

**ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH**  
**TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN**



# **TRỰC QUAN HÓA DỮ LIỆU**

Đề tài:

## **TRỰC QUAN VÀ PHÂN TÍCH TÌNH HÌNH DÂN SỐ TẠI VIỆT NAM**

### **THÀNH VIÊN NHÓM**

Cao Nguyễn An Sơn – 18127196

Dương Ngọc Nguyên Phương – 18127187

Võ Văn Quốc Huy – 18127113

Bùi Thành Long – 18127132

**Môn học: Trực Quan Hóa Dữ liệu**

**Thành phố Hồ Chí Minh – 2022**

---

# MỤC LỤC

---

MỤC LỤC .....	2
Thông tin nhóm .....	3
Thông tin đề tài .....	4
1. Tên đề tài .....	4
2. Mô tả dữ liệu .....	4
Trực quan và phân tích dữ liệu .....	9
3. Dashboard về mật độ dân số .....	9
4. Dashboard về thành phần dân số .....	12
5. Dashboard về dân số và kinh tế đất nước .....	16
Phân tích .....	19
Tham khảo .....	21

---

## Thông tin nhóm

---

MSSV	Họ và tên
18127196	Cao Nguyễn An Sơn
18127187	Dương Ngọc Nguyên Phương
18127113	Võ Văn Quốc Huy
18127132	Bùi Thành Long

# Thông tin đề tài

## 1. Tên đề tài

Trực quan và phân tích tình hình dân số tại Việt Nam trong các năm vừa qua.

Lấy ý tưởng từ việc dân số hiện đang là một trong những vấn đề đáng quan tâm tại thế giới nói chung và Việt Nam nói riêng. Do tầm quan trọng và mức độ ảnh hưởng sâu rộng, có thể nói là ít nhiều tác động trực tiếp đến tăng trưởng kinh tế của nước nhà, các vấn đề về giáo dục, y tế, môi trường dẫn theo nhiều hệ lụy đáng quan ngại...

## 2. Mô tả dữ liệu

Ở đề tài lần này, nhóm chọn Trực quan hóa về dân số tại đất nước Việt Nam

Nhóm có bộ dữ liệu thông qua trang của Tổng Cục Thống Kê Việt Nam:

<https://www.gso.gov.vn/>



Trong đây dữ liệu khá đa dạng và đầy đủ được tổng hợp thành các file: csv, xlsx... và cho phép người dùng tải về sử dụng với định dạng người dùng mong muốn.

## Số liệu thống kê

Đơn vị hành chính, đất đai và khí hậu	
Dân số và lao động	<ul style="list-style-type: none"> <li>• Diện tích, dân số và mật độ dân số phân theo địa phương</li> <li>• Dân số trung bình phân theo giới tính và thành thị, nông thôn</li> <li>• Dân số trung bình phân theo địa phương, giới tính và thành thị, nông thôn</li> <li>• Tỷ số giới tính phân theo thành thị, nông thôn</li> <li>• Tỷ số giới tính của dân số phân theo địa phương</li> <li>• Tỷ số giới tính của trẻ em mới sinh phân theo vùng</li> <li>• Tỷ suất sinh thô, tỷ suất chết thô và tỷ lệ tăng tự nhiên của dân số phân theo thành thị, nông thôn</li> <li>• Tỷ suất sinh thô, tỷ suất chết thô và tỷ lệ tăng tự nhiên của dân số phân theo địa phương</li> <li>• Tổng tỷ suất sinh phân theo thành thị, nông thôn</li> <li>• Tổng tỷ suất sinh phân theo địa phương</li> <li>• Tỷ suất chết của trẻ em phân theo giới tính và phân theo thành thị, nông thôn</li> <li>• Tỷ suất chết của trẻ em dưới 1 tuổi phân theo địa phương</li> <li>• Tỷ suất chết của trẻ em dưới 5 tuổi phân theo địa phương</li> <li>• Tỷ lệ tăng dân số phân theo địa phương</li> <li>• Tỷ suất nhập cư, xuất cư và di cư thuần phân theo địa phương</li> <li>• Tuổi thọ trung bình tính từ lúc sinh phân theo giới tính và phân theo vùng</li> <li>• Tỷ lệ dân số từ 15 tuổi trở lên biết chữ phân theo giới tính và phân theo thành thị, nông thôn</li> <li>• Tỷ lệ dân số từ 15 tuổi trở lên biết chữ phân theo địa phương</li> <li>• Số cuộc kết hôn phân theo địa phương</li> <li>• Số vụ ly hôn đã xét xử phân theo địa phương và phân theo cấp xét xử</li> <li>• Lực lượng lao động từ 15 tuổi trở lên phân theo giới tính và phân theo thành thị, nông thôn</li> <li>• Lực lượng lao động từ 15 tuổi trở lên phân theo địa phương</li> <li>• Lao động từ 15 tuổi trở lên đang làm việc hàng năm phân theo loại hình kinh tế</li> <li>• Lao động từ 15 tuổi trở lên đang làm việc hàng năm phân theo giới tính và phân theo thành thị, nông thôn</li> <li>• Lao động từ 15 tuổi trở lên đang làm việc hàng năm phân theo giới tính</li> <li>• Lao động từ 15 tuổi trở lên đang làm việc hàng năm phân theo nhóm tuổi</li> <li>• Lao động và cơ cấu lao động từ 15 tuổi trở lên đang làm việc hàng năm phân theo ngành kinh tế</li> </ul>
Tài khoản quốc gia	
Nông, lâm nghiệp và thủy sản	
Công nghiệp	
Vốn đầu tư	
Doanh nghiệp	
Thương mại, du lịch và chi số giả	
Vận tải và Bưu chính, viễn thông	
Giáo dục	
Y tế, văn hóa, thể thao, mức sống dân cư	

ong-ke/#tab-65f7873fd78c7e60e39

Với những biến đánh dấu \* bạn cần phải chọn ít nhất một giá trị

Cách tính *	Năm *	Phân tổ *
Thông tin Tập PC-AXIS Tập Excel (xls) Tập Excel (xls) với cột mã và nội dung Phân tách bởi dấu Tab kèm tiêu đề Phân tách bởi dấu Tab không kèm tiêu đề Phân tách bởi dấu phẩy kèm tiêu đề Phân tách bởi dấu phẩy không kèm tiêu đề Phân tách bởi khoảng trống kèm tiêu đề Phân tách bởi khoảng trống không kèm tiêu đề Phân tách bởi dấu chấm phẩy kèm tiêu đề Phân tách bởi dấu chấm phẩy không kèm tiêu đề Tập HTML (htm) Tập quan hệ (scb) Biểu đồ dạng tập ảnh PNG Biểu đồ dạng tập ảnh GIF Biểu đồ dạng tập ảnh JPEG Excel (xlsx) Excel (xlsx) with code and text column <b>JSON-stat file (json)</b> Bảng - Giao diện 1	Chọn 21 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817 2818 2819 2820 2821 2822 2823 2824 2825 2826 2827 2828 2829 2830 2831 2832 2833 2834 2835 2836 2837 2838 2839 2840 2841 2842 2843 2844 2845 2846 2847 2848 2849 2850 2851 2852 2853 2854 2855 2856 2857 2858 2859 2860 2861 2862 2863 2864 2865 2866 2867 2868 2869 2870 2871 2872 2873 2874 2875 2876 2877 2878 2879 2880 2881 2882 2883 2884 2885 2886 2887 2888 2889 2890 2891 2892 2893 2894 2895 2896 2897 2898 2899 2900 2901 2902 2903 2904 2905 2906 2907 2908 2909 2910 2911 2912 2913 2914 2915 2916 2917 2918 2919 2920 2921 2922 2923 2924 2925 2926 2927 2928 2929 2930 2931 2932 2933 2934 2935 2936 2937 2938 2939 2940 2941 2942 2943 2944 2945 2946 2947 2948 2949 2950 2951 2952 2953 2954 2955 2956 2957 2958 2959 2960 2961 2962 2963 2964 2965 2966 2967 2968 2969 2970 2971 2972 2973 2974 2975 2976 2977 2978 2979 2980 2981 2982 2983 2984 2985 2986 2987 2988 2989 2990 2991 2992 2993 2994 2995 2996 2997 2998 2999 3000 3001 3002 3003 3004 3005 3006 3007 3008 3009 3010 3011 3012 3013 3014 3015 3016 3017 3018 3019 3020 3021 3022 3023 3024 3025 3026 3027 3028 3029 3030 3031 3032 3033 3034 3035 3036 3037 3038 3039 3040 3041 3042 3043 3044 3045 3046 3047 3048 3049 3050 3051 3052 3053 3054 3055 3056 3057 3058 3059 3060 3061 3062 3063 3064 3065 3066 3067 3068 3069 3070 3071 3072 3073 3074 3075 3076 3077 3078 3079 3080 3081 3082 3083 3084 3085 3086 3087 3088 3089 3090 3091 3092 3093 3094 3095 3096 3097 3098 3099 3100 3101 3102 3103 3104 3105 3106 3107 3108 3109 3110 3111 3112 3113 3114 3115 3116 3117 3118 3119 3120 3121 3122 3123 3124 3125 3126 3127 3128 3129 3130 3131 3132 3133 3134 3135 3136 3137 3138 3139 3140 3141 3142 3143 3144 3145 3146 3147 3148 3149 3150 3151 3152 3153 3154 3155 3156 3157 3158 3159 3160 3161 3162 3163 3164 3165 3166 3167 3168 3169 3170 3171 3172 3173 3174 3175 3176 3177 3178 3179 3180 3181 3182 3183 3184 3185 3186 3187 3188 3189 3190 3191 3192 3193 3194 3195 3196 3197 3198 3199 3200 3201 3202 3203 3204 3205 3206 3207 3208 3209 3210 3211 3212 3213 3214 3215 3216 3217 3218 3219 3220 3221 3222 3223 3224 3225 3226 3227 3228 3229 3230 3231 3232 3233 3234 32	

