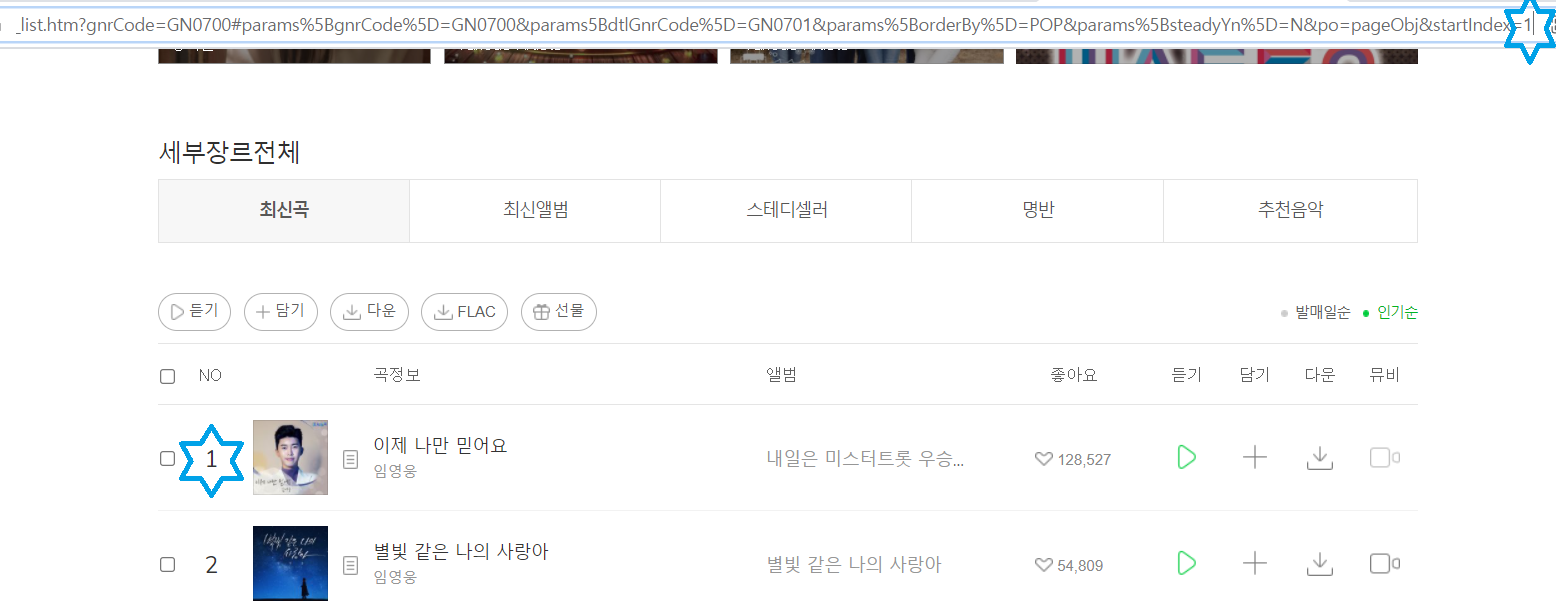
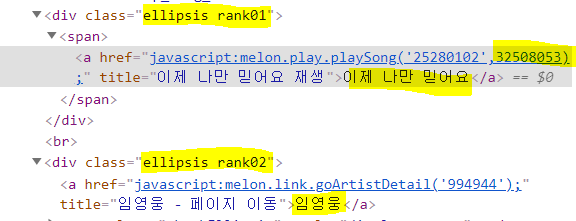
1. **곡 번호 수집**
2. **곡 번호로 가사 수집**
3. **19금 노래들 곡 번호 수집**
4. **19금 노래들 곡 번호로 가사 수집**
5. **기존 데이터에 19금 노래 가사 추가**
6. **결측치 제거**
7. **중복제거, 모든 장르 데이터 통합**
8. **겹치는 장르 묶어주기**
9. **훈련, 테스트 데이터 나누기**
10. **전처리**

