
코로나 거리두기 리포트 천안시

코로나 거리두기

제 1부 천안시의 2년의 기록

사용 데이터

제 1부 천안시의 2년의 기록

1. 2019년 03월부터 2019년 04월까지 천안시의 유동인구 데이터
2. 2020년 03월부터 2020년 04월까지 천안시의 유동인구 데이터
3. 천안시 코로나 확진자 데이터
4. 천안시 날씨 데이터

성별, 연령대별 유동인구 수

제 1부 천안시의 2년의 기록

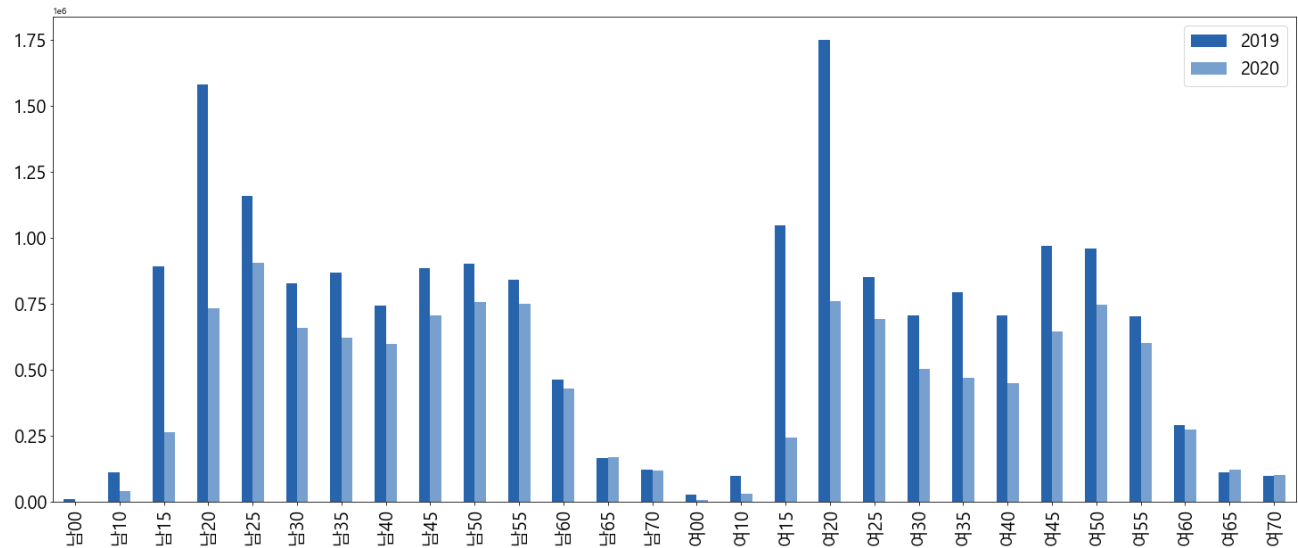
아래의 그래프는 2019년 03월에서 04월까지의 유동인구와 2020년 03월에서 04월까지의 유동인구를 연령대, 성 별로 표시한 그래프이다.

전반적으로 2019년의 유동인구보다 2020년의 유동인구가 적음을 확인할 수 있다.
2019년의 유동인구 수의 합은 약 158,413,589 명으로 확인되었고 2020년의 유동인구 수의 합은 122,854,457 명이다.
또한 2019년과 2020년의 연령대별 유동인구는 남자가 여자보다 비율이 높고, 20세에서 25세 사이의 인구의 유동인구 수가 가장 많다.

2020년 평균 유동인구 수와 2019년 평균 유동인구 수의 차이

35,559,132명 = -22.45%의 감소율

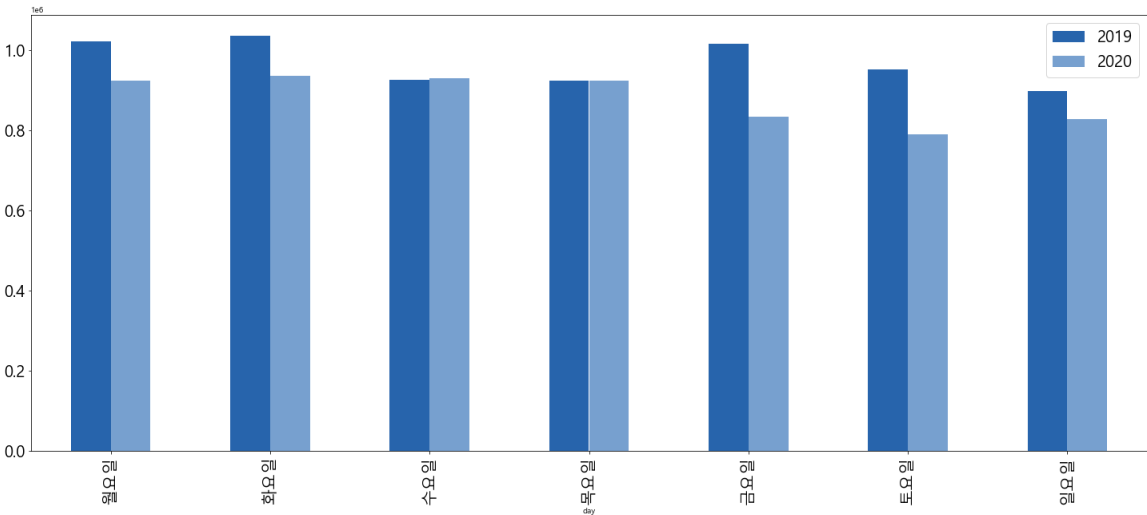
전체 연령대별 유동 인구 수



요일별 유동인구 수

제 1부 천안시의 2년의 기록

전체 요일별 유동 인구 수

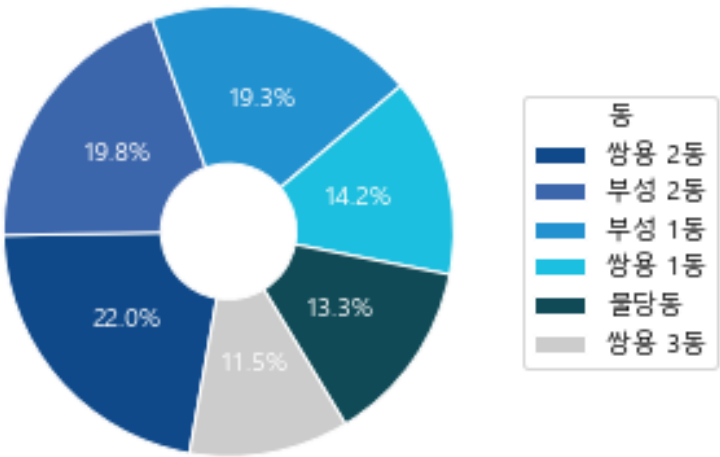


행정구역 별 거주 인구 수

제 1부 천안시의 2년의 기록

천안시의 5개의 행정동을 기준으로 성별, 연령대별로 유동인구 수를 비교해보았다.
전체 인구는 쌍용 2동, 부성 2동, 부성 1동, 쌍용 1동, 불당동, 쌍용 3동 순으로 인구가 많다는
것을 확인할 수 있다.

동 별 인구 비교



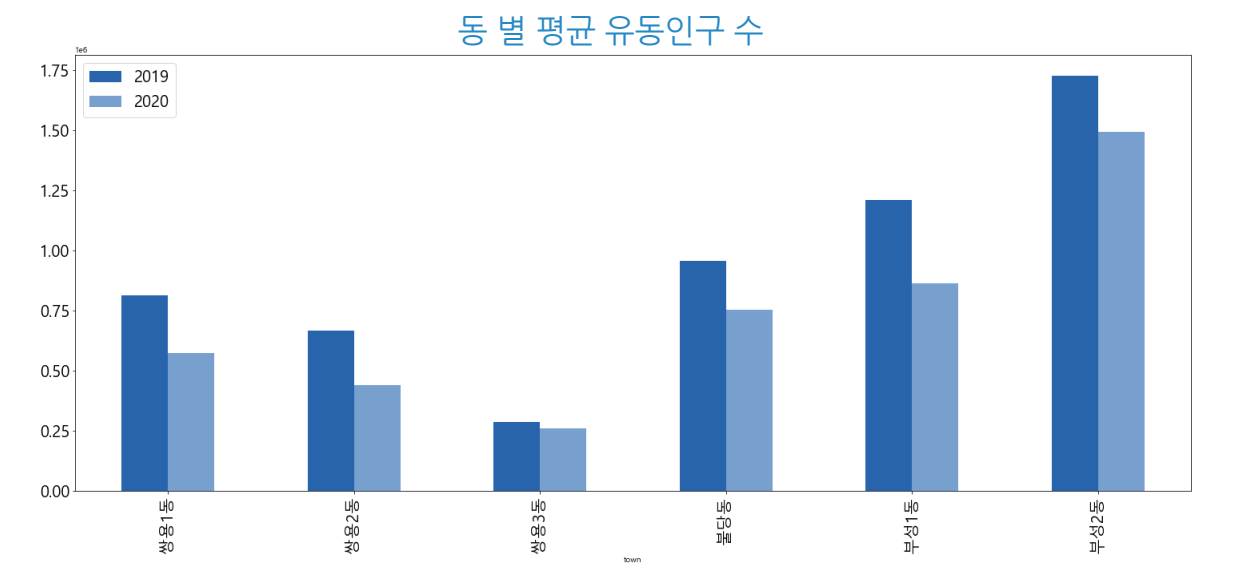
순위	동 이름	인구 수
1	쌍용 2동	42491
2	부성 2동	38238
3	부성 1동	37447
4	쌍용 1동	27544
5	불당동	25656
6	쌍용 3동	22181

행정구역 별 유동인구 수

제 1부 천안시의 2년의 기록

천안시의 6개의 행정 구역의 평균 유동인구 수를 비교해 표로 나타냈다.
천안시의 행정동의 유동인구 수와는 다른 결과를 보였다. 인구로 따져보았을 때는 쌍용 1동의 인구가 가장 많지만 유동인구 수로 비교했을 때는 부성 2동의 유동인구 수가 2019년, 2020년 모두 유동인구 수 가 많은 것을 알 수 있다.

이는 행정구역별로 산업단지인지 혹은 거주 지역인지에 따라 유동인구가 달라질 수 있다.
각 행정 구역별 특징들을 다음 장에서부터 세부적으로 다룰 것이다.



순위	동 이름	유동 인구 수 (2019)	유동인구 수 (2020)
1	부성 2동	1726805	1493208
2	부성 1동	1209643	864795
3	불당동	956064	753710
4	쌍용 1동	812425	573482
5	쌍용 2동	666723	441935
6	쌍용 3동	285969	260529

쌍용 1동 유동인구 분석

제 1부 천안시의 2년의 기록

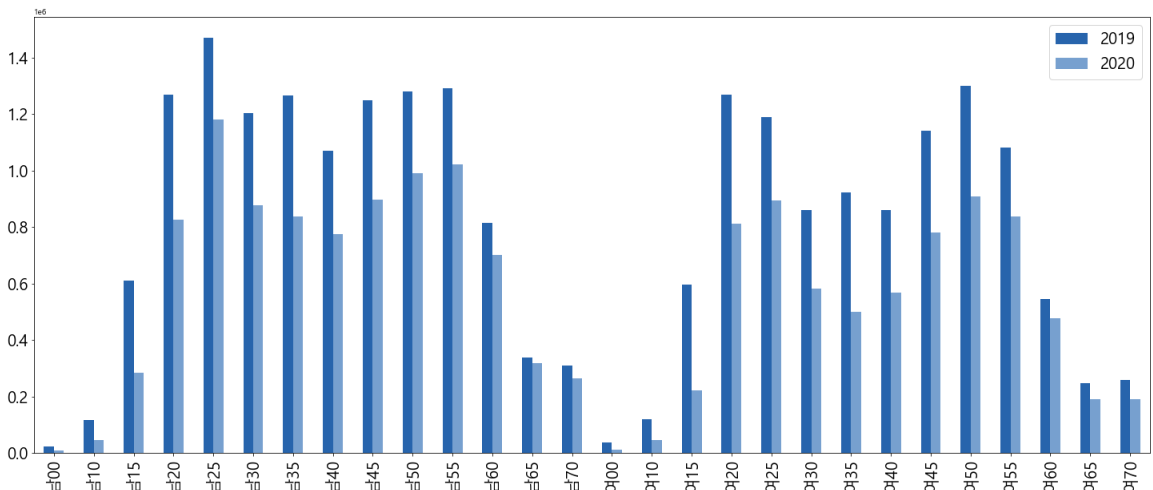
아래의 그래프는 쌍용 1동의 2019년 03월에서 04월까지의 유동인구와 2020년 03월에서 04월까지의 유동인구를 연령대, 성 별로 표시한 그래프이다.