Khi nhóm tải về và sử dụng, tìm hiểu thì thấy dữ liệu chưa được phù hợp theo mong muốn nên đã có thực hiện bước tiền xử lý dữ liệu để có được các thông tin nhóm mong muốn.

```

[4] 1 import pandas as pd
    2 import numpy as np
    3 import seaborn as sns
    4 import matplotlib.pyplot as plt
    5 from functools import reduce

[5] 1 def melt_df(df):
    2     df_ = df.copy()
    3     index = df_.index
    4     df_ = df_.transpose().reset_index().melt(id_vars= 'index', value_vars = index)
    5
    6     return df_

1 df1 = pd.read_excel('Crude birth rate, crude death rate and natural increase rate by province.xlsx', index_col = 0, skiprows=[0,1,2])
2 df1.index = df1.index.str.strip()
3 df1.head(5)

```

	2005	2007	2008	2009	2010	2011	2012	2013	2014	2015	...	2011.2	2012.2	2013.2	2014.2	2015.2	2016.2	2017.2	2018.2	2019.2	Prel. 2020.2
WHOLE COUNTRY	18.6	16.9	16.7	17.6	17.1	16.6	16.9	17	17.2	16.2	...	9.7	9.9	9.9	10.3	9.4	9.2	8.1	7.8	10	10.24
Red River Delta	17.2	16.2	16.1	17.6	16.7	16.6	16.5	16.2	18.1	16.2	...	9.2	8.8	8.4	11	8.8	8.7	7.9	7.6	11.5	11.1
Ha Noi	16.3	16.5	16.9	19.2	18.8	18.6	17.1	16.5	18.9	16.3	...	11.8	9.9	9.2	12.3	9.1	8.8	9.6	8.6	13.4	12.47
Ha Tay	18.2	16.9	..	..	..	..	..	..	..	..	...	..	..	..	..	..	..	..	..	..	..
Vinh Phuc	18.7	17.5	17.9	19	18.8	17.7	18.7	18	20.3	17.7	...	10.5	11.4	11	13.6	11.2	13.4	10.3	8.8	11.3	11.79

5 rows x 45 columns

```

[8] 1 attributes_1 = np.array_split(df1.columns, 3)

1 df1_1 = melt_df(df1[attributes_1[0]])
2 df1_1['index'] = df1_1['index'].str.lstrip('Prel. ')
3 index = df1_1['index']
4 df1_1.columns = ['Year', 'Province', 'Crude birth rate']
5 df1_1

```

	Year	Province	Crude birth rate
0	2005	WHOLE COUNTRY	18.6
1	2007	WHOLE COUNTRY	16.9
2	2008	WHOLE COUNTRY	16.7
3	2009	WHOLE COUNTRY	17.6

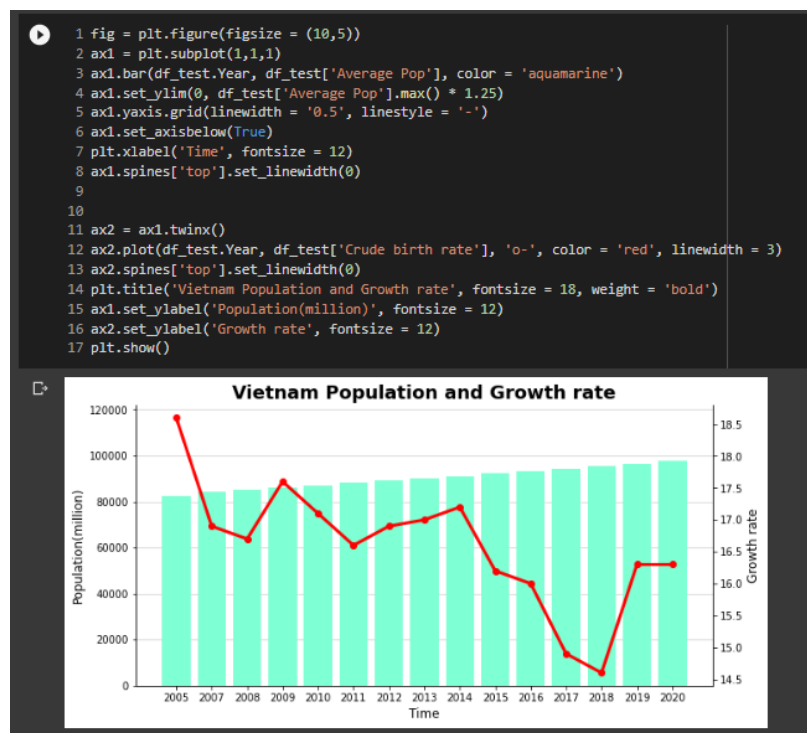
Trong phần tiền xử lý này, nhóm sử dụng các thư viện:

- Pandas
- Numpy
- Seaborn
- Matplotlib

Sau khi tiền xử lý dữ liệu thì nhóm sẽ chuyển qua dạng CSV và lưu các file lại để tiện cho phân bài làm.

Percentage of trained employed workers at 15 years of age and above by age group.csv															
Search (Alt+Q)															
File Home Insert Page Layout Formulas Data Review View Help															
Clipboard Font Alignment Number Styles															
A1															
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1		Year	Age	Percentage of trained											
2	0	2009	15-19	2.2											
3	1	2010	15-19	1.5											
4	2	2011	15-19	1.7											
5	3	2012	15-19	1.9											
6	4	2013	15-19	2.2											
7	5	2014	15-19	2											
8	6	2015	15-19	1.8											
9	7	2016	15-19	2											
10	8	2017	15-19	1.4											
11	9	2018	15-19	1.8											
12	10	2019	15-19	1.3											
13	11	2009	20-24	16.7											
14	12	2010	20-24	15.9											
15	13	2011	20-24	18.2											
16	14	2012	20-24	21.2											
17	15	2013	20-24	25.1											
18	16	2014	20-24	25.9											
19	17	2015	20-24	29.1											
20	18	2016	20-24	28.5											
21	19	2017	20-24	27.8											
22	20	2018	20-24	26											
23	21	2019	20-24	25											
24	22	2009	25-29	23.3											
25	23	2010	25-29	24.1											
26	24	2011	25-29	25.3											
27	25	2012	25-29	27.3											
28	26	2013	25-29	29.7											
29	27	2014	25-29	30.6											

Bên cạnh đó nhóm cũng sẽ thử trực quan lên bằng code trực tiếp trên Colab để xem thử dữ liệu đã như mong muốn hay chưa.



Ngoài ra, để vẽ tháp tuổi (tháp dân số) nhóm có sử dụng thêm nguồn dữ liệu từ WorldBank. Ở đây, ta sẽ lấy các thông tin về dân số theo các nhóm 5 tuổi (5-year age group) tại Việt Nam qua các năm.

The screenshot shows the World Bank DataBank interface. The sidebar on the left allows for selecting variables from the 'Database' and 'Country' categories. The 'Country' section is set to 'Vietnam'. The 'Series' section shows 67 available and 34 selected series. The 'Time' section shows 91 available and 10 selected time periods. The main preview table displays population estimates and projections for Vietnam from 2016 to 2021, categorized by age group and sex.

