Milka Wafula

Data Foundation:Survey Project

Data visualisation using: **Google Spreadsheets**

Bar chart

Histagram

Scatter plot

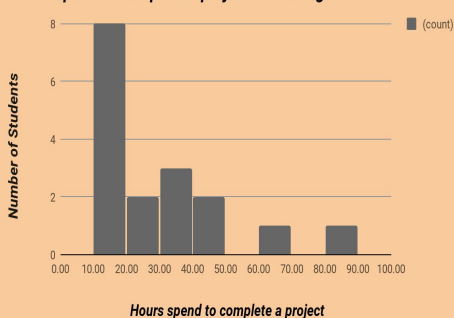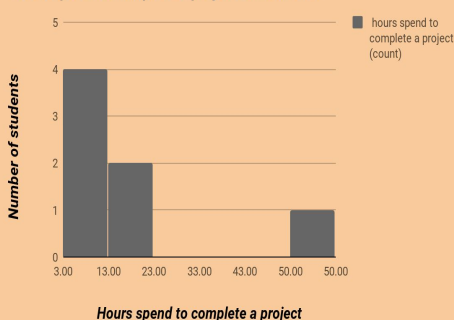*Box plot (Python)*

07.Nov.2018

UDACITY

Project Completion hours by Education Level

Hours spend to complete a project : Nanodegree

Hours spend to complete a project :Associates

Hours spend to complete a project - Bachelors

Hours spend to complete a project -Master

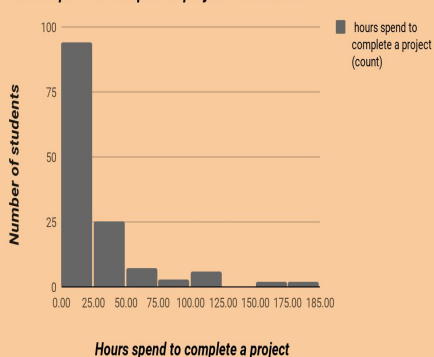**Histogram for project completion hour by:**

Bachelors Vs Masters and Nanodegree vs Association

Both distributions is right-skewed so the mean for each is higher than the median.
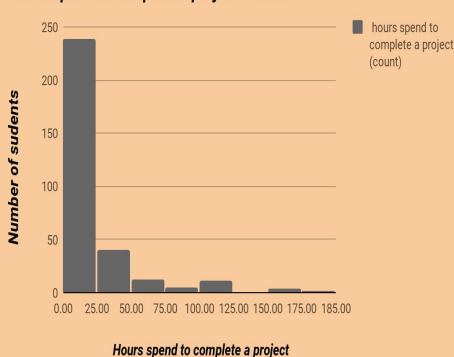
Bachelors and Masters mean for each group does appear to be similar with 9 hours, therefore there is no difference in average project completion hours for Bachelors and Masters students.However, the standard deviation for master is lower with 28 hrs, while bachelor 37 hrs with a high variance that could be due to outliers of more than 150 hours.

