**Vietnam population status 2011-2020**

**I. Data Information**

* Process Step: Search, download, correct, structure and update data
* Attributes - Unit:
* Population density: people/km2
  + Average population: Thousand people
  + Sex ration: Male/100 female
  + Population growth rate: %
  + Labor\_15+: Thousand people
* License: Creative Commons Attribution Share-Alike

**II. User Story**

I made the user story table to analyze the purpose for each position. The best way to analyze, I realized that I put myself as the reader/ the audience. We analyze data to tell and make any person get our idea, not talk to our self.

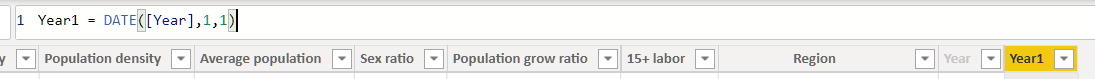
The user guide I am using Agile methodology.

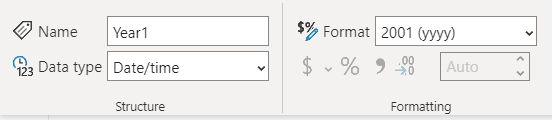
**III. Data Visualization**

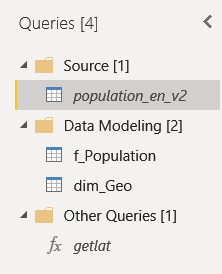
I am using Power Bi to clean data and visualize data. The positive of Power Bi is ***required outer should login to Power Bi service*** before read this. It is not a barrier because I can ***provide my individually account to you***, but I suggest you contact to me via my email or phone number. Futhermore, I willing to open with HR headhunt so please contact to me if you interested.

1. **Data Prepare**

The data source has not much complex , it can modify and prepare data by Power Querry on Power Bi. I create one column name ‘Year1’ in data source by Table Tool to definite ‘Year’ column. Power Bi completely undiscoverd this column is the date because of missing days, month in data source. Restruturing the column is Date/Time and formating the data type for this column is YYYY.After that, I hided the column ‘Year’ in data model.

