



# 식습관 스몰 데이터 분석을 통한 장트러블 극복기

팀 푸푸

고은솔, 김민주, 배예슬, 송혜정, 이선영, 정보경

# | 목차

I. 프로젝트 소개

II. 데이터 소개

III. 분석 방법 / 결과

IV. 가설 검증 / 예측

# | 가설



음식



쾌변

# | 가설

Spicy



Cold



Coffee



Alcohol



Milk



## | 퀴즈



하루 동안 화장실에서 보낸 시간이 가장 긴 사람은 누구일까요?

## | 데이터 수집

- 수집 기간: 8/2 ~ 10/2 (62일)
- 데이터 개수: 총 1316개

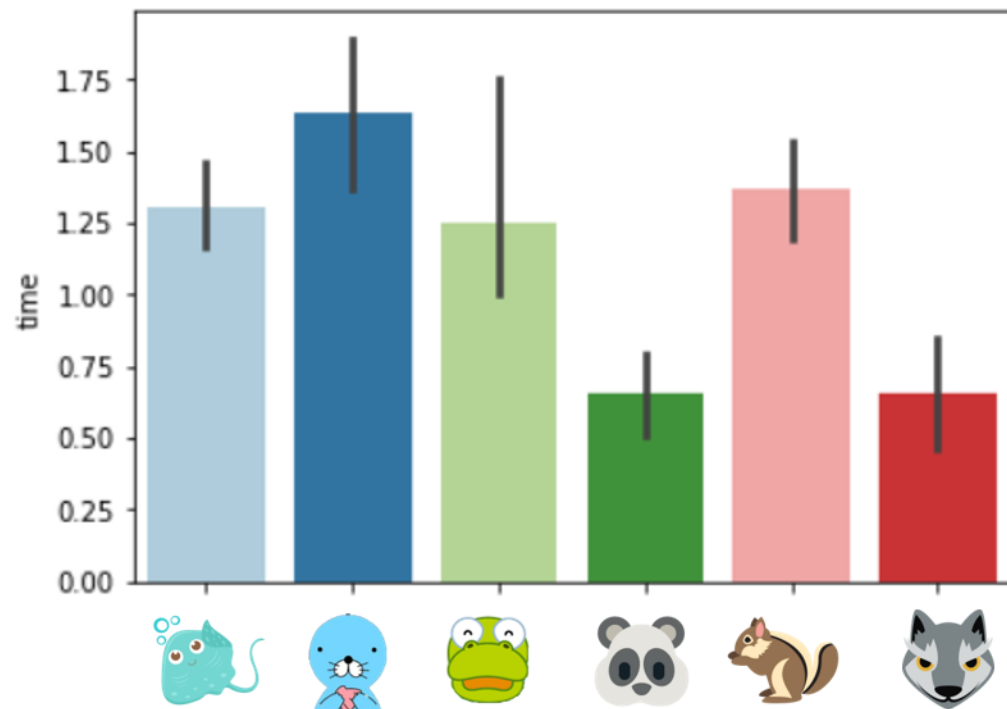
| 닉네임 | 날짜         | 시간    | 식사 메뉴                | 맵기 | 매운거 | 차가운거 | 커피 | 탄산 | 술       | 유제품, 두유    | 물   | 채소, 과일 | 유산균 | 특이사항 |
|-----|------------|-------|----------------------|----|-----|------|----|----|---------|------------|-----|--------|-----|------|
| 가오리 | 2019-08-02 | 22:00 | 오징어, 수박, 새우깡, 맥주 두모금 | 0  |     | -    |    | -  | 맥주 20ml | -          | 100 | 수박     | 0   |      |
| 가오리 | 2019-08-02 | 19:30 | 만두국, 김치참치치즈밥         | 2  |     | -    |    | -  |         | 슬라이스치즈 1/2 | 300 | -      | 0   |      |
| 가오리 | 2019-08-03 | 12:00 | 달강정                  | 0  |     | -    |    | -  |         | -          | 500 | -      | 0   |      |
| 가오리 | 2019-08-03 | 15:00 | 쌀햇도그                 |    |     | -    |    | -  |         | -          | 200 | -      | 0   |      |
| 가오리 | 2019-08-03 | 20:30 | 생선구이                 | 0  |     | -    |    | -  |         | -          | 300 | 오이 콩나물 | 0   |      |
| 가오리 | 2019-08-03 | 9:30  | 만두국, 김치참치치즈밥         | 2  |     | -    |    | -  |         | 슬라이스치즈 1/2 | 300 | -      | 1   |      |
| 가오리 | 2019-08-04 | 11:00 | 누룽지                  | 0  |     | -    |    | -  |         | -          | 300 |        | 0   |      |

식사 관련 데이터 | 940개

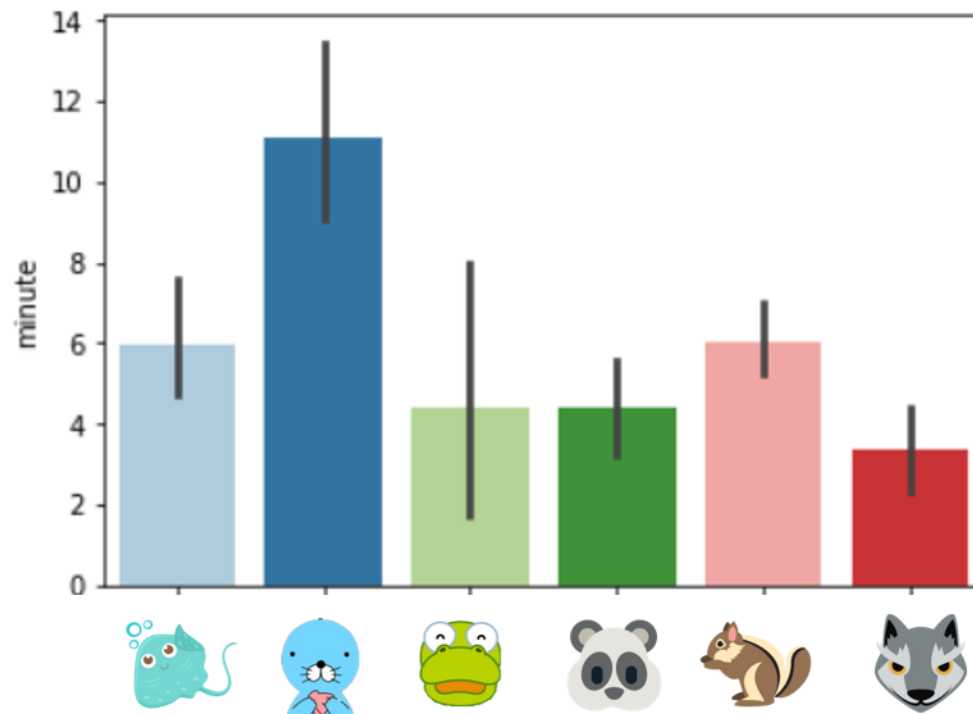
| 닉네임 | 날짜         | 화장실 간 시각 | 걸린 시간 | 쾌변 / 애매 / 시도 | 쾌변 여부                 | 쾌변 패턴 | 생리 |
|-----|------------|----------|-------|--------------|-----------------------|-------|----|
| 가오리 | 2019-08-03 | 9:00     | 5     | 애매           | 아주 약간 배 아팠는데 아주 약간 찜찜 |       |    |
| 가오리 | 2019-08-03 | 9:50     | 3     | 애매           | 또 조금 배 아프고 여전히 찜찜     |       |    |
| 가오리 | 2019-08-03 | 19:00    | 10    | 애매           | 배아픔...약한🔥             |       |    |
| 가오리 | 2019-08-04 | 10:30    | 10    | 애매           | 아주 약간 배 아팠는데 아주 약간 찜찜 |       |    |
| 가오리 | 2019-08-05 | 10:00    | 2     | 쾌변           | ○                     |       |    |
| 가오리 | 2019-08-06 | 8:30     | 2     | 쾌변           | ○                     |       |    |
| 가오리 | 2019-08-06 | 9:00     | 5     | 쾌변           | ○                     |       |    |

