



항만 선박 대기시간 예측 및 운영 시스템 개선 제안

Vessel Trackers

류민승, 김신우, 김성원, 한규동

목차

항만 선박 대기 시간

예측 모델링 및 운영 시스템 개선 제안

서론

- 01 주제 선정 배경
- 02 주요 용어 및 프로세스
- 03 데이터 수집
- 04 데이터 자료 구조
- 05 분석 프로세스
- 06 선행 연구 분석

본론

- 07 입출항 프로세스 및 시점 재정의
- 08 1차 데이터 전처리
- 09 Target 항구 선정
- 10 울산항 데이터 추가 수집 및 2차 전처리
- 11 EDA
- 12 Feature Engineering
- 13 Model Fitting
- 14 Model Prediction

결론

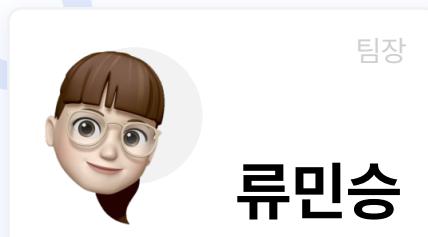
- 15 대기 감소 효과 및 운영시스템 개선 방안

부록

- 01 Tableau 시각화
- 02 한계점
- 03 참고문헌
- 04 일정

팀원 소개

Vessel Trackers



팀장

류민승

- 프로젝트 프로세스 관리
- 데이터 수집 및 시각화
- EDA
- 모델링
- Feature Engineering
- 분석 결과 활용 시뮬레이션

<https://everyday-joyful.tistory.com>

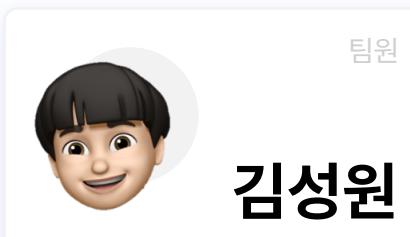


팀원

김신우

- 데이터 수집
- 데이터 전처리 (1, 2차)
- 모델링
- GCP 데이터 적재 및 활용
- 분석 결과 활용 시뮬레이션

<https://datapilots.tistory.com>

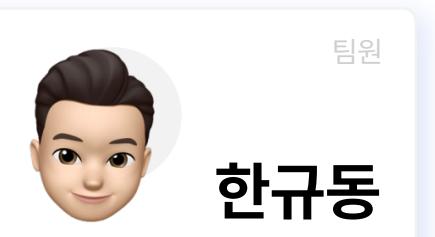


팀원

김성원

- 데이터 수집 자동화
- 데이터 전처리 (1차)
- EDA
- GCP 데이터 적재 및 활용
- 대시보드 작성

<https://castle1data.tistory.com/>



팀원

한규동

- 데이터 수집 및 시각화
- EDA
- Feature Engineering
- 모델링
- 모델 결과 활용 시뮬레이션

<https://bmil2011s.tistory.com/>

1. 주제 선정 배경

대한민국 항만 경쟁력 확보의 핵심 지표인 선박 대기시간 예측 및 관리의 중요성 증가

1. 항만 경쟁력 확보를 위한 효율적 운영 필요

국제 물동량 회복과 국내 해상 물동량 비중

- 21년 3분기 이후 코로나 팬데믹 이전 세계 물동량으로 회복됨
- 우리나라 수출입 물류비중 중 해양 운송이 전체의 **99.7%** 차지
- 항만의 역할이 과거와 마찬가지로 더욱이 중요하므로,

항만 경쟁력 확보를 위한 데이터 기반

항만 운영 프로세스의 선행 필요성 대두

99.7%의 중요성, 해운을 바로알자

등기고 | ① 입력 2022-02-02 1850 | ② 수정 2022-02-02 1850 | 링크 4

바다에 대한 인식도 조사에 따르면 우리 국민들은 수산물, 해양환경 등에 대해서는 관심도가 높으나 해운에 대해서는 상대적으로 관심도가 낮다. 그러나 실제 해운이 국민경제와 세계경제에서 차지하는 비중은 절대적이다. 정부통계에 따르면 작년 우리나라 총 수출입 물동량 9억5800만톤 중 해상 물동량이 9억5500만톤을 기록하며 그 비중이 99.7%를 차지하고 있다.

2. 선박 대기시간의 예측의 경제적 편익

물류 비용 절감의 핵심, 선박 대기시간

- 항구 운영 효율성 향상 및 선박 회전율 증가에 따른 **물류 비용 절감 및 항구 경쟁력 증대**
- 선박 대기 간 발생하는 **추가 비용 절감**
- 광양항 항만자동화 테스트베드 구축(KDI, 2022)의 예비 타당성 조사에서 **선박 대기 관련 편익이 전체의 83% 차지(약, 214억)**

〈표 1-1〉 항만개발사업의 경제적 편익항목

표준지침 제1판(2000)	표준지침 제2판(2001)	표준지침 제3판(2014)
• 선박대기비용 절감효과	• 선박대기비용 절감효과	• 선박대기비용의 절감
• 선박재항비용 절감효과	• 선박재항비용 절감효과	• 선박재항비용 절감
• 하역비용 절감효과	• 하역비용 절감효과	• 하역비용 절감효과
• 내륙운송비용 절감효과	• 내륙운송비용 절감효과	• 선박운항비용의 절감효과
	• 화물운송시간 절감효과	• 내륙운송비용 절감효과
	• 토지조성효과	• 화물운송시간 절감효과
		• 흰적화물 수입증대효과
		• 토지조성효과

자료: KMI(2022), 항만부문 경제적 편익항목의 개선방안 연구

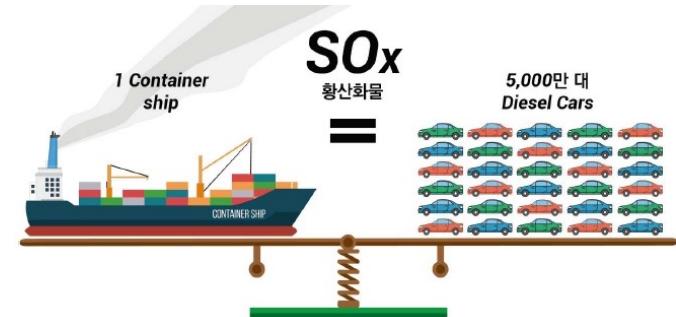
1. 주제 선정 배경

선박 대기시간 예측 정확도 향상이 대기 오염 저감에 미치는 영향

3. 선박 대기(정박) 간 발생하는 오염물질 절감

선박 대기 시간 감소의 환경오염 저감 효과

- 정박 시 연료소비량은 선박 정지 시 전기, 펌프, 그 외 선박 **설비의 가동을 위해 드는 연료**, 또는 **항내에서 양하/적하 작업이 없을 때 소비하는 연료**의 소비량을 포함
- 일반적으로 정박 시 발생하는 연료 소비량은 운항 시 발생하는 소비량의 **20%**로 산정 (보정계수 : 0.2)
- 선박 대기시간 감축 시, 선박 대기 간 발생하는 오염물질 절감 가능**



접안 시 연료 소비량(ton, 화물선)

접안 시 연료소비량(ton) (Eq. 2)

$$= \sum (\text{톤급별 입출항대수} \times \text{운항거리(km)}) / \text{톤급별 연료경제(km/kL)}$$

대기 시 연료 소비량(ton, 화물선)

정박 시 연료소비량(ton) (Eq. 3)

$$= \text{톤급별 입출항대수} \times \text{연료소비계수(ton/일)} \times \text{정박시간(일/회)} \times \text{보정계수(0.2)}$$

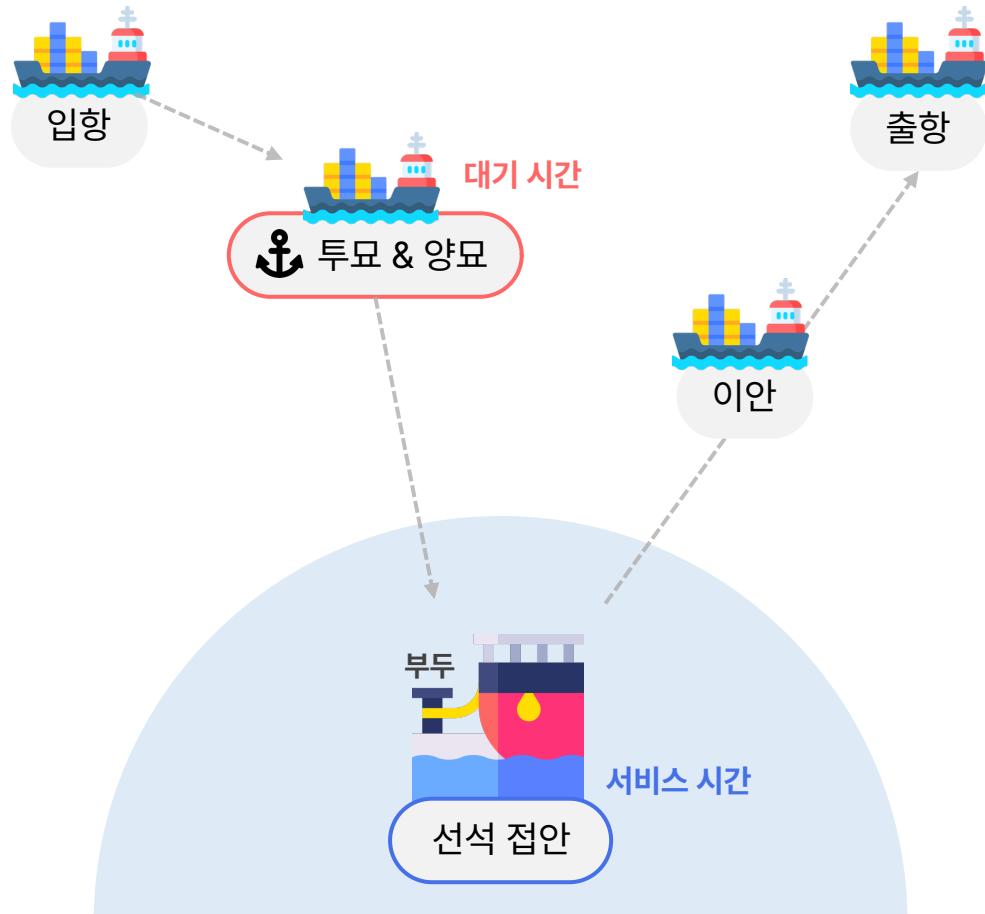
선박 종류 별 오염물질 배출량

Sources	(Unit: ton/yr)					
	CO	NO _x	SO _x	PM-10	PM-2.5	VOC
Passenger ships	692	7,361	1,249	255	233	259
Cargo ships	8,033	85,768	36,699	4,904	4,447	2,970
Fishing vessels	45,641	58,564	519	1,772	1,698	14,773
Total	54,366	151,693	38,467	6,931	6,378	18,002

2. 주요 용어 및 프로세스

선박 입항부터 출항까지의 기본 프로세스의 구성과 개념

선박 입출항 프로세스



용어	설명
투묘 / 양묘	<ul style="list-style-type: none">투묘 : 닻을 내리는 작업양묘 : 닻을 올리는 작업
묘박 (Anchor)	<ul style="list-style-type: none">묘박 : 선박이 해상에서 닻을 내리고 운항을 정지하는 것묘박 장소 : 묘박지
접안(Berth) / 이안(Unberth)	<ul style="list-style-type: none">접안 : 선박이 하역 등을 위해 부두에 배를 대는 것하안 : 선박이 부두에서 떠나는 것
선석*	<ul style="list-style-type: none">항내에서 선박을 계선시키는 시설을 갖춘 접안 장소하나의 부두에 여러 개의 선석이 존재
대기 시간	<ul style="list-style-type: none">통상적으로 선석 부족으로 정박지에 대기한 시간 (투묘시간에서 양묘시간까지 차이를 사용하여 계산)
서비스 시간	<ul style="list-style-type: none">통상적으로 접안하여 하역 등 서비스를 부두에서 실시하고 떠나기까지의 시간으로 정의 (접안시간에서 이안시간까지 차이를 사용하여 계산)

