

# Project Statement

With the specified breakpoint of 50k dollars using different visualizations a number of factors affecting income have been identified. These factors are gender, education, work class. After making different visualization to understand the data we see that education has a direct affect on income because higher income range showed higher education. Gender also plays a huge role as we can observe the males get higher income compared to females at same age. Work class also affects the income as distribution of wages is different across sectors.

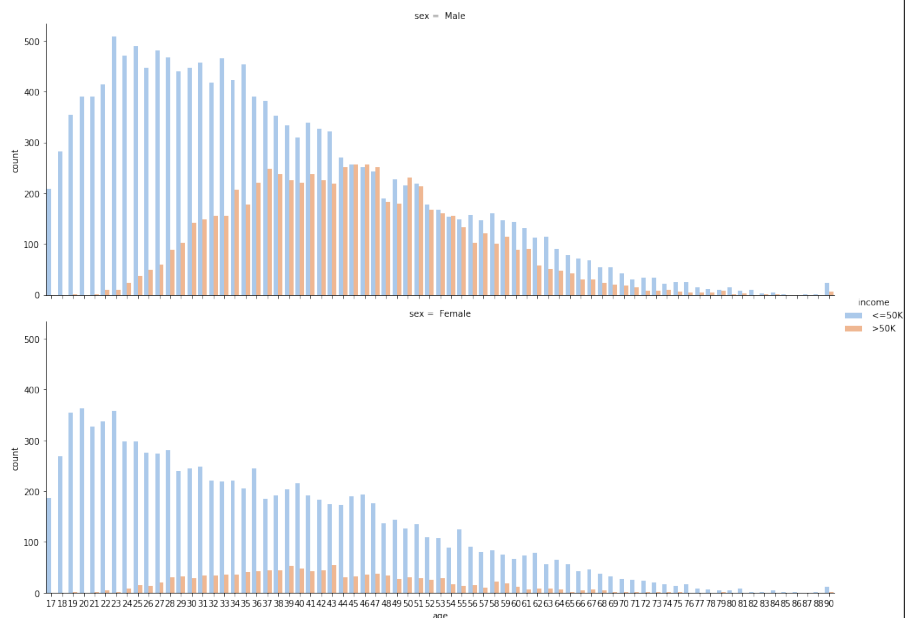
Apoorva Rastogi - CSE 578 Data Visualization

## Effect of gender and age on Income

Here we can see that gender and age are both factors which affect income. Majority of people earning more than 50k are 30-60 and men. For men after 40 there are similar no of people in both income category while the difference in earning is quite large for women at any age.

Narration: After going through the data provided the first visualization I would like to present is based on age and gender vs Income. Here red represents more than 50k whereas blue represents less than 50k. We can clearly see that a large majority of male population earns more than 50k where the count reaches above 200 at many age instances compared to female population where the count is not able to take 100 even. Also, majority of people earning more than 50k are between 30-60 for both the gender. If we notice the red graph in both the gender categories we can see then the graph for males has a steep increase in comparison to women which can tell us that more men are able to get increment in their salary compared to women.

