

대한민국 송도 만세

미디어의 영향을 통해 알아보는 송도의 변화



0. 목차

1. 연구배경 및 목적

1.1 연구배경

1.2 프로젝트 목적

2. 데이터 소개

2.1 데이터 수집

2.2 변수설정

3. 데이터 분석

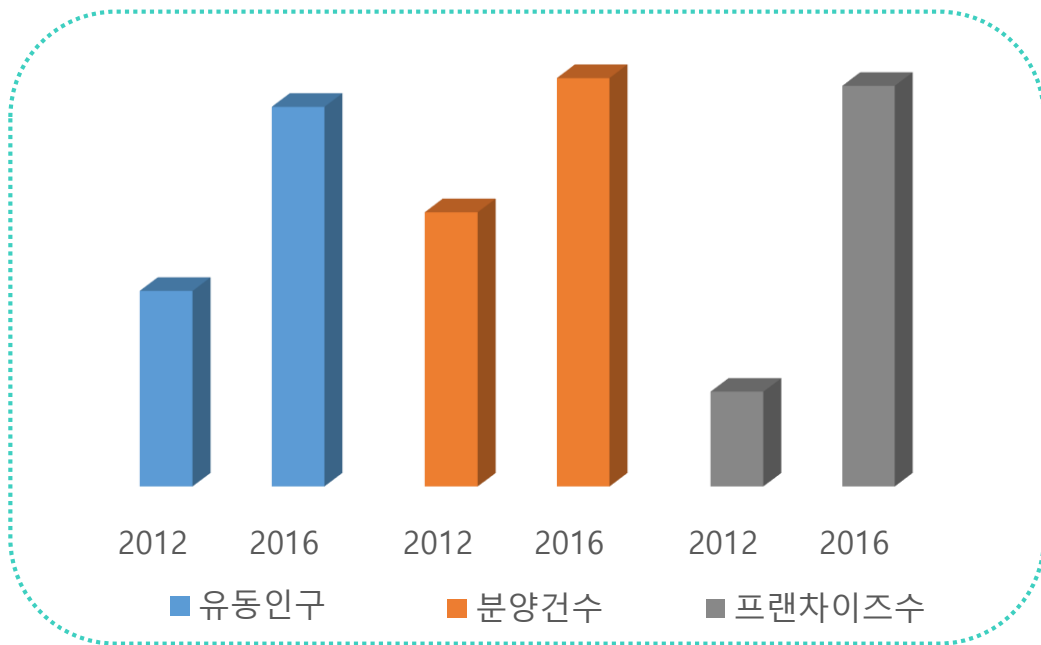
3.1 데이터 분석 개요

3.2 회귀분석

3.3 회귀분석 모델 해석

1. 연구배경 및 목적

1.1 연구배경



송도의 변화과정 (유동인구 수)

'대한·민국·만세' 덕 본 송도... 유령도시에서 핫플레
이스로
송도국제도시 아파트, '삼둥이 효과'에 쑥쑥 크네
송도주민 "송일국 효과가 정부 부동산 대책
보다 낫다"