3. 데이터 수집

PORT-MIS 항만 데이터 및 해양 구역별 기상 정보, 시설사용허가 데이터 수집 프로세스

구분	데이터	대상 항구	수집 년도
타겟 항구 선정 시	시설사용허가현황 및 입출항현황	부산, 울산, 인천, 광양, 평택	22
타겟 항구 선정 이후 (울산항)	전체 데이터	울산	20,21,22

데이터명	부산	울산	인천	광양	평택
시설사용허가현황	92,726	50,237	27,866	19,386	14,345
입출항현황	44,305	48,808	29,777	23,213	16,693

*타겟 선정 시, 항구별 데이터 rows 개수

데이터명	수집 방법	상세	컬럼	형태	rows 개수	출처
시설사용허가현황	파일 다운로드	PORT-MIS 시설사용허가현황	순번, 입항횟수, 호출부호, 시설사용횟수, 신고톤수, 선박명, 신청시설, 신청일시, 사용목적명(접안대기)	xlsx	144,823	해양수산부 PORT MIS
입출항현황	파일 다운로드	PORT-MIS 선박 입출항현황	순번, 입항횟수, 호출부호, 선박명, 입항일시, 출항일시	xlsx	148,648	해양수산부 PORT MIS
선박제원정보	동적 페이지 크롤링	PORT-MIS 선박 운항 선박제원정보	선박국적, 선박국적코드, 총톤수, 재화중량톤수, 선박제원(길이,m)	xlsx	7,442	해양수산부 PORT MIS
울산항 선박제원정보	파일 다운로드	울산항 선박제원정보	선박국적, 선박국적코드, 총톤수, 재화중량톤수, 선박 총 길이, 선박 너비	csv	22,795	울산항만공사
관측지점별 기상센서정보	API 크롤링	관측지점별 기상센서정보API	기관명, 지점명, 지점코드, 관측요소(풍향, 풍속, 기온, 습도, 기압)	json	25,329	국립해양측위정보원

*타겟 선정 이후, 울산항 데이터 rows 개수

4. 데이터 자료 구조

선박 고유호출번호와 입항횟수를 기준으로 5개의 데이터셋 결합

ERD



5. 분석 프로세스

항만 대기시간 예측 및 운영시스템 개선 프로세스

항구 선정

데이터 수집 및 병합(1차)

시설사용허가현황 dataset
입출항 현황 dataset

데이터 전처리 (1,2차)

- 대기율, 입출항 프로세스 재정의
- 입출항 사건 별 데이터셋 재구성



울산항 선정

데이터 병합 및 전처리(3차)

선박제원정보 dataset
해양기상정보 dataset

타겟 항구 EDA

울산항 EDA

- 울산항 물동량 분석
- 울산항 액체화물 비중 분석
- 선박용도별(액체/비액체)
대기율&대기시간 분석
- 선박용도별(액체/비액체) 상관 분석

액체화물 선박 단독 모델링

항만 대기시간 예측

Feature Engineering

파생변수 생성

- 선석 기준
- 선박 기준

Model Prediction

RMSE 평가

XGBoost
CatBoost
LightGBM

예측 결과 활용

대기 감소 효과

온실가스 배출량 지표 재구성
선박대기비용 지표 재구성

운영시스템 개선

선석별 점유율 개선 제안



항만 관리 및 운영 효율성 제고

6. 선행 연구 분석

Target 항구 선정을 위한 선행연구(대기율) 검토 및 데이터 부재 확인

■ 대기시간과 대기율

대기율(WR) = 대기시간(T_w) / 서비스시간(T_s)

- 항만의 주요 서비스 지표인 선박 대기율
- 선박 대기율의 정확한 측정과 최적값 유지가 중요함

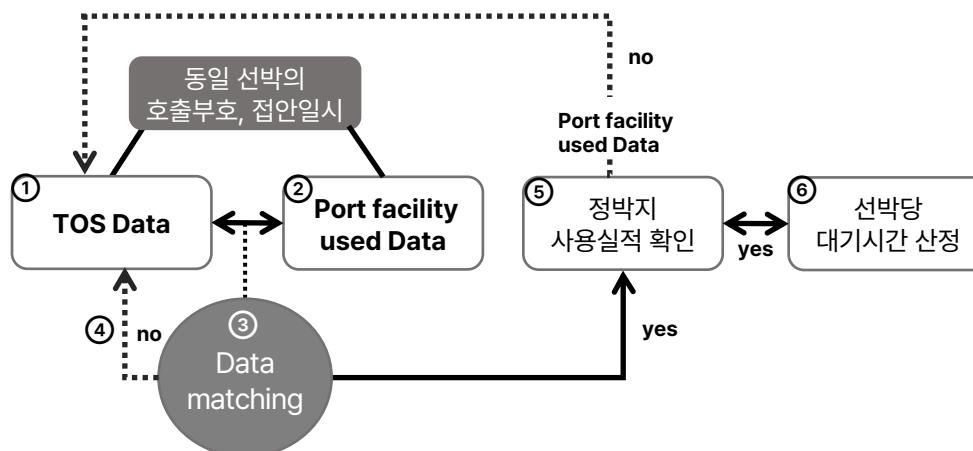


그림7. 실제 대기시간 산정 프로세스

**출처 : 오재용(Jae-Yong Oh);김혜진(Hye-Jin Kim). (2020). 정박 대기시간 예측을 위한 Port-MIS 데이터 분석. 한국항해항만학회 학술대회 논문집, 2020(1), 3-4.

■ 선행연구 활용 데이터 검토

- 선박의 시설사용목적에 따른, 시작 시간과 종료 시간 데이터 활용 (TOS 교통데이터, 2016)
→ 선박 상태(현황)별 세분화된 시점 및 지속 시간을 확인 가능

항만코드	선박명	호출부호	항차	시설사용목적	시설코드	시설명	시설사용시작시간	시설사용종료시간
020	MSC VALERIA	3FFK4	06	Berth Waiting	WAW-01	Anchorage W-1	16.12.24. 23:00	16.12.25. 5:00
020	MSC VALERIA	3FFK4	06	Berth Waiting	MSN-09	New port Terminal2	16.12.25. 6:00	16.12.26 7:00

표7. 부산항 항만시설사용실적 자료

Issue

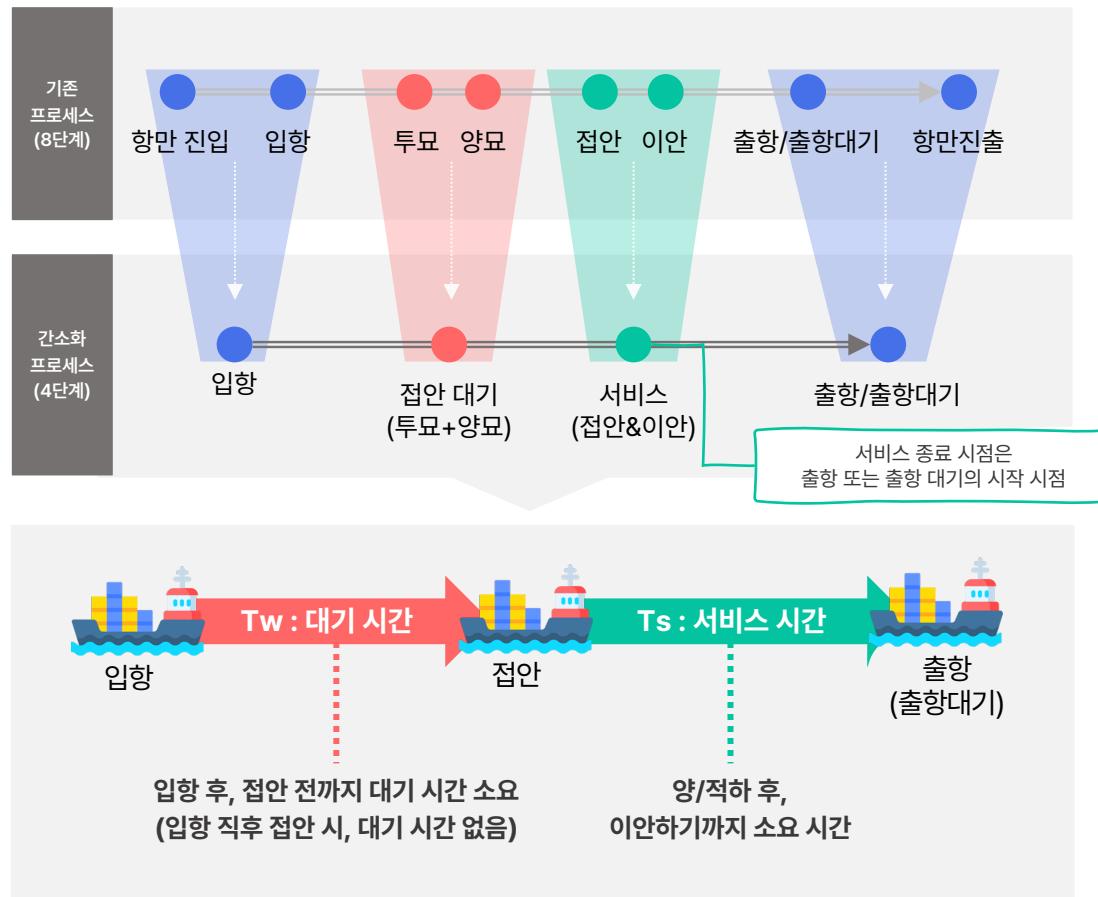
프로젝트 시점(2023.11) 기준 수집 데이터에
시설사용목적별 명확한 시점 및 지속 시간 부재

