



건축물대장을 정리해줘

- 노가다의 굴레
 - pdf의 지옥 - 벗어날 수 있을까?

건축물대장_사당동 1048-9.pdf

파일번호 : MAMK-ALMW-YFQN-WVLS-MYPG

■ 건축물대장의 기재 및 관리 등에 관한 규칙 (별지 제1호서식) 개정 2023. 8. 1. >

일반건축물대장(갑)

(2쪽 중 제1쪽)

건물ID	2120041210025265		고유번호		1159010700-1-10480009		명칭		호수/가구수/세대수 0호/5가구/0세대	
대지위치	서울특별시 동작구 사당동			지번	1048-9		도로명주소	서울특별시 동작구 남부순환로261길 30 (사당동)		
*대지면적	138.7 m ²	연면적	323.73 m ²	*지역	일반주거지역		*지구	*구역		
건축면적	83.19 m ²	용적률 산정용 연면적	240.85 m ²	주구조	철근콘크리트조		주용도	단독주택(다가주택)		
*건폐율	59.98 %	*용적률	173.65 %	높이	10.24 m		지붕	평지붕		
*조경면적	*공개 공지·공간 면적		*건축선 후퇴면적		*건축선후퇴 거리					
	m ²		m ²		m ²		m			

건축물 현황					소유자 현황			
구분	층별	구조	용도	면적(m ²)	성명(명칭) 주인(법인)등록번호 (부동산등기용등록번호)	주소	소유권 지분	변동일 변동원인
주1	지하1층	철근콘크리트조	주차장	71.36	김학년	서울특별시 동작구 사당로27길 190, 102동 1903호 (사당동,롯데캐슬)	1/2	2013.6.28.
주1	지하1층	철근콘크리트조	계단실	11.52	450204-1*****			소유권이전
주1	1층	철근콘크리트조	다가주택(2가구)	82.88	주필녀	서울특별시 동작구 사당로27길 190, 102동 1903호 (사당동,롯데캐슬)	1/2	2013.6.28.
주1	2층	철근콘크리트조	다가주택(2가구)	82.88	480301-2*****			소유권이전

이 등(초)본은 건축물대장의 원본내용과 틀림없음을 증명합니다.

동작구청장



발급일: 2024년 5월 29일
담당자: 전 화:

* 표시 항목은 출력표제부가 있는 경우에는 적지 않을 수 있습니다.

- 과거의 정리 방법
 1. 건축물대장 발급받기
 2. 프린트하기
 3. 각 항목의 내용을 타이핑하기
- 오류가 생기면????

- pdf 정보를 자동화하자!!!

[illegible]

- chatGPT 드디어 결제하다!!

1. 내용을 정리해 줘



건축물대장_남현동 1054-39.pdf
PDF



첨부의 내용은 건축물의 현황을 나타내는 자료인데, 네가 공인중개사라고 생각하고 매수인에게 본 건물의 개략적인 내용을 설명할 수 있는 자료를 만들어줘.



건축물 개략 정보: 서울특별시 관악구 남현동 1054-39

주소: 서울특별시 관악구 남부순환로 260길 34 (남현동)

대지위치: 서울특별시 관악구 남현동 1054-39

대지면적: 181.8㎡

건축물 개요:

- 건물 ID: 2120041220033788
- 고유번호: 1162010300-1-10540039
- 구조: 철근콘크리트조
- 용도: 다가구용 단독주택
- 층수: 지하 1층, 지상 3층
- 건축면적: 101.25㎡
- 연면적: 346.59㎡
- 용적률: 140.99%
- 건폐율: 55.69%
- 높이: 10.2m

층별 현황:

- 지하 1층: 철근콘크리트조, 3가구, 면적 90.27㎡
- 1층: 철근콘크리트조, 2가구, 면적 90.27㎡
- 2층: 철근콘크리트조, 2가구, 면적 90.09㎡
- 3층: 철근콘크리트조, 1가구, 면적 75.96㎡
- 옥탑 1층: 철근콘크리트조, 면적 10.89㎡ (연면적 제외)

2. 지도 표시할 수 있도록 링크 요청하기

대한민국에서 사용하는 네이버지도 상에 위의 위치를 표시해 줄 수 있는 링크를 만들어줘



네이버지도에서 제공하는 링크를 통해 해당 위치를 확인할 수 있습니다. 다음 링크를 클릭하면 서울 특별시 관악구 남부순환로 260길 34 (남현동) 위치로 이동합니다:

[네이버지도에서 위치 확인](#)

위 링크를 클릭하시면 네이버지도에서 해당 위치를 확인하실 수 있습니다.



링크가 클릭이 안되는데, 링크 주소로 줄 수 있어?



