

ARIMA 모델을 활용한 상품의 적정 재고량 분석 보고서





동아대학교
DONG-A UNIVERSITY

목차

ARIMA 모델을 활용한
상품의 적정 재고량 분석 보고서

1. 데이터 분석
2. 적정 재고량 산정
3. 계절성을 고려한 모델 예측
4. 휴일 유무에 따른 예측 값 변화

DATA
SCIENCE
LABS

01) 데이터 정보 파악

- 두 데이터는 각각 MG0921과 MG1342의 **생산량, 출고량, 현재고량**을 **생산일자** 별로 기록한 데이터이다.
- 두 데이터의 개수는 MG0921은 1013개, MG1342는 1113개이다.
- 두 데이터 모두 2018-01-02부터 2023-10-12까지의 기간동안 수집한 데이터이다.

생산일자	상품	생산량	출고량	현재고량
2018-01-08	MG0921	42000	6000	48650
2023-09-27	MG0921	0	6000	91310
2023-10-12	MG0921	22530	0	126270

<MG0921 데이터의 일부>

생산일자	상품	생산량	출고량	현재고량
2018-01-08	MG0921	42000	6000	48650
2023-09-27	MG0921	0	6000	91310
2023-10-12	MG0921	22530	0	126270

<MG0921 데이터의 일부>

02) 데이터 정상성 확인

■ 시계열 데이터 정상성

- 평균과 분산이 일정하고 추세와 계절성이 존재하지 않을 경우, 해당 시계열 데이터는 **정상성을 가지는 데이터**이다.
- 정상성을 가지지 않은 시계열 데이터의 경우 예측이 잘 이루어지지 않기 때문에 **Time Series Decompose**와 **ACF, PACF** 그래프를 통해 데이터가 정상성을 가지는지 확인해야 한다.

-시계열 데이터: 일정한 시간 동안 수집 된 일련의 순차적으로 정해진 데이터 세트의 집합
-추세(Trand): 시계열 데이터가 시간에 따라 증가, 감소, 일정 수준을 유지하는 경우
-계절성(Seasonality): 일정한 빈도로 패턴이 반복되는 경우

02) 데이터 정상성 확인

Time Series Decompose

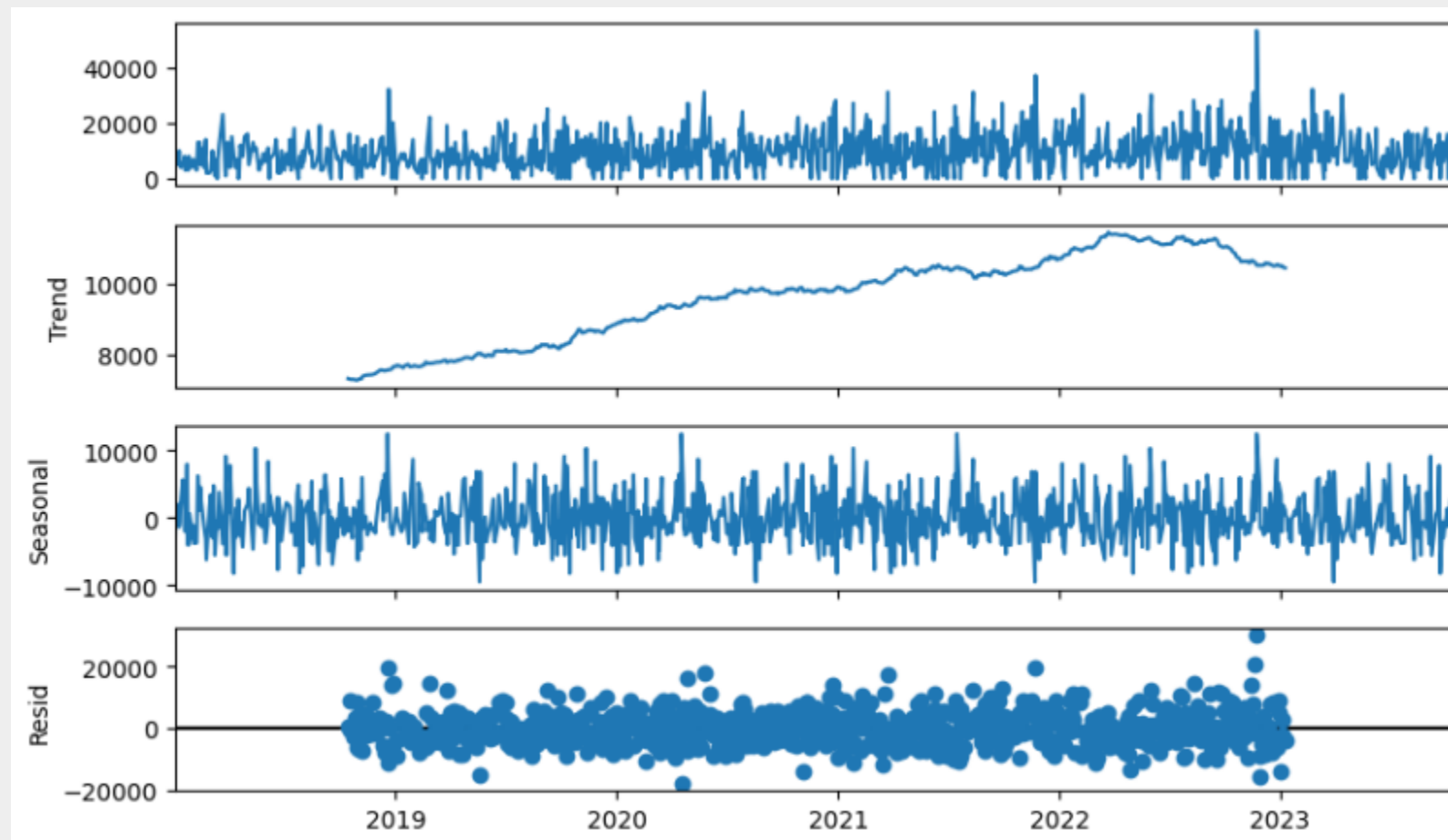
- 시계열 데이터를 추세, 계절성, 잔차로 분해하는 방법을 말한다.
- 데이터의 상승/하락하는 움직임을 추세로 추출하고, 추세를 추출한 후 주기적으로 반복되는 scale은 계절성으로 추출한다. 추세와 계절성을 추출한 후 남은 것을 잔차로 둔다.
- 모델을 통해 예측값을 구할 때는 각 데이터의 '출고량' 데이터만을 사용할 예정이기 때문에 해당 부분만 데이터를 분해할 예정이다.

-잔차(Residual): 데이터를 통해 예측된 값과 실제 값 사이의 오차

02) 데이터 정상성 확인

Time Series Decompose

MG0921과 MG1342 출고량의 Seasonal Decompose 결과에서 추세가 존재함을 확인했다.



<MG0921 출고량>



<MG1342 출고량>

02) 데이터 정상성 확인

■ 상관도표 (Correlogram)

- 데이터들의 정상성 만족유무에 대한 더 정확한 판단을 위해 ACF와 PACF를 사용한다.
- ACF와 PACF를 그래프로 표현했을 때 정상성을 만족하는 상관관계는 0과 가까운 값을 가진다.

■ ACF (Autocorrelation Function, 자기상관함수)

- k시간 단위로 구분된 시계열 관측치 y_t 와 y_{t+k} 간의 상관관계를 측정하는 함수

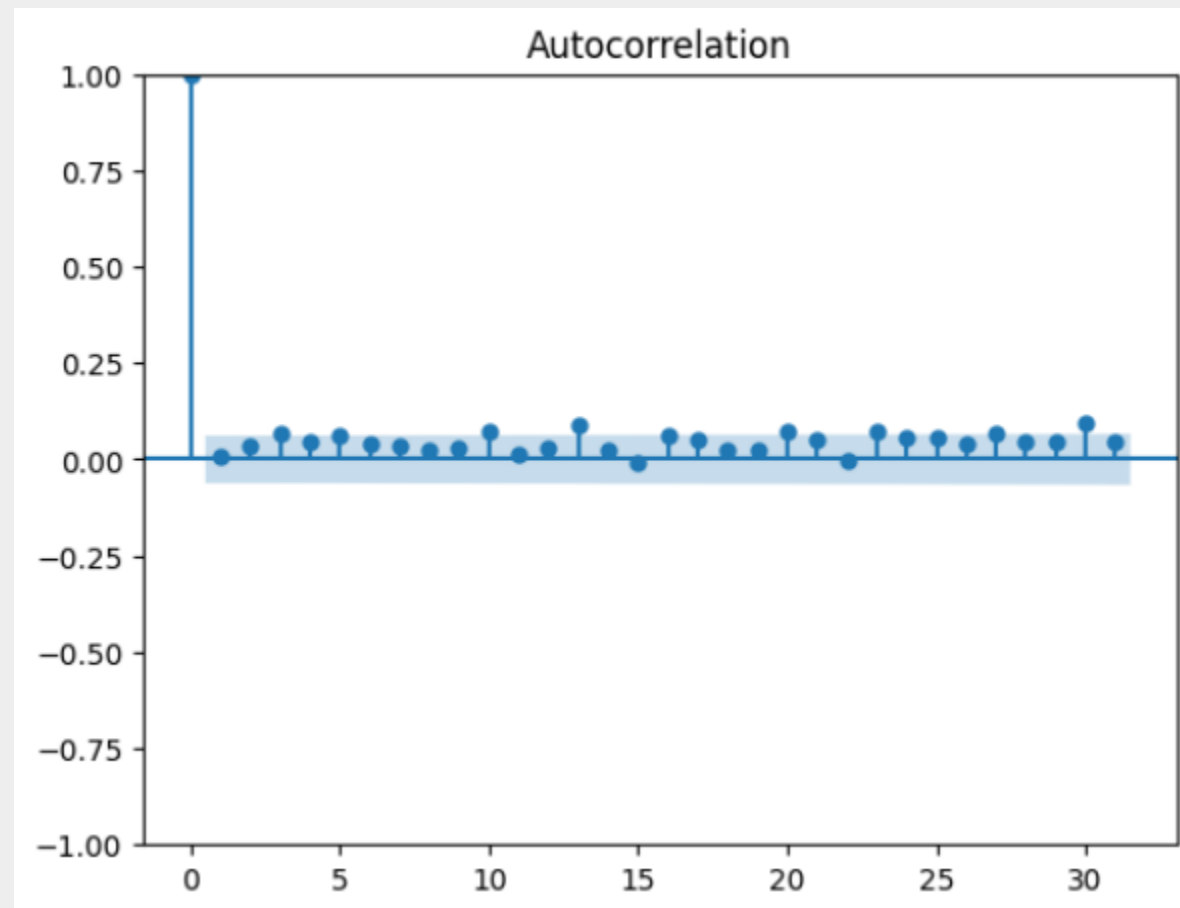
■ PACF (Partial Autocorrelation Function, 편자기상관함수)

- ACF 이후의 추가적인 분석이 필요할 경우 사용하는 함수
- y_t 와 y_{t+k} 간의 상관관계를 측정하는건 ACF와 같지만 다른 시간대의 관측값의 영향은 고려하지 않는다.

