| get # L get imp for for In [104 # 불러오 import | pandas as pd |
|--|---|
| import import import import import mpl.rcF plt.rcF import #%% | matplotlib.pyplot as plt matplotlib as mpl numpy as np geopandas as gpd Params['font.family'] = 'NanumGothic' Params['axes.unicode_minus'] = False matplotlib.font_manager as fm |
| import os.chdi 출신 In [105 # 분석배 | ir('C:\Python\Syntex\따로 공부\출산율') 난율이 현재 사회에 얼마나 큰 이슈인지 알아보기 |
| Out[105 0 1 | 기사제목 '주4일 근무' 충남도, 24시간 365일 완전 돌봄 약속 '파리지앵'들이 사라지는 이유[조은아의 유로노믹스] 중시 '출산율 제고' 방안 내놔냉동난자 사용 보조생식술 지원 첫 시행 |
| 3 4 65 "≌ 66 | 2030년 세계 인구 전환점이 온다 군포시 '필수 가임력 검사비' 지원출산율 높이기 총력 엄마 놀이터에 친구가 없어"놀이 기구 대신 자리 잡는 어르신 의자들 경북 유일 신생아 집중 치료 시설 개소저출생 극복 디딤돌 |
| 67 68 69 70 rows 3 | "경력단절 우려에 출산율 하락노동시장 이중구조 개혁해야" '10명 출산' 모성 영웅 칭호북한도 '저출산' 고민 여당發 '부가세 인하론'5년 전 정부에선 "2%P 인상" 권고 나왔다 × 1 columns |
| 그래서 어떤게 | 는 기삿거리가 쏟아지는걸 알수 있습니다. 현 상황이 어떤지 문제점인지 알아보고자 합니다. 물을 알아보도록 하겠습니다. |
| file_a file_a ax = fi file_a Out[106 | <u>리문제인 출산을 확인</u> = pd.read_excel('출생아수_합계출산을_자연증가_등_20240401151505.xlsx') # 출산을 = file_a.set_index("기본항목별") ile_a.loc["출생아수(명)"] 2019 2020 2021 2022 2023 기본항목별 |
| 자연증 조출생 자연증가 | 출생아수(명) 302676.000 272337.000 260562.000 249186.000 230000.00 중가건수(명) 7566.000 -32611.000 -57118.000 -123753.000 -122800.00 생물(천명당) 5.900 5.300 5.100 4.900 4.50 사물(천명당) 0.100 -0.600 -1.100 -2.400 -2.400 배출산물(명) 0.918 0.837 0.808 0.778 0.72 |
| In [107 # 그래프 plt.xla plt.yla plt.tit ax.plot | 출생성비(명) 105.500 104.800 105.100 104.700 105.10 |
| 30000 29000 | 5년간 출생 그래프 00 - 00 - |
| 28000 (B) 小 公 和 26000 25000 | |
| 24000 23000 | 00 - 2019 2020 2021 2022 2023 년도 |
| past = today = decreas birth_r print(f | 변부터 2023년까지 출산물 ax.iloc[0] # 과거 se = xt.iloc[0] # 과거 se = totay.values - past.values # 감소한 인구 rate = np.array((decrease)/ past.values*100,dtype=np.float16) # 퍼센트 f'2019년부터 2023년까지 감소한 출산물은 {birth_rate}%입니다.') 터 2023년까지 감소한 출산물은 [-24.02]%입니다. 생아들에 수는 굉장히 빠른 속도로 줄어들고있습니다. |
| ax_a = ax_a['ax_a.se ax_a | 마나 감소한 수 ex.name = "년도" ax.reset_index(drop=False) 갑소수'] = ax_a['출생아수(명)'].sub(ax_a.loc[ax_a['년도'] == '2019', '출생아수(명)'].iloc[0]) et_index('년도', inplace=True) 출생아수(명) 감소수 |
| 2021 2022 | 302676.0 0.0 272337.0 -30339.0 260562.0 -42114.0 249186.0 -53490.0 |
| In [110 #%% | ax_a.drop(labels='출생아수(명)',axis=1) gure(figsize=(12, 8)) lot() lot() |
| | size 1200x800 with 0 Axes> - 출생아수(명) - 감소수 |
| 150000 - 100000 - 50000 - | |
| -50000 - o - | 2019 2020 2021 2022 2023 년도 - 감소수 |
| -10000 - -20000 - -30000 - | |
| -40000 - -50000 - -60000 - -70000 - | |