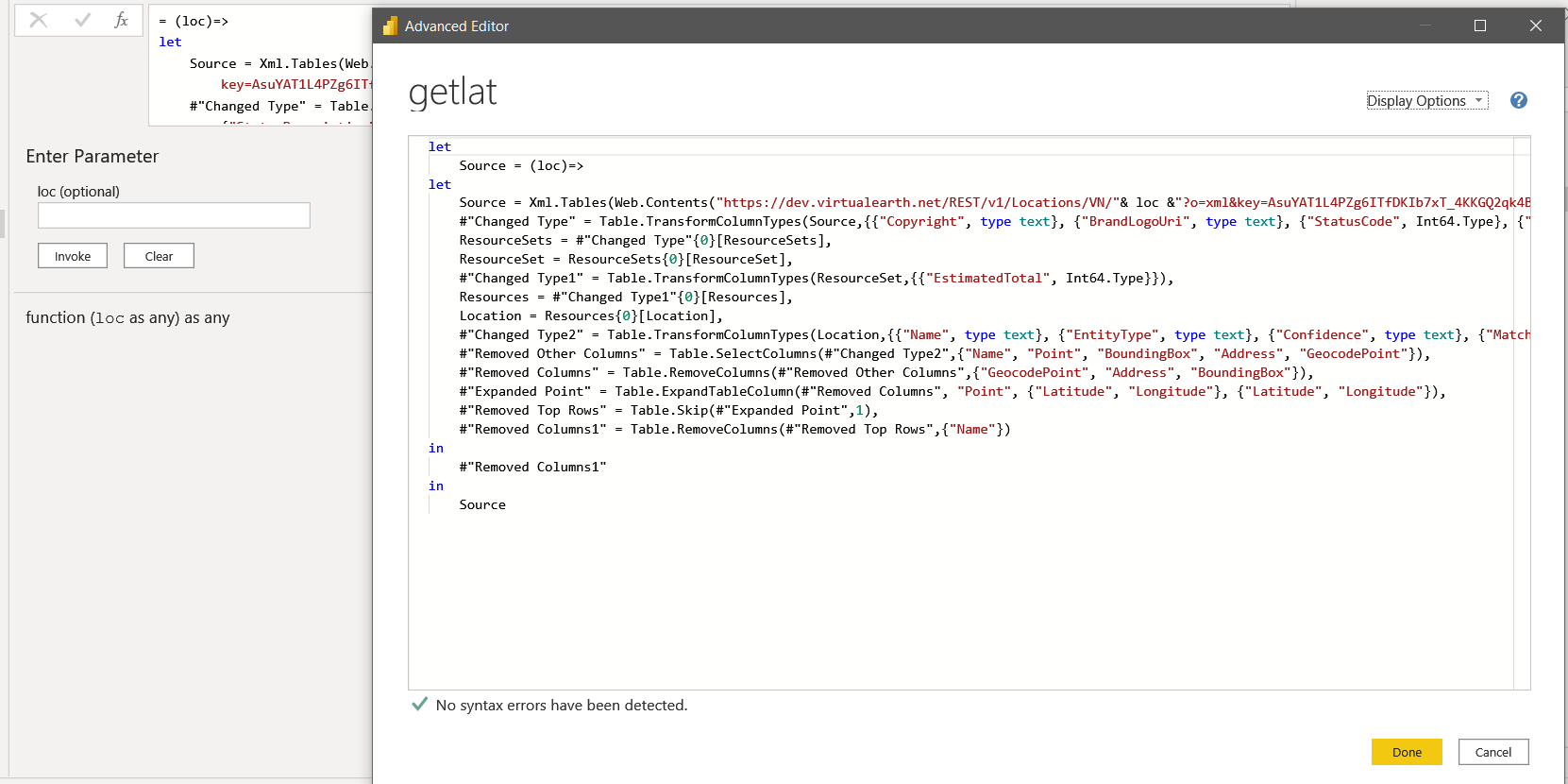






I need created the Vietnam map to provide for user Geography who I discussed in User guide table on Section II User Story. I created the function getting latitude and longtitude for provincies and cities in Vietnam.

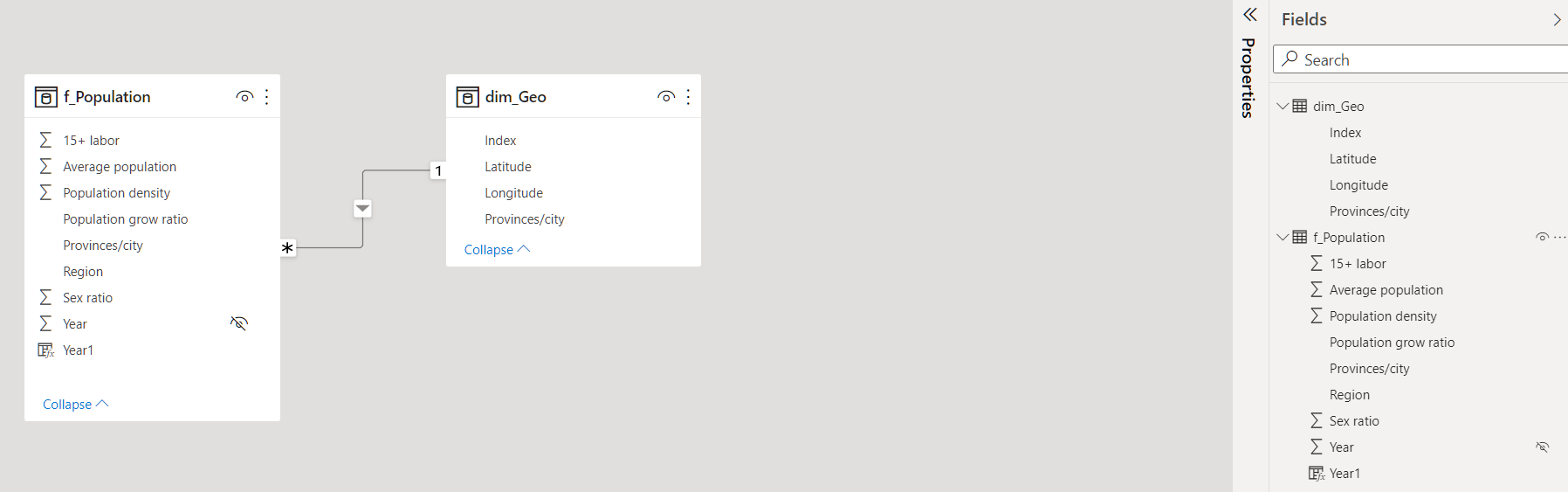
This function using Bing map services. The function using crawling data from Bing map web then transform to statics data. I changed it to function by using DAX code.