전반적으로 2019년의 유동인구보다 2020년의 유동인구가 적음을 확인할 수 있다.
2019년의 평균 유동인구 수의 합은 22,747,904명으로 확인되었고 2020년의 유동인구 수의 합은 약 16,057,481명이다.
또한 2019년과 2020년의 연령대별 유동인구는 남자가 여자보다 비율이 높고, 20세에서 25세 사이의 인구의 유동인구 수가 가장 많다.

2020년 평균 유동인구 수와 2019년 평균 유동인구 수의 차이

6,690,423명 = -29.4%의 감소율

쌍용 1동 연령대별 유동 인구 수



쌍용 1동

제 1부 천안시의 2년의 기록

유동인구 감소율은 29.4%

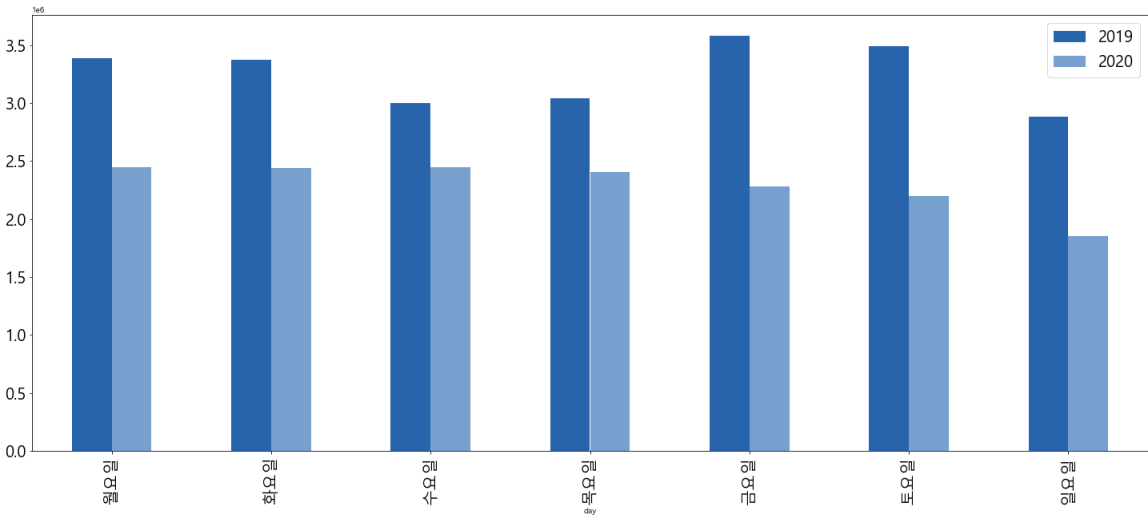
쌍용 1동의 경우 29.4%이다. 쌍용 1동의 경우 대형교회가 많이 밀집되어 있다. 또한 천안 내 2차 의료기관인 천안 충무병원, 노인 복지관, 재활센터 등이 존재한다. 그렇기에 유동인구 감소율이 높은 것으로 판단되어 진다.



쌍용 1동 요일별 유동인구 수

제 1부 천안시의 2년의 기록

쌍용1동 요일별 유동 인구 수



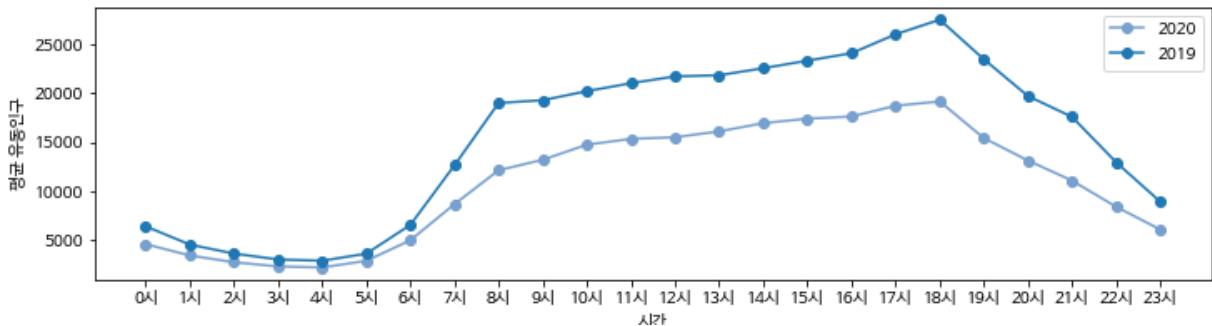
쌍용 1동 시간대별 유동인구 수

제 1부 천안시의 2년의 기록

쌍용 1동의 경우 앞서 확인 했듯이 상업 지역이다. 천안 내 2차 의료기관인 천안 충무병원, 노인 복지관, 재활센터 등이 존재했다. 또한 주거지역보다는 상업지역이기에 8시부터 유동인구가 점점 늘어나고 저녁시간인 18시대 최대를 찍고 점점 유동인구가 줄어드는 형태를 취한다.

2019년과 2020년 모두 시간대별로 같은 형태의 그래프를 그렸다. 하지만 2020년은 2019년에 비해 매 시간대 유동인구가 감소했다.

쌍용1동 시간대별 유동 인구 수

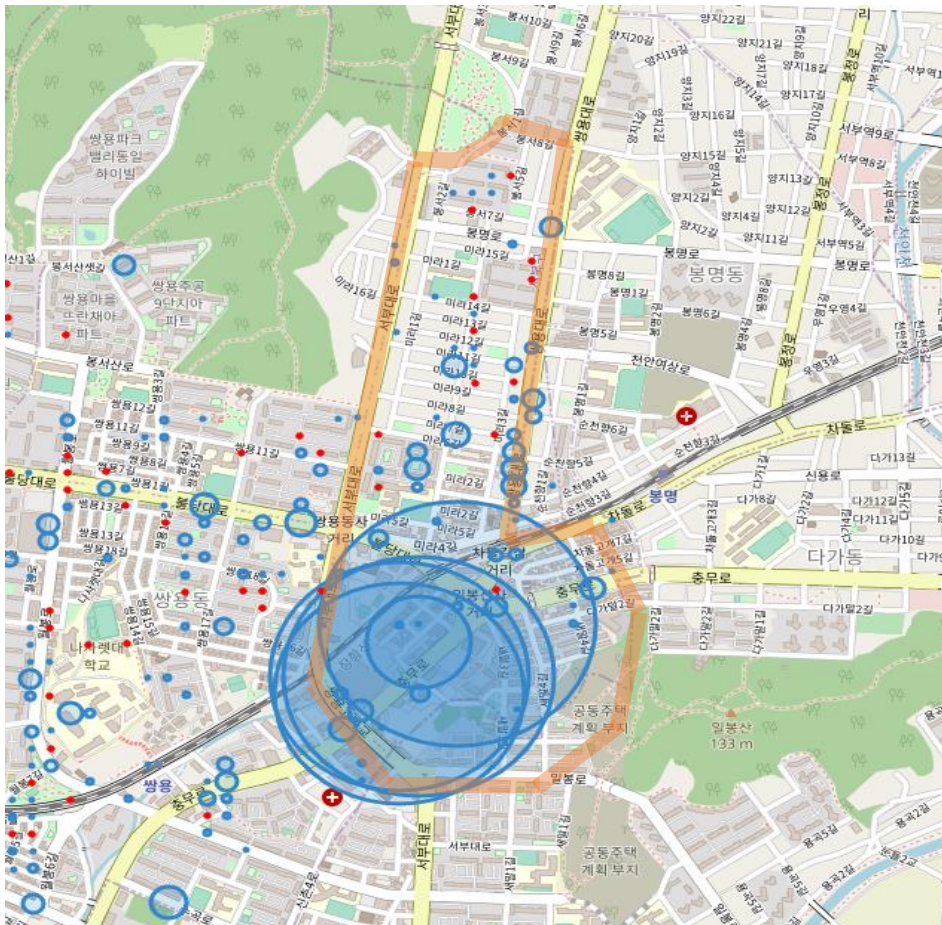


쌍용 1동 Circle map

제 1부 천안시의 2년의 기록

쌍용 1동의 지도 위에 위경도별로 (2019년 유동인구 수 - 2020년 유동인구)의 값을 이용해 양수 값이라면 빨간색 원을 그려 표시했다. 즉, 2020년 유동인구 수가 더 늘어났다는 것을 뜻한다. 반대로 음수 값이라면 파란색 원을 그려 표시했다. 즉, 2019년 유동인구 수가 더 늘어났다는 것을 뜻한다. 이때 각 원의 크기는 (2019년 유동인구 수 - 2020년 유동인구)의 값이 커지면 원의 크기가 커진다.

쌍용1동 Circle marker

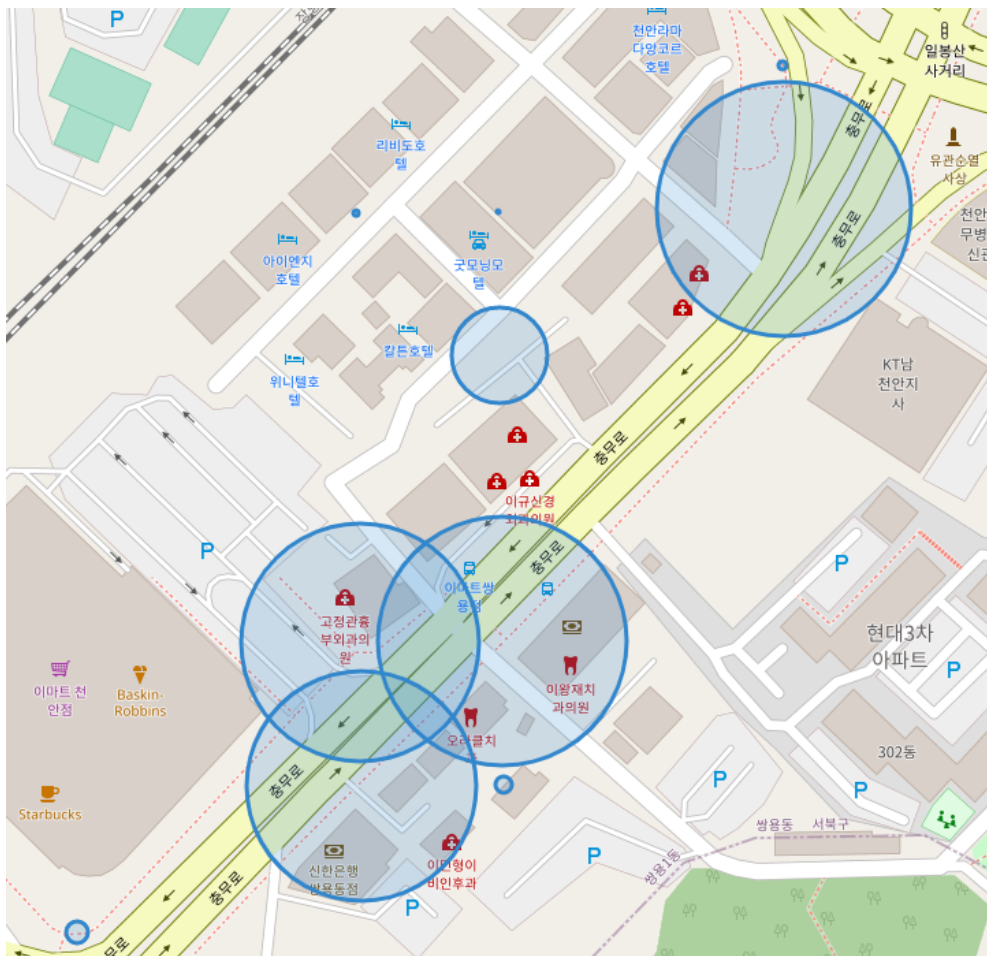


쌍용 1동 Circle map

제 1부 천안시의 2년의 기록

(2019년 유동인구 수 - 2020년 유동인구 수)가 양수 값을 가지는데 가장 큰 원을 가지는 부분을 확대했을 때, 병원이나 마트가 위치한 상업지구가 2019년이 2020년에 비해 유동인구가 급격히 감소한 지역이다.

쌍용1동 Circle marker



쌍용 1동 Circle map

제 1부 천안시의 2년의 기록

(2019년 유동인구 수 - 2020년 유동인구 수)가 음수 값을 가지는 부분을 확대했을 때, 아파트와 같은 주거시설이 2019년 유동인구에 비해 2020년 유동인구가 증가한 부분이다. 이때, 주목할 점은 각각 아파트 동 외에 아파트 내 단지 시설은 오히려 2019년 유동인구가 더욱 많은 걸 확인할 수 있다.

