출처 : [통신사 고객 이탈(11.1.3+) (ibm.com)](https://community.ibm.com/community/user/businessanalytics/blogs/steven-macko/2019/07/11/telco-customer-churn-1113)

*Telco 고객 이탈* 데이터에는 3분기의 캘리포니아 고객 7043명의 집 전화 및 인터넷 서비스를 제공한 가상의 통신 회사에 대한 정보가 포함되어 있습니다. 어떤 고객이 서비스를 떠났거나, 머물렀거나, 가입했는지를 나타냅니다. 각 고객에 대한 여러 중요한 인구 통계와 만족도 점수, 이탈 점수 및 고객 평생 가치(CLTV) 지수가 포함됩니다.

업데이트 일자 : 2019년 7월 11일

* **통신사 이탈:** 이 샘플 대시보드는 다양한 요인을 기반으로 가상의 통신 회사의 고객 이탈을 추적합니다. 이탈 레이블 열은 고객이 지난 달에 퇴사했는지 여부를 나타냅니다. 다른 열에는 위치, 월별 요금, 서비스 및 고객 평생 가치가 포함됩니다. 위치: *팀 콘텐츠 > 샘플 > 대시보드.*
* **분기별 이탈 업데이트:** 이 샘플 스토리는 가상의 통신 회사에서 고객 이탈의 분기별 변화와 다음 분기의 목표를 결정하기 위해 가장 높은 계약 및 위치를 보여줍니다. 이탈 레이블 열은 고객이 지난 분기 내에 떠났는지 여부를 나타냅니다. 위치: *팀 콘텐츠 > 샘플 > 스토리*.
* **우편번호별 고객 이탈 정보:** 이 샘플 보고서는 샘플 대시보드 'Telco churn' 및 샘플 스토리 'Quarterly churn update'에 대한 드릴스루 대상 보고서입니다. 위치: *팀 콘텐츠 > 샘플 > 보고서.*
* **통신사 이탈 관계:** 이 샘플 탐색은 다양한 요인을 기반으로 가상의 통신 회사의 고객 이탈을 추적합니다. 이탈 레이블 열은 고객이 지난 달에 퇴사했는지 여부를 나타냅니다. 다른 열에는 위치, 월별 요금, 서비스 및 고객 평생 가치가 포함됩니다. 위치: *팀 콘텐츠 > 샘플 > 탐색 분석.*
* **통신사 고객 이탈:** 이 샘플 데이터 모듈은 가능한 다양한 요인을 기반으로 가상의 통신 회사의 고객 이탈을 추적합니다. 이탈 열은 고객이 지난 달에 퇴사했는지 여부를 나타냅니다. 다른 열에는 성별, 피부양자, 월별 요금 및 각 고객이 보유한 서비스 유형에 대한 정보가 포함된 많은 열이 포함됩니다. 출처: IBM. 위치: *팀 콘텐츠 > 샘플 > 데이터*. *Telco 고객 이탈* 데이터 모듈은 5개의 업로드된 파일로 구성됩니다.
  + Telco\_customer\_churn\_demographics.xlsx
  + Telco\_customer\_churn\_location.xlsx
  + Telco\_customer\_churn\_population.xlsx
  + Telco\_customer\_churn\_services.xlsx
  + Telco\_customer\_churn\_status.xlsx

## 

## 데이터

[**기본**](https://community.ibm.com/community/user/businessanalytics/blogs/steven-macko/2018/09/12/base-samples-for-ibm-cognos-analytics) 샘플이 설치되면 *팀 콘텐츠 > 샘플 > 데이터로* 이동하면 *Telco 고객 이탈*이라는 데이터 모듈이 표시됩니다. 여기에는 5 개의 테이블이 있습니다.

* 인구 통계
* 위치
* 인구
* 서비스
* 상태

각 표는 아래에 설명되어 있습니다.