**출처 : 컨테이너 터미널의 이론적 대기율과 실제 대기율 비교에 관한 연구: 부산항 신항 A 터미널을 대상으로

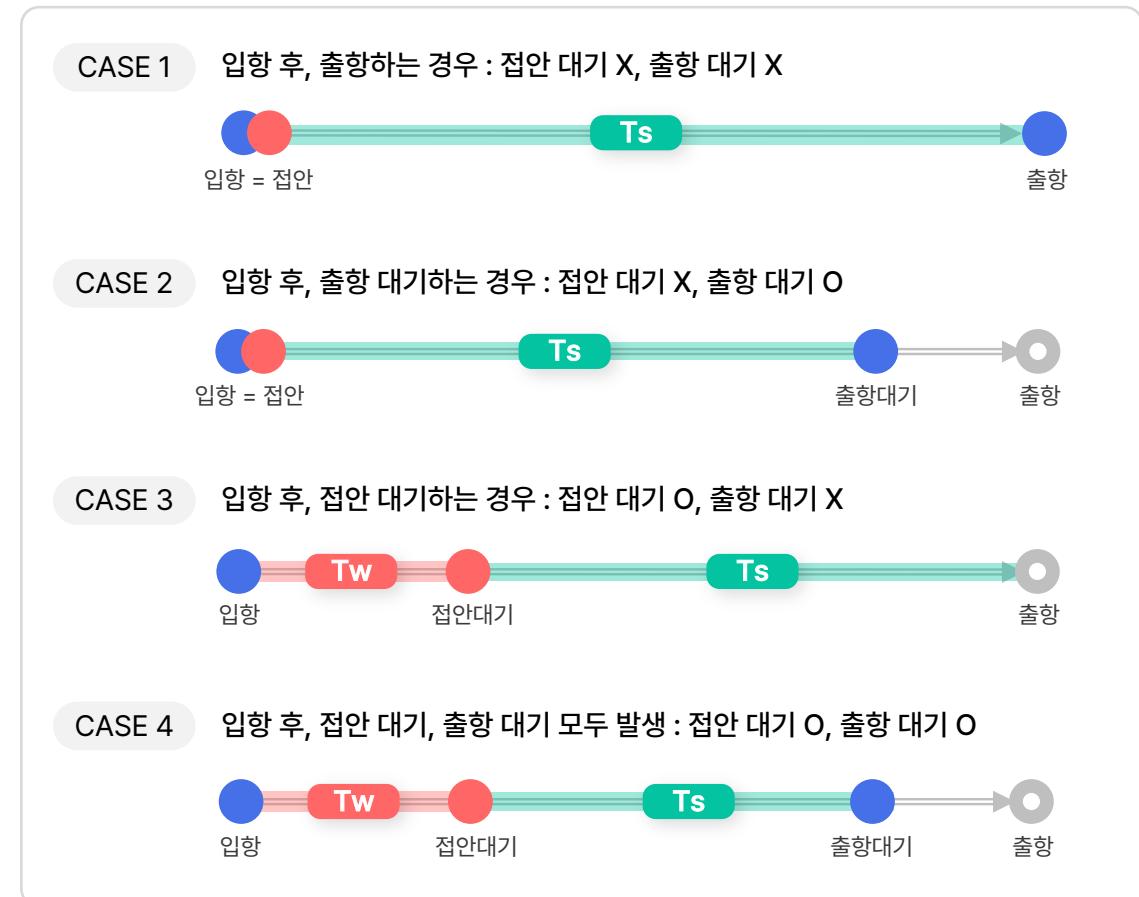
7. 입출항 프로세스 및 시점 재정의

입출항 프로세스 간소화 및 Case별 서비스/대기 시점 재정의

대기율 산출을 위한 입출항 프로세스 간소화



Case 별 서비스/대기 기준 시점 정의



8. 1차 데이터 전처리

선박 호출부호 기준, 입출항 일시를 기준으로 데이터셋 재구성

1. 사용목적명 → 입출항 일시 기준으로 데이터 재구성

* rows : 143,915						
호출부호	선박명	사용목적명	입항일시	Service_Time_Start	Service_Time_End	Service_Time
000308	제21금진호	양하	2020-03-23 17:10	2020-03-23 17:10:00	2020-03-24 15:35:00	0 days 22:25:00
000308	제21금진호	사용목적명	2020-03-09 12:35:00	2020-03-09 12:35:00	2020-03-12 00:00:00	2 days
000308	제21금진호	접안대기	2020-01-05 11:30	2020-01-06 06:15:00	2020-01-06 14:50:00	0 days 08:35:00
000308	제21금진호	적하	2020-01-05 11:30	2020-01-05 11:30:00	2020-01-06 14:50:00	1 days 03:20:00
000308	제21금진호	양하	2020-02-27 08:50	2020-02-27 08:50:00	2020-02-27 16:30:00	0 days 07:40:00

시설사용목적 기준

입출항 1건 당
사용신청이 발생한 횟수만큼
대기시간 & 서비스 시간 발생

입출항 일시 기준

입출항 1건 당
대기시간 & 서비스 시간 산출

- (1) 사용목적명 '접안대기', '출항대기'만 고려
- (2) 선별된 사용목적 건에 대해 '접안 대기시간', '출항 대기시간' 일괄 적용

2. 호출부호, 입항일시 기준 개별 사건 추출

* rows : 72,114							
호출부호	입항일시	출항일시	접안_대기시간_분	출항_대기시간_분	Service_Time_Start	Service_Time_End	계선장소_코드
0 000308	2020-01-05 11:30:00	2020-01-06 14:50	1125	0.0	2020-01-06 06:15:00	2020-01-06 14:50:00	MBO
1 000308	2020-02-27 08:50:00	2020-02-27 16:30	0	0.0	2020-02-27 08:50:00	2020-02-27 16:30:00	MB2
2 000308	2020-03-09 12:35:00	2020-03-12 15:40	0	0.0	2020-03-09 12:35:00	2020-03-12 15:40:00	MB1
3 000308	2020-03-23 17:10:00	2020-03-24 15:35	0	0.0	2020-03-23 17:10:00	2020-03-24 15:35:00	MB2
4 000552	2021-07-03 16:20:00	2021-07-05 04:15	0	0.0	2021-07-03 16:20:00	2021-07-05 04:15:00	MLJ

Groupby Column	Aggregation Column	Aggregation Function	Result
호출부호, 입항일시	접안_대기시간_분	max	<ul style="list-style-type: none"> '접안 대기' 미발생 시, 0(분) '접안 대기' 발생 시, '대기시간(분)'
	Service_Time_Start	max	<ul style="list-style-type: none"> '접안 대기' 미발생 시, 입항일시 '접안 대기' 발생 시, 접안대기 종료 시점
	Service_Time_End	min	<ul style="list-style-type: none"> '출항 대기' 미발생 시, 출항일시 '출항 대기' 발생 시, 출항대기 종료 시점
	출항일시	first	
	계선장소_코드		
	계선장소_숫자		
	계선장소명		
	선박용도		<ul style="list-style-type: none"> 종복 Record 중 첫번째

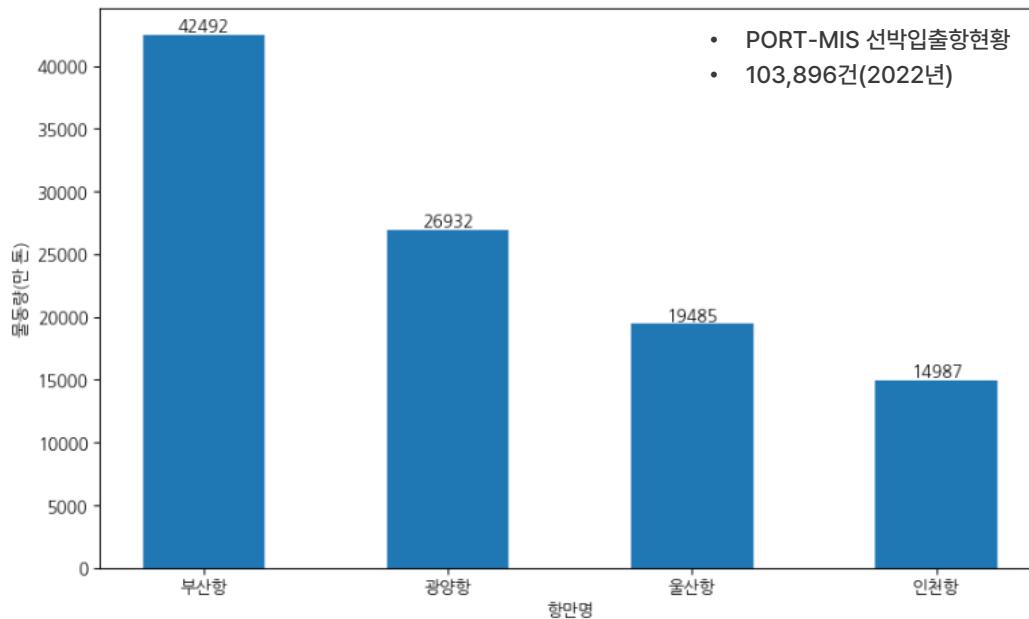
9. Target 항구 선정

분석 항구 선정을 위한 평균 선박 대기율 분석 결과, 14.3%로 가장 높은 **울산항** 선정

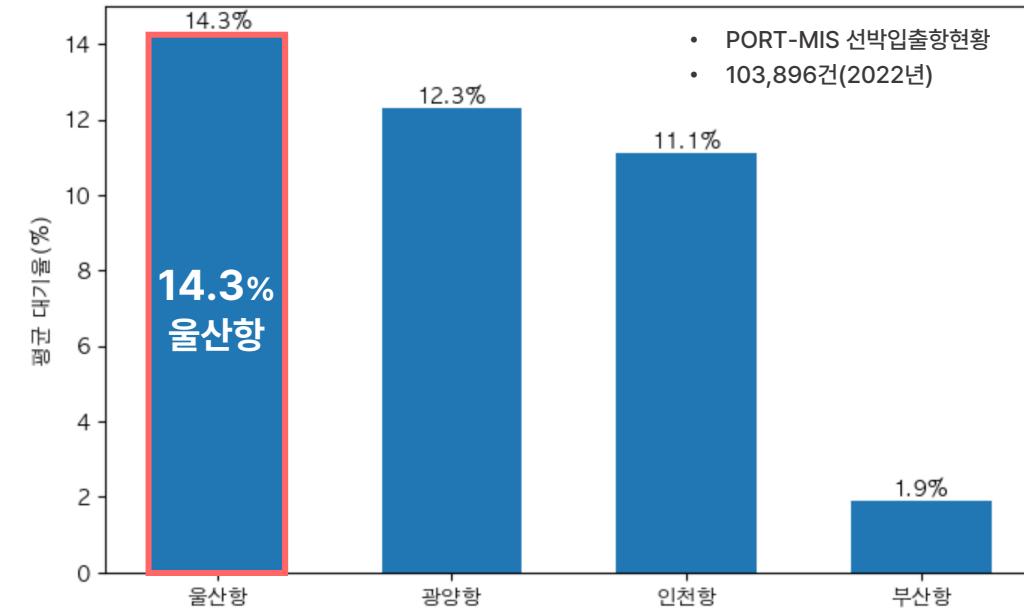
물동량이 높고 대기율이 많이 발생하는 항만의 개선 필요성 높음

■ 항구 후보지 선정 : 부산, 광양, 울산, 인천

▶ 선정 기준 : 2022년 기준 물동량이 많은 4개 항만



■ 후보 항만 대기율 분석 결과



10. 울산항 데이터 추가 수집 및 2차 전처리

울산항 선박제원정보 & 해양 기상 데이터 병합 및 전처리

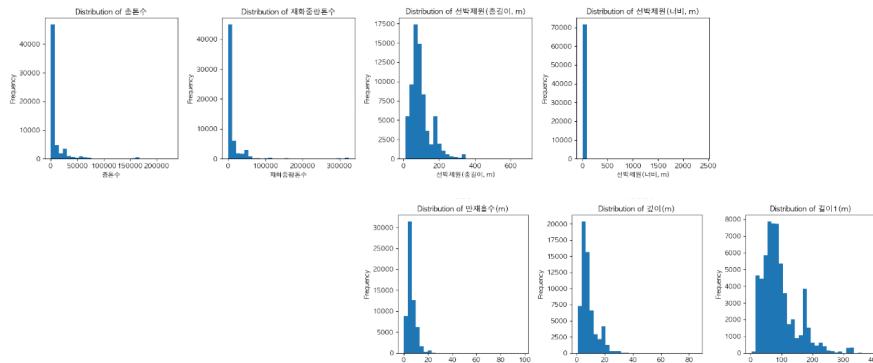
피쳐 특성에 따른 결측치 대체(중앙값, 선형 보간)

1. 선박제원정보 결측치 처리 : 선박용도별 중앙값으로 대체

	총톤수	재화중량톤수	선박제원(총길이, m)	선박제원(너비, m)	만재출수(m)	깊이(m)	길이1(m)
0	NaN	NaN	73.2	12.5	NaN	NaN	NaN
1	NaN	NaN	73.2	12.5	NaN	NaN	NaN
2	NaN	NaN	73.2	12.5	NaN	NaN	NaN
3	NaN	NaN	73.2	12.5	NaN	NaN	NaN
4	141.0	NaN	28.5	8.8	3.5	3.5	28.5
...
65499	31200.0	0.0	206.7	29.7	17.0	20.3	206.7
65500	25742.0	27190.0	193.3	27.0	8.5	15.3	193.3
66189	NaN	NaN	57.8	9.8	NaN	NaN	NaN
66452	12383.0	10055.0	142.0	22.2	8.3	13.6	142.0
71766	28517.0	47047.0	182.5	32.2	12.2	19.1	173.1