쌍용1동 Circle marker



쌍용 2동 유동인구 분석

제 1부 천안시의 2년의 기록

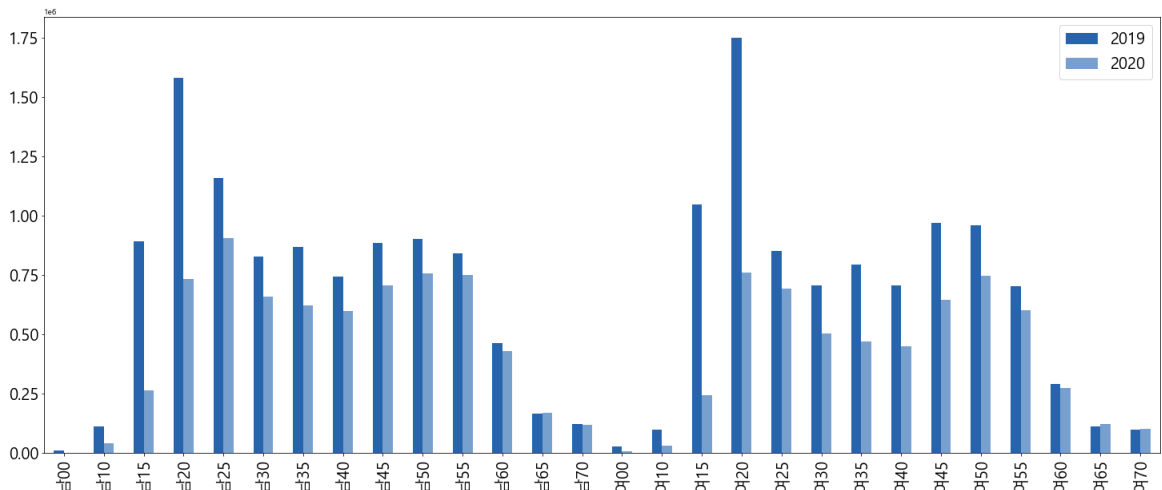
아래의 그래프는 쌍용 2동의 2019년 03월에서 04월까지의 유동인구와 2020년 03월에서 04월까지의 유동인구를 연령대, 성 별로 표시한 그래프이다.

전반적으로 2019년의 유동인구보다 2020년의 유동인구가 적음을 확인할 수 있다.
2019년의 유동인구 수의 합은 약 18,668,233 명으로 확인되었고 2020년의 유동인구 수의 합은 12,374,179 명이다.
또한 2019년과 2020년의 연령대별 유동인구는 남자가 여자보다 비율은 크게 차이가 없음을 확인할 수 있다, 또한 20세에서 25세 사이의 인구의 유동인구 수가 가장 높게 나타났다.

2020년 평균 유동인구 수와 2019년 평균 유동인구 수의 차이

6,294,054명 = -33.72%의 감소율

쌍용 2동 연령대별 유동 인구 수



쌍용 2동 유동인구 1위는 20대

제 1부 천안시의 2년의 기록

쌍용 2동의 경우, 쌍용역(나사렛대) 및 나사렛 대학교가 위치해 있다.그렇기에 20대의 유동인구가 가장 높은 비율을 차지했다.

또한 쌍용2동은 대부분 아파트로 이루어져 있는 주거지역이고, 초, 중, 고등학교가 위치했다. 쌍용수산물시장, 롯데마트, 도서관, 이마트 등이 위치 되어있다.

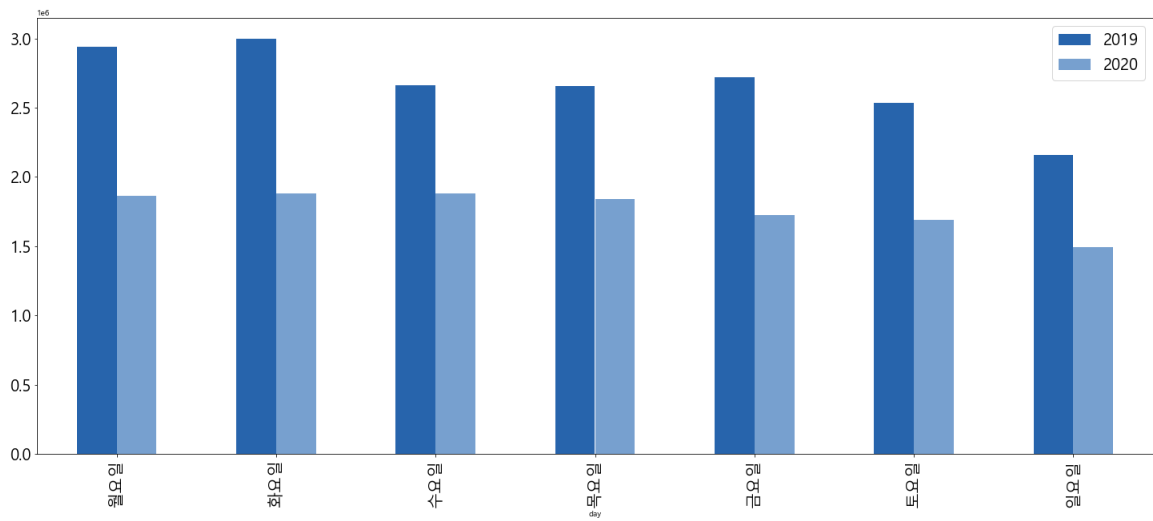
그렇기에 COVID-19 이후 온라인으로 변환된 나사렛 대학교의 영향과 청소년들의 비율이 많아 유동인구 수가 급감했다.



쌍용 2동 요일별 유동인구 수

제 1부 천안시의 2년의 기록

쌍용2동 요일별 유동 인구 수



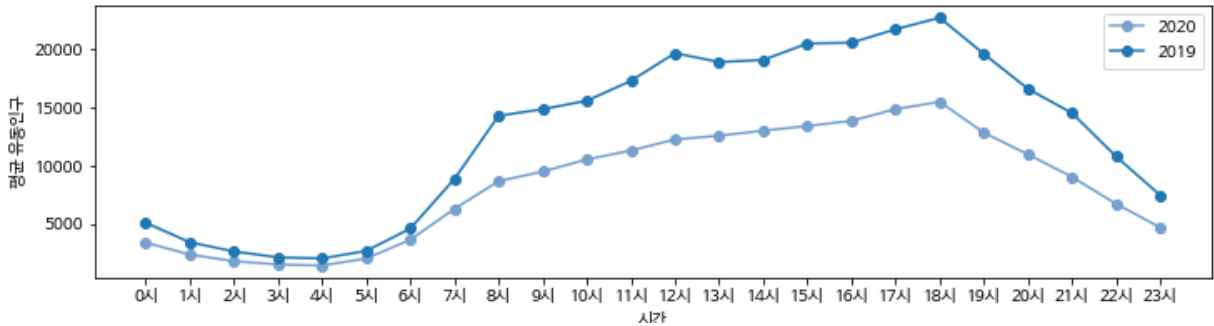
쌍용 2동 시간대별 유동인구 수

제 1부 천안시의 2년의 기록

쌍용 2동의 경우 앞서 확인 했듯이 주거 지역이다. 초등학교, 중학교, 고등학교와 나사렛 대학교가 위치해있다. 그렇기에 8시부터 유동인구가 점점 늘어나고 저녁시간인 18시대 최대를 찍고 점점 유동인구가 줄어드는 형태를 취한다.

2019년과 2020년 모두 시간대별로 같은 형태의 그래프를 그렸다. 하지만 2020년은 2019년에 비해 매 시간대 유동인구가 감소했다.

쌍용2동 시간대별 유동 인구 수

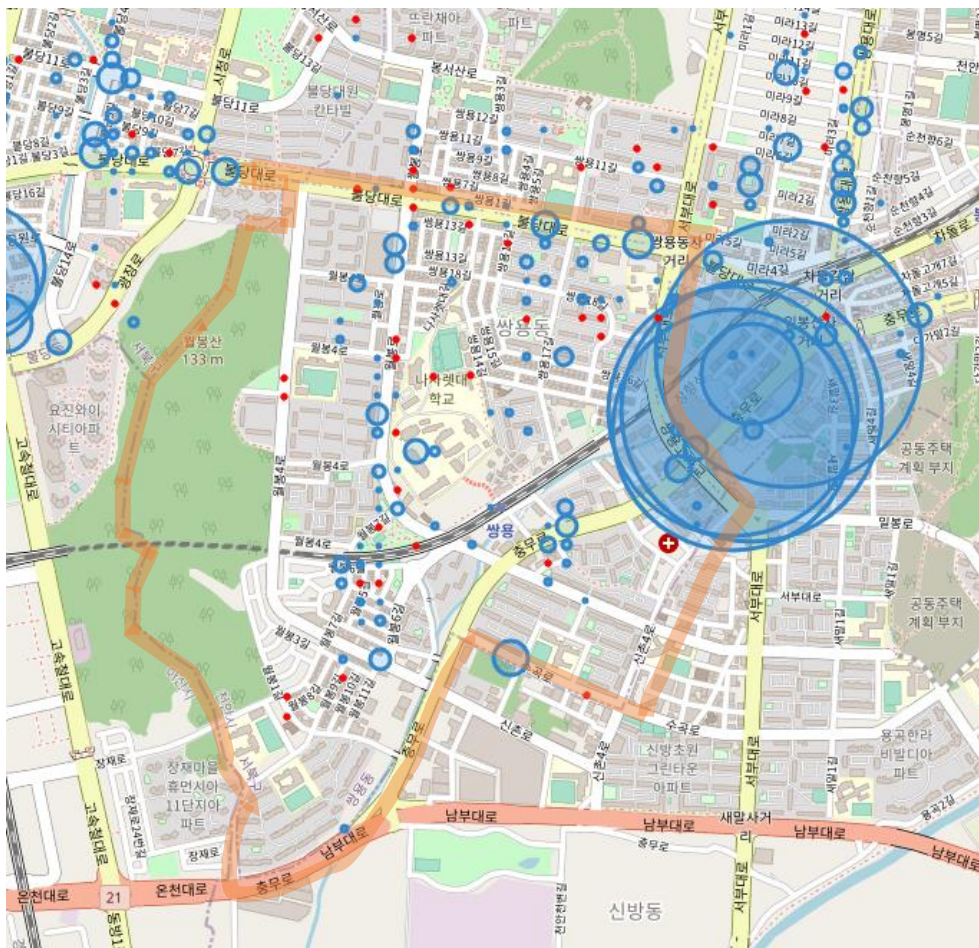


쌍용 2동 Circle map

제 1부 천안시의 2년의 기록

쌍용 2동의 지도 위에 위경도별로 (2019년 유동인구 수 - 2020년 유동인구)의 값을 이용해 양수 값이라면 빨간색 원을 그려 표시했다. 즉, 2020년 유동인구 수가 더 늘어났다는 것을 뜻한다. 반대로 음수 값이라면 파란색 원을 그려 표시했다. 즉, 2019년 유동인구 수가 더 늘어났다는 것을 뜻한다. 이때 각 원의 크기는 (2019년 유동인구 수 - 2020년 유동인구)의 값이 커지면 원의 크기가 커진다.

쌍용2동 Circle marker



제 1부 천안시의 2년의 기록

쌍용1동 Circle marker



쌍용 3동 유동인구 분석

제 1부 천안시의 2년의 기록

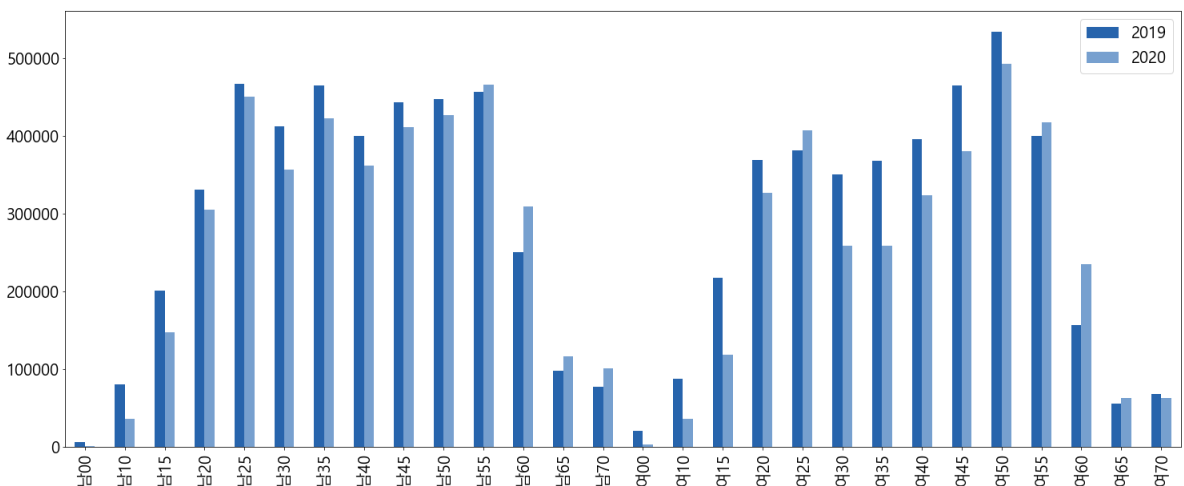
아래의 그래프는 쌍용 3동의 2019년 03월에서 04월까지의 유동인구와 2020년 03월에서 04월까지의 유동인구를 연령대, 성 별로 표시한 그래프이다.

