



# 웨어러블 기기를 통한 삼성헬스 데이터 시각화

## 한국폴리텍대학 스마트 금융과

### 01. 프로젝트 개요

#### 개발 배경 및 목적

- 웨어러블 기기와 삼성헬스 앱에서 추출한 사용자의 데이터를 분석해 건강정보를 시각화
- 수집한 사용자의 데이터를 분석하여 사용자의 건강상태 및 활동정보 등을 알림

#### 프로젝트 특징

- 1) 웨어러블 기기에서 추출된 사용자 실데이터 활용 및 재정의(삼성헬스 데이터)
- 2) 분석에 필요한 데이터 시각화(심박수, 걸음 수, 스트레스 등)

### 02. 개발 환경 및 기술

#### 개발 환경

- 운영 체제: Windows 10
- 개발 언어: Python 3.7  
Java 1.8  
MariaDB 10.4
- 개발 도구: Jupyter lab,  
eGovFrameWork 3.9  
AWS  
Docker

#### 개발 흐름도

- SamSung Health Data  
- 사용자의 삼성헬스 데이터 매일 추출



- Data ETL  
- Pandas 라이브러리 활용 데이터 정제 및 DB적재



- Web 구축 및 데이터 시각화  
- 전자정부표준프레임워크와 부트스트랩, 구글차트 API를 통해 데이터를 시각화



- AWS & Docker  
- AWS와 Docker를 활용하여 소스코드 및 구축된 사이트 운영서버 가동



### 03. 결과물



#### 삼성헬스 데이터 ETL

- 차후 분석을 위해 헬스 데이터 수집 및 DB 적재  
- 사용자 별로 다르게 추출된 데이터를 ETL을 통해 일관된 데이터로 정제하여 통합 후 DB 적재

```
# time 데이터 형식 (2020, 07, 14 형식) timestamp 형식을 2020-07-14 형식으로 통일)
def dateToDatetime(date):
    date = str(date)
    result = None
    if re.search("\d{4}, \d{2}, \d{2}", date) is not None: # 시간 형식이 2020, 07, 08, ~인 경우에 대해서만
        splitDate = [each.strip() for each in date.split(',')]
        splitDateLen = len(splitDate)

        yearMonthDay = "-".join(splitDate[:splitDateLen - 1])
        hourMinSec = [each for each in splitDate[-1].split(':')]

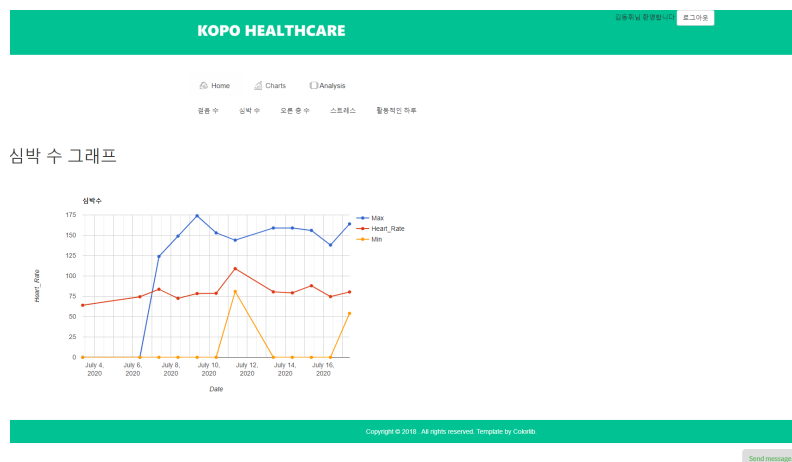
        if "오전" in hourMinSec[0]:
            hourMinSec[0] = str(int(hourMinSec[0].split("-")[-1]) + 12)
        elif ("오전" or "오전 12") in hourMinSec[0]:
            hourMinSec[0] = str(int(hourMinSec[0].split("-")[-1]))

        hourMinSec = ":".join(hourMinSec)
        result = yearMonthDay + " " + hourMinSec
    elif re.search("\d{4}-\d{2}-\d{2}", date) is not None:
        #splitDate = [each.strip() for each in date.split('-')]
        #splitDateLen = len(splitDate)
        #if splitDateLen == 2:
        #    result = str(date).split('.')[0]
        #else:
        #    #result = date
        #    #result = date
        elif len(str(date)) <= 8 and re.search("UTC", date) is None: # time_offset / longest_idle_time(-1) 일때
            if not math.isnan(float(date)):
                result = convertMilli(int(float(date)))
            elif len(str(date)) == 13: # 13 digits unixtime / day_time / set_time / start_date / date / update_time
                result = datetime.fromtimestamp(int(date)/1000).strftime("%Y-%m-%d %H:%M:%S")
            else: # UTC+0900, time_offset
                result = date
    return result
```



#### 웹 페이지 시각화

- 정제된 실데이터를 구축한 웹 페이지 시각화  
- DB에 저장된 데이터를 구글 차트 API를 활용하여 웹페이지에 사용자 별 시각화



### 04. 의의 및 기대효과



#### 의의

- 삼성헬스를 통해 수집된 사용자의 데이터를 그래프로 확인하여 사용자의 건강 및 심박수, 스트레스를 일자 별 제공



#### 기대효과

- 차후 분석과 함께 사용자의 건강상태를 실시간으로 감시하여 건강상태 및 응급상황을 알림
- 코로나 여파로 인한 웨어러블 기기 및 앱을 통해 회사 및 단체에서 실시간 구성원 상태 확인