Nanodegree and Association mean for each group does appear to be similar with 4 hours, therefore there is no difference in average project completion hours for these students.However the standard deviation for association is lower with 16 hours than Nanodegree with 23 hours therefore has high variance that could be due to outliers of more than 150 hours.
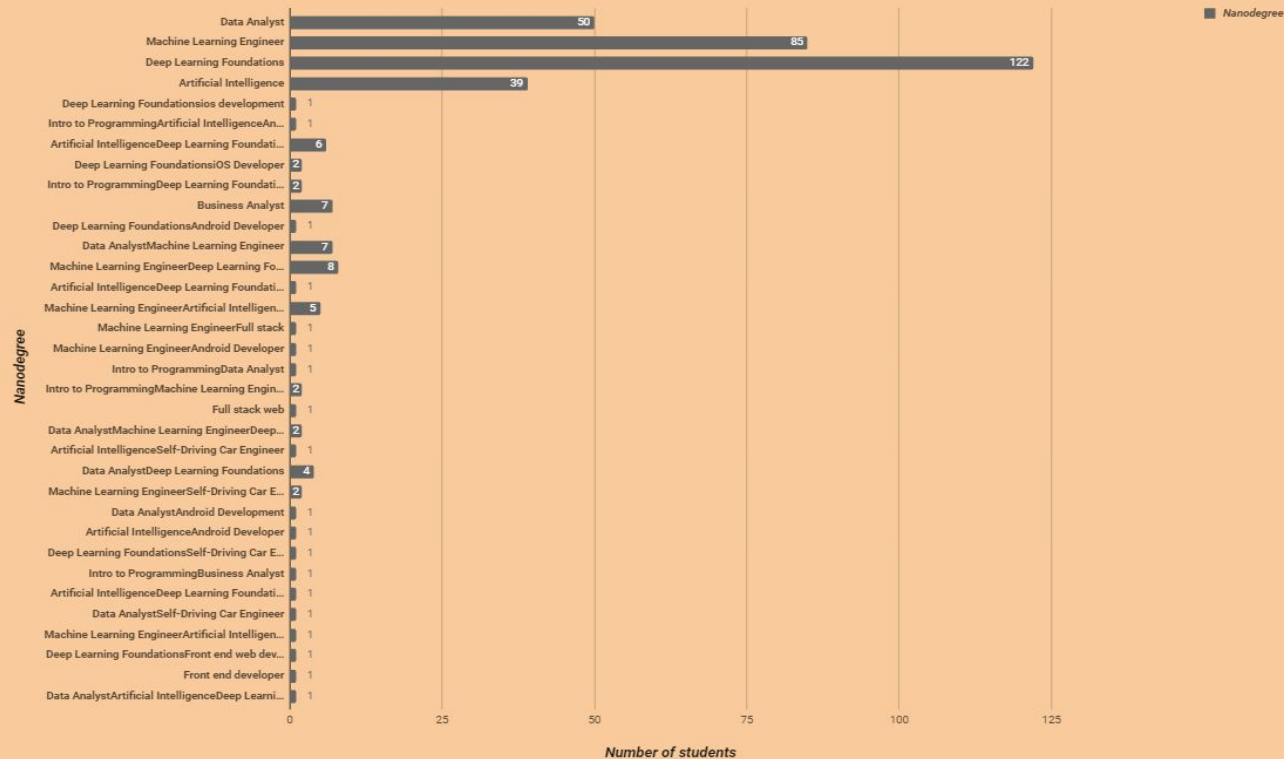
The mode for Bachelors, Masters, Nanodegree and Association is 10 hours,Therefore this means all students spend approximately 10 hours to complete a project.
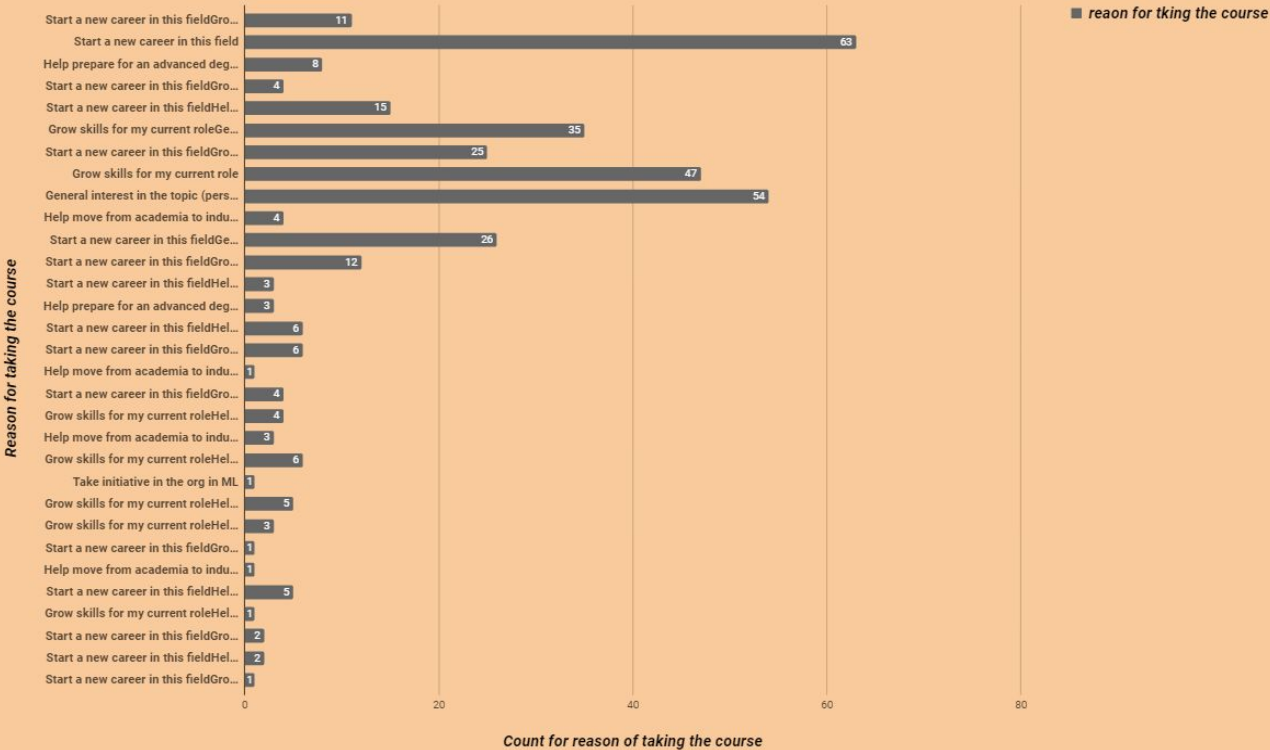
The most demand Nanodegree

