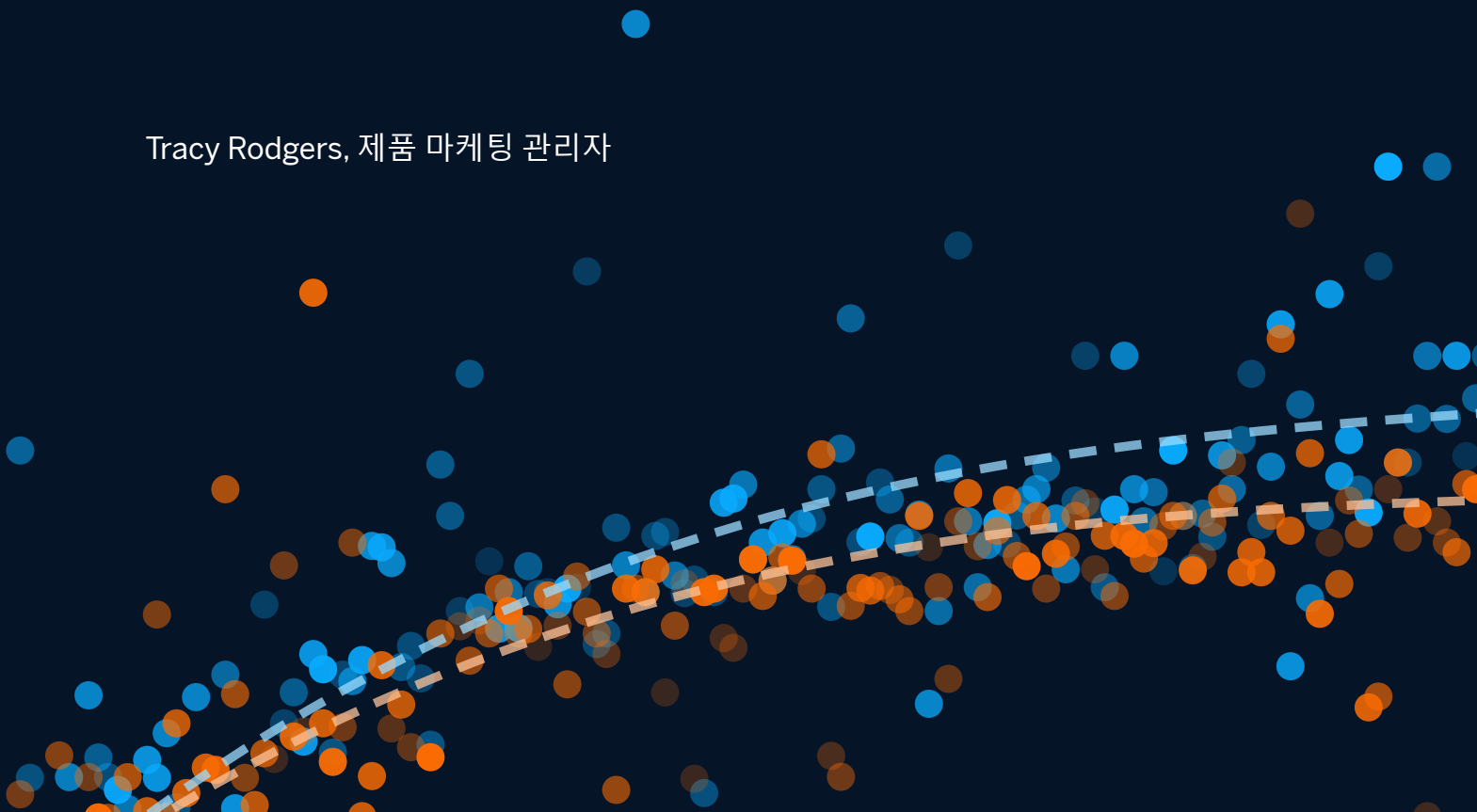




Tableau를 사용한 고급 분석

Tracy Rodgers, 제품 마케팅 관리자



소개

IT 환경에는 두 가지 부류의 사용자가 존재했습니다. 프로그래밍 방법을 아는 사람들이 있는가 하면 고급 분석 기술이 능력 밖이라고 생각하는 사람들이 있었습니다. 그리고 R이나 Python 또는 SAS를 배운 사람들이 있는가 하면 까다로운 작업을 누군가가 처리해줘야 하는 사람들이 있었습니다. Tableau는 인간의 지능을 진정으로 높이려면 사용자의 기술 능력 수준에 관계없이 다양한 기능을 제공해야 한다고 믿고 있습니다. 그리고 거기엔 고급 분석 기능도 포함됩니다.