쾌변 관련 데이터 | 376개

# | Barplot

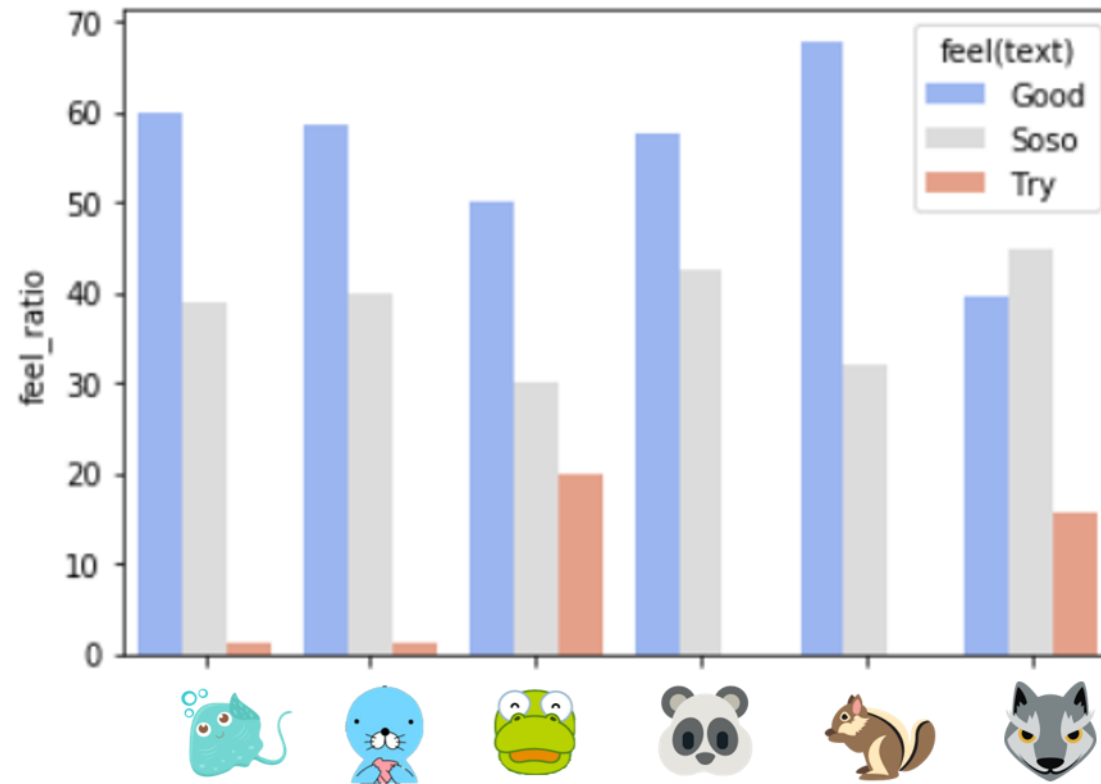


일 평균 화장실 간 횟수



일 평균 화장실에서 보낸 시간

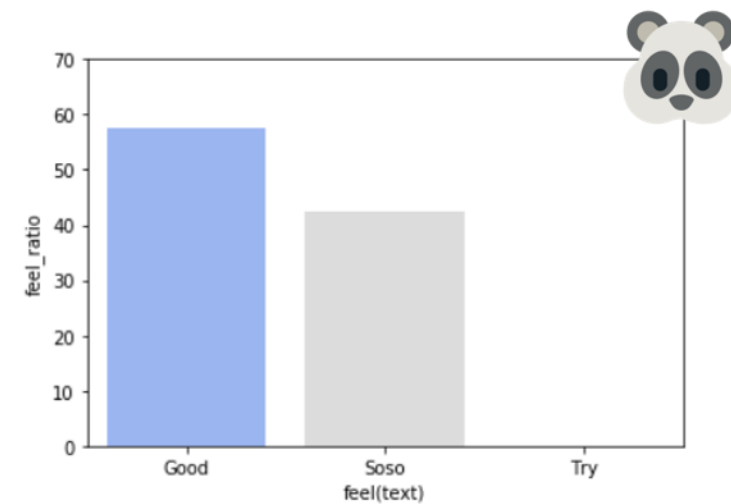
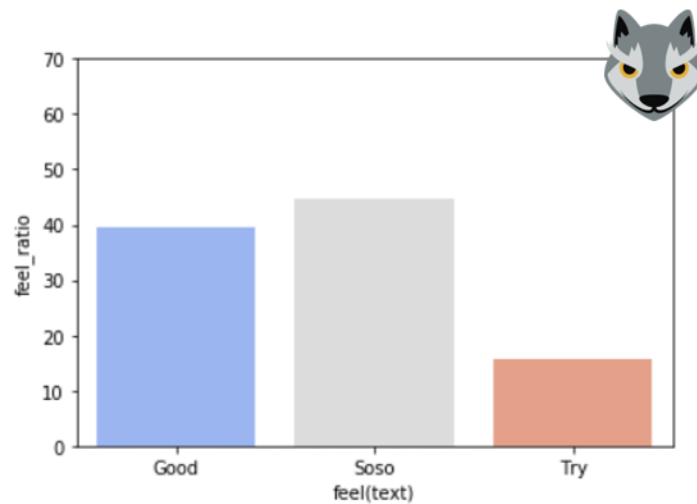
