이름: \_최윤영\_\_\_\_\_\_\_

**R 프로그래밍 Quiz**

빨간색 상자 안에 실행 결과 스크린샷해서 첨부해주세요.

최종적으로 RScript 파일과 Google docs와 함께 제출하시면 됩니다.

mpg data 관련해서 참고

|  |  |
| --- | --- |
| **mpg data** | **설명** |
| displ | 배기량 |
| hwy | 고속도로 연비 |
| cty | 도시 연비 |
| manufacturer | 제조 회사 |
| class | 차종 |

1. **Data preprocessing**

**문제 1번**

**1.1** ‘chevrolet’, ‘ford’, ‘honda’ 자동차의 고속도로 연비 (hwy) 평균을 알고자 한다. 이 회사들의 데이터를 추출한 후 각 회사의 고속도로 연비 (hwy) 전체 평균을 출력하여라.

|  |
| --- |
|  |

**1.2** ‘cty’와 ‘hwy’를 더한 ‘평균 연비 변수’인 ‘mean\_ch’를 data frame에 추가하여라. (하나로 연결된 dplyr 구문으로 실행하도록 할 것)

|  |
| --- |
|  |

1. **ggplot**

**문제 2번**

midwest(ggplot2 데이터, 미국 지역별 인구통계정보 자료)에서 전체 인구와 아시아인 인구 간의 관계를 파악하고자 한다.

**2.1** x축에 전체 인구(poptotal), y축에 아시아인 인구(popasian)로 된 산점도를 그려라( geom\_point() 이용).

|  |
| --- |
|  |

**문제 3번**

**3.1** 아래 그림과 같은 형태의 data frame을 생성하라. 변수이름은 fl, fuel\_kind, price\_fl로 하여라.

|  |  |  |
| --- | --- | --- |
| **fl** | **연료종류** | **가격)갤런당 USD)** |
| c | CNG | 2.35 |
| d | disel | 2.38 |
| e | ethanol E85 | 2.11 |
| p | premium | 2.76 |
| r | regular | 2.22 |

|  |
| --- |
|  |

**3.2** 위에서 구성한 data frame을 이용하여 mpg에 price\_fl(연료가격) 변수를 추가한 p\_mpg dataframe을 구성하여라.

(hint: merge(dataframe\_1, dataframe\_2, by =”variable”)

여기서 by 라는 조건은 어떤 변수를 기준으로 합칠것인지를 의미한다)

참고

|  |  |
| --- | --- |
| price\_fl(연료가격) 변수를 추가하기 전 | price\_fl(연료가격) 변수를 추가 한 후 |
|  |  |

|  |
| --- |
|  |