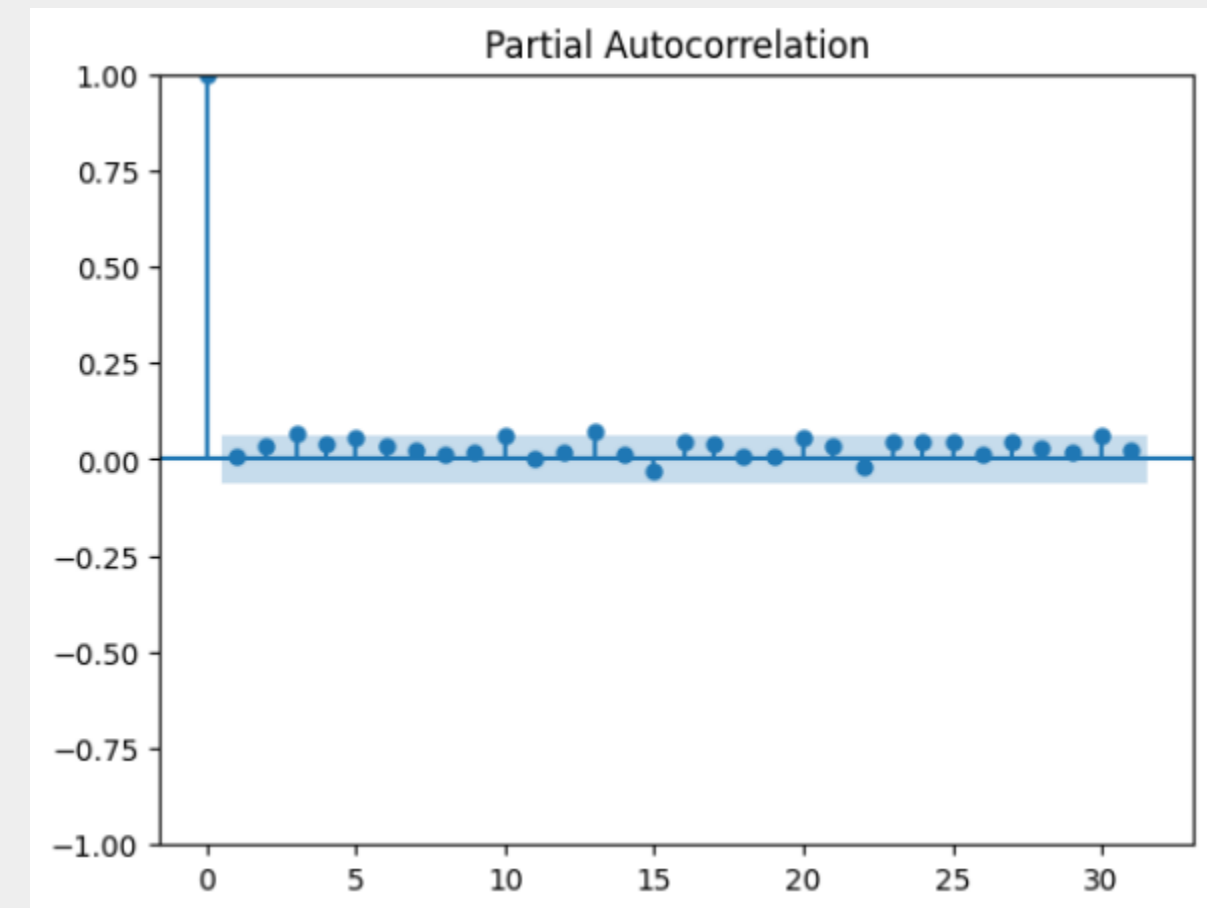
02) 데이터 정상성 확인

MG0921의 ACF, PACF

조금씩 0을 벗어난 데이터가 보이지만 관측값 사이의 상관관계가 거의 없음을 알 수 있다.



<MG0921의 ACF 그래프>

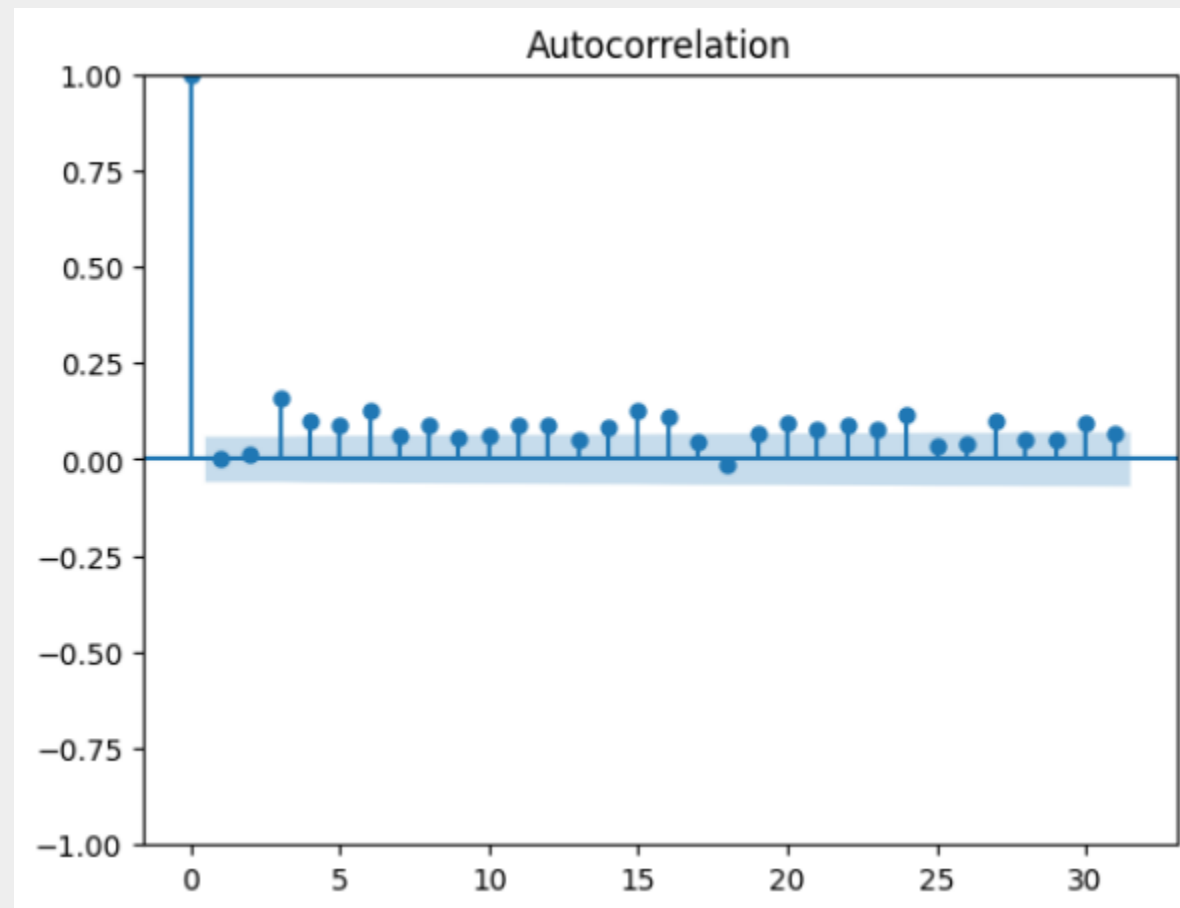


<MG0921의 PACF 그래프>

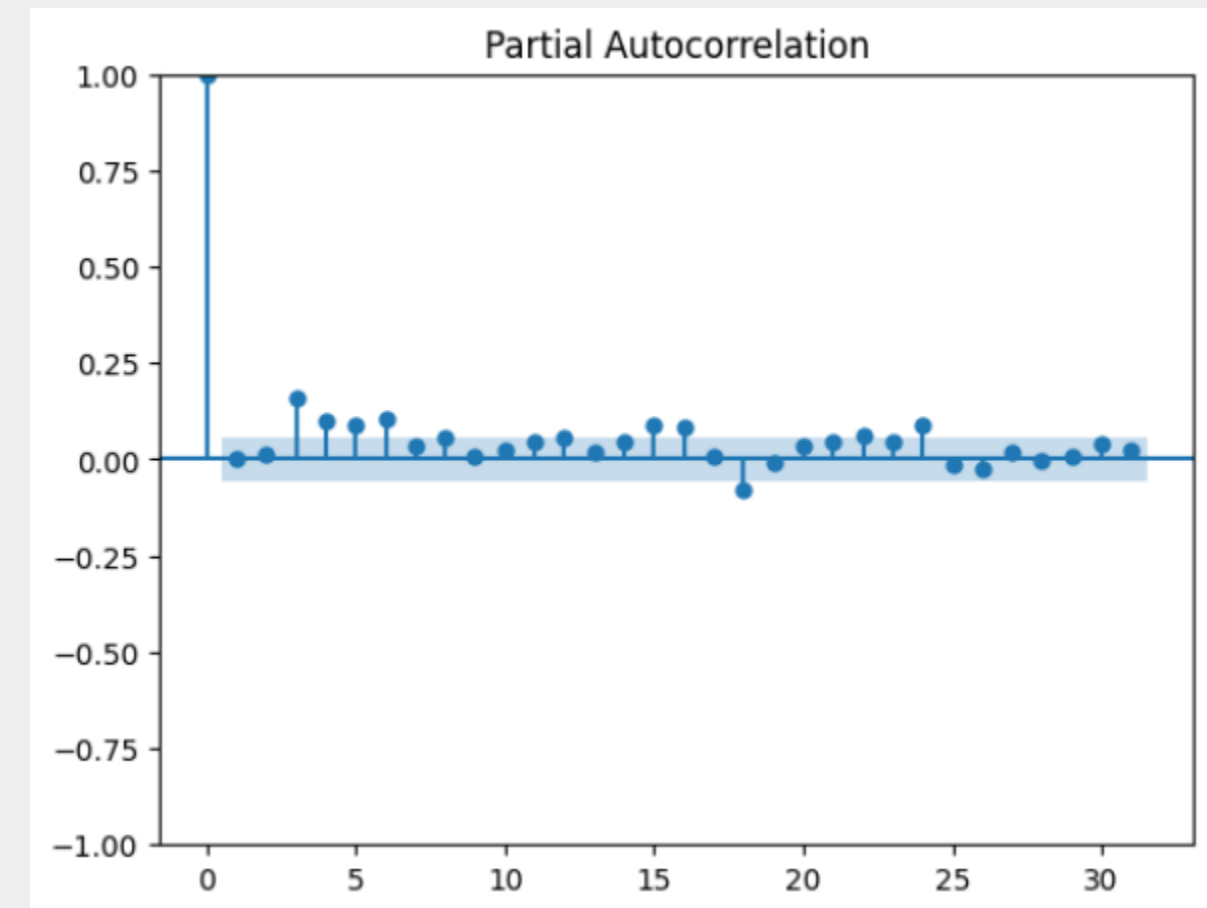
02) 데이터 정상성 확인

MG1342의 ACF, PACF

MG0921보다 관측값 사이에 더 많은 상관관계가 존재함을 알 수 있다.



<MG1342의 ACF 그래프>



<MG1342의 PACF 그래프>

01) ARIMA 모델을 활용한 데이터 예측

ARIMA 모델

- d차 차분한 데이터와 AR(p) 모델과 MA(q)모델을 결합한 모델
- **ARIMA(p,d,q)**로 표현하며, p는 과거 관측값의 개수, d는 차분 횟수, q는 예측 오차의 개수를 의미한다.

