ĐẠI HỌC BÁCH KHOA HÀ NỘI

TRƯỜNG CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

──────── \* \* \* ────────



BÀI TẬP LỚN

QUẢN TRỊ DỮ LIỆU VÀ TRỰC QUAN HÓA

**PHÂN TÍCH DỮ LIỆU DÂN SỐ VIỆT NAM**

Lớp : 136856

Giáo viên hướng dẫn : TS. Trần Việt Trung

Nhóm sinh viên : Nhóm 8

|  |  |
| --- | --- |
| Họ và tên | MSSV |
| Đào Trọng Hiếu | 20183739 |
| Trần Phú Xuân | 20194357 |
| Bùi Trọng Huy | 20183765 |

***Hà Nội, tháng 2 năm 2023***

MỤC LỤC

[PHÂN CÔNG CÔNG VIỆC 5](#_Toc126808694)

[CHƯƠNG 1. GIỚI THIỆU 7](#_Toc126808695)

[1.1. Động lực 7](#_Toc126808696)

[1.2. Nguồn dữ liệu 7](#_Toc126808697)

[1.3. Công nghệ 8](#_Toc126808698)

1.4. Cách sử dụng……………………………………………………………………

[CHƯƠNG 2. TRỰC QUAN HÓA 9](#_Toc126808699)

[2.1. Trang Overview 9](#_Toc126808700)

[2.2. Trang Khảo sát dân số 10](#_Toc126808701)

[CHƯƠNG 3. NHẬN XÉT 13](#_Toc126808703)

[3.1. Dân Số Phân bổ 13](#_Toc126808704)

[3.2. Tốc độ tăng dân số 17](#_Toc126808705)

[KẾT LUẬN 19](#_Toc126808706)

LỜI NÓI ĐẦU

*Xã hội ngày này, giữa dân số và sự phát triển có mối quan hệ khăng khít với nhau, vừa dàng buộc vừa hỗ trợ lẫn nhau. Muốn tăng trưởng và phát triển kinh tế thì yếu tố cơ bản là nguồn lực mà nguồn lực thì gắn liền với sự biến đổi dân số, đồng thời sự biến đổi dân số còn gắn liền với sự phát triển về nguồn lực.*

*Để tiếp cận với học phần “Trực quan hóa dữ liệu” chúng em lựa chọn đề tài “Phân tích dữ liệu dân số Việt Nam từ năm 2011-2021”. Dữ liệu được tải ở Tổng cục thống kê, được bọn em tiền xử lý rồi đưa vào vẽ biểu đồ.*

*Project này được chia thành 3 phần: Giới thiệu (Giới thiệu về động lực, nguồn dữ liệu, công nghệ, …), Trực quan hóa (Mô tả, giải thích các bảng biểu), Nhận xét (Những tri thức thu được từ dữ liệu).*

*Do thời gian và khả năng có hạn nên không thể tránh khỏi sai sót. Vì vậy nhóm chúng em mong có thể nhận được những đóng góp thẳng thắn từ phía thầy và các bạn.*

*Cuối cùng, nhóm 8 xin chân thành cảm ơn thầy* ***Trần Việt Trung*** *đã hướng dẫn chúng em trong suốt thời gian làm báo cáo này.*

*Trân trọng*

*Nhóm 8*

PHÂN CÔNG CÔNG VIỆC

**Bảng phân công công việc**

|  |  |  |  |
| --- | --- | --- | --- |
| Họ và tên | MSSV | Công việc thực hiện | Đánh giá |
| Đào Trọng Hiếu | 20183739 | * Nhóm trưởng * Làm slide và thuyết trình * Trực quan hóa dữ liệu | Hoàn thành tốt công việc được giao |
| Trần Phú Xuân | 20194357 | * Tìm nguồn dữ liệu * Thu thập, tiền xử lý dữ liệu | Hoàn thành tốt công việc được giao |
| Bùi Trọng Huy | 20183765 | * Thu thập, tiền xử lý dữ liệu * Trực quan hóa dữ liệu | Hoàn thành tốt công việc được giao |
| Cả nhóm | | * Viết báo cáo | Hoàn thành tốt công việc được giao |

# GIỚI THIỆU

## Động lực

Sau mấy chục năm thực hiện quyết liệt chính sách dân số - kế hoạch hóa gia đình, công tác dân số đã đạt được nhiều thành tựu quan trọng - được Liên hợp quốc đánh giá là một quốc gia thành công trong công tác dân số và xóa đói giảm nghèo. Việt Nam trở thành một khuôn mẫu trong các nước đang phát triển thực hiện xuất sắc các công tác này. Tuy nhiên, những thành công về công tác dân số của mới chỉ là bước đầu và chưa bền vững. Những thành công ấy mới thể hiện ở những con số thống kê cân đo đong đếm về số lượng chứ chưa phải là những chỉ báo tương quan các tiêu chí, các mức độ về chất lượng.

