Udacity Students Survey Analysis

study project for DFND

by Tatiana Kurilo

github.com/TianaQ

Who are the respondents?

752 participants,
the most frequent answers

Interests

55% - to start a new career
55% - general interest in the field
41% - to grow skills for the current role

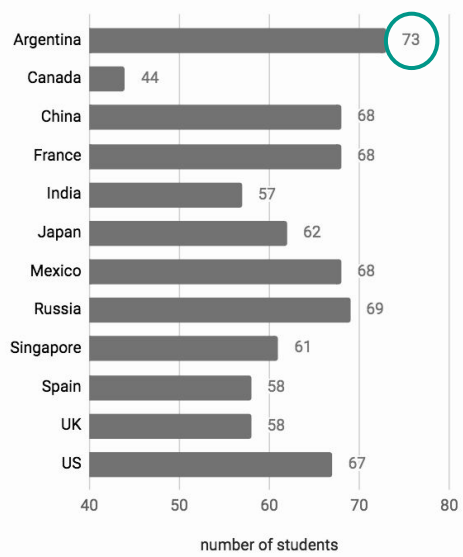
Nanodegrees

39% - Deep Learning Foundations
31% - Machine Learning Engineer
21% - Data Analyst

Education

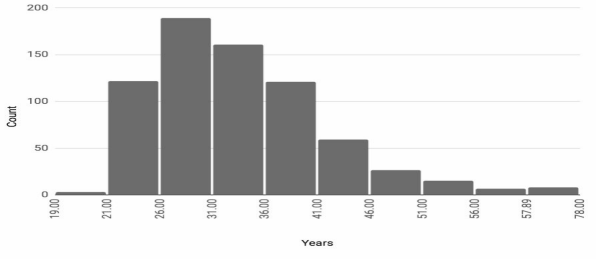
42% - Masters
38% - Bachelors
10% - PhDs

Countries



Age

range from **19** to **78** years,
32-**33** on average
middle **50%** - from **27** to **37** years



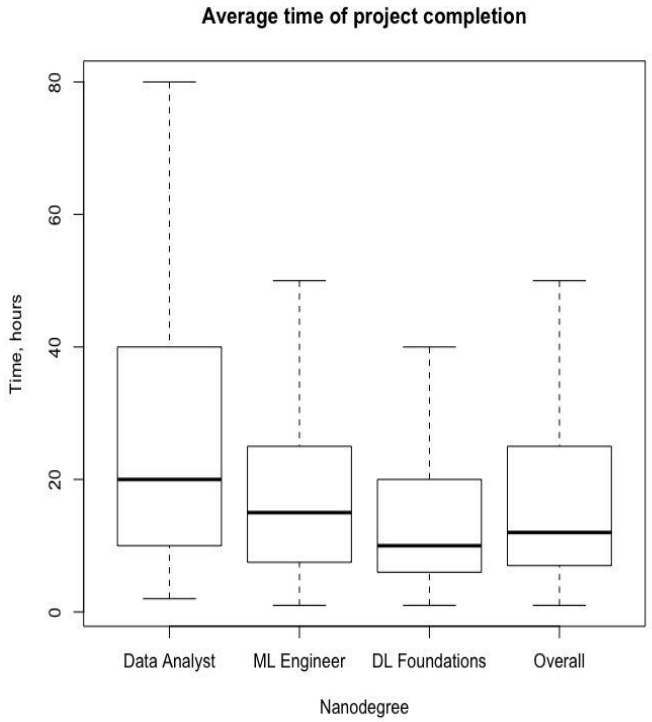
Employment

82% - employed
30% from Technology & Internet
38% work as individual contributors
25% - software engineers
20% - data-related occupations

Found out about Udacity

59% - via Google
20% - via word of mouth

Time to complete projects by Nanodegree



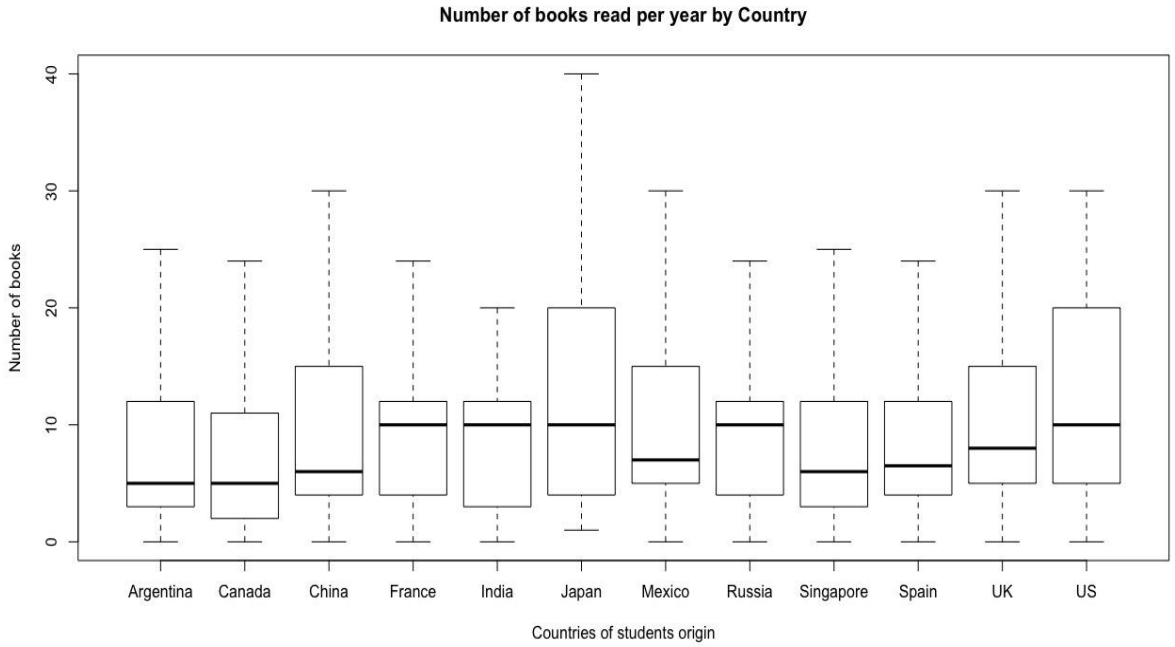
For the purpose of readability the outliers are omitted on the chart.