AR(Auto Regression, 자기회귀)

- 과거 관측값의 선형결합으로 해당 변수의 미래값을 예측하는 모형

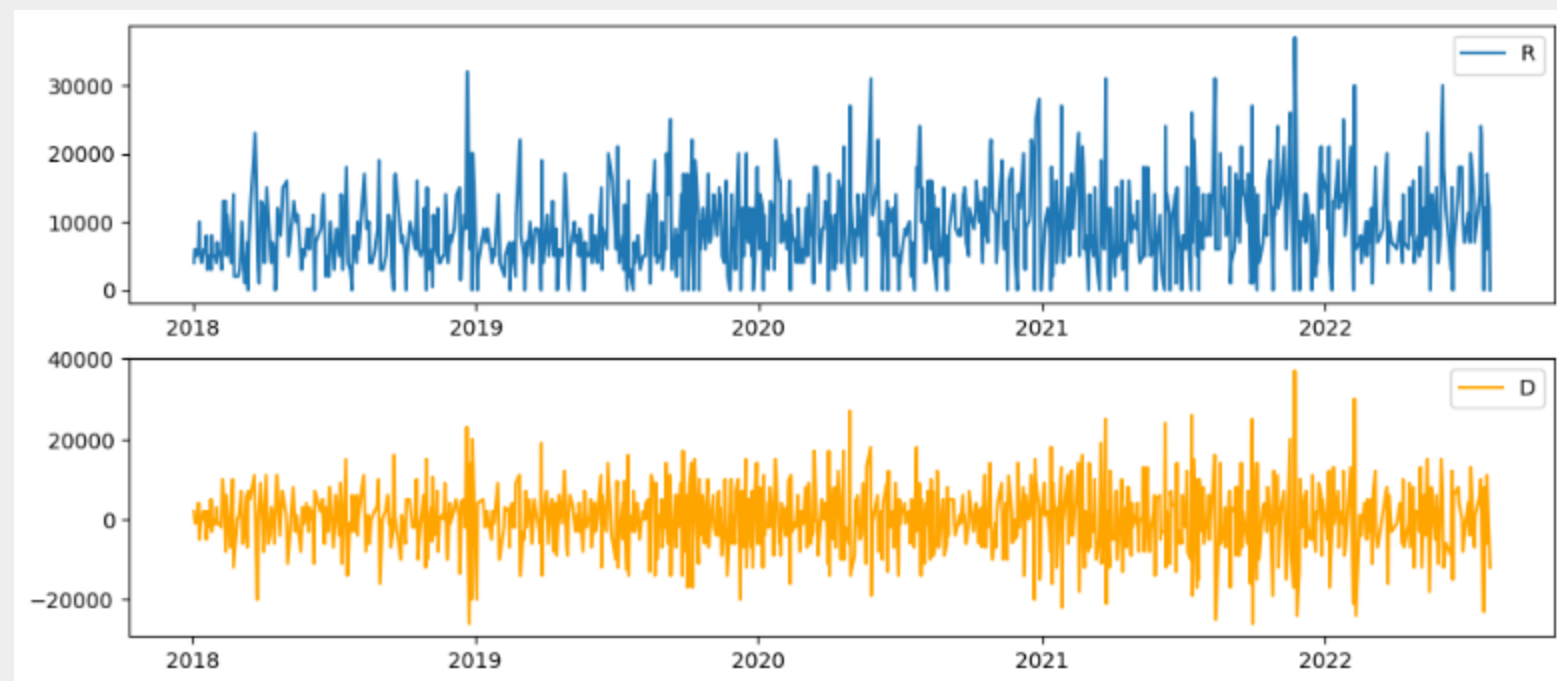
MA(Moving Average, 이동 평균)

- 예측 오차를 이용하여 미래를 예측하는 모형

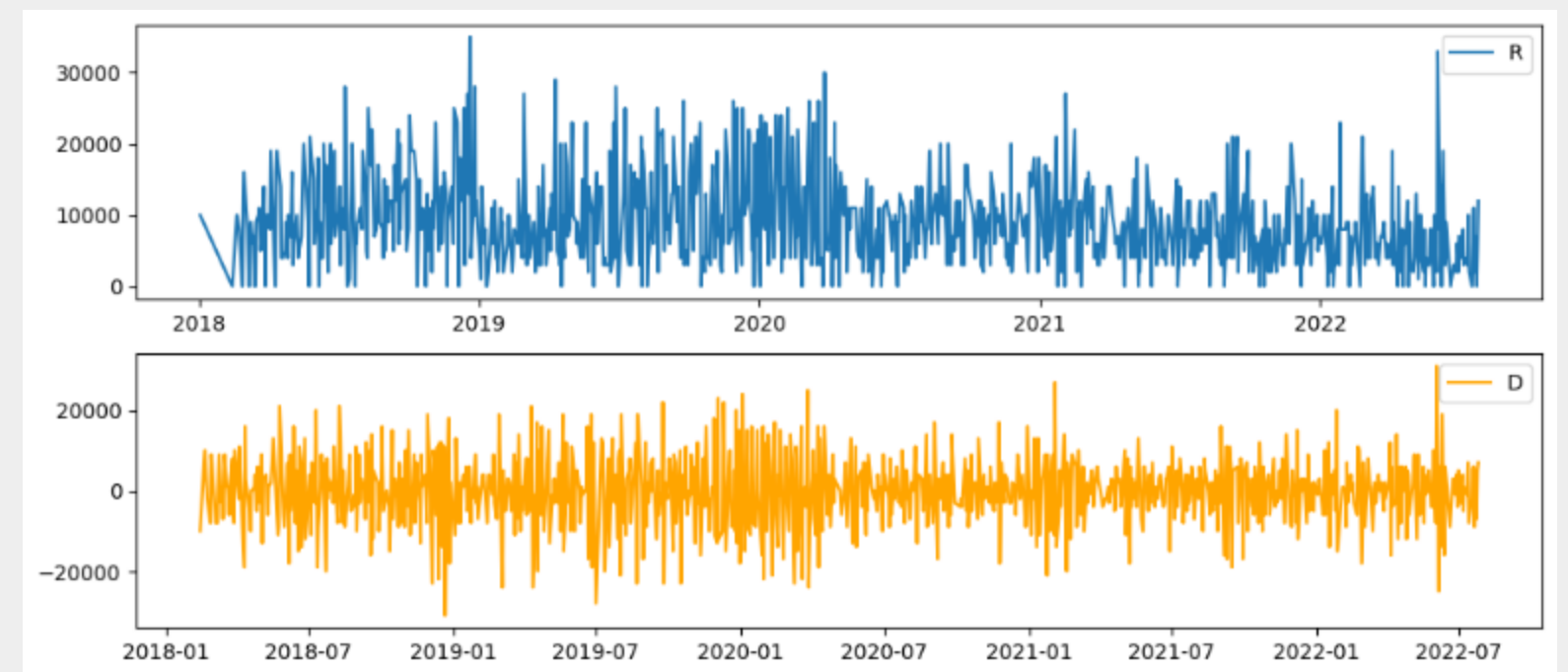
01) ARIMA 모델을 활용한 데이터 예측

차분(differencing)

- 정상성이 없는 시계열 데이터에서 추세와 계절성을 제거해 정상성을 가진 데이터로 변환하는 기법
- 사용하는 두 데이터 모두 추세가 존재하기 때문에 차분을 통해 제거할 필요가 있다.



<MG0921 원본 데이터(R), 차분데이터(D) 비교>



<MG1342 원본 데이터(R), 차분데이터(D) 비교>

01) ARIMA 모델을 활용한 데이터 예측

ARIMA 모델 최적화

- 데이터 예측을 수행할 때 가장 최적화된 파라미터를 구하기 위해 **grid search** 방법을 사용한다.
- 각각의 파라미터에 들어갈 값의 범위를 제한한 후, 해당 범위 안의 조합에서 **가장 낮은 AIC 점수를 획득한 모델이 최적 모델로 선정**된다.

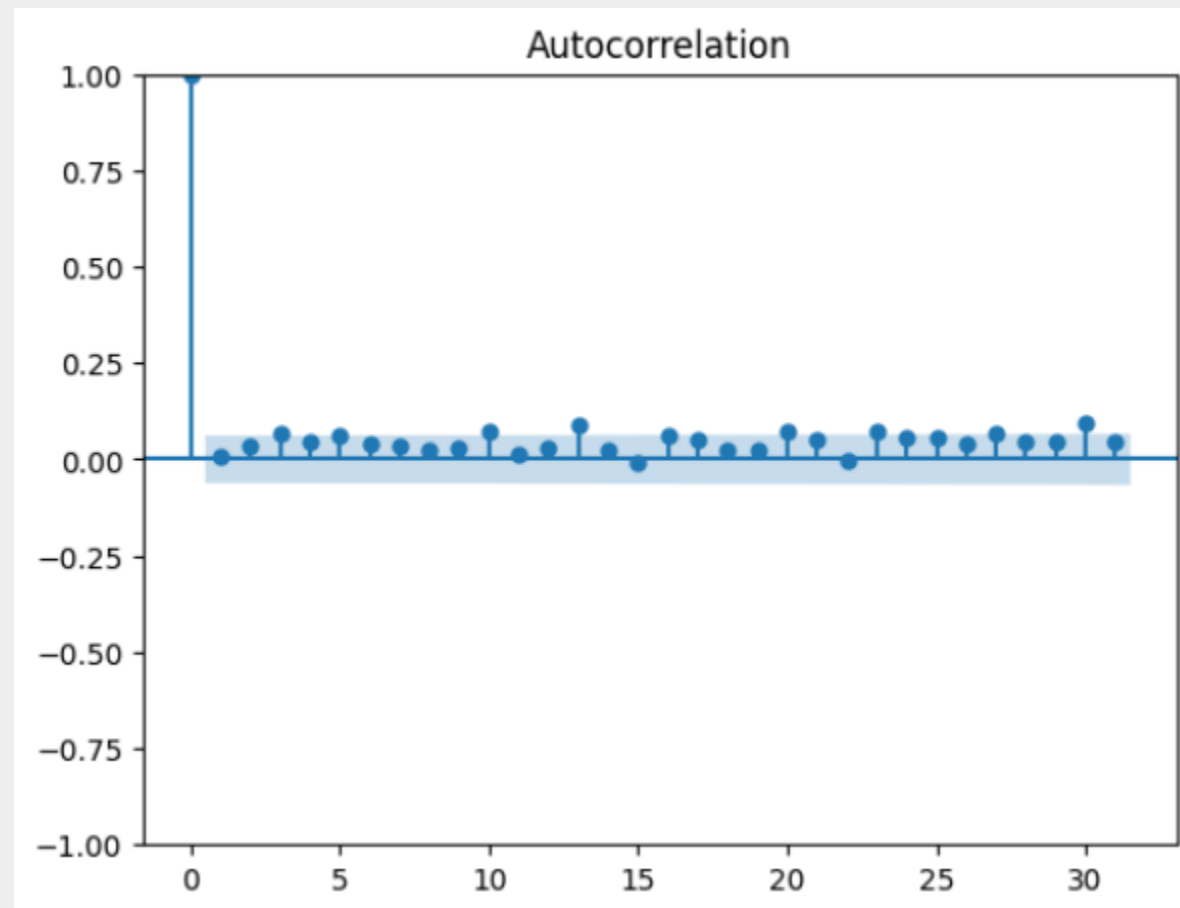
AIC(Akaike's information Criterion)

- 모델 비교를 통해 통계 모델이 얼마나 자료를 잘 설명할 수 있는지를 고려하는 값이다.
- 주로 통계 모델에서 자주 사용하는 지표이며, 값이 낮을수록 좋은 모델로 판단한다.
- 최소의 정보 손실을 갖는 모델을 가장 데이터와 적합한 모델로 선택한다.