	2016	2017	2018	2019	2020	2021
Population ages 00-04, female	3,624,203.0	3,657,631.0	3,696,744.0	3,724,360.0	3,727,629.0	3,725,000.0
Population ages 00-04, male	4,048,570.0	4,086,242.0	4,134,160.0	4,166,596.0	4,164,842.0	4,157,000.0
Population ages 05-09, female	3,446,610.0	3,490,900.0	3,524,741.0	3,552,394.0	3,581,342.0	3,613,000.0
Population ages 05-09, male	3,839,020.0	3,900,114.0	3,942,398.0	3,973,148.0	4,004,759.0	4,041,000.0
Population ages 10-14, female	3,180,542.0	3,217,278.0	3,264,586.0	3,315,778.0	3,364,138.0	3,410,000.0
Population ages 10-14, male	3,433,097.0	3,498,408.0	3,578,005.0	3,660,286.0	3,734,036.0	3,799,000.0
Population ages 15-19, female	3,241,392.0	3,170,575.0	3,139,291.0	3,131,578.0	3,136,022.0	3,153,000.0
Population ages 15-19, male	3,443,333.0	3,371,222.0	3,342,166.0	3,343,413.0	3,364,847.0	3,402,000.0
Population ages 20-24, female	4,121,738.0	3,922,564.0	3,693,102.0	3,481,953.0	3,320,020.0	3,200,000.0
Population ages 20-24, male	4,318,004.0	4,113,340.0	3,878,227.0	3,663,198.0	3,500,183.0	3,378,000.0
Population ages 25-29, female	4,249,886.0	4,292,171.0	4,314,426.0	4,290,227.0	4,205,672.0	4,067,000.0
Population ages 25-29, male	4,408,512.0	4,454,171.0	4,476,922.0	4,451,047.0	4,363,621.0	4,222,000.0

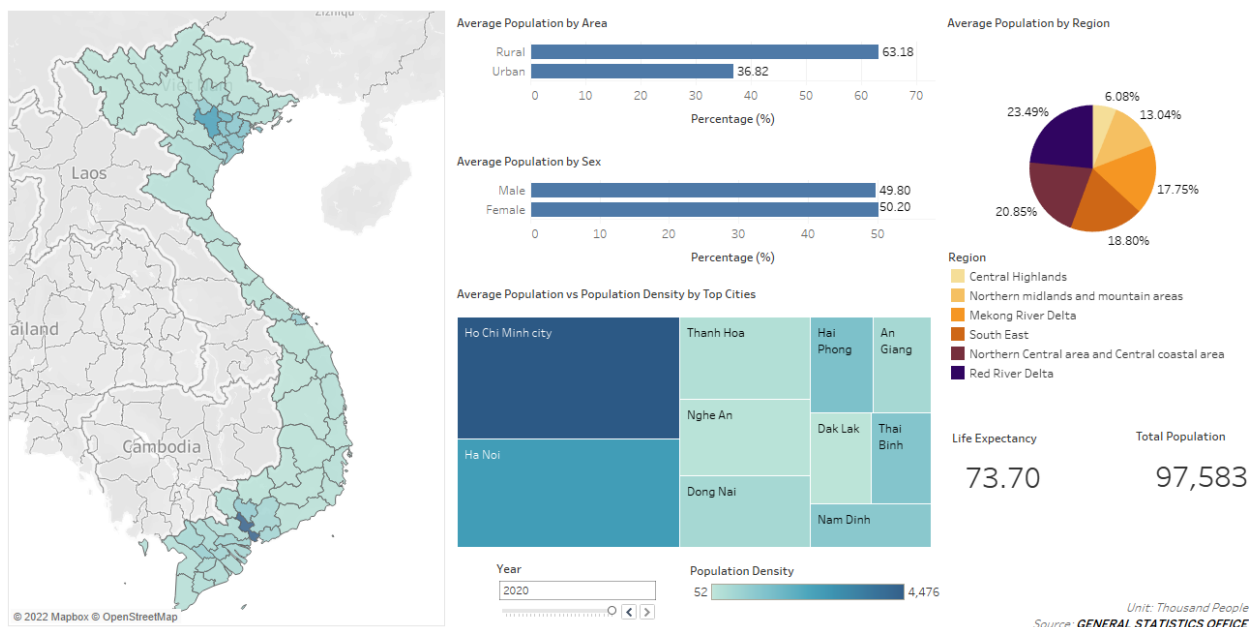
Dữ liệu này, sau khi tải về có thể sử dụng trực tiếp Tableau để tiền xử lý rồi sau đó trực quan hóa.



