**2022.08.14 일요일**

**DB table에서 query문 작성**

1. device id와 date time column을 이용하여, select문 작성하여 일자 별 data select

ex) device id = 51, date time = 2019.10.24 has 552 rows in set

the number of users \* the number of dates 만큼 query 작성.

* 위의 data들을 이용하여 python 코드를 돌려서 stay point 추출해야 함.

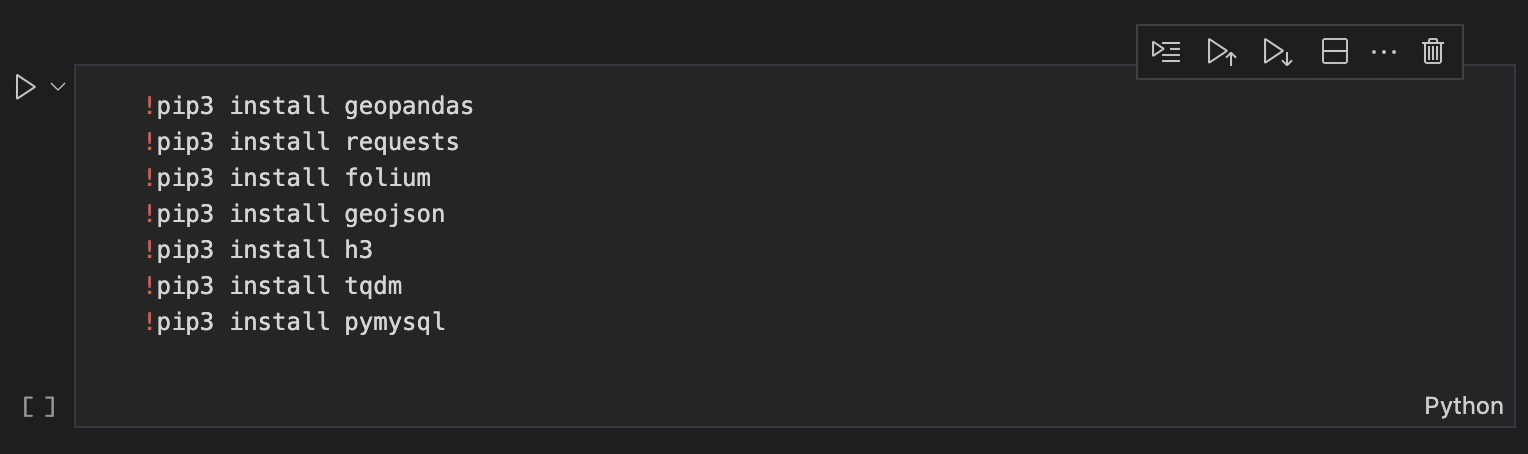
select 문으로 row data 찾고 나서, 해당 data들을 이용할 수 있는 방법이 무엇?

1. python과 MySQL database를 연동하는 방법 = pymysql library 사용

**새로운 staypoint.ipynb 파일 생성하여 코드 작성하니 skmob import 오류가 발생**

1. conda skmob environment를 activate해도 해당 오류가 해결되지 않음.

그러나 기존에 만들었었던 tutorial.ipynb에서는 코드가 잘 실행 됨. (dependency 오류인가?)



코드상에서 위의 library들을 모두 설치 함.

**query문을 작성하여 뽑아낸 data를 이용하여 stay point를 추출 함**

1. stay point를 추출하기 전 TrajDataFrame에 data를 넣고 folium interactive map에 plot

지도이(가) 표시된 사진

자동 생성된 설명

위의 그림은 10월 24일 51번 user의 trajectory data를 그린 것

**2022.08.15 월요일**

**user의 관심사를 추출하기 위해서, stay point들을 통해 추출한 GPS 좌표가 실제 지도상에 어느 위치에 있는지 확인해야 함.**

1. **관심사에 대한 특정**

stay point들을 추출 한 다음, stay point들을 clustering하여 군집들을 생성함.