Chính vì lẽ đó, trong bối cảnh toàn cầu hóa và hội nhập quốc tế, cần nhận diện một cách chính xác thực trạng dân số nước ta hiện nay và những nét đặc trưng cơ bản của nó, để có thể đưa ra những dự báo đúng về xu hướng biến đổi dân số trong thời gian tới, và từ đó, xây dựng một chiến lược dân số mới phù hợp.

Nhóm chúng em quyết định thu thập và trực quan hóa dữ liệu đề tài này để tìm ra những tri thức thú vị về tình hình dân số Việt Nam trong 10 năm kể từ 2011 đến 2021.

## Nguồn dữ liệu

Nguồn dữ liệu được bọn em lấy ở “Tổng cục thống kê”, là 1 trang web của chính phủ nghiên cứu về các đề tài an sinh xã hội của Việt Nam. Có rất hiều dữ liệu nhưng chúng em sẽ sử dụng các dữ liệu chính sau:

* Dân số phân theo địa phương
* Tống độ phát triển dân số của địa phương
* Dòng người nhập cư
* Dòng người di cư
* Tỉnh lệ trẻ được sinh ra và số lượng người chết đi



Hình 1 – Trang web Tổng cục thống kê

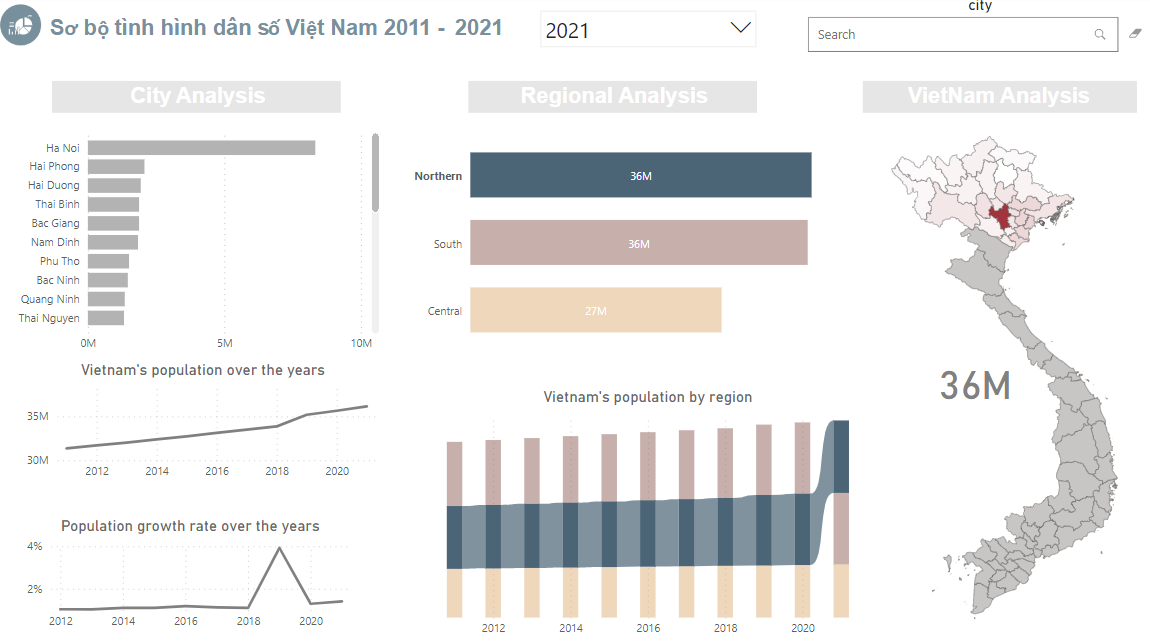
## Công nghệ

Nhóm sử dụng những công nghệ chính sau:

* Python để tiền xử lý dữ liệu
* Phần mềm Power BI để hỗ trợ xử lý và trực quan hóa dữ liệu với số lượng lớn

## Cách sử dụng.

Hầu hết tất cả các vị trị trên Dashboard đều là filter. Muốn xe kĩ phần nào thì có thể click chuột và vùng đó sẽ hiển thị dữ liệu riêng của vùng đó. Ví Dụ:



**Hình 2- Hình ảnh sau khi click và vùng Northern**

* Có 2 thanh chọn năm và chọn thành phố để show ra các biểu đồ tương ứng được chọn.
* Các biểu đồ đường ban đầu sẽ hiển thị dữ liệu của Việt Nam, khi tùy chọn vào các vùng sẽ hiện thị dữ liệu của riêng vùng đó.