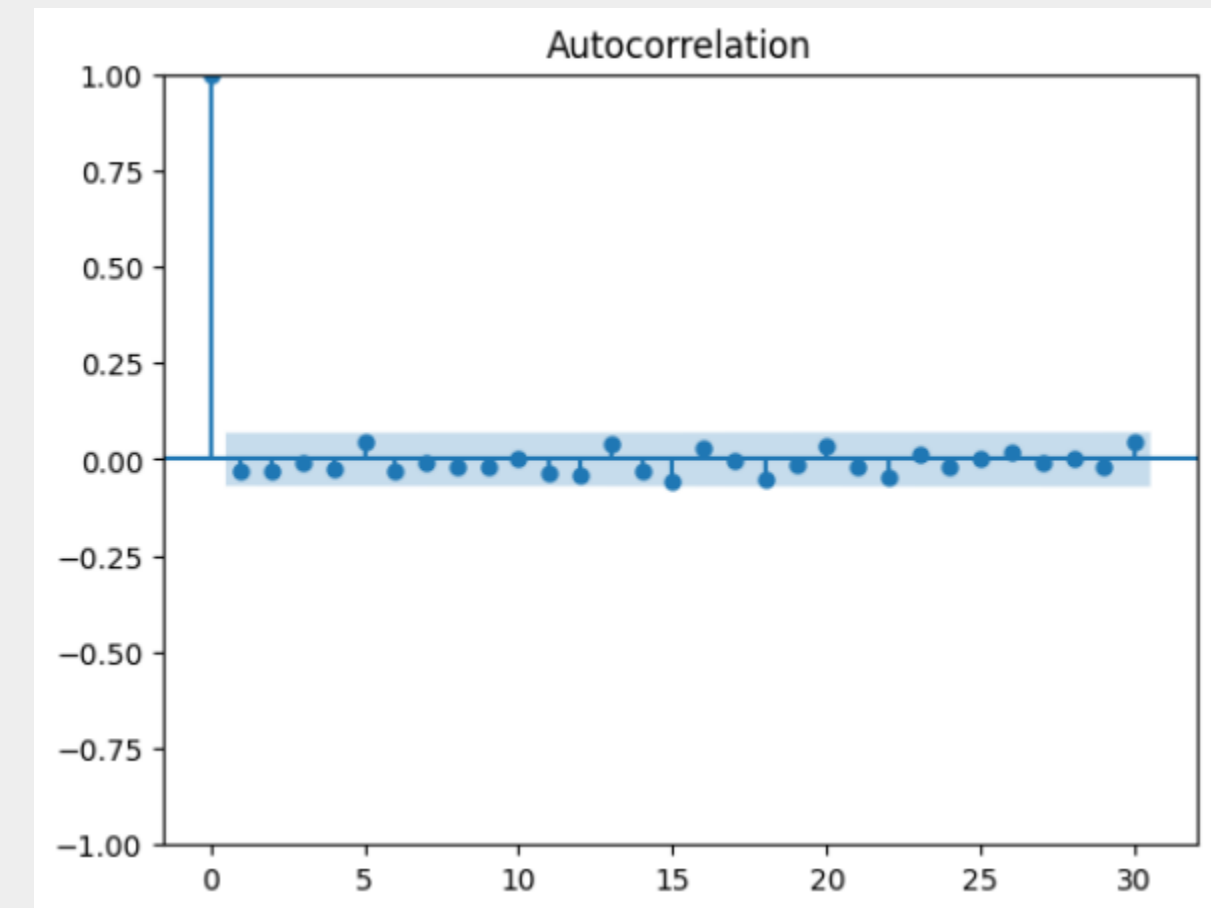
01) ARIMA 모델을 활용한 데이터 예측

ARIMA 모델 최적화

MG0921의 기존 ACF와 최적 ARIMA 모델 적용후 ACF 비교



<MG0921의 ACF 그래프>

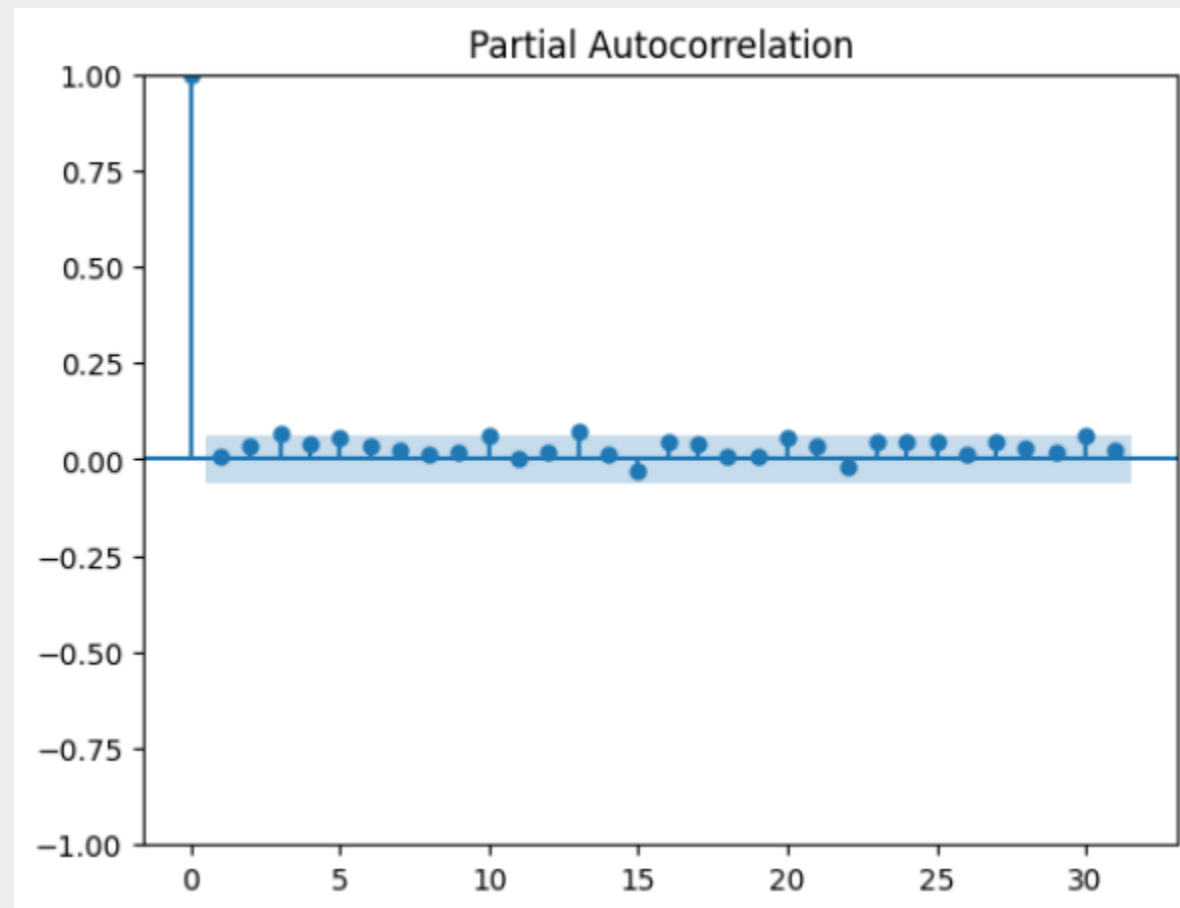


<최적 모델 적용 후 MG0921의 ACF 그래프>

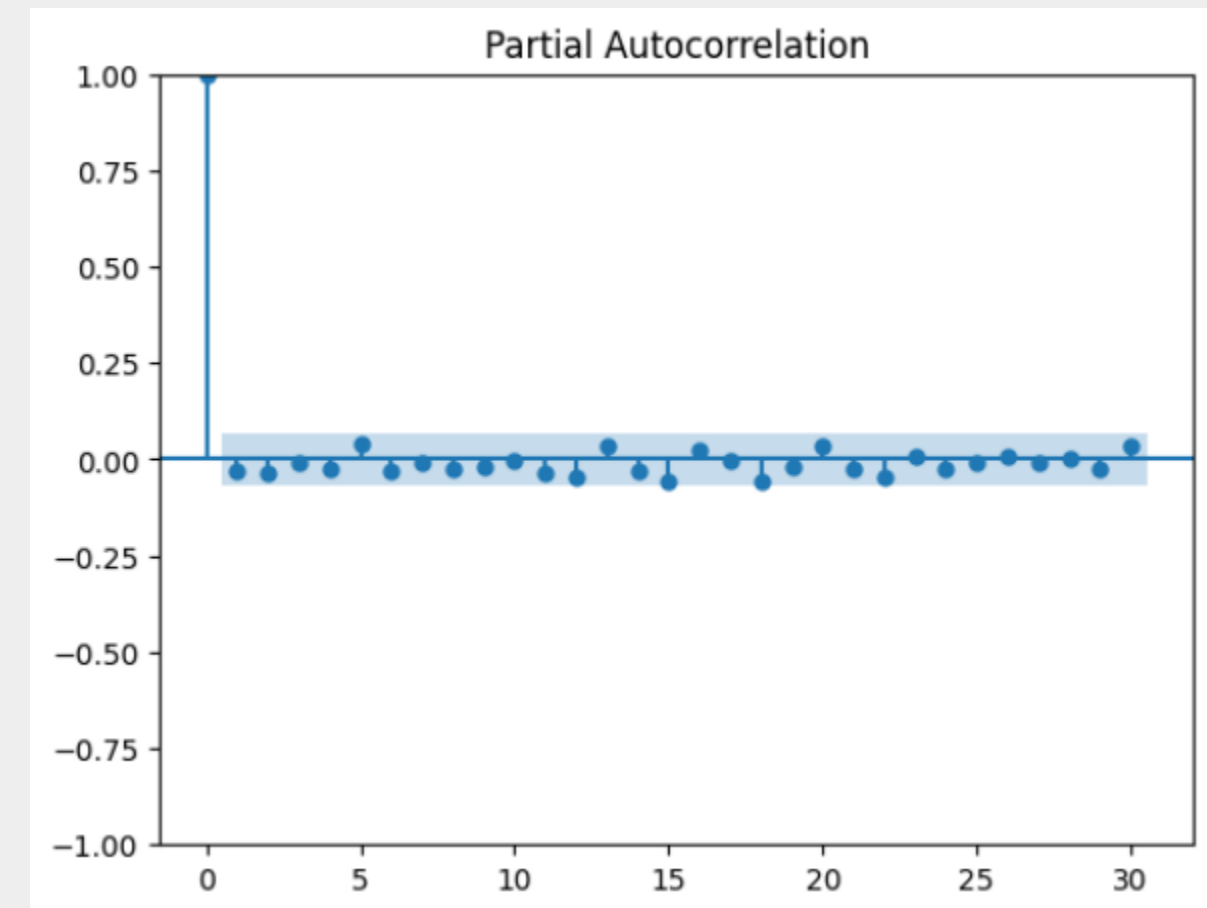
01) ARIMA 모델을 활용한 데이터 예측

ARIMA 모델 최적화

MG0921의 기존 PACF와 최적 ARIMA 모델 적용후 PACF 비교



<MG0921의 PACF 그래프>

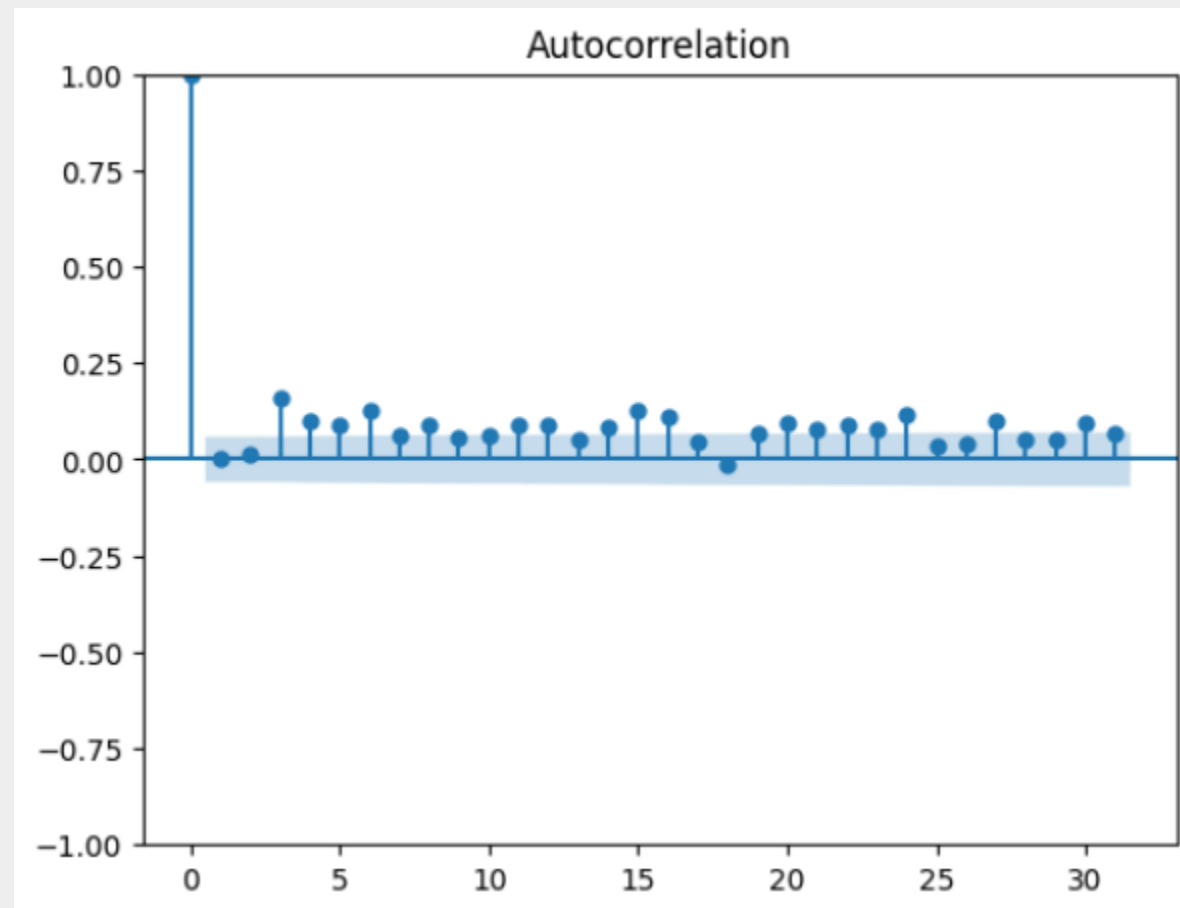


<최적 모델 적용 후 MG0921의 PACF 그래프>

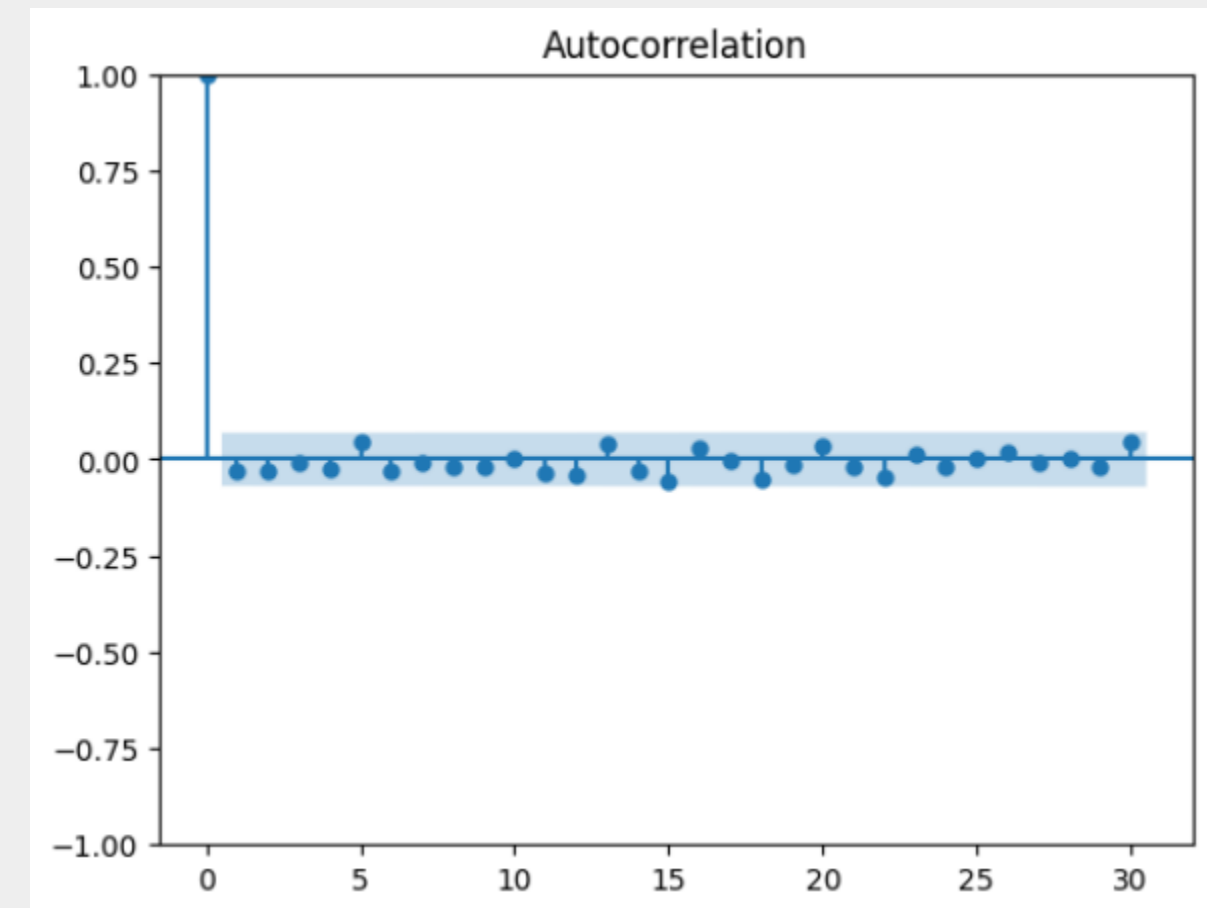
01) ARIMA 모델을 활용한 데이터 예측

ARIMA 모델 최적화

