

- 1 Lab. Microsoft Power BI를 이용한 데이터 군집화하기
- 2
- 3 1. Microsoft Power BI Desktop 설치하기
- 4 1) <https://powerbi.microsoft.com/ko-kr/get-started/>
- 5 2) 주황색 [무료 다운로드] button click한다.
- 6 3) "Microsoft Store을(를) 여시겠습니까?" 에서 [Microsoft Store 열기] button click한다.
- 7 4) Microsoft Store로 이동한다.
- 8 5) Power BI Desktop App에서 [설치] button click하여 download한다.
- 9 6) [실행] button click한다.
- 10 7) 처음 실행할 때에는 [Welcome to Power BI Desktop] 계정등록을 한다.
- 11 8) 각 항목에 값을 입력한 후 [I accept] 체크하고 [Done] click한다.
- 12
- 13
- 14 2. 데이터 군집화하기
- 15 1) Power BI로 앞 Lab에서 작업한 Data Query를 불러온다.
- 16 -[홈] Tab > [데이터 가져오기] > Excel > C:/workspace/2016년 3분기 상장회사 손익계산서.xlsx 선택
- 17 한다.
- 18 2) [탐색 창]에서 '_2016_3분기보고서_02_손익계산서_20190418'을 선택하면 오른쪽에 미리보기가 나타난다.
- 19 3) 해당 sheet를 check하고 [로드] button click한다.
- 20 4) Power BI Desktop 오른쪽 목록 중 [시각화] > [분산형]를 선택하고 [필드]에서 '회사명'을 선택해서 [시각화]
- 21 의 [자세히]가 회사명을, [필드]에서 'Revenue'를 선택하여 [시각화]의 [X축]이 'Revenue'가 선택되게 하고, 마지
- 22 막으로 [필드]에서 'Gross Profit'을 선택하여 [시각화]의 [Y축]이 'Gross Profit'에 맞추게 한다.
- 23 5) 그러면 화면의 왼쪽에 산점도를 통한 데이터 군집화 그래프가 나타난다.
- 24 6) Chart의 우하단을 마우스로 적당한 크기로 늘린다.
- 25 7) 우하단 끝부분의 [...]을 클릭하여 나타나는 menu에서 [자동으로 클러스터 찾기]를 click한다.
- 26 8) [클러스터] 창이 나타나면 [이름]은 '매출액 대비 손익 기준 클러스터'로, [필드]는 '회사명'으로, [설명]도 '매출액
- 27 대비 손익 기준 클러스터'로, 마지막으로 [클러스터 수]는 자동으로 선택한 후 [확인] button click한다.
- 28 9) 참고로 [클러스터 수] 즉, 군집갯수는 선택 옵션이기에 이를 지정하지 않으면 Power BI가 최적의 군집갯수를 찾아
- 이를 기준으로 data를 군집화해준다.
- 10) 현재 cluster수는 4개임을 알 수 있다. 각 군집은 고유의 색으로 표현된다.
- 11) 군집화된 그룹은 페이지의 제일 오른쪽 [필드]에서 확인할 수 있는데, 자세한 군집화 내역을 보기 위해서 [필드] 영
- 역의 '매출액 대비 손익 기준 클러스터'에 right-mouse click해서 [클러스터 편집]을 선택하면 된다.
- 12) 현재 Power BI를 통한 군집화는 어떤 알고리즘을 사용했는지에 대한 정보는 제공해 주지 않는다는 단점이 있다.
- 13) 따라서 군집화 과정에 대한 탐색이 필요한 경우에는 Power BI에서 제공하는 자동 군집화보다는 알고리즘을 직접
- 선택할 수 있는 다른 방법을 사용하는 것이 좋다.