선박제원 정보 피쳐	총톤수
	재화중량톤수
	선박제원(총길이, m)
	선박제원(너비, m)
	만재출수(m)
	깊이(m)
	길이1(m)

- Object 타입의 쉼표 제거 후, 숫자형으로 변환(에러값은 NaN)
- 선박제원정보별 분포 확인 → 선박용도별 '중앙값'을 대푯값으로 채택



2. 기상 데이터 결측치 처리 : 선형 보간

	풍속(m/s)	최대파고(m)
풍향(deg)	유의파고(m)	
GUST풍속(m/s)	평균파고(m)	
현지기압(hPa)	파주기(sec)	
습도(%)	파향(deg)	
기온(°C)	수온(°C)	

해양 기상 피쳐	풍속(m/s)	최대파고(m)
	풍향(deg)	유의파고(m)
	GUST풍속(m/s)	평균파고(m)
	현지기압(hPa)	파주기(sec)
	습도(%)	파향(deg)
	기온(°C)	수온(°C)

- 결측값의 **시계열 데이터가 일치하는 행을 제외**하고 선형 보간(method = 'time')
- 중복된 (날짜)행을 무시**하고 내부 데이터만을 기준으로 보간(limit_area = 'inside')
- 보간을 양방향으로 적용하여 **결측값 앞뒤 방향으로** 값을 보간(limit_direction='both')

[Python Code]

```
df.set_index('Datetime', inplace=True)
df[columns_to_interpolate].interpolate
(method = 'time', limit_direction = 'both', limit_area = 'inside', inplace=True)
```

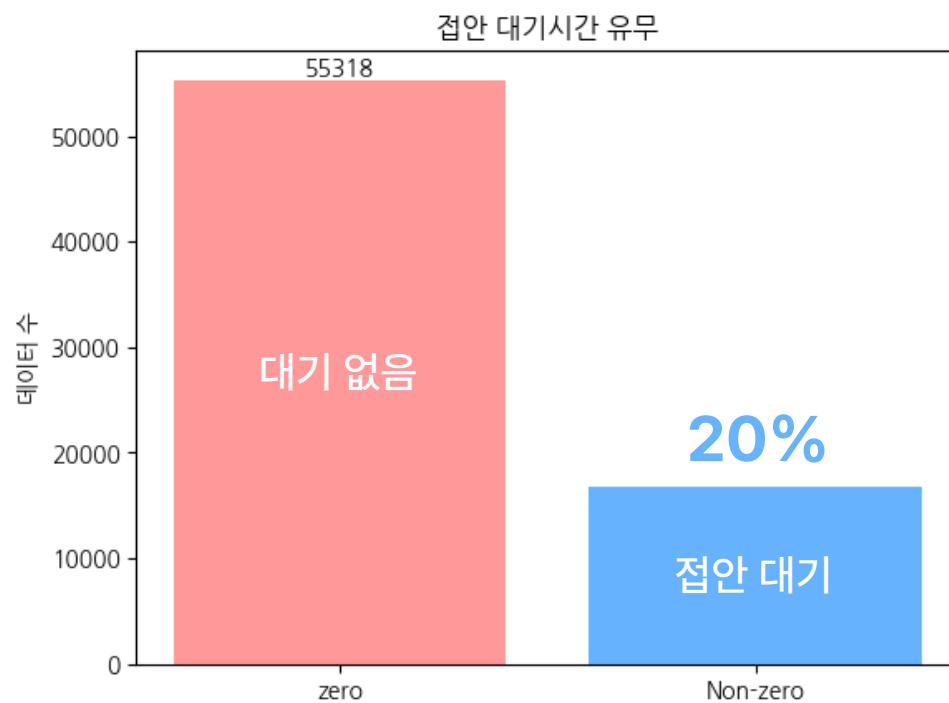
11. EDA

울산항의 선박 대기 발생 비율 20%로 분석 필요성 확인

선박 용도별 방문 횟수 분석 결과 석유화학 제품 운반선 비중 86%

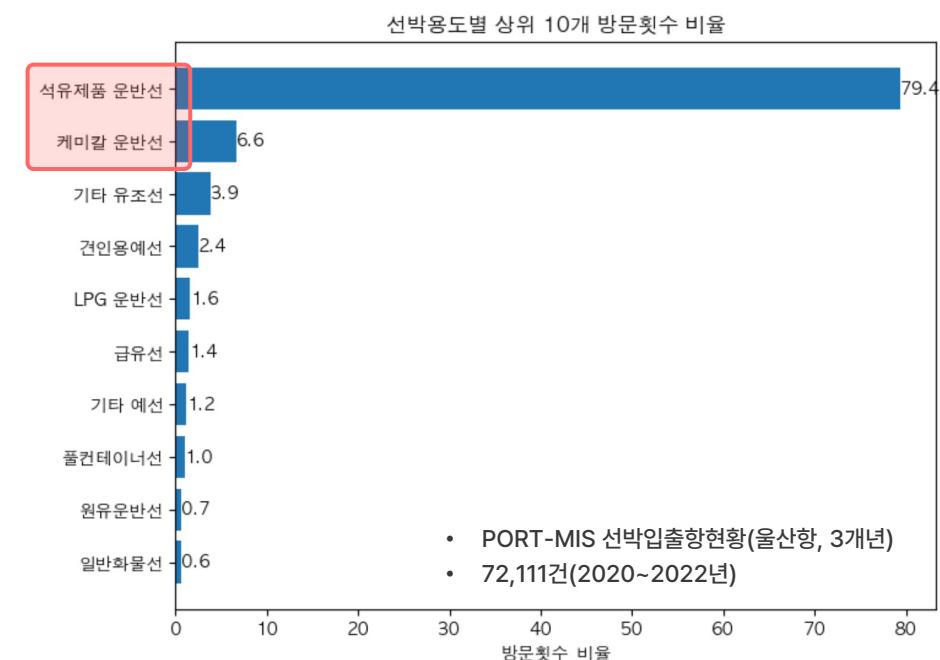
선박 대기시간 발생 비율

전체 입항 건(72,111건) 중
약 20%(16,793건)의 접안 대기 발생



울산항 선박 용도별 방문횟수

울산항은 지역 산업 특성상 석유제품 운반선(69.4%)과
케미칼 운반선(6.6%)이 약 86%를 차지하고 있음



액체화물 선박의 대기시간 감소 시, 울산항 전체 대기시간 감소에 효과적

울산항의 높은 액체화물 선박 수용 비중(70% 이상)을 확인

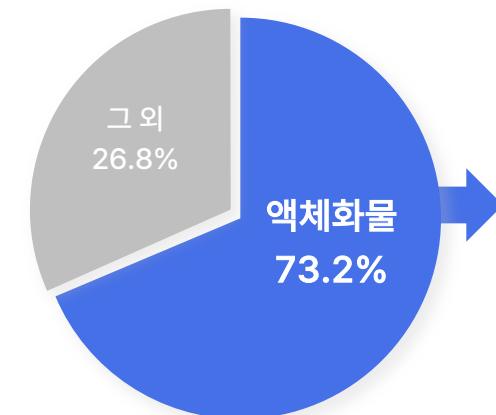
PORT-MIS 화물 품목 코드 기준, 액체화물 필터링

액체 화물 종류	원유	원유(역청유), 석유
	석유제제품	석유제제품
	석유가스 등	석유가스 및 기타가스
	케미칼 등	화학공업생산품
		동·식물성 유지류



액체 화물 관련 선박 (7종)		그 외 선박
석유제제품 운반선	원유운반선	풀컨테이너선
케미칼 운반선	케미칼가스 운반선	견인용예선
기타 유조선	LNG 운반선	일반화물선
LPG 운반선	-	기타 예선, 시멘트운반선 등 총 35건

울산항 액체화물 선박 방문 현황 및 관련 자료



** 울산 석유화학산업 현황, 전국 기준(2021)

- 울산항은 국내 최대의 산업지원항으로, 방문 선박 용도의 **약 73%**가 액체화물
- 대한민국 1위 석유화학 생산지로, **106조원** 가치의 석유화학 제품 생산
- 아시아 최대 석유화합산업단지 내 SK-에너지, S-OIL 등

세계적 규모의 정유공장 보유

**울산 투자환경 HP "석유화학산업" (2023.11.17 기준 데이터)

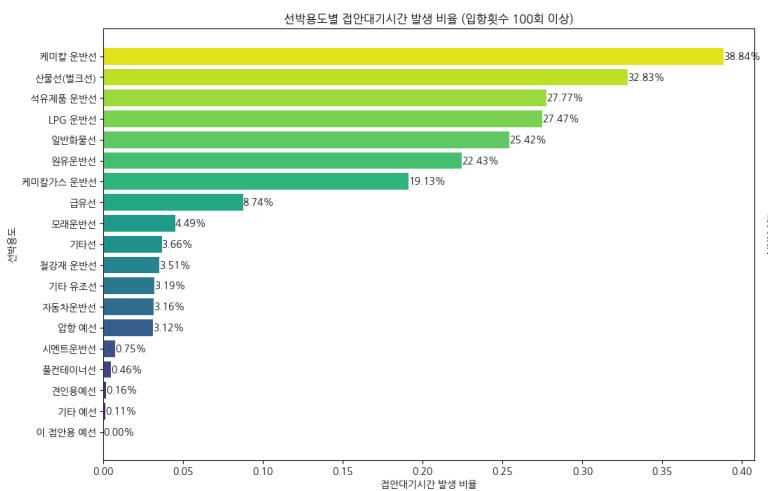
11. EDA

울산항 대기시간 예측시 액체화물 선박 집중 필요성 높음

액체화물 선박의 대기시간 및 대기율 분석 결과, 기타 선박 용도 대비 높은 수치를 확인

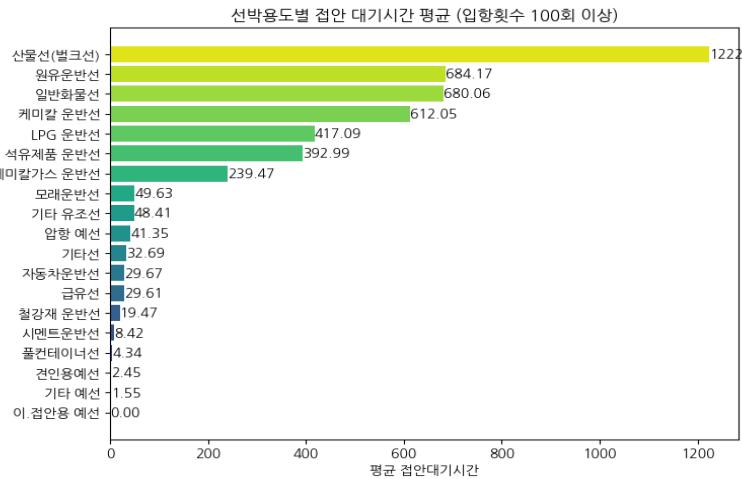
선박용도별 접안대기시간 및 대기율 분석 (입항횟수 100회 이하 선박용도 제외)

