

Fitbit Data

Youngseok Lee
yslee@cs-cnu.org

Fitbit Data

- 1일 데이터

- 기간: ~ 10월 말 (예정),
- 1일 단위의 합
- 항목
 - 걸음수: steps, distance, calories,
 - 활동정도분류: minutesSedantary, minutesLightlyActive, minutesFairlyActive, minutesVeryActive,
 - 수면: timeInBed, minutesAsleep, minutesAwake, awakeningsCount
 - 몸무게: weight, bmi, fat

- Intraday 데이터 (분단위)

- 5분 단위 heart rate
- 15분 단위 steps

데이터 수집 방법

- Python script
 - <https://github.com/orcasgit/python-fitbit>
- Google script
 - <https://github.com/simonbromberg/googlefitbit>
- R script
 - fitbitScraper: <http://blog.corynissen.com/2015/01/r-package-to-download-fitbit-data.html>
 - <https://github.com/corynissen/fitbitScraper>
- 기타
 - Java, ...
 - Fitbit API 활용이면 OK



```
import fitbit
import ConfigParser
import json
```

```
#Load Settings
parser = ConfigParser.SafeConfigParser()
parser.read('config.ini')
```

```
CI_id          = parser.get('Login Parameters', 'C_KEY')
CI_client_secret = parser.get('Login Parameters', 'C_SECRET')
CI_access_token  = parser.get('Login Parameters', 'ACCESS_TOKEN')
CI_refresh_token = parser.get('Login Parameters', 'REFRESH_TOKEN')
```

```
authd_client = fitbit.Fitbit(CI_id, CI_client_secret, oauth2=True, access_token=CI_access_token,
refresh_token=CI_refresh_token)
```

```
intradayS = authd_client.intraday_time_series('activities/steps', base_date='today',detail_level='15min')
f = open('datadumpSteps.json', 'w')
json.dump(intradayS, f)
```

```
intradayH = authd_client.intraday_time_series('activities/heart', base_date    = 'today', detail_level = '15min')
f = open('datadumpHeart.json', 'w')
json.dump(intradayH, f)
```



```
install.packages("fitbitScraper")  
library("fitbitScraper")
```

```
cookie <- login(email="corynissen@gmail.com", password="mypassword")  
# 15_min_data "what" options: "steps", "distance", "floors", "active-minutes", "calories-burned"  
df <- get_15_min_data(cookie, what="steps", date="2015-01-21")  
library("ggplot2")  
ggplot(df) + geom_bar(aes(x=time, y=data, fill=data), stat="identity") +  
  xlab("") + ylab("steps") +  
  theme(axis.ticks.x=element_blank(),  
        panel.grid.major.x = element_blank(),  
        panel.grid.minor.x = element_blank(),  
        panel.grid.minor.y = element_blank(),  
        panel.background=element_blank(),  
        panel.grid.major.y=element_line(colour="gray", size=.1),  
        legend.position="none")
```

```
# daily_data "what" options: "steps", "distance", "floors", "minutesVery", "caloriesBurnedVsIntake"  
df <- get_daily_data(cookie, what="steps", start_date="2015-01-13", end_date="2015-01-20")  
ggplot(df) + geom_point(aes(x=time, y=data))
```