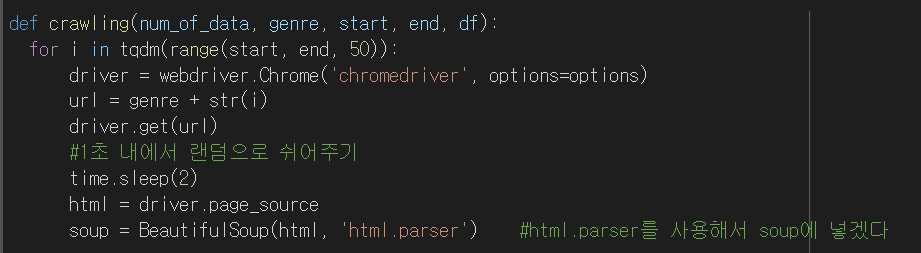
**곡 번호 수집**

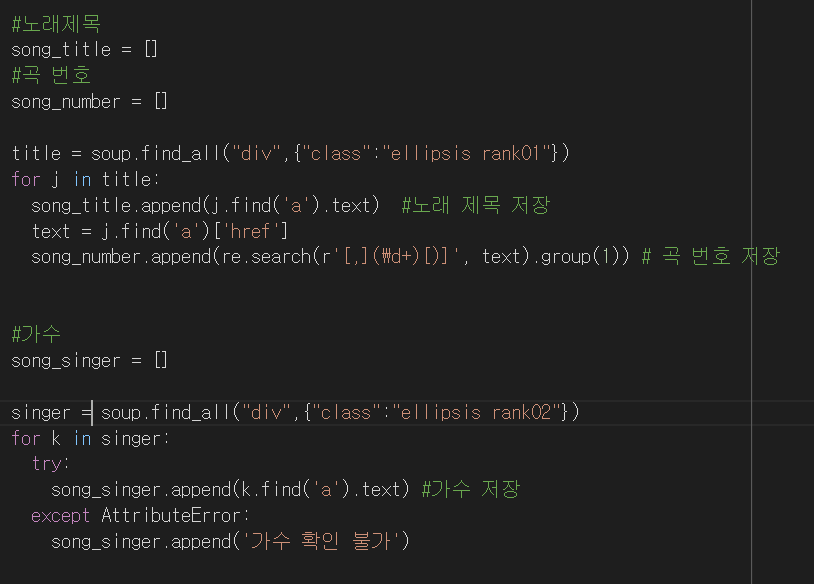


주소 형태

‘<https://www.melon.com/genre/song_list.htm?gnrCode=GN0700#params%5BgnrCode%5D=GN0700&params5BdtlGnrCode%5D=GN0701&params%5BorderBy%5D=POP&params%5BsteadyYn%5D=N&po=pageObj&startIndex=>’ + 1, 51, 101 , ,,,,