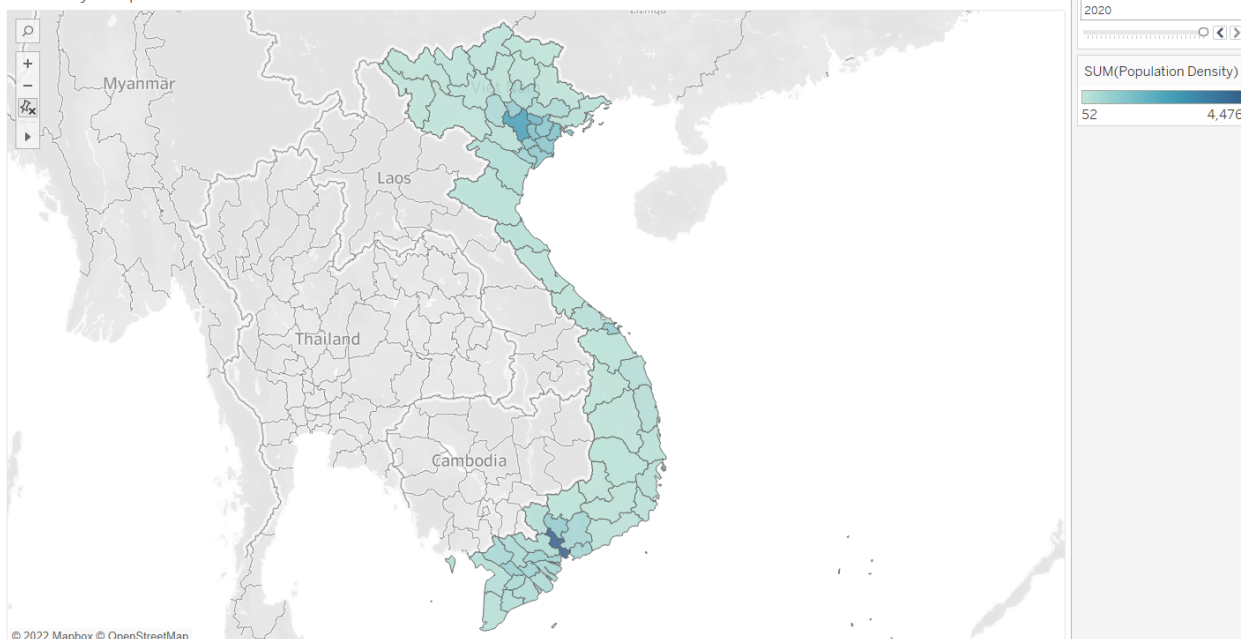
# Trực quan và phân tích dữ liệu

## 3. Dashboard về mật độ dân số

POPULATION DENSITY AND AVERAGE POPULATION IN VIETNAM



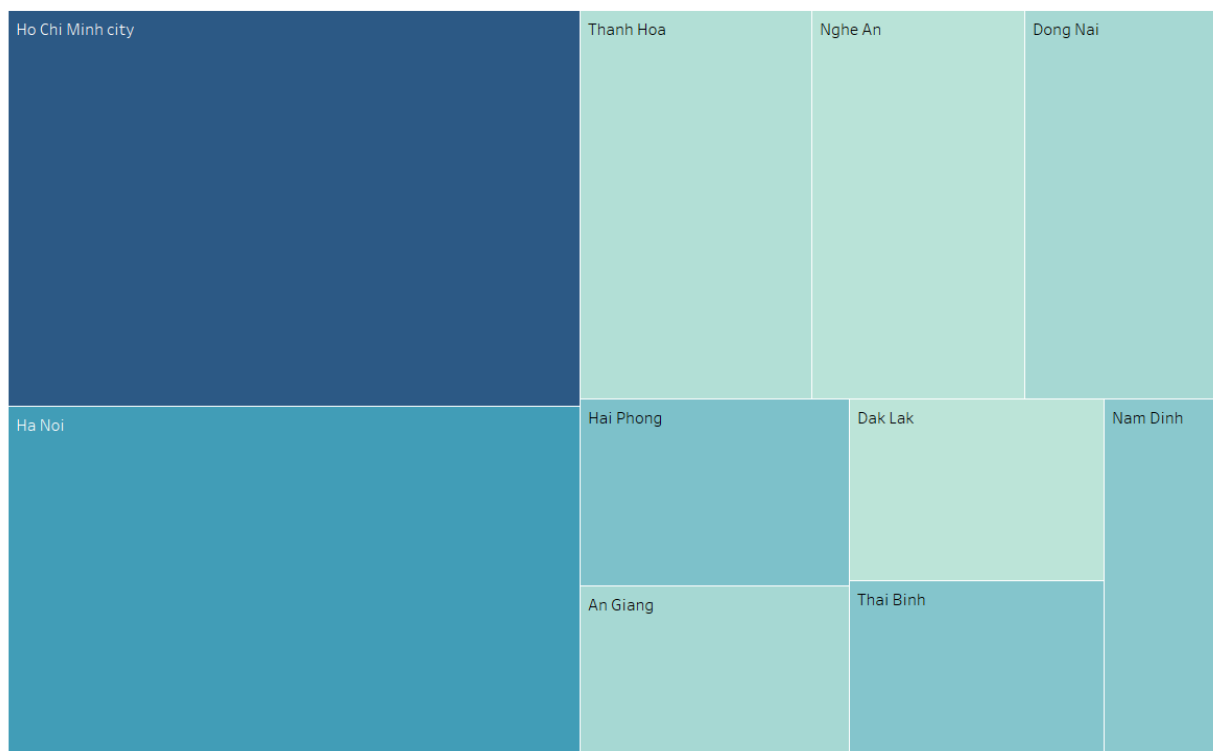
Density Map



Biểu đồ Bản đồ thể hiện mật độ dân số ở các tỉnh khắp nước Việt Nam. Qua đó có thể thấy dân số Việt Nam đang tập trung ở các thành phố lớn như Hồ Chí Minh với mật độ lên đến 4.476 hay cụ thể là 4.476 dân /km<sup>2</sup>. Đây là con số thể hiện thành phố Hồ Chí Minh đang dẫn đầu về mật độ dân số ở nước ta (Số liệu năm 2020). Bên cạnh đó, ta còn thấy rằng các

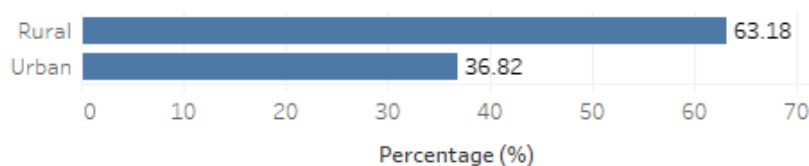
tỉnh thành có mật độ dân số cao thường tập trung ở xung quanh hai thành phố lớn nhất nước chính là TP Hồ Chí Minh và Hà Nội. Với công cụ Tableau, chúng ta có thể điều khiển và quan sát sự thay đổi mật độ dân số của các tỉnh, thành phố qua các năm với thanh kéo bên phải (hiện trên ảnh là đang để năm 2020).

Average Population vs Population Density by Top Cities

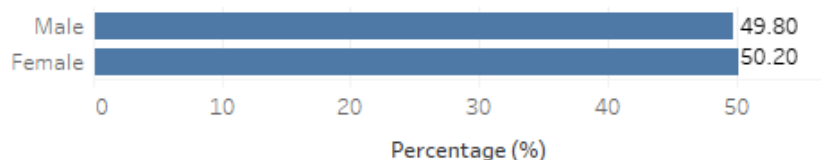


Nhóm có sử dụng thêm biểu đồ Tree Map để thể hiện sự tương quan giữa dân số trung bình và mật độ dân số của các tỉnh thành. Mật độ dân số tiếp tục sử dụng màu sắc để thể hiện, còn dân số trung bình ta sẽ sử dụng kích thước của ô trong biểu đồ. Ta thấy, hai thành phố Hồ Chí Minh và thủ đô Hà Nội vừa có dân số trung bình cao nhất cũng như mật độ dân số cũng lần lượt đứng vị trí cao nhất. Ở các tỉnh thành còn lại, ta thấy rõ dù dân số trung bình có cao nhưng chưa chắc có mật độ dân số cao. Ví dụ như trường hợp của Thanh Hóa và Nghệ An có dân số trung bình cao hơn nhưng có mật độ dân số thấp hơn hai tỉnh thành có diện tích nhỏ hơn nhưng hiện có nhiều các khu công nghiệp là Đồng Nai và Hải Phòng.

Average Population by Area



Average Population by Sex

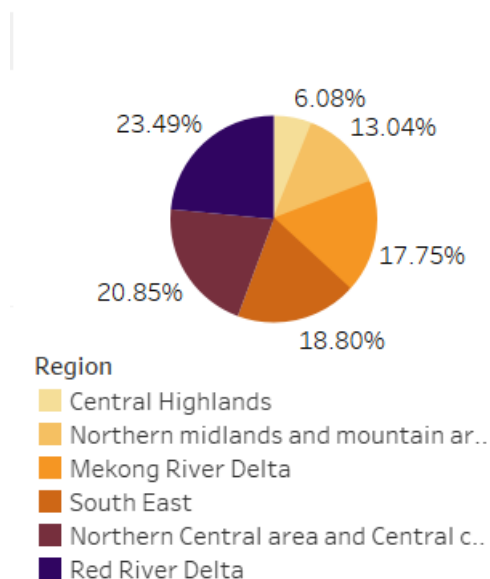


Tuy mật độ dân số ở các thành phố lớn cao nhưng nước Việt Nam ta đa phần dân số đều ở các vùng nông thôn là chính, chiếm tỉ trọng đến 63.18%.

Việc dân số phân bố tập trung như trên có vẻ chưa đồng đều giữa các tỉnh với nhau.

Bên cạnh đó ta cũng có thể thấy tỷ lệ Nam/Nữ trong dân số của chúng ta cũng khá đồng đều, trong đó Nữ chiếm nhiều hơn khoảng 50.2%.

Average Population by Region



Với biểu đồ tròn trên cho thấy được việc phân bố dân số ở nước ta đa phần là các vùng nông nghiệp chiếm tỉ trọng lớn. Như vùng Đồng bằng sông Hồng chiếm tỉ trọng cao nhất với 23.37% dân số trên cả nước trong khi vùng Trung nguyên thì chỉ có 6.08% dân số tại đây, đa phần là các dân tộc thiểu số và dân tộc đã sống từ trước.

Ngoài ra, ở dashboard này ta còn có hai thông tin về tuổi thọ trung bình và dân số trung bình của nước ta. Hai thông tin này sẽ thay đổi tương ứng khi ta thay đổi thanh kéo (slider) điều chỉnh thông tin năm.

Life Expectancy

73.70

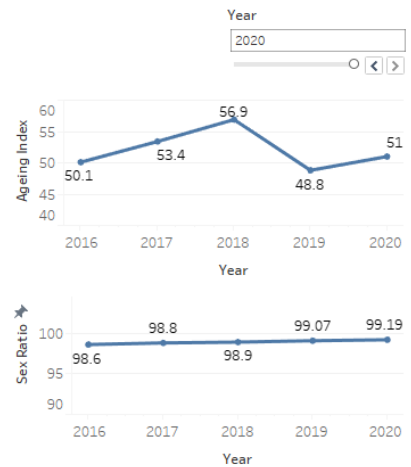
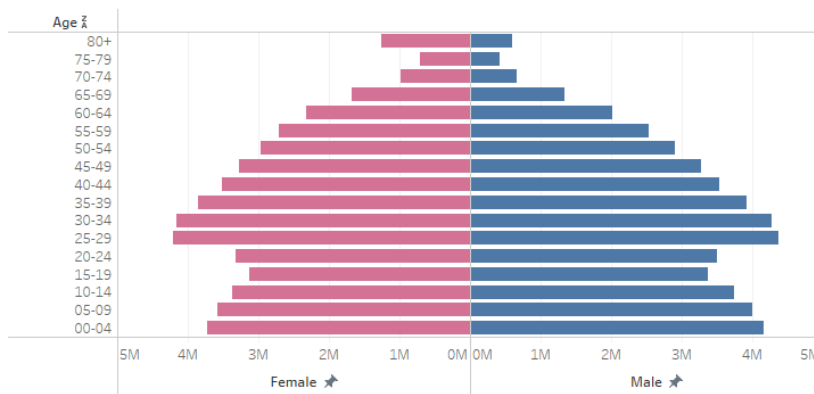
Total Population

