



Global Airline Network Analysis

Cai Xue	G1901924K
Cheng Huaihui	G1902364D
Luo Sijia	G1902374B
Tang Yuqun	G1901924K

1. Introduction

In today's globalized world, airplanes are playing an increasingly essential roles as major transportation for human and commercial goods. This advanced transportation technology brings the convenience of trade and boosts mobility of people. Therefore, exploring the global airlines and airports data is not only meaningful but also necessary.

2. Data Source

The 'airlines' data has 6162 records, 8 attributes, namely Airline ID, Name, Alias, IATA, ICAO, Callsign, Country and Active.

Show 5 entries

Search:

	airline_id	name	alias	iata	icao	callsign	country	active
1	-1	Unknown	\N	-	N/A	\N	\N	Y
2	1	Private flight	\N	-	N/A			Y
3	2	135 Airways	\N		GNL	GENERAL	United States	N
4	3	1Time Airline	\N	1T	RNX	NEXTIME	South Africa	Y
5	4	2 Sqn No 1 Elementary Flying Training School	\N		WYT		United Kingdom	N

The 'airports-extended' data has 7750 records, 14 attributes, namely Airport_ID, Airport_Name, City, Country, IATA, ICAO, Latitude, Longitude, Altitude, Timezone, DST, Tz, Type and Source.

Show 5 entries

Search:

	Airport_ID	Airport_Name	City	Country	IATA	ICAO	Latitude	Longitude	Altitude	Timezone	DST
1	1	Goroka Airport	Goroka	Papua New Guinea	GKA	AYGA	-6.08168983459	145.391998291	5282	10 U	1
2	2	Madang Airport	Madang	Papua New Guinea	MAG	AYMD	-5.20707988739	145.789001465	20	10 U	1
3	3	Mount Hagen Kagamuga Airport	Mount Hagen	Papua New Guinea	HGU	AYMH	-5.82678985595703	144.296005249023	5388	10 U	1
4	4	Nadzab Airport	Nadzab	Papua New Guinea	LAE	AYNZ	-6.569803	146.725977	239	10 U	1
5	5	Port Moresby Jacksons International Airport	Port Moresby	Papua New Guinea	POM	AYPY	-9.44338035583496	147.220001220703	146	10 U	1

The 'routes' data has 135326 records, 10 variables, namely airline, airline_id, source_airport_id, destination_airport_id, codeshare, stops, equipment, id, Airport_type and Airport.

Show 5 entries

Search:

	airline	airline_id	source_airport	source_airport_id	destination_airport	destination_airport_id	codeshare	stops	equipment
1	2B	410	AER	2965	KZN	2990		0	CR2
2	2B	410	ASF	2966	KZN	2990		0	CR2
3	2B	410	ASF	2966	MRV	2962		0	CR2
4	2B	410	CEK	2968	KZN	2990		0	CR2
5	2B	410	CEK	2968	OVb	4078		0	CR2

The 'countries of the world' data has 227 records, 20 variables, namely Country, Region, Population, Area, Pop. Density, Coastline, Net migration, Infant mortality, GDP, Literacy, Phones, Arable, Crops, Other, Climate, Birthrate, Deathrate, Agriculture, Industry and Service.

3. Research Questions

Our project focus on airports and airlines of each country. And it is divided into six questions: first of all, the global airport distribution; secondly, the global airline route; thirdly, which country has most airports; fourthly, which country has most airlines; fifthly, the relationship between the number of airports and that of airlines; finally, the airlines between Singapore and other countries.

