

## 분석 배경

물류센터 생산성은 제한된 시간 내 얼마나 많은 출고 작업을 처리 했는지로 측정되며, 그중 핵심은 단연 '피킹' 작업이므로, '피킹' 효율성에 중점을 두고, 물류센터 운영 데이터를 분석할 필요가 있음

## 분석 방법

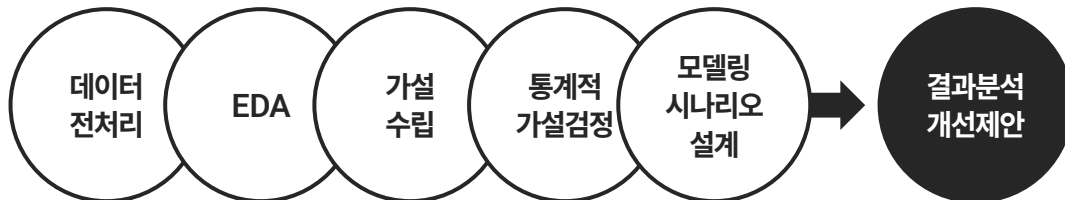
## 활용 데이터

데이터명: Mega Star Distribution Centre dataset (출처: 캐글)

주요 변수: 상품별 주문 빈도, 위치 분산도, 카테고리별 주문 현황, 피킹 기록 등

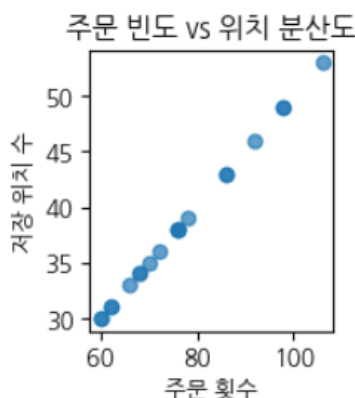
주요 컬럼: 상품, 로케이션, 스택프, 카테고리, 주문 빈도, 피킹 시간 등

## 분석 구조



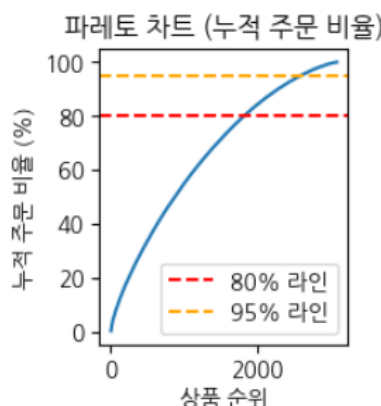
## 분석 과정

## 주요 변수 상관관계



주문 빈도 상위 20개 상품은 위치 분산도와 강한 양의 상관관계, 현재 피킹 시 동선 낭비가 발생하고 있음 → 출고 작업량 저하 발생

## 파레토 차트



전체 주문의 약 80%가 상위 20%의 상품에서 발생하는 파레토 분포를 따름 → 주문 빈도에 따라(ABC 등급), **진열 최적화 필요성이 도출**

## 분석 결과

## 가설 수립

EDA 주요 발견사항 요약:

- 전체 상품 수: 3,079개
- 평균 상품당 주문 횟수: 15.1회
- 최고 주문 횟수: 106회
- 고빈도 상품들이 평균 38.4개 위치에 분산 배치
- 상위 20% 상품이 전체 주문의 39.2% 차지
- ABC 분류: A등급 1,812개(58.9%), B등급 773개(25.1%), C등급 494개(16.0%)

## 1. 문제 정의

주문 빈도 데이터 미반영 진열 ▶ 피킹 시 이동 거리 증가  
▶ 피킹 생산성 저하 ▶ 운영 효율성 저하 ▶ 개선 필요

## 2. 가설 수립

귀무가설( $H_0$ ): ABC 등급 기반 진열 방식과 기존 방식 간 피킹 효율성에는 유의한 차이가 없다.

대립가설( $H_1$ ): 두 방식 간 피킹 효율성에 유의한 차이가 존재한다.