전반적으로 2019년의 유동인구보다 2020년의 유동인구가 적음을 확인할 수 있다.
2019년의 유동인구 수의 합은 8,007,132 명으로 확인되었고 2020년의 유동인구 수의 합은 7,294,815 명이다.
또한 2019년과 2020년의 연령대별 유동인구는 남자가 여자보다 비율은 크게 차이가 없음을 확인할 수 있다, 주목할 점은 남자 50세 이상부터는 2020년도의 유동인구가 늘었고, 여자는 55세에서 60세 사이의 유동인구가 늘었다.
뿐만 아니라 다른 동에 비해 감소율이 적었다.

2020년 평균 유동인구 수와 2019년 평균 유동인구 수의 차이

712,317명 = -8.9%의 감소율

쌍용 3동 연령대별 유동 인구 수



쌍용 3동

제 1부 천안시의 2년의 기록

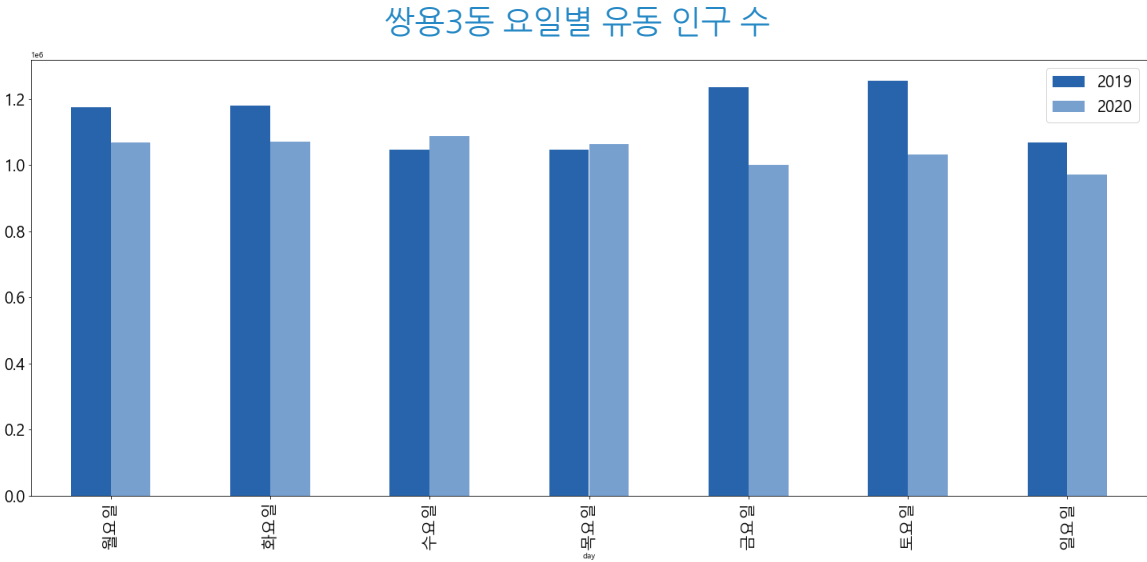
유동인구 감소율은 8.9%

쌍용 3동의 경우 아파트와 어린이집 및 우체국이 위치한 주거지역과 골프연습장으로 크게 두 구역으로 나뉘는 것을 확인했다. 전반적으로 남성 50대 이상 연령대에서 2020년 유동인구가 더욱 늘어난 이유는 골프 연습장이 가장 큰 영향을 미쳤다고 해석했다.



쌍용 3동 요일별 유동인구 수

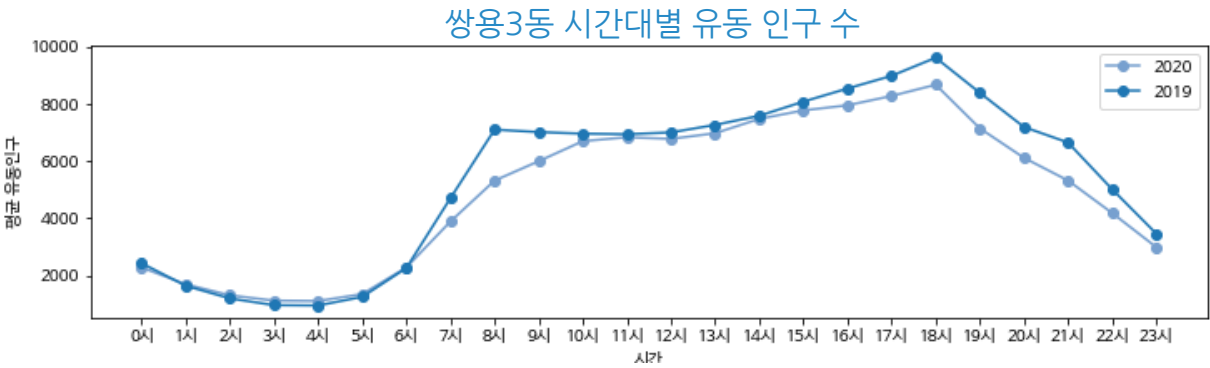
제 1부 천안시의 2년의 기록



쌍용 3동 시간대별 유동인구 수

제 1부 천안시의 2년의 기록

쌍용 3동의 경우 앞서 확인 했듯이 주거 지역으로 재택근무로의 변화 등으로 인해 집에서 생활하는 사람들이 늘어났다. 그렇기에 2019년과 2020년의 유동인구 변화의 큰 차이가 없는 형태를 보이고 있다.

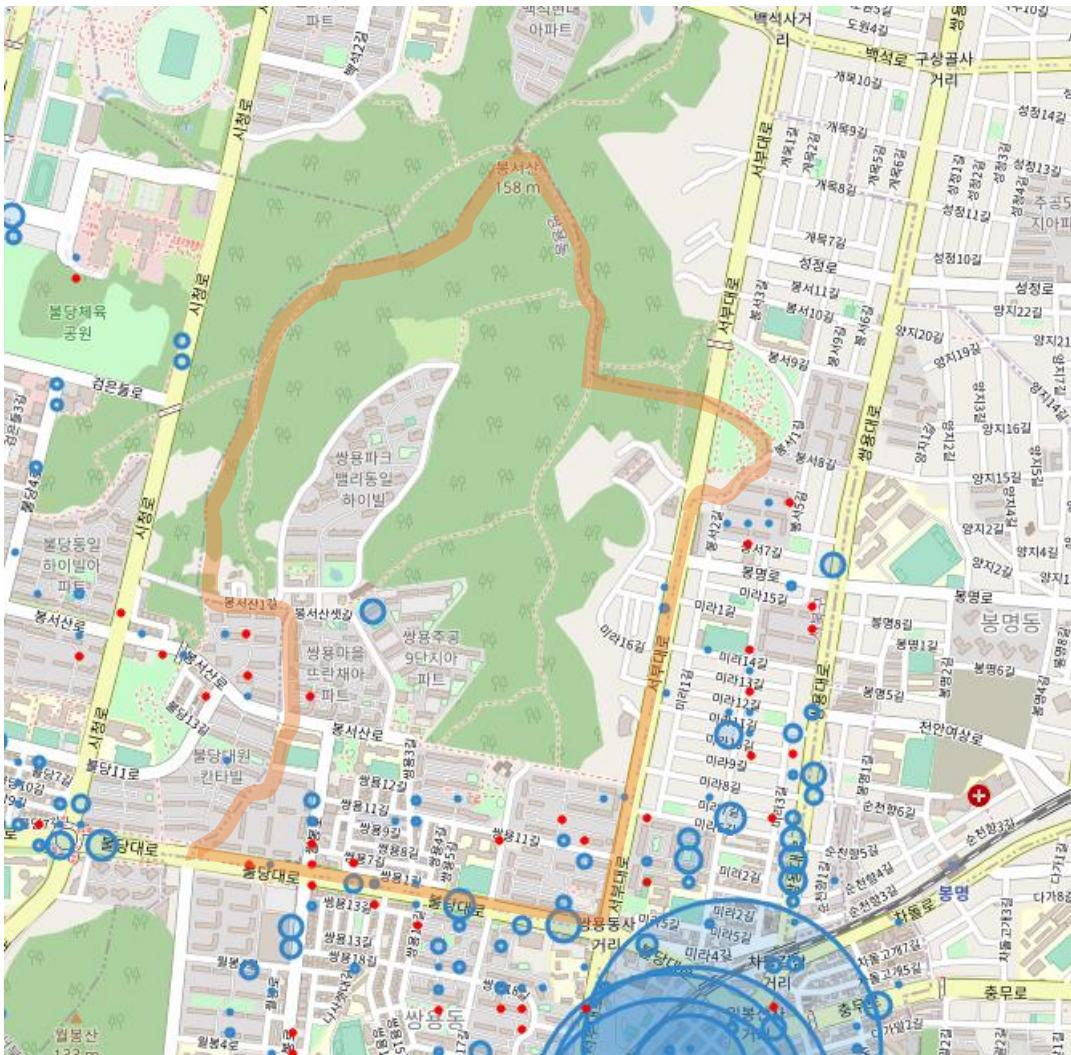


쌍용 3동 Circle map

제 1부 천안시의 2년의 기록

쌍용 3동의 지도 위에 위경도별로 (2019년 유동인구 수 - 2020년 유동인구)의 값을 이용해 양수 값이라면 빨간색 원을 그려 표시했다. 즉, 2020년 유동인구 수가 더 늘어났다는 것을 뜻한다. 반대로 음수 값이라면 파란색 원을 그려 표시했다. 즉, 2019년 유동인구 수가 더 늘어났다는 것을 뜻한다. 이때 각 원의 크기는 (2019년 유동인구 수 - 2020년 유동인구)의 값이 커지면 원의 크기가 커진다.

쌍용2동 Circle marker



제 1부 천안시의 2년의 기록

(2019년 유동인구 수 - 2020년 유동인구 수)가 양수 값을 가지는데 가장 큰 원을 가지는 부분을 확대했을 때, 아파트 내의 단지 시설에 양수값을 가지는 파란색 원이 존재했다. 뿐만 아니라 도로 위 또한 차로 이동하는 유동인구로 파란색 원을 가졌다.

빨간색 원을 가지는 부분을 확대했을 때는 초등학교와 아파트 특정동에 빨간색 원이 존재했다.

쌍용1동 Circle marker



불당동 유동인구 분석

제 1부 천안시의 2년의 기록

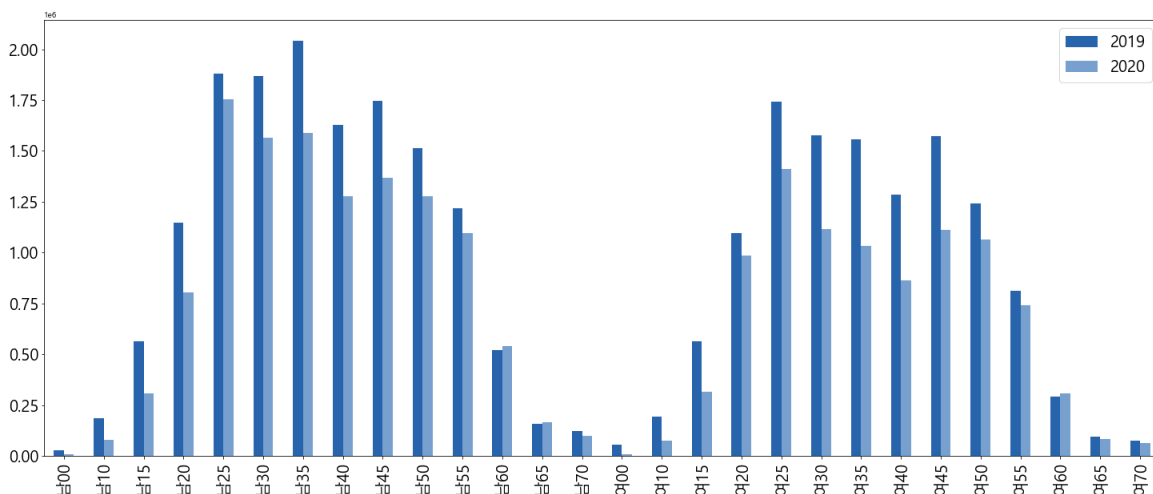
아래의 그래프는 불당동의 2019년 03월에서 04월까지의 유동인구와 2020년 03월에서 04월까지의 유동인구를 연령대, 성 별로 표시한 그래프이다.