1. **Data Modeling**

First, I created data model including fact\_Population table and dimension\_Geo. The f\_population table is collected main fields in data source to analyze data, the d\_Geo created to connect and constrain to f\_population table.

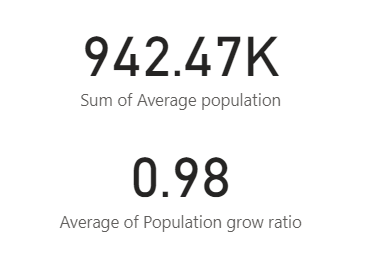
The primary key is ‘Provinces/Cities’ column in fact\_Population and the foreign key is the same name column in dim\_Geo table.



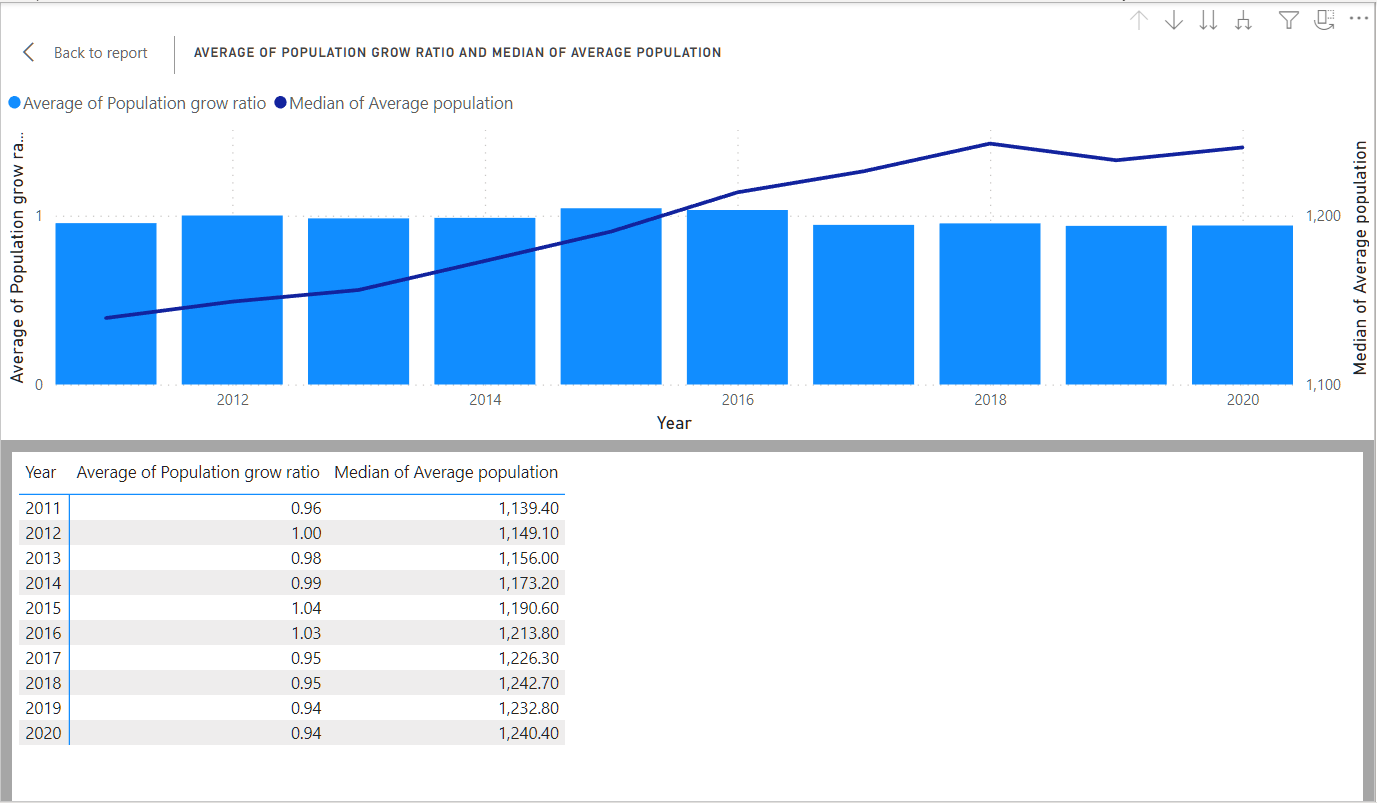
**IV. Data Analysis**

This section I present by using Vietnamese. Actually, I am Vietnamese people so that I proud of speaking Vietnamese and easy for me to depth analysis. Kindly understand for me.

