

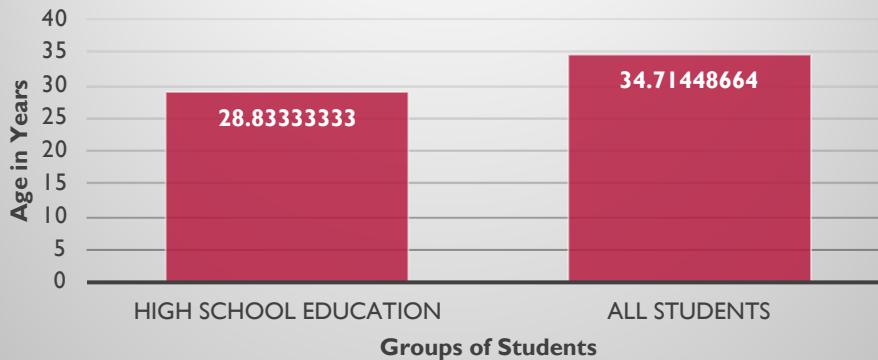
# JACKSON SAGER – RESOURCES

---

- All data cleaning was done with the help of the linked video alongside the instructor. Changes are highlighted in blue and include the average age column. On this column I got rid of students in the age range of 0-3 or 4.
- I rewatched course videos and used the following link: <https://www.exceltip.com/excel-formula-and-function/averageifs-sumifs-and-countifs-functions-in-microsoft-excel.html>

# WHAT IS THE AVERAGE AGE OF ALL STUDENTS, VS. THE AVERAGE AGE OF STUDENTS WITH A HIGH SCHOOL EDUCATION OR LOWER?

## All nanodegree student ages vs. Ages of students with high school education or lower



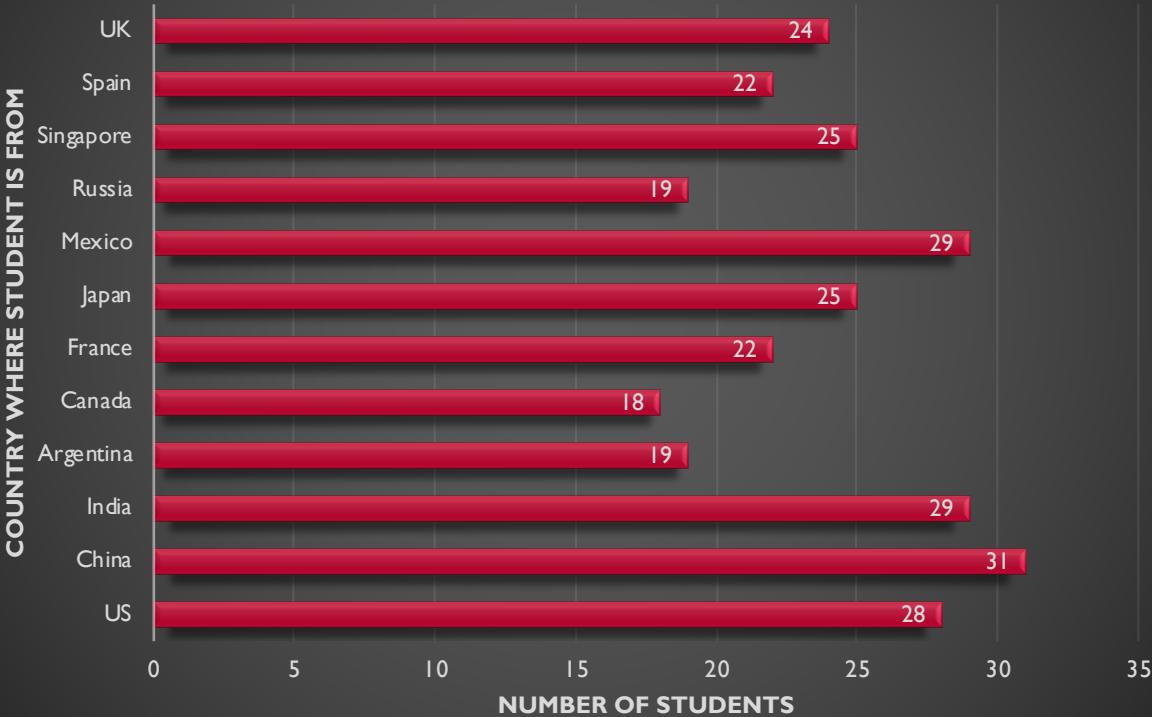
- This column chart highlights the average age of all students vs. the average age of students with a high school diploma or lower.  
- The mean of “All Students” is 34.7 years, while the mean of the “High School Ed. & Below” group is 28.8 years. The difference in ages is likely due to the time it takes to achieve more advanced degrees and therefore the whole population is older when averaged. The mode of “All” is 31 years, while for “HS & Below” it’s 23. The most common ages for each group again highlight how much older the “All” group is. The median for the “All” group is 33, with half of students being older and half younger while the “HS” group has a median of 24. While there may be some older outliers in the “HS” group, only half of them are older than 24.