MG1342의 기존 ACF와 최적 ARIMA 모델 적용후 ACF 비교



<MG1342의 ACF 그래프>

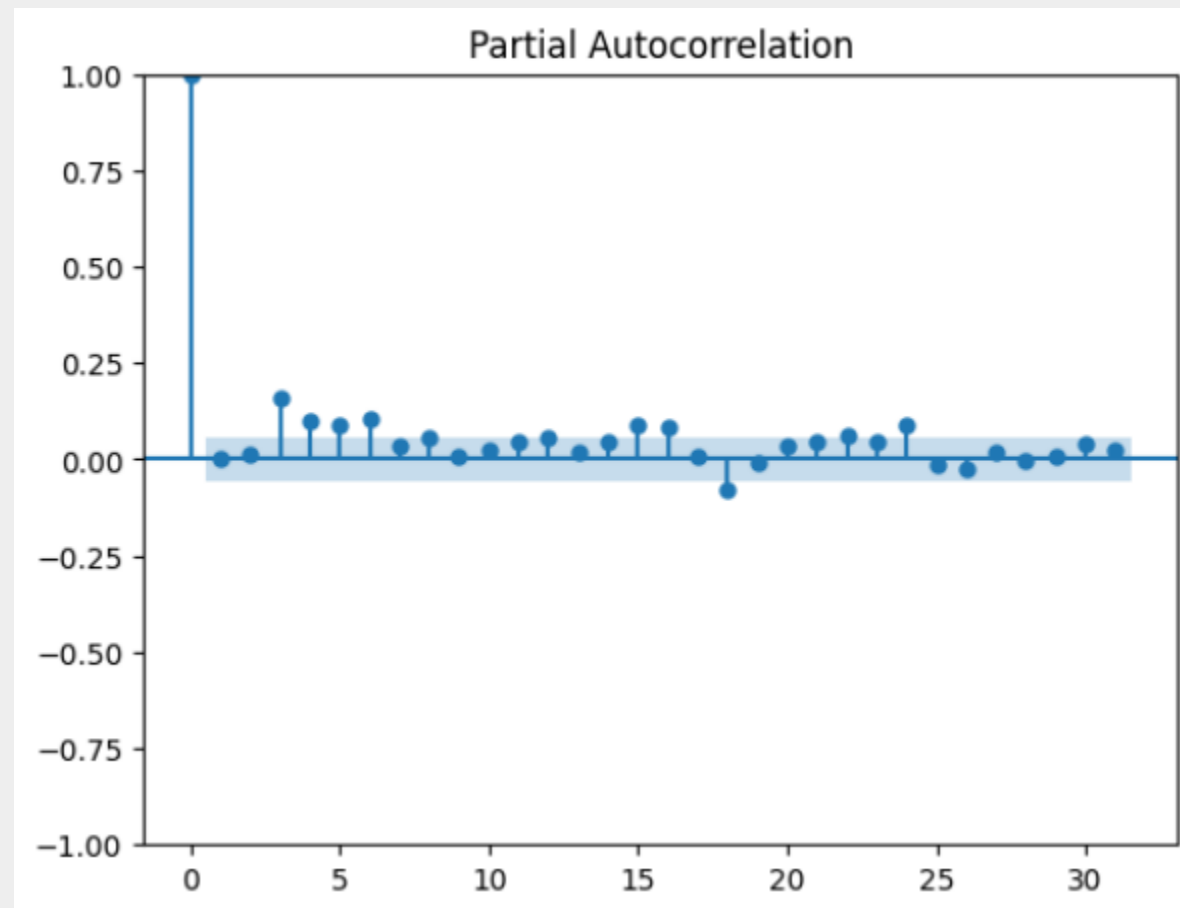


<최적 모델 적용 후 MG1342의 ACF 그래프>

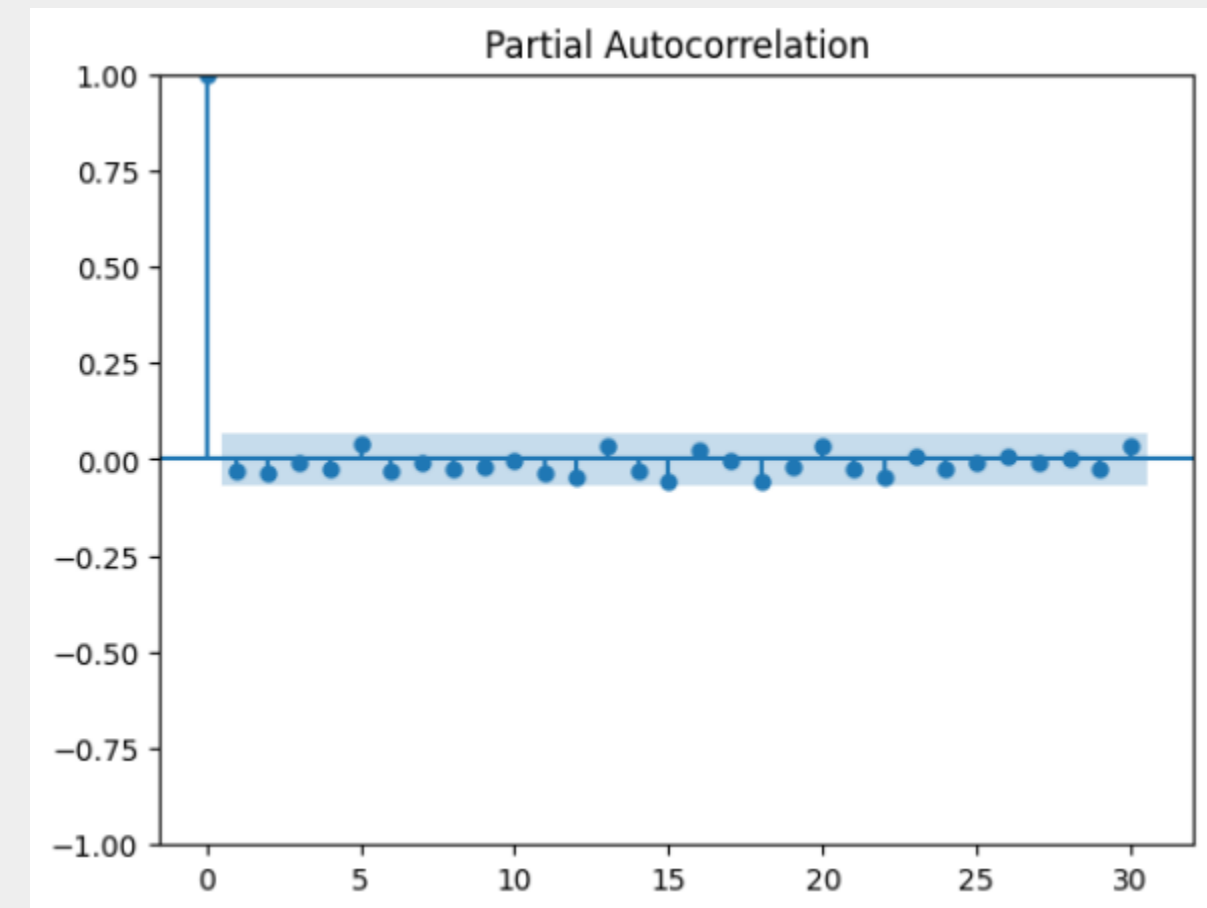
01) ARIMA 모델을 활용한 데이터 예측

ARIMA 모델 최적화

MG1342의 기존 PACF와 최적 ARIMA 모델 적용후 PACF 비교



<MG1342의 PACF 그래프>

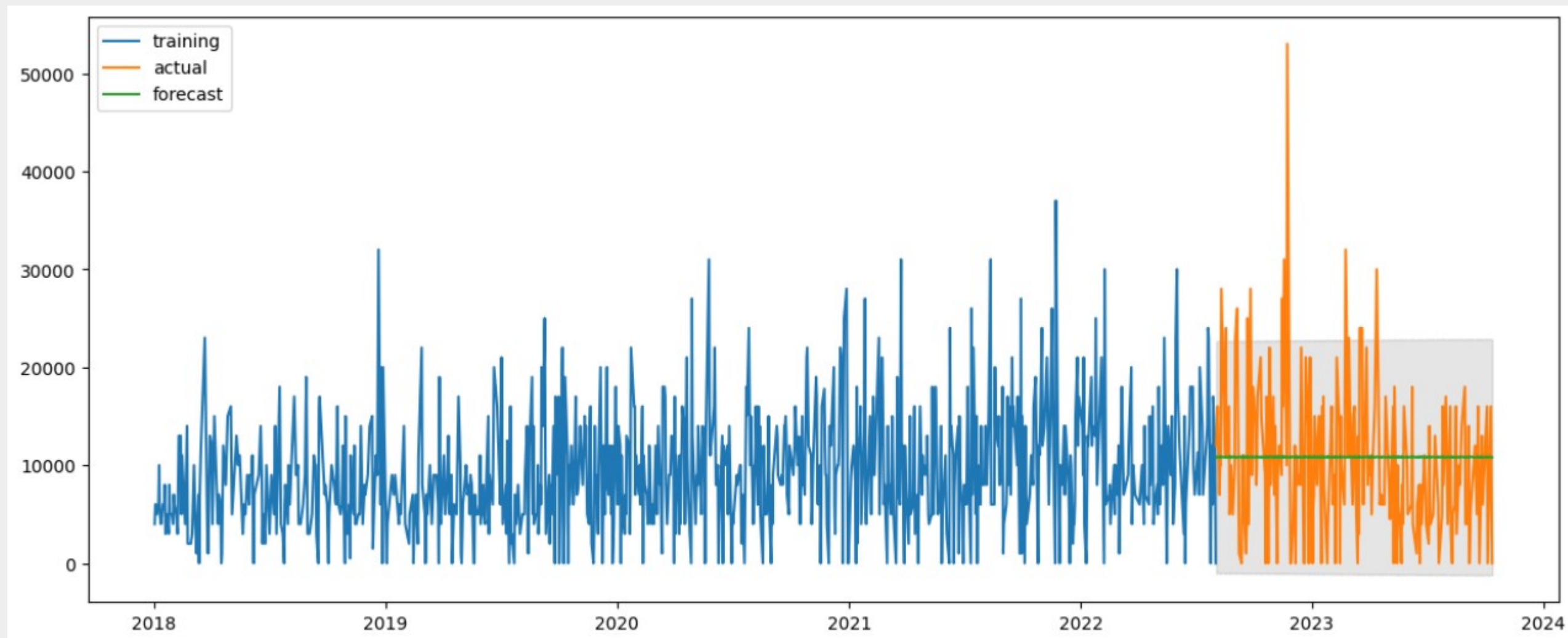


<최적 모델 적용 후 MG1342의 PACF 그래프>

01) ARIMA 모델을 활용한 데이터 예측

ARIMA를 활용한 데이터 예측

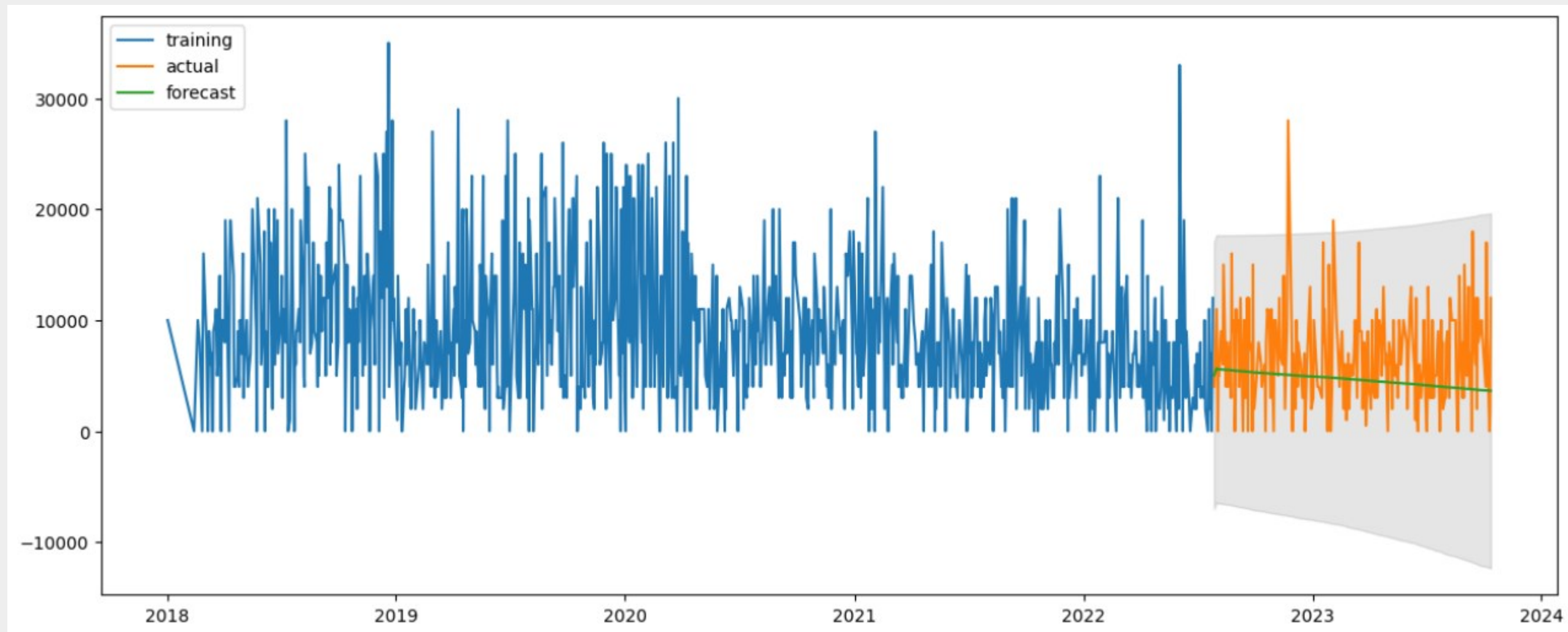
MG0921 데이터를 8:2로 나누어 각각 훈련 데이터와 검증 데이터로 사용하여 예측한 결과



01) ARIMA 모델을 활용한 데이터 예측

ARIMA를 활용한 데이터 예측