Trong phần này, mình sẽ trình bày bằng ngôn ngữ tiếng Việt. Thực sự, mình là người Việt Nam nên mình tự hào trình bày bằng tiếng Việt và cũng dễ dàng hơn cho mình khi phân tích sâu dữ liệu này.



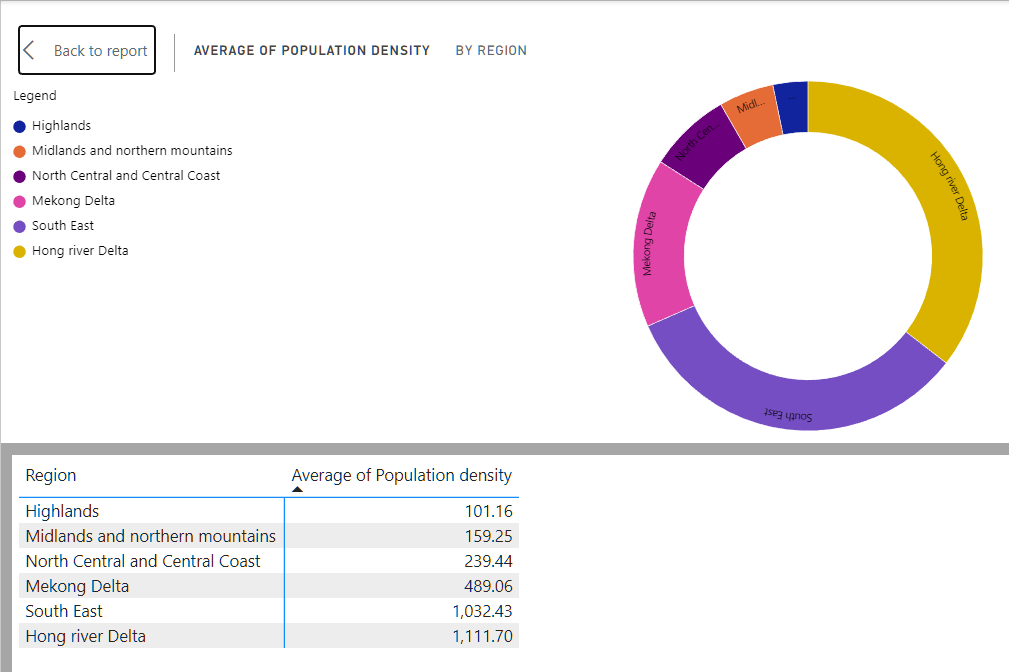
Số liệu cho thấy tổng bình quân dân số tại Việt Nam từ năm 2011 đến năm 2022 là 942.472.000 người tức dân số hiện nay rơi vào khoảng 94.247.000 người. Trong đó tỷ lệ tăng dân số bình quân là 0.98% tức nghĩa rằng cứ mỗi một người sẽ có thêm 0.98 người nữa được sinh ra.