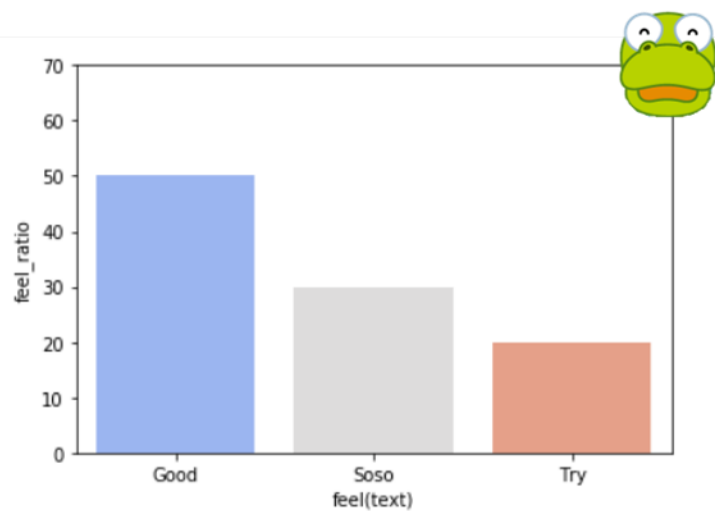
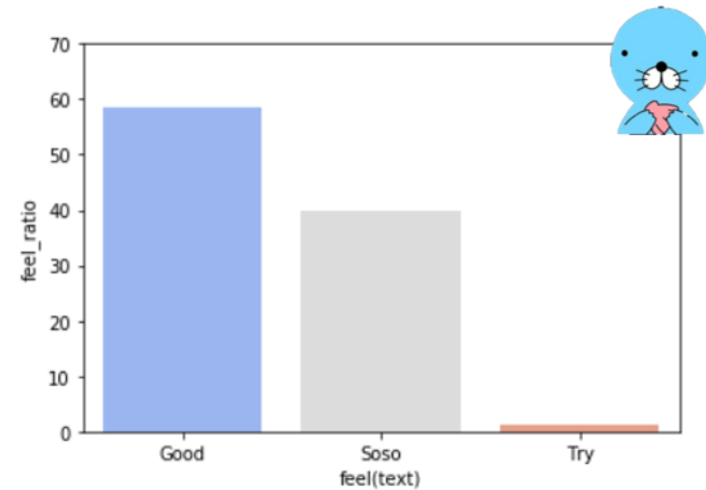
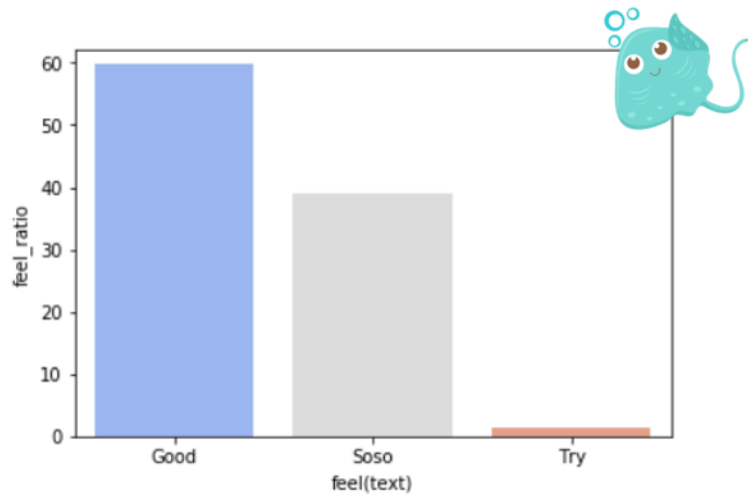
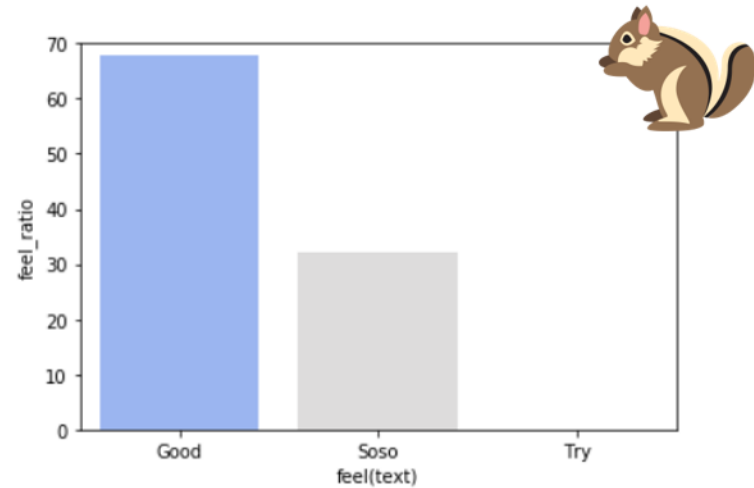
## | Barplot