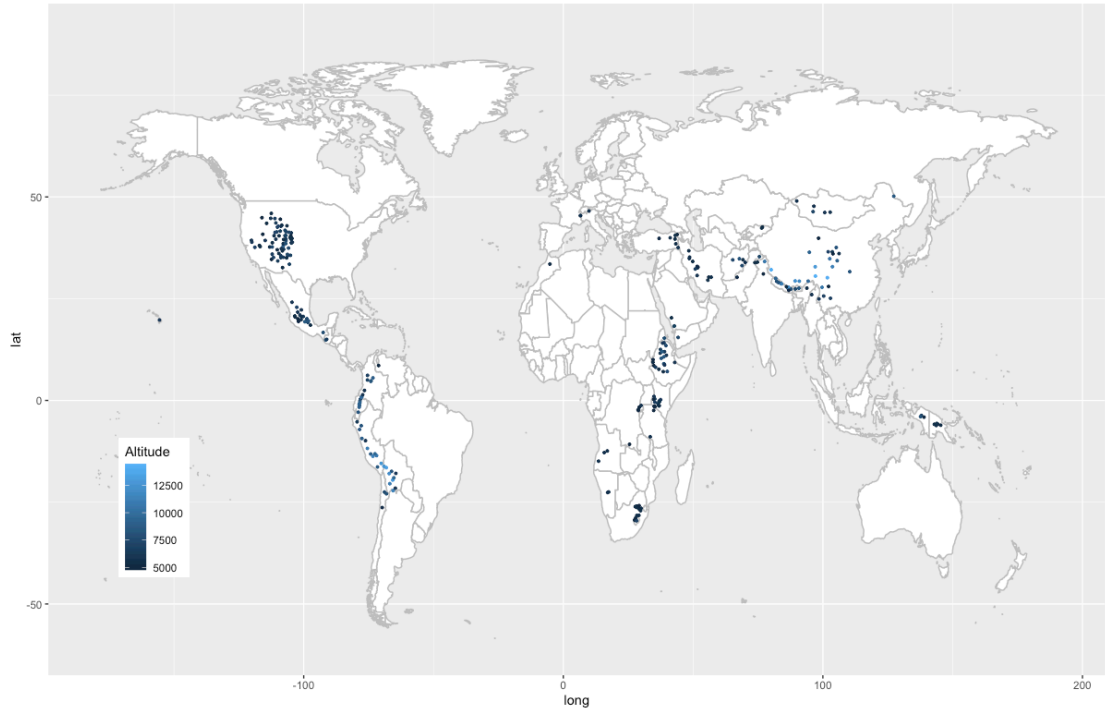
4. Global Airport Distribution



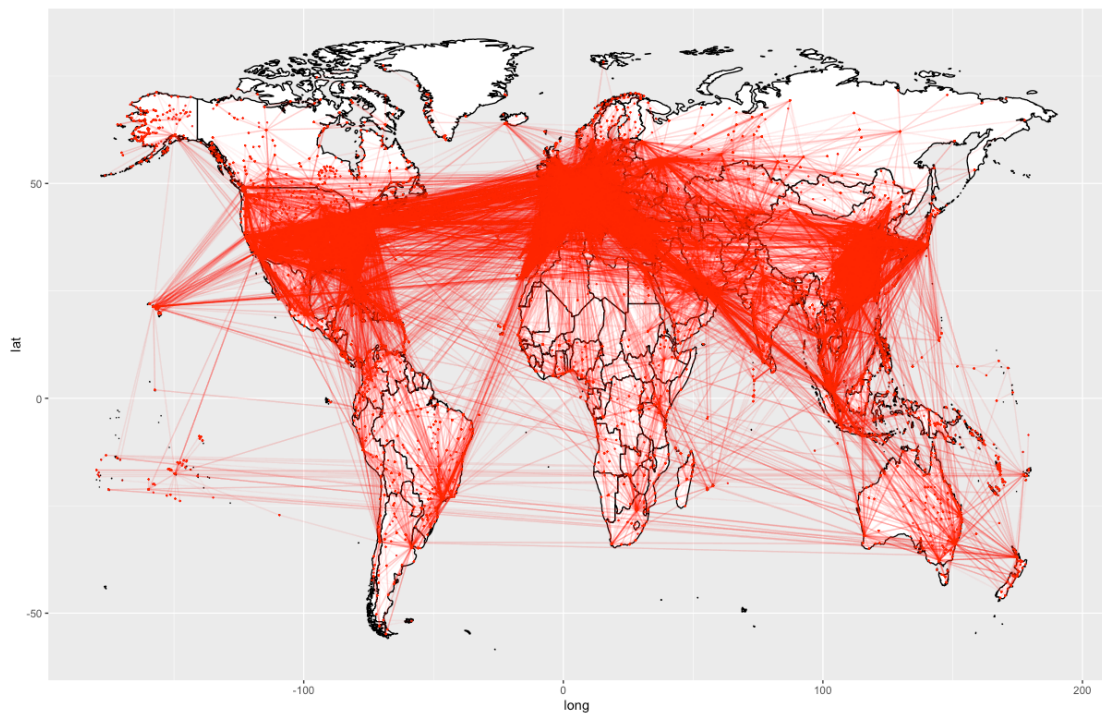
After pre-processing the datasets, we first graph to see the distribution of all those global airports. Based on the map, it's clear to see that, in this half of the earth, the airport density is much higher in Europe than that in Africa, which makes sense. In addition, the graph provides the latitude and longitude coordinate of each airport, as well as the name. Overall, based on the datasets, there are 7750 airports on the earth in the year of 2007.