## 3. 평가지표 정의

피킹 효율성 지표: 이동거리, 작업시간, 시간 당 피킹수량

진열 효율성 지표: 위치 접근성, 위치 집중도

운영 효율성 지표: 전체 처리시간, 작업자 생산성, 오류율

## 통계적 가설 검정

- ① 정규성 검정 → ② 등분산 검정 →  
③ 일원분산분석 → ④ 사후 검정 →  
⑤ 효과 크기 검정 → ⑥ t-검정 → 채택

A등급 개선 효과 t-검정:

- t-통계량: 118.6702
- p-value: 0.0000
- 유의성:  $p < 0.001$  (매우 유의함)
- 결론: ABC 최적화가 A등급 상품의 위치 집중도를 유의하게 개선함

B등급 개선 효과 t-검정:

- t-통계량: 164.0964
- p-value: 0.0000
- 유의성:  $p < 0.001$  (매우 유의함)
- 결론: ABC 최적화가 B등급 상품의 위치 집중도를 유의하게 개선함

## 통계 검정 절차 및 결과값 예시

6단계를 거쳐 진행, 모든 주요 지표에 대해 유의미한 차이가 확인돼, 귀무가설을 기각하고 대립가설을 채택

→ ABC 등급 기반 진열 전략이 **기존 대비 유의미한 효율 차이가 있음**

### 인사이트

#### 모델링 설계

- 상품별 주문 빈도를 기반으로 ABC 등급을 예측하는 단변량 모델 설계
- 사용된 특성: 주문 빈도(order\_frequency), 보조 파생 특성: 주문\_밀도, 주문\_등급

모델	테스트 정확도	교차검증 평균
Random Forest	0.9740	0.9663 ± 0.0066
Logistic Regression	0.9594	0.9659 ± 0.0091
SVM	0.9594	0.9659 ± 0.0091

주문 빈도 기반 ABC 예측 모델 설계 결과,

💡 Random Forest가 가장 높은 정확도(97.4%)를 기록하여 최종 채택하여 진행

#### 시나리오 설계

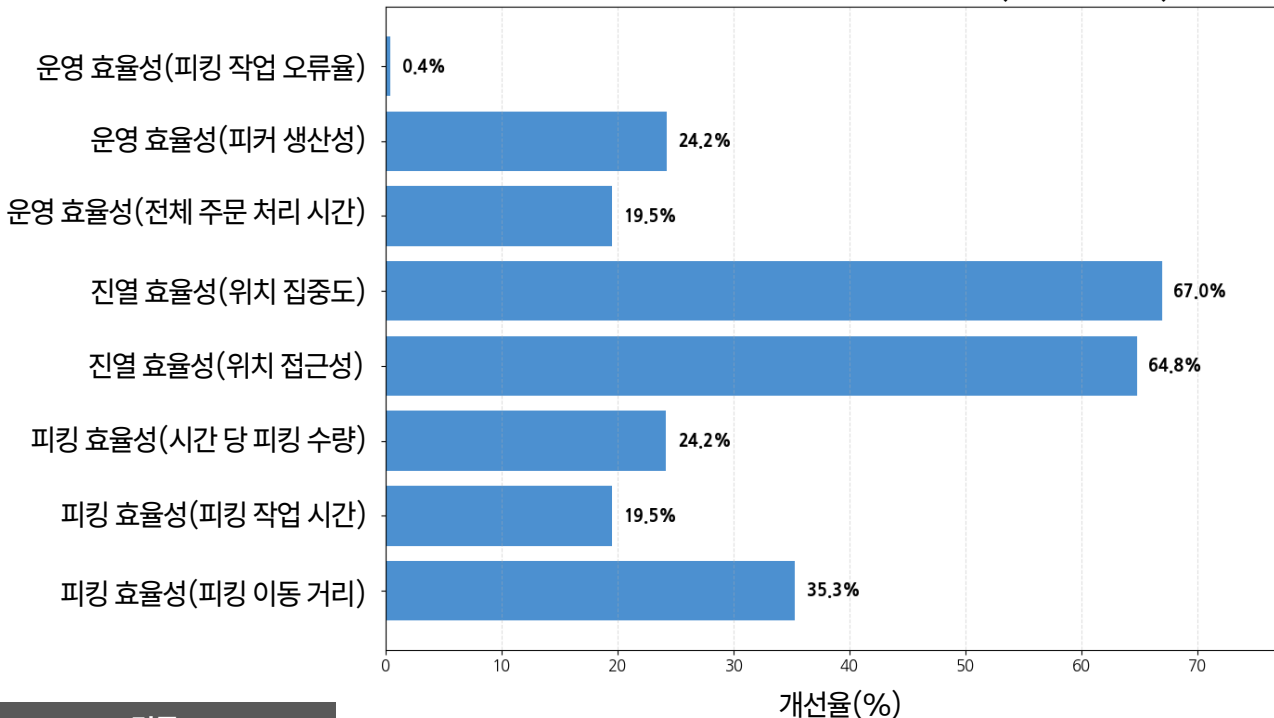
- 데이터셋에는 구획화된 상품 Zone 정보가 없어 진열 방식이 무작위일 것으로 판단
- 현재 진열 방식은 임의(Random), 변경 진열은 ABC 등급 기반으로 정의하여 양측의 피킹 효율 차이를 비교하는 시나리오를 설계