97,583

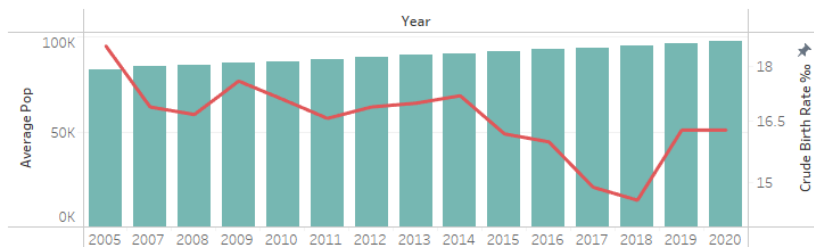
## 4. Dashboard về thành phần dân số

### POPULATION OF VIETNAM

Age Pyramid

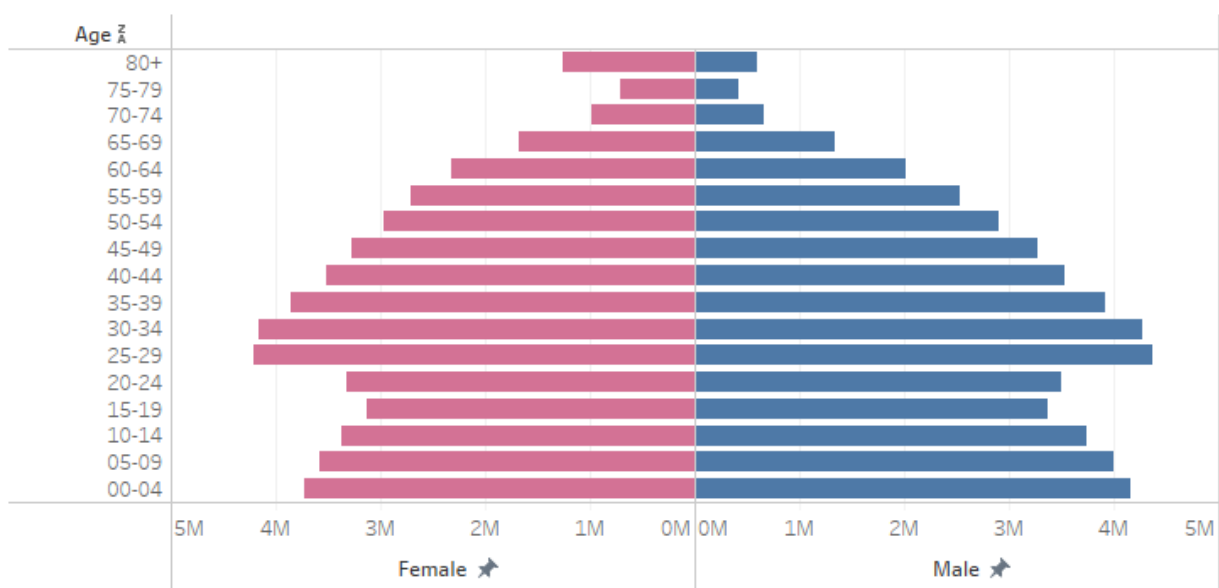


Average Population vs Crude Birth Rate



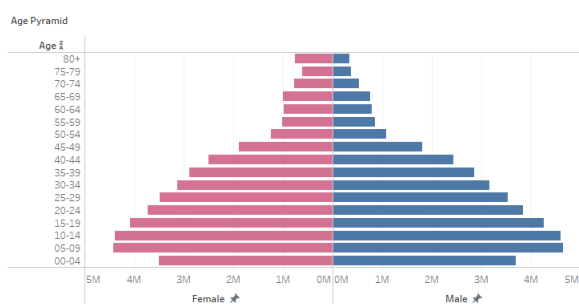
Source: GENERAL STATISTICS OFFICE, WORLD BANK

Age Pyramid

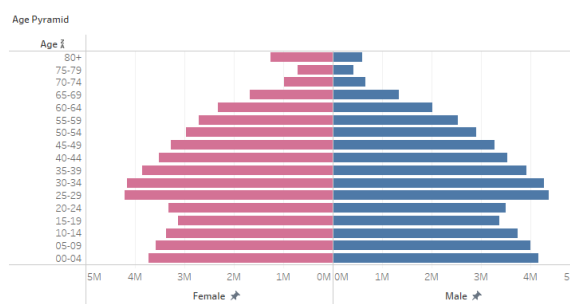


Ở biểu đồ kim tự tháp về tháp tuổi của dân số, ta có thể thấy cấu trúc dân số theo độ tuổi và giới tính. Có thể thấy đây là tháp tuổi đặc trưng của Việt Nam trong giai đoạn nước đang phát triển với tháp tuổi kiểu mở rộng: đáy rộng, đỉnh nhọn và cạnh thoải. Qua đó chứng tỏ tỉ suất sinh cao, và tuổi thọ trung bình chưa được cao.

Với tháp tuổi trên, thấy được thuận lợi và khó khăn trong cơ cấu dân số theo độ tuổi với việc phát triển kinh tế - xã hội nước ta. Với việc dân số trẻ cao, Việt Nam sẽ có nguồn lao động đông, bổ sung lao động lớn và thu hút các ngành kinh tế sử dụng lao động và sự đầu tư nước ngoài.



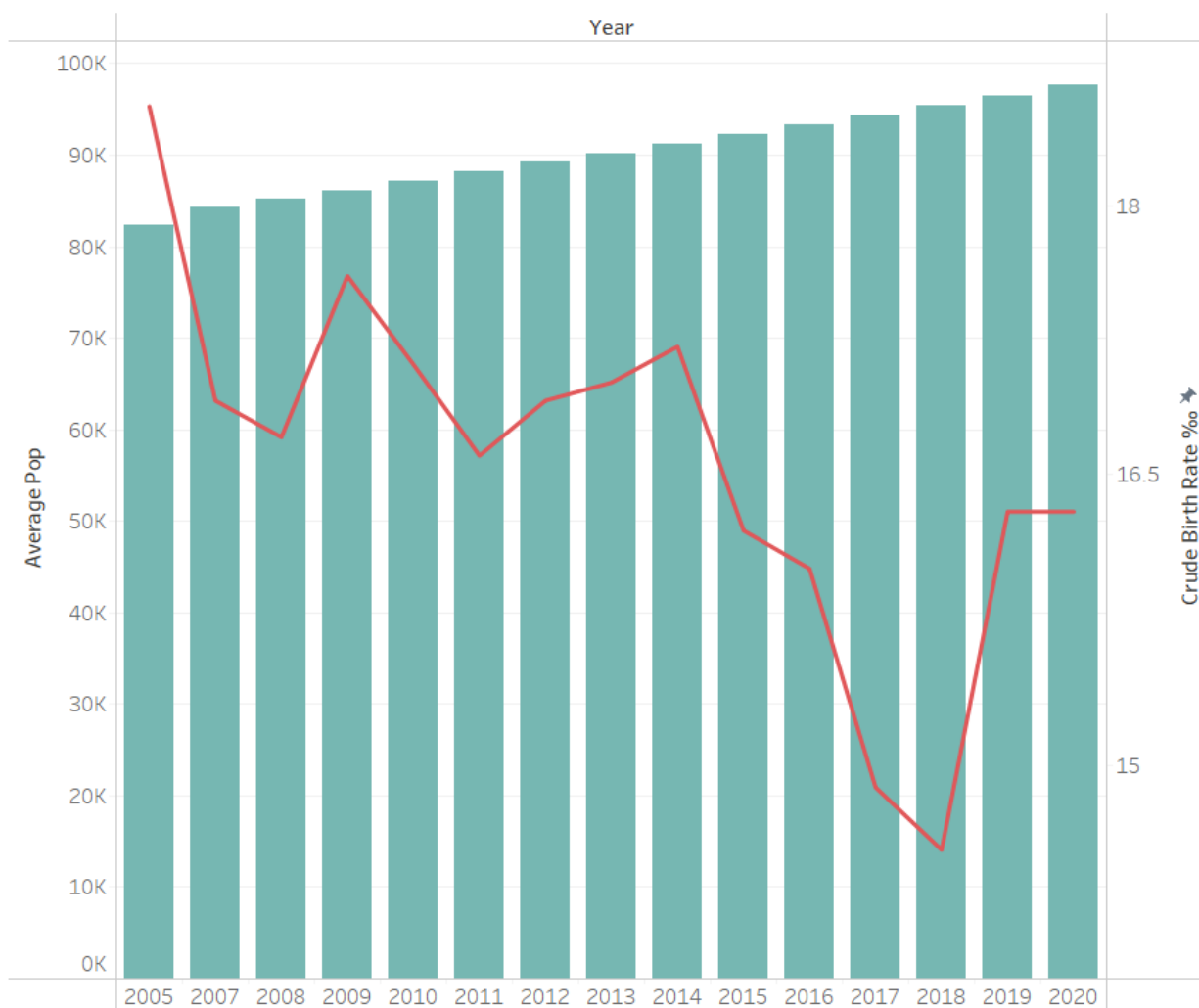
*Tháp tuổi Việt Nam năm 2000*



*Tháp tuổi Việt nam năm 2020*

Tuy nhiên, khi ta so sánh tháp dân số hiện tại và tháp dân số vào năm 2000. Ta thấy rõ sự thay đổi khi phần đáy tháp tuổi có xu hướng bị bóp lại thể hiện xu hướng giảm sinh ở nước ta. Đồng thời, phần đỉnh tháp có xu hướng lớn hơn. Đây thể hiện sự già hóa dân số ở nước ta. Tuy nhiên, ta thấy được phần lớn dân số vẫn nằm trong dân số trẻ và dân số thuộc độ tuổi lao động và đây chính là cơ cấu “dân số vàng”, yêu cầu nước ta cần tận dụng triệt để vì đây là cơ cấu hiếm hoi, không quay trở lại (nếu có phải cần ít nhất 100-200 năm). [1]