MG1342 데이터를 8:2로 나누어 각각 훈련 데이터와 검증 데이터로 사용하여 예측한 결과



02) ARIMA를 통해 적정 재고량 산출

ARIMA를 활용한 적정 재고량 선정

- 앞선 모델을 활용하여 이후 1주일 분량의 출고량을 예측한다.
- 적정 재고량은 예측된 출고량의 평균에 2와 3을 각각 곱하여 계산한다.
- 계산된 적정 재고량을 소수점 첫째 자리에서 반올림한 값은 다음과 같다.

	MG0921	MG1342
2주	17732	14304
3주	26598	21456

01) 계절성을 고려한 SARIMAX 모델

SARIMAX 모델

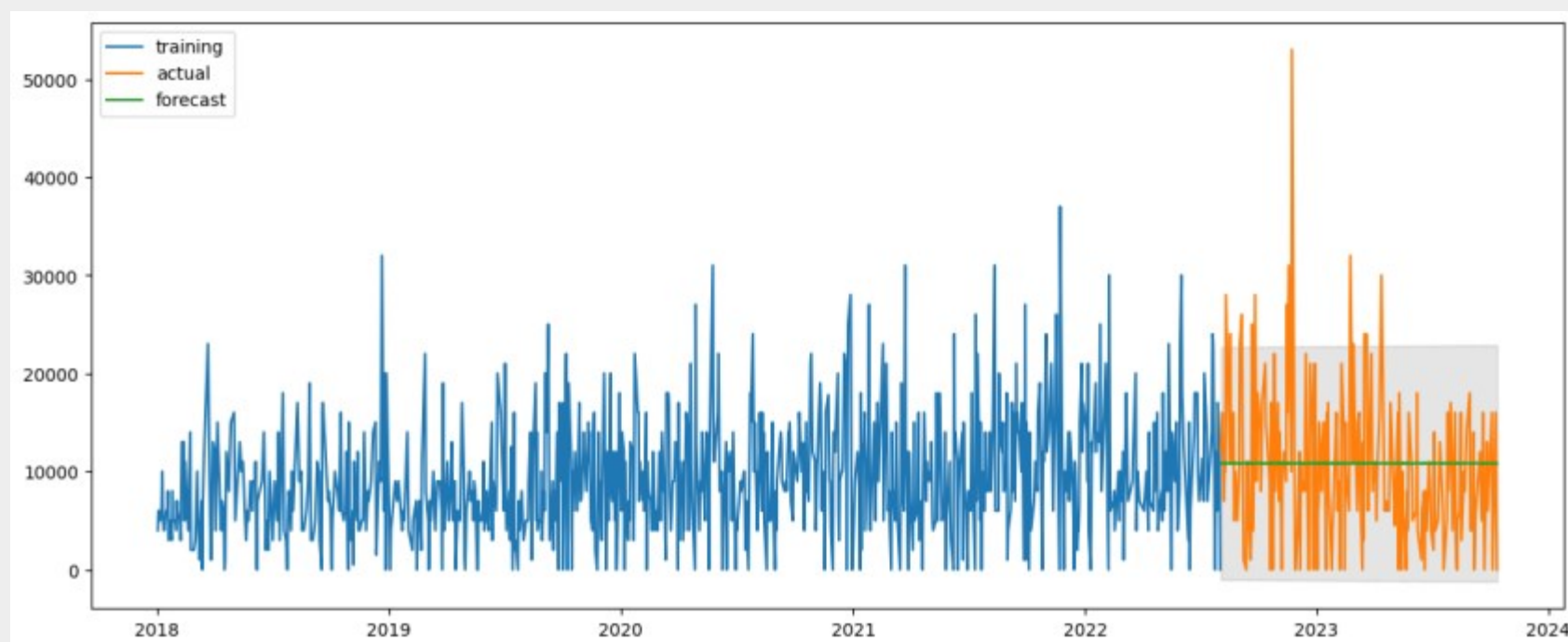
- 계절성을 반영하도록 ARIMA 모델을 확장시켜 만든 모델이다.
- (p,d,q) 에 계절성을 포함한 파라미터 (P,D,Q) 를 추가로 사용한다.
- 증가하는 추세를 보였던 MG0921의 예측 데이터의 평균값은 증가하고, 감소하는 추세를 보였던 MG1342의 예측 데이터의 평균값은 감소하였음을 알 수 있다.

