

IITP XML 파일 보고서

작성일	2025-08-14
작성자	임성묵, 황유민
소속	데이터 사이언스팀
관련 문서 및 링크	agent 생성 python 코드

요약

- 수정된 사항을 반영하여 엑셀 파일 만든 후 XML 파일로 연동하여 `population.xml` 파일 작성.
- 작성된 `population.xml` 파일로 `agent.xml` 파일을 생성하기 위한 python 코드 작성.
- `Population_2.xml` 파일과 `population_3.xml` 파일을 바탕으로 `agent.py` 코드를 사용하여 `agent.xml` 파일 100 개씩 생성

내용

- 기존 행동 특성 중 성별, 나이, 1인 가구를 삭제함.
- 일반인, 학생(초등학생, 중학생, 고등학생)을 제외한 직업 인구는 구조 인원으로 배치함.
- 기준 low, high 를 구분하였으나, $high = 1 - low$ 이므로, low 비율만 작성하여 파일 생성함.
- 울주군과 기장군의 범위로 생성한 `population_2.xml` 파일과 해운대구를 추가한 범위로 생성한 `population_3.xml` 파일을 바탕으로 agent 파일을 각각 100 개씩 생성. (`agent_2.zip, agent_3.zip`)
- 실증 지역의 인구수가 많아, 100 명으로 지정하여 테스트를 진행함. (`agent_test_1.xml`)
- 실제 인구수로 테스트한 파일이 `agent_file_1.xml ~ agent_file_100.xml`로, 전체 인원에 대해 행동 특성 부여함.
- 울주군과 기장군을 기준으로 할 경우 총 인구수는 약 40 만 명으로, `agent.xml` 파일 크기가 약 230MB 임.
- 해운대구를 합칠 경우 총 인구수는 약 77 만 명으로, `agent.xml` 파일 크기가 약 450MB 임.
- `agent.py` 파일에 대한 설명
 - `population.xml` 파일을 불러와 `text` 태그 내부에 `total_population` 태그에 작성된 인구수를 가져옴.

pebbrous

- `for` 루프를 사용하여 원하는 개수 (`number_of_files_to_create`) 만큼 파일을 반복적으로 생성.
생성되는 파일 이름은 자동 카운팅되어 저장됨.
- `personal_categories_min_max` 딕셔너리를 순회하며 각 직업에 대해 `min`과 `max` 사이의 임의의 실수를 생성.
- 각 파일을 생성할 때마다 대피자의 특성 (`behavior`, `injury`)에 대한 `low` 비율이 `min`과 `max` 범위 내에서 새로운 난수로 생성됨. 따라서 각 `agent` 파일은 모두 서로 다른 무작위의 특성 분포를 가짐.
- `personal`, `behavior`, `injury` 등 각 태그에 있는 데이터를 파싱하여 `python` 딕셔너리에 저장.
이는 `n` 개의 파일을 생성하기 전에 단 한 번만 실행됨.