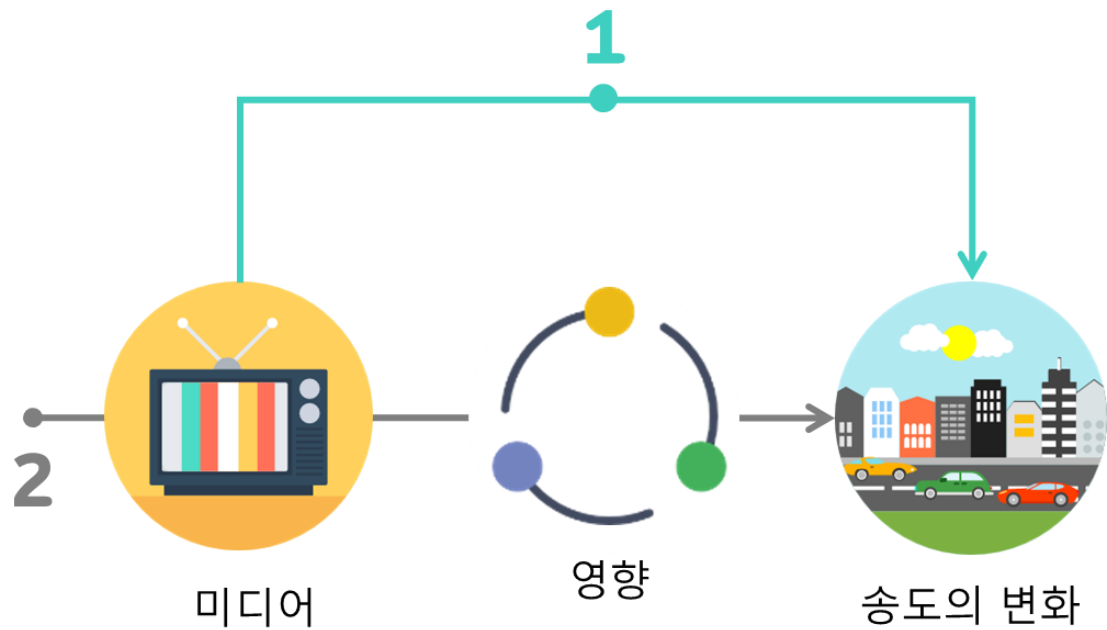
김창 기자 ▾

송도-삼둥이 관련 기사

과연 송도의 성장에는 미디어의 영향이 있었을까?
(슈퍼맨이 돌아왔다)

1. 연구배경 및 목적

1.2 프로젝트 목표



1 미디어가 송도에 영향을 주었는지 확인하고

2 그 영향으로 사람들의 인식변화가 송도에 어떤 변화를 불러왔는지 확인해본다

2. 데이터 소개

2.1 데이터 수집

부동산 관련 데이터
유동 인구 데이터
지역 인허가 기관 데이터
인구통계 데이터 등

오픈 데이터

송도라는 도시의 변화 과정을
공신력 있는 데이터를 기반으로
분석하기 위해

네이버 블로그 데이터
(포스팅 제목, 일자, 내용 등)
네이버 검색 데이터
(블로그 및 신문 기사 검색 량)

크롤링 데이터

송도에 대한 사람들의 인식 변화와 송
도에 관련된 키워드를 알아보기 위해

삼둥이편
영상 분석 데이터
(분량, 노출장소 등)

영상 데이터

미디어에 노출된 송도와 삼둥이를
정량적으로 알아보기 위해

기간 : 2012.01.01 ~ 2017.12.31 ‘슈퍼맨이 돌아왔다’ 삼둥이 편 방영 일자를 기준으로 (2014.07.05 ~ 2016.02.06)
방영 전, 중, 후 상황을 비교하기 위해 방영일자 만큼의 전, 후 기간을 살펴보려 한다

2. 데이터 소개

2.2 변수 선정

구분	Variable	Variable Label
1	회차	슈퍼맨이 돌아왔다_삼둥이편 진행 회차
2	삼둥_전체	슈돌 전체 방송 분량 중 삼둥이편이 차지하는 비율
3	송도노출	삼둥이편 방송 중 송도가 나오는 시간 (단위 : 분)
4	Blog_Songdo	네이버 블로그에서 키워드 '송도' 검색량
5	Article_Triplet	네이버 기사에서 키워드 '삼둥이' 검색량
6	floating	송도 유동인구수 (단위 : 명)
7	사업체수	송도에서 운영중인 업소 갯수
8	송도인구	송도 인구수 (단위 : 명)
9	미분양	송도 아파트 미분양 호수

3. 데이터 분석

3.1 데이터 분석 개요

분석 목적	사용 데이터		분석 기법	Tool
Q1. 미디어가 송도에 어떤 영향을 미쳤을까?	부동산 통계정보(2012 - 2017) 출처: 한국 감정원		Linear Regression	R
	슈퍼맨이 돌아왔다 삼둥이편(34화 - 116화)			
	송도 유동인구 데이터(2012-2016) 출처: 인천교통공사			
	송도 인구통계 데이터 출처: 통계청			
	지역 인허가 기관 정보(2012 - 2017) 출처: Business Open Data			
	'송도' 블로그, 기사 검색 데이터 (2012 - 2017) 출처: 네이버			
Q2. 그 영향으로 송도는 어떻게 변했을까?	'송도' 블로그 검색 데이터 (2012 - 2017) 출처: 네이버	Crawling – Python 띄어쓰기 – R library (kospacing) 명사추출 – R library (KoNLP)	Word Cloud	D3
			Bump Chart	R_library(ggplot2)/plotly
	지역 인허가 기관 정보(2012 - 2017) 출처: Business Open Data	위경도 추출 – Python package(geocoder)	Mapping	D3

3. 데이터 분석

3.2 회귀분석

lm(Blog_Songdo ~ 회차 + 삼등_전체+송도노출, data = dat_su)

1단계	Estimate	Std. Error	t value	Pr (> t)
Intercept	1071.126	238.135	4.498	2.33e-05 ***
회차	8.001	2.178	3.674	0.000434 ***
삼등_전체	42.669	4.483	9.518	9.51e-15 ***
송도노출	14.787	7.995	1.850	0.068122 .

