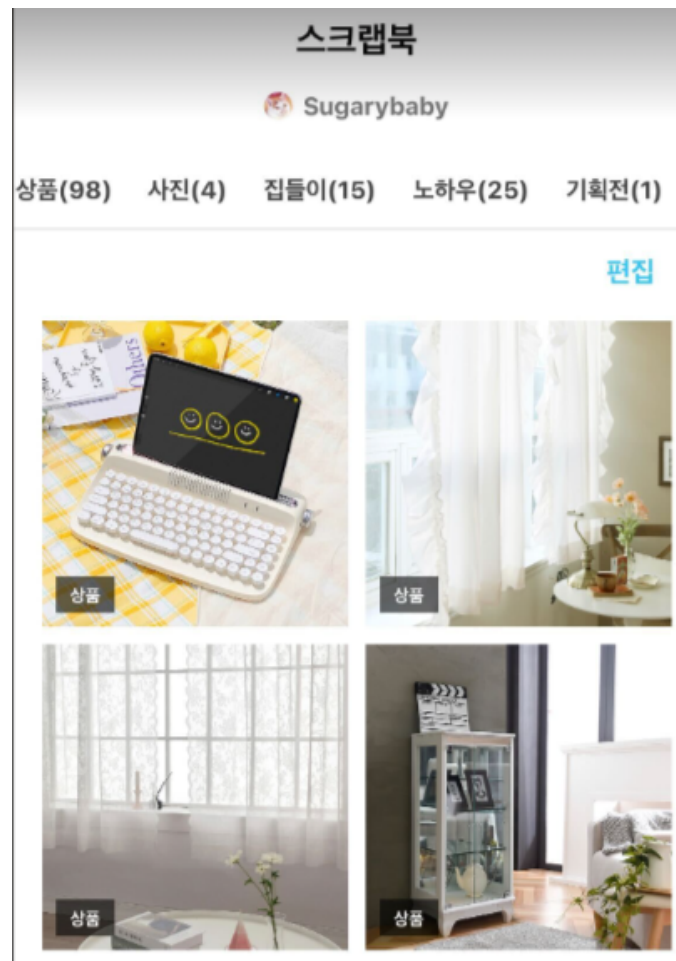
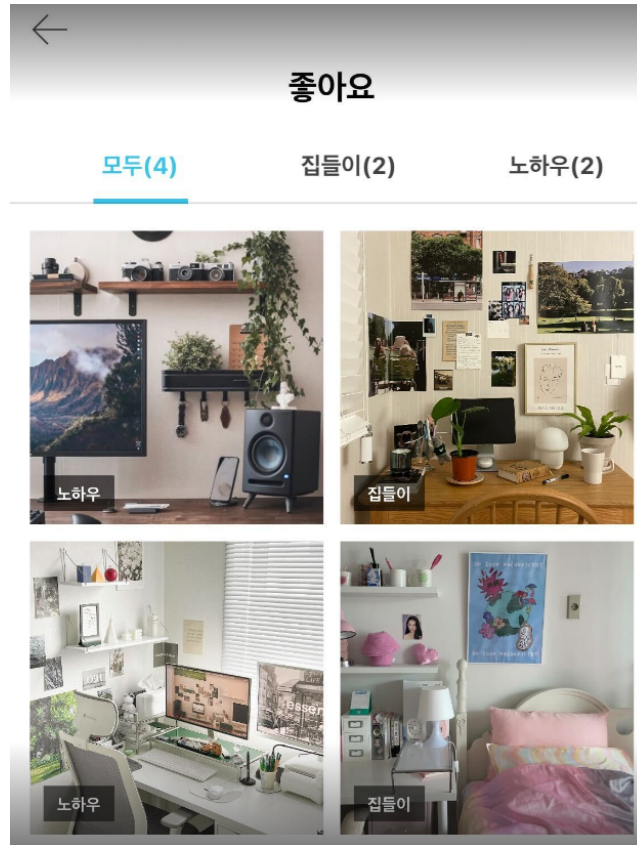


# 집들이컨텐츠 full 데이터 (housewarming2.csv) EDA

오늘의집 어플에서 좋아요와 스크랩의 의미 + 가중치 산정





- 집들이 콘텐츠 : 좋아요와 스크랩 모두 나중에 모아보기 가능
- 상품 : 좋아요 기능이 없고 스크랩만 가능
- 좋아요와 스크랩 모두 다시보기 & 섹션 분류 기능이 있지만,
- 일반적으로 좋아요보다는 스크랩이 조금 더 '나중에 참고할 용도'에 가까울 것이라 추측
- 또한 상품까지 섹션 분류가 가능한 스크랩 기능이 좋아요보다 더 상위 가치를 가진 기능이라고 추측해, '인기도' 지표 산정 시 스크랩에 1, 좋아요에 0.5의 가중치를 두었음
- 결과 해석시에도 스크랩이 높은 경우 나중에 참고할 목적, 좋아요는 포스팅에 대한 호감 표시 목적이 더 크다고 보았음

