Week 9 Homework

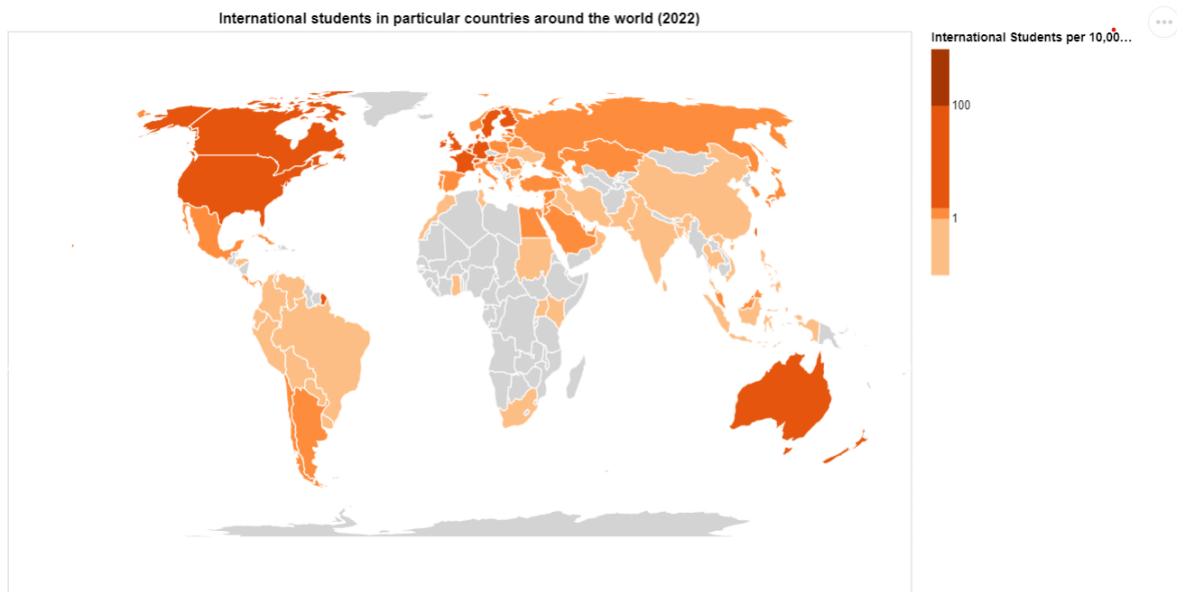**Name:** Jason Marshall
Student ID: 29730341
Lab #: 01_OnCampus
Tutor name: Clair Pan

Publicly accessible web page:

https://jason-marshall-git.github.io/FIT3179_Week9_Homework/

Screen capture of the map:

One short bullet point for each of the following items:
- **The domain of your visualisation**
  - My domain is universities and where they stand and rank among the world. This will be helpful for students looking for universities to study after they finish highschool, or if they are looking to get a degree.
- **The visualised dataset (attribute types, source and author, etc.)**
  - The dataset consists of:

university	year	rank_display	score	country	city	region	type	research_output	student_faculty_ratio	international_students	size	faculty_count
categorical	ordinal	quantitative	quantitative	categorical	categorical	categorical	categorical	ordinal	quantitative	quantitative	ordinal	quantitative

  - The attributes that will be used for the map are country and international_students, so we can see the distribution of international students in universities globally. This dataset was created by Padhma Muniraj and open source on kaggle:

- **Data transformation that you applied (if any), such as normalisation by area or population.**
  - I normalised the data by population. I found a world population dataset and left outer joined it to my current dataset so that I could get the population for each country. I then used the calculate transform function to calculate the international students per 10,000 population in each country. I then created a threshold scale to accommodate for the new transformed data, and added this data to the tooltip when the user scrolls over the chart.
- **A justification for the type of map idiom used. For example, explain why you chose to create a proportional symbol map instead of a choropleth map or a dot map.**
  - I created a choropleth map as it is easier to see the differing international_students values and compare them to the other countries on the map to identify correlations, trends or outliers in the data. I did not create a dot map as this would be more beneficial if I was using a more precise region attribute such as 'city' where it would specify certain positions in each country. However, this wouldn't be too beneficial as each country in the dataset does not include a large amount of city data making it redundant to use, and harder to see on a world map.