Đánh giá đầu tiên chúng ta có thể thấy rằng cơ cấu dân số Việt Nam suốt từ năm 2011 đến năm 2020 nhìn chung là tăng liên tục qua chỉ số trung vị của bình quân dân số (Median of Average Population) được biểu thị bằng biểu đồ đường trong biểu đồ trên. Năm 2011 đến năm 2020 tăng từ 1,139.40 lên 1,240.40 tức tăng 101.40 tương đương khoảng 1.08 lần. Năm có mức độ trung vị thấp nhất là năm 2014 khi điểm trung vị là 1,139.40 và năm cao nhất là năm 2018 mới điểm trung vị là 1,242.70.

Bình quân tỷ lệ tăng trưởng dân số (Average of Population growth ratio) tăng giảm không đồng đều và có dấu hiệu đi xuống sau khi đạt đỉnh vào năm 2016 với tỷ lệ 1.03%. Trong vòng 5 năm trở lại đây tỷ lệ tăng trưởng dân số bình quân giao động trong khoảng 0.94% đến 0.95%.

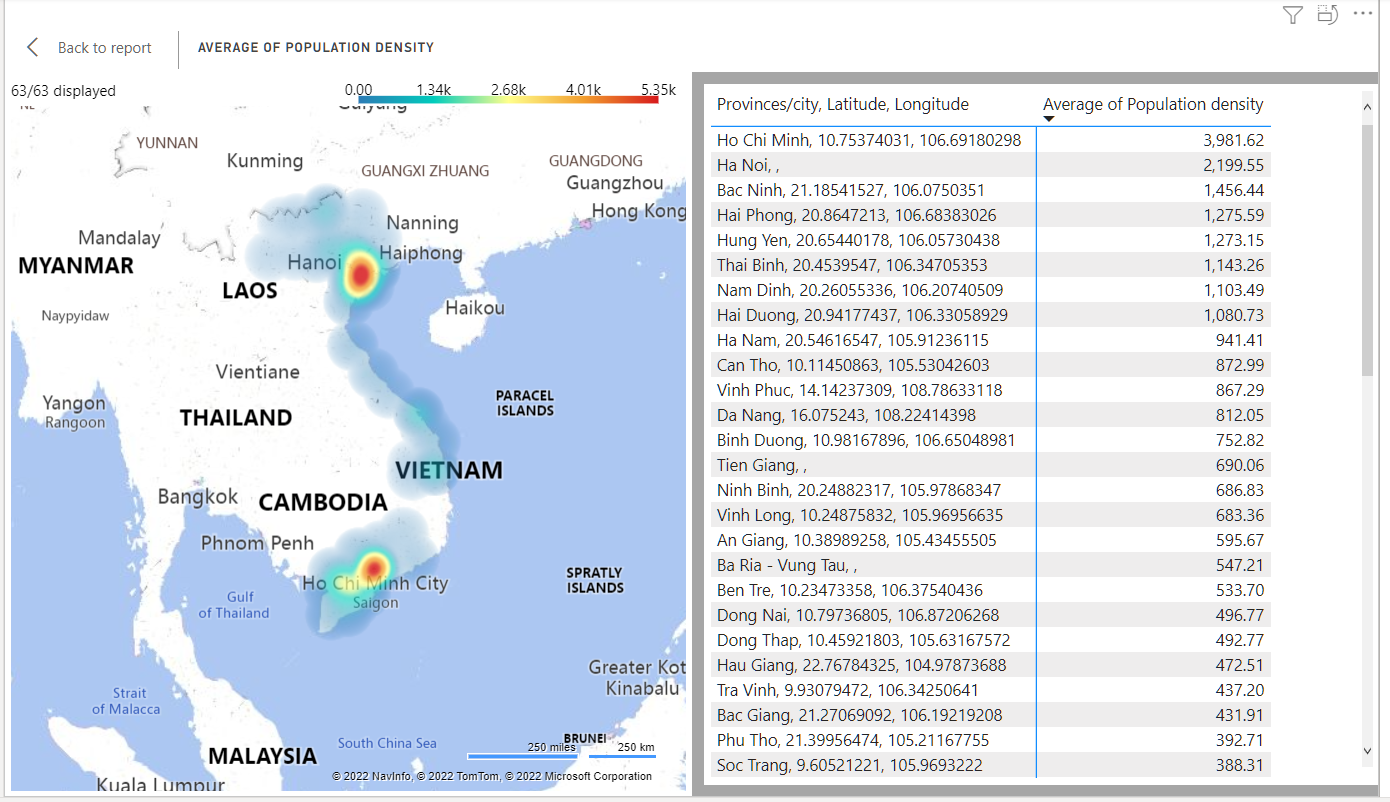
Sự tỷ lệ nghịch của hai chỉ số này cho thấy dân số Việt Nam đang có dấu hiệu chậm phát triển. Điểm trung vị của dân số trung bình vẫn tăng nhưng đó là hệ quả của các thế hệ trước đã và đang phát triển nhưng bình quân tốc độ tăng trưởng dân số lại giảm và đi ngang điều này cho thấy dân số Việt Nam đang ít sinh nở hơn so với thế hệ trước.