항목	내용
현재 진열(current_zone)	무작위로 로케이션에 임의(랜덤)진열 방식
변경 진열(optimal_zone)	상품의 ABC 등급에 따라 구획화된 Zone A/B/C에 매핑 및 진열 방식
Zone 기준 설정	입구 거리 기반: A(10m), B(25m), C(40m) 접근성: A > B > C
평가 목적	진열 방식 변경에 따른 피킹 이동 거리 및 효율성 비교 평가 분석

### 결과 분석

#### 시뮬레이션 결과 분석

ABC 등급 기반 진열 방식 시뮬레이션 결과(지표별 개선율)



#### 결론

ABC 등급 기반 진열 프로세스는 기존 대비 **피킹 효율성을 35.3% 개선** 하고, **전체 처리 시간을 19.5% 단축** 시킬 수 있어 물류센터 운영 효율성 향상에 실질적 기여가 가능

# 배송 지연, 사전에 대응 할 수 있을까?

- 배송 지연 탐지 및 실시간 알림 시스템 설계 프로젝트

## 분석 배경

고객의 신뢰도와 구매자의 경험을 좌우하는 **배송 지연**은 단순한 시간 초과 문제가 아닙니다. 반드시 AI 기술과 데이터 기반의 솔루션을 찾을 필요가 있는 **핵심 문제**입니다.

또한, 배송 운영 담당자가 실시간 모니터링 시 **사전 배송 지연 예측 판단의 어려움**이 있습니다.

- 확인하고 판단해야 할 배송 운영 지표가 많음.
- 실시간으로 빠르게 변하는 지역 상황을 인지 하는데 한계가 있음
- 외부 변수(날씨, 집회 등)로 인해 지표의 기준이 일관되지 않음

이에 따라, **이상 탐지** 기법을 활용하여 배송 지연을 사전에 예측하여 **조기 대응**할 수 있는 실시간 알림 시스템을 설계할 필요가 있습니다.