# TRỰC QUAN HÓA

Toàn bộ phần trực quan hóa được chia thành 2 trang:

* Trang Overview: Cho cái nhìn tổng quan về dân số
* Trang Khoản sát ra tăng dân số: Cho biết về tình trạng ra tăng dân số của địa phương

Sau đây nhóm em sẽ đi qua từng trang một và giải thích cách sử dụng.

## Trang Overview



Hình 2 – Trang Overview

Trang Overview sẽ bao gồm 6 biểu đồ. Về cơ bản trang sẽ chứa các thông tin về tình hình dân số trong 10 năm trở lại của Việt Nam. Gồm 3 phần chính từ trái sang:

* Dân số theo thành phố
* Dân số theo vùng miền ( Bắc, Trung Nam)
* Dân số cả nước

Dòng đầu tiền bao gồm tiêu đề , 1 slicer, và 1 thanh tìm kiếm. Slicer để chọn năm và thanh tìm kiếm để lựa chọn thành phố.

Phần City Analysis: Bao gồm 3 loại biểu đồ:

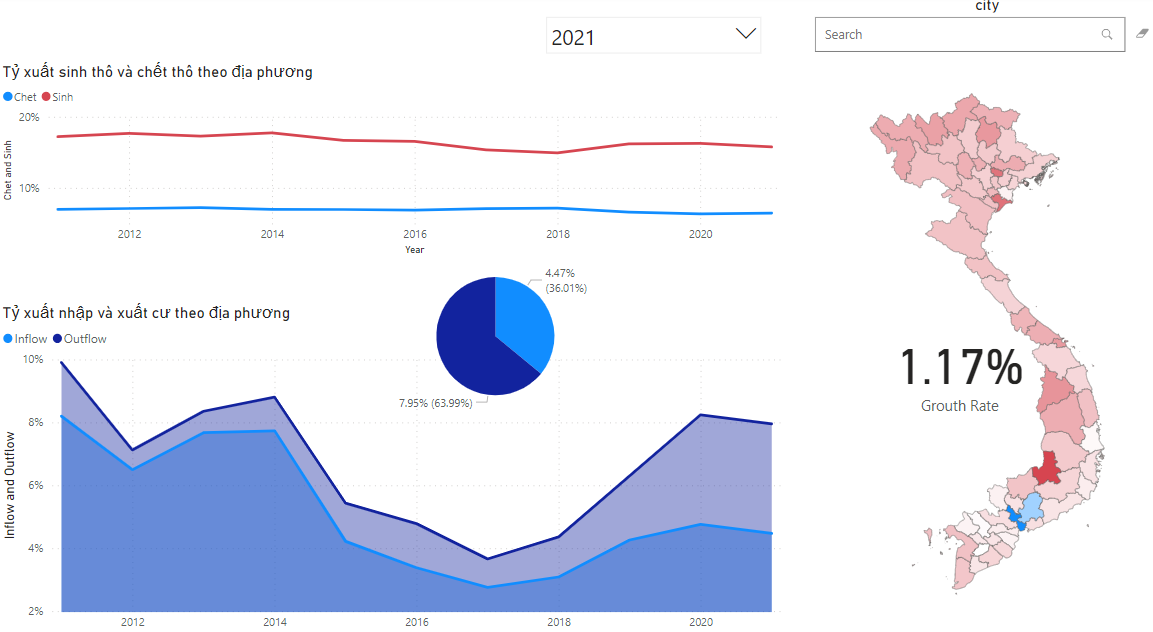
* Biểu đồ cột thống kê dân số của địa phương
* Biều đồ đường chỉ diễn biến dân số của địa phương hoặc cả Việt Nam từ 2011 – 2021
* Biểu đồ đường chỉ % ra tăng dân số của địa phương hoặc cả Việt Nam từ năm 2011 – 2021

Phần Reginal Analysis: Bao gồm 2 biểu đồ:

* Biểu đồ cột thống kê dân số theo miền
* Biểu đồ Ribbon chart thống kê tình hình dân số của các miền qua các năm

Phần VietNam Analysis: Biểu đồ nhiệt thể hiện tình trạng phân bổ dân cư của Việt Nam