Residual standard error: 444.5 on 79 degrees of freedom

Multiple R-squared: 0.5446

F-statistic: 31.49 on 3 and 79 DF, p-value: 1.715e-13

3. 데이터 분석

3.2 회귀분석

lm(사업체수 ~ Blog_Songdo + Article_Triplet + log(floating), data = dat_exc)

2단계	Estimate	Std. Error	t value	Pr (> t)
(Intercept)	-61.5	14.7	-4.18	8.81e-05 ***
Blog_Songdo	0.15	0.02	8.55	3.03e-12 ***
Article_Triplet	0.14	0.08	1.75	0.0843 .
log(floating)	21.85	4.15	5.27	1.65e-06 ***

Residual standard error: 5.848 on 65 degrees of freedom

Multiple R-squared: 0.89, Adjusted R-squared: 0.89

F-statistic: 182.1 on 3 and 65 DF, p-value: < 2.2e-16

3. 데이터 분석

3.2 회귀분석

`lm(formula = 지가지수 ~ 사업체수 + log(송도인구) + log(미분양), data = dat_exc)`

3단계	Estimate	Std. Error	t value	Pr (> t)
(Intercept)	39.82	16.23	2.45	0.016854 *
사업체수	0.08	0.02	3.64	0.011037 *
log(송도인구)	4.82	1.51	3.19	0.002199 **
log(미분양)	-0.18	0.035	-5.1	3.17e-06 ***

Residual standard error: 0.4629 on 65 degrees of freedom

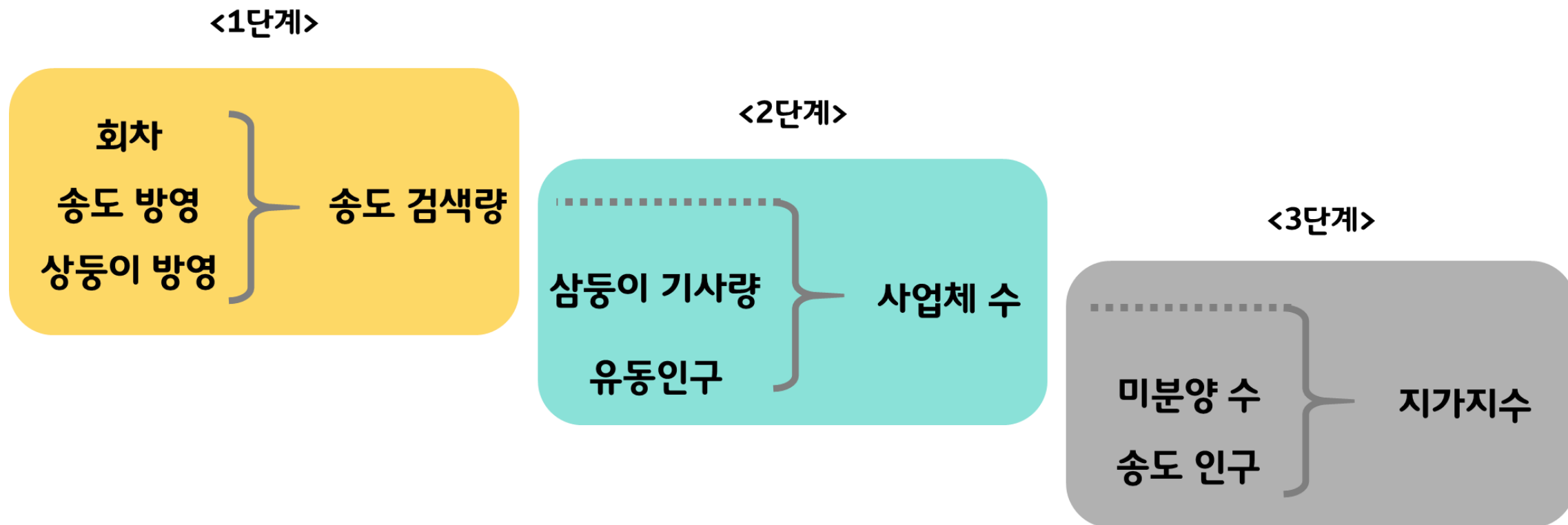
Multiple R-squared: 0.9751, Adjusted R-squared: 0.9739

