

Raw Data :
제주도의 2019년 9월 1일 ~ 30일의 버스 운행 데이터
Size = 415,423 개

Raw Data에서 Train, Test, Zero df로 나눔

Zero_df :
Raw Data 중
1회 운행된 정류소 제외
(예측모델 생성불가)
Size = 0.2

Test Data:
Size = 0.2

Train Data (1) :
Category Columns의 고유데이터 포함
Size = 0.8

Train Data에서 Validation Data로 나눔

Validation Data :
Train Data(1)에서
랜덤하게 검증용
데이터 추출
Size = 0.2

Train Data (2) :
Category Columns의 고유데이터 포함
Size = 0.8

생성된 Model을 Train과 Test Data에 적용

Model 생성 및 검증

Train Data와 Validation Data로 Model 생성