전반적으로 2019년의 유동인구보다 2020년의 유동인구가 적음을 확인할 수 있다.
2019년의 유동인구 수의 합은 26,769,779명으로 확인되었고 2020년의 유동인구 수의 합은 21,103,876 명이다.
또한 2019년과 2020년의 연령대별 유동인구는 남자가 여자보다 비율은 크게 차이가 없음을 확인할 수 있다,

2020년 평균 유동인구 수와 2019년 평균 유동인구 수의 차이

5,665,903명 = -21.17%의 감소율

불당동 연령대별 유동 인구 수

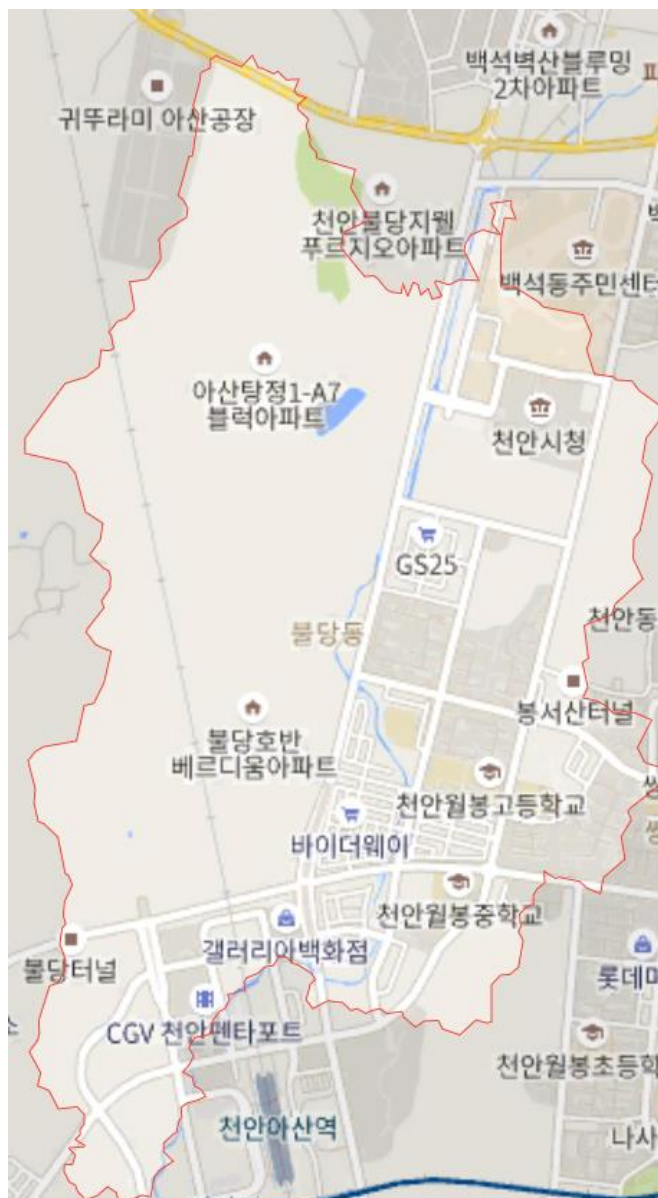


불당동

제 1부 천안시의 2년의 기록

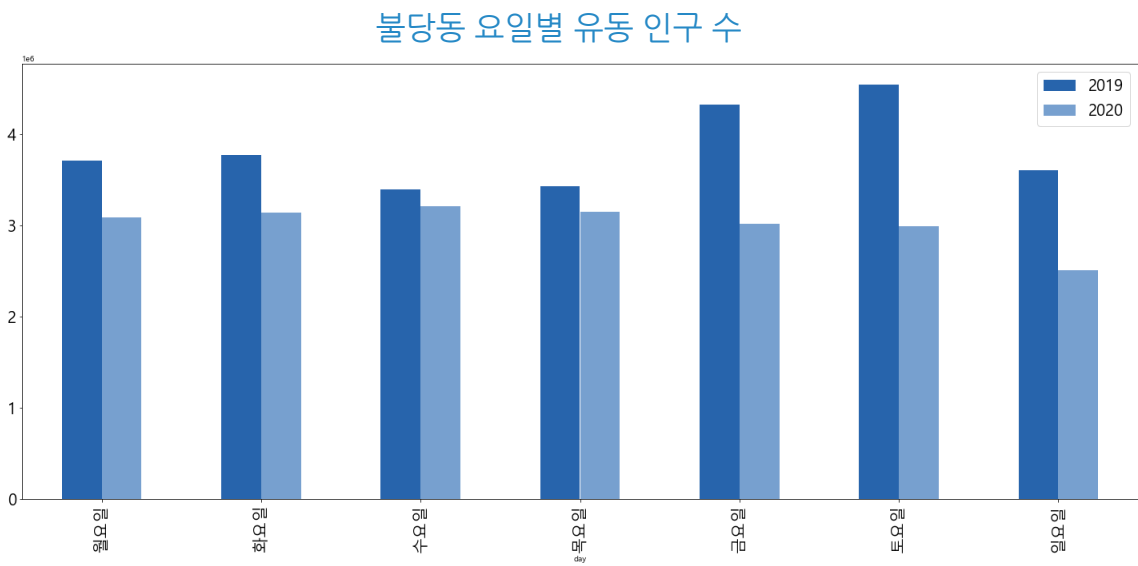
유동인구 감소율은 21.17%

불당동의 경우 카페거리가 존재하고 천안시청, 천안 교육 지원청을 제외하고는 대부분 아파트 및 주변 상가 지역이다. 하지만 갤러리아 백화점 및 CGV 등 문화 여가 생활을 즐길 수 있는 곳이기에 유동인구 감소율이 높을 것으로 판단되어 진다.



불당동 요일별 유동인구 수

제 1부 천안시의 2년의 기록

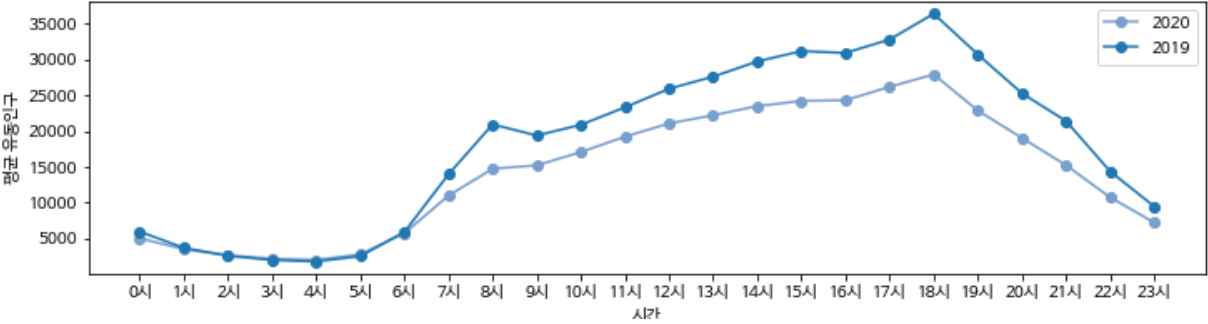


불당동 시간대별 유동인구 수

제 1부 천안시의 2년의 기록

불당동의 경우 앞서 확인 했듯이 상업 지역이다. 코로나 바이러스 이후에 상업지역에는 최소한으로 움직이게 하는 정부 거리두기 정책으로 2020년의 유동인구가 2019년의 유동인구보다 전체적으로 감소하는 형태를 보인다.

불당동 시간대별 유동 인구 수

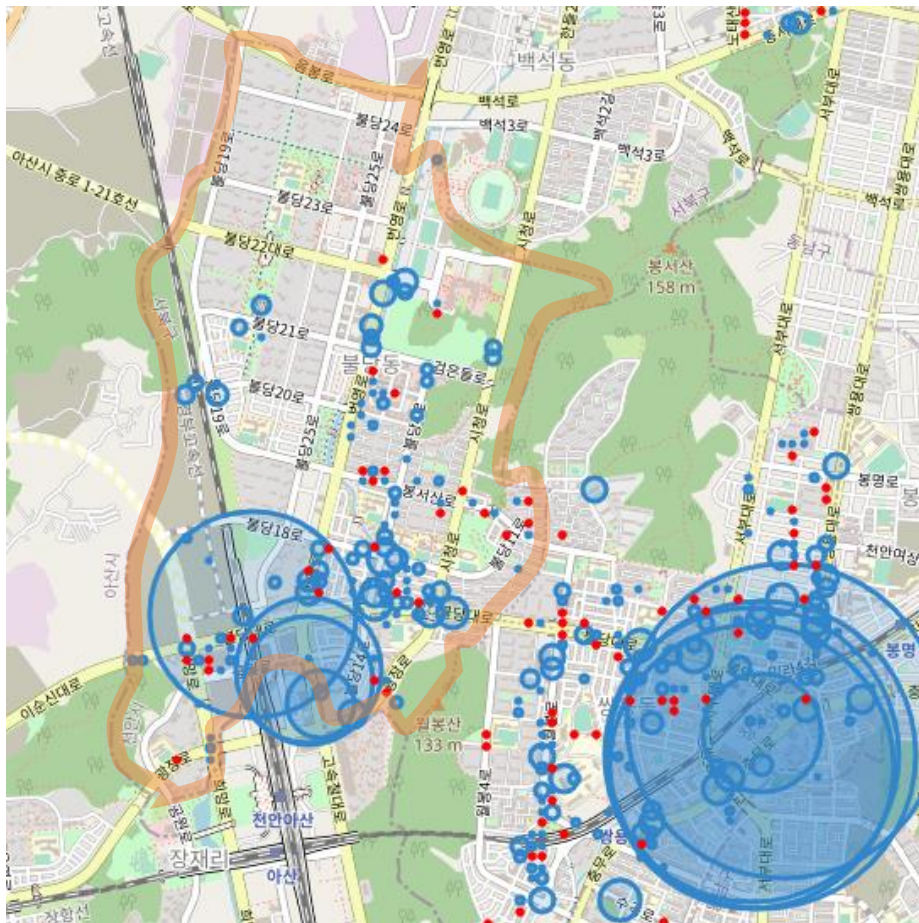


불당동 Circle map

제 1부 천안시의 2년의 기록

불당동의 지도 위에 위경도별로 (2019년 유동인구 수 - 2020년 유동인구)의 값을 이용해 양수 값이라면 빨간색 원을 그려 표시했다. 즉, 2020년 유동인구 수가 더 늘어났다는 것을 뜻한다. 반대로 음수 값이라면 파란색 원을 그려 표시했다. 즉, 2019년 유동인구 수가 더 늘어났다는 것을 뜻한다. 이때 각 원의 크기는 (2019년 유동인구 수 - 2020년 유동인구)의 값이 커지면 원의 크기가 커진다.

쌍용2동 Circle marker

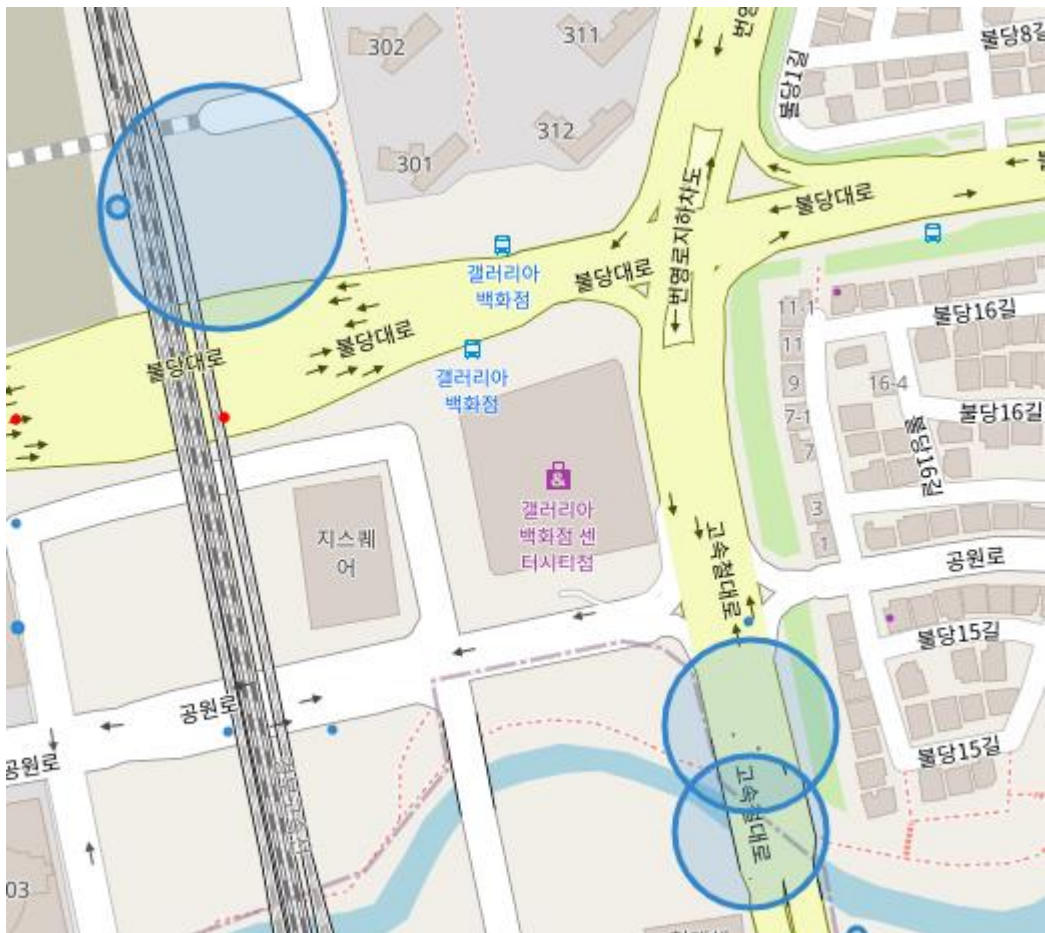


불당동 Circle map

제 1부 천안시의 2년의 기록

(2019년 유동인구 수 - 2020년 유동인구 수)가 양수 값을 가지는데 가장 큰 원을 가지는 부분을 확대했을 때, 고속도로에 양수값을 가지는 파란색 원이 존재했다. 그를 통해 차를 이용해 이동하는 유동인구도 많이 감소했다는 것을 알 수 있다.