To further explore our datasets, we want to see the airports that are located over 10,000 feet altitude. Here are the 22 airports that are located over 10,000 feet. They are mainly distributed in the mountainous areas such as Tibet.

Airports located over 5,000 feet altitude



Global Airline Routes

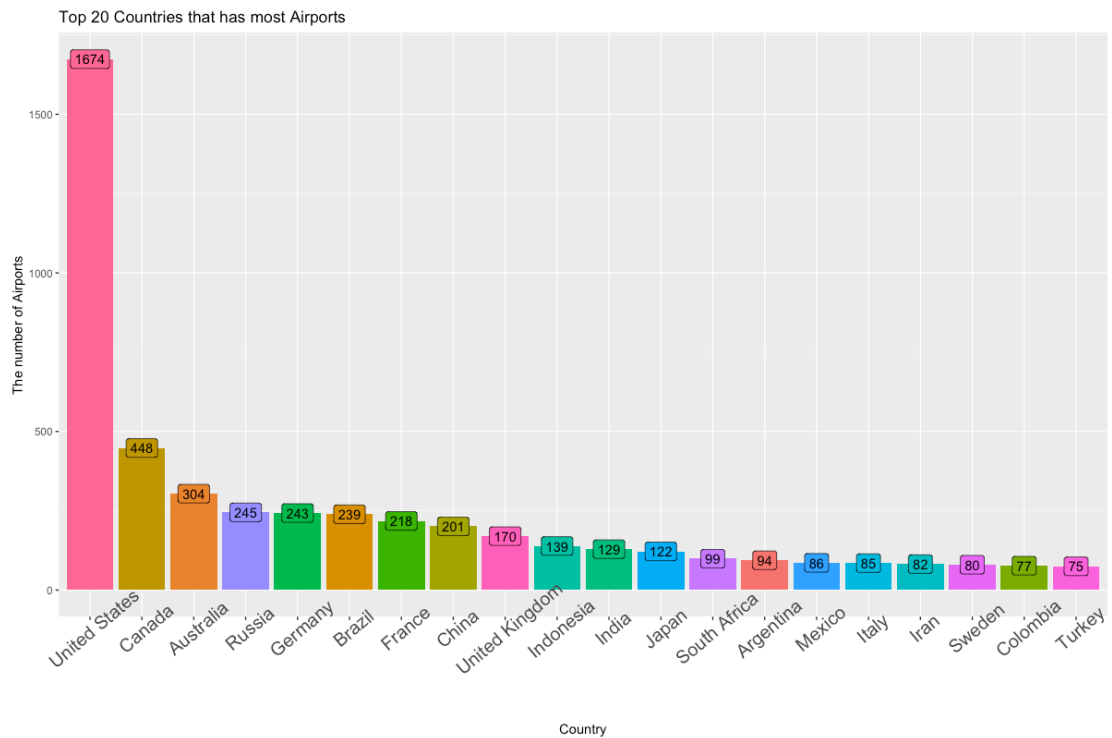


Next, we want to take a look at the global airline routes. To do that, we connected the airports that have direct flights with each other. Based on the map, we can tell that Europe, south America and East Asia have really busy airline traffic. Meanwhile, there are places that still don't have direct flights with each other. For example, there is no direct flight from Japan to Polynesian islands in year 2007.

However, there is one mistake in this map that we can't find out a way to solve it. As we can see from the map, there are no line plots across Pacific Ocean, which is obviously not true. For

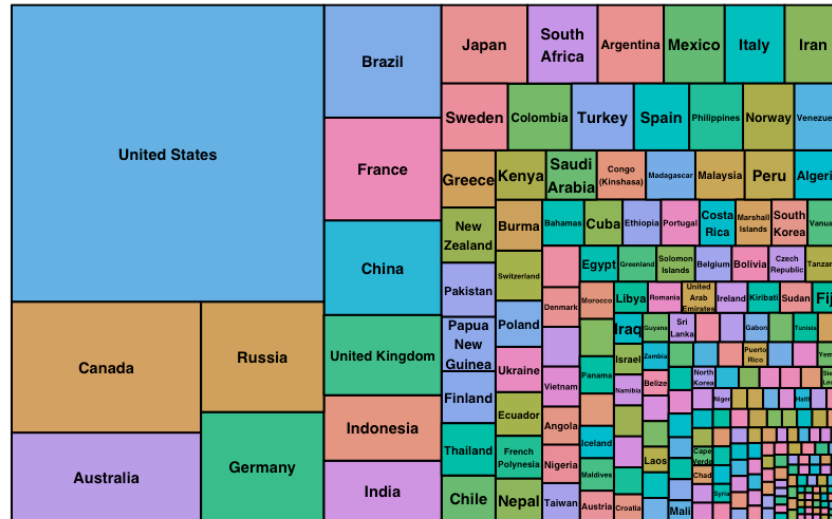
instance, if we take flights from Beijing to the United States, the flight should go across the Pacific Ocean, but it doesn't appear to be that in this map.

