1. sales.csv 파일은 2014년 1년간 3507명의 구매정보에 관한 데이터이고 silver\_demo.csv는 각 고객의 개인정보를 포함하고 있다. 각 카테고리 중분류(CT2)별로 아래의 변수를 계산하시오.

* 총 구매수량
* 총 구매금액
* 구매횟수
* 평균 구매금액
* 구매한 적이 있는 고객의 수
* 각 구매고객의 평균연령
  1. 위의 6개 변수의 단변량, 이분량 분포를 살피고 변환이 필요한 경우 변환하시오.
  2. Heatmap과 군집분석 결과를 함께 관찰하여 어떤 특성을 가지는 중분류들이 같은 군집에 속하게 되는지 살피시오.
  3. 나이팅게일 차트, 별그림, 별그림을 사용하여 중분류 별 특징을 파악하시오.

1. 2015\_baseball.csv는 2015년 시즌 각 타자의 성적을 포함하고 있다. 각 변수의 설명은 아래와 같다. 평행좌표 플롯을 활용해 선수들의 성적 패턴에 어떤 경향이 있는지 알아보려 한다.
   1. 전체 데이터의 평행좌표 플롯을 그리고 선수기록들 간의 경향이 있는지 파악하시오.
   2. 팀별 평행좌표 플롯을 그리고 팀별 타자의 성적에 어떤 경향이 있는지 서술하시오.

|  |  |
| --- | --- |
| [AVG](javascript:sort('HRA_RT');) | 타율 |
| [G](javascript:sort('GAME_CN');) | 경기 |
| [PA](javascript:sort('PA_CN');) | 타석 |
| [AB](javascript:sort('AB_CN');) | 타수 |
| [R](javascript:sort('RUN_CN');) | 득점 |
| [H](javascript:sort('HIT_CN');) | 안타 |
| [2B](javascript:sort('H2_CN');) | 2루타 |
| [3B](javascript:sort('H3_CN');) | 3루타 |
| [HR](javascript:sort('HR_CN');) | 홈런 |
| [TB](javascript:sort('TB_CN');) | 루타 |
| [RBI](javascript:sort('RBI_CN');) | 타점 |
| [SAC](javascript:sort('SH_CN');) | 희생번트 |
| [SF](javascript:sort('SF_CN');) | 희생플라이 |