1. 대기 발생 비율



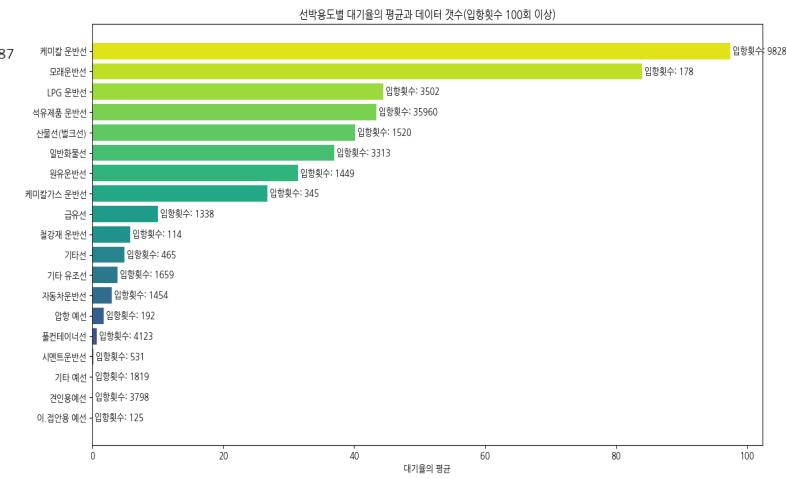
액체화물선 접안대기 발생률
평균 22% 이상

2. 평균 대기 시간



액체화물의 접안대기시간
평균 300분 내외로 상위에 속함

3. 평균 대기율



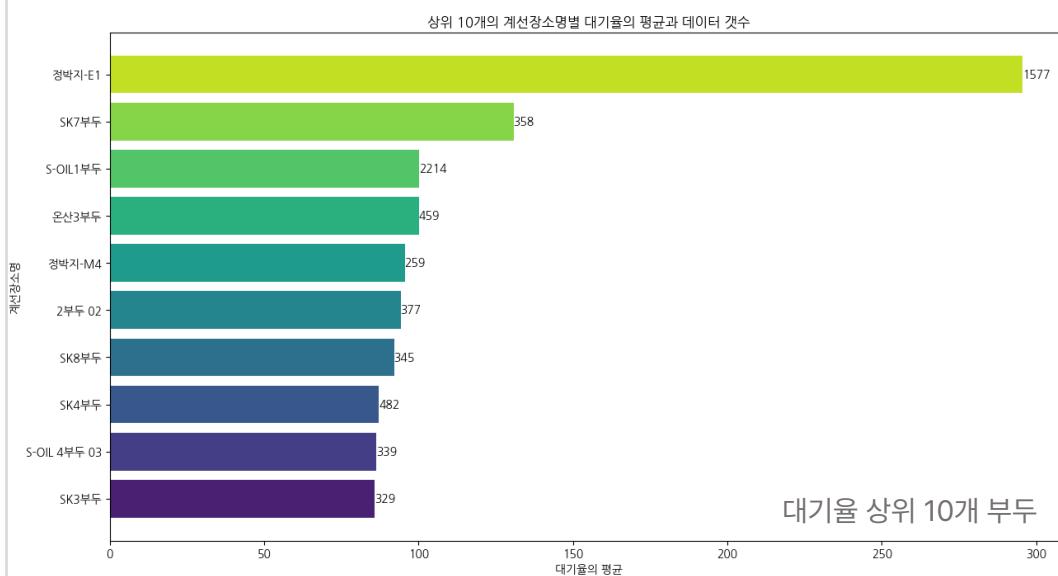
액체화물의 대기율
45%로 타 화물 대비 상위에 속함

액체화물 선박의 '분(Minute)' 단위 대기시간 예측 진행 결정

울산항 계선장소 기준 대기율이 높은 3곳의 평균 대기율 90%, 주요 취급화물은 액체화물

계선장소별 평균 대기율

정박지-E1 을 제외한 상위 부두의 평균 대기율 약 90%



대기율 상위 3위 부두의 주요 취급 화물

- 상위 3개의 부두의 주요 취급 화물이 **액체화물**(연료, 화학물질)
- 정박지-E1는 선박들이 임시 경유지로 추정됨 (울산항 공식 데이터 부재)

구분	SK 7부두	S-Oil 1부두	온산 3부두
소재지	울산광역시 남구 용감동 수역	울산광역시 울주군 온산읍 처용리 수역	울산광역시 울주군 온산읍 산암리 143-20
준공년도	1990	1991	1991
준공주체	민간	민간	국가
안벽길이	370m(강관돌핀)	391m(돌핀)	230m(잔교식)
수심	15m	11m	12m
접안능력	130,000DWT / 1척 20,000DWT / 1척	40,000DWT / 1척 20,000DWT / 1척	20,000DWT / 1척
하역능력	25,000Bbls	965천 톤	816천 톤
운영사	SK에너지(주)(usmdumuy)	오드펠터미널코리아 (usmdw23)	고려아연(주)(usmbo3)
주요취급화물	유류, 케미칼 등	화학공업생산품, 석유정제품 등	석유정제품, 광석 등

**울산항만공사. "부두현황". 검색일 : 2023.11.17

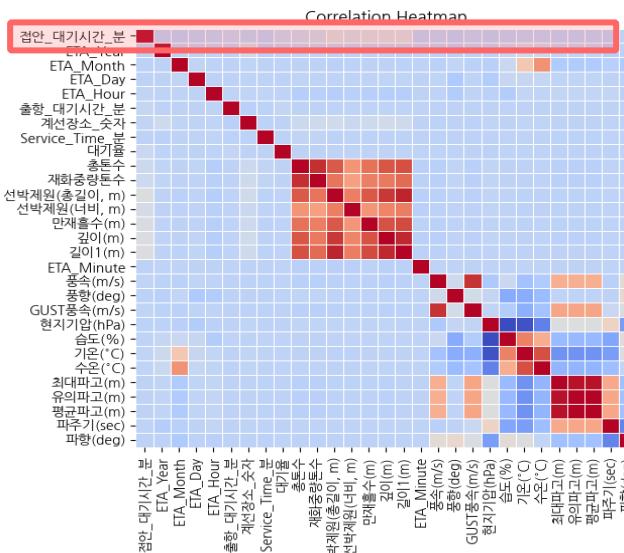
11. EDA

상이한 선박 데이터 존재로 인한 변수 간 상관성 약세 확인

상대적으로 상관계수가 큰 선박 제원과 밀접한 선박 용도로 데이터 분리 시도 필요

변수 간 상관분석

▶ 수치형 피쳐와 접안 대기시간과의 단순 상관관계



모든 변수들의 상관계수 절댓값은 0.2 미만

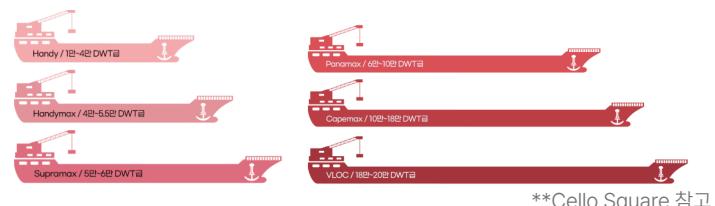
▶ 상관계수(절댓값) 상위 10개 Features

변수명	상관계수
선박제원(총길이, m)	0.161
길이1(m)	0.158
깊이(m)	0.157
만재흘수(m)	0.143
선박제원(너비,m)	0.119
재화중량톤수	0.078
총톤수	0.064
대기율	0.058
출항_대기시간_분	0.035
ETA_HOUR	0.016

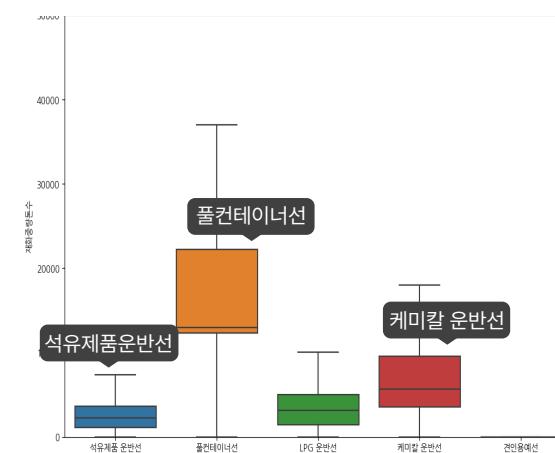
선박의 크기 관련 피쳐(길이, 깊이)에서
양의 상관관계를 확인 (0.15 ~ 0.16)

■ 선박 크기(재화 중량 톤수)는 선박용도에 따라 다름

선박 크기를 나누는 대표 기준 → 재화중량톤수



선박 용도별 재화중량톤수 차이가 있음



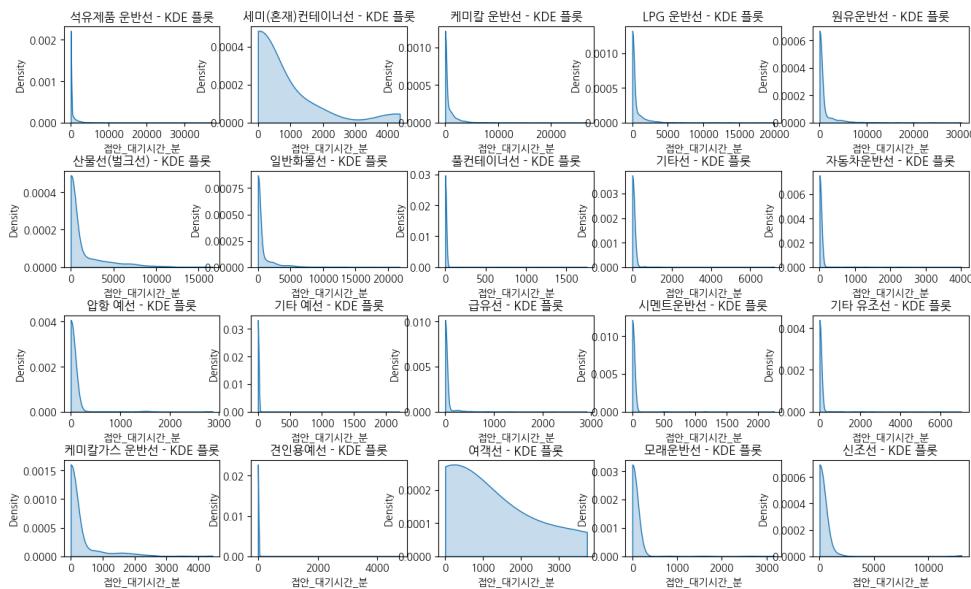
11. EDA

모델 정확도 향상을 위한 선박 유형별 분류 후 변수 상관성 재확인

선박용도(액체/비액체)에 따른 분류가 변수 상관성을 가장 잘 개선함

1차 시도 : 선박 용도별 대기시간 분포를 활용한 그룹화 (K-S 검정)

- 외형적으로 3개의 그룹 확인 (장시간 대기한 호출부호 건의 분포에 의함)
- 분포 동일성 검정 결과 전체 중 95%가 같은 분포로 확인
→ 대기시간이 0인 레코드가 전체의 80%로 검정에 큰 영향을 미침



2차 시도 : 화물 종류 분류(액체화물 / 비액체화물) 후 상관분석

- 상관계수 상위 5개 피쳐 확인,
분류 전 대비 평균 38%의 상관계수 상승
→ 액체화물 / 비액체화물 개별 모델링 고려

화물 종류 분류 전 상관계수

변수 명 (상위 변수)	상관계수
접안_대기시간_분	1.000
선박제원(총길이, m)	0.161
길이1(m)	0.158
깊이(m)	0.157
만재흘수(m)	0.143
선박제원(너비,m)	0.119
재화중량톤수	0.078
총톤수	0.064
대기율	0.058
출항_대기시간_분	0.035
ETA_HOUR	0.016

화물 종류 분류 후, '액체화물' 상관계수

변수 명 (상위 변수)	상관계수
접안_대기시간_분	1.000
선박제원(총길이, m)	0.239
길이1(m)	0.227
깊이(m)	0.201
만재흘수(m)	0.187
선박제원(너비,m)	0.133
재화중량톤수	0.187
총톤수	0.169
대기율	0.058
출항_대기시간_분	0.031
ETA_HOUR	0.010

12. Feature Engineering

선행연구를 토대로 선박 대기시간에 영향을 미치는 변수 추가

선석/선박으로 기준을 나누어 파생변수 생성

기본 변수 (선박용도 : 액체화물 한정)

기준 호출부호(선박ID) 입항 1건		* Rows : 51,001건	
선박 제원 정보	ETA_Year	기상정보	풍속
	ETA_Month		풍향
	ETA_Day		GUST풍속
	ETA_Hour		현지기압
	총톤수		습도
	재화중량톤수		기온
	선박_총길이		수온
	선박_너비		최대파고
	선박_만재흘수		유의파고
	선박_깊이		평균파고
선박 대기 관련 지표	선박_길이1		파주기
	대기율		파향
선박 대기 시간		Target Label	접안_대기시간_분