### 인구 통계

**고객 ID:** 각 고객을 식별하는 고유 ID입니다.

**개수:** 보고/대시보드에서 필터링된 집합의 고객 수를 합산하는 데 사용되는 값입니다.

**성별:** 고객의 성별 : 남성, 여성

**연령:** 회계 분기가 종료된 시점의 고객의 현재 연령(년)입니다.

**고령자:** 고객이 65세 이상인지 여부: 예, 아니요

**기혼여부:** 고객이 기혼인지 여부(예, 아니요)를 나타냅니다.

**부양 가족:** 고객이 부양 가족과 함께 살고 있는지 여부를 나타냅니다(예, 아니요). 부양 가족은 자녀, 부모, 조부모 등이 될 수 있습니다.

**부양가족 수:** 고객과 함께 살고 있는 부양가족의 수를 나타냅니다.

### 위치

**고객 ID:** 각 고객을 식별하는 고유 ID입니다.

**개수:** 보고/대시보드에서 필터링된 집합의 고객 수를 합산하는 데 사용되는 값입니다.

**나라:** 고객의 주 거주지 국가입니다.

**상태:** 고객의 주 거주지 상태입니다.

**도시:** 고객의 주 거주지 도시입니다.

**우편번호:** 고객의 주 거주지 우편번호입니다.

**위경도 :** 고객의 주 거주지의 위도와 경도를 합한 것입니다.

**위도:** 고객의 주 거주지의 위도입니다.

**경도:** 고객의 주 거주지의 경도입니다.

### 인구

**아이디:** 각 행을 식별하는 고유 ID입니다.

**우편번호:** 고객의 주 거주지 우편번호입니다.

**인구:** 전체 우편 번호 지역에 대한 현재 인구 추정치입니다.

### 서비스

**고객 ID:** 각 고객을 식별하는 고유 ID입니다.

**개수:** 보고/대시보드에서 필터링된 집합의 고객 수를 합산하는 데 사용되는 값입니다.

**회계 분기:** 데이터가 파생된 회계 분기(예: Q3)입니다.

**지인 추천 여부 :** 고객이 친구나 가족을 이 회사에 추천한 적이 있는지 여부를 나타냅니다(예, 아니요).

**추천 횟수:** 고객이 수행한 현재까지의 조회 수를 나타냅니다.

**임기(개월):** 위에 지정된 분기 말까지 고객이 회사에 근무한 총 개월 수를 나타냅니다.

**최신 마케팅 제안 :** 해당하는 경우 고객이 수락한 마지막 마케팅 제안을 식별합니다. 값에는 없음, 제품 A, 제품 B, 제품 C, 제품 D 및 제품 E가 포함됩니다.

**전화 서비스 가입 여부 :** 고객이 회사의 집 전화 서비스에 가입했는지 여부를 나타냅니다.

**평균 월간 장거리 요금:** 위에 지정된 분기 말까지 계산된 고객의 평균 장거리 요금을 나타냅니다.

**중복회선 가입여부:** 고객이 회사의 여러 전화 회선에 가입하는지 여부를 나타냅니다(예, 아니요).

**인터넷 서비스:** 고객이 회사의 인터넷 서비스에 가입했는지 여부를 나타냅니다(아니요, DSL, 광섬유, 케이블).

**평균 월간 GB 다운로드:** 위에 지정된 분기 말까지 계산된 고객의 평균 다운로드 볼륨(GB)을 나타냅니다.

**온라인 보안:** 고객이 회사에서 제공하는 추가 온라인 보안 서비스에 가입했는지 여부를 나타냅니다(예, 아니요).

**온라인 백업:** 고객이 회사에서 제공하는 추가 온라인 백업 서비스에 가입했는지 여부를 나타냅니다(예, 아니요).

**장치 보호 계획:** 고객이 회사에서 제공하는 인터넷 장비에 대한 추가 장치 보호 계획에 가입했는지 여부를 나타냅니다(예, 아니요).

**프리미엄 기술 지원:** 고객이 대기 시간이 단축된 회사의 추가 기술 지원 플랜에 가입하는지 여부를 나타냅니다(예, 아니요).

**스트리밍 TV:** 고객이 인터넷 서비스를 사용하여 타사 공급자의 TV 프로그램을 스트리밍하는지 여부를 나타냅니다(예, 아니요). 회사는 이 서비스에 대해 추가 요금을 부과하지 않습니다.

**스트리밍 영화:** 고객이 인터넷 서비스를 사용하여 타사 공급자의 영화를 스트리밍하는지 여부를 나타냅니다(예, 아니요). 회사는 이 서비스에 대해 추가 요금을 부과하지 않습니다.

**스트리밍 음악:** 고객이 인터넷 서비스를 사용하여 타사 공급자의 음악을 스트리밍하는지 여부를 나타냅니다(예, 아니요). 회사는 이 서비스에 대해 추가 요금을 부과하지 않습니다.

**무제한 데이터:** 고객이 무제한 데이터 다운로드/업로드를 위해 추가 월 사용료를 지불했는지 여부를 나타냅니다.