Phân bổ dân cư theo vùng miền thì chúng ta tạm chia Việt Nam thành 6 vùng: Đồng bằng sông Hồng (Hong River Delta), Đông Nam Bộ (Southeast), Đồng bằng Sông Cửu Long (Mekong Delta), Tây Nguyên (Highlands), Bắc trung Bộ và duyên hải miền Trung (North Central and Central Coast) và Trung du miền núi Bắc Bộ (Midlands and Northern Mountains).

Dân cư phần lớn tập trung tại 2 vùng lớn nhất nước là Đồng bằng sông Hồng và Đông Nam Bộ cũng là nơi có hai thành phố lớn nhất cả nước là TP.HCM và TP.Hà Nội. Khu vực thưa dân nhất là Tây Nguyên và Trung Du miền núi Bắc Bộ. Nhìn chung khu vực đồi núi vẫn là khu vực khó khăn , kém phát triển về kinh tế, điều kiện sống nên dân cư không tập trung nhiều là điều tự nhiên.

Tiếp theo chúng ta xem xét tiếp mật độ dân số phân bổ theo từng tỉnh và thành phố cụ thể để hiểu rõ hơn vấn đề.

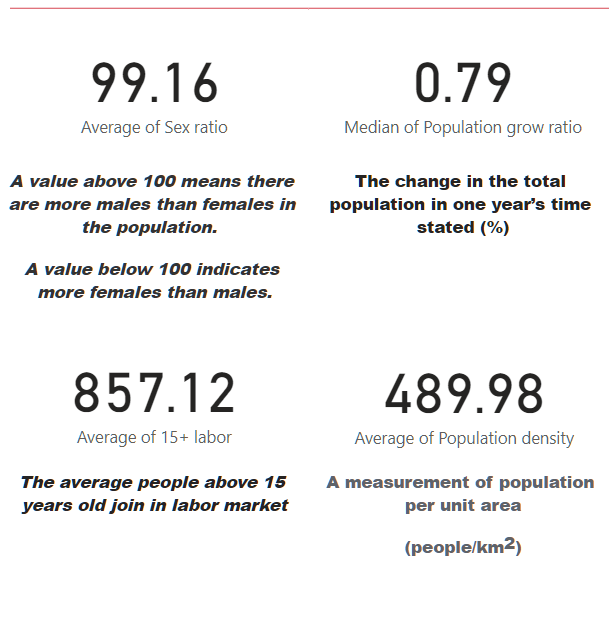


Trên Việt Nam ta hiện nay 5 tỉnh thành phố có mật độ dân cư trung bình lớn nhất là TP.HCM (3.981,62) , Hà Nội (2.199,55) , Bắc Ninh (1.456,44), Hải Phòng (1.275,59), Hưng Yên (1.273,15). Chúng ta có thể thấy rằng mật độ dân số 1.000 người trên 1 km2 là mức rất cao trong khi mức trung bình của Thế giới là 54 người trên 1 km2. Điều này cũng lý giải tại sao đời sống trong các độ thị lớn thường ngột ngạt, kẹt xe, khói bụi.