개인별 쾌변 만족 비율

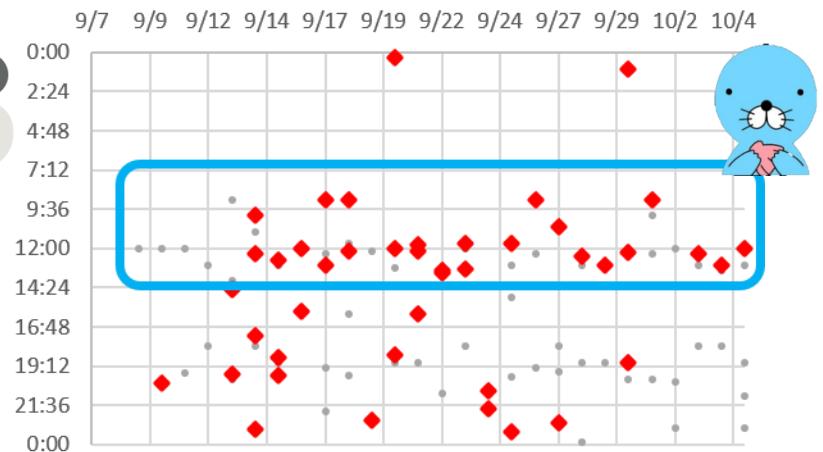
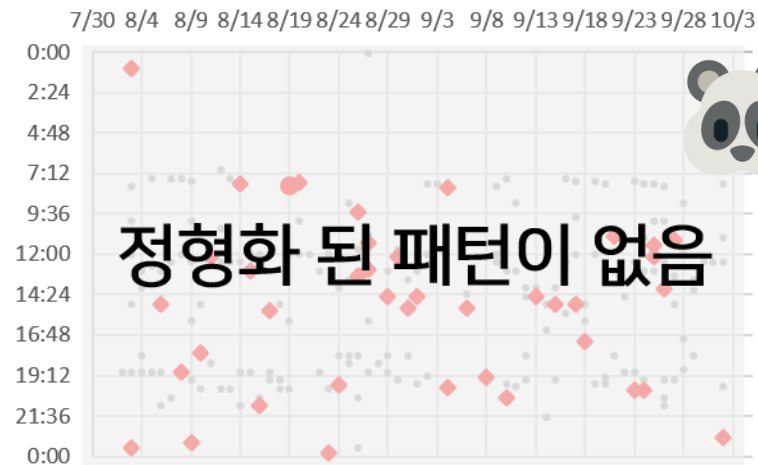
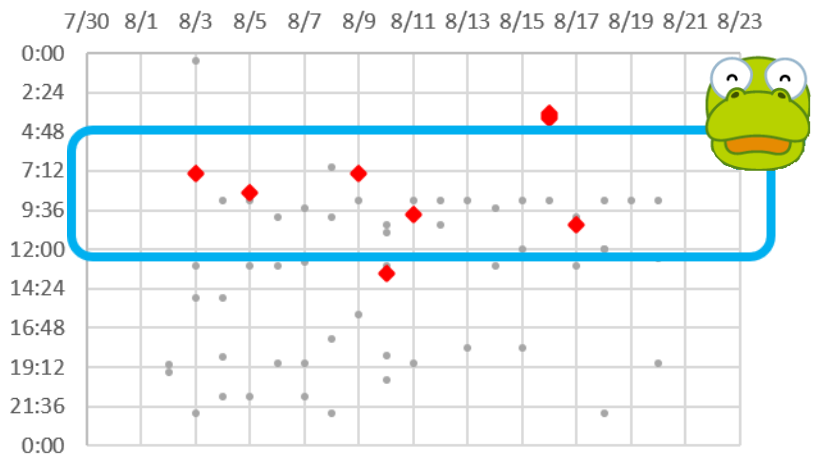
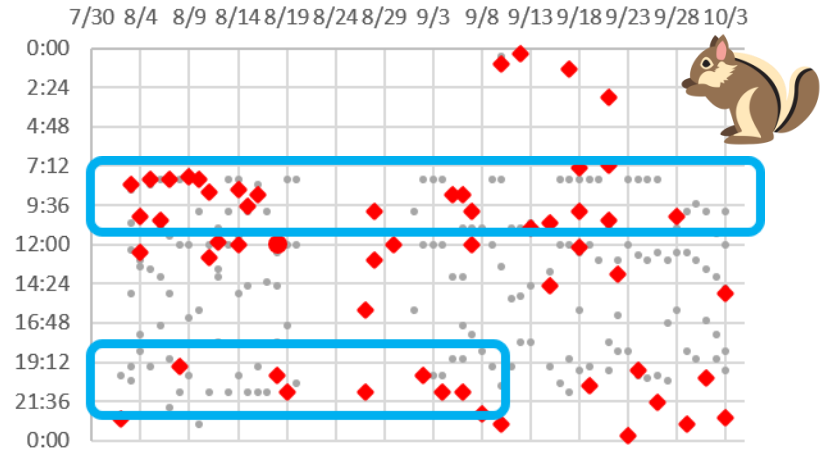
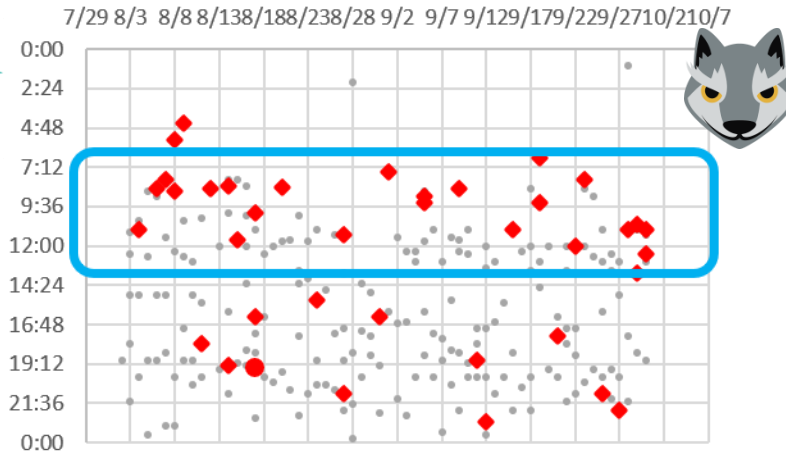
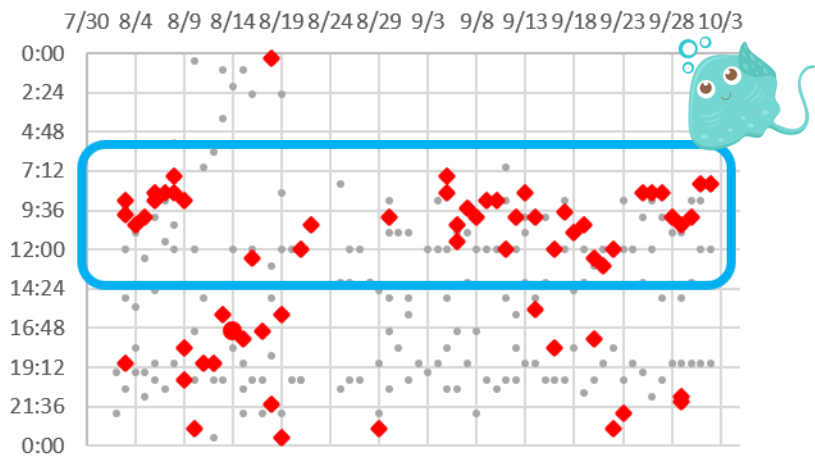