- The two groups have relatively large standard deviations at 8.34 for “All” and 8.09 for “HS.” These numbers were calculated using standard deviations of population for “All” and of a segment for “HS.” When adjusted to a population for the “HS” group the numbers are slightly more similar, however both standard deviations are far from the mean, creating a more spread group of data.
- The range of ages for the “All” group is 59 and the range for the “High School and below group” is 27. By looking at these numbers I can assume there are some relatively extreme outliers in the “All” group and that the data is not bunched together.

\*This data is from Survey Respondents and is not from the entire Udacity Student population\*

# HOW MANY STUDENTS FROM EACH COUNTRY ARE STUDYING IN THE DEEP LEARNING FOUNDATIONS NANODEGREE?

**Deep Learning Foundations Students By Country**



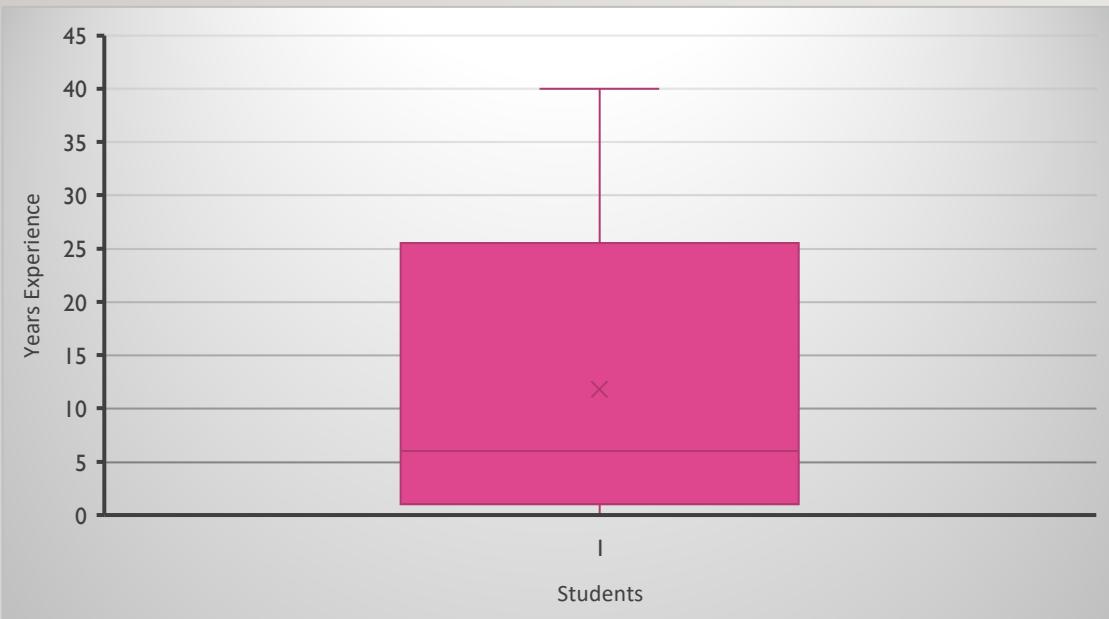
Here is a bar chart showing the number of students from each country enrolled in the Deep Learning Foundations Nanodegree.

The median is 24.5 and is very close to the mean. The most common number of students, or mode, was 29, 22 and 19. Thus we have a multimodal set. All three of these numbers of students were represented by two countries. The mean was 24.25. This number is relatively close to both the minimum and maximum values in our data set.