Most common Nano-degree

● From the horizontal bar chart,most students chose to do the course **Deep learning foundation**
● The leading most demand nanodegree has 122 counts.

Reason for taking the course

Reason for taking the course

Reason for taking the course	Count
Start a new career in this fieldGro...	11
Start a new career in this field	63
Help prepare for an advanced deg...	8
Start a new career in this fieldGro...	4
Start a new career in this fieldHel...	15
Grow skills for my current roleGe...	35
Start a new career in this fieldGro...	25
Grow skills for my current role	47
General interest in the topic (pers...	54
Help move from academia to indu...	4
Start a new career in this fieldGe...	26
Start a new career in this fieldGro...	12
Start a new career in this fieldHel...	3
Help prepare for an advanced deg...	3
Start a new career in this fieldHel...	6
Start a new career in this fieldGro...	6
Help move from academia to indu...	1
Start a new career in this fieldGro...	4
Grow skills for my current roleHel...	4
Help move from academia to indu...	3
Grow skills for my current roleHel...	6
Take initiative in the org in ML	1
Grow skills for my current roleHel...	5
Grow skills for my current roleHel...	3
Start a new career in this fieldGro...	1
Help move from academia to indu...	1
Start a new career in this fieldHel...	5
Grow skills for my current roleHel...	1
Start a new career in this fieldGro...	2
Start a new career in this fieldHel...	2
Start a new career in this fieldGro...	1

reaon for tking the course

Count for reason of taking the course

- From the horizontal bar chart,most students chose to do the course **to start a new careers in the field,**
- The leading reason has 63 counts.

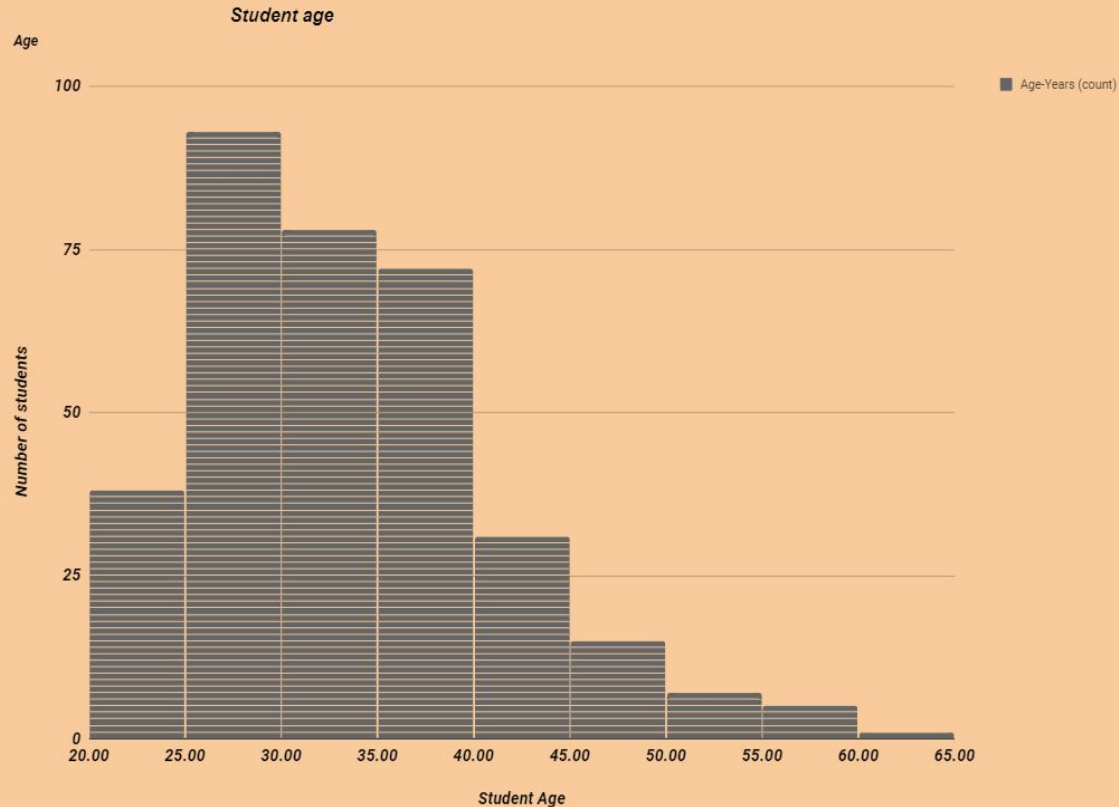# Percentage of Students by Country of origin