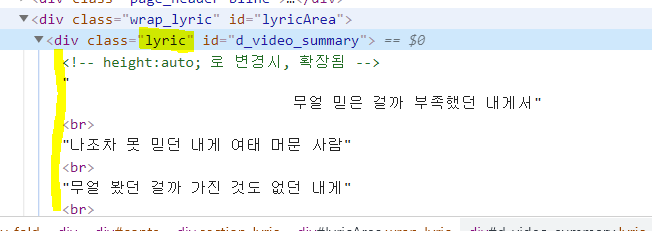
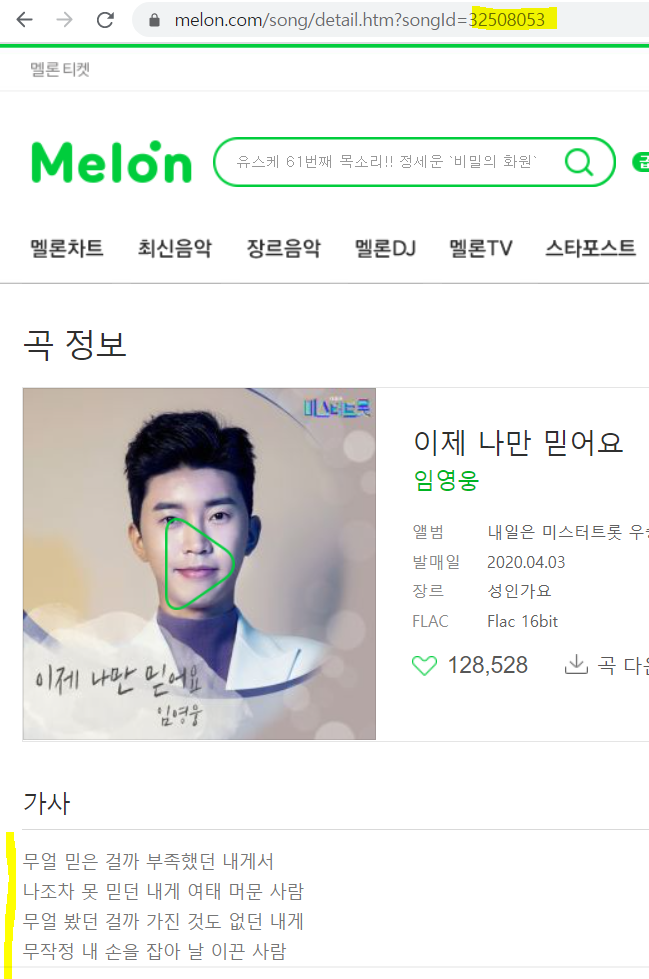


