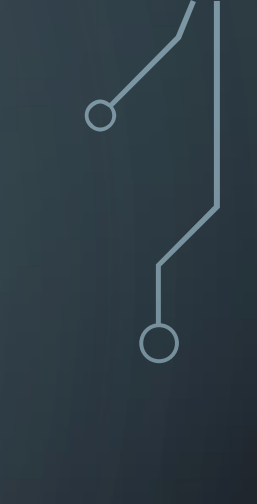
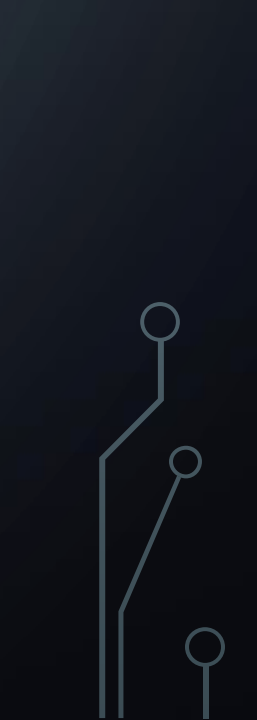
A glowing lightbulb with a circuit board overlay. The lightbulb is illuminated, casting a warm glow. The circuit board is a faint, light blue overlay on the left side of the image, featuring various lines and nodes. The background is a soft, out-of-focus blue.

INCOME PREDICTOR

ALBERTO TORRES


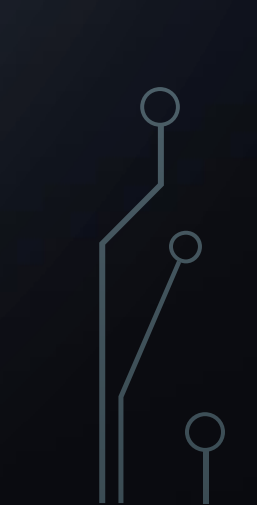


DECISIONS

- Crossroad about going to college.
 - Is the income after college worth it?
 - Predicting model to forecast annual income for individuals with a college education.
- 
- 
- 



A TOOL FOR EDUCATIONAL INSTITUTIONS

- With the predictive model, institutions can show students or potential students that in the long term a college education is the best route.
 - The model will try to predict if the individual has the potential to make more than 50 thousand a year.
 - Shows individuals a forecast of the future.
- 
- 



