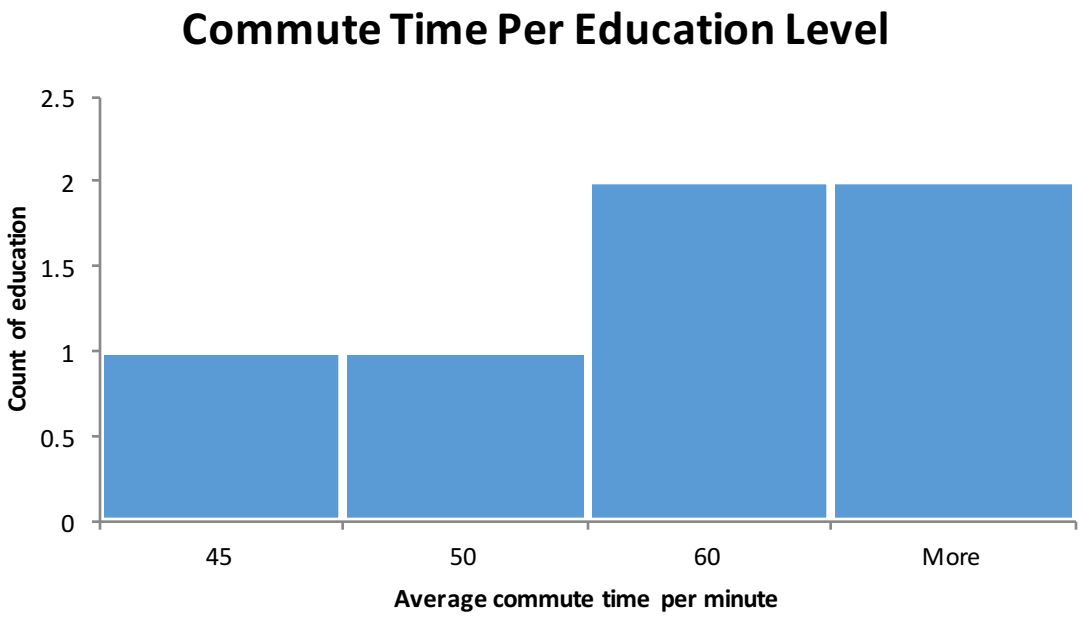
Statistical Questions to be answered

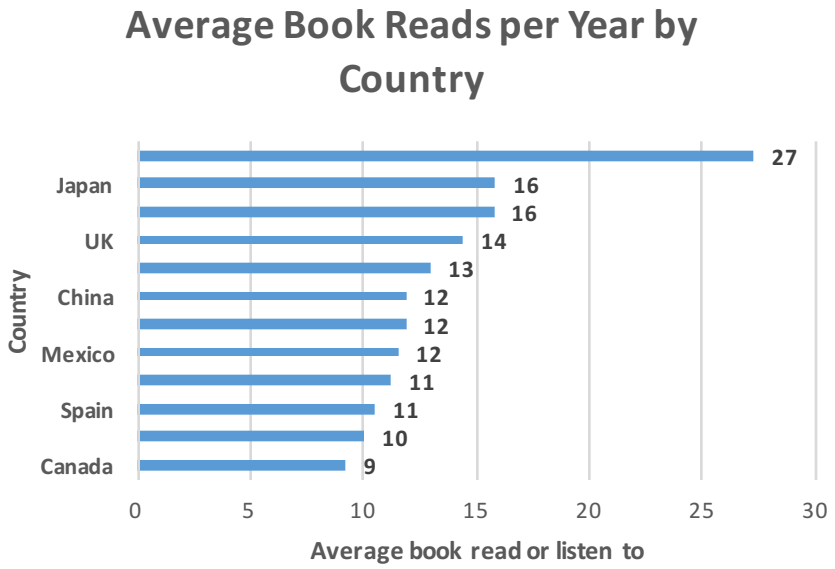
- How much is the average commute time for PHD students?
- How much is the average reading time per country?
- How much is the average time of reading per age ?
- What is the most common way people hear about Udacity ?