# | Barplot

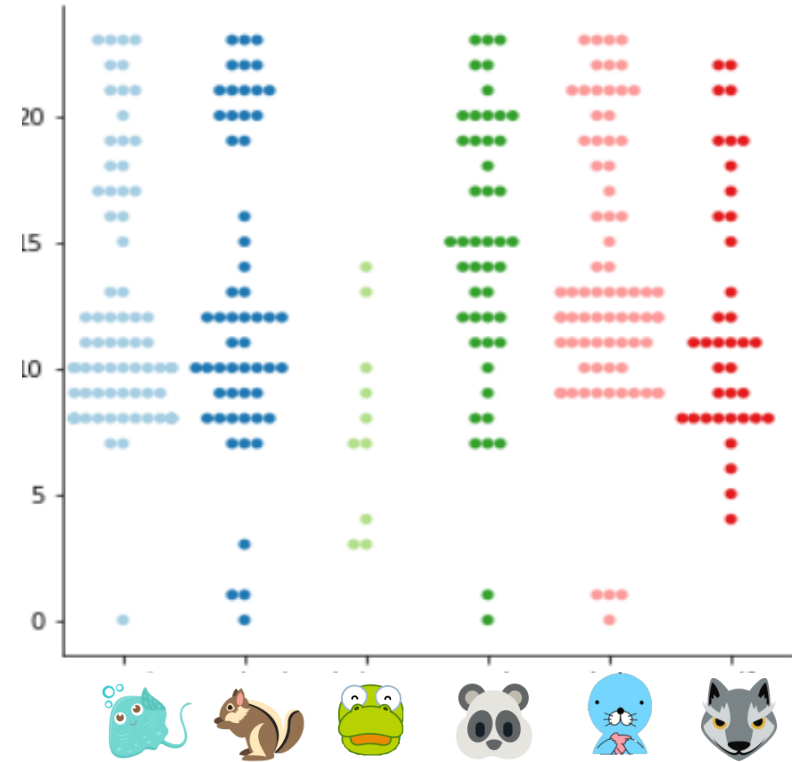
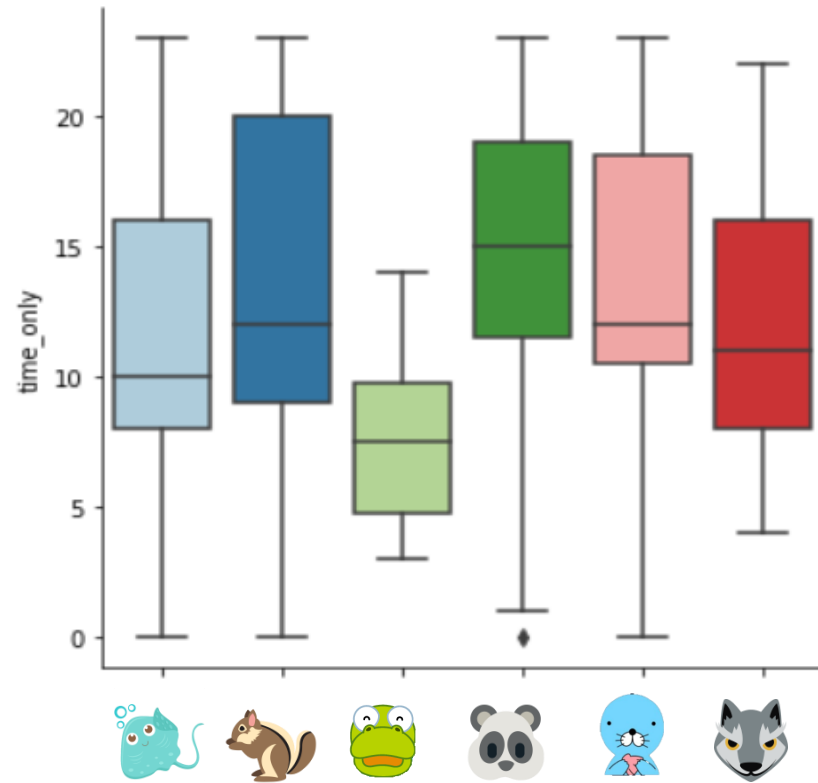


개인별 캐변 만족도 비율

# | Scatterplot

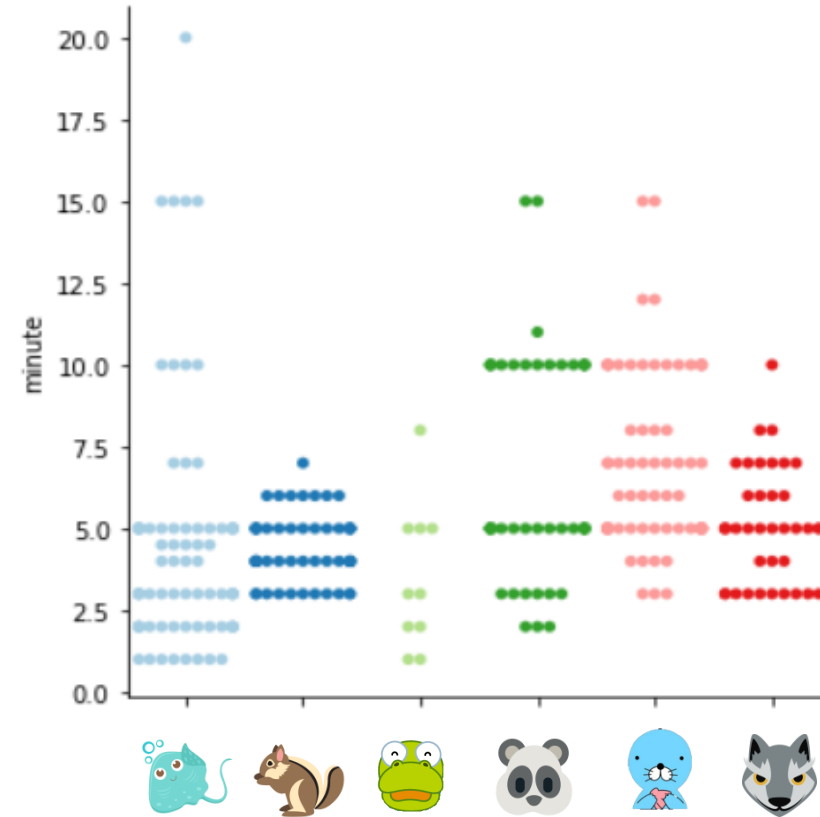
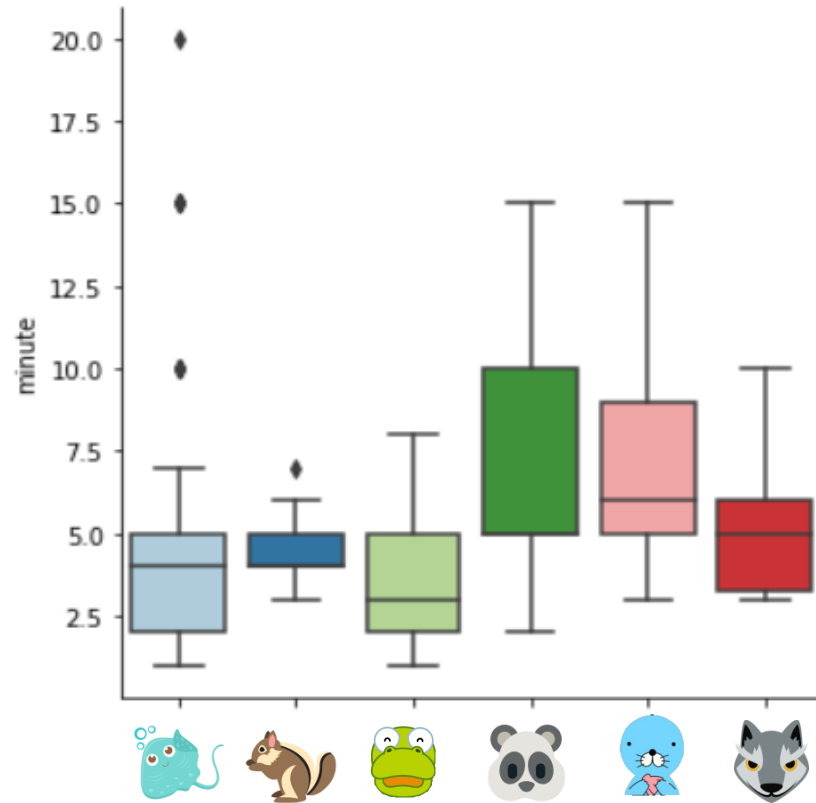