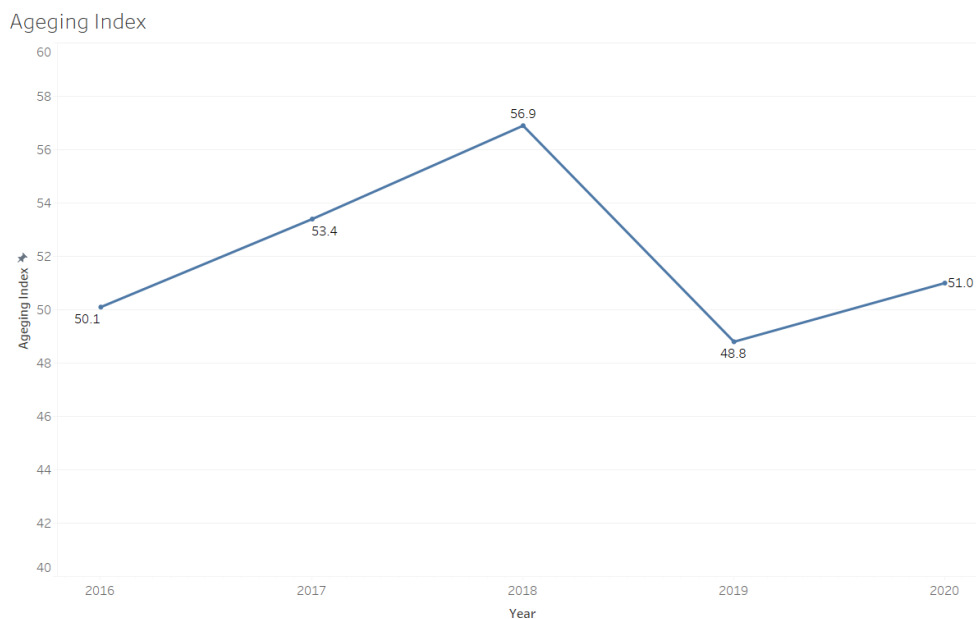
Average Population vs Crude Birth Rate



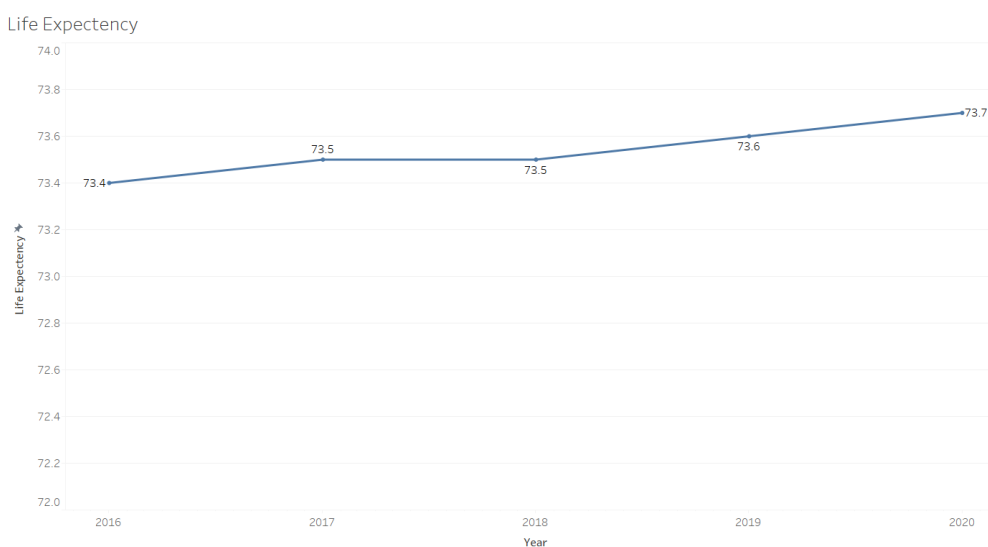
Qua hằng năm, dân số trung bình tăng thêm khoảng 1 triệu dân hàng năm. Quy mô dân số Việt Nam ngày càng lớn, tốc độ gia tăng đều hằng năm. Trong khi tỉ suất sinh thô của Việt Nam đang có xu hướng giảm nhanh do chính sách dân số triệt để của nước ta thực hiện tốt, tâm lý người dân ngại sinh đẻ cũng như người dân được nâng cao nhận thức về việc kế hoạch hóa gia đình.

Việc phân bố giới tính tại Việt Nam luôn có sự chênh lệch giữa nam và nữ. Chỉ số Sex Ratio được tính bằng tổng số nam / 100 người nữ. Con số này chưa có sự giảm đi mà lại gia tăng chậm rãi qua hằng năm, cho thấy xu hướng tăng.





Tiếp theo là chỉ số già hóa, được đo bằng cách tính số lượng người trên 60 tuổi trên một trăm người dưới 15 tuổi. Dân số Việt Nam đang có xu hướng già hóa khi mà chỉ số lão hóa trung bình tăng đều từ 2016 – 2018 nhưng rồi giảm đột ngột vào năm 2019 nhưng sau đó lại có xu hướng tăng lên.

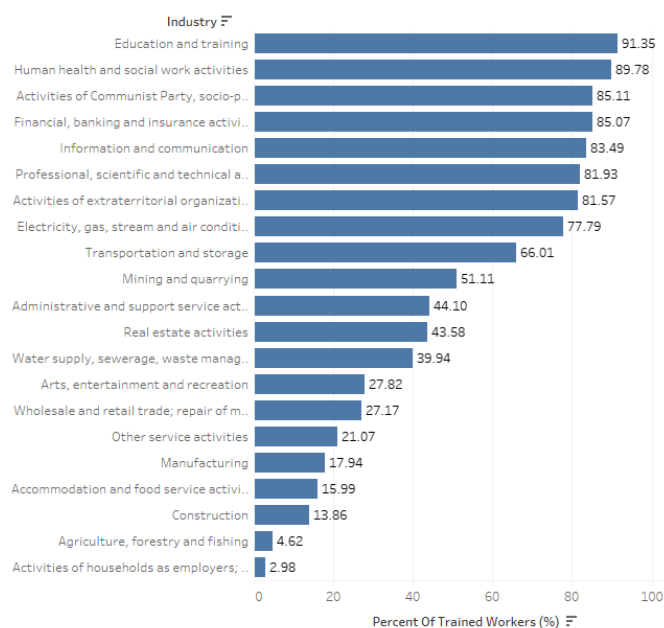


Tuổi thọ trung bình của dân số Việt Nam ngày càng tăng cao cho thấy được sự phát triển của đất nước về kinh tế - xã hội, y tế và chăm sóc sức khỏe. Cho thấy được đời sống và chất lượng sống của con người ngày càng tốt hơn.

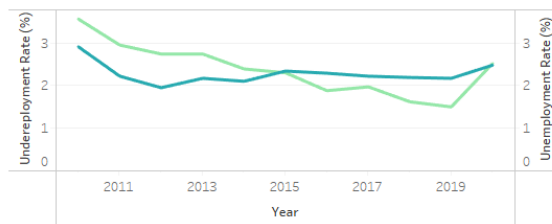
## 5. Dashboard về dân số và kinh tế đất nước

