

# UDACITY DATA FOUNDATIONS NANODEGREE

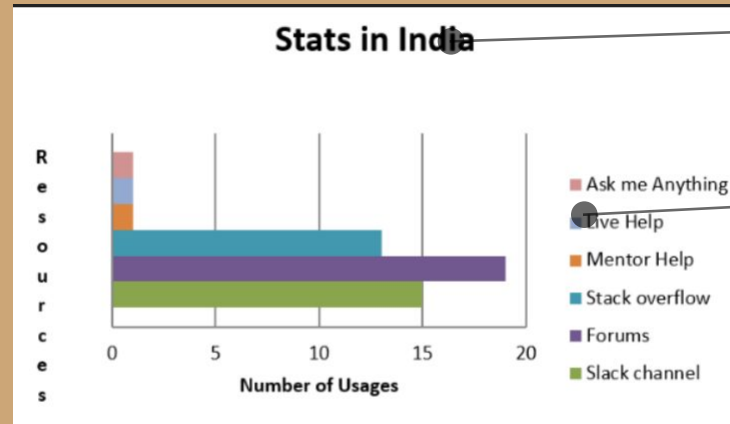
## PROJECT 2 - ANALYZE SURVEY DATA



: You are definitely going in the right direction. Overall, you are doing a great job here but there were a few issues that need to be fixed. I'm trying my best to provide helpful feedback. Hope it helps you in your learning process :)

Submitted by,  
Thirumalai devi ramya

# Which resource is found to be helpful in India ?



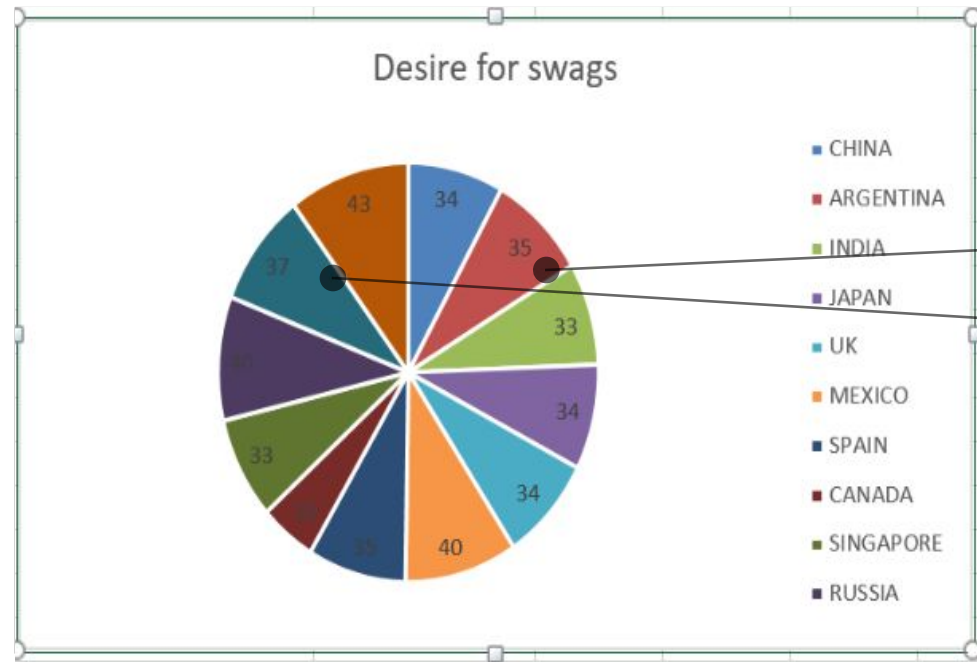
To answer the question such as **Which resource is most helpful when you got stuck in your nanodegree program in India?** I have created a **Bar chart**, with Usage count in X-axis and Resources in Y-axis.

From the Data visualization it is clear that, Mostly *Forums* are used in India with Slack and Stack overflow engaging the following positions.  
I calculated an analysis for India as I was very eager to know details about my birth country .

: Required: Please be more descriptive in your titles. Stats of what exactly?

: Required: There's no need to use so many colors here. Please remove the legends, stick to one color and add axis labels that align with the columns.

# Which country ranks highest in desire for swags?



To answer the question such as **Which country ranks the highest in desire for swags?** I have created a **Pie chart**.

The pie chart shows that US ranks the highest, the following position is shared by Mexico and Russia.

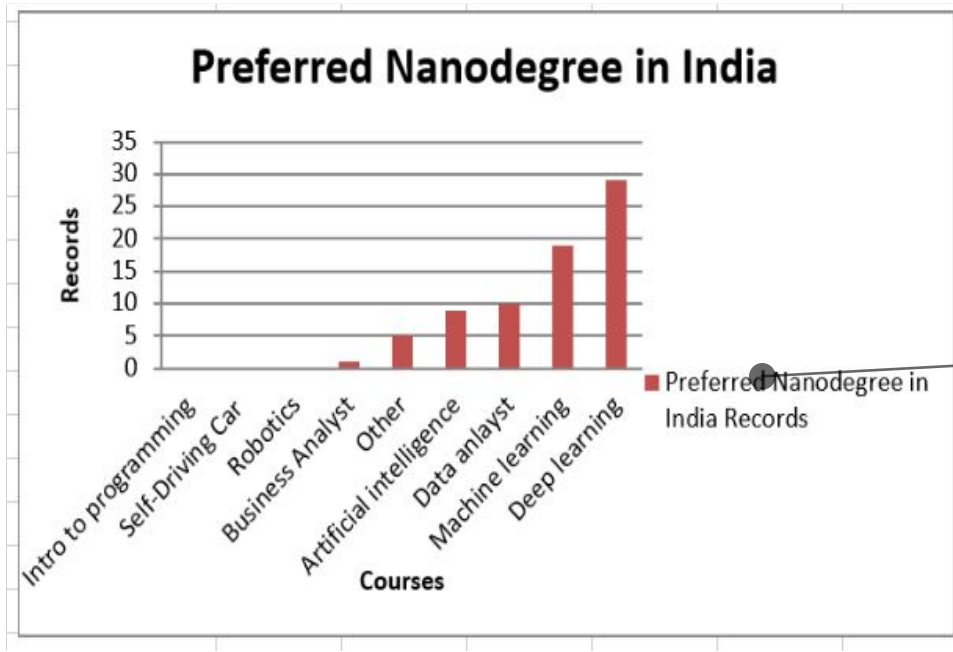
Canada ranks the lowest as it has the minimum number of students who wants to buy swags.