| 약 5 년 출산율 | 2019 2020 2021 2022 2023 년도 2020 2021 2023 년도 2020 2021 2023 년 전 출산율 감소는 -24.02% 합계 출산율은 0.72%인것을 알수 있습니다. |
| 접근해. 다음어 현재 <i>(</i> | 로가상승률이 가장큰 요인이라고 생각하고 보았습니다. 에 알아볼것은 신혼,혼인건수부터 연령별 출산율을 통해서 상황에서 어떤게 문제인지 조금더 자세히 알아보도록 하겠습니다. |
| file_c In [115 # 신혼부 by_year | r_all = file_b[['행정구역별','2019','2020','2021','2022']].set_index('행정구역별') r = by_year_all.loc['행정구역별':'전국'] # 매해 전국 신혼 부부수 |
| 행정구역 행정구역 전 In [116 # 2022년 Decreas | 변별 신혼부부 수 (A) 선혼부부 수 (A) 선혼부부 수 (A) 선혼부부 수 (A) 선혼부부 수 (B) 선조 1260117 1183750 1101455 1032253 선 - 227864= 감소한 수 -227864 se_in_newlyweds = by_year.iloc[1,3] - by_year.iloc[1,6] |
| Decreas # [(202 Decreas print(f print(f 2019년부터 2019년도 In [117 # 맞벌이 Dual_in | se_in_newlyweds = by_year.iloc[1,3] - by_year.iloc[1,0] 22 - 2019) + 2019] × 100 = 감소율 - 18.08 5e_rate_of_newlyweds = np.array(Decrease_in_newlyweds / by_year.iloc[1,0] * 100,dtype=np.float16) f'2019년부터 2022년까지 줄어든 인구수: {Decrease_in_newlyweds}') f'2019년도부터 2022년까지에 감소율은: {Decrease_rate_of_newlyweds}') 너 2022년까지 줄어든 인구수: -227864 부터 2022년까지에 감소율은: -18.078125 기 비율 ncome_ratio = file_b[['행정구역별','2019.2','2020.2','2021.2','2021.2','2022.2']].set_index('행정구역별').loc['행정구역별':'전국'] |
| Dual_in # 주택보 Homeown Dual_in | ncome_ratio_A = Dual_income_ratio.iloc[1,3] - Dual_income_ratio.iloc[1,0] 보유 비율 mership_ratio= file_b[['행정구역별','2019.4','2020.4','2021.4','2022.4']].set_index('행정구역별':'전국'] ncome_ratio 2019.2 2020.2 2021.2 2022.2 |
| 전 In [118… print(f | 변국 46.3 49 51.7 53.8 f'2019년부터 2022년까지 총 {Dual_income_ratio_A}% 증가했습니다.') 터 2022년까지 총 7.5% 증가했습니다. |
| | 변발 |
| <pre>x = By_ y = By_ plt.fig xy = pl plt.tit plt.tex plt.tex plt.tex</pre> | = file_c.iloc[[0,1]].set_index('행정구역별') # 전국 해당연령의 출산수 _age.iloc[0] _age.iloc[1] gure(figsize=(12, 8)) lt.bar(x.index,y.values,color = ['skyblue','r','orange','pink']) tle('5년간 출산 연령대',fontsize=20) xt(-0.8,15, "20-24세",fontsize=15) xt(0.2,43, "25-294",fontsize=15) xt(2.5,80, "30-344",fontsize=15) xt(2.5,80, "35-3940 &',fontsize=15) |
| plt.xla for red hei plt plt.yla plt.xti Out[120 ([<matp <matp<="" td=""><td>abel("년도", fontsize=18) ct in xy: ight = rect.get_height() t.text(rect.get_x() + rect.get_width()/2.0, height, '%.1f' % height, ha='center', va='bottom', size = 18) abel("여자인구 1천명당 명", fontsize=18) icks(['2019','2020','2021','2022'],['2019','2020','2021','2022']) cplotlib.axis.XTick at 0x23596683c90>, cplotlib.axis.XTick at 0x2359663f290>, cplotlib.axis.XTick at 0x23596d0f90>, cplotlib.axis.XTick at 0x23596d0f90>,</td></matp> | abel("년도", fontsize=18) ct in xy: ight = rect.get_height() t.text(rect.get_x() + rect.get_width()/2.0, height, '%.1f' % height, ha='center', va='bottom', size = 18) abel("여자인구 1천명당 명", fontsize=18) icks(['2019','2020','2021','2022'],['2019','2020','2021','2022']) cplotlib.axis.XTick at 0x23596683c90>, cplotlib.axis.XTick at 0x2359663f290>, cplotlib.axis.XTick at 0x23596d0f90>, cplotlib.axis.XTick at 0x23596d0f90>, |
| <mat<sub>l [Text Text Text</mat<sub> | plotlib.axis.XTick at 0x23596d65110>], (0.0, 0, '2019'), (4.0, 0, '2020'), (4.0, 0, '2020'), (5.0, 0, '2020'), (5.0, 0, '2020')) (5.0, 0, '2020') (12.0, 0, '2022')]) 5년간 출산 연령대 86.2 |
| 이루고 계속되 | 20~24세 7.1 6.2 5.0 4.1 2019 2020 2021 2022 년도 보면 출산은 대체적으로 30대가 주를 1 있지만 그조차 줄어들고 있는 추세고 1는 인구 감소로 인한 젊은 총의 수는 줄어들고 |
| 젋은 <i>층</i> 또한 전 결혼 비 | 은 층이 감당해야 할 부양층들의 수가 늘어 더욱 더 등은 부담감이 강해질것으로 예상 할 수 있습니다. 전체통계 자료를 확인해보면 비율,주택보유 비율,출산율은 현저히 떨어지고 있는 반면에 |
| In [122 # 월급과 file_d Earned_ year19 year23 # 5년간 | 도 부담감은 증가하는것을 알수 있습니다. |
| years_5 Earned_ Out[122 가계수지 | 5 = np.array((year23-year19) / year19 * 100,dtype=np.float16)income_by_year 2019.1/4.1 2020.1/4.1 2021.1/4.1 2021.1/4.1 2021.1/4.1 2023.1/4.1 지한목별 |
| 그 In [123 print(f 5년간 근토 In [125 # <i>5년간</i> | 임수 (명) 2.61 2.53 2.49 2.44 2.39 |
| file_e populat Populat Percent populat | 당안 서울 인근 지역이 전체 인구에 몇퍼센트 차지하는지 = pd.read_excel('행정구역_시군구, 별 인구수_2404043155513.xlsx') tion = file_e.iloc[[0,1,2,10]].set_index('0540740[4]군구)별') tion_near_Seoul = population.iloc[2] + population.iloc[3] # 서울과 근교 경기도의 함 tage_of_population_near_Seoul = Population_near_Seoul*100/population.iloc[1] # 전체 인구수에 차지하는 서울근교에 비율 tion 2019 2020 2021 2022 2023 2024 描(시군구)별 |
| | 경(시군구)별 총인구수 (명) 후인구수 (명) 후인구 |
| Out[126 2019 2020 2021 2022 2023 2024 dtype: | 44.29862 44.5609 44.685206 44.747734 44.845022 44.8603 |
| korea_m # 데이트 populat # 인구 대 | 지도 읽어오기 map = gpd.read_file('C:/Python/Syntex/따로 공부/출산율/TL_SCCO_CTPRVN.json') map['CTP_KOR_NM'] = korea_map['CTP_KOR_NM'].astype(str) 터 읽어오기 tion_data = pd.read_excel('C:/Python/Syntex/따로 공부/출산율/24년_지역별_인구수.xlsx') 데이터를 한국 지도에 포함시키기 _data = korea_map.merge(population_data, how='left', left_on='CTP_KOR_NM', right_on='행정구역(시군구)별') |
| # 설정 merged_ plt.tex ax.set_ | <pre>x = plt.subplots(figsize=(10, 8))</pre> |
| ♂ | 24년 지역별 인구 밀도 (서울.경기) 전체비율에 44% |
| | -1.0 -0.8 -0.6 -0.4 |
| | - 0.2 - 는 감소하는 추세에도 |
| 만약 수 In [129 # 단위는 # 전국 : file_f file_g file_h | 수도권 지방 = pd.read_excel('평균매매가격_단독.xlsx') = pd.read_excel('평균매매가격_아파트.xlsx') = pd.read_excel('평균매매가격_연립다세대.xlsx') |