# | Boxplot



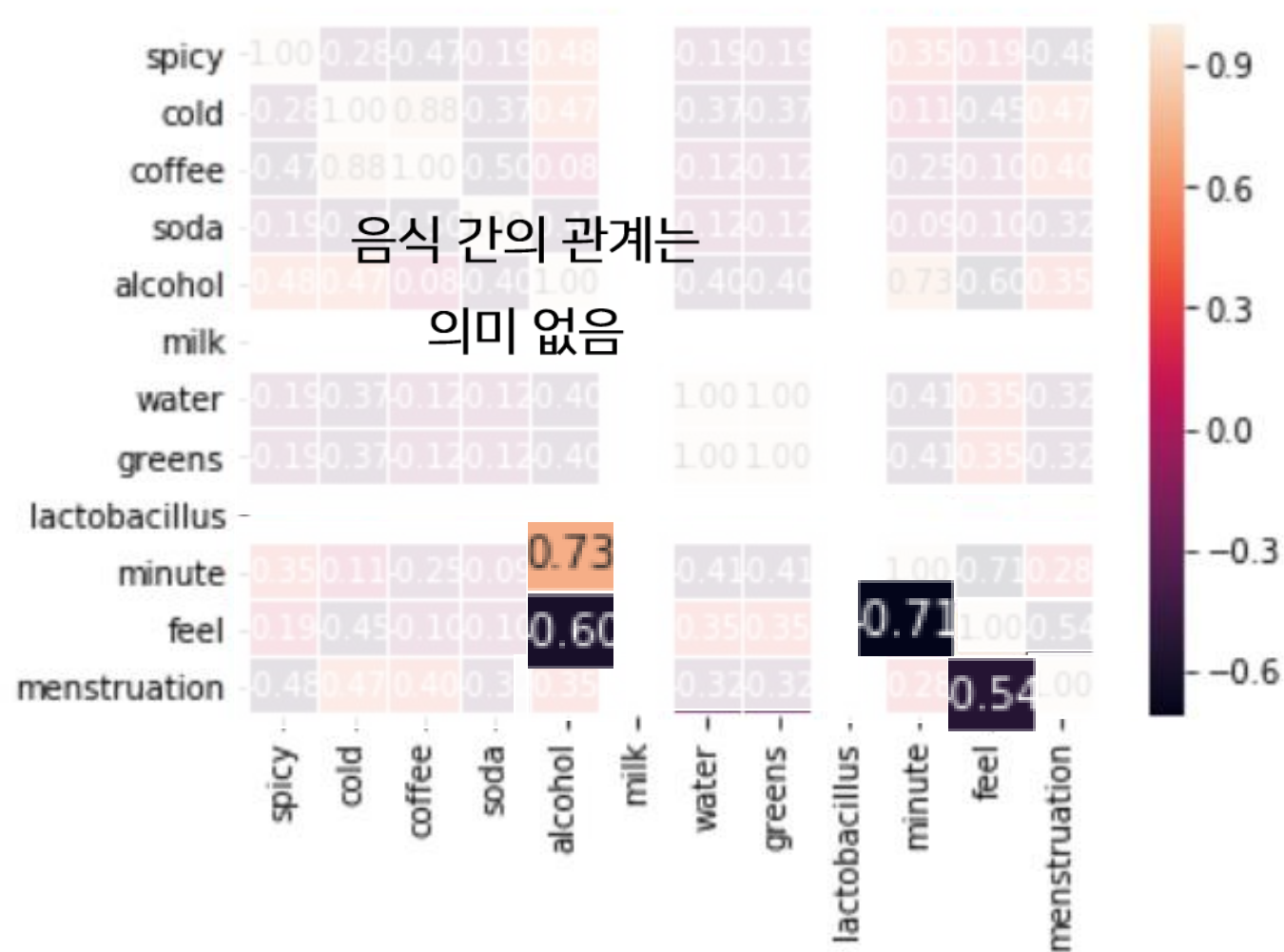
화장실 간 시간대

# | Boxplot

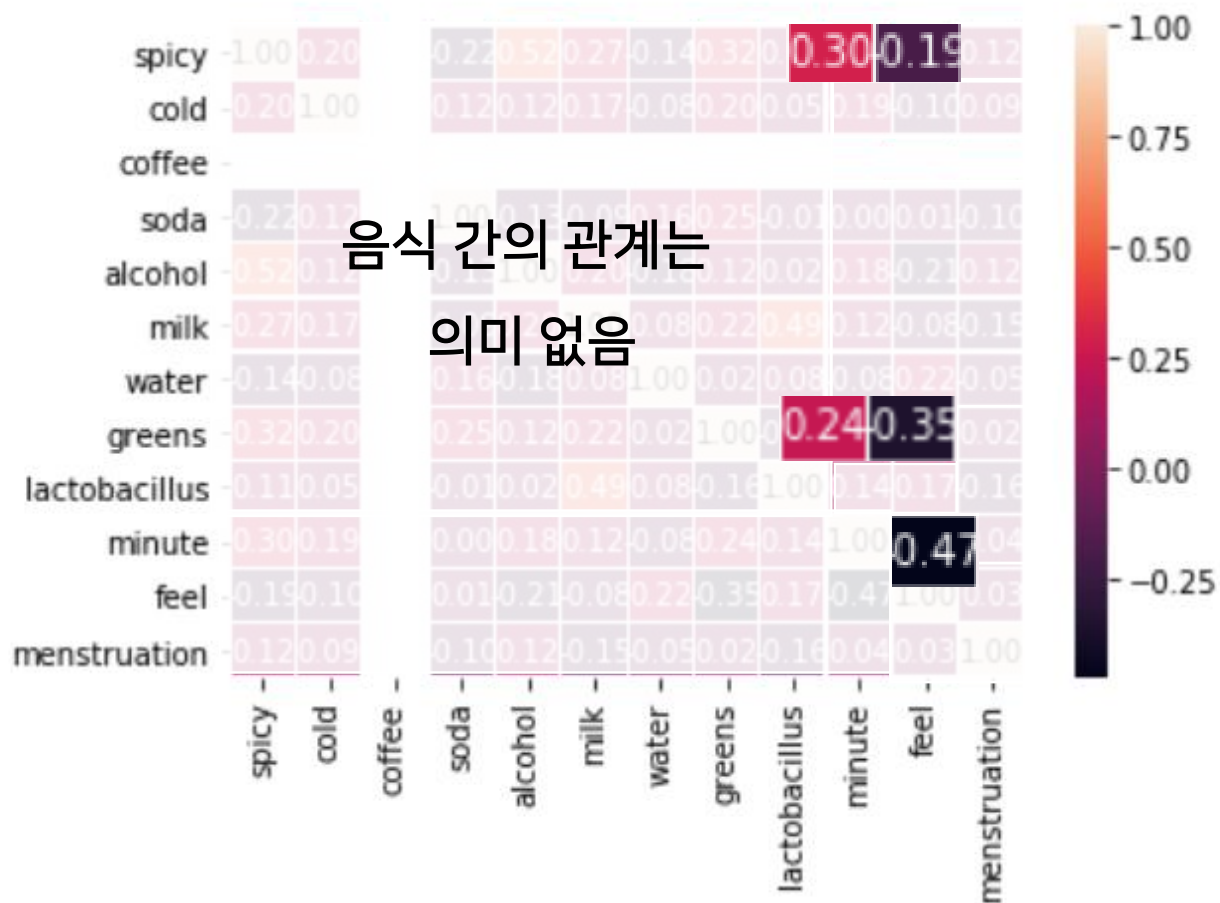


화장실에서 보낸 시간

# | Heatmap 🐸



# | Heatmap & Linear Regression



## OLS Regression Results

|                   |                       |                     |                          |
|-------------------|-----------------------|---------------------|--------------------------|
| Dep. Variable:    | feel                  | R-squared:          | 0.224                    |
| Model:            | Adj. R-squared: 0.205 |                     |                          |
| Method:           |                       |                     |                          |
| Date:             | Fri, 18 Oct 2019      | Prob (F-statistic): | 3.05e-05                 |
| Time:             | 11:22:39              | Log-Likelihood:     | -52.241                  |
| No. Observations: | 85                    | AIC:                | 110.5                    |
| Df Residuals:     | 82                    | BIC:                | 117.8                    |
| Df Model:         | 2                     |                     |                          |
| Covariance Type:  | nonrobust             |                     |                          |