Gender VS Income VS Age

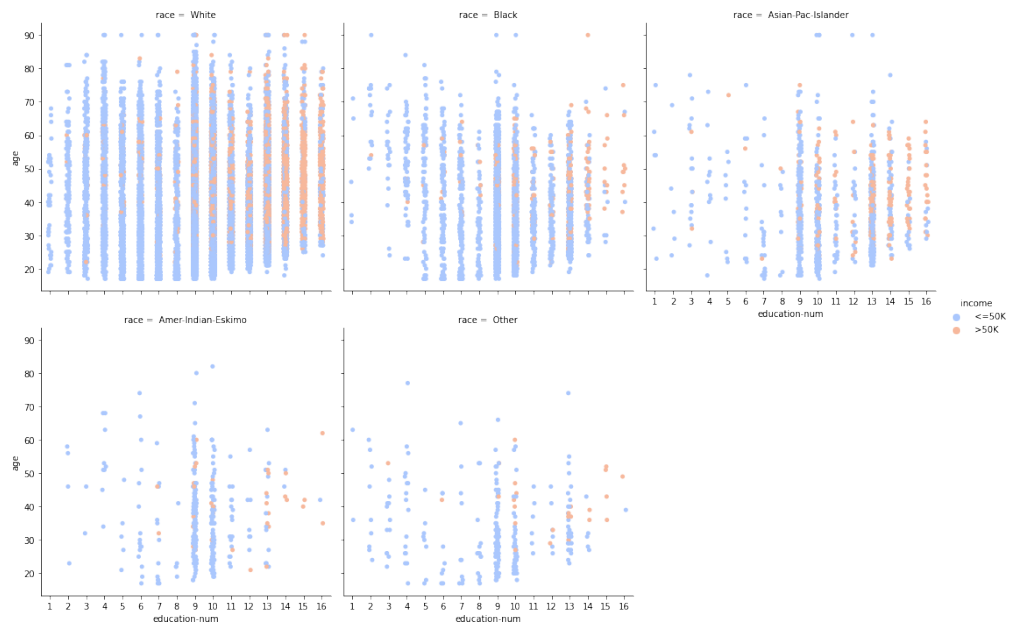


## Effect of Education -Num and Race on Income

Here we see that education-num is a strong factor affecting income. We can see that the majority of data points are for white race. However we can notice that for any race majorly the people earning more than 50k have education-num greater then or equal to 13.

Narration: The second visualization I would like to present is based on Race vs Education-num vs Income. We can observe that education-num has a direct impact on the earning of people. Majority of people with education num greater than 13 have income greater than 50k. Although more education does not guarantee the your income will be more than 50k as there are data points which indicate that some people with more education num are still earning less than 50k, it does help indicate that more education-num helps with more income. As for the race there is imbalance of data instances as we do not have equal data instance for all races. Therefore, it is not possible to say any effect of race on income. We can say that people from races are earning more than 50k.

Education-Num VS Income VS Race

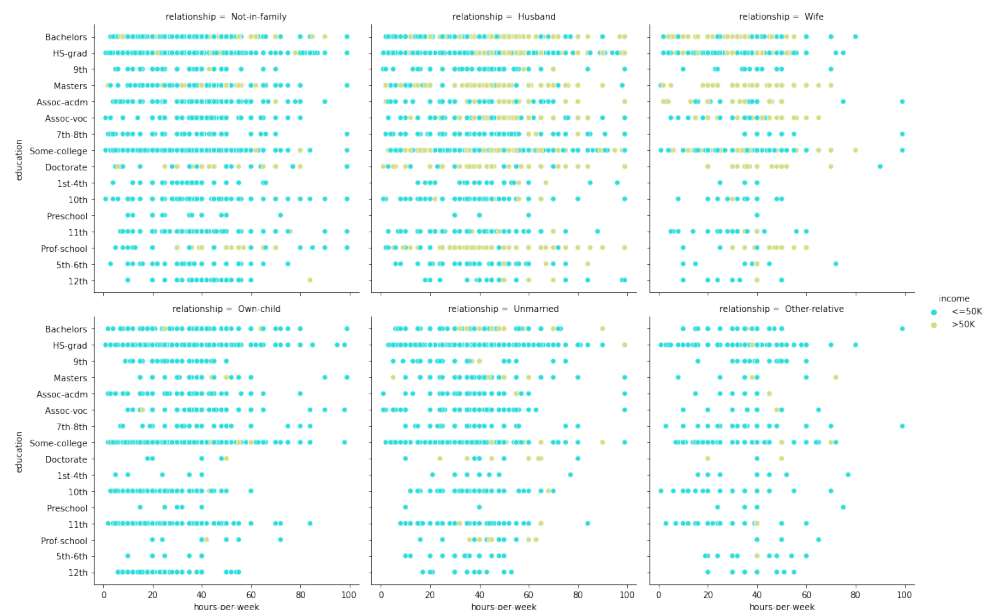


## Effect of Relation, Education on Income

We have more number of higher salary instances in masters, doctorate and prof school. Majorly people work 20-60 hours a week. There are a lot of people earning more than 50k in husband or wife relation compared to any other relation