## 1. 가족형태별 인기도 분석 / 오늘의집에서는 어떤 가족형태 관련 콘텐츠를 더 만들어야 할까?

행 레이블	개수 : 가족형태	합계 : 조회대비인기도
싱글라이프, 부모님과 함께 사는 집	4	19.0
아기가 있는 집, 취학 자녀가 있는 집, 부모님과 함께 사는 집	1	14.7
<b>싱글라이프</b>	971	<b>13.5</b>
신혼부부, 아기가 있는 집	44	10.8
아기가 있는 집, 취학 자녀가 있는 집	33	10.8
<b>취학 자녀가 있는 집</b>	564	<b>10.7</b>
<b>부모님과 함께 사는 집</b>	341	<b>10.7</b>
<b>신혼부부</b>	1955	<b>10.6</b>
아기가 있는 집, 부모님과 함께 사는 집	2	9.4
아기가 있는 집	727	9.0
기타	250	7.5
취학 자녀가 있는 집, 부모님과 함께 사는 집	5	5.3
싱글라이프, 신혼부부	1	3.1
신혼부부, 부모님과 함께 사는 집	2	2.9
<b>총합계</b>	<b>4900</b>	<b>11.0</b>

- 스크랩에 1, 좋아요에 0.5의 가중치를 두고 스크랩\*좋아요 로 '인기도' 지표를 산출함
- 각 가족형태별 조회 대비 인기를 봤을때 '싱글라이프' 가 눈에 띄게 높았음
- 그 외에 취학 자녀가 있는 집, 부모님과 함께 사는 집, 신혼부부 등은 비슷한 수치를 보임
- 결론 : **신혼부부의 콘텐츠가 가장 많지만, 인기도가 더 높은것은 싱글라이프**이므로 **오늘의집은 싱글라이프 관련 콘텐츠를 지금보다 더 많이 생산하면 좋을듯**

## 2. 어떤 작업이 가장 인기있을까? / 오늘의집은 어떤 작업 관련 콘텐츠를 더 만들어야 할까?

행 레이블	개수 : 작업	합계 : 조회대비인기도	합계 : 조회대비스�크랩	합계 : 조회대비좋아요
반셀프	863	10.7	2.6%	30.0%
셀프•DIY	<b>2354</b>	10.4	<b>2.9%</b>	25.9%
전문가	1683	<b>11.9</b>	2.7%	<b>30.6%</b>
<b>총합계</b>	<b>4900</b>	<b>11.0</b>	<b>2.8%</b>	<b>28.2%</b>

- 콘텐츠 수 자체는 셀프&DIY가 가장 많았음 → '싱글라이프' 가 가장 인기있는 것과 관련이 있지 않을까? (싱글들은 돈이 없는경우가 많으니까)
- 조회 대비 인기도와 좋아요는 전문가 > 반셀프 > 셀프 순이었음 → 전문가의 손길이 닿은 (예쁜) 포스팅이 좋아요를 많이 받는 것을 알 수 있다
- 조회 대비 스크랩은 셀프가 가장 많았음. → 셀프 인테리어를 올리는 사람의 비중이 많으므로, 참고하기 위해 스크랩을 많이 한 게 아닐까?
- 결론 : 작업 포스팅 수는 셀프/DIY가 가장 많지만 **인기도와 스크랩 차이를 봤을 때 전문가와 반셀프 작업 또한 충분한 수요가 있다고 판단됨.** 셀프/DIY를 메인으로 하되, **전문가/반셀프 포스팅 콘텐츠도 지금보다 더 적극적으로 생산하면 좋을듯**

