**BSAN53800 Visualizing Information - Assignment 5 (6 points)**

After reviewing Week5 lectures, please complete the following steps.

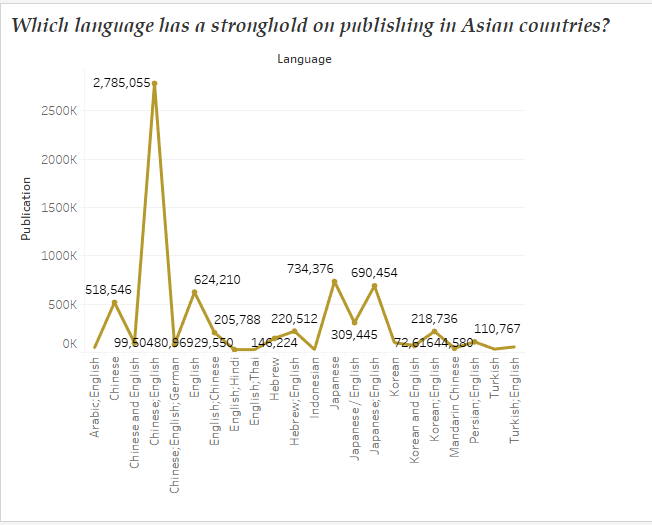
* Download a dataset from any public data sources (e.g., data.gov, data.worldbank.com).
  + You can reuse the dataset from the previous assignments.
* Visualize the multiple variables (3 or more variables) using Tableau.
* Visualize the distribution of one or two variables.
* Optimize the charts based on the principles of good statistical graphics from E.R . Tufte, and describe how you optimized (to optimize) the charts.
* Using the Story function in Tableau, briefly explain any insights you found from the datasets.
* Submit the results in the MS word format.

# SOLUTIONS

Link to the downloaded dataset (loan dataset) - https://www.kaggle.com/code/jadelhelm/r-analysis-eda-cleaning-corrupt-columns/data

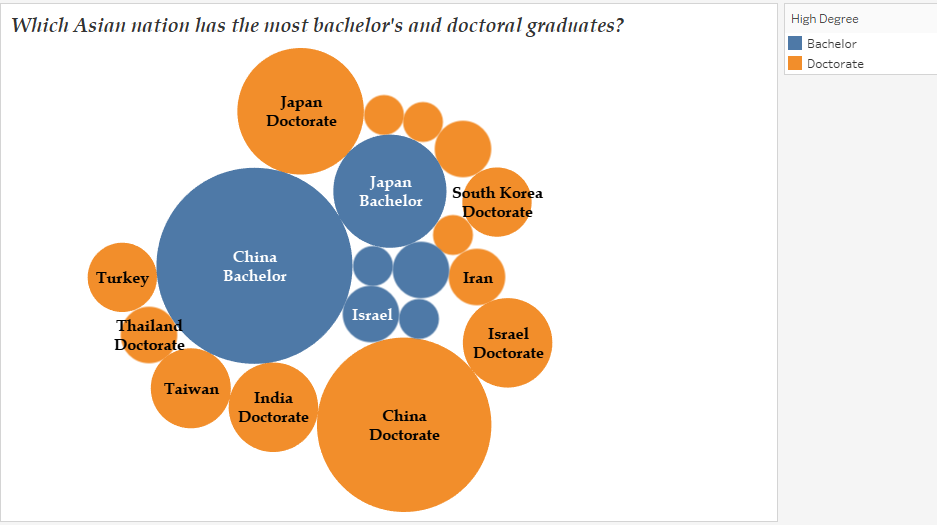
**Asian Top 100 Universities**

**TWO VARIABLES (PUBLICATION VS LANGUAGE)**

****

In the above line graph, we can see how widely spoken Asian languages are. In the publishing industry, it's obvious that Chinese-English pair is at the top of the food chain. But the Turkish-English language pair is the least popular combination in book distribution. Thus, most of the publications are done with the combination of Chinese-English pair.

**MULTIPLE VARIABLES ( COUNTRY VS HIGH DEGREE VS EDUCATION)**

****

We can see which Asian nation is best-known for having the highest concentration of college graduates with advanced degrees by looking at the densely packed bubbles of their educational attainment in the preceding graph. China has the most bachelor's and doctoral degree holders, while Japan comes in second.