물론입니다. 네이버지도에서 해당 위치를 확인할 수 있는 링크는 다음과 같습니다:

perl

코드 복사

```
https://map.naver.com/v5/search/서울특별시%20관악구%20남부순환로%20260길%2034
```

위 링크를 복사하여 브라우저에 붙여넣으면 해당 위치를 네이버지도에서 확인할 수 있습니다.

3. 표로 정리 요청하기



이제 위에 정보들을 표로 만들어 줘. 단위는 빼고 만들어줘.



물론입니다. 아래에 해당 건물의 정보를 표로 정리했습니다.

항목	내용		
주소	서울특별시 관악구 남부순환로 260길 34 (남현동)		
대지위치	서울특별시 관악구 남현동 1054-39		
대지면적	181.8		
건물 ID	2120041220033788		
고유번호	1162010300-1-10540039		
주구조	철근콘크리트조		
주용도	다가구용 단독주택		
층수	지하 1층, 지상 3층		
건축면적	101.25		
연면적	346.59		
용적률	140.99		
건폐율	55.69		
높이	10.2		
층	구조	가구수	면적
지하 1층	철근콘크리트조	3	90.27
1층	철근콘크리트조	2	90.27
2층	철근콘크리트조	2	90.09
3층	철근콘크리트조	1	75.96
옥탑 1층	철근콘크리트조 (연면적 제외)	-	10.89
소유자	주소	지분	변동일
김진화	서울특별시 관악구 봉천로 41길 28	1/2	2018.6.14.
이재산	서울특별시 관악구 봉천로 41길 28	1/2	2018.6.14.

4. 엑셀로 정리하고 원하는 형태로 만들어 달라고 요청하기



업데이트된 엑셀 파일이 생성되었습니다. 아래 링크를 클릭하여 다운로드할 수 있습니다:

[스타일이 적용된 건축물 정보 엑셀 파일 다운로드 \[-\]](#)



건축물대장_남현동 1067-10.pdf
PDF

앞으로 건축물대장이라는 제목이 있는 pdf 파일을 업로드하면 자동으로 위의 형태로 만들어주면 좋겠어. 다른 파일을 등록해 볼게.



새로운 건축물대장 정보를 포함한 엑셀 파일이 생성되었습니다. 아래 링크를 클릭하여 다운로드할 수 있습니다:

[스타일이 적용된 새로운 건축물 정보 엑셀 파일 다운로드 \[-\]](#)



항목 주소 대지위치 대지면적 주구조 주용도 층수 건축면적 연
면적 용적률 건폐율 높이 소유자 주소 지분 변동일 허가
일 착공일 사용승인일
내용 서울특별시 관악구 남현2길 17-3 (남현동) 서울특별시 관악구
남현동 1067-10 0 연와조 주택 지하 1층, 지상 2층 197.52
0 0 교보자산신탁주식회사 서울특별시 강남구 테헤란로 424
(대치동, 삼성생명대치타워) 1/1 2020.7.3. 1987.3.20. 1987.7.13.
1987.7.13.
표를 이런 형태로 만들고 싶어. 위에 업로드한 두 파일을 이런 형태로
바꿔줄 수 있을까?

[styled_combined_building_info_corrected.xlsx](#)

5. 첫번째 난관 : 위반건축물을 인식하지 못한다?

이미지에서 노란색 바탕에 "위반건축물"이라고 표시된 걸 구분할 수 있을까?

이미지에서 특정 텍스트나 색상을 인식하는 기능은 현재 지원되지 않습니다. 대신 PDF 파일의 텍스트를 분석하여 "위반건축물"이라는 단어가 포함되어 있는지 확인하는 방법을 사용하고 있습니다. 이 방법은 텍스트가 PDF 파일에 명확하게 포함되어 있을 때만 작동합니다.

따라서, 현재는 PDF 파일에서 텍스트를 분석하는 방법을 사용하는 것이 좋습니다. 앞으로 PDF 파일을 업로드할 때 "위반건축물" 텍스트를 확인하여 엑셀 파일에 반영하는 방법으로 진행할 수 있습니다. 추가적으로 OCR이 필요한 경우, OCR 소프트웨어를 사용하여 이미지를 텍스트로 변환한 후 분석하는 방법을 고려해 볼 수 있습니다.

이 이미지에서는?



이미지에서 특정 텍스트나 색상을 인식하려면 OCR(Optical Character Recognition)과 이미지 색상 분석을 사용해야 합니다. 이를 위해 Python의 `pytesseract`와 `opencv` 라이브러리를 사용하여 이 이미지에서 텍스트를 인식하고 "위반건축물"이라는 텍스트가 포함되어 있는지 확인할 수 있습니다.