### 3. 가족형태별로는 어떤 작업이 인기있을까? / 오늘의집에서는 가족형태별로 어떤 작업 콘텐츠를 생산하고 추천하면 좋을까?

행 레이블	개수 : 작업	합계 : 조회대비인기도	합계 : 조회대비스크랩	합계 : 조회대비좋아요
☐싱글라이프	971	13.5	3.1%	25.5%
반셀프	61	16.5	3.2%	30.1%
셀프•DIY	819	13.4	3.1%	24.9%
전문가	91	12.7	2.9%	28.9%
☐취학 자녀가 있는 집	564	10.7	2.4%	32.0%
반셀프	105	6.8	2.3%	31.1%
셀프•DIY	152	8.6	2.5%	29.6%
전문가	307	13.0	2.4%	33.4%
☐부모님과 함께 사는 집	341	10.7	3.3%	24.9%
반셀프	37	12.2	2.7%	30.5%
셀프•DIY	270	10.6	3.5%	23.4%
전문가	34	8.8	2.6%	32.8%
☐신혼부부	1955	10.6	2.7%	28.3%
반셀프	428	10.8	2.7%	28.9%
셀프•DIY	770	7.5	2.5%	26.7%
전문가	757	13.0	2.8%	29.2%
☐아기가 있는 집	727	9.0	2.5%	31.1%
반셀프	185	9.6	2.5%	31.2%
셀프•DIY	198	6.8	2.3%	30.1%
전문가	344	9.7	2.6%	31.5%
☐기타	250	7.5	2.7%	31.4%
반셀프	26	14.8	3.0%	33.0%
셀프•DIY	105	7.1	2.9%	27.6%
전문가	119	5.3	2.4%	35.2%
총합계	4808	11.0	2.8%	28.1%

가족형태	포스팅 수	조회대비 인기도	조회대비스크랩	조회대비좋아요	생산/추천하면 좋은 콘텐츠
싱글라이프	셀프/DIY	반셀프	반셀프	반셀프	반셀프
취학자녀가 있는 집	전문가	전문가	셀프/DIY	전문가	전문가
부모님과 함께	셀프/DIY	반셀프	셀프/DIY	전문가	방 ONLY : 셀프 집 전체 : 반셀프
신혼부부	셀프/DIY	전문가	전문가	전문가	전문가
아기가 있는 집	전문가	전문가	전문가	전문가	전문가

- 포스팅 수는 셀프/DIY가 높았지만, 조회대비 인기도는 전문가가 가장 높음
- 특히 결혼 후~자녀 라이프사이클부터는 전문가 포스팅이 조회 대비 인기도가 높음 → 집을 마련하고 본격적인 인테리어를 하는 수요 때문이 아닐까? 혹은 전문가 포스팅의 퀄리티가 높아서?
- 조회대비 좋아요에서 싱글만 '반셀프' 의 비중이 높았음 → 이유가 뭘까?
- 신혼부부는 포스팅 수에 셀프/DIY가 많은것에 비해, 스크랩과 좋아요는 모두 전문가를 선호

- 아기가 있는 집은 포스팅 수, 스크랩, 좋아요 모두 전문가를 선호함 → 아기가 태어나 큰 집으로 옮기면서 전체를 전문가에게 맡기는 경향이 있지 않을까? 혹은 아기가 있어 셀프 인테리어를 진행하기 어려울수도

#### 4. 평수와 분야별 평균예산은? / 평수와 분야별로, 사람들의 관심을 얻을 수 있는 '숫자' 를 발굴해보자

행 레이블	개수 : 평수구간	평균 : 예산
5	212	173
리모델링	4	200
부분공사	4	250
홈스타일링	204	171
15	284	935
건축	2	8,000
리모델링	48	3,235
부분공사	17	965
홈스타일링	217	359
25	607	2,665
건축	4	14,875
리모델링	356	3,644
부분공사	51	1,406
홈스타일링	196	965
35	697	3,443
건축	4	21,750
리모델링	412	4,760
부분공사	114	1,562
홈스타일링	167	1,040

45	139	6,592
건축	8	30,663
리모델링	96	6,236
부분공사	11	3,845
홈스타일링	24	1,250
55	58	7,276
건축	3	27,867
리모델링	39	8,156
부분공사	9	2,011
홈스타일링	7	314
65	20	13,705
건축	4	50,925
리모델링	8	6,513
부분공사	1	2,000
홈스타일링	7	2,329
75	23	20,304
건축	7	36,571
리모델링	9	14,111
부분공사	2	2,050
홈스타일링	5	15,980
총합계	2040	3,137