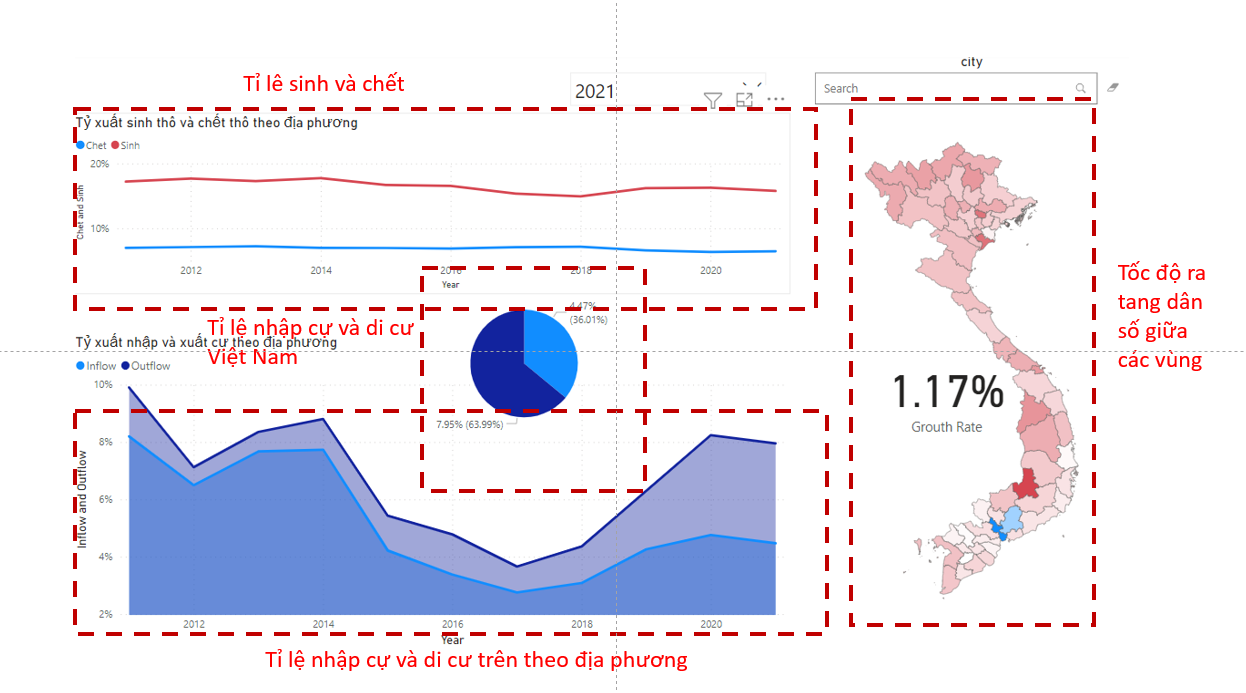
## Trang Khảo Sát ra tăng dân số



Hình 3 – Trang Khảo sát ra tăng dân số

Về cơ bản, trang này sẽ chứa tất cả thông tin liên quan tốc độ ra tăng dân số của các địa phương và toàn quốc. Bao gồm 4 thuộc tính chính là:

* Tỉ lệ sinh
* Tỉ lệ tử vong
* Dòng người nhập cư
* Dòng người di cư



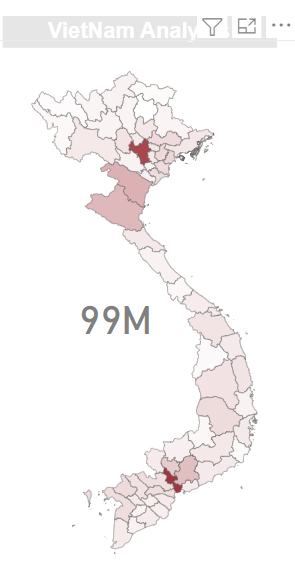
Hình 4 – Chi tiết

* Ở Bản đồ Việt Nam, khu vụng có màu đỏ là tỉ lệ dương, khu vực màu xanh là tỉ lệ âm

# NHẬN XÉT

Trong phần này, nhóm chúng em sẽ tiến hành đánh giá dữ liệ để đưa ra một số tri thức hữu ích từ dữ liệu. Phần nhận xét chi thành 2 phần: Phần về mật độ dân số và Phần về tốc độ ra tăng dân số.

## Dân số

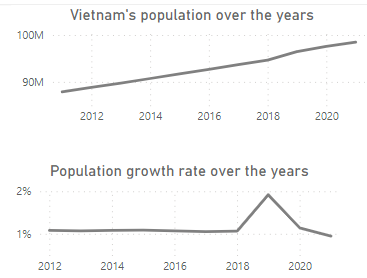


**Hình 5 – Phân bổ dân số Việt Nam.**

* Nhìn chung thì trong 10 năm trở lại bản phân bổ dân số vẫn không thay đổi nhiều. Dân số vẫn có xu hướng tập trung ở các thành phố lớn như Hà Nội và HCM
* Ngoài các thành phố lớn, Dân số có xu hướng chuyển tới làm việc tại các thành phố có nhiều vốn đầu tư nước ngoài như Bắc Ninh, Lâm Đồng ( Hình 6)