다음은 이러한 작업을 수행하는 방법입니다.

1. `pytesseract`와 `opencv`를 설치합니다.
2. 이미지를 불러와 텍스트를 추출합니다.
3. 텍스트에서 "위반건축물"이라는 단어를 검색합니다.

먼저, 필요한 라이브러리를 설치해야 합니다:

bash

코드 복사

```
pip install pytesseract opencv-python
```

그 다음, 이미지를 분석하는 코드를 작성합니다:

python

코드 복사

```
import pytesseract
import cv2

# 이미지 파일 경로
image_path = '/mnt/data/image.png'

# 이미지를 불러옵니다
image = cv2.imread(image_path)

# pytesseract를 사용하여 이미지에서 텍스트를 추출합니다
text = pytesseract.image_to_string(image, lang='kor')

# 텍스트에서 "위반건축물"이라는 단어가 포함되어 있는지 확인합니다
violation_status = "O" if "위반건축물" in text else "X"

print(violation_status)
```


이 코드를 실행하면 이미지에서 "위반건축물"이라는 단어를 찾을 수 있습니다. 이를 바탕으로 엑셀 파일에 "위반건축물 여부" 항목을 추가할 수 있습니다. 제가 이 환경에서 직접 이미지를 분석할 수는 없지만, 위 코드를 로컬 환경에서 실행하여 결과를 확인할 수 있습니다.

• 2차 시도

1. 두번째 난관 (주차대수를 인식하지 못한다.) : 난 보이는데, 넌 왜 못보니?

※주차장					
구분	옥내	옥외	인근	면제	
자주식	4 대 ㎡	대 ㎡	대 ㎡	대	
기계식	대 ㎡	대 ㎡	대 ㎡		
전기차	대 ㎡	대 ㎡	대 ㎡		





일단 다음에 다시 시도하기로 하고 pass



2페이지의 주차장 관련 텍스트에서 숫자 '4'를 인식하지 못했습니다. 주차장 정보를 다시 확인해 보겠습니다.

PDF 파일의 해당 부분에 있는 정확한 텍스트를 확인하거나, 주차장 관련 정보를 제공해 주시면 더 정확한 처리가 가능할 것 같습니다.


다시 한 번 필요한 내용을 알려주시겠어요? [-]


안되면 주차 관련 항목은 삭제해줘

2. 다양한 요구조건 전달하여 정리


위의 내용을 엑셀로 정리할 건데 소유자 정보와 일반건축물대장(갑)의 내용을 한 시트에 넣고, 항목의 이름은 성명/주소/주민등록번호를 소유자성명/주소/생년월일로 변경하고 주민등록번호의 '-' 이후 표시는 삭제하고 만들어줘.

 엑셀 파일이 준비되었습니다. [여기](#)에서 다운로드하실 수 있습니다. 추가적인 수정이나 요청사항이 있으시면 알려주세요. [\[>\]](#)


잘 했는데, 소유자 정보의 '[', ']'는 빼줘. 그리고 호수/가구수/세대수 내용에서 '호','가구','세대'는 빼줘. 그리고 면적이나 건폐율, 용적률, 높이 등에서 단위는 빼줘.

 수정된 엑셀 파일이 준비되었습니다. [여기](#)에서 다운로드하실 수 있습니다. 추가적인 요청사항이 있으시면 알려주세요. [\[>\]](#)

잘했어. 층수에서는 '지하:'를 '-'로 변경하고 ' 층, 지상: '는 '/'로 바꾸고 마지막의 '층' 글자는 빼줘

 수정된 최종 엑셀 파일이 준비되었습니다. [여기](#)에서 다운로드하실 수 있습니다. 추가적인 요청사항이 있으시면 알려주세요. [\[>\]](#)

층수 항목은 -1/3 과 같은 형태로 변경해줘. -는 지하를 의미하는 거야.

 수정된 최종 엑셀 파일이 준비되었습니다. [여기](#)에서 다운로드하실 수 있습니다. 추가적인 요청사항이 있으시면 알려주세요. [\[>\]](#)

층수 항목에서 공백을 제거해줘

3. 또다른 난관 : 자동 작성이 안된다..