파생변수 (선박용도 : 액체화물 한정)

기준 호출부호(선박ID) 입항 1건		* Rows : 51,001건			
이전_입항선박수	연도별_선석점유율	시설연평균_재화중량톤수	연평균_총입항건수	선박_연평균_서비스시간	선박_연평균_대기시간
7	1.201603	1449.208295	477.000000	653.765957	0.000000
8	1.305933	6888.822695	217.000000	1245.671296	0.000000
9	0.872777	2011.979829	299.333333	625.326923	0.000000
10	1.201603	1449.208295	477.000000	696.228018	0.000000
11	1.970294	5312.995470	541.333333	1683.624383	239.645475
12	4.143219	2073.888986	1729.000000	860.698065	662.278253
13	1.114249	2816.871812	215.333333	152.526224	0.000000
선석 기준 변수 생성		선박 기준 변수 생성			
시간대별 누적 선박 합계	1일 기준, 시간대 별 선석에 입항한 누적 선박 수	선박 평균 서비스 시간	선박(호출부호) 별 3년 평균 서비스 시간		
연평균 선석 점유율	3년 평균 선석 점유율 (= 선석 총 서비스 시간/선 석 서비스 일 수)	선박 평균 대기 시간	선박(호출부호) 별 3년 평균 대기 시간		
연평균 처리량	선석의 3년 평균 처리 재화 톤 수				
연평균 접안 척수	선석의 3년 평균 접안 선박 수	**참고 : 선박 평균 대기 시간은 예측 Target 값(대기시간) 을 포함하여, 독립변수에서 제외, 시각화만 고려			

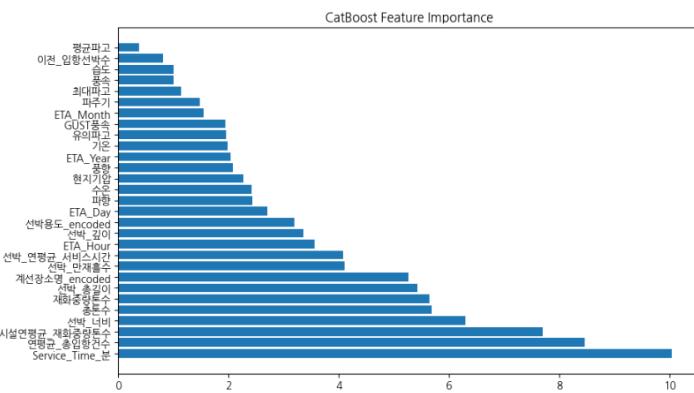
13. Model Fitting

Boosting 기반 3개 모델 결과 확인

액체화물 접안대기시간(분) 예측 결과, 평가지표 오차가 가장 작은 LGBM 채택

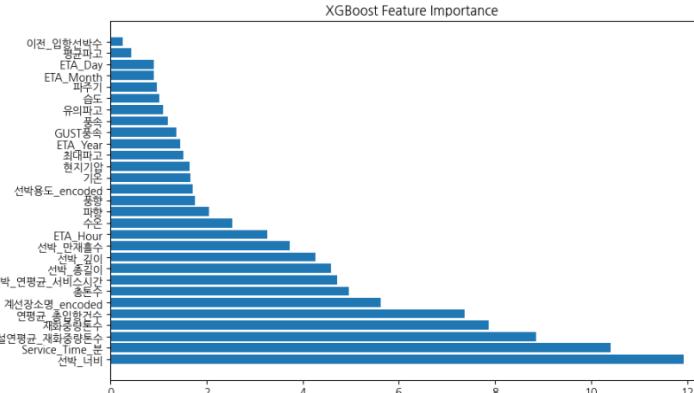
CatBoost

- * 테스트 RMSE (최적 모델): 1096.2605242253724
- * 테스트 MAE (최적 모델): 518.9263379848721



XGBoost

- * 테스트 RMSE (최적 모델): 1050.1287844480644
- * 테스트 MAE (최적 모델): 508.1890787616101

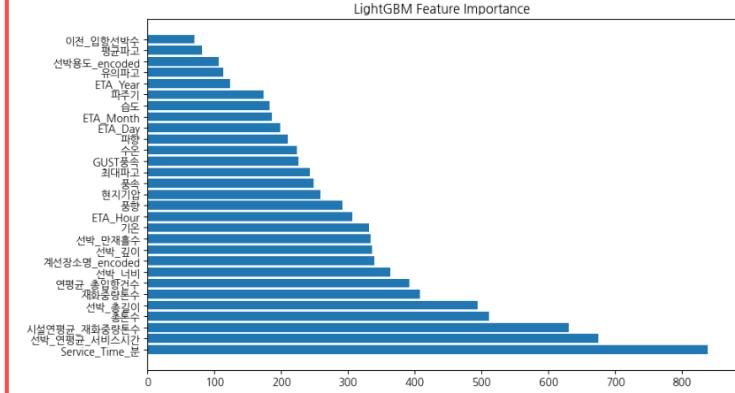


- 범주형 데이터와의 상호작용을 잘 처리하여 정확도와 효율성을 개선
- 모델 복잡성 제어 및 과적합 방지를 위한 L2 정규화
내재, 다른 트리기반 모델보다 뛰어난 일반화 성능
- 대기시간 예측과 같은 복잡한 회귀 문제에 적용 가능

채택

LightGBM

- * 테스트 RMSE (최적 모델): 1028.9960742121095
- * 테스트 MAE (최적 모델): 499.05209516601593



- 많은 피쳐가 포함된 데이터 세트에 적합
- 손실을 가장 크게 줄일 수 있는 노드를 우선적으로 분할하는 리프-와이즈 트리 성장 전략을 사용하여 학습 속도를 단축
- 효율적인 메모리 사용

14. Model Prediction

선박, 선석, 기상정보 기반 예측 모델을 활용한
대기시간 예측 결과 예시 및 대기 감소 효과 지표 도출

대기시간 예측 결과 예



대기 감소 효과 지표 산출식

대기 시, 온실가스 배출량

$$E = \sum (FC \cdot EF)$$

- FC : 연료 소비량(tonnes)
= 연료 소비 계수(ton/minute X 대기시간(minute) X 보정계수(0.2))
- EF : 온실가스 배출계수(kg/tonnes)
→ MDO/MGO 기준 채택

* 정박 시, 온실가스 배출량(IPCC, 2006)을 대기 시, 온실가스 배출량으로 재구성

선석 별 선박 대기비용

$$(Vt \times Wt) \times (WC + FC + EC)$$

- Vt : t년도 선석에서의 선박 입항 수
 - Wt : t년도 선석의 선박당 평균 대기시간(분)
 - WC : (액체화물)의 분당 평균 체선비용
 - FC : (액체화물)의 분당 평균 체화비용
 - EC : (액체화물)의 분당 평균 환경비용
-
- 체선 비용 : 선박이 항구에 체선 시, 발생하는 비용(항구 이용료, 접안료, 대기료 등)
 - 체화 비용 : 선박에 화물을 싣거나 내릴 때 발생하는 비용(화물가치, 인력, 장비 등)
 - 환경 비용 : 선박 운영 시, 환경에 미치는 비용 (대기 시, 온실가스 배출량 사용)

* 선박 대기비용(일)(조아현, 2023)을 액체화물 대기비용(분) 산출식으로 재구성

15. 대기 감소 효과 및 운영시스템 개선 방안

대기시간 예측 결과, 선박 입항 후 즉시 접안 및 서비스 이용 시

대기 감소 효과 및 항만 운영 시스템 개선 방안

선박 비용 및 환경 측면



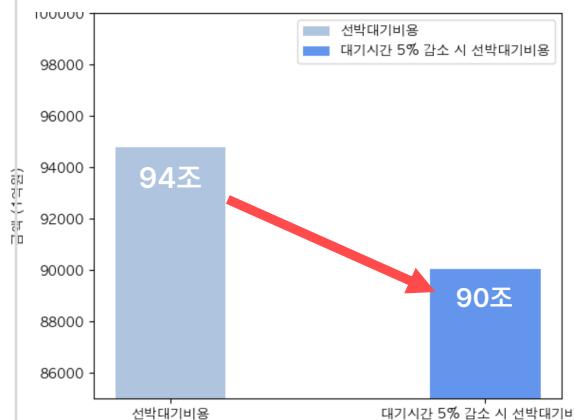
효율적인 체선 및 체화 비용 계획
→ 선박 운영비 감소
→ 에너지 사용 최적화



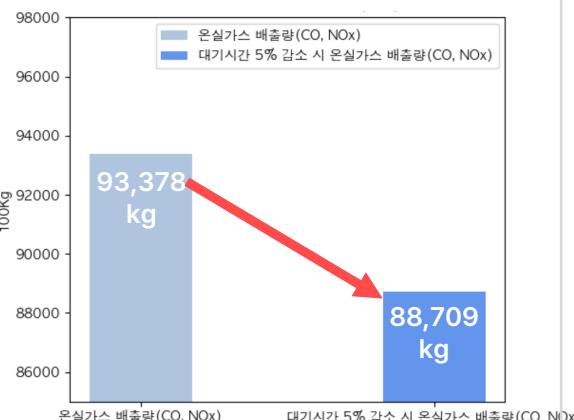
선박 운항 시간 최적화
→ 대기 중 오염물질 배출 감소

연간 대기시간(전 선석) 5% 절감 시 효과

선박대기비용 4조 738억 원 감소



온실가스 배출량 4,668,901 kg 감소



항만 시스템 개선 : 선석 점유율 제고 & 유지 방안



부두 입출항 운영 시간 최적화
→ 연료 및 인력 비용 감소
→ 터미널 이용률 제고



접안 대기시간 감소
→ 선박 운항 스케줄 조정 용이
→ 항만 이용 만족도 증가

대기 중 선박의 제원 정보 및 화물 정보와 빈 선석 정보를 결합(Match)하여 선석 재배정을 통해 점유율 제고

고객 대기 확률을 낮추며
터미널 이용률의 상위 유지를 위한 점유율 80% 달성



**컨테이너 터미널의 서비스 수준 평가지표에 대한 고찰(김창곤, 2016)

15-1. 선박 대기 비용 산출식 도출

선박 대기시간 감소시, 대기비용 산출식 재구성

$$BH = (V' \times W' - V \times W) \times (WC + FC + EC)$$

정박 시 연료소비량 기준식

▶ 기준식 출처 : 선박대기시간 예측모형 개발을 위한 방법론 연구(2023)

- BH : t년도 **선박대기비용 절감효과**
- Vt(Vt') : with case(without case)의 t년도 **선석에서의 선박 입항 수**
- Wt(Wt') : with case(without case)의 t년도 **선석의 선박당 평균 대기시간(분)**
- WC : (액체화물)의 **분당 평균 체선비용**
- FC : (액체화물)의 **분당 평균 체화비용**
- EC : (액체화물)의 **분당 평균 환경비용** → 표준선형 > 액체화물(시간→분)로 변경, 추산

체선비용

체화비용

환경비용

WC 체선비용

= 고정비 + 변동비(유류비)

선박크기(DWT)	시간당비용	일일비용
10,000	383000	9183000
30,000	742000	17811000
50,000	1010000	24235000
150,000	1959000	47006000
200,000	2330000	55910000

FC 시간당 체화비용

= (화물가치 x 사회적 할인율) / 365 * 24

화물가치(kg당 화물가치(원))

- 유류 600(수입), 694(수출)의 평균값 **647**로 계산
- 수입, 수출 품목 데이터에서 구분 어려워, 수입 수출의 평균값으로 화물가치 산정
- 화물가치 = (재화중량톤수/1000) * 647