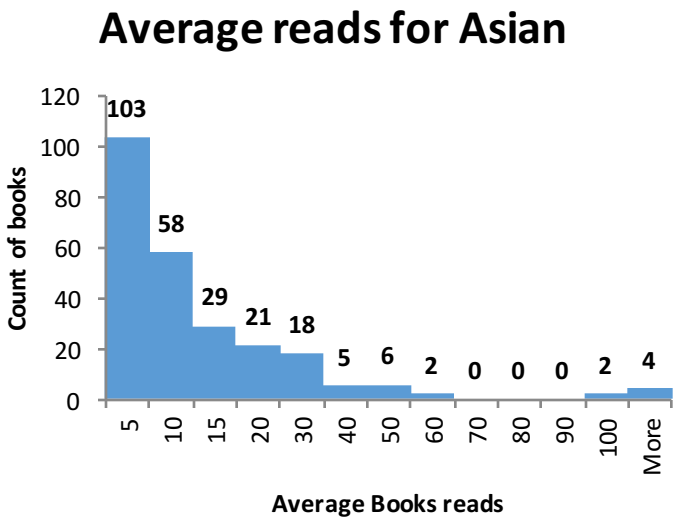
COMMUTE TIME BY EDUCATION

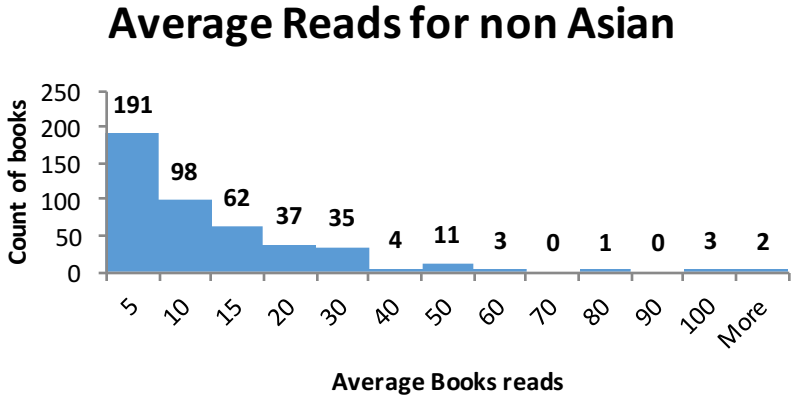


Answering the question of what is the lowest average commute time by education level require me to clean the data by identifying the outliers and replace it with average of commute time of all participants.
Since it is categorical data, I used the histogram to analysis the set. As we noticed, right-skewed distribution which mean that the median is > mean (57 > 56)