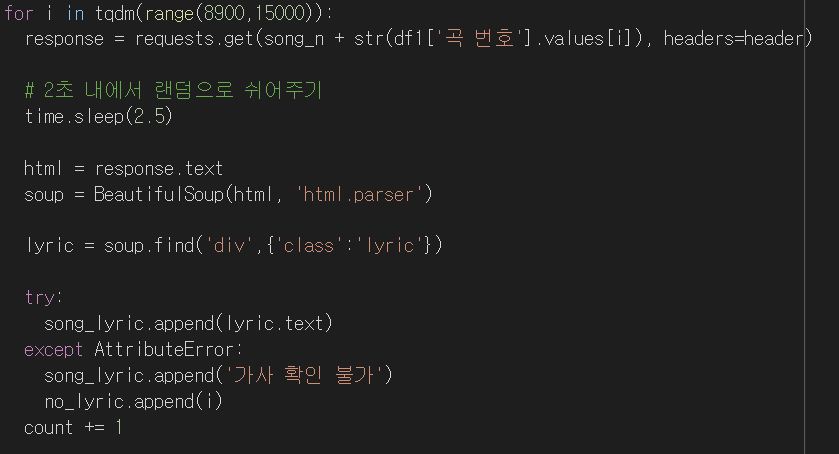


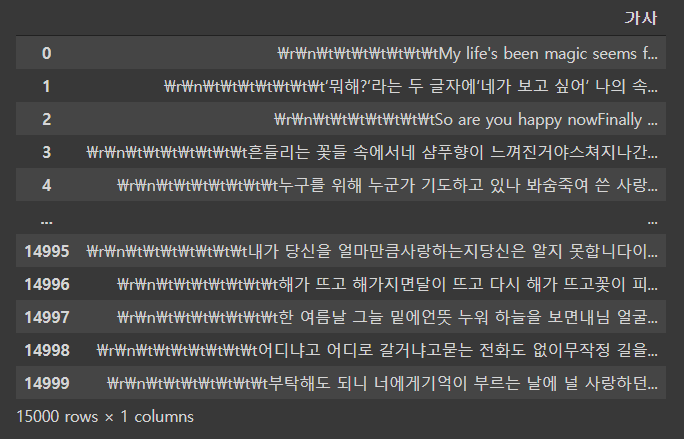




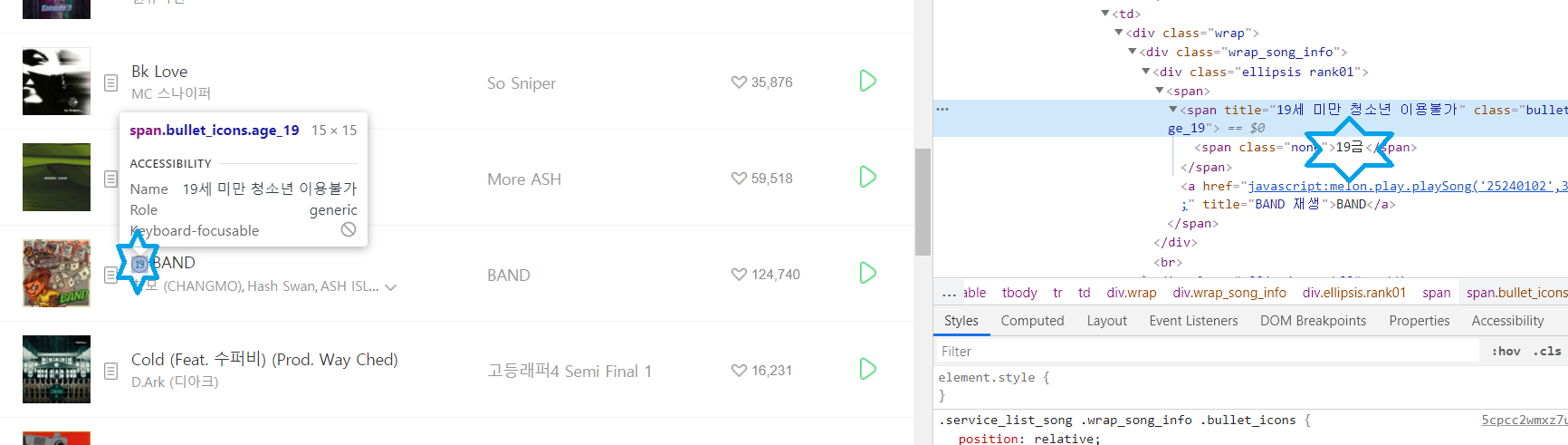
**곡 번호로 가사 수집**

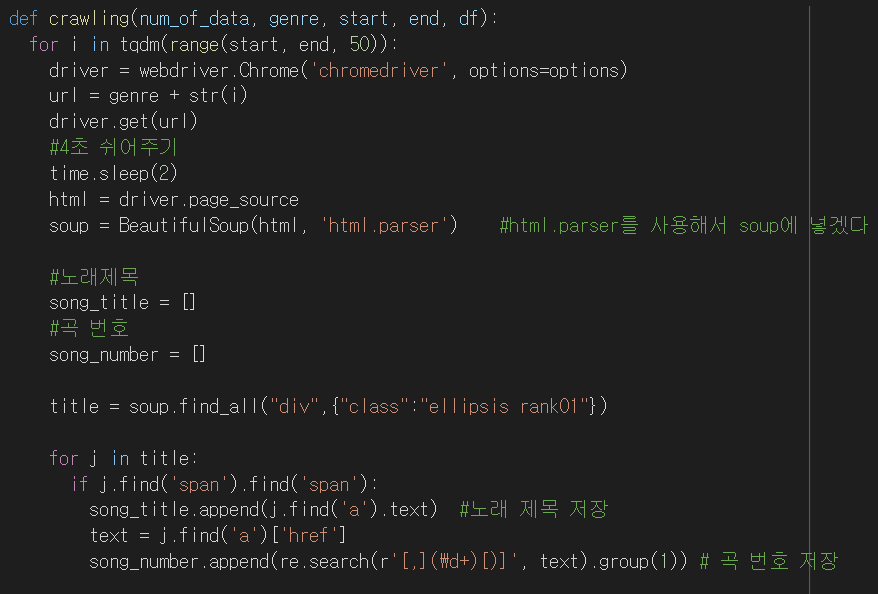


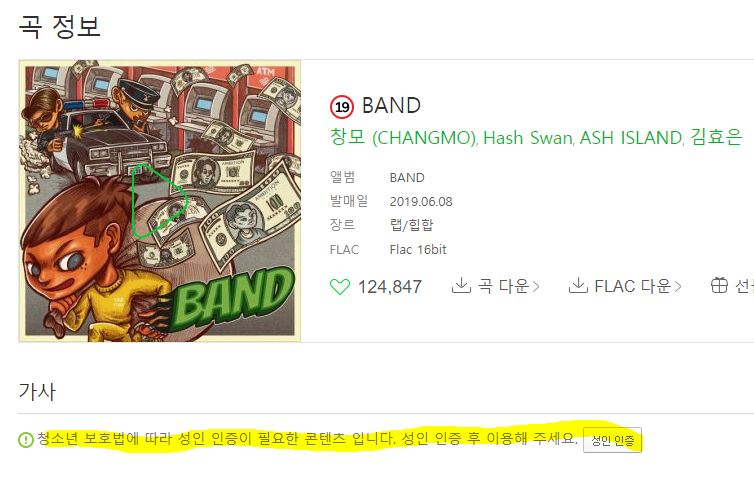




**19금 노래들 곡 번호 수집**

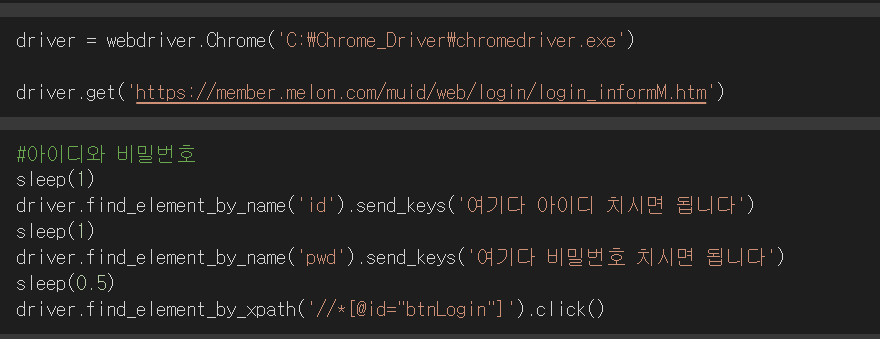