The standard deviation is 4.18. This tells us that the data is moderately spread out from the mean. The range is 13 and no one country is an extreme outlier in this group.

\*This data is from Survey Respondents and is not from the entire Udacity Student population\*

# WHAT IS THE 5 NUMBER SUMMARY FOR STUDENTS AND THEIR LEVEL OF EXPERIENCE?



Mean	8.04530744
Standard Deviation	7.15712823
Mode	1
Range	40

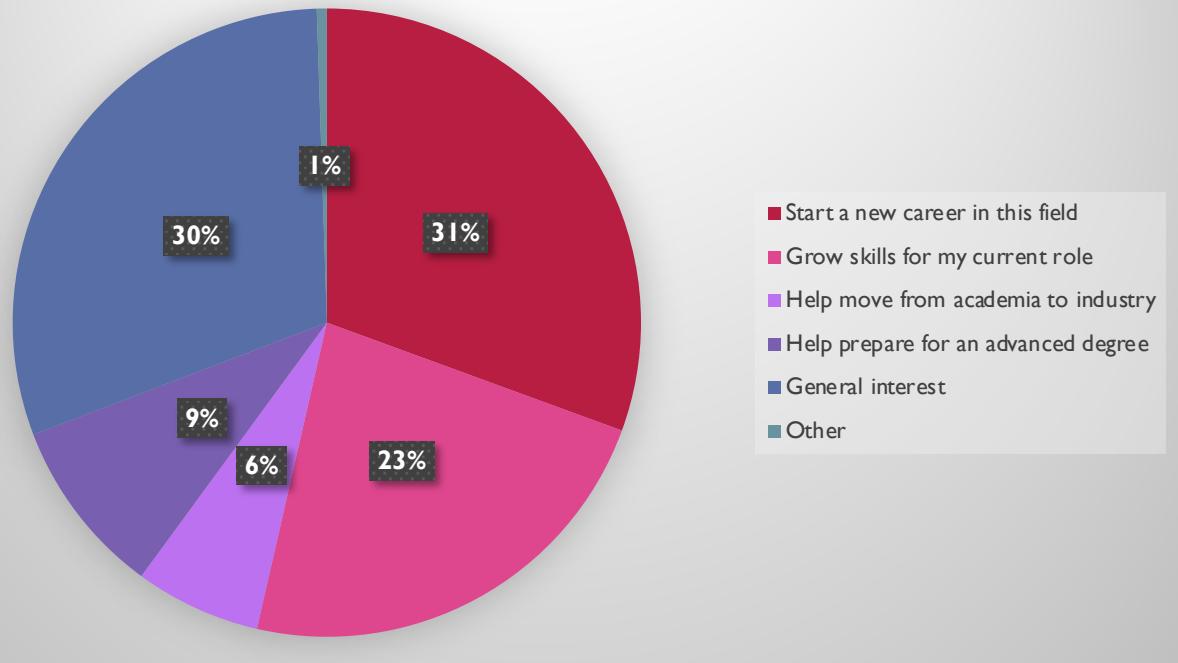
Minimum	0
First Quartile	2
Median	6
Third Quartile	11
Maximum	40

To the left, I've included a box and whisker plot intended to highlight the 5 number summary of the students and their experience. The mean, or average years of experience for students is roughly 8 years. The median is 6 years with half of students having more experience and half having less. Because the median is greater than the mean and the mode, the data distribution is skewed right. The most common value overall, or mode is 1, meaning most students have only 1 year of experience. Because some students have zero years experience and at least one student has 40 years work experience, our range is rather large at 40. A high variability of experience gives us a standard deviation of 7.15.

\*This data is from Survey Respondents and is not from the entire Udacity Student population\*

# WHY DO STUDENTS PURSUE THE NANODEGREE PROGRAM?

What motivates students to pursue a nanodegree?



The pie graph highlights the reasons students chose to pursue a nanodegree by percentage.

The largest volume of students chose to pursue a nanodegree because they desire to be in a new career field. However, students could choose multiple fields to answer from.

\*This data is from Survey Respondents and is not from the entire Udacity Student population\*