사회적 할인율 = 4.5%

화물의 시간가치는 화물 금액 가치에 이자율을 적용한 화물의 재고비용 관점에서 추정

- 컨테이너 화물의 체화비용 원단위를 산정하기 위해 경제성 분석의 사회적 할인율 4.5%를 연 이자율로 적용
- **우리나라 항만에서의 체선·체화 시간비용 재추정(2002)

온실가스 배출량(CO, NOx) 기준 사회적 비용

CO(배출량) / 1000 * 158

(온실가스 배출량에서 구한 CO)

NOx(배출량) * 15835

(온실가스 배출량에서 구한 NOx)

15-2. 대기중 온실가스 배출량 산출식 도출

대기 시 연료 소비량 산출식을 분/입출항 1건 기준으로 재구성

대기 시 연료 소비량(분/입출항 1건) = 연료 소비 계수(ton/minute) X 대기시간(minute) X 보정계수(0.2)

■ 선박 배출 대기오염 물질

- 일산화탄소 (CO)
- 질소산화물 (NOx)
- 이산화황 (SO2)
- 미세먼지 (PM)
- 휘발성유기 화합물(NMVOCs)

온실가스

*온실가스에 해당하는 CO, NOx 만 고려(EEA, 2013)

■ 선박 온실가스 배출량 산정 방법

온실가스 배출량
산정 방법

(IPCC 가이드라인 유류통계, 2006)

- Tier 1 연료소비량 및 배출계수 이용 선택
- Tier 2 연료소비량 및 국가 고유 배출계수
(엔진별, 운항모드 별, 연료별로 분류를 이용하는 경우)

Tier 1 산정식

$$E = \sum (FC \cdot EF)$$

- E : 온실가스 배출량(tonnes)
- FC : 연료 소비량(tonnes) → 정박 시 연료 소비량
- EF : 온실가스 배출계수(kg/tonnes)

*선박부분 온실가스 배출량 산정에 관한 연구(2010)

대기 시 연료소비량 기준식

▶ 정박 시 연료소비량 산출식을 재구성

(산출식 출처 : 선박 대기오염 물질 배출량 산정 및 배출 기여도 분석연구, 2021)

정박 시
연료 소비량 = 톤급별 입출항 대수 X 연료 소비 계수(ton/day) X 정박시간(일/회) X 보정계수(0.2)

▶ 연료 소비 계수(ton/day) : 1일 연료소비량(톤) 기준

Table 4. Fuel consumption coefficient of ship gross tonnage.

Gross tonnage	Fuel consumption coefficient (ton/day)	Gross tonnage	Fuel consumption coefficient (ton/day)
≤100	16.363	15,000~20,000	33.763
100~500	16.563	20,000~25,000	38.763
500~1,000	17.013	25,000~30,000	43.763
1,000~3,000	18.263	30,000~50,000	56.263
3,000~5,000	20.263	50,000~60,000	71.263
5,000~7,000	22.263	60,000~75,000	83.763
7,000~10,000	24.763	75,000~100,000	103.763
10,000~15,000	28.763	Above 100,000	116.263

온실가스 배출계수(배출량/연료)

▶ 온실가스 배출 계수(kg/ton fuel) : MDO/MGO 채택

Table 1. Emission factors for ship emissions.

Fuel type	CO	NO _x	SO _x	PM-10	PM-2.5	VOC	(Unit: kg/ton fuel)
Gasoline	573.9	9.4	205*	9.5	9.5	181.5	
MDO/MGO	7.4	78.5	205*	1.5	1.4	2.8	* MDO : 해양 디젤 오일
BFO	7.4	79.3	205*	6.2	5.6	2.7	* MGO : 해양 가스 오일

*S: Sulfur content in fuel (%)

* 선박 대기오염 물질 배출량 산정 및 배출 기여도 분석연구, 2021 / (EEA, 2013)

15-3. 대기 감소 효과

선박 대기 비용 & 온실가스 절감효과

정박 시간 감소에 따른 선박 대기 비용 및 온실가스 배출량 감소 효과 시뮬레이션

대기시간(분) 예측에 따른 대기시간 감소 효과



효율적인 체선 및 체화 비용 계획
→ 선박 운영비 감소
→ 에너지 사용 최적화



선박 운항 시간 최적화
→ 대기 중 오염물질 배출 감소



부두 입출항 운행 시간 최적화
→ 연료 및 인력 비용 감소

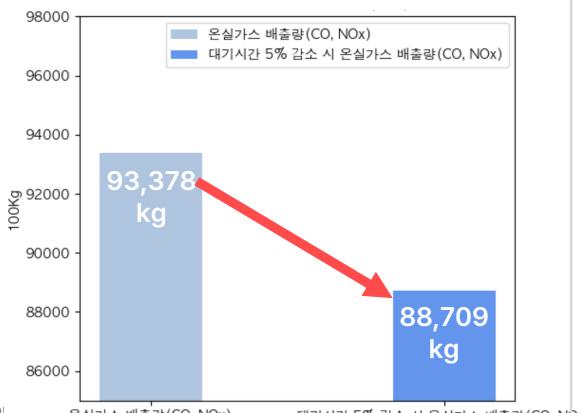
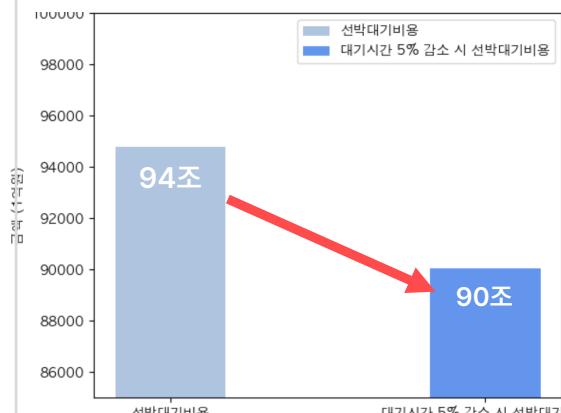


접안 대기시간 감소
→ 선박 운항 스케줄 조정 용이
→ 항만 이용 만족도 증가

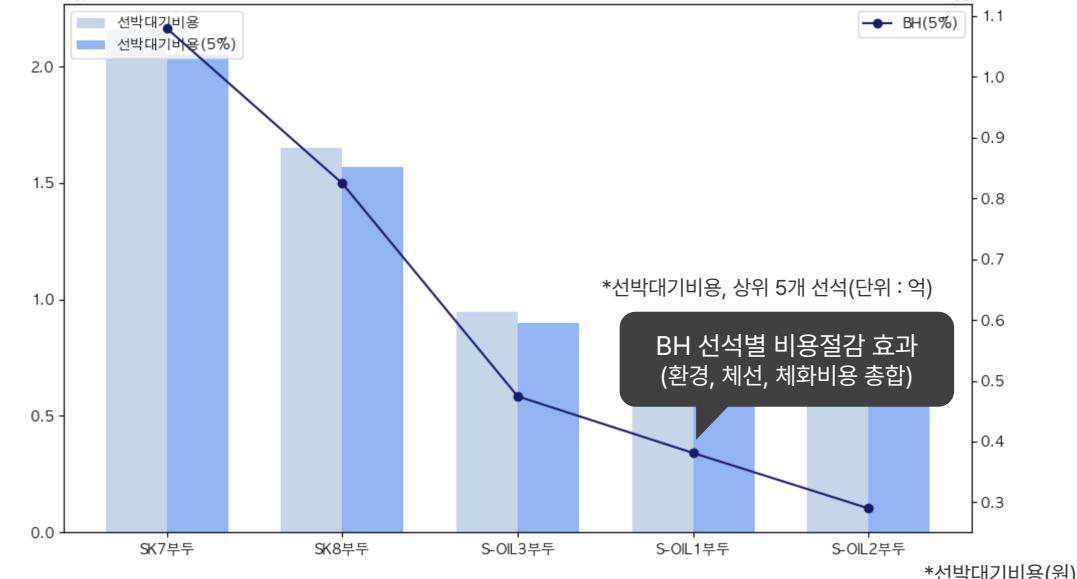
연간 대기시간(전 선석) 5% 절감 시 효과

선박대기비용 4조 738억 원 감소

온실가스 배출량 4,668,901 kg 감소



선석별 대기시간 5% 절감 시 편의 산출



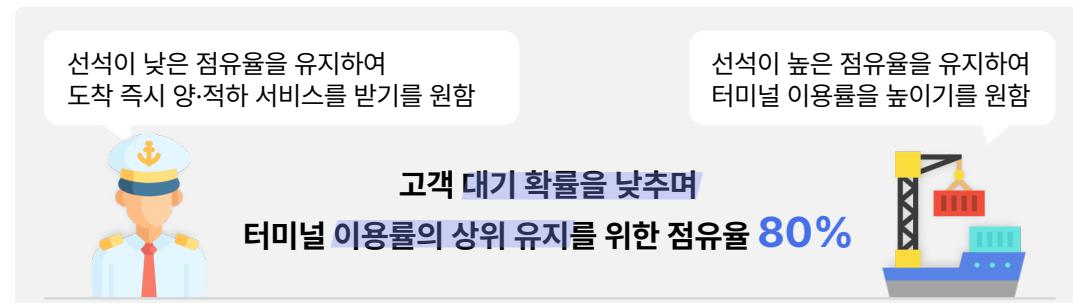
계선장소명	입항수	접안_대기시간_분	분당_환경비용	분당_체화비용	분당_체선비용	선박대기비용
SK7부두	89	2607.853933	1.603911e+09	7.697442e+09	24448.314607	21588404.0
SK8부두	82	2215.329268	1.389779e+09	7.689313e+09	23369.715447	16492850.0
S-OIL3부두	274	1230.817518	4.333394e+08	2.379561e+09	11497.080292	9486376.0
S-OIL1부두	728	1098.497253	2.457833e+08	7.075403e+08	7527.060439	7623846.0
S-OIL2부두	906	594.661148	2.342495e+08	8.426135e+08	7987.711552	5801782.0

15-4. 운영 시스템 개선 방안

선석 점유율 개선을 통한 항만 관리 편의성 & 효율성 제고

전 선석 점유율 80% 유지를 위한 운영 시스템 제안

■ 전 선석 점유율의 적정 수준(80%) 유지 필요성

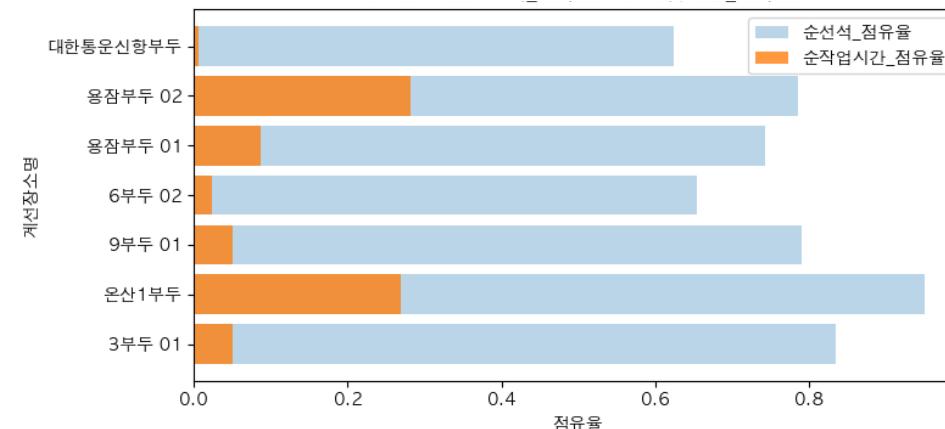


2022년 선석별 대기율 & 점유율
(점유율 대비 대기율 기준, 상위 7개 선석 ; 정박지 제외)

