

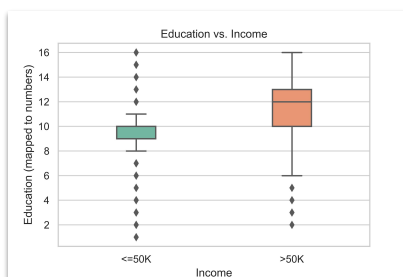
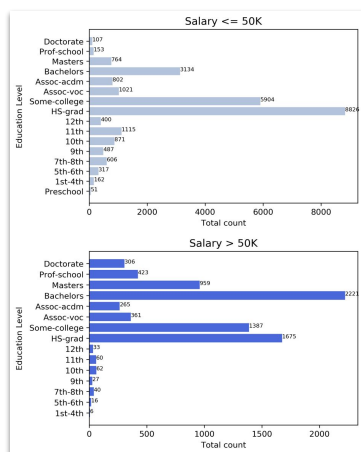
# Scatterplotter (Team 38)

## Executive Summary Report

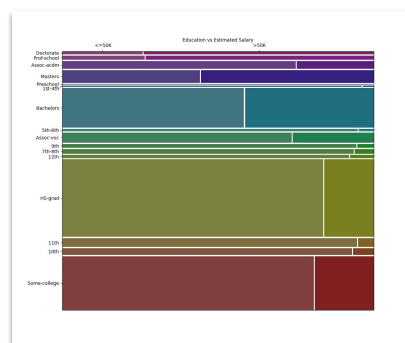
## Project Statement

- The goal of this project is to visualize attributes in the dataset to determine which attributes are the best predictor of income level among the population.
- UVW College is looking to bolster enrollment by utilizing our findings as marketing strategy.
- For the purpose of this project we utilized exploratory data analysis techniques to visualize 8 attributes, age, sex, education, marital status, race, capital gain, hours worked per week and occupation.
- Shown in the following slides are the top 4 attributes which we chose as the best predictors of income level; Education, Age, Capital Gain and Sex.

## Income vs Education



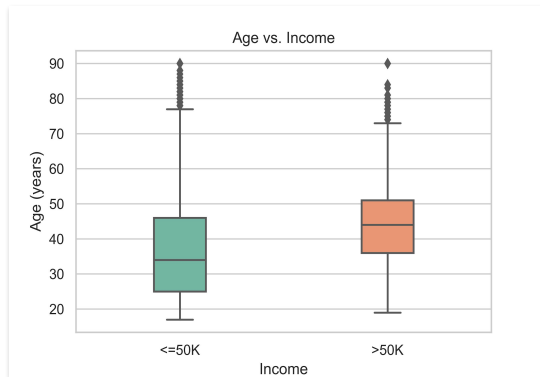
Another perspective of the same data. Higher education, higher earnings.



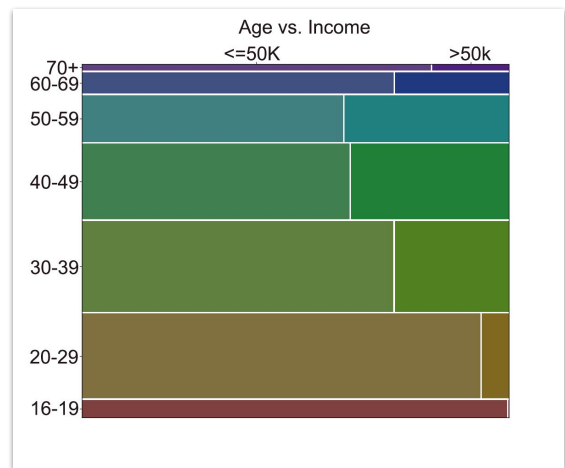
Majority of the individuals with education level of hs-grad, or some-college earn a salary <= \$50k  
Majority of the individuals with education level of bachelors and higher earn a salary > \$50k

Individuals with a HS degree demonstrate a large portion of the salaries <= \$50k.  
Individuals with a Bachelor's degree demonstrate a large portion of the salaries > \$50k.

# Income vs Age



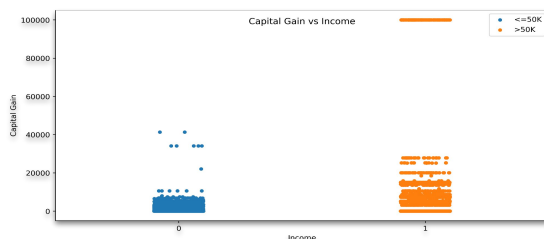
34 years is median age of those who earn  $\leq \$50k$ . Interquartile range for age is 25 to 46 for income earners of  $\leq \$50k$ .