- 예산에 영향을 미치는 요소로는 공간(아파트/오피스텔 등), 공사분야(리모델링/홈스타일링 등), 가족형태(싱글/신혼부부 등), 평수, 작업(전문가, 셀프 등) 이 있음
- 예산 0, 1 로 된 값은 미리 제거함
- 모든 요소를 다 반영할경우 각각의 표본이 너무 작아지는 문제가 있음
- **평수가 공간과 가족형태를 대체할 수 있다고 추측하고, 공사분야가 작업을 대체할 수 있다고 추측해서 평수/분야별로 평균예산을 구해보았음**
- “이 평수의 이 작업에는 이정도 예산이 든다” 라는 의미보다는, **“이 평수에서 이 작업을 하는 사람들이 어느정도의 예산을 사용하고 있는가”** 라고 해석하는 게 더 좋을 듯
- 예를 들어 5평의 평균 홈스타일링 예산은 171만원 → 포스팅을 본 사람들이 인지하고 있는 평균 예산이 171만원이라고 볼 수 있으므로, 해당 숫자를 이용하면 “100만원으로 5평 스타일링하기” 처럼 더 관심을 끌 수 있는(5평 홈스타일링에 100만원은 많이 저렴하다고 생각할테니) **컨텐츠 가이드라인**을 잡을 수 있지 않을까...?

## 5. 연도별 인기있었던 스타일은?

## 6. 가족형태별 인기있는 스타일은?

## 7. 분야별로는 어떤 작업이 가장 인기있을까?

행 레이블	개수 : 작업	합계 : 조회대비인기도	합계 : 조회대비스크랩	합계 : 조회대비좋아요
▣ 건축	279	9.8	2.4%	41.1%
반셀프	47	11.4	2.4%	43.4%
셀프•DIY	5	11.1	2.6%	37.2%
전문가	227	9.3	2.4%	40.5%
▣ 리모델링	1729	12.7	2.7%	30.0%
반셀프	434	11.9	2.7%	30.3%
셀프•DIY	69	12.7	2.6%	33.6%
전문가	1226	13.0	2.7%	29.7%
▣ 부분공사	340	7.2	2.5%	28.0%
반셀프	157	6.8	2.5%	27.5%
셀프•DIY	57	9.9	2.7%	28.1%
전문가	126	6.5	2.4%	28.6%
▣ 홈스타일링	2552	10.3	2.9%	25.9%
반셀프	225	10.4	2.7%	28.0%
셀프•DIY	2223	10.3	2.9%	25.5%
전문가	104	8.7	2.6%	28.6%
총합계	4900	11.0	2.8%	28.2%

- 공사 규모가 커질수록(홈스타일링 → 부분공사 → 리모델링 → 건축) 셀프 → 반셀프 → 전문가를 선호하는 경향이 있음
- 건축의 경우 조회대비 스크랩 비율은 낮지만, 좋아요 비율이 타 분야보다 높음. 실질적인 수요는 적지만 규모가 크고 공이 들어간 콘텐츠라 좋아요를 많이 눌러주는듯?

## 8. 연도와 가족형태별로 집들이 콘텐츠는 얼마나 많아졌을까? / 2022년 남은 기간동안 더 늘려야 하는 콘텐츠는?

행 레이블	개수 : 가족형태	증가율
싱글라이프	971	
2019년	44	
2020년	233	530%
2021년	366	157%
2022년	421	115%
취학 자녀가 있는 집	564	
2019년	18	
2020년	134	744%
2021년	231	172%
2022년	232	100%
부모님과 함께 사는 집	341	
2019년	10	
2020년	69	690%
2021년	141	204%
2022년	155	110%
신혼부부	1955	
2019년	86	
2020년	560	651%
2021년	642	115%
2022년	857	133%
아기가 있는 집	727	
2019년	34	
2020년	238	700%
2021년	229	96%
2022년	290	127%
총합계	4558	

- 2022년의 경우 남은 기간(10/8~12/31)을 고려해 예상되는 최종 포스팅 수를 구했음
- 콘텐츠는 코로나가 본격적으로 시작되는 2020년에 가장 폭발적으로 성장하기 시작해서 매년 늘어나는 중
- 2021 → 2022로 오면서 신혼부부, 아기가 있는 집만 전년대비 콘텐츠량 증가율이 높아질 것으로 추측, 나머지 가족형태는 증가율이 둔화세를 보임
- 가설 1에서 신혼부부보다 싱글라이프 콘텐츠가 더 인기있음을 알았으므로, **2022년의 남은 기간동안 싱글라이프 콘텐츠 늘리기에 집중하면 좋을 듯**

## 9. 연도와 스타일별로 집들이 콘텐츠는 얼마나 많아졌을까? / 2022년 남은 기간동안 더 늘려야 하는 콘텐츠는? or 2023년 유행 스타일 예측?

### <질문>

- 쪼갬 표본 수가 전체 데이터 대비 몇% 이하가 될때 제거하는 것이 합리적인지
- 스타일 중복되는것 전처리 부분
- 예산 값 0인것들 전처리 부분