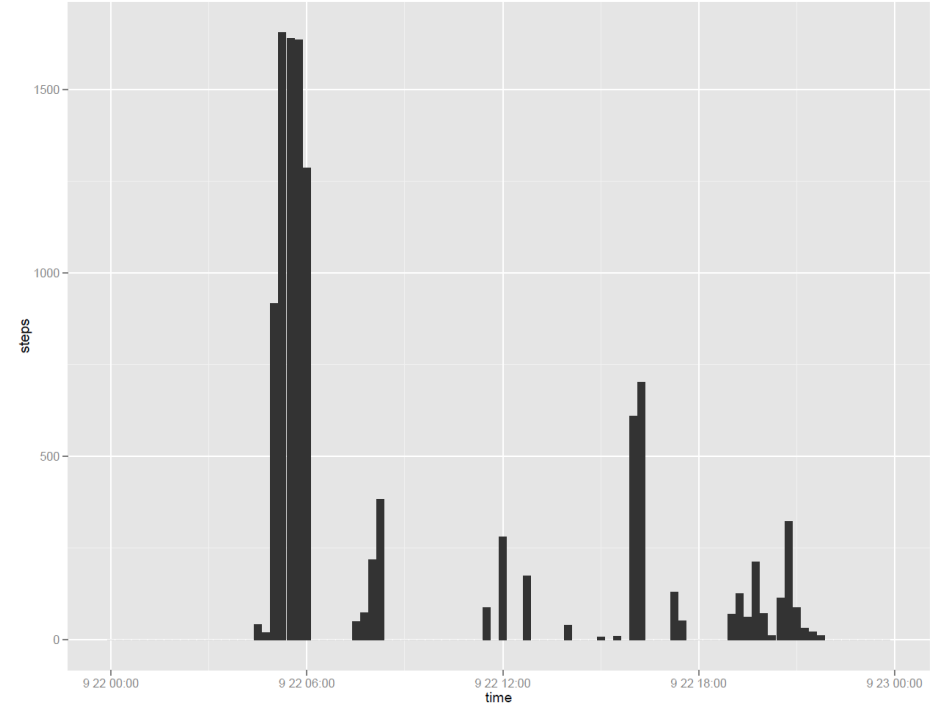
스마트와치 데이터

- Fitbit과 동일
 - Daily data
 - Intraday data
- Google fit developer
 - <https://developers.google.com/fit/>
- Apple Healthkit
 - <https://developer.apple.com/healthkit/>
- Samsung sHealth
 - <http://developer.samsung.com/health>

채점기준

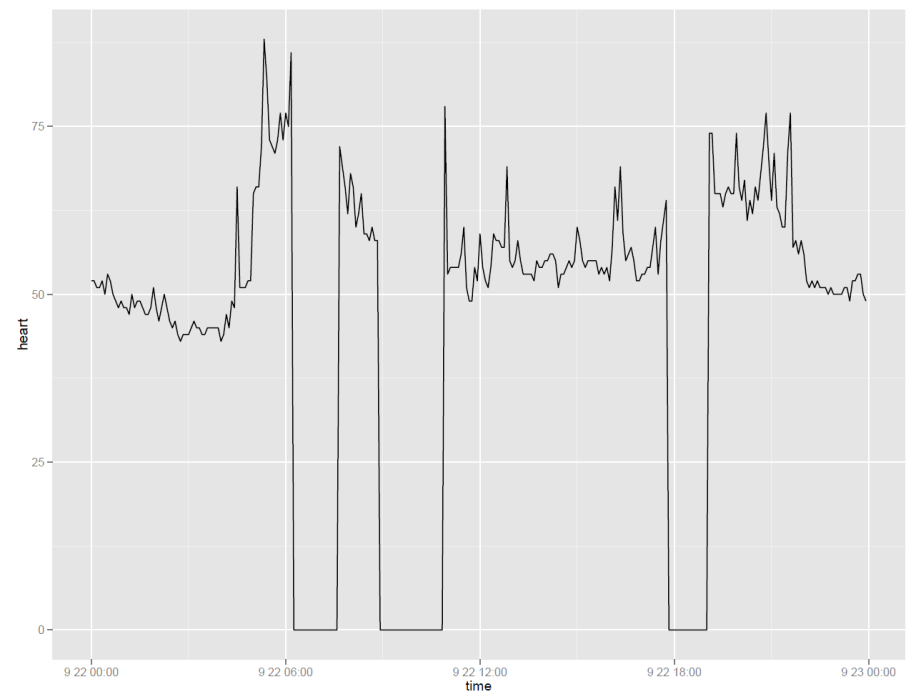
- 데이터무결성
 - 데이터의 오류가 없어야함: 예) 미착용
- 많은 양의 데이터
 - 1일 데이터: 기록된 날짜
 - 최소 1개월 이상
 - Intraday 데이터: 분단위 기록 건수
 - 최소 1개월 이상
 - 많은 필드들의 기록들
 - 예) 몸무게관련

Steps



2015.9.22
Steps / 15 min
Heart rate / 5 min

Heart Rate



제출방법

- Csv file
 - 클래스룸 제출
 - Google spreadsheet 조교 공유
- Source code
 - Googlescript
 - Python/...
 - R

예고

- 30명의 데이터를 이용하여 무엇을 할 것인가?
 - 파일 저장
 - Local filesystem, database, NoSQL, Hadoop filesystem
 - Classification
 - 생활패턴: Early bird ? Night owl ?
 - 건강척도: healthy ? Unhealthy ?
 - Clustering
 - K means clustering
 - Identification
 - feature vector 를 이용하여 개인을 식별할 수 있을까?
 - Recommendation
 - 어떤 추천 서비스가 사람들에게 유익할까?