5. Which Country has the most Airports?



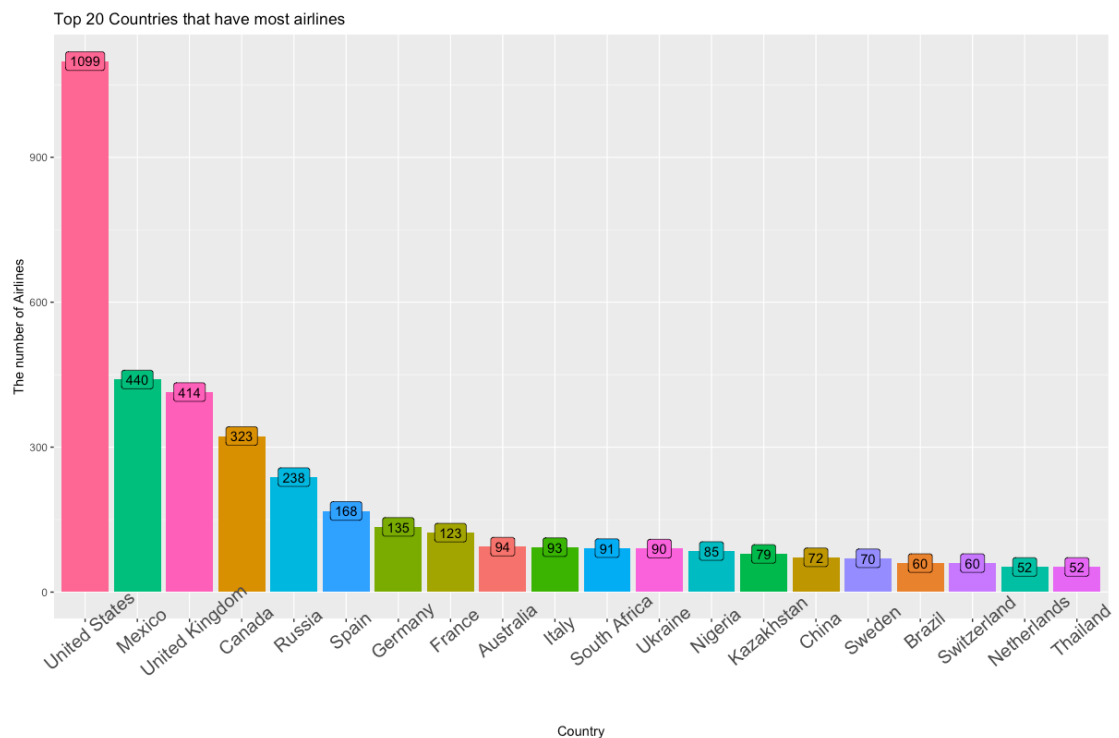
United States has by far the most airports. Probably this is because united states has many military bases around the world. By the way, nations with bigger territories, such as Russia, Canada has many airports because they need them to have access to remote cities. However, small countries such as Japan also ranked in top 20 countries. The number of airports are affected by how large a country is and how good the economy is, we can say.

Overall Number of Airport owned by each Nation



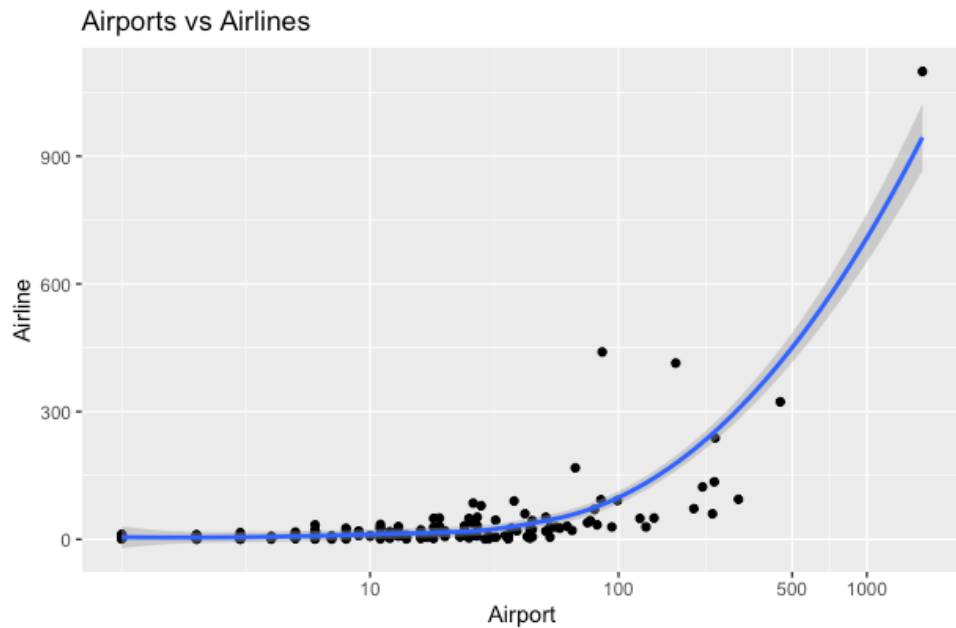
Above is the tree map, analysis is similar to the previous bar chart.

6. Which Country has the most Airlines?



United States has by far the most airlines. However, countries like Japan does not appear in this ranking. Those countries have restricted number of airlines.