Student Country of origin
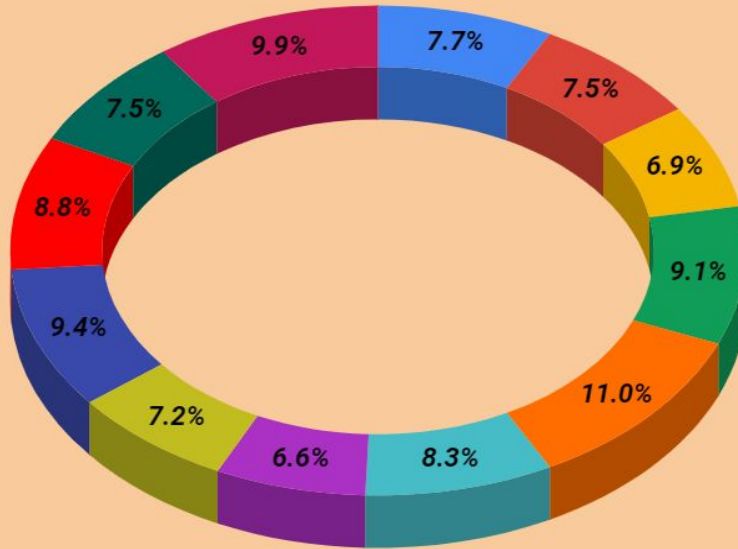
- From the pie chart, the highest percentage of students come from **Argentina** with **11%.** followed by:
- Russia 9.9 % (Second)
- UK 9.4 % (Third)
- Singapore 9.1% (Forth)
- Mexico 8.8% (Fifth)

Highest Student age range

- **Students age from the survey have** mean 34 years similar as medium at 33 years while the mode is 29 years,therefore this is a symmetric normal distribution.

- We can visualise the highest age range is between age **25-30** years.

Relationships on Hours spend to complete project vs Hours of sleep per night,Hours spent sitting per day,Hours per week spend applying skills learned

Hours spend to compete project vs Hours of sleep per night,Hours spend sitting per day,Hours per week spend applying skills learned
Scatter Chart

★ Hours of sleep per night
★ hours spend sitting per day
★ hours per week spend applying skills learned