| # 서울 : file_i file_j file_k # 5년간 In [130 # 증감률 def Nat | 경기 = pd.read_excel('서울.경기_평균매매가격_단독.xlsx') = pd.read_excel('서울.경기_평균매매가격_아파트.xlsx') = pd.read_excel('서울.경기_평균매매가격_여립다세대.xlsx') # 구하는 함수 tionwide(df): anges = [] # 중감률 r i in df.index: |
| ret In [131 single Metropo province | r i in df.index: val1 = df.iloc[i,1] # 2019년도 val1_1 = df.iloc[i,5] # 2023년도 change = np.array((val1_1 - val1) / val1 * 100, dtype=np.float16) # 공식 changes.append(change) turn changes = Nationwide(file_f) # (전국, 수도권, 지방 5년간 증가을 - 단독) olitan_area = Nationwide(file_g) # (전국, 수도권, 지방 5년간 증가을 - 아파트) ce = Nationwide(file_h) # (전국, 수도권, 지방 5년간 증가을 - 연립다세대) |
| single_ Metropo provinc In [132 Increme Increme | _A = Nationwide(file_i) # (서울,경기 5년간 증가율 - 단독) olitan_area_B = Nationwide(file_j) # (서울,경기 5년간 증가율 - 아파트) ce_C = Nationwide(file_k) # (서울,경기 5년간 증가율 - 아파트) ce_C = Nationwide(file_k) # (서울,경기 5년간 증가율 - 연립다세대) ent_rate_A = pd.DataFrame(data = [single, Metropolitan_area, province], columns = ['전국', '수도권', '지방권'], index = ['단독', '아파트', '연립다세대']) ent_rate_A 전국 수도권 지방권 단독 24.44 20.84 19.5 |
| 아파 연립다세 In [133 Increme Increme | 라트 30.48 35.1 24.17 세대 21.44 33.3 9.14 ent_rate_B = pd.DataFrame(data = [single_A,Metropolitan_area_B,province_C],columns = ['서울','경기'],index = ['단독','아파트','연립다세대']) ent_rate_B |
| 아파 연립다세 | 서울 경기 한동 19.3 26.84 타로 28.78 48.0 대한 32.72 29.95 |
| In [135 # 写계년 def xyb ind col val fig | # 자료 그래프 함수 bar(df): dex = df.index lums = df.columns lues = df.values g, ax = plt.subplots(figsize=(10, 10)) r_width = 0.25 d = range(len(colums)) |
| ax. ax. ax. ax. ax. | c = range(len(colums)) |
| A = xyb | t.show() bar(Increment_rate_A) bar(Increment_rate_B) 5년간에 지역별 매매가격 증가율 다동하는 이라는 이라는 이라는 이라는 이라는 이라는 이라는 이라는 이라는 이라 |
| 30- 25- 多化会 化約 10- 50- 50- | 전국 지형권 지역 5년간에 지역별 매매가격 증가율 |
| | 지역 지역 장크를 확인하고 1 전문 24 27% 정도 모르겠어 반경 모은 연구수가 있는 경기 기존 |
| 적게는 이런 데 사람들(알수 있 - 분스 - 투 | : 26.84% 많게는 48%가량 오른것을 확인해 볼수 있습니다. 데이터로 가지고 유추해 볼수 있는것은 데 노동소득만으로는 결혼,자녀에 부담감이 올라갈수밖에 없는것을 있습니다. 선 활용: 과거보다 현재에 2030에 부담감이 해결되지 않으면 출산율은 오르지 않을 것입니다. 부담감을 즐여주기위해 정책은 필수라고 생각합니다. |
| - 3 | 부부 혜택 증가하는 정책: 주거지에 대한 불안을 해소하기 위해 민간분양,매매대출같은 정책을 강화 양육비용 지원(출산장려금,아동수당)같은 정책에 조건 완화 |

대한민국 출산율에 관한 통계분석으로 해결방안 알아보기

데이터 분석

- 데이터 선정:

1. 국가 공개포털 KOSIS

- 근로소득 정보

- 지역별 인구수

- 연령별 출산율 - 합계 출산율

2. 한국부동산원 부동산통계정보시스템(R-ONE)

• 월급과 주택에 비교로 현재 심적부담이 어느정도로 강해졌는지 알아보기

- 서울·경기 평균 매매가

- 인구밀도 - 신혼 부부

- 거주지 평균 매매가

• 현재 출산율이 얼마 얼마나 떨어졌는가?

• 5년간 출산 연령대는 어떻게 되는가?

if 'google.colab' in sys.modules:

- 데이터 분석과 해석:

• 신혼부부에 수는 어떻게 되는가?

맞벌이 비율주택 소유율

In [103... **import** sys