## 분석 방법

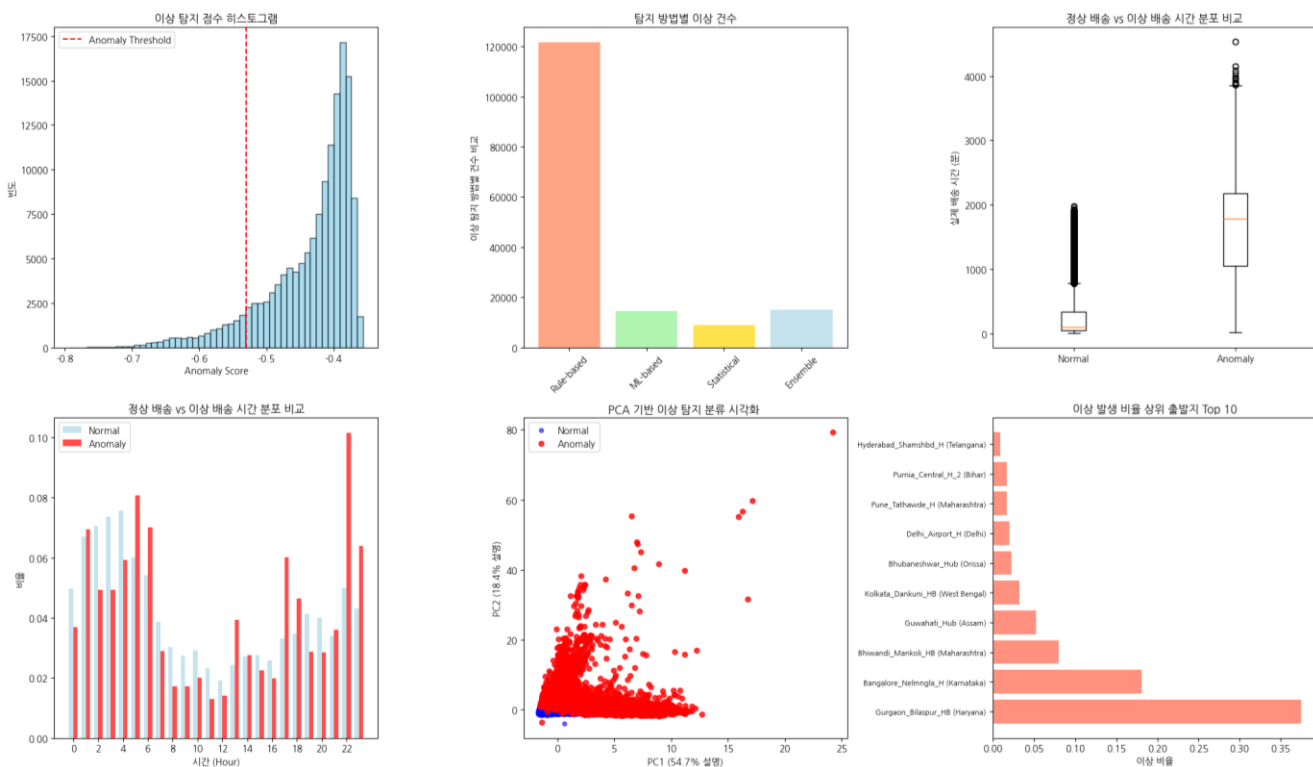
- 데이터: Delhivery Logistics Dataset (Kaggle, 144,867건 배송 이력)
- 주요 피쳐: 배송 시간, 출발지, 목적지, 운송 수단, 거리 등 총 9개
- 이상 탐지 기법: 규칙 기반 조건, Isolation Forest, Z-score를 조합한 앙상블 기법 활용

## 분석 과정

① 지연 기준 설정 → ② 피쳐 엔지니어링 → ③ 이상 탐지 모델 구성 →  
④ 이상 점수 계산 및 분포 분석 → ⑤ 이상 탐지 기반 배송 패턴 분석

## 분석 결과

### 이상 탐지 기반 배송 패턴 분석



- 이상 탐지 기반 분석을 통해 전체 데이터의 약 10.5%가 비정상 배송으로 분류된 점 확인  
배송 시간, 출발지, 운송 유형, 시간대 등 주요 특성에서 정상/이상 그룹 간 유의미한 차이가 관측됨
- 통계 분석 결과, 이상 그룹은 평균 배송 시간이 현저히 길며, 심야 시간대(22시 전후) 및 특정 출발지(Gurgaon, Bangalore 등)에 집중되는 경향을 보임
- 분석 결과를 바탕으로 **배송 지연 발생 조건에 대한 예측 기반 대응 전략 수립의 필요성**이 도출됨

# 배송 지연, 사전에 대응 할 수 있을까?

- 배송 지연 탐지 및 실시간 알림 시스템 설계 프로젝트

## 시스템 설계

### 배송 지연 기준 및 단계 분류

#### 배송 지연 판단 기준

-  $\text{actual\_time} - \text{osrm\_time} > \text{기준값}$

기준값:  $\text{delay\_min} > 60\text{분}$  또는  $\text{delay\_ratio} > 1.5$  또는  $Z\text{-score} > 3$  중 하나 이상 충족