7. Airports vs Airlines



From this plot we can find out that there is a positive correlation between the number of airport and airline.

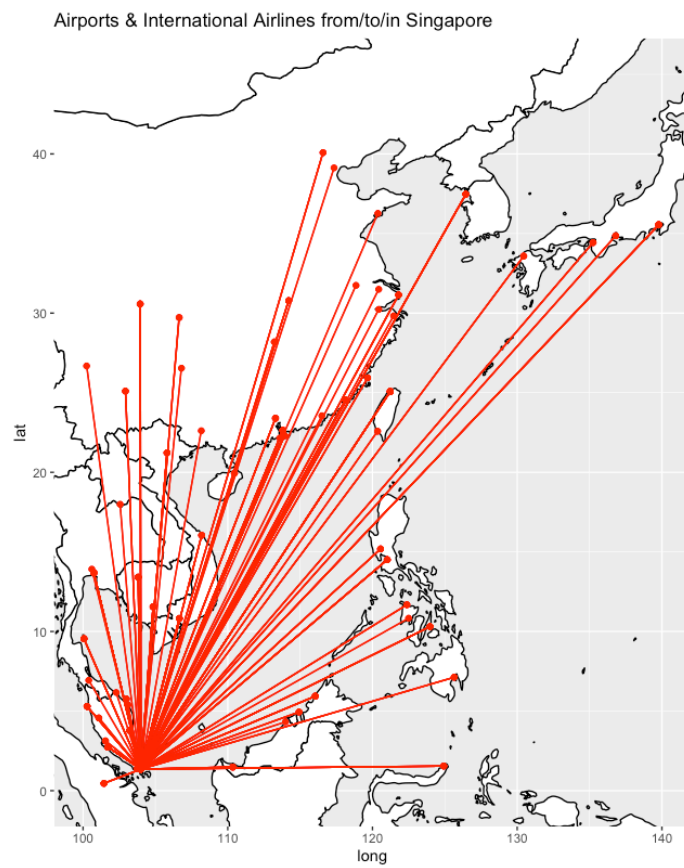
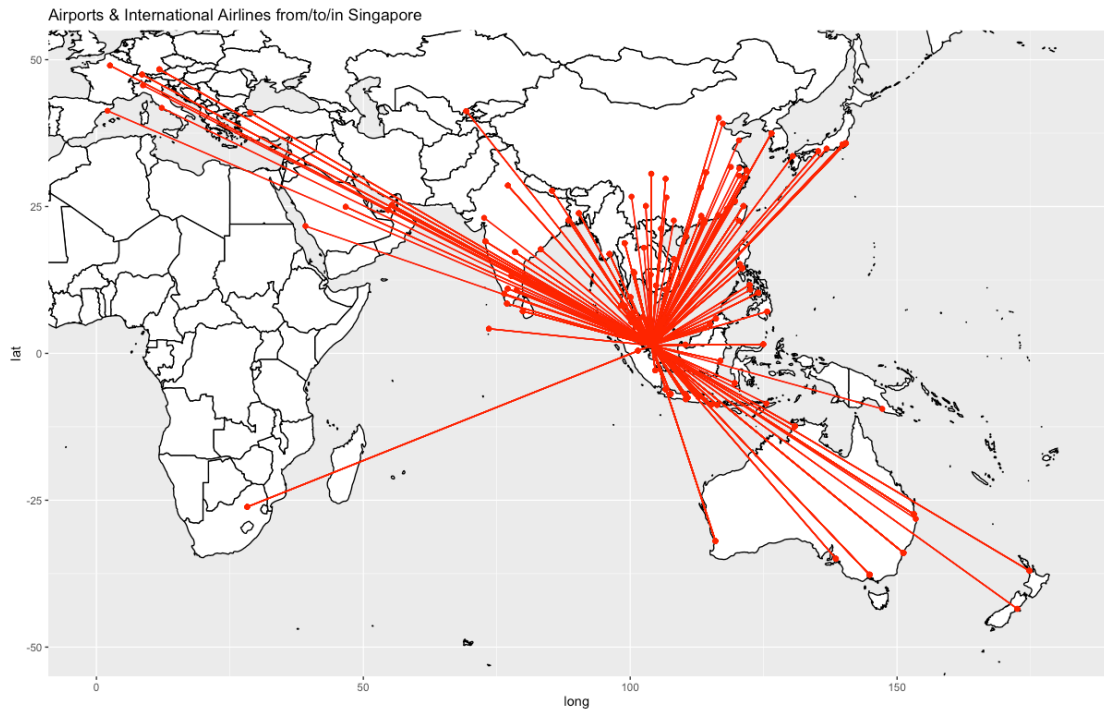
And we can find out several outliers, such as Mexico, which has the second most airlines, but number of airports is not in the top 10. And Australia, with the third most airports, but the number of airline ranks 9th.