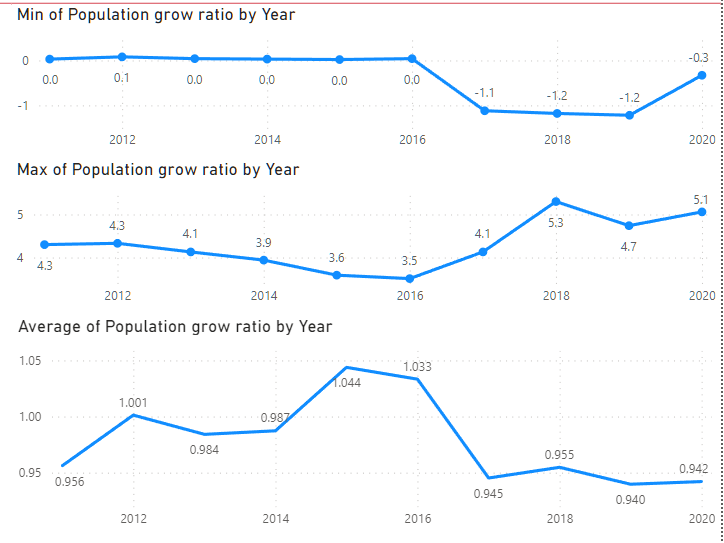
Bên cạnh đó, Việt Nam có 5 tỉnh thành phố có mật độ thấp nhất gồm Lai Châu (47.44), Kon Tum (51.82) , Điện Biên (58.00), Bắc Kạn (64.40), Cao Bằng (78.56). Các tỉnh này đều thuộc vùng trung du miền núi Bắc Bộ với điều kiện và phương tiện sống cũng như khí hậu khắc nghiệt.

Dưới đây là một vài con số tổng quát về tỷ lệ chênh lệch giới tính, số lượng bình quân người lao động trên 15 tuổi, mật độ dân số bình quân, mức trung vị của tốc độ tăng trưởng dân số hằng năm.

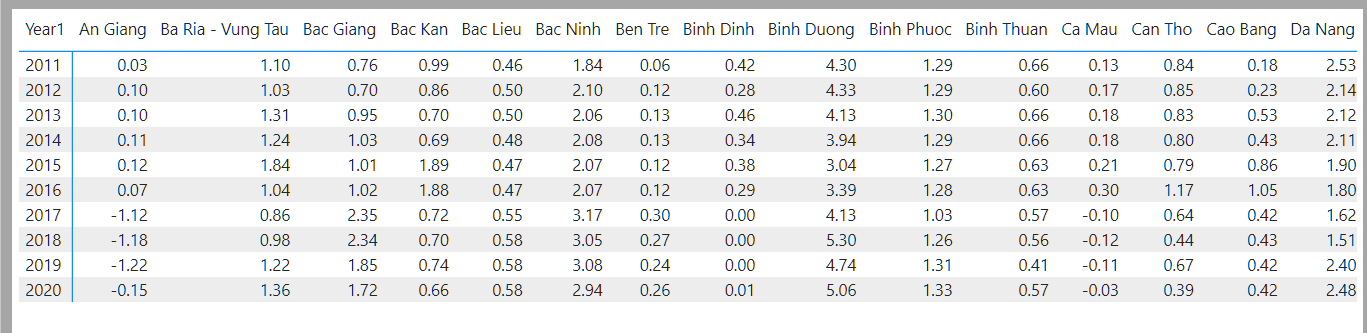
Trong đó có thể thấy tỷ lệ chênh lệch giới tính của Việt Nam đang ở mức ổn định là 99.16 , tức nghĩa là có 99.16 người đàn ông trên 100 người nữ. Nếu tỷ lệ này vượt quá 100 thì đang có sự mất cân bằng giới tính với số lượng người nữ nhiều hơn người nam. Đây cũng là con số tốt khi Việt Nam không có quá nhiều sự chênh lệch giới tính đặc biệt rằng Việt Nam là một nước chịu ảnh hưởng của văn hóa Nho giáo.