분류 단계	분류 내용
위험 단계	- 이상 탐지 조건 3가지 이상 충족 - 이상 점수 하위 5% 이내 ( $\text{anomaly\_score} < 5\% \text{ quantile}$ ) - 연속 지연 3회 이상 발생
주의 단계	- 이상 탐지 조건 2가지 충족 - 이상 점수 5~10% 범위 - 고위험 출발지 또는 **피크 시간대(5시/일요일 등)**에 지연 발생



위험 단계 배송 건: 8623건 (6.0%)



주의 단계 배송 건: 6541건 (4.5%)

### 실시간 알림 시스템 설계

- 주요 운영 지표 배송 지연율, 고위험 주문 수, 평균 배송 시간, 이상 탐지 알림 수
- 시간대별 핵심 성과 지표: 시간별 배송 완료율, 시간별 배송 지연율
- 경고 임계값: - 주의 단계 배송 건: 0.15  
- 위험 단계 배송 건: 0.25  
- 연속 지연 횟수 기준: 3회

## 프로세스 설계

### 운영 대응 매뉴얼 설계

#### 즉시 조치 대응 방안

- 주의 단계 발생: 해당 배송 기사에게 연락하여 작업 중 특이사항 및 지연 사유 확인
- 위험 단계 발생: 배송 지연 이슈 최소화를 위한 대체 배송 경로 또는 인력 배정 즉시 실행
- 기술적 이슈 동반 시: 시스템 이상 여부 확인, 기술팀 긴급 공유, 백업 절차 전환 및 고객 공지

#### 단계별 대응 체계 방안

- Level 1: 실시간 운영 지표 모니터링 담당자 → 15분 이내 실시간 대응 진행
- Level 2: 배송 운영 실무 담당자 → 30분 이내 운영 상 조치 수행
- Level 3: 배송 운영 리더급 담당자 → 60분 이내 운영 대응 판단 및 실행

## 결론

### 프로젝트 결론 요약

1. 전체 데이터의 약 10.5%는 이상 배송으로 분류됨
2. 특정 출발지/시간대/운송 유형에서 이상 건 집중 발생
3. 이상 탐지 기반 조건으로 '위험' 및 '주의' 단계 기준 수립
4. 실시간 알림 및 대응 매뉴얼 기반 운영 체계 설계

→ 향후 시뮬레이션 기반 개선 효과 검증 및 운영 지표 추적 필요

## 서울시 따릉이, 언제 고장이 가장 잦을까?

- 2024년 8월 따릉이 고장율 분석 프로젝트

## 분석 배경

서울시 공공자전거 '따릉이'는 시민들의 일상적 교통수단이자 여가수단으로 자리잡았으며, 누적 이용건수가 2억 건에 육박하는 등 이용률이 급증했습니다. 특히 퇴근시간대와 주중 출퇴근 시간에 집중된 사용 패턴이 나타났고, 비수기 대비 성수기의 이용 격차도 뚜렷했습니다.

하지만 따릉이의 고장 빈도 증가는 단순한 불편을 넘어 시민들의 안전 문제까지 유발하고 있으며, 고장이 잦은 지역에서는 자전거 이용 회피 현상까지 발생하고 있습니다. 이러한 고장 문제는 여름철 집중되는 계절성 패턴과도 연관되어 있으며, 폭염 대응과 정비 전략 수립이 시급합니다.

## 분석 목적

2024년 8월 고장률 분석을 통해 9월 고장률 급증의 원인을 진단하고, 효율적인 대응 전략 및 정비 방식 개선안 도출로 2025년 9월 고장률 안정화를 목표로 함.

## 분석 과정

## 1. 데이터 수집 및 정제

- 2023년~2024년 자전거 고장 신고 내역, 대여·반납 이력, 대여소 위치 정보를 통합하여 고장률 산출 가능한 형태로 정제
- 총 60만건 이상 데이터를 활용해 연도/월/자치구 단위 고장률 계산