****

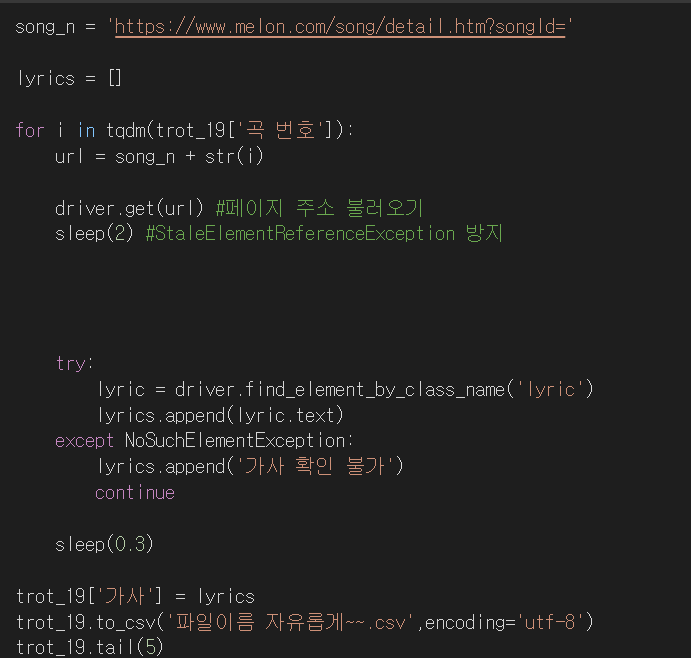
****

****

**19금 노래들 곡 번호로 가사 수집**

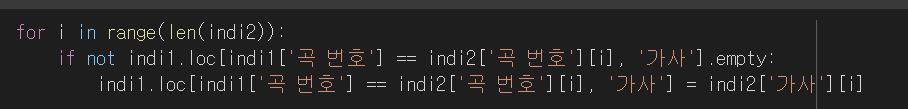
**멜론 로그인**

****

****

****

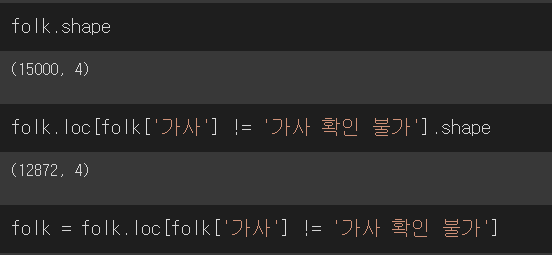
**기존 데이터에 19금 노래 가사 추가**

****

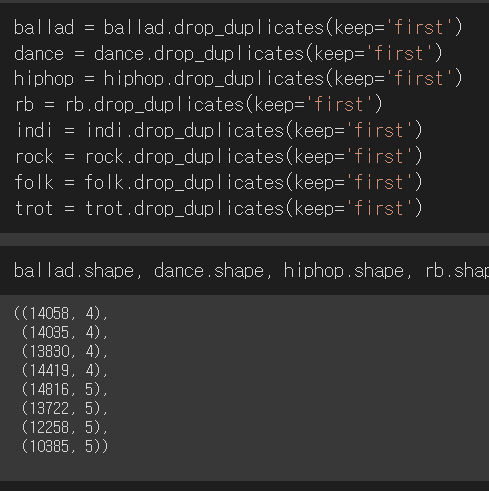
기존 데이터의 곡 번호 중