|                   | coef                  | std err             | t P> t  [0.025 0.975]    |
| Intercept         | 2.8841                | 0.083               | 34.730 0.000 2.719 3.049 |
| minute            | -0.0683               | 0.014               | -0.096 -0.040            |
| menstruation      | 0.0674                | 0.137               |                          |

# | Backward Elimination



9개의 변수



후방 제거 방법 선택

# | Backward Elimination

```
pm1 = ['spicy', 'cold', 'soda', 'alcohol', 'milk', 'water', 'greens', 'lactobacillus']  
backward_minute(pm1, fs_1d[fs_1d['id'] == 'gaori'])
```

('spicy+soda+alcohol+milk+water+greens+lactobacillus', 0.0630560144737009)

```
pm2 = ['spicy', 'soda', 'alcohol', 'milk', 'water', 'greens', 'lactobacillus']  
backward_minute(pm2, fs_1d[fs_1d['id'] == 'gaori'])
```

('spicy+soda+milk+water+greens+lactobacillus', 0.07286882643059989)

```
pm3 = ['spicy', 'soda', 'milk', 'water', 'greens', 'lactobacillus']  
backward_minute(pm3, fs_1d[fs_1d['id'] == 'gaori'])
```

('spicy+soda+water+greens+lactobacillus', 0.0813544221919793)

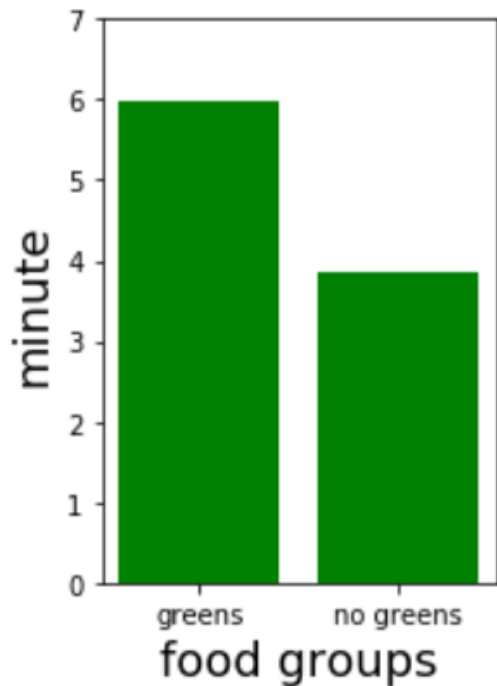
```
pm4 = ['spicy', 'soda', 'water', 'greens', 'lactobacillus']  
backward_minute(pm4, fs_1d[fs_1d['id'] == 'gaori'])
```

('spicy+water+greens+lactobacillus', 0.09271542753997108)