**계약:** 고객의 현재 계약 유형(월별, 1년, 2년)을 나타냅니다.

**종이 없는 청구:** 고객이 종이 없는 청구를 선택했는지 여부를 나타냅니다(예, 아니요).

**지불 방법:** 고객이 청구서를 지불하는 방법을 나타냅니다: 은행 인출, 신용 카드, 우편 수표

**월별 요금:** 회사의 모든 서비스에 대한 고객의 현재 총 월별 요금을 나타냅니다.

**요금 합계:** 위에 지정된 분기 말까지 계산된 고객의 총 요금을 나타냅니다.

**총 환불액:** 위에 지정된 분기 말까지 계산된 고객의 총 환불 금액을 나타냅니다.

**총 추가 데이터 요금:** 위에 지정된 분기 말까지 요금제에 지정된 것보다 더 많은 추가 데이터 다운로드에 대한 고객의 총 요금을 나타냅니다.

**총 장거리 요금:** 위에 지정된 분기 말까지 요금제에 지정된 것보다 긴 거리에 대한 고객의 총 요금을 나타냅니다.

### 상태

**고객 ID:** 각 고객을 식별하는 고유 ID입니다.

**개수 :** 보고/대시보드에서 필터링된 집합의 고객 수를 합산하는 데 사용되는 값입니다.

**회계분기:** 데이터가 파생된 회계 분기(예: Q3)입니다.

**만족도 점수:** 회사에 대한 고객의 전반적인 만족도 등급은 1(매우 불만족)에서 5(매우 만족)까지입니다.

**만족도 점수 레이블:** 점수의 텍스트 버전(1-5)을 텍스트 문자열로 나타냅니다.

**고객 상태:** 분기 말에 고객의 상태(Churned, Stay 또는 Joined)를 나타냅니다.

**이탈 레이블:** 예 = 고객이 이번 분기에 회사를 떠났습니다. 아니요 = 고객이 회사에 남아 있습니다. 이탈 가치와 직접적인 관련이 있습니다.

**이탈 값:** 1 = 고객이 이번 분기에 회사를 떠났습니다. 0 = 고객이 회사에 남아 있습니다. Churn Label과 직접 관련이 있습니다.

**이탈 점수:** 예측 도구 IBM SPSS Modeler를 사용하여 계산되는 0-100 사이의 값입니다. 이 모델은 이탈을 유발하는 것으로 알려진 여러 요인을 통합합니다. 점수가 높을수록 고객이 이탈할 가능성이 높아집니다.

**이탈 점수 범주:** 0-10, 11-20, 21-30, 31-40, 41-50, 51-60, 61-70, 71-80, 81-90 및 91-100 범주 중 하나에 변동 점수를 할당하는 계산

**CLTV 크랙:** 고객 생애 가치. 예측된 CLTV는 기업 공식과 기존 데이터를 사용하여 계산됩니다. 가치가 높을수록 고객의 가치가 높아집니다. 가치가 높은 고객은 이탈을 모니터링해야 합니다.

**CLTV 카테고리:** CLTV 값을 2000-2500, 2501-3000, 3001-3500, 3501-4000, 4001-4500, 4501-5000, 5001-5500, 5501-6000, 6001-6500 및 6501-7000 범주 중 하나에 할당하는 계산입니다.

**이탈 카테고리:** 고객의 이탈 이유에 대한 높은 수준의 범주: 태도, 경쟁자, 불만족, 기타, 가격. 그들이 회사를 떠날 때, 모든 고객은 퇴사 이유에 대해 묻습니다. Churn Reason과 직접 관련이 있습니다.

**이탈 이유:** 고객이 회사를 떠나는 구체적인 이유. 이탈 카테고리와 직접 관련이 있습니다.

—-

[인구통계] - 고객ID - 개수 - 성별 - 연령 - 고령자 여부 - 결혼여부 - 부양가족 여부 - 부양가족 수

[위치] - 고객ID - 개수 - 나라 - 상태 - 도시 - 우편번호 - 위경도 - 위도 - 경도 인구 - 아이디 - 우편번호 - 인구추정치

[서비스] - 고객ID - 개수 - 회계분기 - 지인 추천 여부 - 추천 횟수 - 임기 - 최신 마케팅 제안 - 전화서비스 가입여부 - 평균 월간 장거리 요금 - 중복회선 가입여부 - 인터넷 서비스 - 평균 월간 GB 다운로드 - 온라인 보안 - 온라인 백업 - 장치 보호 계획 - 프리미엄 기술 지원 - 스트리밍 TV - 스트리밍 영화 - 스트리밍 음악 - 무제한 데이터 - 계약 - 종이 없는 청구 - 지불 방법 - 월별 요금 - 요금 합계 - 총 환불액 - 총 추가 데이터 요금 - 총 장거리 요금

[상태] - 고객ID - 개수 - 회계분기 - 만족도 점수 - 만족도 점수 레이블 - 고객상태 - 이탈레이블 - 이탈값 - 이탈 점수 - 이탈 점수 범주 - CLTV 크랙 - CLTV 카테고리 - 이탈 카테고리 - 이탈 이유