	MG0921_ARIMA	MG0921_SARIMAX	MG1342_ARIMA	MG1342_SARIMAX
2주	17732	23037	14304	11018
3주	26598	34556	21456	16527

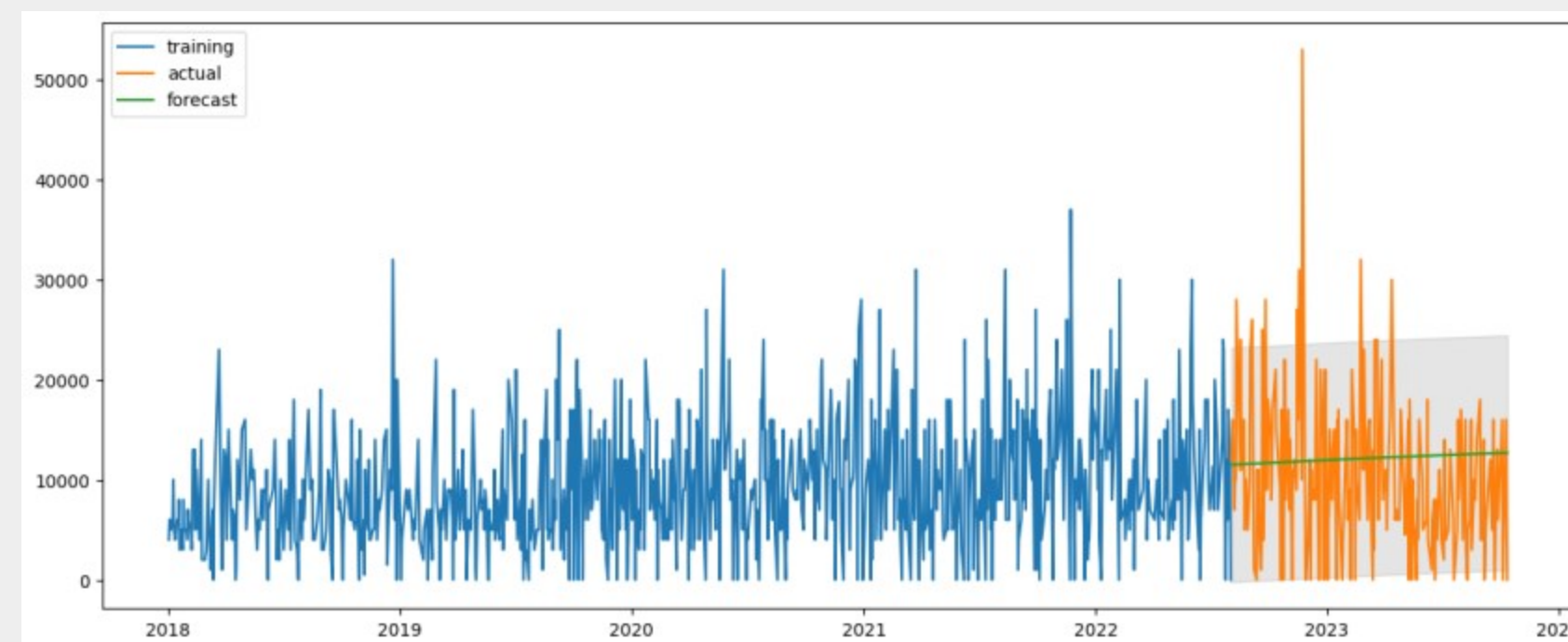
01) 계절성을 고려한 SARIMAX 모델

SARIMAX 모델

MG0921



<ARIMA를 활용한 예측데이터 시각화>

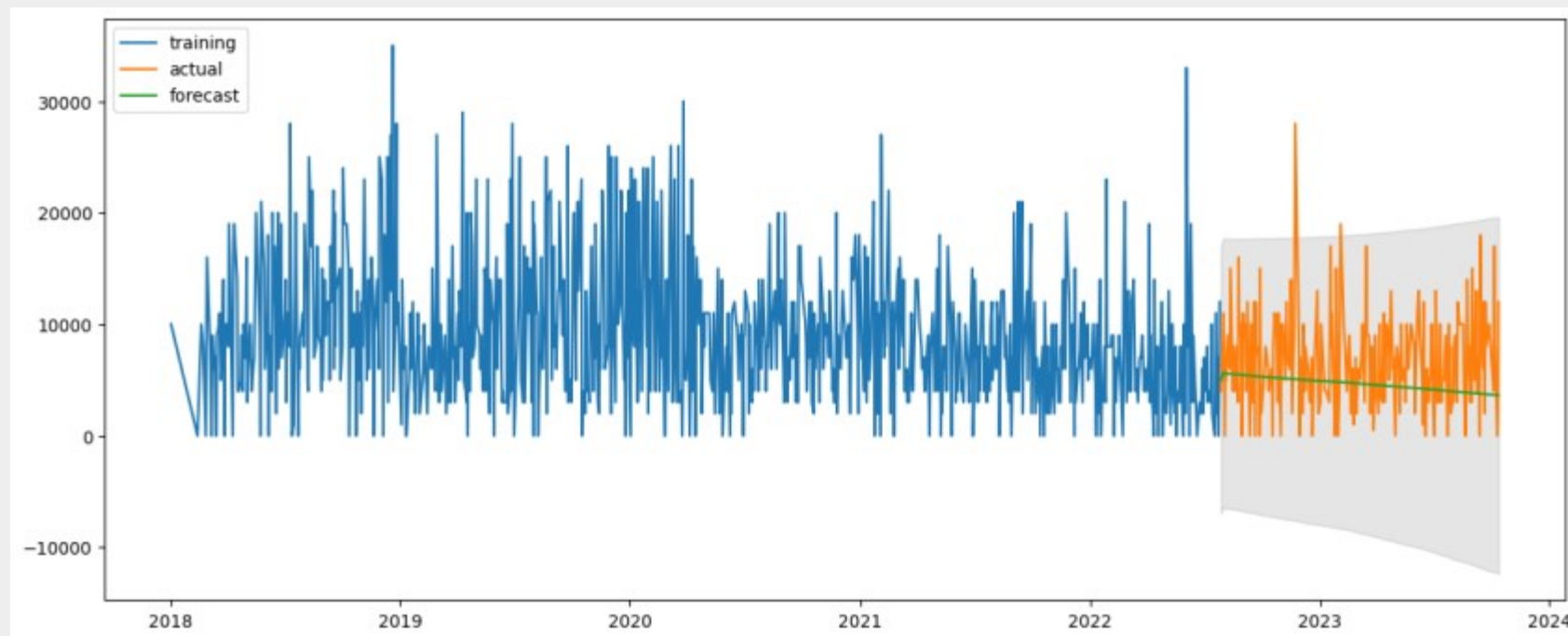


<SARIMAX를 활용한 예측데이터 시각화>

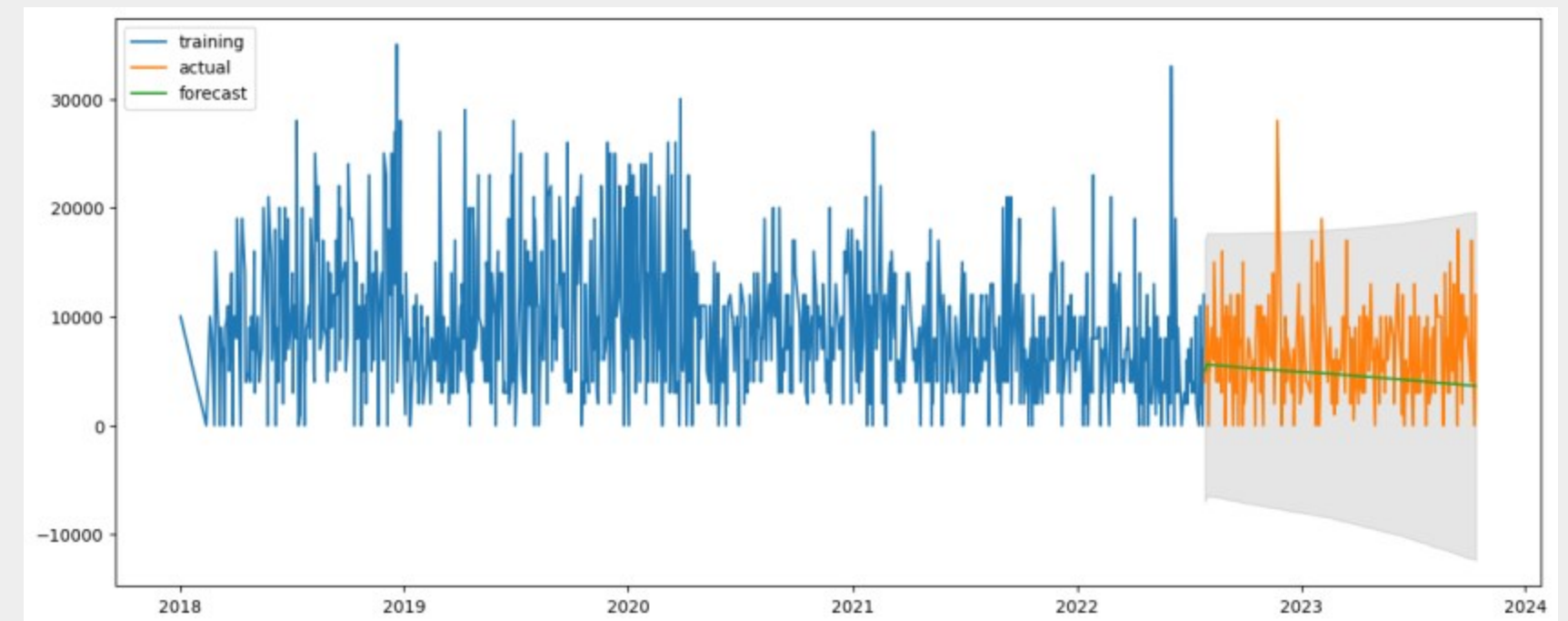
01) 계절성을 고려한 SARIMAX 모델

SARIMAX 모델

MG1342



<ARIMA를 활용한 예측데이터 시각화>



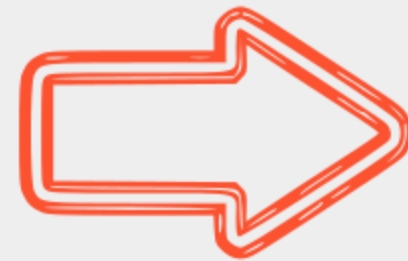
<SARIMAX를 활용한 예측데이터 시각화>

01) 휴일 데이터 획득

휴일인 날의 데이터 보간

- 휴일인 경우 생산량과 출고량의 변화가 없기 때문에 데이터에 업데이트되지 않는다.
- 휴일인 날짜에 생산량과 출고량이 0인 데이터를 넣어 데이터셋을 보간한다.

생산일자	제품	생산량	출고량	현재고량
2018-01-02	MG0921	0	4000	29650
2018-01-03	MG0921	0	6000	23650
2018-01-04	MG0921	0	6000	17650
2018-01-05	MG0921	0	5000	12650
2018-01-08	MG0921	42000	6000	48650
...
2023-09-27	MG0921	0	6000	91310
2023-10-04	MG0921	0	16000	75310
2023-10-05	MG0921	44430	0	119740
2023-10-10	MG0921	0	16000	103740
2023-10-12	MG0921	22530	0	126270



생산일자	제품	생산량	출고량	현재고량
2018-01-01	MG0921	0	0	33650
2018-01-02	MG0921	0	4000	29650
2018-01-03	MG0921	0	6000	23650
2018-01-04	MG0921	0	6000	17650
2018-01-05	MG0921	0	5000	12650
...
2023-10-08	MG0921	0	0	119740
2023-10-09	MG0921	0	0	119740
2023-10-10	MG0921	0	16000	103740
2023-10-11	MG0921	0	0	103740
2023-10-12	MG0921	22530	0	126270