NUMBER OF BOOKS READ BY COUNTRY



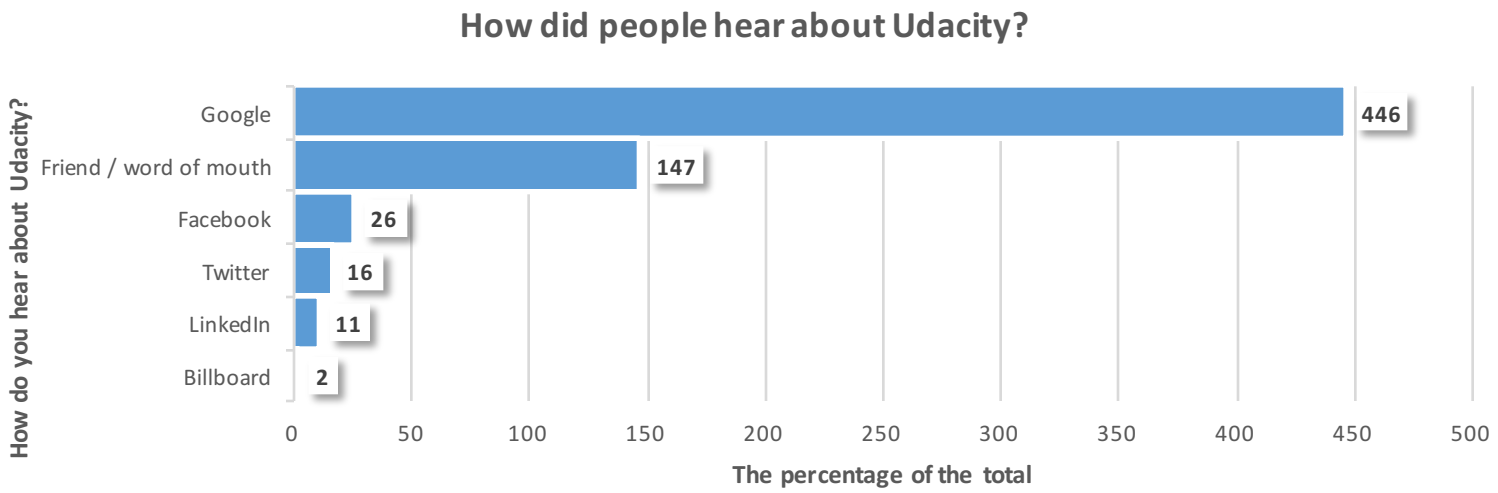


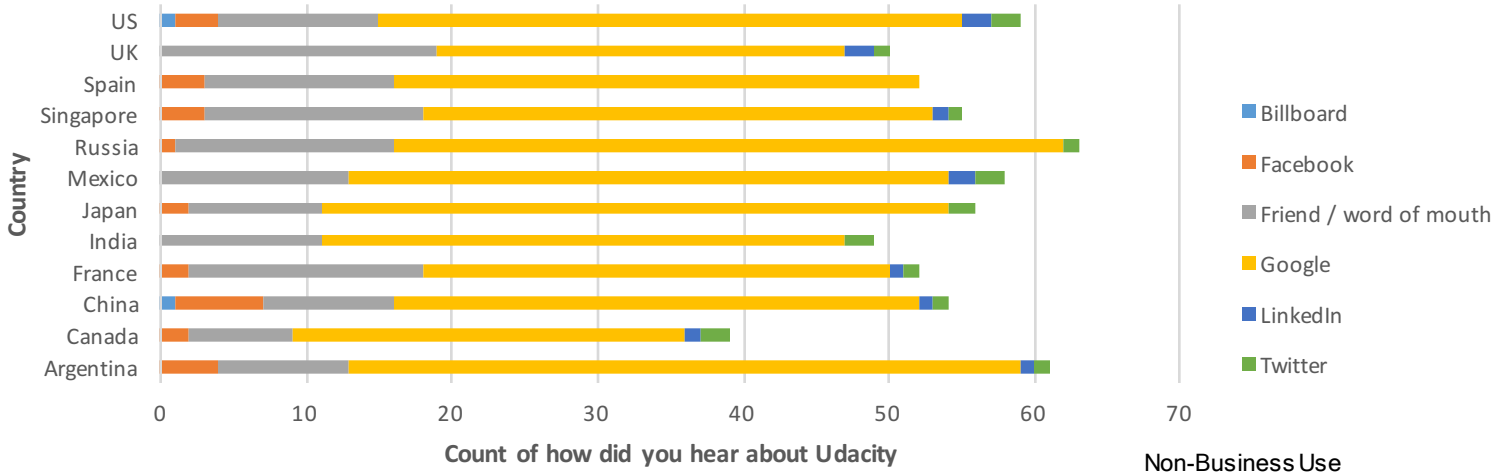


STDV for all readers	29
Mean for all readers	13
STDV for Asian readers	45
STDV for non-Asian readers	17

Answering the question of what is the average number of books read by country require me to clean the data by identifying the outliers and replace it with average of number of books read of all participants. From the chart we can it is right skewed distribution. A right-skewed will have the mean to the right of the median which indicate that median is less than mean. Since the STDV for Asian reader is 45 books and for non-Asian is 17 books after cleaning the data which indicate that the variability in number of books read by Asian is larger. We may conclude that Asian are better reader than non-Asian

THE MOST COMMON WAY PEOPLE HEAR ABOUT UDACITY





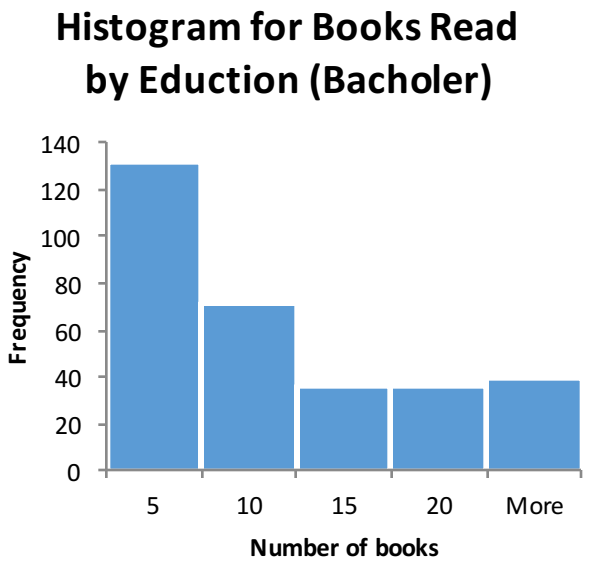
From the chart we can see clearly that students mostly hearing about Udacity from Google. It was interesting to see that 'Words of mouth' has been selected the most from student in France which may indicate Udacity is trend there.

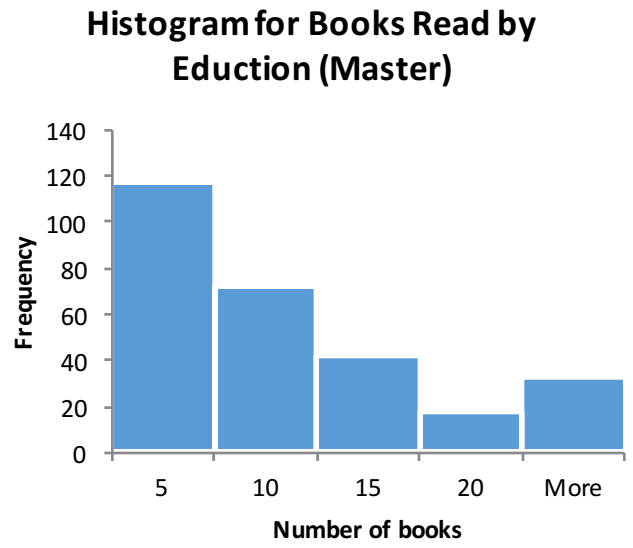
Non-Business Use

NUMBER OF BOOKS READ BY EDUCATION

Number of Books Read

	Bacholer	Master
Minimum	0	0
Q1	4	4
Q2	6.5	8
Q3	15	15
Maximum	120	300





From the chart we can it is right-skewed distribution. A right-skewed will have the mean to the right of the median which indicate that median is less than the mean.