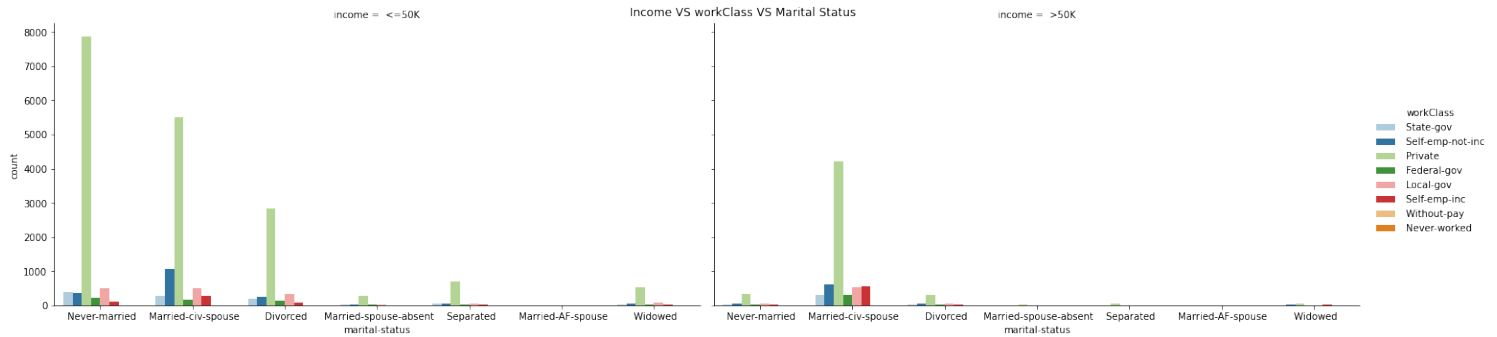
Narration: The third visualization I would like to present is based on Relation vs Education vs Income. Not in family and Husband has many more data points compared to any other relation. Husband and wife relation has most number of people earning more than 50k. From the visualization we can see that people with prof-school, doctorate, masters have majority of people earning greater than 50k and after those are people with Bachelors , some college , HS-grad with good no of people earning more than 50k.

Education VS Income VS Relation



## Effect of Marital-Status and Work Class on Income

People earning more than 50k either belong to never-married, married-civ-spouse or divorced as well as majority of data points. Married civ spouse is earning more than 50k in other sectors also apart from private. Private sector has the largest number of people as well as most number of people with salary greater than 50k.



**Narration:** The fourth and final visualization I have compared WorkClass with Marital Status and Income. We see the majority of people work in the private sector compared to any other sector. There are almost similar if not same number of people earning more than 50k for married-civ-spouse across sectors. There are some people in private sector earning more than 50k for never married and divorced. Majority of the data points belong to three marital statuses: namely never married, married-civ-spouse and divorced. Although the majority of people earning more than 50k belong to the private sector, the distribution is more skewed. Out of the total people working in the private sector, the percentage of people earning more than 50k is less compared to other sectors.

## Conclusion

From the above drawn visualizations we can draw the following conclusions.

- Males earn more than females. Based on the steep increase in the income graph we can also say that they get increments faster than females.
- Education has a direct effect on income because as we observed from the visualizations, mostly people earning higher salaries belong to higher education levels.
- WorkClass also comes up as an important factor in determining income. All sectors except private have almost similar numbers of people in both categories. Private, with the most people working in it, has the least percentage of people earning more than 50k. In terms of numbers, private has the highest number of people earning 50k.