02) 휴일 유무가 가지는 영향 확인

■ 휴일 유무에 따른 적정 재고량의 변화

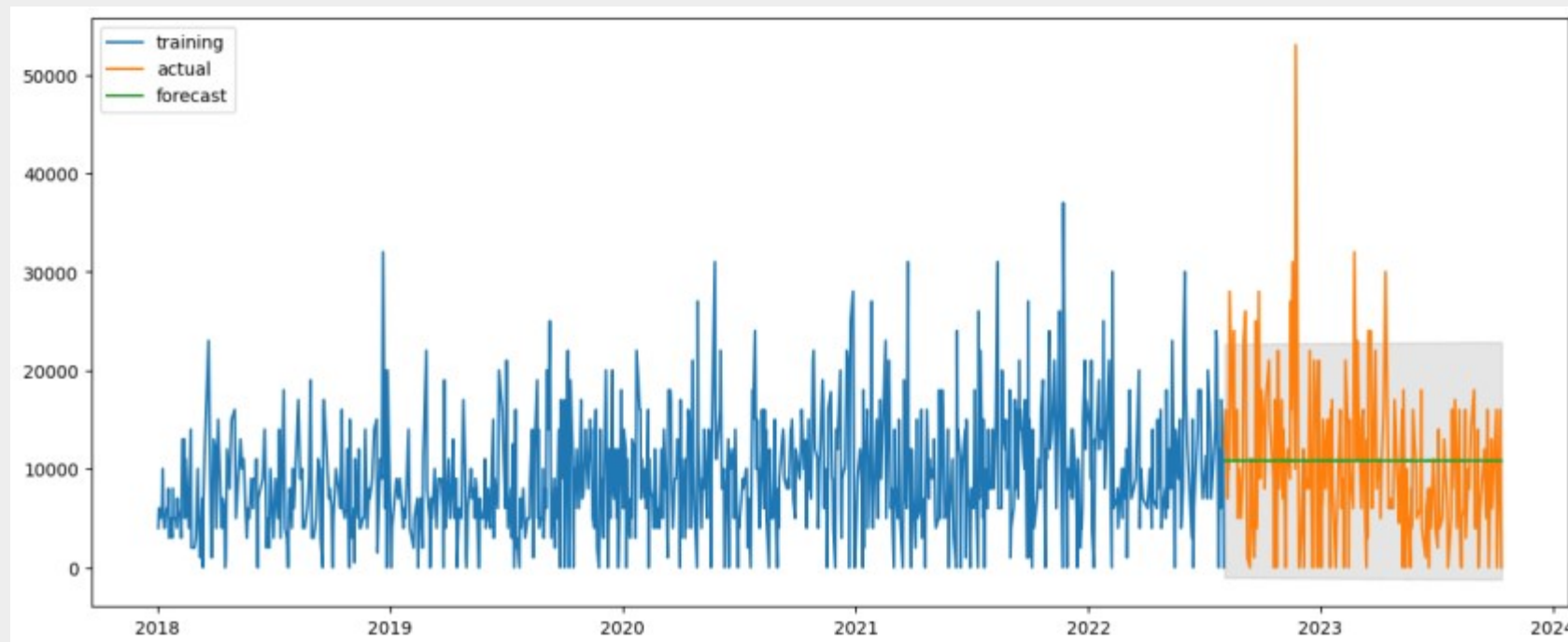
- 휴일이 포함된 경우 예측 데이터의 평균값이 눈에 띄게 줄었음을 확인할 수 있다.
- 이는 휴일을 위해 보관된 '0' 값이 학습에 영향을 주었기 때문이다.
- 따라서 적정 예측값을 올바르게 구하기 위해서는 휴일 데이터를 제거하는 것이 바람직하다.

	MG0921	MG0921_holiday	MG1342	MG1342_holiday
2주	17732	6395	14304	7525
3주	26598	9593	21456	11288

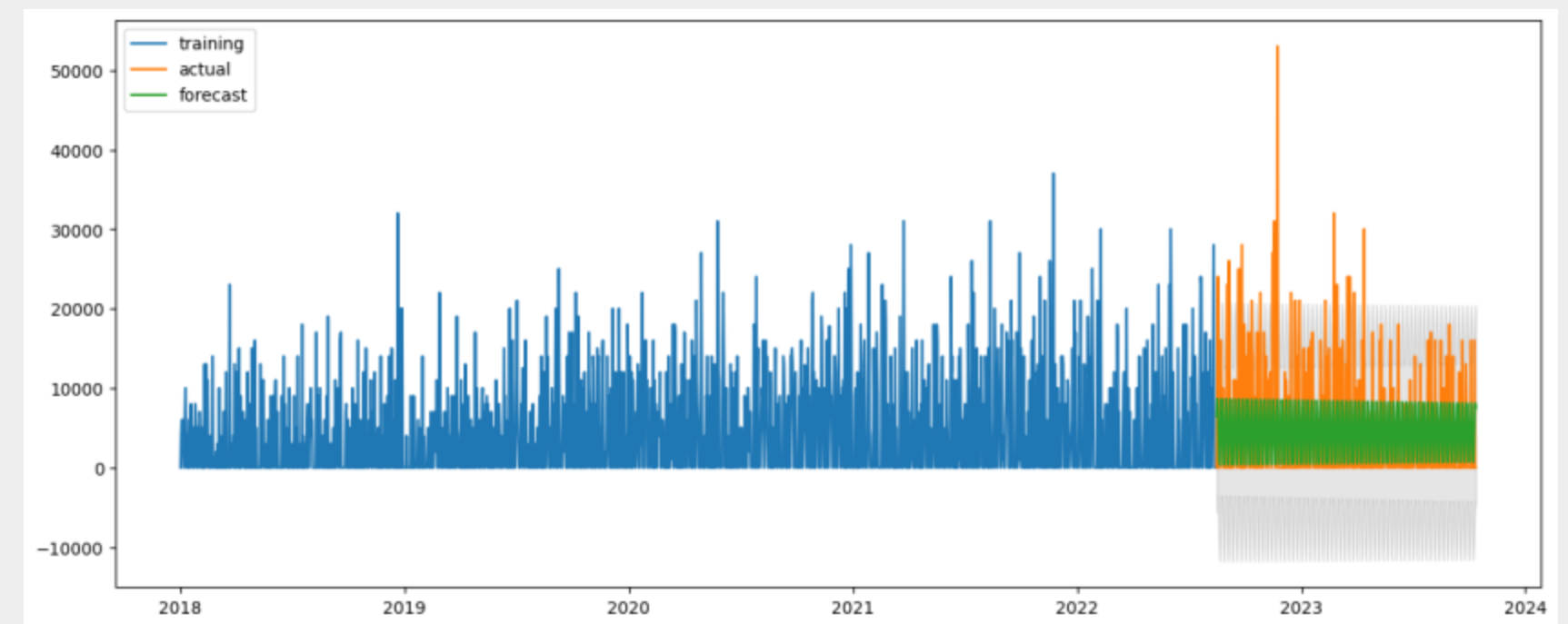
02) 휴일 유무가 가지는 영향 확인

휴일 유무에 따른 예측 데이터 변화

MG0921



<휴일이 포함되지 않은 데이터의 예측 결과>

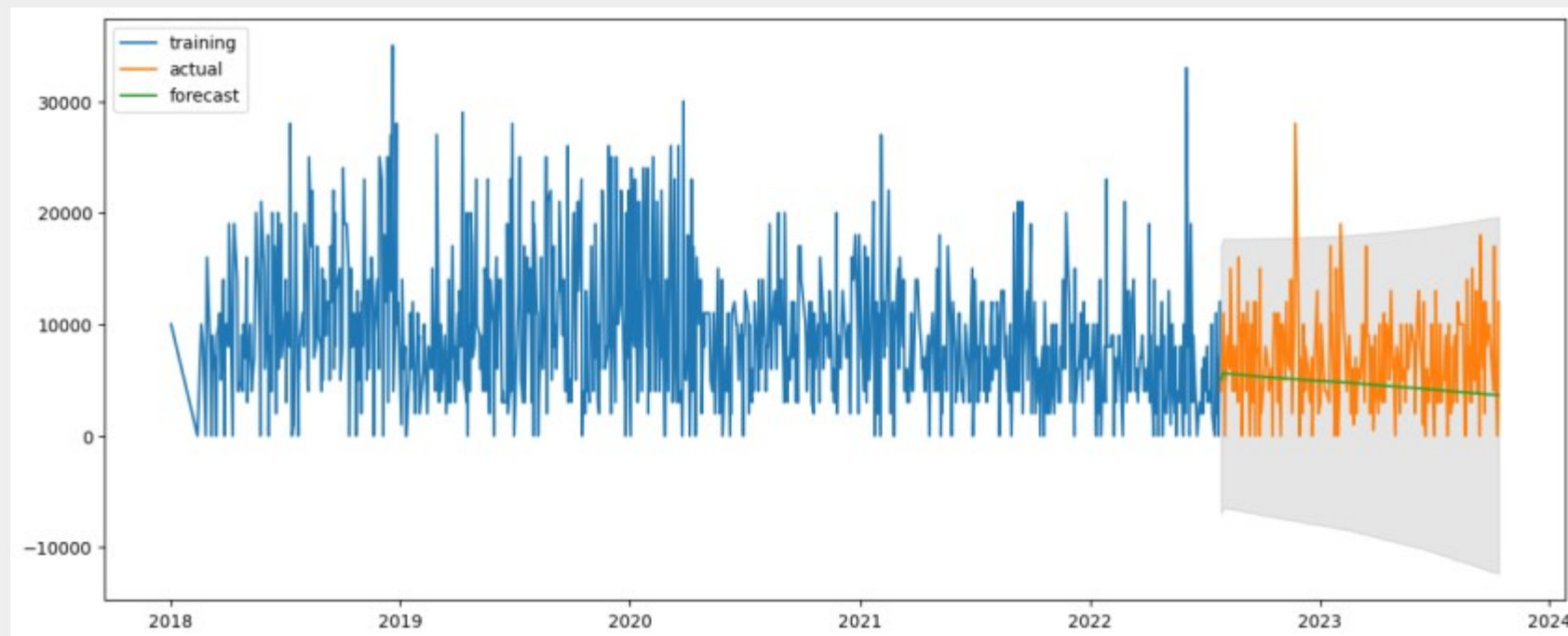


<휴일이 포함된 데이터의 예측 결과>

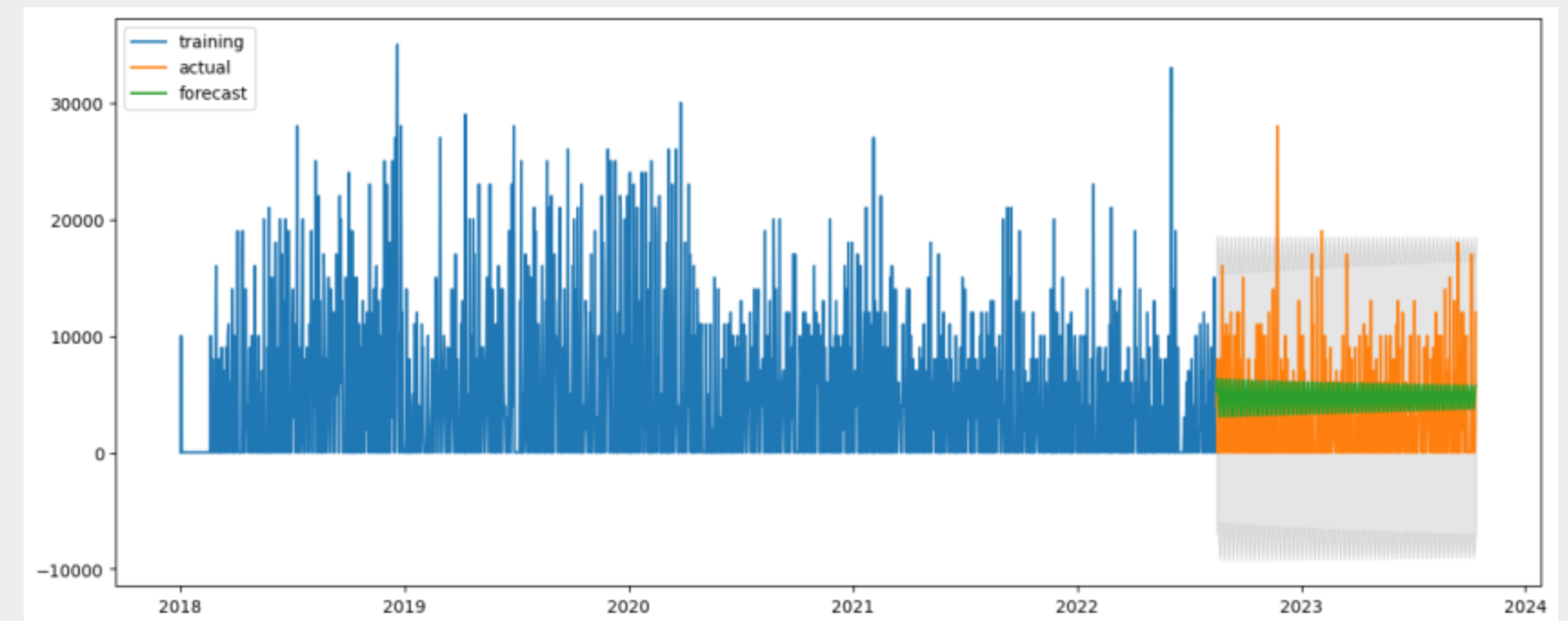
02) 휴일 유무가 가지는 영향 확인

■ 휴일 유무에 따른 예측 데이터 변화

■ MG1342



<휴일이 포함되지 않은 데이터의 예측 결과>



<휴일이 포함된 데이터의 예측 결과>



ARIMA 모델을 활용한 상품의 적정 재고량 분석 보고서



동아대학교
DONG-A UNIVERSITY



DATA
SCIENCE
LABS