생성된 군집들에 대한 빈도수를 통해 사용자들의 관심사를 특정할 수 있음.

각 cluster를 대표하는 stay point를 추출하는 것이 중요하다.

* 각 user마다 생성된 cluster의 개수는 서로 다를 것이다,.

직장을 다니는 경우에는 회사가 제 2의 stay point가 될 것이지만, 그렇지 않은 경우에는 다른 point가 선택될 것이기 때문임.

* **cluster 빈도 / 전체 stay point의 개수 비율**이 의미가 있다.

해당 비율이 가장 높을 것으로 예상되는 지점은 집일 것임.

* 일반적으로 사람들은 일주일 단위로 생활 패턴이 바뀐다.

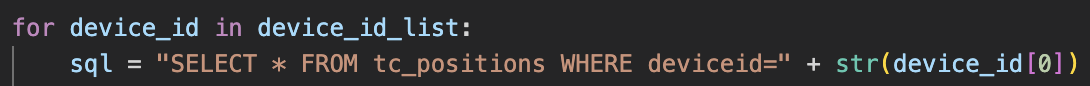
일주일 단위로 생성된 stay point를 분석해야 사람들의 관심사를 파악하기 쉬울 듯.

* 일주일 단위로 데이터를 뽑아올 수 있게 query를 잘 작성해야 함.

**2022.08.16 화요일**

**scikit mobility library를 이용하여 각 device id에 들어있는 trajectory data로부터 stay point를 추출하고, 일자별로 저장한다.**

1. table에서 device id 별로 trajectory data를 불러온다.



1. 불러온 trajectory data로부터 stay point를 추출한다.

텍스트이(가) 표시된 사진

자동 생성된 설명

stay point를 추출할 때 최적의 parameter 값은 경험적으로 구해야 함!!

1. 추출한 stay point를 2차원 dictionary에 저장한다.

첫 번째 index는 device id를 가리키고, 두 번째 index는 date time을 가리킨다.

* stdf 각 row에 접근하는 방법을 찾아야함

**2022.08.17 수요일**

**현재 2차원 dictionary 안에 trajectory data frame 형태로 저장되어 있어, 파일에 쓰는 것이 불가능**

1. trajectory data frame 자체를 csv 파일로 저장 하는 방법.

stay point의 date time을 확인하여 날짜가 같으면 하나의 csv 파일로 묶음.

* stay\_point\_dict에 device\_id를 key 값으로 갖는 dictionary 생성.
* 해당 dictionary안에 date time을 key 값으로 갖는 list 생성.
* list 안에 해당 날짜의 trajectory data frame 형식의 stay point 저장.

텍스트이(가) 표시된 사진

자동 생성된 설명

* trajectory data frame 형식의 각각의 row 들을 하나의 frame으로 만듦.

**2022.08.18 목요일**

**텍스트이(가) 표시된 사진

자동 생성된 설명**

**stay point를 csv 파일로 저장하면 위와 같이 저장되는 문제 발생. 띄어쓰기가 w 처리 되는 듯.**

* trajectory data frame을 list 형태로 변경하여 처리하면 됨.

**각 device id당 전체 stay point를 하나의 csv 파일로 저장하고, 일자 별 stay point를 별도의 csv 파일로 저장함.**

1. scikit mobility 라이브러리를 이용하여 구한 stay point를 list 형태로 변경하고, 하나의 csv 파일로 저장함.

텍스트이(가) 표시된 사진

자동 생성된 설명

만약 디렉토리 내 staypoint/{deviceid} 폴더가 없으면 해당 디렉토리를 만들어준 다음, 해당 폴더에 전체 stay point를 하나의 csv 파일로 저장해 준다.

1. 각 유저의 일별 stay point를 하나의 csv 파일로 저장한다.

텍스트이(가) 표시된 사진

자동 생성된 설명

텍스트이(가) 표시된 사진

자동 생성된 설명