F-statistic: 847.3 on 3 and 65 DF, p-value: < 2.2e-16

3. 데이터 분석

3.3 회귀분석 모델 해석

Analasys Diagram



3. 데이터 분석

3.3 회귀분석 모델 해석

<1단계>	<2단계>	<3단계>
삼둥이편 회차가 진행될 수록	송도 게시물 수가 증가할 수록	송도의 사업체 개수가 증가할 수록
삼둥이 편에서 송도가 노출 될수록	삼둥이 기사량이 증가할 수록	미분양 호수가 감소할 수록
‘슈돌’ 내에서 삼둥이 편 비율이 증가할 수록	송도의 유동인구가 증가할 수록	송도의 인구가 증가할 수록
송도 관련 게시물수가 증가했다	송도의 사업체 개수가 증가했다	송도의 지가지수가 상승했다
‘슈퍼맨이 돌아왔다’로 인해 송도에 대한 관심 증가	송도 변화 확인	송도의 부동산 가치 상승

“1,2,3단계를 거쳐 슈퍼맨이 돌아왔다의 방영이 송도의 지가지수와 **상관이 있다**는 것을 확인하였다.”



songdotriplet.github.io



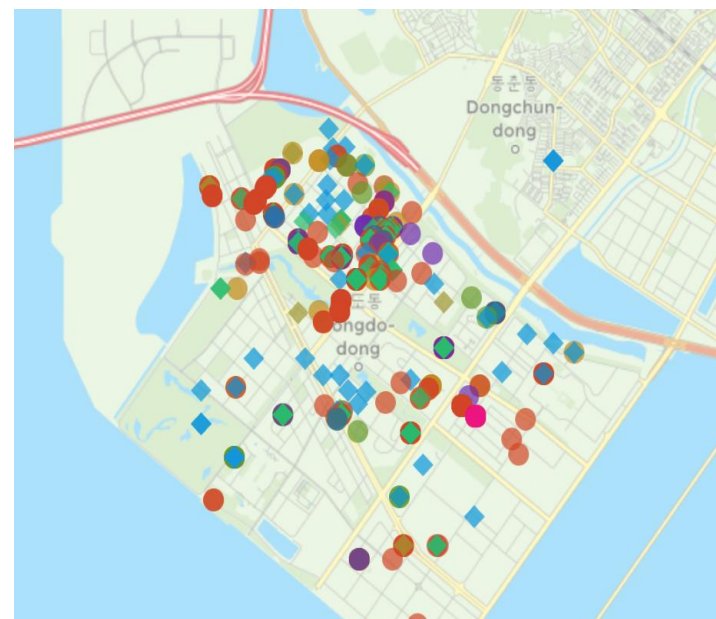
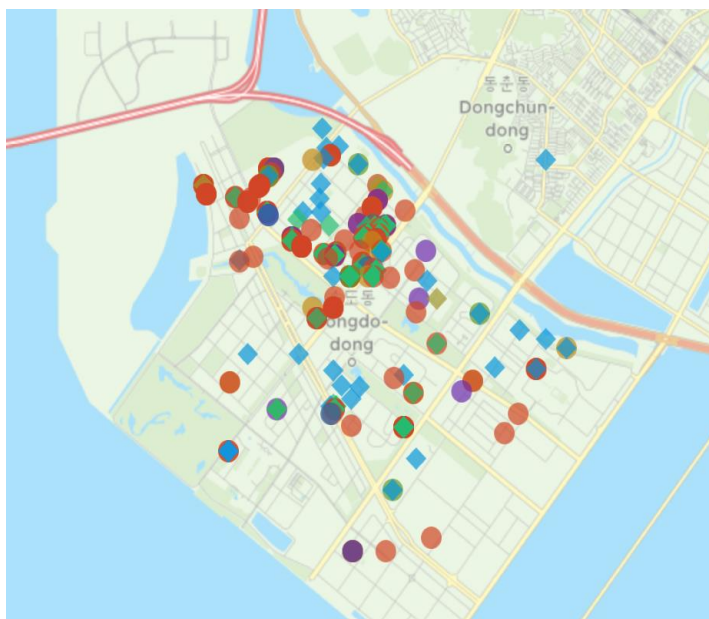
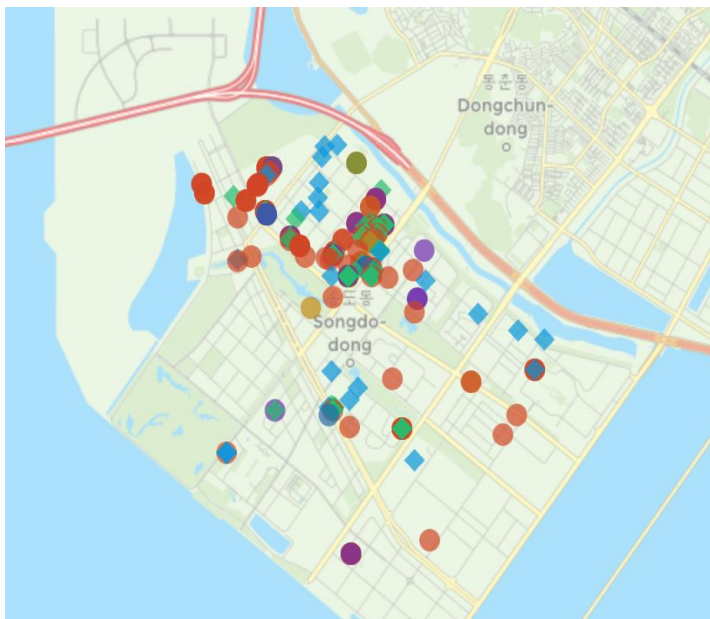
3. 데이터 분석

