

# DOORDASH가 SNOWFLAKE와의 경험을 통해 데이터 및 BI 역량을 최대로 끌어올리다

<b>회사</b>	DoorDash	<b>목표</b>	직원 및 판매자의 데이터 가용성을 높여 데이터의 영향을 극대화한다
<b>위치</b>	캘리포니아, 샌프란시스코	<b>웹사이트</b>	doordash.com
<b>사업 영역</b>	모바일 배달		

## SNOWFLAKE 이점

- 별도의 스토리지 및 컴퓨팅을 갖춘 멀티 클러스터 공유 데이터 아키텍처
- 무한에 가까운 확장성을 갖춘 완전 관리형 인프라
- 라이브 데이터의 안전한 공유를 위한 Snowflake 데이터 마켓플레이스

DoorDash는 4,000개 이상의 도시에서 고객과 기업을 연결합니다. ‘Dasher’로 알려진 판매자 및 현지 배달 기사 네트워크를 통해 DoorDash는 테이크아웃 음식, 식료품, 가정 필수품을 온디맨드 방식으로 배달할 수 있습니다. DoorDash는 내부 이해 관계자와 판매자에게 통찰을 제공하기 위해 많은 양의 운영 데이터와 웹 및 모바일 앱의 데이터를 수집하고 분석합니다.

## 문제: 데이터 가용성 향상

DoorDash의 레거시 데이터 아키텍처는 데이터 기반 문화와 보조를 맞출 수 없었습니다. 소비자, 판매자 및 Dasher 데이터는 데이터 사일로 전체에 분산되어 운영 보고 및 비즈니스 의사 결정에 문제를 야기했습니다. 하루 70,000개의 대시보드 쿼리와 3,000개의 임시 쿼리로 인한 리소스 경합은 SLA 누락과 오래된 데이터 문제로 이어졌습니다.

데이터 모델링은 시간이 많이 들고 복잡했습니다. DoorDash의 비즈니스 인텔리전스 이사인 Marta Vovchenko는 이렇게 말했습니다. “우리의 목표는 단순한 음식 배달이 아닌 최고의 배달 플랫폼이므로 새로운 비즈니스 라인의 성과를 측정할 수 있는 데이터 모델을 몇 개월 동안의 개발 없이 신속히 만들어야 했습니다.”

잡은 파이프라인 오류는 DoorDash의 데이터 가용성 이슈를 악화시키고, 관리 병목 현상을 일으키고, BI 팀이 새로운 데이터 소스를 통합하는데 방해가 되었습니다.

결과적으로 IT 팀은 비즈니스를 지원하기 위해서, 레거시 기술 스택에 의해 생기는 문제를 극복하는 데 과도한 시간과 노력을 들여야 했습니다.

## 솔루션: 확장 가능한 배달 통찰 플랫폼

DoorDash는 현대적인 데이터 인프라의 필요성을 깨닫고 Snowflake 플랫폼으로 눈을 돌렸습니다.

Snowflake의 멀티 클러스터 공유 데이터 아키텍처는 DoorDash의 모든 데이터, 사용자, 워크로드를 이전보다 두 배 빠른 속도로 처리하도록 확장되었습니다. Snowflake의 커넥터, 드라이버, 프로그래밍 언어의 네트워크는 600개의 ETL 작업 마이그레이션을 가속화했습니다. ETL 작업은 23% 더 빠르게 완료되어 BI 팀이 SLA 보고를 99.7% 처리할 수 있었습니다.

무한에 가까운 확장성을 갖춘 Snowflake의 완전 관리형 인프라 덕분에 팀은 데이터 분석 및 모델링에 집중할 수 있었습니다. Vovchenko는 “Snowflake의 탄력성 덕분에 우리가 어떤 문제를 겪든 궤도를 유지할 수 있습니다.”라고 말했습니다.

“Snowflake로의 마이그레이션은 BI의 성장과 DoorDash에서의 데이터 영향력 확대에 필수적인 요소로 있어 왔습니다.”