## 2. 8월 기준 분석 및 등급 분류

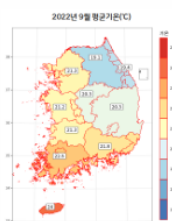
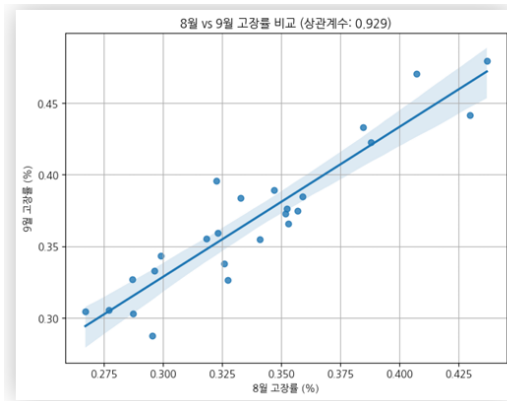
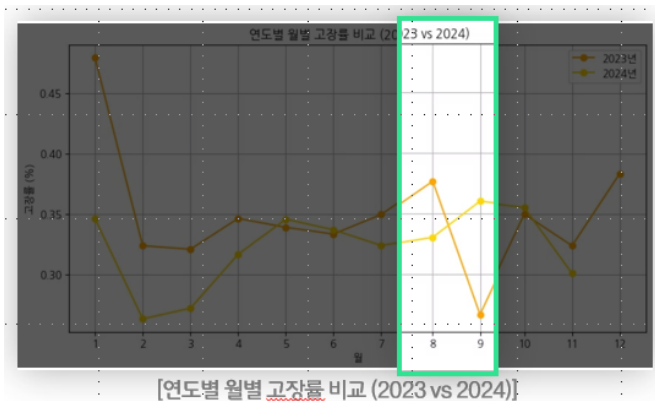
- 자전거별 고장이력과 반납 데이터를 매칭해 자치구 단위 고장률을 계산
- 고장률에 따라 각 자치구를 등급화하여 '따릉이 레전드' 등 계층화된 분석 구조 도입

## 3. 고장 부위 식별 및 우선순위 도출

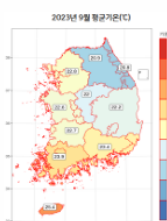
- 고장 사유별 데이터를 정리하고, 주요 부위별(예: 페달, 바퀴, 브레이크 등) 고장 빈도를 파악하여 유지보수 방향 제시

## 분석 결과

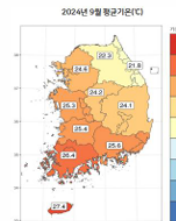
## 원인 분석



24.7°C

2024년 9월 평균기온  
역대 1위

6.0

2024년 9월 폭염일수  
역대 1위

4.3

2024년 9월 열대야일수  
역대 1위

## 서울시 따릉이, 언제 고장이 가장 잦을까?

- 2024년 8월 따릉이 고장율 분석 프로젝트

## 원인 분석

- 8월 정비의 중요성: 8월 고장률과 9월 고장률 간 상관관계수 0.929  
→ 8월 정비 수준이 9월 고장률에 강한 영향, 8월 대비 정비 강화 필요성 확인
- 2024년 9월은 역대 최고 기온 기록: 기온 상승에 따라 타이어, 체인, 브레이크 고장 등 급증  
→ 평균기온과 고장건수 간 뚜렷한 정(+) 상관관계 확인
- 서울시가 2024년 도입한 민간 정비팀 '따릉이포' 운영으로 전반적 고장률은 개선  
→ 다만 여름철 피크 기간에는 여전히 고장신고 급증  
→ 사후 정비 시스템 한계 존재: 사용자 신고 이후 수거 및 수리까지 지연 발생

## 결론 제시

## 탄력적 사전 정비팀 운영안

## 자치구별 등급 관리 시스템 도입

고장 데이터 분석 및 자치구별 고장률 기반 등급 부여



여름철 고장률 급증에 대응하기 위해, 고장 예방과 신속 대응이 가능한 '사전 정비팀' 운영을 제안합니다. '사전 정비팀'은 고장률 기반 등급 분류를 통해 우선 정비 지역을 선정하고, 타이어·체인 등 주요 부품을 집중 점검합니다. 출퇴근 시간대를 고려한 정비 인력 운영으로 시민 불편을 최소화할 수 있습니다.

## 시각화

## 따릉이 고장 관리 대시보드

