

Project_1

Choi Pro

Table of contents

1 난이도 하	1
2 난이도 중	2
3 난이도 상	3



아래는 R을 이용한 데이터 전처리 예제입니다. [문제와 답](#)만 있습니다. 풀이 과정을 코드로 제출하세요. 제출 결과에 따라 다음 과정의 Project 2 난이도 조정이 됩니다.

defect 데이터 셋 다운로드

5월부터 9월까지 일별 생산 물량이 10000개에서 나오는 불량데이터이다.

세로선불량(defect_a), 가로선불량(defect_b), 명점불량(defect_c), 암점불량(defect_d)

```
library(tidyverse)

df <- read.csv("../data/defect.csv")
head(df)
```

	defect_a	defect_b	defect_c	defect_d	Month	Day
1	41	190	7	67	5	1
2	36	118	8	72	5	2
3	12	149	13	74	5	3
4	18	313	12	62	5	4
5	NA	NA	14	56	5	5
6	28	NA	15	66	5	6

1 난이도 하

1.1.a.a

1. 열별 결측치가 몇 개가 있는지 표시하라.

```
defect_a defect_b defect_c defect_d    Month    Day
      37        7        0        0        0        0
```

1.1.a.b

2. 월별 **defect_a**의 평균과 **defect_c**의 표준편차를 구하시오. (소수 첫째자리까지)

```
# A tibble: 5 × 3
  Month defect_a_평균 defect_c_표준편차
  <int>      <dbl>          <dbl>
1     5      23.6           3.5
2     6      29.4           3.8
3     7      59.1           3.1
4     8      60            3.4
5     9      31.4           3.5
```

1.1.a.c

3. 명점불량과 암점 불량을 더해서 픽셀불량(**pixel_defect**)이라는 새로운 열로 만들어라. 이 때 픽셀불량이 가장 많이 나온 날은 몇월 몇일, 몇개가 나왔는지 표시하라.

pixel_defect = defect_c + defect_d

```
Month Day pixel_defect
1     8  28           107
```

```
Month Day pixel_defect
1     8  28           107
```

2 난이도 중

2.1.a.a

4. **defect_b**이 150 이상인 날 중에 8월~9월 총 몇일이나 되는가

```
n
1 38
```

2.1.a.b

5. **defect_a** 가 결측치가 있는 날 중에 월별 **defect_c**의 중간값을 구하시오.(**defect_c** 중간값)

```
# A tibble: 5 × 2
  Month defect_c_중간값
```

	<int>	<int>
1	5	14
2	6	9
3	7	11
4	8	12
5	9	13

2.1.a.c

6. 월별 defect 종류별 평균 갯수를 구하시오 (정수로 표시하시오)

```
# A tibble: 5 × 5
  Month defect_a_평균갯수 defect_b_평균갯수 defect_c_평균갯수 defect_d_평균갯수
  <int>          <dbl>          <dbl>          <dbl>          <dbl>
1     5             24             24             24             24
2     6             29             29             29             29
3     7             59             59             59             59
4     8             60             60             60             60
5     9             31             31             31             31
```

3 난이도 상

3.1.a.a

7. 일별 제품 1개당 나오는 총 불량율을 계산하여 월별 최대값을 구하라. (NA는 0으로 계산하라.)

total_defect = (defect_a+defect_b+defect_c+defect_d)/10000

```
# A tibble: 5 × 2
  Month max_defect
  <int>      <dbl>
1     5    0.0432
2     6    0.0466
3     7    0.0492
4     8    0.049
5     9    0.0378
```

3.1.a.b

8. 총 불량갯수가 최소인 날은 몇월 몇일인가? (NA는 0으로 계산하라.)

total_defect = (defect_a+defect_b+defect_c+defect_d)/10000

```
Month Day
1     7  1
```

3.1.a.c

9. 선불량 (가로선+세로선)이 **pixel** 불량(명점+암점) 보다 많은 날은 몇일 인가?(NA는 0으로 계산하라.)

```
n
1 126
```