버스 노선 데이터분석

곽재현

**I. 소개**

버스 노선의 정류장의 승하차 승객데이터를 분석하고 시각화를 통해 다양한 인사이트를 확보하는 프로젝트입니다.

**II. 제작 동기**

저는 버스를 타면서 어떤 정류장에는 타는 승객이 적은데 내리는 승객이 왜 많을가, 혹은 마을버스가 도시버스보다 승객이 더 많은 이유는 뭘까 하는 궁금중을 가지고 있었습니다. 그래서 저는 이 궁금중을 해결하기 위해 이 프로젝트를 시작했습니다.

**III. 제작 과정**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **준비** | **조사** | 티머니 교통카드 통계자료(2023년 데이터) 사용. | | |
| 인터넷 정보를 통해 조사 | | |
| **기술** | **내용** | * 개발환경: Google Colab, Google Chrome(브라우저) * 기타: GIt, 구글 드라이브 * 사용된 파이썬 모듈: Matplotlib, Pandas | |
| **학습 방법** | * 모두의데이터(책) | |
| **제작**  **과정** | [2023년 3월] | | | 아이디어 고도화 및 프로젝트 계획 |
| [2023년 4월] | | | 프로젝트 초기 아이디어 변경 |
| [2023년 5월] | | | 데이터 조사 및 확보 |
| [2023년 6월] | | | github 사용을 시작함.  csv 모듈에서 Pandas 모듈로 변경 |
| [2023년 7월] | | | 일부 추가적 기능을 하는 코드 셀을 추가  코드 개선 |
| [2023년 8월] | | | 주피터랩(Jupyterlab)에서 구글 코랩(Google Colab)으로 작업환경 변경 |

**V. 기능 소개**

|  |  |
| --- | --- |
|  |  |
| 승차 / 하차 비율이 가장 높은 버스 노선의 정류장 계산 코드  계산 후 그 정류에 대한 파이차트를 표시함. | 버스 노선을 입력 받은 후 그 버스 노선에서 정류장 중 승차 / 하차 비율이 가장 높은 곳은 계산하고 파이차트를 표시함 |
|  | 입력받은 버스노선의 모든 정류장의 승하차 승객을 파이차트로 표시함. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **VI. 주요 코드 (위의 기능 소개 중 세번째 것)**   |  | | --- | |  | | **Pandas를 pd로, matplotlib의 pyplot을 plt로 Python 모듈을 불러옴.**  **month\_data와 f\_name을 month\_import(사용자 함수)등에 사용하기 위해 값을 만듦.** |  |  | | --- | |  | | **month\_import함수를 지정함.**   * **사용자에게 월을 입력 받아 그 월에 해당하는 데이터를 불러옴.** * **그 다음 버스 노선을 입력받고 그 버스 노선에 해당한 데이터를 df2로 데이터프레임을 만듦(만약 버스 노선이 csv파일에 없으면 ValueError을(를) 표시함.** |  |  | | --- | |  | | **버스노선의 정류장별 승하차 승객 수를 파이차트로 표시함.** | |