**비즈니스 모델링 2**

**Assignment 1**

**Due: 11/11, 2020 (수요일), 11:59PM**

**(Total Points: 9)**

**Instructions:** 다음 문제들을 잘 읽고, 요구되는 코드를 작성하고 설명하시오. 코드와 설명을 포함한 모든 답안을 ‘HW1\_김일주.r’ 하나의 파일에 저장하여 제출하시오. 설명을 요하는 문제의 경우 주석을 사용하도록 하고, 코드와 설명이 몇번 문제에 대한 답인지를 반드시 표시하시오. 늦은 제출은 받지 않으니, 제출시한에 맞추도록 유의하시오. 반드시 자신들의 이해를 충분히 보이는 방법으로, 자신들의 Words를 사용하여 답안을 작성하시오. 질문있을 시 언제든지 [ijkim@mju.ac.kr](mailto:ijkim@mju.ac.kr)로 연락하시오.

1. **승차식 잔디깎이 기계를 만드는 이 제조회사는 대대적인 판촉 캠페인을 벌이기 전에 잠재된 우수고객을 발견하고자 한다. 특히, 이 제조사는 개별 가구들을 Income(가계소득, 단위: 1천 달러)과 Lot Size(대지면적, 단위: 1천 평방피트)를 기준으로 잠재고객 여부(owner/nonowner)를 판별하려고 한다. 마케팅 전문가는 Riding Mowers.csv파일 안에 들어 있는 24개 랜덤표본가구 데이터를 사용한다.**
2. **(3pts) R을 사용하여 Lot Size와 Income의 관계를 결과변수인 owner/nonowner에 의해 컬러-코드화 된 산점도를 그리시오. 형식이 잘 갖춰진 플롯(읽기 쉬운 라벨(label)과 범례(legend)를 포함)을 만드시오. 이 산점도를 보고 owner/nonowner의 분포를 나머지 두 변수와 함께 설명하시오.**

**Owner는 Lot\_Size와 Income이 비례하는 모양으로 분포되어 있다. 즉, Lot\_Size가 클수록 Income이 증가하는 모양으로 오른쪽 상단에 많이 분포 되어있다. Nonowner는 owner에 상대적으로 왼쪽 아래에 분포되어 있으며, 역시 비례하는 모양으로 분포되어 있다. 하지만 owner와는 다르게 Los\_Size가 작을수록 Income이 감소하는 모양으로 분포되어 있다.**

1. **LaptopSalesJanuary2008.csv 파일은 런던 소재의 한 컴퓨터 체인점의 2008년 1월 노트북 판매 매출데이터이다. 이것은 2008년도 전체 매출데이터의 일부이다. 아래의 문제에서 요구하는 플롯을 그리고, 그와 관련된 질문들에 답하시오.**
2. **(3pts) 매장(Store.Postcode)별 평균 소매가격(Retail.Price)을 보여주는 막대차트를 그리시오. 평균 소매가격이 가장 높은 매장은 어느 곳인가? 반대로 가장 낮은 평균 소매가격은 어떤 매장인가? (Hint: aggregate함수를 통하여 먼저 막대차트를 그리기 위한 데이터를 만듬.)**

**평균 소매가격이 가장 높은 매장은 N17 6QA이고, 평균 소매가격이 가장 낮은 매장은 W4 3PH이다.**

1. **(3pts) 매장별 소매가격을 더 잘 비교하려면 병력 박스플롯을 그리시오. (1)에서 찾은 두 매장의 가격을 비교해 보시오. 두 매장의 가격분포가 어떤 차이점이 있는가? 어떤 설명이든 좋으니 본인이 발견한 차이점을 서술해보시오.**

**N17 6QA의 중위값, 1사분위값, 3사분위값이 W4 3PH보다 높았고, W4 3PH에 이상치가 더 많았다. 그리고 두 매장의 최솟값은 비슷했으나 최댓값은 N17 6QA가 더 높았다.**

**그리고 N17 6QA는 50단위로 가격이 분포되어 있는 반면에 W4 3PH는 100단위로 분포 되어있었다.**