8. The flight situation in Singapore

In this part , we are going to focus on global flights engaging Singapore to find some insights After filtering all the airports engaging flight to Singapore, we found there are 121 airports meet the requirement. The figure below show 4 sample airports.

	Airport_Name	Country	City	Latitude	Longitude
1	Adelaide International Airport	Australia	Adelaide	-34.9449996948242	138.531005859375
2	Auckland International Airport	New Zealand	Auckland	-37.0080986023	174.792007446
3	Sardar Vallabhbhai Patel International Airport	India	Ahmedabad	23.0771999359	72.6346969604
4	Amsterdam Airport Schiphol	Netherlands	Amsterdam	52.3086013794	4.76388978958

Using R we can get a visualization of the flights engaging Singapore. The figure below shows the result of the visualization.



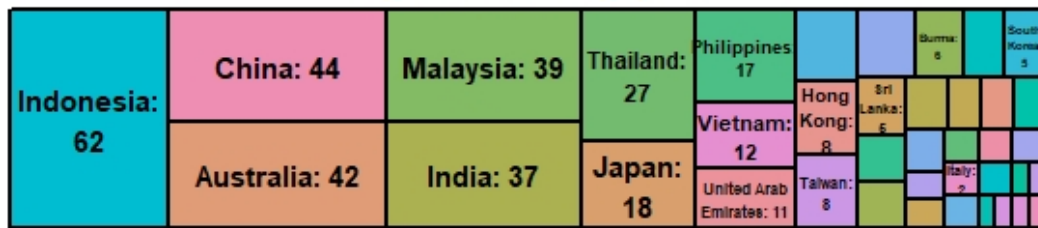
Weight

- 20
- 40
- 60

1. All the flights engaging Singapore are mainly on the Pacific Ocean. Because the dataset was collected years ago and there were not many direct route taking off / arriving Singapore flying over other oceans which means much longer distance.
2. Most flight engaging Singapore are related to countries in Asia ,like India ,Philippines ,China ,Malaysia and Indonesia. That means Singapore has a much tighter relation with the countries in Asia than those in other continents.

Indonesia: 63	China: 45	Australia: 42	India: 37	Japan: 18	Vietnam: 12	Taiwan: 9	South Korea: 6	Sri Lanka: 5
		Malaysia: 39	Thailand: 28	Philippines: 17	United Arab Emirates: 11	Hong Kong: 8		
					Cambodia: 8	Germany: 7		
						Burma: 6		

Flights from Singapore



By the tree map, we can find that the most frequent flight Singapore is the flight between Indonesia and Singapore. The second place is the flights between China and Singapore which has the similar number to the flights between Australia and Singapore. Besides, the number of flights from place A to place B is almost the same as the number of flights from place B to place A which is also in accordance with the common sense.

9. The flight network in Asia

Known from the insights we find out before, we may build a sense that the connection between the countries in Asia is tight and complex. So in this part, we try to find out the flight network for countries in Asia which may help. The picture below shows the visualization result.