계선장소명	대기율	점유율_대비_대기율	순선석_점유율	순작업시간_점유율
대한통운신항부두	37.291280	59.776746	0.623843	0.005199
용잠부두 02	41.934932	53.298436	0.786795	0.281935
용잠부두 01	34.554158	46.528676	0.742642	0.086642
6부두 02	27.974880	42.711830	0.654968	0.023652
9부두 01	27.780958	35.111823	0.791214	0.050550
온산1부두	32.414845	34.084249	0.951021	0.269456
3부두 01	26.274584	31.481801	0.834596	0.051003

- 순작업시간 점유율 = 선석 작업시간 / 정상적인 선석 작업시간(360일)
- 순선석 점유율 = 선석작업시간/총선석시간

2022년 계선장소별 순선석_점유율과 순작업시간_점유율



1. 비정상적 범위(240일 이하)의 총 선석 작업 시간 정상화

- 항만 내 선석 정보(선박용도, 안벽 길이 등)를 종합적으로 파악하여 **높은 대기율의 선박이 해당 선석(개선 대상; 대체 선석)에 접안하도록 선석 운영 & 배치 시스템 정비**
→ 접안대기중인 선박들의 대기율을 낮추고, 서비스 시간을 높여 항만 운영 효율성 제고

2. 선석 설치가 필요한 케이스 수집

- 대기중 선박의 대체 선석이 부족한 경우(타겟 선석과 대체 선석의 특성 불일치),
대기가 발생하는 선박 특성 조사 후, 특성 별 선석의 추가 설치를 고려

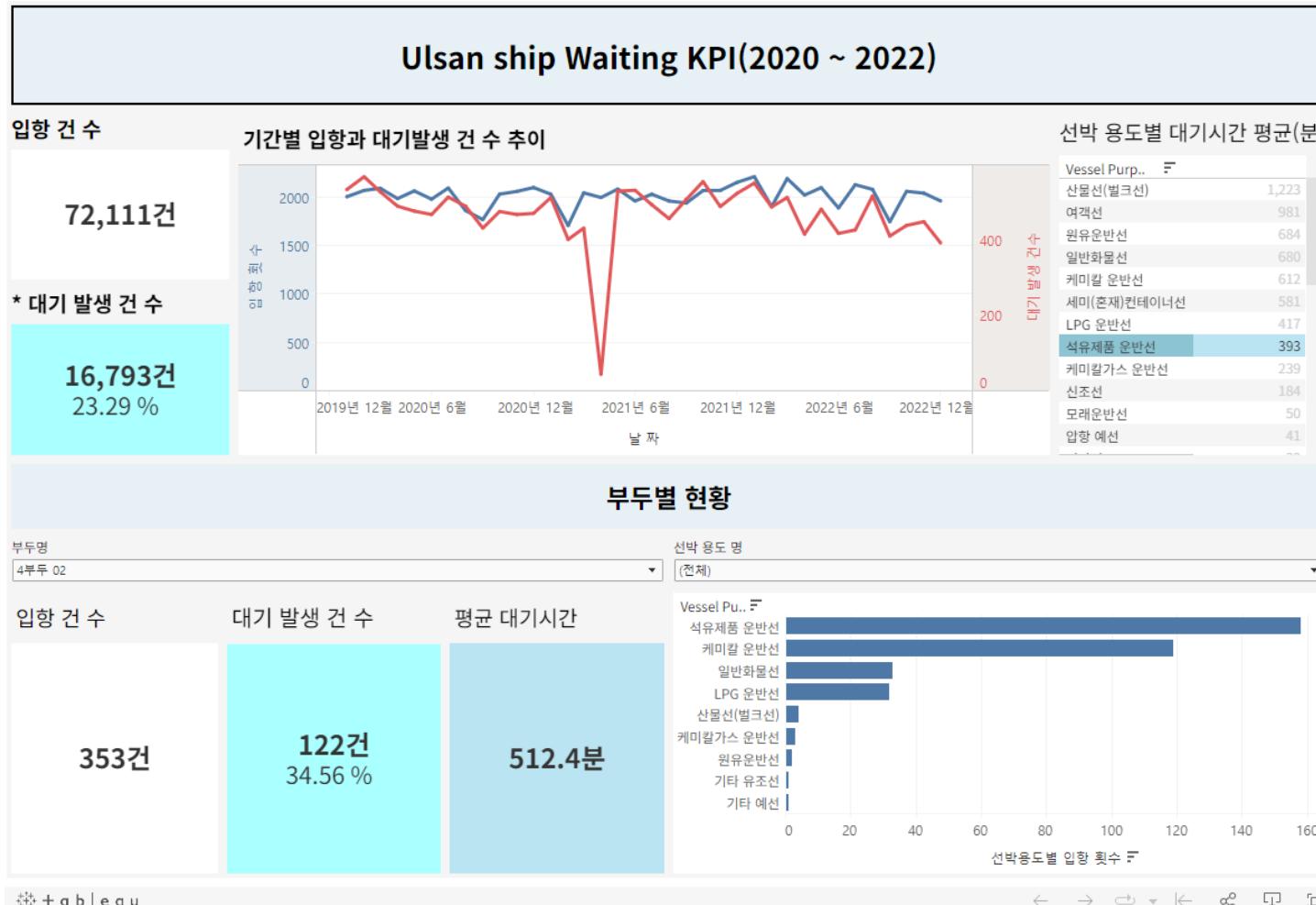
**컨테이너 터미널의 서비스 수준 평가지표에 대한 고찰(김창곤, 2016)

**선사비용과 항만비용을 동시에 고려한 선석점유율 비교 분석(한국해운물류학회, 2007)

부록1. Tableau 시각화

대시보드 소개

선박 용도별 대기시간 EDA 등 울산항만 분석 자료



부록2. 한계점

데이터 수집 한계점 및 개별 선석 결과 지표 예측 부재

■ 데이터 수집

- **기준 데이터 수집**
 - PORT MIS 입출항 정보, 시설사용허가현황 외 부두별 관제현황 데이터로 교차 검증 미흡
 - 정확한 선박의 입출항 프로세스 시점 추정 어려움
(예) 급유 실시/미실시, 양하와 적하를 다른 선석에서 실시, 양하와 적하를 다른 선석에서 다회 실시 등
- **부두 종류별/개별 선석 운영 데이터**
 - 개별 선석별 데이터 수집의 한계로 개별 참조만 진행함
- **선박 대기발생 데이터**
 - 선박의 감속운항 또는 항계 밖 표류, 기항 취소 등
PORT MIS 데이터 내에서 관측 불가능한 대기시간이 존재
 - 대기 발생하는 모든 경우를 포괄하지 못함

■ 선석 개별 예측 미실시

- 액체화물 선박 기준 예측을 진행하여 선박이 아닌 선석 기준으로 대기시간 예측을 실시하지 않음

■ 모델 적용의 한계

- 울산항의 데이터만을 기준으로 진행한 분석이므로 다른 항만에 즉시 적용에 어려움이 있음

■ 결과 지표 산정

- 체선비용 : 수집한 데이터 기준, 장기기회비용 & 회귀계수 산정 불가
 - 논문 프로세스 대체
 - 개별 선박 연료 사용량 데이터가 부재하여 추정된 연료량을 사용
- 체화 비용
 - 화물가치의 수입/수출을 구분이 어려워 개별 화물에 대한 화물가치 산정 한계
- 환경비용
 - 총톤수, DWT 기준으로 환경비용 추산이 어려움
 - 전체 온실가스 중 CO, NOx 만 산정함

부록3. 참조 문헌

■ 참조 논문

- 김은수, 김근섭. (2016). "국내 항만의 선박 대기율 실증 분석 연구." *한국항해항만학회지*, 40(1), 35-41.
- 박병인, 배종욱, 박상준. (2009). "선박당 평균대기비용에 의한 항만의 서비스 수준 평가." *한국항만경제학회지*, 25(4), 185-205.
- 조아현, 이화섭, 김보경, 김성아, 이종필, 한장협, 천민수. (2023). "선박대기시간 예측모형 개발을 위한 방법론 연구." *한국해양수산개발원*.
- 장영태, 성숙경. (2002). "우리나라 항만에서의 체선 체화 시간비용 재추정." *한국항해항만학회지*, 26(4), 383-390.
- 정장표. (2022). "PORT-MIS 자료를 활용한 부산항 선박대기오염물질 배출량 산정 및 활용." *한국대기환경학회 학술대회논문집*.
- 설성희, 정성운, 성미애, 임재현, 여소영, 진형아. (2021). "선박 대기오염물질 배출량 산정 및 배출 기여도 분석연구." *한국대기환경학회지(국문)*, 37(2), 324-337.
- Parolas, I. (2016). "ETA prediction for containerships at the Port of Rotterdam using Machine Learning Techniques" (Master's thesis). [Delft University of Technology].
- 박상국. (2014). "벌크 전용부두의 최적 서비스 수준에 관한 연구." *충실파워포트*.
- Yoon, J.-H., Kim, D.-H., Yun, S.-W., Kim, H.-J., & Kim, S. (2023). "Enhancing Container Vessel Arrival Time Prediction through Past Voyage Route Modeling: A Case Study of Busan New Port." *Journal of Marine Science and Engineering*, 10(2), 1-29.
- Kim, H., & Dang, B. H. (2020). "Estimation of Ship Emission based on Data of Ship's In and Out of Port." *Journal of Korean Institute of Intelligent Systems*, 30(6), 453-458.
- 최상진, 박성규, 장영기, 이희관, 황의현, 봉춘근. (2010). "선박부문 온실가스 배출량 산정에 관한 연구." *대한교통학회지*, 28(6), 33-42.
- 이정욱, 이향숙. (2021). "인천항의 대기오염물질 배출량 산정 연구." *한국항만경제학회지*, 37(1), 143-157.
- 박상국. (2017). "항만에서 최적 생산성 및 서비스 수준 관리를 위한 소프트웨어 개발." *한국항해항만학회지*, 41(3), 137-148.
- Syed, M. A. B., Ahmed, I. (2023). "A CNN-LSTM Architecture for Marine Vessel Track Association Using Automatic Identification System (AIS) Data." 1(1), 0.
- 김창곤. (2007). "선사비용과 항만비용을 동시에 고려한 선석점유율 비교 분석." *해운물류연구*, (55), 69-88.
- Shin, H., Kim, C., & Kim, H. (2022). "Prediction of container vessel waiting time in Busan port using LSTM neural network." *Journal of Advanced Marine Engineering and Technology*, 46(6), 394-401.
- Pel, R. (2022). "Vessel ETA Prediction." Master's Thesis, Utrecht University, Graduate School of Natural Sciences.
- 이정훈, & 박남규. (2018). "컨테이너 터미널의 이론적 대기율과 실제 대기율 비교에 관한 연구: 부산항 신항 A 터미널을 대상으로." *Journal of Korea Port Economic Association*, 34(1), 69-82.

■ 기타 자료

- 울산항만공사. 2022. 2022 울산항 통계연감. https://www.upa.or.kr/bbs/view.do?bbsId=BBS_0000000000000057&mld=001004006005000000&dataId=17112
- 울산항만공사. "부두현황". <https://upa.or.kr/safe/port/i-391/list.do> 검색일 : 2023.11.17
- 울산 투자환경 홈페이지. "석유화학산업". <https://www.investkorea.org/us-kr/cntnts/i-1111/web.do>. 검색일 : 2023.11.17

부록4. 일정

WBS

대분류	구분	1주차 (11/01 ~ 11/07)	2주차 (11/08~11/14)	3주차 (11/15~11/21)	4주차 (11/22~11/28)	5주차 (11/29~12/05)
기획	상세 기획 마무리					
데이터 수집	데이터 수집 및 Pipeline 구축					
데이터 전처리	EDA(변수 특징 확인)					
	Feature Engineering					
	최종 기능별 변수 선정 및 변수 변환					
모델링 및 평가	선박 대기시간 예측 모델 구축					
대시 보드 제작	태블로 대시보드 구현					
PPT 제작	프레젠테이션 자료 제작					

감사합니다