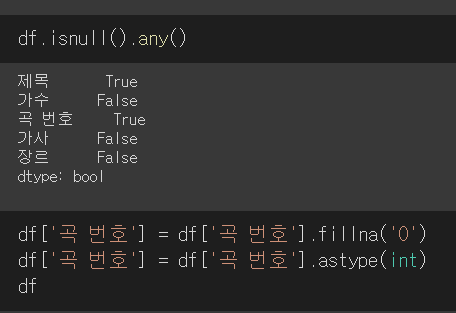
불러온 19금 가사 데이터의 곡 번호와 일치하면 가사 추가

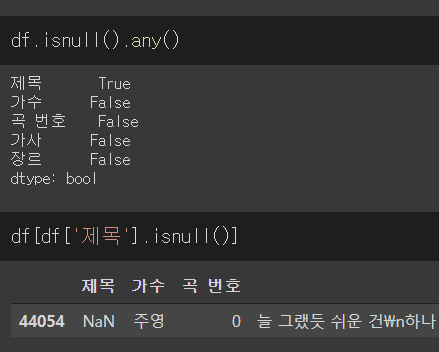
**결측치 제거**



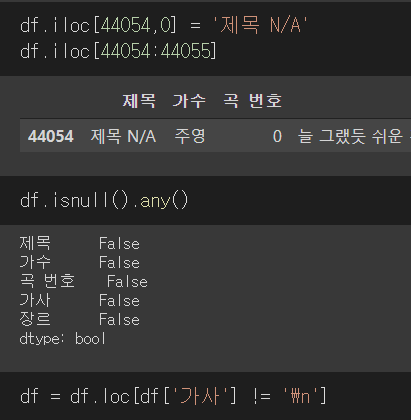
**중복 데이터 제거, 모든 장르 데이터 통합**

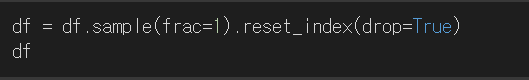








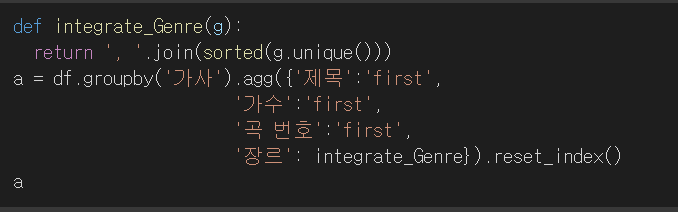