3.4 텍스트 마이닝 & 시각화

Mapping

[카테고리_주거 : 병원, 학교, 피트니스센터]
:해당 사업이 실제 주거인구가 많아질 수록 증가하므로 선정

[카테고리_핫플레이스 : 편의점, 제과점, 영화관, 노래방, 카페, 패스트푸드]
:해당 사업이 유동인구가 많아질 수록 증가하기 때문에 선정



시각화 툴 | D3.js **의의** | 사업체 데이터의 위치와 분포를 가시적으로 확인할 수 있다

분석 내용 | 삼등이편 방영을 기점으로 기간을 전, 중, 후로 나누어 각 카테고리 별 사업체 수의 변화를 분석했다

결과 해석 | 시간이 지날 수록 모든 카테고리의 사업체 수가 증가하는 것을 통해 주거와 여가에 적합한 공간으로 변화한 것을 확인할 수 있다

3. 데이터 분석

3.4 텍스트 마이닝 & 시각화

Word Cloud



시각화 툴 | D3.js **의의** | 단어의 빈도를 시각적으로 표현,
빈도수를 기준으로 크기가 정해지기 때문에 자주 언급된 키워드를 확인 가능하다

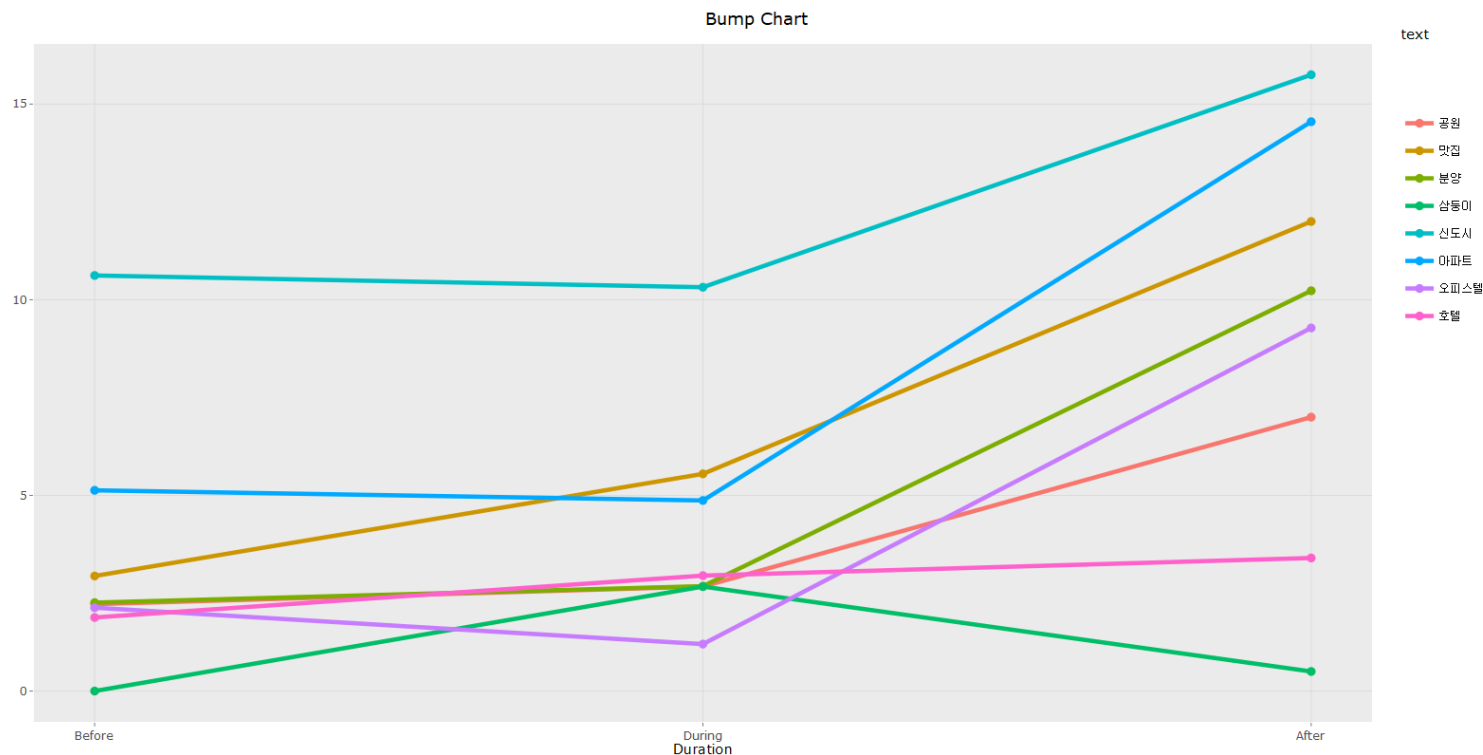
분석 내용 | 삼동이편 방영 전, 후를 비교.
방영 전에는 '신도시/아파트/위치' 등 투자대상으로서의 키워드를 볼 수 있고, 부동산 관련 단어가 눈에 띈다
방영 후에는 '맛집/아울렛' 등 여가 관련 단어가 추가되었다

결과 해석 | '아파트 / 분양, 빌라/ 월세'의 크기가 증가 한 것은 실 거주 목적으로써의 검색 키워드나 늘어난 것이고