**Hình 6 - 5 tình có vốn đầu tư nước ngoài nhiều nhất**



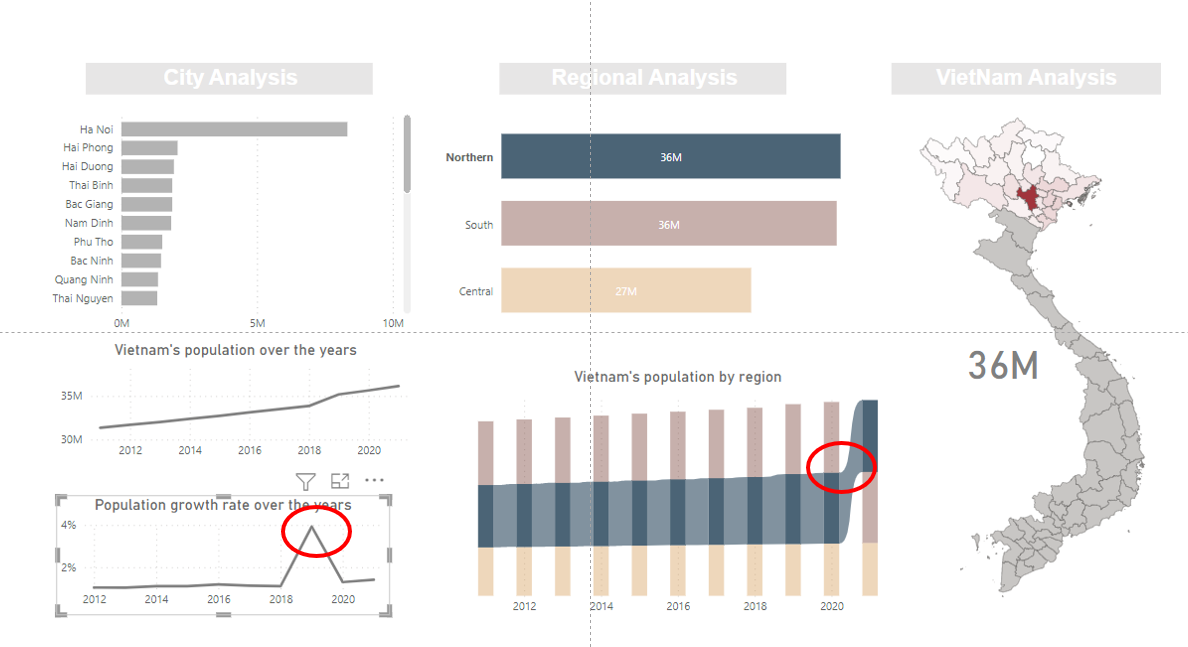
Hình 6 – Dân số Việt Nam trong 10 năm

Nhận xét:

* Tốc độ phát triển dân số Việt Nam vẫn tăng đều qua các năm từ dưới 90 triệu người năm 2011 lên đến gần 100 triệu ngườ năm 2021
* Tốc độ trung bình là khoảng 1% trên năm
* *Đặc biệt năm 2019* có sự ra tăng dân số đáng kinh ngạc, cao gần gấp đôi so với các năm khác . Sự đóng góp này là do miền Bắc năm 2019 có tỉ lệ ra tăng dân số khoảng 4% và cũng từ năm 2019 miền bắc đã vượt lên miền nam trở thành vùng có đông dân cư nhất ( hình 6 )