쌍용1동 Circle marker



불당동 Circle map

제 1부 천안시의 2년의 기록

(2019년 유동인구 수 - 2020년 유동인구 수)가 음수 값을 가지는데 가장 큰 원을 가지는 부분을 확대했을 때, 아파트 내의 특정 동에는 빨간색 원이 있고 공동 단지에는 파란색 원이 존재했다.

쌍용1동 Circle marker



부성 1동 유동인구 분석

제 1부 천안시의 2년의 기록

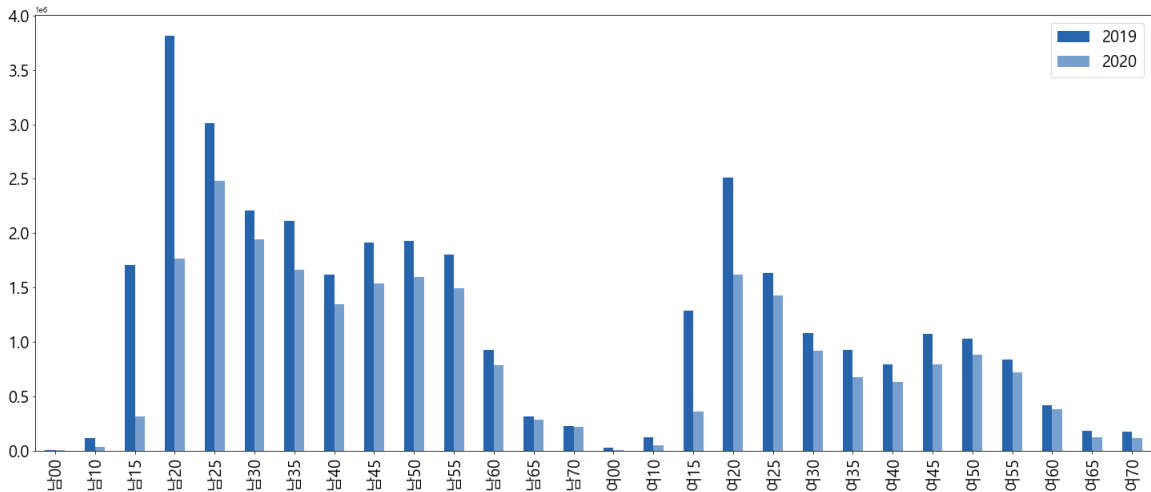
아래의 그래프는 부성 1동의 2019년 03월에서 04월까지의 유동인구와 2020년 03월에서 04월까지의 유동인구를 연령대, 성 별로 표시한 그래프이다.

전반적으로 2019년의 유동인구보다 2020년의 유동인구가 적음을 확인할 수 있다.
2019년의 유동인구 수의 합은 33,870,010 명으로 확인되었고 2020년의 유동인구 수의 합은 24,214,270 명이다.
또한 2019년과 2020년의 연령대별 유동인구는 남자가 여자보다 비율은 크게 차이가 없음을 확인할 수 있다, 또한 20세에서 25세 사이의 인구의 유동인구 수가 가장 높게 나타났다.

2020년 평균 유동인구 수와 2019년 평균 유동인구 수의 차이

9,655,740명 = -28.51%의 감소율

부성 1동 연령대별 유동 인구 수



부성 1동 유동인구 1위는 20대

제 1부 천안시의 2년의 기록

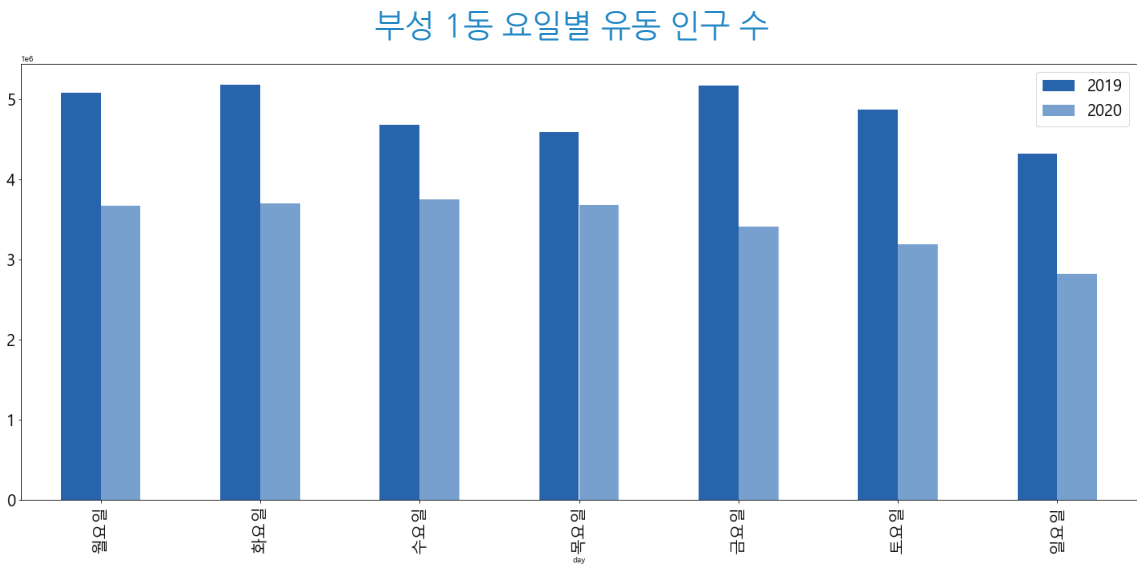
부성 1동의 경우, 한국 기술 교육 대학교 제2 캠퍼스가 위치해 있다. 그렇기에 20대의 유동인구가 가장 높은 비율을 차지했다. 또한 초, 중, 고등학교가 위치했다.

그렇기에 COVID-19 이후 온라인으로 변환된 한국 기술 교육 대학교 의 영향과 청소년들의 비율이 많아 유동인구 수가 급감했다.



부성 1동 요일별 유동인구 수

제 1부 천안시의 2년의 기록

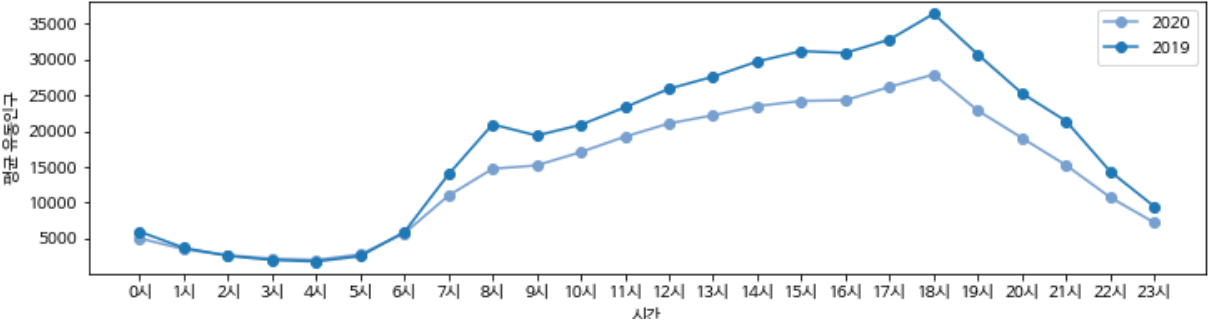


부성 1동 시간대별 유동인구 수

제 1부 천안시의 2년의 기록

부성 1동의 경우 앞서 확인 했듯이 학업 지역이다. 한국 기술 교육대학교 제 2캠퍼스가 위치해있고 뿐만 아니라 많은 학교들이 있다. 코로나 바이러스 이후에 온라인 수업으로 전환되어있기에 2020년의 유동인구가 2019년의 유동인구보다 전체적으로 감소하는 형태를 보인다.

부성 1동 시간대별 유동 인구 수

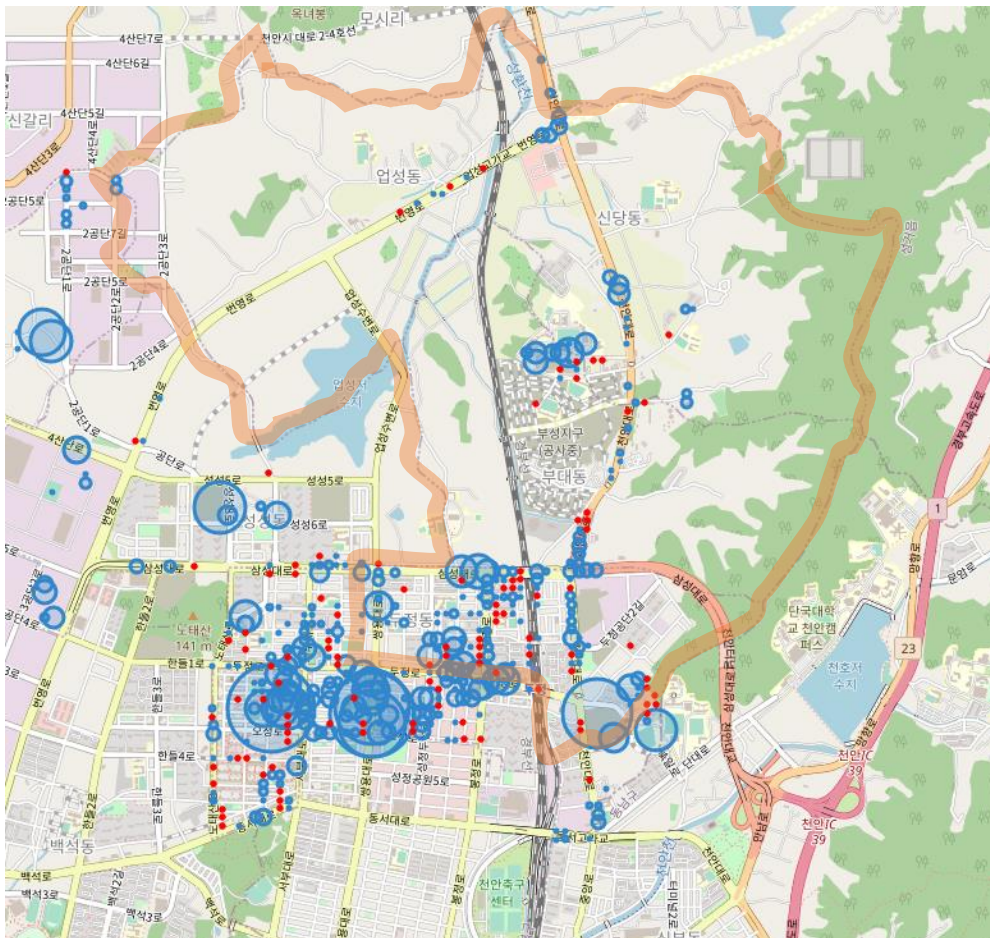


부성 1동 Circle map

제 1부 천안시의 2년의 기록

부성 1동의 지도 위에 위경도별로 (2019년 유동인구 수 - 2020년 유동인구)의 값을 이용해 양수 값이라면 빨간색 원을 그려 표시했다. 즉, 2020년 유동인구 수가 더 늘어났다는 것을 뜻한다. 반대로 음수 값이라면 파란색 원을 그려 표시했다. 즉, 2019년 유동인구 수가 더 늘어났다는 것을 뜻한다. 이때 각 원의 크기는 (2019년 유동인구 수 - 2020년 유동인구)의 값이 커지면 원의 크기가 커진다.