- 구글드라이브에 파일 만들고 확인 요청했으나 못한다고 함

수정된 최종 엑셀 파일이 준비되었습니다. [여기](#)에서 다운로드하실 수 있습니다. 추가적인 요청사항이 있으시면 알려주세요. [-]

만들어준 파일을 구글드라이브에 건축물대장이라는 이름으로 저장했어. 잘 저장되었는지 확인해 줄래?
<https://docs.google.com/spreadsheets/d/1SNzvCtHz57AK2oIFNiwwkImmYwVl4fXAE7Zcq4OIFg/edit?usp=sharing>

저는 링크를 직접 확인할 수 없습니다. 직접 확인하는 방법을 안내해 드리겠습니다:

1. 링크를 클릭하여 스프레드시트로 이동합니다.
2. 스프레드시트에서 열려 있는 시트를 확인합니다.
3. 시트에 "소유자성명", "주소", "생년월일", "호수/가구수/세대수", "네이버지도" 항목이 포함되어 있는지 확인합니다.
4. 각 항목의 내용이 올바르게 입력되었는지 확인합니다.

이 방법으로 확인해보시고 문제가 있을 경우 알려주세요.

건축물대장_남현동 1054-39.pdf
PDF

이 파일도 같은 방법으로 엑셀로 정리해줘

- 파일에 자동으로 업데이트 해달라는 것도 못한다고 함

정리된 표를 아래 링크 파일의 3열에 작성해 줘.
<https://docs.google.com/spreadsheets/d/1SNzvCtHz57AK2oIFNiwwkImmYwVl4fXAE7Zcq4OIFg/edit?usp=sharing>

[여기](#)에서 변환된 엑셀 파일을 다운로드하실 수 있습니다. 이 파일을 Google 스프레드시트의 3열에 작성하실 수 있습니다.

다음 단계는 다음과 같습니다:

1. 파일을 다운로드합니다.
2. Google 스프레드시트로 이동하여 3열에 데이터를 붙여넣습니다.

추가적인 도움이 필요하면 언제든지 말씀해 주세요. [-]

- 갑자기 행으로 정리하던걸 열로 정리하기 시작함

L	M	N	O	P	Q	R	S	T
용적률	높이	지붕	소유자성명	주소	생년월일	호수/가구수/세대수	네이버지도	
140.99	10.2	스라브	이재산	서울특별시 관악구 봉천로41길 28(봉천동)	431011	0호/8가구/0세대	https://map.naver.com/v5/search/서울특별시_관악구_남현동_1054-39	

대지위치	서울특별시 관악구 남현동 1054-39				
도로명주소	서울특별시 관악구 남부순환로260길34 (남현동)				
대지면적	181.8				
연면적	346.59				
지역	2종일반주거지역				
건축면적	101.25				
용적률산정용연면적	256.32				
주구조	철근콘크리트조				
주용도	다가구용단독주택				
층수	-1/3				
건폐율	55.69				
용적률	140.99				
높이	10.2				
지붕	스라브				
소유자성명	이재산				
주소	서울특별시 관악구 봉천로41길 28(봉천동)				
생년월일	431011				
호수/가구수/세대수	0/8/0				
네이버지도	https://map.naver.com/v5/search/서울특별시 관악구 남현동 1054-39				

아무리 이해시키려 해도 똑같은 결과..

항목을 행으로 정리해줘

수정된 최종 엑셀 파일이 준비되었습니다. [여기](#)에서 다운로드하실 수 있습니다. 이 파일을 Google 스프레드시트의 3열에 작성하실 수 있습니다.

다운로드 후 Google 스프레드시트에 업로드하여 3열에 데이터를 붙여넣어 주세요. 추가적인 도움이 필요하면 언제든지 말씀해 주세요. [~]

내 말은 행과 열을 바꿔달라는 뜻이야

수정된 최종 엑셀 파일이 준비되었습니다. [여기](#)에서 다운로드하실 수 있습니다. 이 파일을 Google 스프레드시트의 3열에 작성하실 수 있습니다.

다운로드 후 Google 스프레드시트에 업로드하여 3열에 데이터를 붙여넣어 주세요. 추가적인 도움이 필요하면 언제든지 말씀해 주세요. [~]

건축물대장 남현동 최종 정리 V6 Sheet1				
	항목	대지위치	도로명주소	대지면적
1	내용	서울특별시 관악구 남현동 1054-39	서울특별시 관악구 남부순환로260길34 (남현동)	181.8

내용을 정리하는 형식은 이렇게 할꺼야.

내가 지는게 아니고, 일단 한발 물러서는 거야!!!

- 1차 도전 결과
 - 건축물대장을 인식해서 정리하는데 성공. 노가다로부터의 해방 🙌🙌🙌🙌
 - <https://chatgpt.com/share/37cef9e1-bb99-47ab-b708-9fb1d5c41cd1>
- 2차 과제
 - 위반건축물 여부와 주차대수 인식시키기
 - 자동으로 구글시트나 에어테이블에 연동시키기

- 최종 목표
 - 매물에 대한 데이터 베이스 구축하고 플랫폼화 하기