1. Miền Bắc

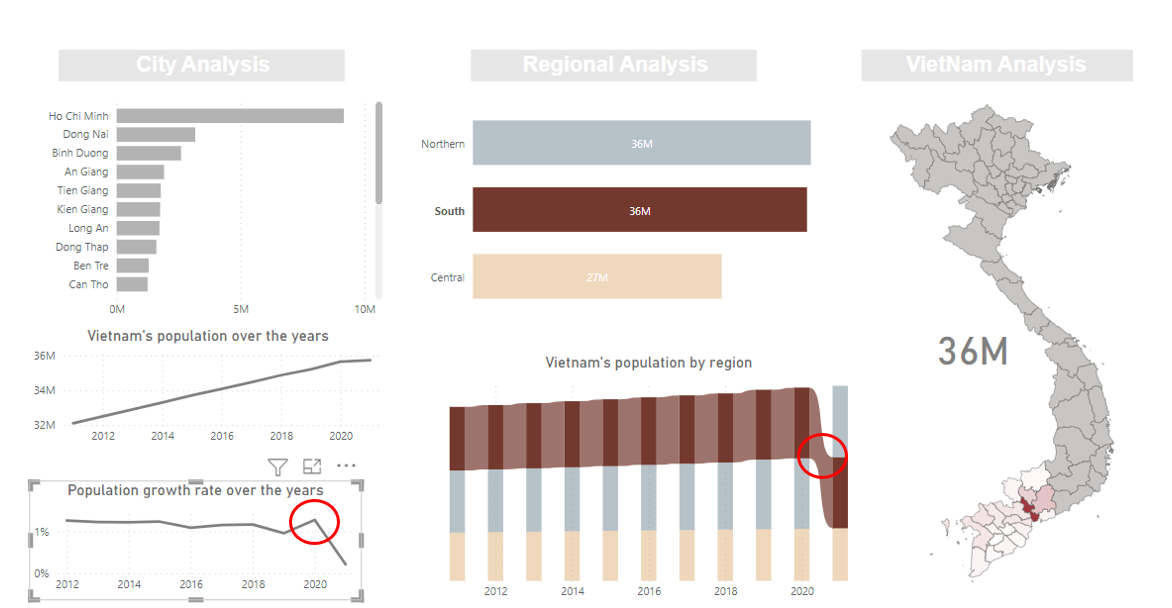
* Dân cư tập trung ở các thành phố lớn và phát triển như Hà Nội, Hải Phòng, Bắc Ninh….
* Dân cư có xu hướng tránh xa các vùng núi gần biên giới phía bắc và tập trung tương đối tại các thành phố ven biển.
* Năm 2019 cho thấy sự ra tăng dân số bất ngờ gần 4% cao gấp 3 lần các vùng miền khác, miền Bắc vươn lên trở thành vùng có dân số lớn nhất cả nước.
* Dự đoán tốc độ ra tăng dân số miền Bắc vẫn diễn ra đều trong các năm tới.



**Hình 6. Dân số miền Bắc**

1. Miền Nam

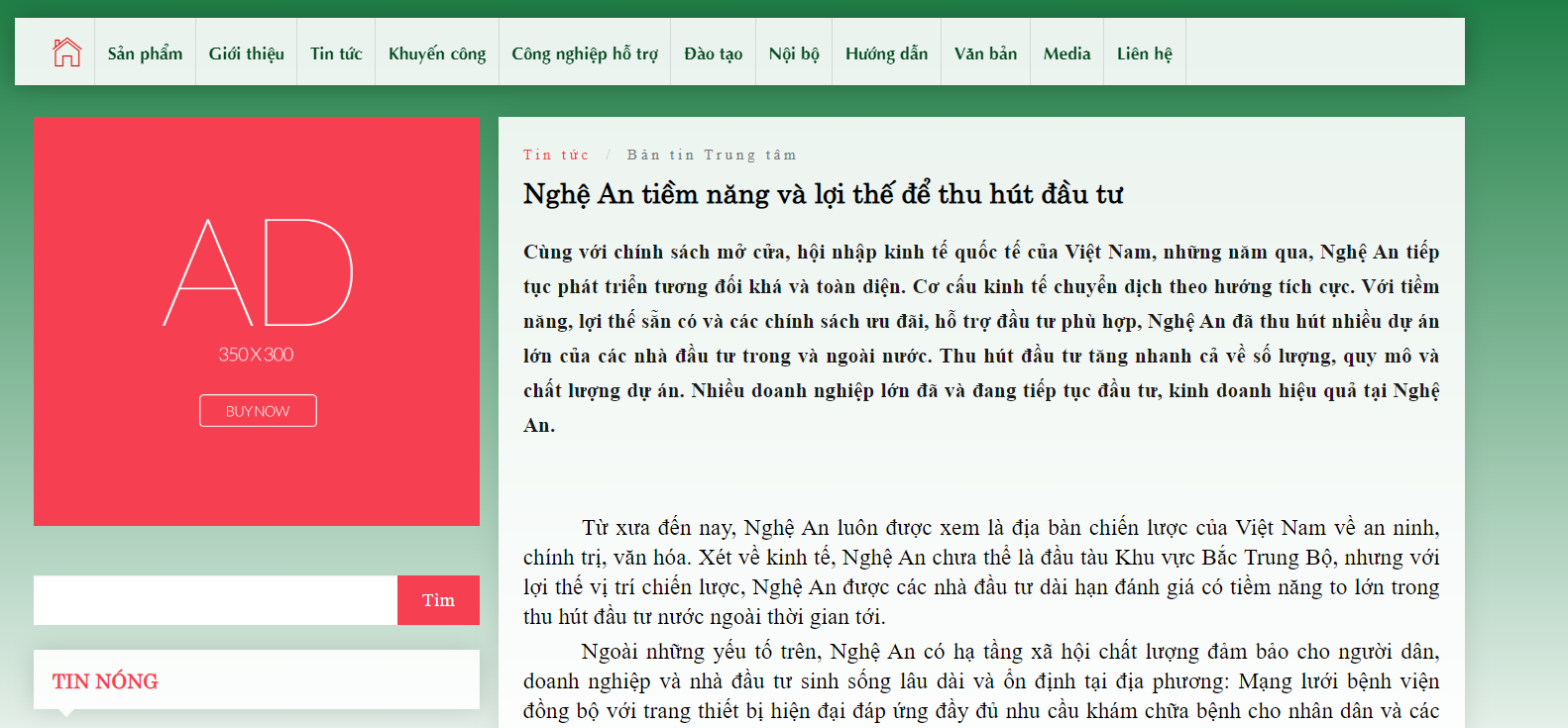
* Nhìn chung thì phân bô dân số phía Nam không khác nhiều so với phía bắc.
* Trái ngược với miền Bắc thì từ năm 2019- 2020 phía nam bắt đầu có sự giảm dân số thập chí tính đến năm 2021 tốc độ ra tăng dân số của miền nam còn ở mức âm.



