|  |  |  |
| --- | --- | --- |
| **| 학습주제** | | * 서울 지하철 이용 데이터를 사용한 데이터 분석 및 시각화 |
| **| 학습목표** | | * Pandas를 활용하여 대용량 데이터를 처리하고, 필요한 정보를 추출하는 방법을 학습한다. * Matplotlib 또는 Seaborn을 사용하여 데이터 시각화를 수행할 수 있다. * 특정 조건에 맞는 데이터를 필터링하고, 분석할 수 있다. |
| **| 지문** | 프로젝트의 두 번째 과업에서는 서울 지하철 이용 데이터를 분석하여 요일별 승객 이용 패턴을 파악하고, 승객 수가 가장 많은 역을 찾아내는 것입니다. 이 분석은 도시 교통의 효율적 운영을 위한 전략적 인사이트를 제공하는 데 목적이 있습니다.  데이터 설명  다음은 서울 지하철 승객 수 데이터를 포함한 가상의 데이터셋입니다. 각 행은 한 역의 하루치 승객 수를 나타내며, 각 열은 해당 날의 특정 정보를 나타냅니다.  Date: 날짜 (YYYY-MM-DD 형식)  Station: 지하철 역 이름  DayOfWeek: 요일 (월, 화, 수, 목, 금, 토, 일)  Passengers: 하루 동안의 승객 수 | |
| **| 학습 문제** | 1. 주어진 지하철 이용 데이터를 Pandas 데이터프레임으로 로드하세요. 그리고 로드된 데이터의 처음 5행을 출력하세요. 2. 요일별로 승객 수의 총합을 계산하세요. 그리고 계산한 결과를 출력하세요. 3. Matplotlib과 Seaborn을 이용하여 요일별 승객 수 변화를 시각화하세요. 이때 한글 텍스트가 포함된 그래프를 정상적으로 출력할 수 있도록 Matplotlib의 폰트를 설정하세요. 4. 금요일에 평균 이용자 수가 가장 높은 역의 이름과 평균 이용자 수를 출력하세요. | |