쌍용2동 Circle marker

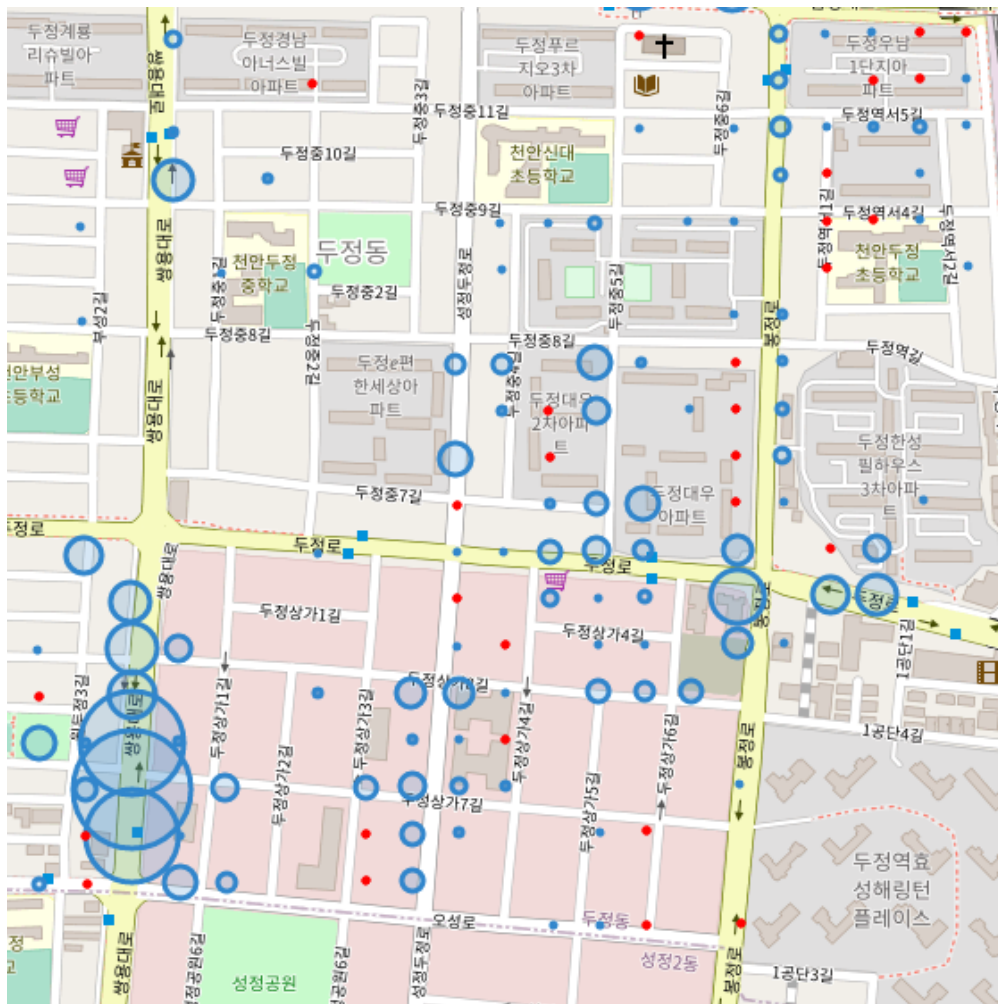


부성 1동 Circle map

제 1부 천안시의 2년의 기록

(2019년 유동인구 수 - 2020년 유동인구 수)가 양수 값을 가지는데 가장 큰 원을 가지는 부분을 확대했을 때, 도로에 양수값을 가지는 파란색 원이 존재했다. 또한 '두성상가로'에는 상가가 많이 즐비한 부분 또한 2019년의 유동인구가 더 많았다는 것을 알 수 있다. 빨간색 원을 가지는 부분은 아파트와 같은 주거지역이었다.

부성1동 Circle marker



부성 2동 유동인구 분석

제 1부 천안시의 2년의 기록

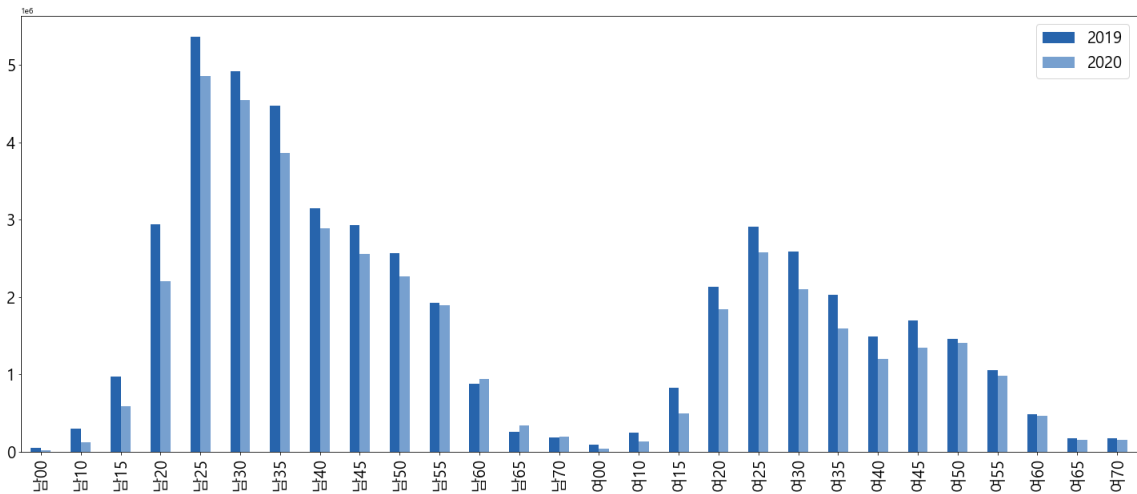
아래의 그래프는 부성 2동의 2019년 03월에서 04월까지의 유동인구와 2020년 03월에서 04월까지의 유동인구를 연령대, 성 별로 표시한 그래프이다.

전반적으로 2019년의 유동인구보다 2020년의 유동인구가 적음을 확인할 수 있다.
2019년의 유동인구 수의 합은 48,350,531 명으로 확인되었고 2020년의 유동인구 수의 합은 41,809,836 명이다.
또한 2019년과 2020년의 연령대별 유동인구는 남자가 여자보다 비율보다 높은 것을 확인할 수 있다, 또한 20세에서 25세 사이의 인구의 유동인구 수가 가장 높게 나타났다.

2020년 유동인구 수와 2019년 유동인구 수의 차이

6,540,695명 = -13.53%의 감소율

부성 2동 연령대별 유동 인구 수



부성 2동 유동인구 분석

제 1부 천안시의 2년의 기록

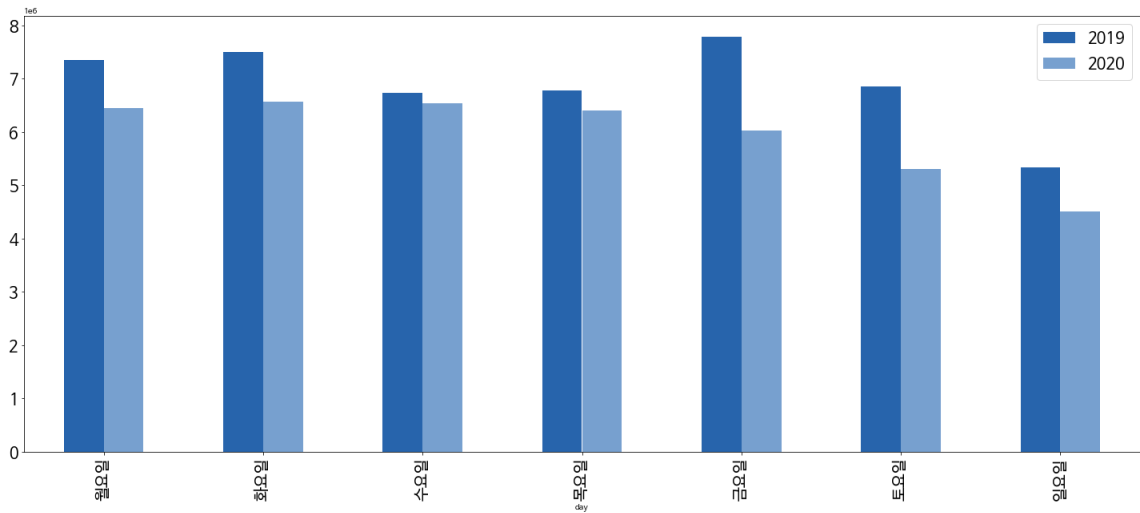
부성 2동의 경우, 제 2/3 일반산업 단지, 롯데 푸드 천안공장, 쿠첸 천안공장, 삼성 SDI 천안 사업장, 삼성 디스플레이 천안사업장 등이 위치해 있다. 그렇기에 20-30대의 남성 유동인구의 비율이 높았다는 것을 알 수 있다.



부성 2동 요일별 유동인구 수

제 1부 천안시의 2년의 기록

부성 2동 요일별 유동 인구 수

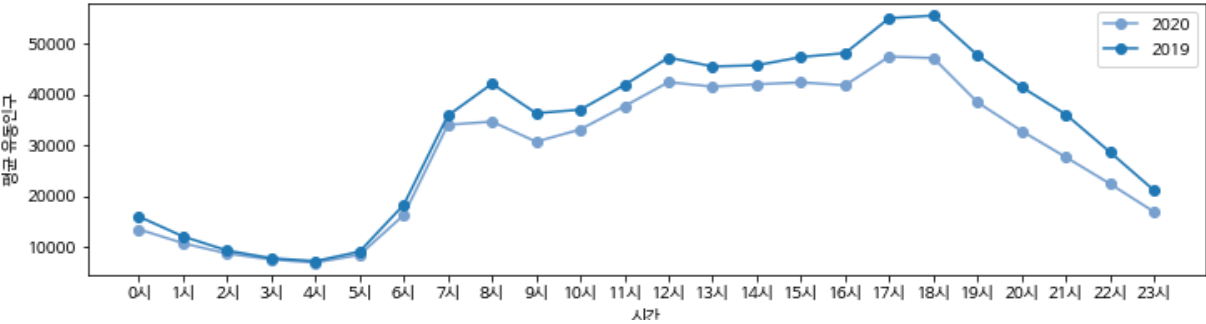


부성 2동 시간대별 유동인구 수

제 1부 천안시의 2년의 기록

부성 2동의 경우 앞서 확인 했듯이 공업단지이다. 쿠첸 천안공장, 삼성 SDI 천안 사업장 외에도 많은 공장이 있는 지역이다. 대부분의 회사도 재택 근무를 권장했으나 공장이기에 실제로 재택근무를 진행하는 직원의 수가 적었다. 그렇기에 시간대별 유동인구 형태는 크게 다르지 않았다. 또한 2020년이 2019년에 비해 유동인구 수가 감소하긴 했으나 크게 감소하지는 않았다.

부성 2동 시간대별 유동 인구 수

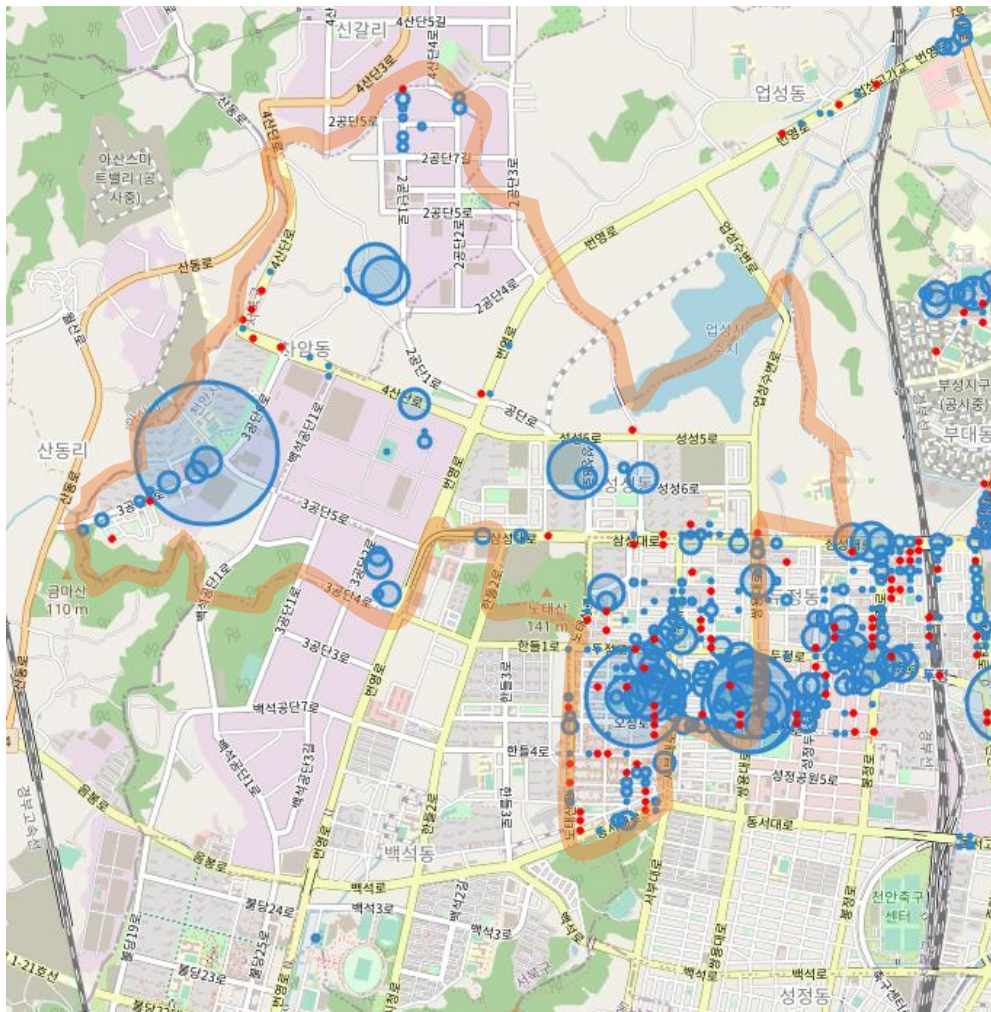


부성 2동 Circle map

제 1부 천안시의 2년의 기록

부성 2동의 지도 위에 위경도별로 (2019년 유동인구 수 - 2020년 유동인구)의 값을 이용해 양수 값이라면 빨간색 원을 그려 표시했다. 즉, 2020년 유동인구 수가 더 늘어났다는 것을 뜻한다. 반대로 음수 값이라면 파란색 원을 그려 표시했다. 즉, 2019년 유동인구 수가 더 늘어났다는 것을 뜻한다. 이때 각 원의 크기는 (2019년 유동인구 수 - 2020년 유동인구)의 값이 커지면 원의 크기가 커진다.