### POPULATION AND ECONOMICS

Trained Workers by Industry



Unemployment vs Underemployment Rate



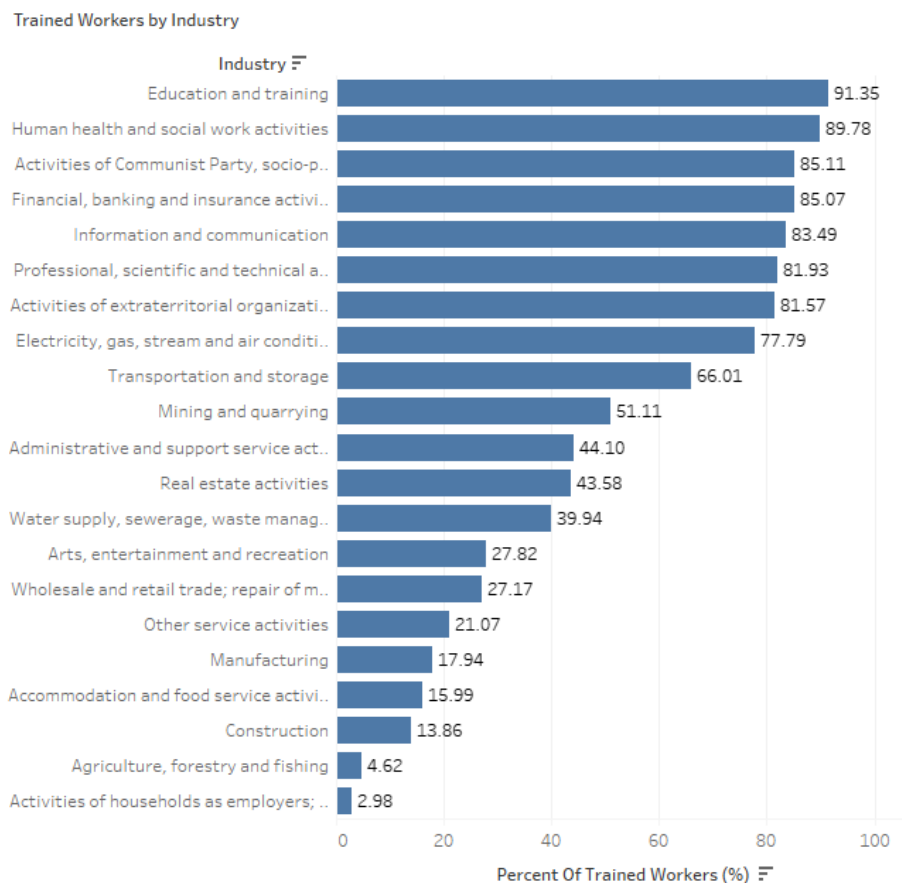
Trained Workers Rate (And Forecast)



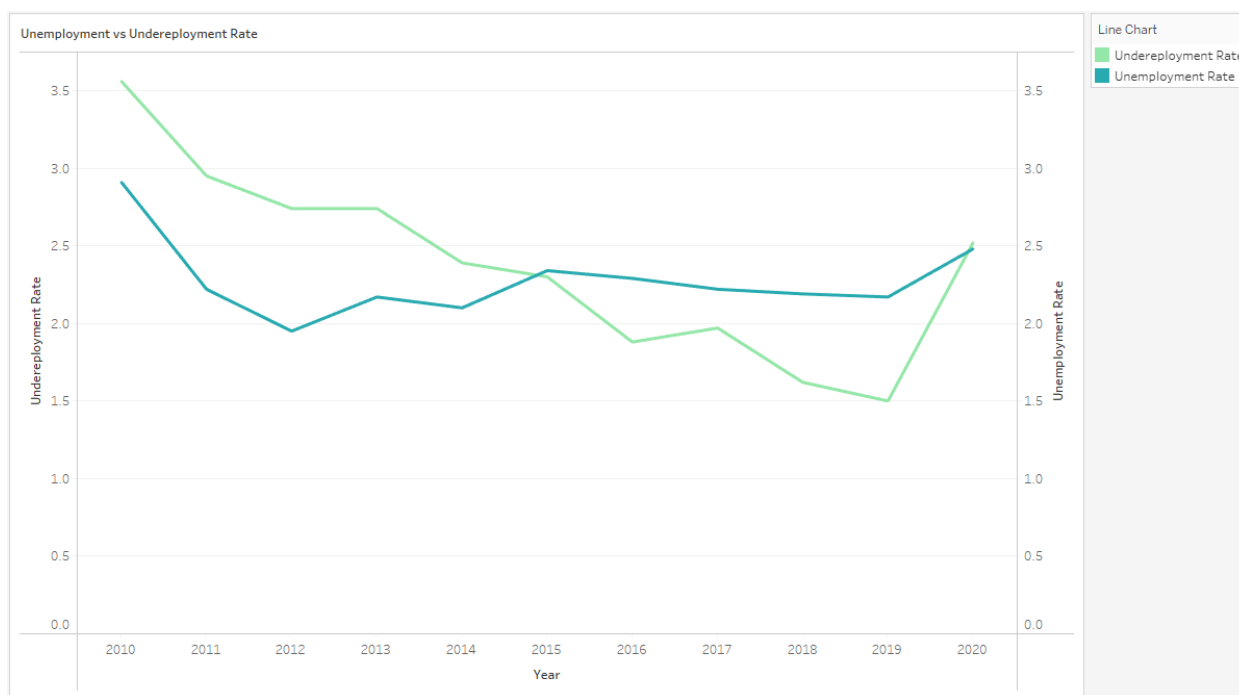
Year  Line Chart  
☐ Underemployment Rate ☐ Unemployment Rate  
 Forecast indicator  
☐ Actual ☐ Estimate

Source: GENERAL STATISTICS OFFICE



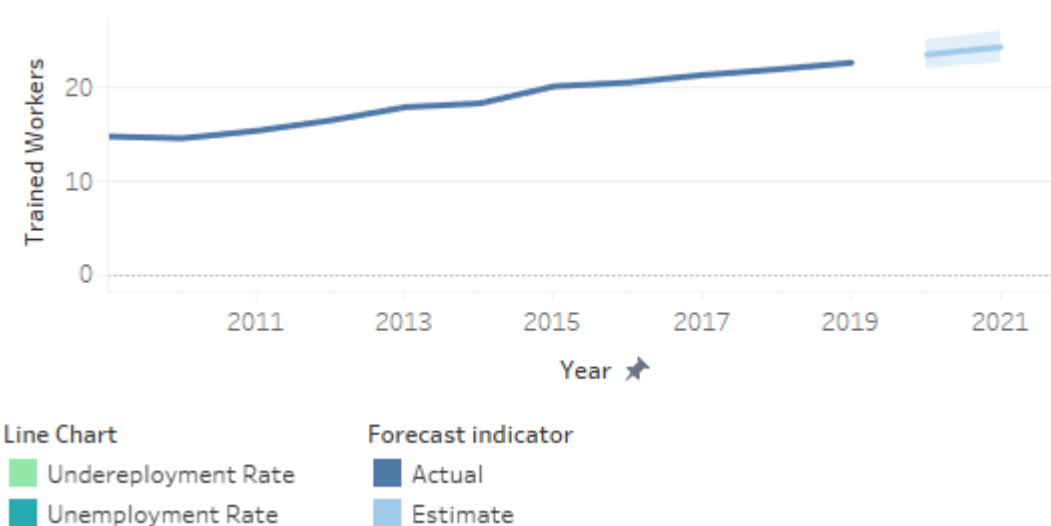


Biểu đồ trên thể hiện phần trăm người làm việc được đào tạo qua các nhóm ngành. Ta có thể thấy được, các nhóm ngành liên quan đến Giáo dục và đào tạo có tỉ lệ người tham gia được đào tạo là lớn nhất. Cũng dễ hiểu vì ngành giáo dục cần có kiến thức để đào tạo các học viên sau. Tiếp theo là về sức khỏe và hoạt động xã hội cũng là lĩnh vực liên quan đời sống hằng ngày của người dân. Việc có đến 91.35% người làm việc được đào tạo trong lĩnh vực giáo dục và 89.78% người làm việc được đào tạo trong lĩnh vực sức khỏe cũng cho thấy rằng Việt Nam có sự phát triển trong kinh tế - xã hội, cho thấy được đời sống và chất lượng cuộc sống của con người ngày càng tốt hơn.



Tiếp theo là biểu đồ về tỉ lệ thiếu việc làm và thất nghiệp. Như ta thấy, cả 2 tỉ lệ đều giảm dần trước năm 2020 cho thấy rằng, đất nước ta đang cải thiện tốt tình trạng thiếu việc làm. Việc tỉ lệ thất nghiệp ở năm 2013 và 2015 tăng nhẹ do sự chuyển dịch cơ cấu kinh tế nông nghiệp của đất nước chậm [2]. Nhưng đến năm 2020, cả 2 tỉ lệ thiếu việc làm và thất nghiệp tăng do hậu quả của dịch Covid-19 ảnh hưởng đến nhu cầu việc làm của người dân.