This measure is based on the amount of students who answered they want to buy swags.

: Required: Pie charts should always display percentages. If you want to display totals, try a bar chart instead.

: Required: Unfortunately, this plot isn't working the way you intended. Pie charts are great for compositions with just a few categories, however, in this case, having so many categories affects the clarity of the chart and it's very difficult to understand the message you tried to communicate. You have a few alternatives. You could show the top 5 categories and group the rest in "others" or you could change your approach and treat this as a comparison with a bar chart and totals.

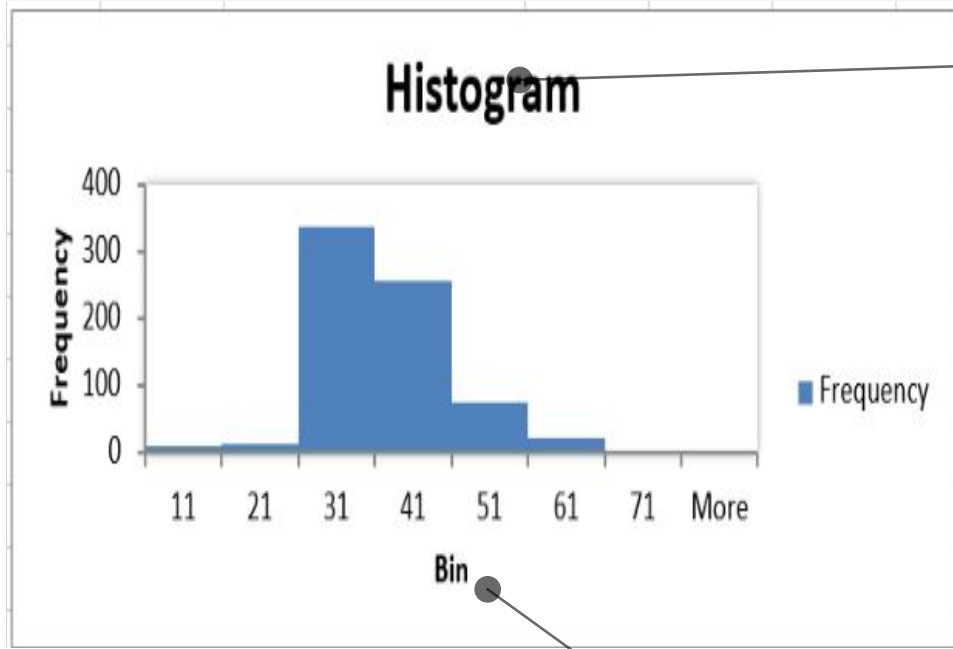
# Which Nanodegree is preferred in India?



To answer the question such as **Which Nanodegree is preferred in India?** I have created a **Column chart**, with Courses in X-axis and Records in Y-axis. From the visual it is clear that Deep learning is the most preferred nanodegree in India. Machine learning is the second preferred nanodegree. It is also clear that courses like Intro to programming, Robotics and Self driving car nanodegree are leastly preferred.

: Required: I understand why you are using this but legends in plots like this are not really necessary. Please remove it

# What is the age of Udacity students?



To answer the question such as **How much hours of sleep does Udacity students prefer?** I have created a **Histogram**, with Bin in X-axis and Frequency in Y-axis. From the analysis it can be said that people between age of 21-41 are more interested in learning courses from Udacity. It may be due to career enrichment, exploring new fields, changing careers.

: Required: Please be more descriptive in your title and use something different instead of Histogram

: Required: Please use the standard deviation somewhere to create insights! Just finding it isn't enough, you need to use it to explain something in its context.

In order to achieve this, you need to look into quantitative data. For example, you could try to find out the "normal" sleeping hours of the students after cleaning the data (mean +/- 1 SD)

: Required: The range should be used to generate insights. The range is the calculation of the difference between the lowest and highest value (MAX - MIN).

Acknowledgement: This data is from survey respondents and is not from the entire Udacity student population.

Once you have the range you could try to explain something about the outliers or the spread.

Note that you should use quantitative (numerical) data for this.

: Required: Please use all 3 measures of center to generate some insights in your slides. Simply finding and stating the mean, mode and median is not enough, you need to use them to explain something in your finding

A good way to use them would be to look for a normal distribution in the data (look into sleeping hours!) and then specify what happens to measures of center in such situation.

: Required: The axis titles must be descriptive. Please use something different instead of just "Bin"

: Thanks for limiting your findings to the survey respondents only