쌍용2동 Circle marker



부성 2동 Circle map

제 1부 천안시의 2년의 기록

(2019년 유동인구 수 - 2020년 유동인구 수)가 양수 값을 가지는데 가장 큰 원을 가지는 부분을 확대했을 때, 도로에 양수값을 가지는 파란색 원이 존재했다. 또한 코스트코 천안점에서 2019년의 유동인구가 2020년 유동인구에 비해 큰 값을 가지는 것을 알 수 있었다. 빨간색 원을 가지는 부분은 주거지역에 빨간색 원이 존재했다.

부성1동 Circle marker



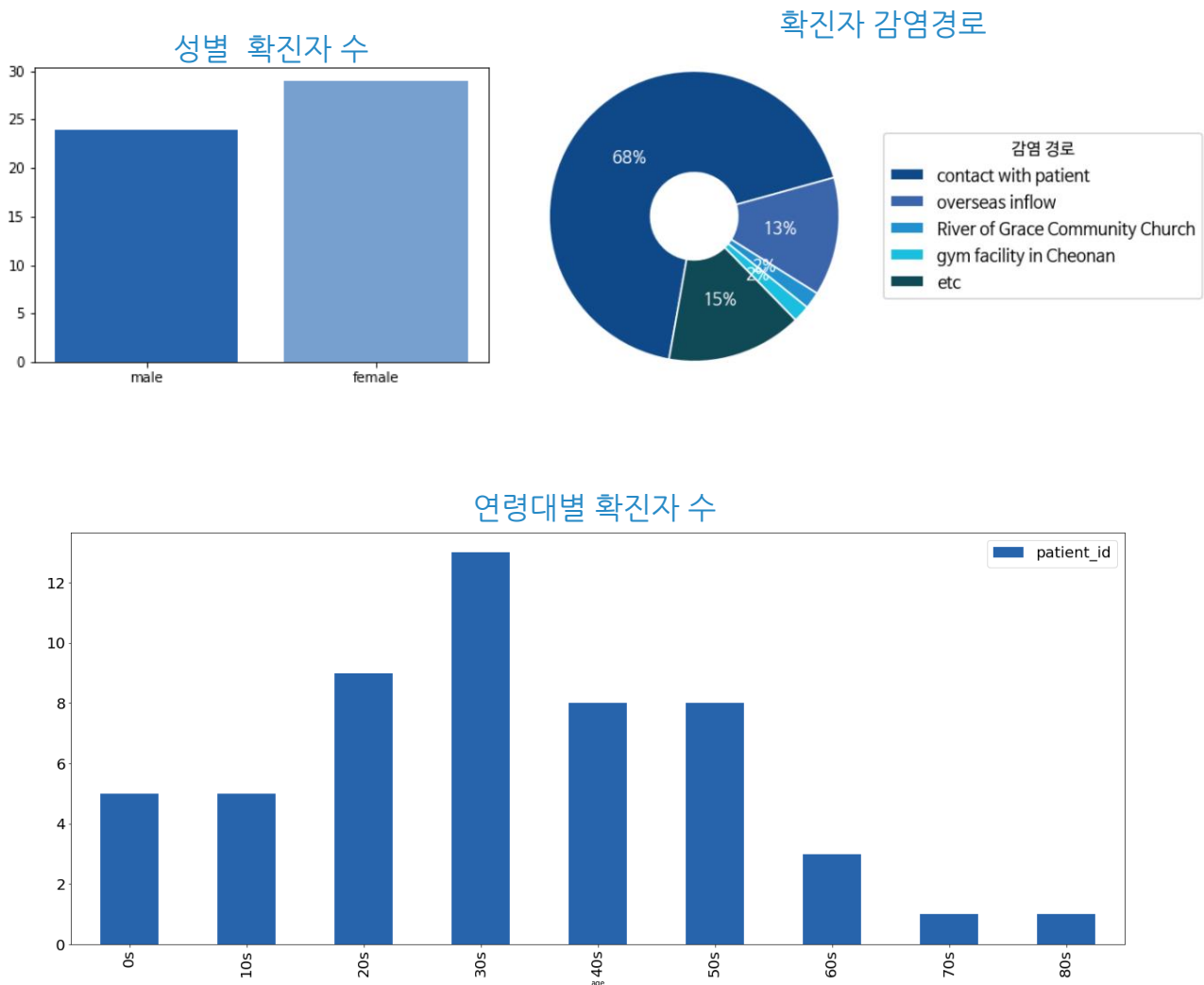
코로나 거리두기

제 2부 거리두기의 효과

코로나 확진자 기본적인 정보

제 2부 거리두기의 효과

아래의 첫 번째 그래프는 2020년 03월 01일 부터 2020년 04월 01일까지의 코로나 확진자 수에 대한 정보이다. 첫 번째 그래프는 성별에 따른 확진자 수이다. 남성은 24명 여성 29명으로 여성이 더 높은 비율을 차지했다.
두 번째 원형 그래프는 확진자의 감염경로의 비율을 시각화한 자료이다. 약 68%가 확진자와의 접촉으로 1위를 차지했고 2위는 해외유입이다.
마지막 그래프는 연령대별 확진자 수를 그래프로 나타낸 것이다. 30대가 가장 높은 비율이다.

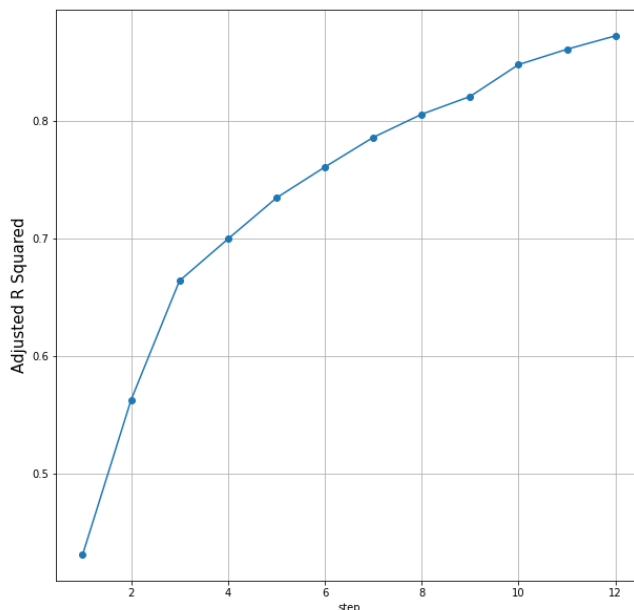


코로나 확진자 회귀 모델 feature

제 2부 거리두기의 효과

확진자와 유동인구 사이의 관계를 밝혀내기 위해서 사용하는 데이터는 weather dataset과 확진자 정보와 유동인구의 데이터 셋이다. 이때, 유동인구 데이터 셋이 확진자 수에 영향을 미치는 요인이 당일의 유동인구가 당일의 확진자 수에 영향을 미치지 않는다고 가정했다. 그렇기에 날씨와 유동인구 사이의 간격을 두었다. 1일 전 유동인구 수, 2일 전 유동인구 수처럼 구성했다. 잠복 기간이 2주인 것을 확인해 간격을 1일부터 14일까지 새로운 변수를 선언했다. 그 결과, 유동인구수 14개, 최고 기온 14개, 최저 기온 14개, 평균 기온 14개, 습도 14개, 풍량 14개로 84개의 설명변수와 1개의 반응 변수로 구성했다. 하지만 14일 전 기간의 데이터는 존재하지 않아 처음부터 14개의 변수를 제거하고 47개의 데이터셋으로 분석을 진행했다.

‘전진 선택법’을 사용해 변수 12개를 선택했다. 변수는 'avg_temp6', 'max_temp3', 'avg_temp8', 'avg_relative_humidity2', 'avg_relative_humidity4', 'avg_relative_humidity1', 'max_temp5', 'avg_relative_humidity12', 'population_log7', 'population_log14', 'max_wind_speed14', 'avg_relative_humidity8'의 변수가 선택되었다. 아래의 그래프는 전진선택법을 이용해 그래프를 그린 결과이다.



전진선택법을 활용했을 때의 R-squared 그래프

코로나 확진자 회귀 모델 결과

제 2부 거리두기의 효과

앞의 단계에서 선택된 설명변수를 이용해 회귀분석을 진행했을 때, 특정 변수들은 p-value 값이 0.05 값이 넘는 것을 확인해 직접 변수들을 제거함으로 최종 설명변수를 구성했다. 그 결과 최종 회귀분석 결과는 아래의 그림과 같다.

확진자 수에 영향을 주는 변수가 크게 기온, 습도, 이 이전 유동인구 수가 영향을 미친다는 것을 도출했다. 특히 유동인구의 경우 7일전 유동인구 수, 14일 전 유동인구 수가 영향을 주지만 7일 전의 경우 회귀 계수가 0.0004, 14일 전의 유동인구 수의 경우 -0.0003으로 나왔다. 실질적으로 유동인구 수가 확진자 수에 큰 영향을 주지 않았다는 결론을 도출했다. 하지만 해당 결과는 천안시로 데이터가 한정되었다는 점과 기간이 3월에서 4월로 코로나 확진자가 천안시에 급격히 증가하는 추세는 아니었다는 점에서 데이터 면에서 한계가 존재했다. 3월에서 4월까지의 데이터가 아닌 최근까지의 데이터셋이 있었다면 더욱 유의미한 결과를 도출할 수 있을 것이라고 기대한다.

OLS Regression Results						
Dep. Variable:	patient_count	R-squared (uncentered):	0.921			
Model:	OLS	Adj. R-squared (uncentered):	0.899			
Method:	Least Squares	F-statistic:	42.96			
Date:	Tue, 15 Dec 2020	Prob (F-statistic):	2.55e-17			
Time:	21:05:45	Log-Likelihood:	-25.106			
No. Observations:	47	AIC:	70.21			
Df Residuals:	37	BIC:	88.71			
Df Model:	10					
Covariance Type: nonrobust						
	coef	std err	t	P> t	[0.025	0.975]
avg_temp6	-0.2239	0.035	-6.385	0.000	-0.295	-0.153
max_temp3	-0.1137	0.026	-4.297	0.000	-0.167	-0.060
avg_relative_humidity2	0.0295	0.006	4.821	0.000	0.017	0.042
avg_relative_humidity4	0.0180	0.007	2.542	0.015	0.004	0.032
avg_relative_humidity1	0.0156	0.007	2.370	0.023	0.002	0.029
max_temp5	-0.0732	0.026	-2.791	0.008	-0.126	-0.020
avg_relative_humidity12	0.0256	0.007	3.418	0.002	0.010	0.041
population_log7	4.204e-06	1.04e-06	4.043	0.000	2.1e-06	6.31e-06
population_log14	-3.867e-06	1e-06	-3.852	0.000	-5.9e-06	-1.83e-06
max_wind_speed14	-0.1344	0.049	-2.757	0.009	-0.233	-0.036
Omnibus:	3.149	Durbin-Watson:	2.073			
Prob(Omnibus):	0.207	Jarque-Bera (JB):	2.559			
Skew:	0.571	Prob(JB):	0.278			
Kurtosis:	3.031	Cond. No.	2.09e+06			

회귀 분석 결과