이는 송도에 대한 실질적인 수요가 증대된 것이라 해석할 수 있다
새로이 등장한 키워드('맛집 /아울렛 등')로 인해 전과 비교해 여가 공간으로 인식 됨을 확인했다

3. 데이터 분석

3.4 텍스트 마이닝 & 시각화

Bump Chart



시각화 툴 | R library(ggplot, Plotly)

의의 | 범프차트는 서로 다른 두 점을 직선으로 연결함으로써 '값이 어떻게 변했는가'를 표현 할 수 있다. 기간 별 변화를 한 차트 안에서 종합적으로 확인할 수 있다

분석 내용 | 가장 상위에 신도시와 아파트가 보이고 분양이나 오피스텔의 경우는 방영 기간을 기점으로 보다 드라마틱하게 상승
맛집이나 호텔 (놀이거리 관련 키워드) 역시 꾸준히 상승

결과 해석 | '삼동미'를 제외한 그래프가 상승하는 것으로 보아 시간이 지남에 따라 사람들에게 송도가 점점 더 관심을 받고 있다는 것을 확인했다

3. 데이터 분석

3.5 결론 도출

회귀분석

Linear Regression

텍스트 분석과 시각화

Text Mining

- 삼둥이 방영이 송도의 인지에 긍정적 영향을 미침
- 송도 인지율 상승은 송도지역 사업체 수 상승과 연관
- 사업체 수와 인구 증가는 지가지수 상승에 영향

삼둥이의
광고 효과
확인

- 삼둥이 방영 후 투자 중심에서 주거 중심의 도시로
- 유명도시에서 핫플레이스로 인식이 변화

송도의
변화
확인



Thank You!

FOR YOUR TIME