COLLEGE ENROLLMENT RATE DOWNFALL

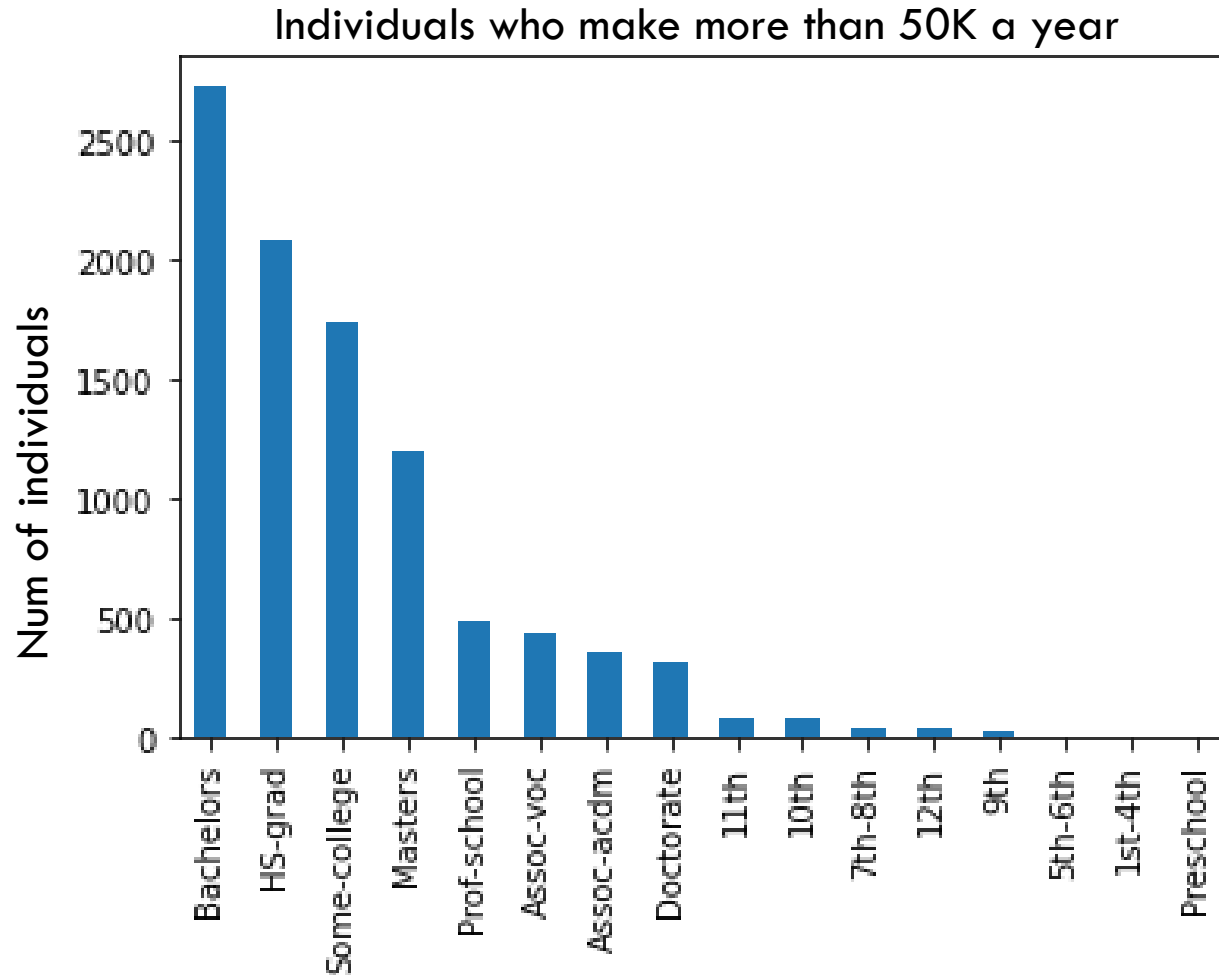
- Nearly one million less students enrolled in college in 2022 then before 2020.
- From 2019-2021 colleges seen undergraduate student enrollment fall by 6.6%.
 - Biggest decline in more than 50 years.

[Americans choose jobs
over college : NPR](#)

SHORT-TERM BENEFITS VS. LONG-TERM BENEFITS OF A DEGREE.