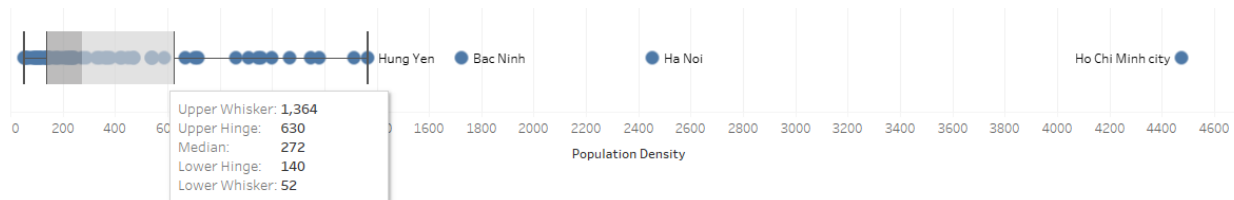
Trained Workers Rate (And Forecast)



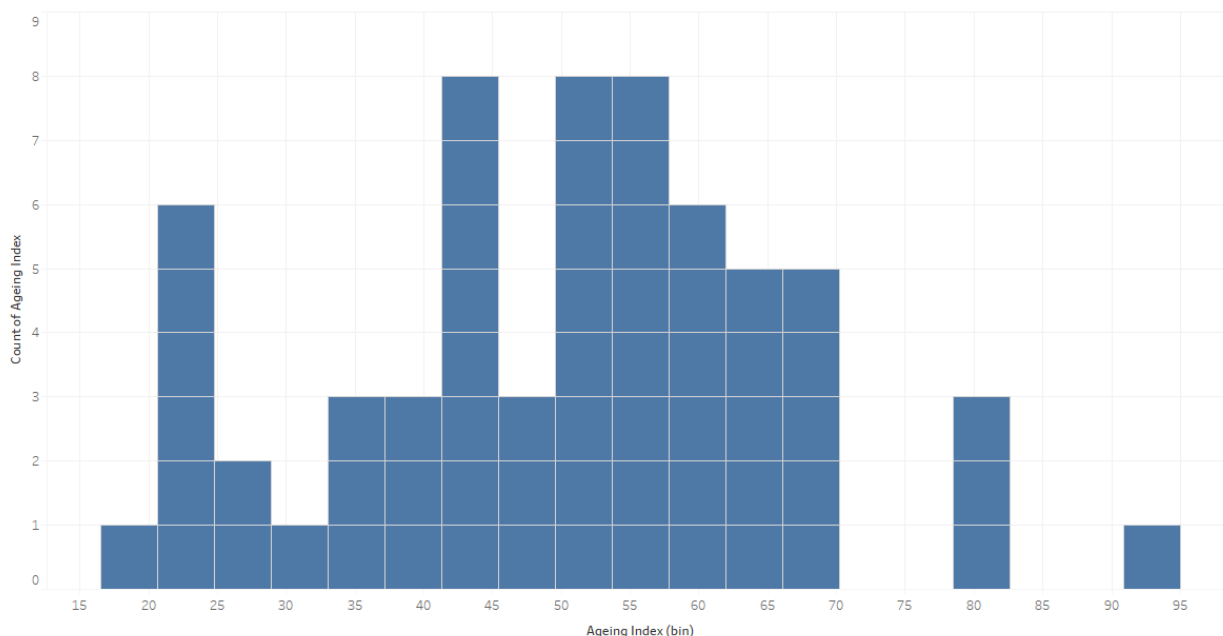
Cuối cùng là dự đoán tỉ lệ đào tạo trung bình trên dân số. Ta thấy, tỉ lệ này luôn tăng dần đều qua các năm cho thấy rằng đất nước ta đang phát triển ổn định đến năm 2019. Và dự đoán của Tableau nếu không có biến động thì con số này sẽ tiếp tục tăng dần đều.

## Phân tích

Nhóm sẽ dùng một số biểu đồ, để phân tích một số trường đơn

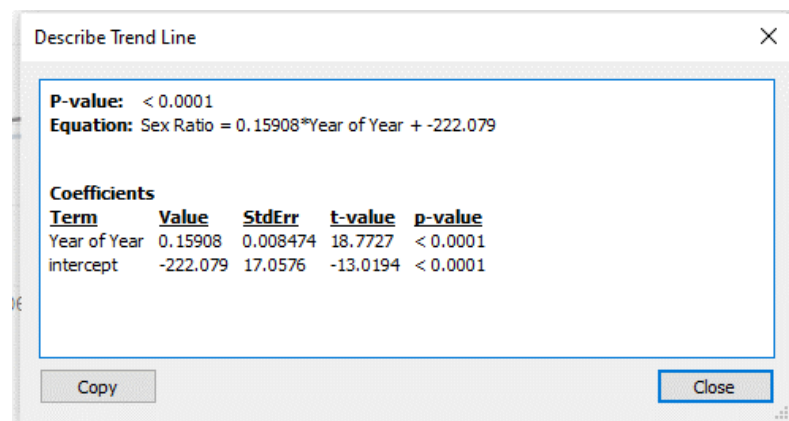
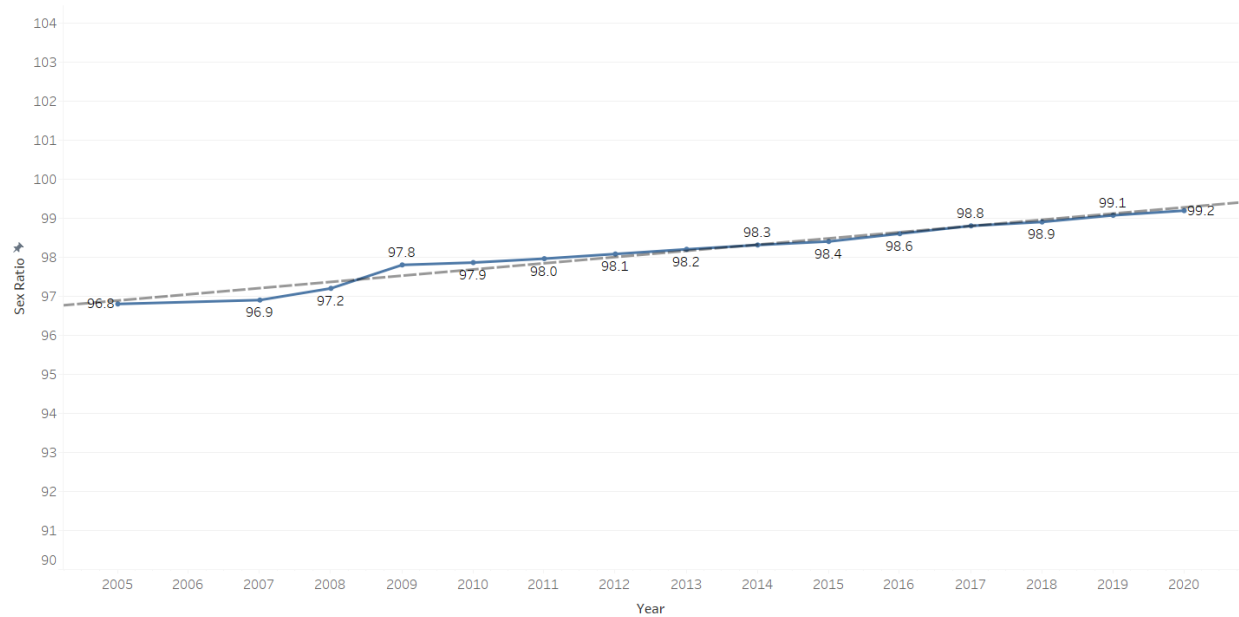


Sử dụng biểu đồ hộp (Box Plot), ta thấy 75% các tỉnh thành có mật độ dân số nhỏ hơn 630 người/ km<sup>2</sup>. Có 3 được xem là ngoại lệ là Bắc Ninh, Hà Nội và Thành phố Hồ Chí Minh, đây chính là những thành phố có mật độ dân số cao nhất nước ta. Do đây là các trung tâm kinh tế - xã hội của vùng cũng như cả nước.



Ở biểu đồ Ageing Index(bin) này cho ta thấy, số lượng khu vực có tỉ lệ người cao tuổi trên 100 người chưa trưởng thành, ở mức 40-60. Con số này cho thấy rằng, dân số nước ta đang chuyển biến theo chiều hướng già hóa. Đặc biệt ở các vùng: Thái Bình, Nam Định, Vĩnh Long và Bến Tre, con số này ở mức hơn 80 cho thấy rằng tình trạng già hóa dân số ở các vùng này đáng báo động.

Sex Ratio



Với mô hình tuyến tính, ta thấy được dữ liệu về giới tính là một tập dữ liệu tuyến tính đơn giản. Đường xu hướng tuyến tính cho thấy sự chênh lệch giới tính giữa nam và nữ đang tăng dần với tốc độ khá ổn định từ những năm sau 2010.

---

## Tham khảo

---

- [1] N. T. Nguyễn, "Cơ cấu “dân số vàng”, nếu không tận dụng, sẽ lỡ thời cơ," Báo Gia đình và Xã hội, 2020. [Online]. Available: [https://moh.gov.vn/chuong-trinh-muc-tieu-quoc-gia/-/asset\\_publisher/7ng11fEWgASC/content/co-cau-dan-so-vang-neu-khong-tan-dung-se-lo-thoi--1](https://moh.gov.vn/chuong-trinh-muc-tieu-quoc-gia/-/asset_publisher/7ng11fEWgASC/content/co-cau-dan-so-vang-neu-khong-tan-dung-se-lo-thoi--1).
- [2] VietStock, "Tỷ lệ thất nghiệp 2015 tăng lên 2.31%," [Online]. Available: <https://vietstock.vn/2015/12/ty-le-that-nghiep-2015-tang-len-231-761-452121.htm>.