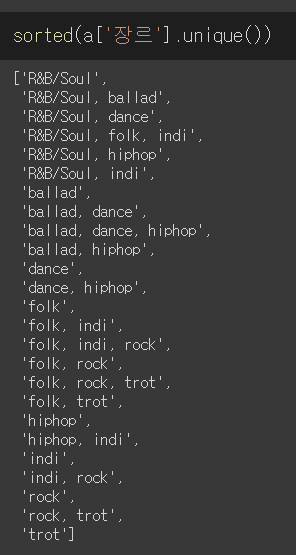


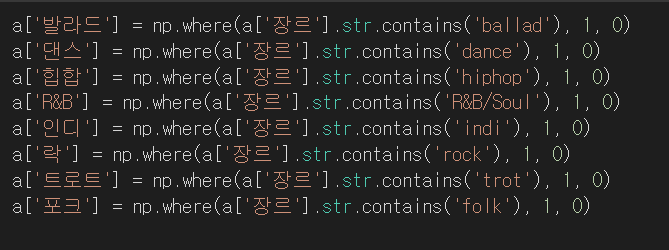


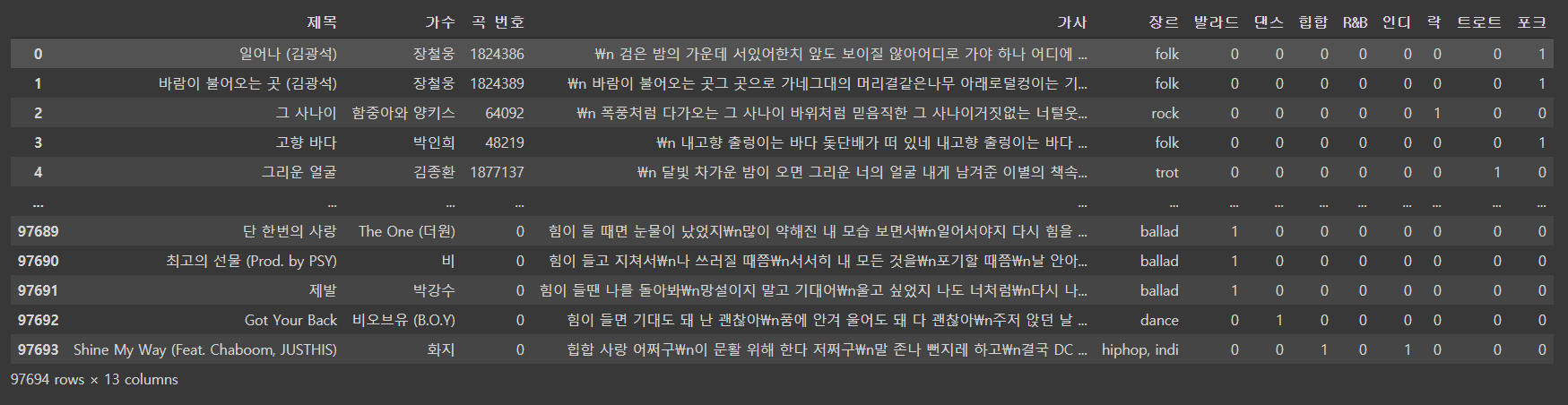
**겹치는 장르 묶어주기**



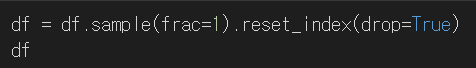


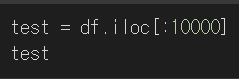


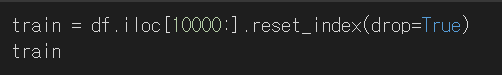




**훈련 / 테스트 데이터 나누기**







**전처리**