**Hình 7. Dân số miền Nam**

1. Miền Trung

* Khác với 2 miền Nam Bắc, thành phố tập trung đông dân nhất không phải là thành phố phát triển nhất là Đà Nẵng mà là 2 thành phố Nghệ An và Thanh Hóa. Do Đà Năng là 1 thành phố du lịch, không thích hợp cho người dân tìm kiếm công việc.
* Dân số vẫn tăng đều qua các năm và không thấy có biến động gì đáng chú ý trong 10 năm trở lại.
* Tỉnh Nghệ An cho thấy sự phát triển dân số vượt bậc so với các khu vực khác ở miền Trung từ 2011-2021. Do Nghệ An được đánh giá có các yếu tố và điều kiện địa lý phù hợp và mức lương ổn định cho đầu tư công nghiệp và ổn định cuộc sống người dân lao động.( hình 9)

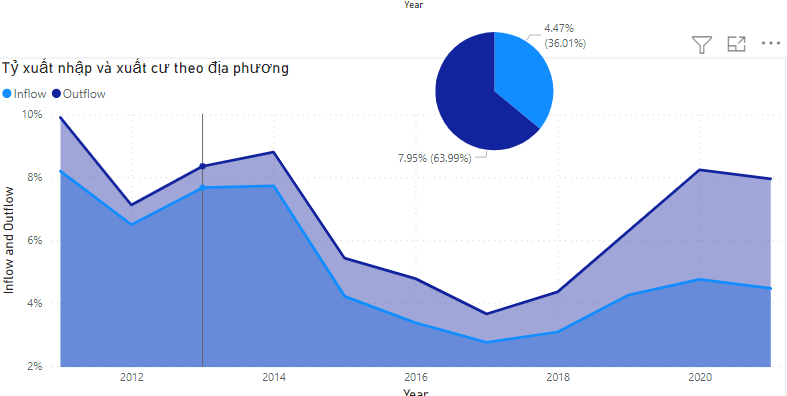


**Hình 9. Tiềm Năng của Nghệ An**

## Tốc độ ra tăng dân số

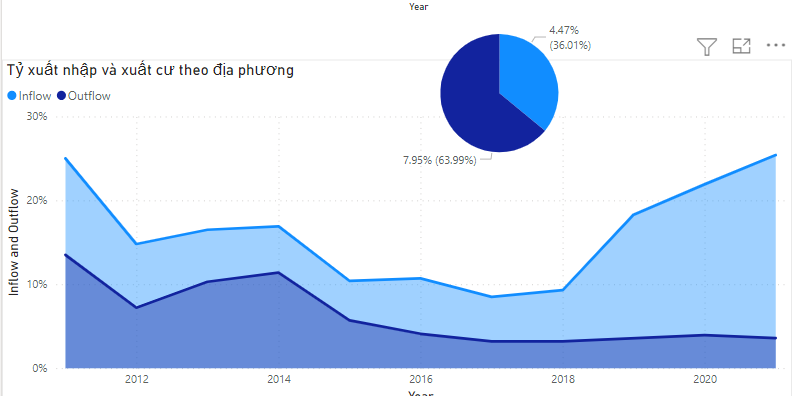
1. Tỉ lệ nhập cư và di cư

* Theo biểu đồ từ 2011-2021 tỉ lệ người dân di cư đi tìm việc tại các thành phố phát triển cao hơn tỉ lệ nhập cư. ( hình 10).