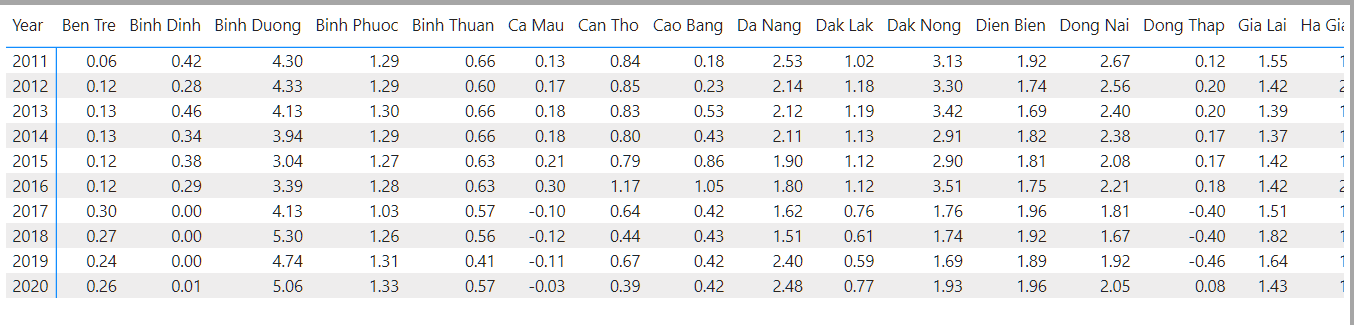
Cuối cùng chúng ta cùng xem xét ba biểu đồ về tỷ lệ tăng trưởng dân số theo mức độ thấp nhất, cao nhất , trung bình qua số liệu mỗi năm như sau.



Trong biểu đồ thể hiện mức độ thấp nhất trong tăng trưởng dân số hằng năm ta có thể thấy tốc độ tăng trưởng dân số đã từng là -1.2% vào năm 2018 tại một số tỉnh thành trên Việt Nam. Bằng một thao tác nhỏ trên biểu đồ ta có thể tìm ra ngay được tỉnh thành phố có tỷ lệ tăng trưởng dân số -1.2% vào năm 2018 là tỉnh An Giang.



Nhưng ở biểu đồ ngay bên dưới cũng vào năm 2018 cho thấy tốc độ tăng trưởng dân số đạt đỉnh cao với tỷ lệ 5.3% , đây là tỷ lệ tăng trưởng dân số cao nhất tại tỉnh Bình Dương vào năm 2018.



Tại biểu đồ cuối cùng ta có thể thấy năm 2019 là năm có tỷ lệ tăng trưởng dân số bình quân thấp nhất trên cả nước với tỷ lệ 0.940% và năm 2014 cả nước ghi nhận bình quân tỷ tăng trưởng dân số cao nhất với tỷ lệ là 1.04%.

Thông qua toàn bộ những phân tích trên ta có thể thấy cơ cấu dân số Việt Nam là cơ cấu dân số đang dần chậm phát triển hơn so với thời kỳ trước với các tỷ lệ tăng trưởng dân số giảm dần trong 5 năm gần nhất từ năm 2018 đến năm 2020. Tỷ lệ cân bằng giới tính ở Việt Nam là bình thường, cơ cấu dân số lao động trên 15 tuổi của Việt Nam ở mức tốt. Tuy nhiên Việt Nam cần cải thiện hơn nữa mức độ phân bổ dân số theo vùng và tỉnh thành phố để phát triển đồng đều toàn vẹn và phát huy sức mạnh của Việt Nam và con người Việt Nam.