The analysis of summary statistics of the distributions of time the respondents reported they spent completing their Nanodegree projects shows that the variability of answers of Data Analyst Nanodegree participants is higher than for other two popular Nanodegrees - Machine Learning Engineer and Deep Learning Foundations, and than the spread of overall distribution. The average time reported by DAND participants is also higher - both for means and medians. Since all the distributions are right-skewed, the median and IQR are more reliable measures, than the mean and standard deviation.

To conclude whether on average it takes students more time to complete projects of DAND, than of MLEND and DLFND, the confidence intervals must be calculated for means using the sample sizes.

	Data Analyst	ML Engineer	DL Foundations	Overall
Q1	10	7.75	6	7
Q2	20	15	10	12
Q3	40	25	20	25
Mean	43.24	38.92	20.68	32.97
St.D.	87.32	94.76	38.4	75.32

Reading by Country

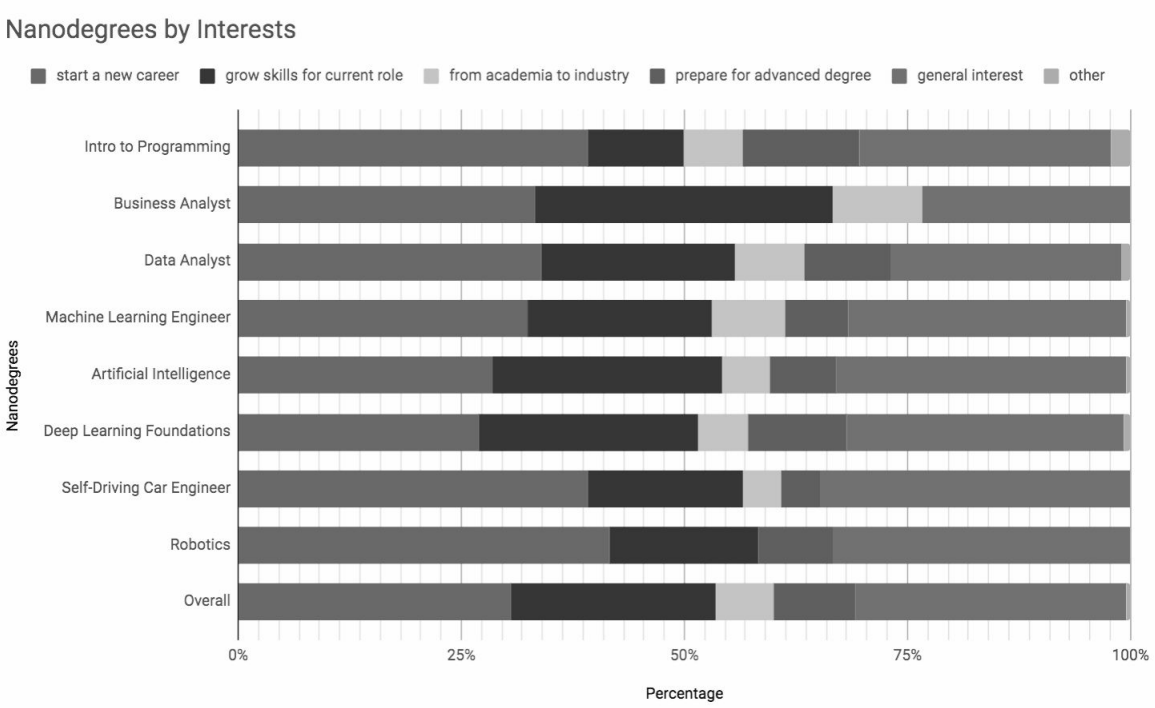


For the purpose of readability the outliers are omitted on the chart.

The distributions of reported number of books, read of listened to by students per year, differs between countries. The highest variability can be seen in answers of Japanese participants, they also have the highest minimum of **1**, instead of **0** for all others.
While median values vary in rather small range - from **5** to **10**, the difference in means is much wider - from **9.25** to **27.26** books per year.

Additional analysis is required to calculate the population parameters and to estimate whether the difference of country means is statistically significant.

Nanodegree by Interest



Though the main three interests:

- start a new career
- grow skills for the current role
- general interest

remain the most noticeable for all nanodegrees, there are slight differences in proportions between NDs.

Thus, for example, **0** respondents from Business Analyst Nanodegree expressed the interest in preparation for an advanced degree, but the proportion of those BAND respondents who'd like to grow skills for the current role, is **25%**, which is comparatively higher that for other NDs.

What must be taken into account is that the number of respondents from BAND is relatively small in comparison with the whole sample (19 of 725), and therefore maybe be not quite representative and therefore require additional data for final conclusions.