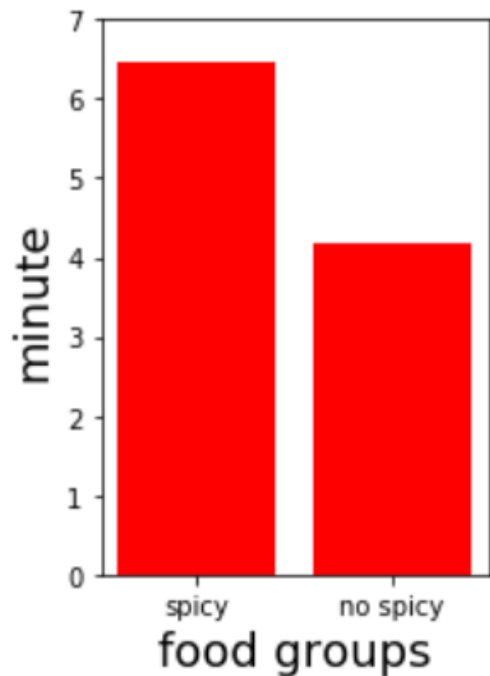


# | 음식 섭취 여부에 따른 결과 비교

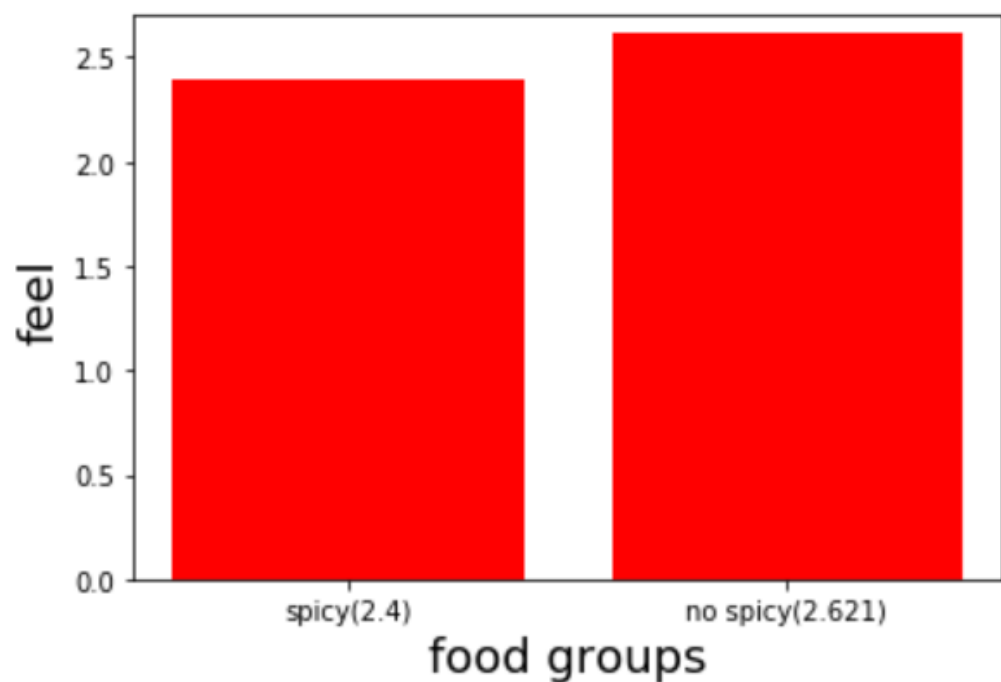
채소



매운 음식



매운 음식



걸린 시간

만족도

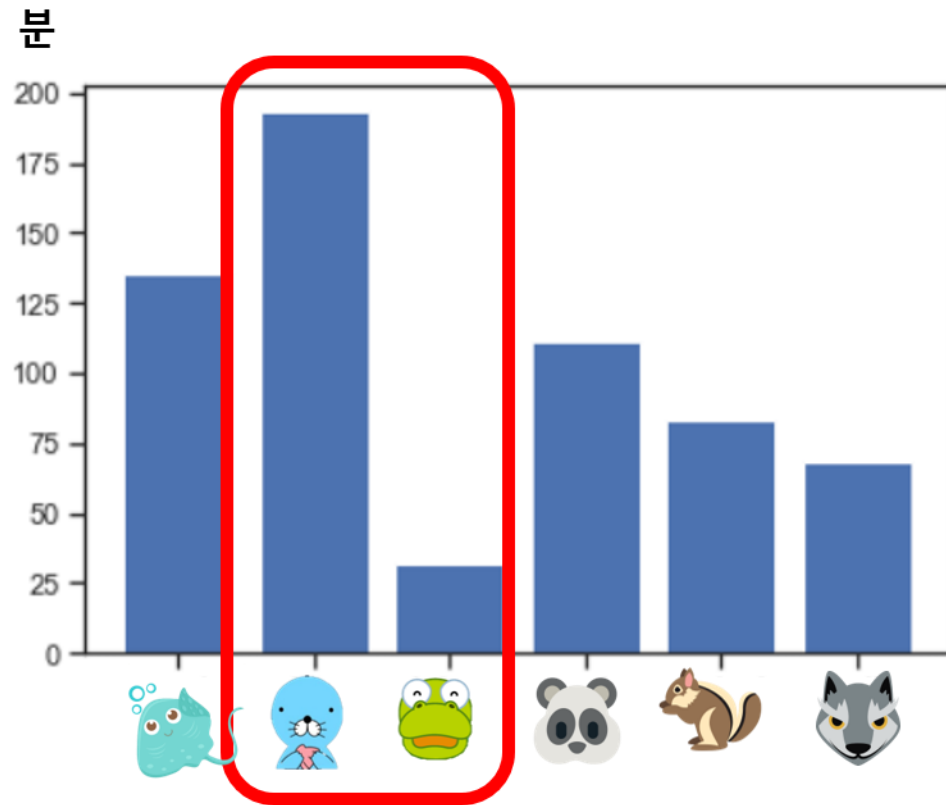
앞으로 50년 동안 얼마의 시간을  
화장실에서 보내게 될까?

## | 퀴즈



화장실에서 보낼 시간이 가장 많은 사람과 가장 적은 사람의 차이는  
**며칠**일까요?

# | 앞으로 50년 동안 화장실에서 보낼 시간



화장실에서 보낸 시간의 총합



$$10.7 \text{ 분 / 일} \\ 10.7 * 365 * 50$$

= 약 135일



$$1.25 \text{ 분 / 일} \\ 1.25 * 365 * 50$$

= 약 15일

**120일 차이 발생**

Q & A