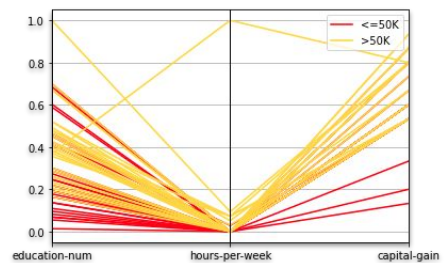
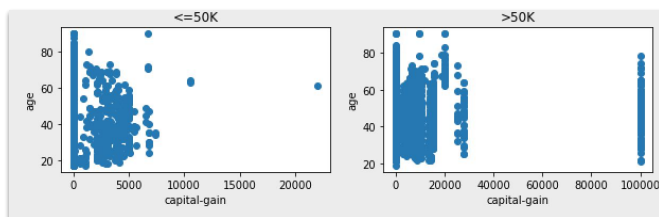
Ages are dominant in  $\leq 50k$  salaries.

Demonstrates the degree at age intervals are dominated. Individuals from ages 30-39 have a higher percentage of salaries above 50k compared to groups of other age intervals.

# Income & Age vs Capital Gain



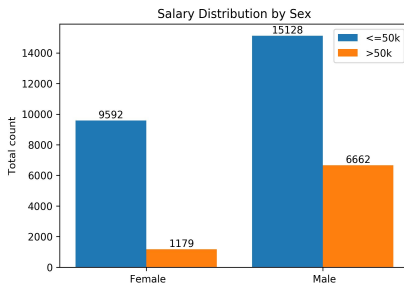
Capital Gain might be factor in earnings  $> \$50k$ . Capital Gains are significantly higher for people earning  $> \$50k$ , \$4000 on average in comparison to income earners of  $< \$50k$  which is \$148 on an average.



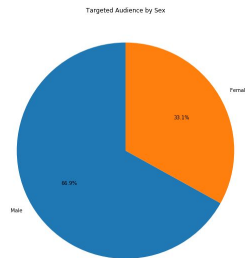
This visualization reiterates that  $> \$50k$  earners have higher education and higher capital gains. Number of hours worked has no major impact.

Independent of the age group, Capital Gains are significantly higher for people earning  $> \$50k$

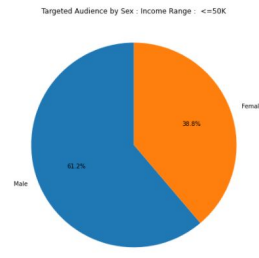
# Income vs Sex



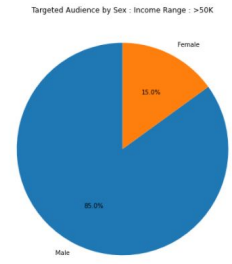
89% of Females and 69% of Males earn <=50K



The total audience consists of 66.9% male population and 33.1% female population.



For the income range of <=50K, the audience includes 61.2% male and 38.8% female.



However, as you can see there is 85% male audience when income range is >\$50k

## Team Roles and Interpretations

### TEAM ROLES:

**James Rodriguez:** Explained project requirements. Analyze attributes and come up with which graphs are best for each one. Proposed the attributes for Mosaic plot. Created Mosaic plots for Education and Age.

**Priyanka Sugawe:** Created shared drive documents. Schedule team meetings. Recommended the attributes for Pie chart and other charts. Wrote Progress Report. Created Pie Charts for Sex with respect to income.

**Quency Hadisubroto:** Came up with the team name. Explained project requirements. Analyze attributes and come up with which graphs are best for each one. Proposed the attributes for the Bar chart. Wrote System Documentation Report. Created Bar Charts for Education and Sex.

**Rakesh Krishnan Thaliyil Veetil:** Schedule team meetings. Analyze attributes and come up with which graphs are best for each one. Wrote System Documentation Report. Created scatterplots for capital gain.

**Sreenivas Nagaraja:** Proposed the features which can be utilized in a box plot. Analyze graphs and decide which graphs are best for attributes chosen. Decide which attributes are best for determining income level. Created box and whisker plot for age and strip plot for capital gain.

**Yingbing Wang:** Explained project requirements and expected results. Presented features that can be used in Parallel Coordinate plot. Calculates mean, median and standard deviation value for education number, and hours of work. Created parallel coordinate plot for capital gain and hours-per-week.

### Interpretations:

From our extensive research we have discovered:

- UVW College should market college courses to all individuals with high school or some college level education in the age group of 25 to 46.

Key statistics:

- 76% of population earn <= 50K
- 89% of females earn <=50K
- 61.2% of Males earn <=50K
- 36.7 years average age, earning <=50K vs. 44.2 years, earning >50K
- "High School Graduation" is the average education level, earning <=50K