—MARTA VOVCHENKO, 비즈니스 인텔리전스 이사, DoorDash

## 결과: 확장 가능한 시장 통찰 및 데이터 모델링

Snowflake로 수집되는 DoorDash의 모든 소비자, 판매자 및 Dasher 데이터는 전 세계 시장 관리자에게 매일 오전 7시까지 공급 및 수요에 대한 최신 통찰을 제공하는 데 사용됩니다.

Snowflake에서 DoorDash의 판매자 포털을 설계하면 판매자에게 주문, 재고, 인력 관리를 위한 데이터 기반 보고서를 제공할 수 있습니다. Vovchenko는 “최종 사용자를 위한 제품으로서 우리 데이터를 제공하는 것은 우리만의 큰 차별화를 이루는 요소입니다.”라고 말합니다.

BI 팀은 Snowflake의 즉각적인 확장성을 통해 별도의 데이터 웨어하우스를 신속하게 가동하고 새로운 비즈니스 라인의 성과를 측정하는 데이터 모델을 개발할 수 있습니다. Snowflake로 구동되는 DoorDash의 마케팅 기여 모델은 Facebook, Google 및 기타 플랫폼의 광고 데이터를 분석해 캠페인과 예산을 최적화합니다. 머신 러닝 알고리즘은 Snowflake에 저장된 데이터를 사용해 고객 개인에게 맞춤형 제품을 추천합니다.

감소된 데이터 관리는 데이터 탐색을 위한 여유 공간을 확보합니다. 예를 들어, DoorDash는 동일 집단별로 인앱 A/B 테스트 결과를 파악하기 위해 Snowflake를 진실 공급원으로 활용합니다.

## 미래: 적은 노력으로 새로운 데이터 세트에 액세스

DoorDash는 Snowflake 데이터 마켓플레이스를 최우선 순위에 둡니다. Vovchenko는 이렇게 말합니다. “Snowflake 데이터 마켓플레이스를 사용하면 데이터 파이프라인 없이 외부 소스의 데이터에 쉽게 액세스할 수 있게 될 것입니다.” 안전한 데이터 공유는 DoorDash가 판매자 포털의 표준 보고보다 데이터 요구 사항이 더욱 확대된 기업 고객에게 서비스를 제공하는 데 훨씬 더 향상된 능력을 갖추게 해 줄 것입니다.

지리 공간 데이터를 Snowflake로 수집하면 매장 위치가 갖는 경제적 영향을 이해하기가 훨씬 용이해질 것입니다. Vovchenko는 이렇게 말했습니다. “시장 관리자는 그러한 통찰을 사용해서 다양한 구성을 분석하고 기하학적 배열을 확장함으로써 수요와 공급에 영향을 미칠 것입니다.”

## 왜 SNOWFLAKE일까요



리소스 경합을 제거해 주는 멀티 클러스터 공유 데이터 아키텍처



복잡한 데이터 모델의 개발을 가속화하는 즉각적인 탄력성



개발을 간소화하는 커넥터, 드라이버, 프로그래밍 언어의 네트워크



데이터 사이언스를 지원하는 머신 러닝 통합



데이터 관리를 간소화하는 완전 관리형 인프라



라이브 데이터 공유를 가능하게 하는 Snowflake 데이터 마켓플레이스

## SNOWFLAKE 소개

Snowflake가 제공하는 데이터 클라우드는 거의 무제한의 규모, 동시성, 성능을 통해 수천 개의 조직이 데이터를 모으는 글로벌 네트워크입니다. 데이터 클라우드 내에서 조직은 사일로된 데이터를 통합하고, 관리형 데이터를 쉽게 검색하고 안전하게 공유하며, 다양한 분석 워크로드를 실행합니다. 데이터나 사용자가 어디에 있든 Snowflake는 여러 공용 클라우드에서 단일하고 원활한 경험을 제공합니다. 이미 데이터 클라우드의 새로운 영역으로 비즈니스를 추진하고 있는 Snowflake 고객, 파트너 및 데이터 공급자에 합류하십시오.

[snowflake.com](https://snowflake.com)