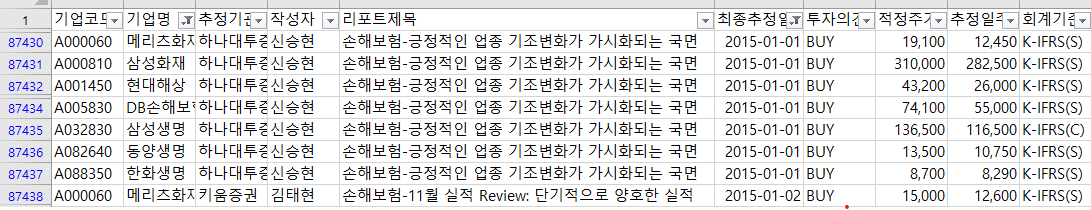
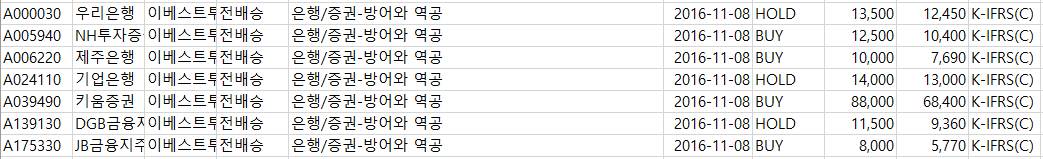
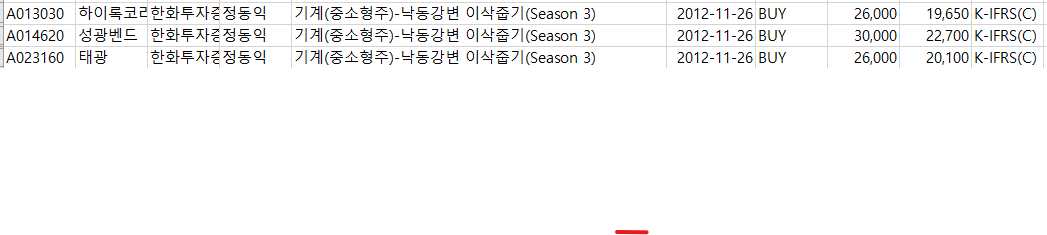
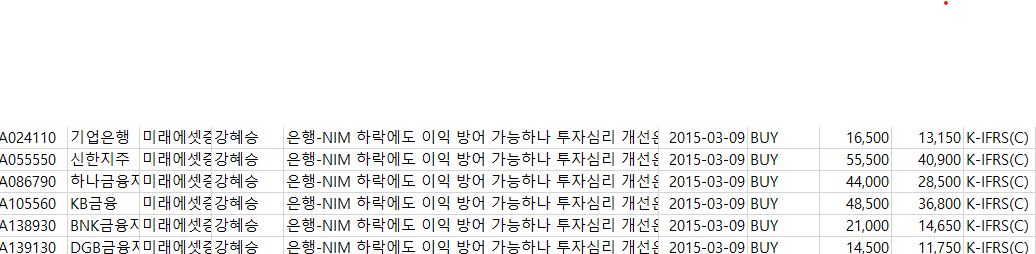
**데이터 이상치 정리**

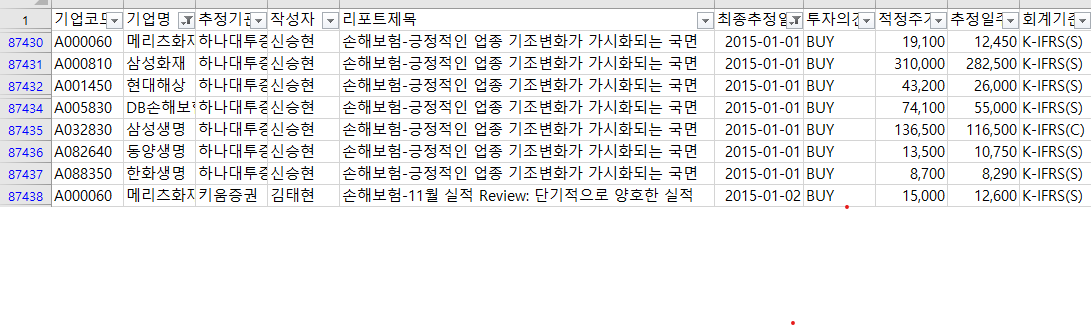
출처: analyst\_daily\_forecast\_09\_19\_annual\_adj\_보고서 제목











**특이한 title 정리**

|  |  |
| --- | --- |
|  |  |
|  | 영단어 Risk는 중요한 단어라고 생각됨 |
|  | - ‘모두투어’라는 기업명 대신 ‘여행-‘ 존재  - 따라서 ‘-‘앞에 모든 단어들은 삭제하는 것이 어떨지..  - 마찬가지로 ‘아시아나항공’ 대신 ‘항공-‘존재.  - ‘-‘앞에 있는 음식료는 제거하는 게 나을지?  - 기업명은 ‘오리온홀딩스’, ‘-‘앞에는 ‘오리온’ |
|  | - 기업명, 한자, 숫자 제거하면 title로 남는게 없음  - 영단어로만 이루어져 있어서 남는게 없음 |
|  | 여기서 Bag은 제거? |
|  | - ‘↑’ ‘↓’는 없는 것으로 보임 (차후에 코딩으로 한번도 확인 필요)  - 대신 Down과 up 존재  - ‘->’는 있지만 그리 중요해 보이지 않음 |
|  | - 물음표 많은 편 |