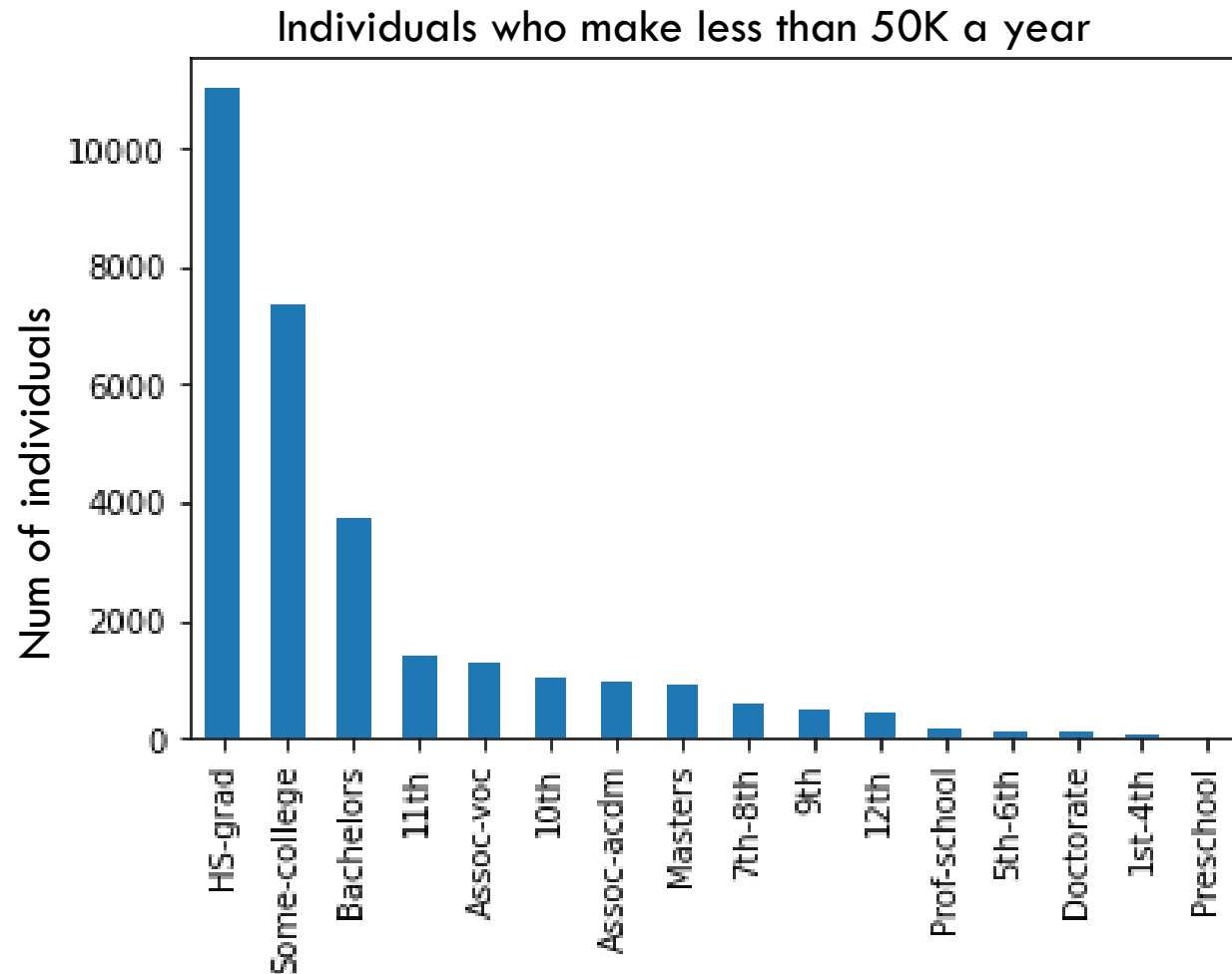
- According to Georgetown University, bachelor's degree holders were found to "earn a median of \$2.8 million during their career, 75% more than if they had only a high school diploma."
- Many high school students go for short-term money rather than a long-term career and high wages.

DATA UNDERSTANDING



Machine learning model:

- Demonstrates who is making more money annually.
- Predictor takes in attributes such as education, relationship, hours worked, etc.
- In Dataset 9.6K/39.4K individuals made over 50K



DATA UNDERSTANDING CONT.

- Georgetown University study, HS-Grads are the ones earning less than 50K a year.
- There's more than 50% HS-grads that make less than 50K then someone that has their bachelor's degree.

PREDICTIVE MODEL

Variables

- Target Variable: Income $>50K$
- Education: Highest level of education individual achieved
- Occupation: Current work position
- Race
- Gender: Male or Female
- Relationship: If individual is a husband, wife, single, etc.
- Hours worked per week

Model

- Random Forest- is learning method for classification, regression and other tasks that operates by constructing a multitude of decision trees at training time.
- Accuracy on dataset: individual's who make $<50K$: 82%
- Accuracy on dataset: individual's who make $>50K$: 64%

NEXT STEPS

- Bigger dataset
- Handle imbalanced dataset



THANK YOU