**Hình 10- Tỉ lệ nhâp cư và di cư toàn quốc**

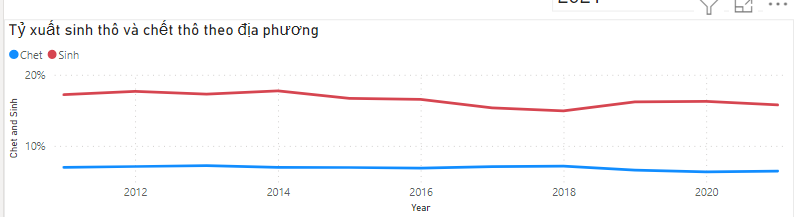
* Các thành phố lớn luôn có tỉ lệ nhập cư cao ( hình 11). VD như TP.HCM.



**Hình 11- Tỉ lệ nhập cư và di cư thành phố HCM**

1. Tỉ lệ sinh và tử vong

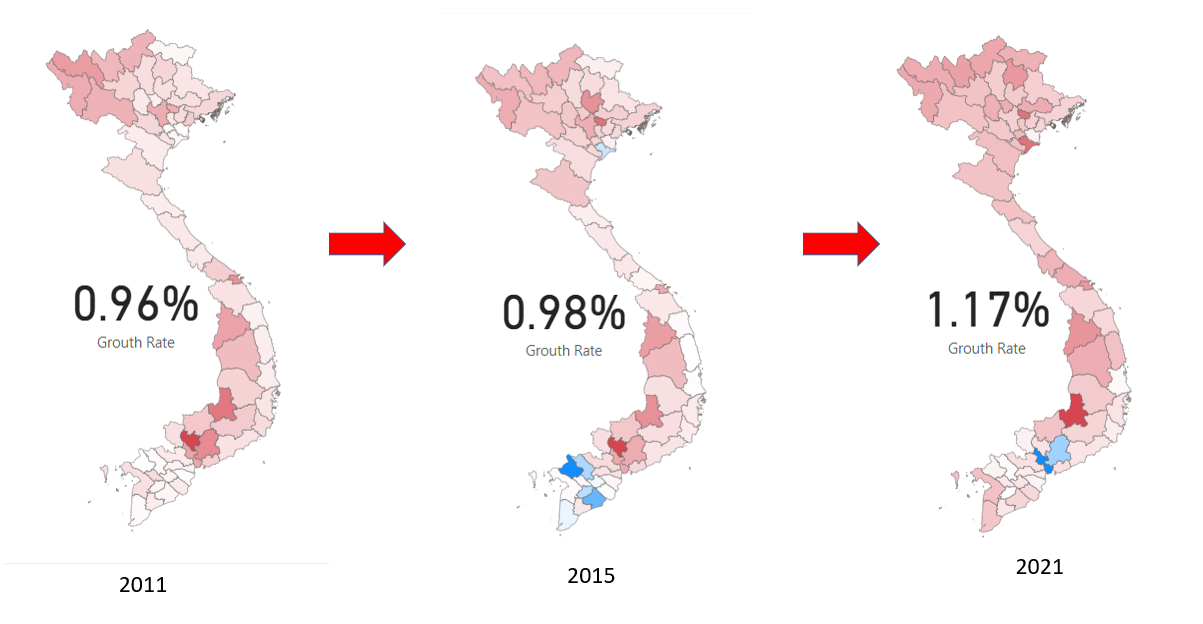
Tỉ lệ sinh vẫn cao hơn tỉ lệ tử vong. Điều này là dễ hiểu do điều kiện cơ sở y tế đã được nâng cao dõ dệt trong việc chăm sóc thai phụ và chữa trị bênh nhân. Điều này tích cực giúp đóng góp vào quá trình phát triển dân số Việt Nam.



**Hình 12 – Tỉ xuất sinh và tử vong theo địa phương 2011-2021**

1. Tốc độ phát triển dân số.

* Tốc độ phát triển dân số vẫn tăng đều đặn.
* Đặc biệt thành phố HCM năm 2021 đánh dấu sự thụt giảm dân số, mức độ tăng dân số về mức âm ( có màu xanh )



Hình 12. Tốc độ tăng dân số Việt Nam

KẾT LUẬN

Sau khi phân tích dữ liệu về Dân số, nhóm em đã tìm ra được những tri thức và kiểm chứng nó bằng dữ liệu. Đến đây thì bài tập lớn của nhóm em đã kết thúc, nhóm em xin chân thành cảm ơn thầy đã giúp đỡ nhóm trong quá trình làm bài tập lớn này.

Chúng em chân thành cảm ơn!

**THE END**