

CMM 데이터 전처리

전제사항

- 현재 작업 사항과 문제점: 파일 끝에 NG는 불량, OK는 정상이고, 아무것도 안 붙은 것은 불량으로 판별
- 데이터, 작업 사항(텍스트 -> 파일별 엑셀 -> 엑셀 집계)

전처리 오류 해결 및 코드 개선

- 지금
- **에러 원인분석:** 현재 문제점에 대해 원인 확인: 86개 항목인데 50개정도 밖에 변환 안됨
 - 예상 원인 1) 다른 파일이랑 공통되는 항목만 들고와서 누락된걸로 추정
 - 예상 원인 2) Smmf 항목은 제거해서 50개 정도 준 것일 수도 있음
 - 예상 원인 3) 편차 값이 없는 것은 제거해서 준 것일 수 있음.
- **기존 파일을 취합하는 코드에 문제 없는지 파악**
 - 어떤 방식으로 항목을 들고와서 취합하는지,
 - 취합할 때 측정 파일의 스케일(shape)이 달라도 특성 같은 것끼리 구분해서 취합하는지.
 - ex) (86, N), (84,N)개 취합시 특성 잘 확인하고 취합하는지
- **전처리 코드 추가 개발:**
 - 현재는 측정 파일에서 '편차' 항목만 들고왔는데 나머지 컬럼(측정값, 기준값, 상한 공차, 하한공차, '편차', 판정)도 추가로 들고 와서 취합하도록 변경
 - 컬럼 포맷 변경: 현재 측정 파일의 "도형"과 "항목"을 쉼표(,)로 구분해 두었는데 언더바(_)로 변경

- 현재: 특성(도형, 항목), ⇒ "원1(I) <상>, D" 컬럼으로 생성

	I	J	
	도형	항목	
3	평면1	평면도	
3	평면1	SMmf	4
5	원1(I) <상>	D	

✓	f_x	원1(I) <상>,D
C	D	
원1(I) <상>,D	원2(I) <중>	

- 변경: 현재: 특(도형_항목), ⇒ "원1(I) <상>_ D"
- 추가 특성 취합시 컬럼 포맷은 측정항목(도형+항목)_컬럼(측정값, 기준값, 상한공차, 하한공차, '편차', 판정)으로 변경
 - ex) 원1(I) <상>_ D_측정값, 원1(I) <상>_ D_기준값, 원1(I) <상>_ D_상한공차

전체 데이터에 데이터 전처리 작업

- 현재 700개 정도에 대해서만 취합 파일이 생성되었는데, 이를 전체 데이터(약 1900개)에 대해서 전처리 필