Frequency

Hours spend to compete project
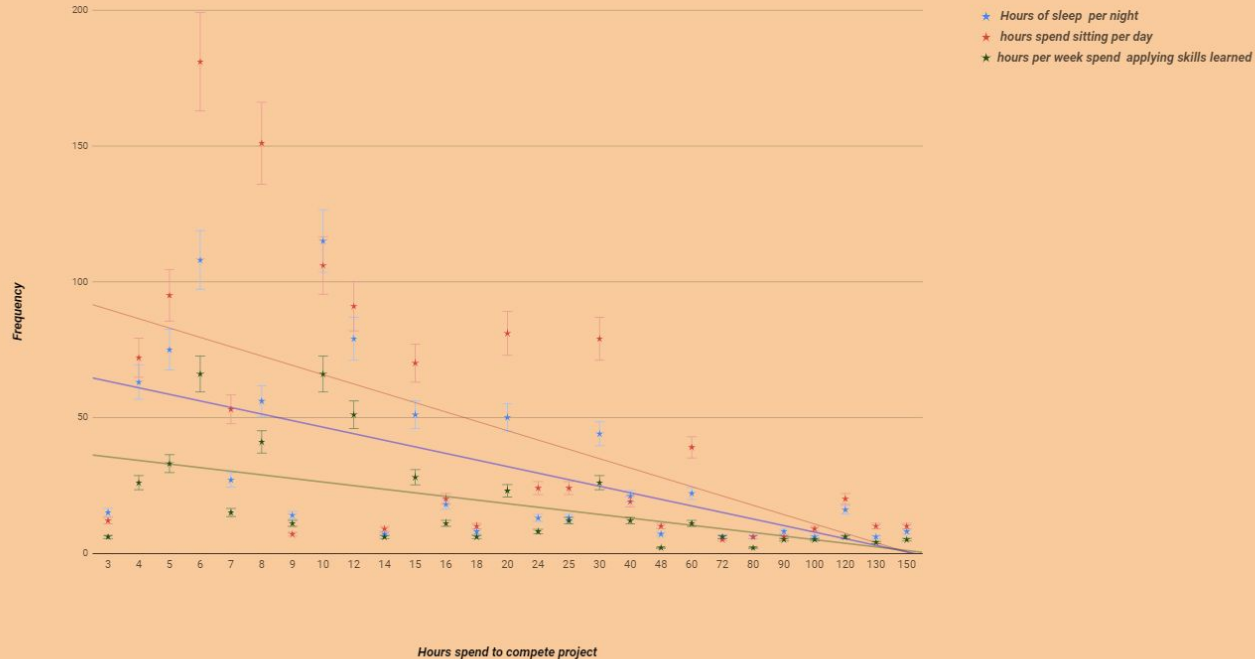
The correlation plot show how four variables for levels of education for **master, bachelors, Association and nanodegree** have no correlated relationship.

As visualized the relationship between Hours spend to compete project and Hours of sleep per night,Hours spent sitting per day, Hours per week spend applying skills learn.

Namely,that Hours spend to complete project increases as Hours of sleep per night,Hours spent sitting per day, Hours per week spend applying skills learn increases.General trend is pretty strong and we can see that these variables are not correlated with Hours spend to compete project.

Therefore, this is a **negative correlation.**

Sleep hours by Education level
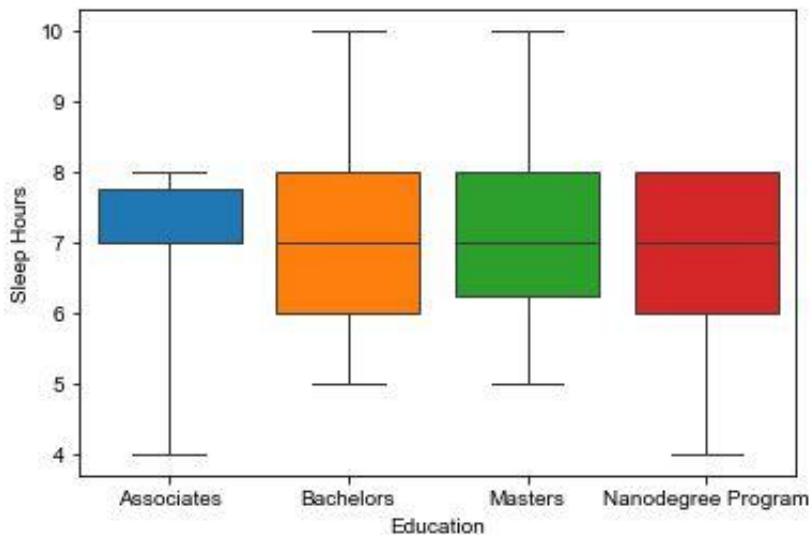
- From the boxplot visualization we can note Associates and nanodegree sleep at a minimum of **4 hour** per night,with a maximum of **8 hours** each.**Q1=7, Q2=7** and **Q3=8 Range=4, IRQ=1** for associates. **Q1=6 ,Q2=6 ,Q3=8 Q3=8 Range=4, IRQ=2** for nanodegree students.Range of

- From the boxplot visualization we can note Bachelors and master sleep at a minimum of **5 hour, 1 hour** per night,respectively, with a maximum of **10 hours** each.**Q1=6, Q2=6** and **Q3=8 Range=5, IRQ=2** for bachelors.**Q1=7 ,Q2=7 ,Q3=5 Q3=8 Range=9, IRQ=1** for master students.

- All students in all education level sleep at the same average of **7 hours** per night.

REFERENCE

Google workbooksheet:
https://docs.google.com/spreadsheets/d/1JLmpFvvubdPxz8MEtayffbRTbCZBuIMJXFqqiSE5sSY/edit?usp=sharing

Github:
https://github.com/Naliaka/UdacityDatascience/blob/master/Datafoudation/Box%20plot.ipynb