Tableau의 심층 분석을 통해 Netflix 및 Spotify와 같은 최첨단 분석 기업에서 데이터 과학 그룹의 워크플로우를 강화하고 있습니다. 몇 번의 클릭만으로 박스 플롯, 트리 맵, 예측 비주얼리제이션, 동질 집단 분석을 만들 수 있습니다. R, MATLAB 및 Python에 연결하고 Tableau를 사용하여 분석을 확장하고 모델 결과를 시각화해 보십시오. 기술 전문가가 아닌 사용자는 이전에는 접근할 수 없었던 문제를 질문할 수 있고, 동시에 데이터 과학자는 좀 더 심층적인 인사이트를 더 빠르게 반복하고 발견하여 더 가치 있는 결과를 얻을 수 있습니다.

본 문서에서는 분석 프로젝트의 모든 단계에서 Tableau가 어떤 도움이 되는지 알아보며 특히 몇 가지 고급 기능을 중점적으로 살펴보겠습니다.

다음과 같은 시나리오와 이를 지원하는 기능을 살펴보도록 하겠습니다.

세분화 및 동질 집단 분석: Tableau를 사용하면 신속하고 유연한 동질 집단 분석에 필요한 직관적인 조사를 원활하게 수행할 수 있습니다. 필요에 따라 다양한 차원에서 데이터의 심층 분석을 통해 여러 가지 관점을 신속하게 살펴볼 수 있습니다. 기계 학습을 통해 세분화 분석을 강화하고자 하는 경우, 자동 클러스터 분석을 통해 차원이 많은 대규모 데이터 집합에서 찾아내기 힘들 수 있는 패턴을 좀 더 수월하게 발견할 수 있습니다.

시나리오 및 가정(what-if) 분석: Tableau의 유연한 프론트엔드와 강력한 입력 기능을 결합하면 계산을 신속하게 수정하고 다양한 시나리오를 테스트할 수 있습니다.

정교한 계산 및 통계 함수: Tableau에는 강력한 계산 언어가 있어, 임의의 계산으로 분석을 보강하고 간결한 식으로 복잡한 데이터 처리도 수행할 수 있습니다. 기본 중첩 집계에서 표준 편차, 모드, 첨도, 기울기, 상관관계, 공분산과 같은 통계 계산까지 가능합니다.

시계열 및 예측 분석: 전 세계 대부분의 데이터는 시계열로 모델링할 수 있으며, Tableau는 기본적으로 풍부한 시계열 분석을 지원하므로 하나의 강력한 UI 내에서 계절성 및 추세를 탐색하고, 데이터를 샘플링하고, 예측 분석을 실행하는 등 다양한 시계열 작업을 수행할 수 있습니다.

외부 서비스 통합: R, MATLAB 및 Python과의 통합은 Tableau 프론트엔드의 성능과 사용 편의성을 제공하기도 하지만, 전문가는 이를 통해 다른 플랫폼의 이전 작업을 활용하고 미묘한 차이가 있는 통계 및 기계 학습 요구 사항을 처리할 수 있습니다.

목차

세분화 및 동질 집단 분석	4
클러스터 분석.....	5
집합 및 집합 작업.....	6
그룹.....	7
가정 및 시나리오 분석	9
매개 변수.....	9
Story points.....	10
정교한 계산	10
세부 수준 식.....	10
테이블 계산.....	11
시계열 및 예측 분석	12
시계열 분석.....	12
예측.....	14
외부 서비스 통합	16
Python, R, MATLAB 통합	16
결론	19
Tableau 정보	20
관련 백서	20
기타 리소스 탐색	20

1. 세분화 및 동질 집단 분석

기능: 클러스터 분석, 집합, 집합 작업, 그룹

비즈니스 사용자와 데이터 전문가가 초기 가설을 생성하기 위해 많은 경우 세그먼트 생성 및/또는 비공식 동질 집단 분석 수행이라는 동일한 방식으로 시작합니다.

다양한 세그먼트에 대한 일련의 기본 질문은 분석가가 데이터를 이해하고 "신용카드로 결제하는 고객이 수표로 결제하는 고객보다 유지율이 높습니까?"와 같은 가설을 검증하는 데 도움이 됩니다. 신속하게 반복할 수 있는 기능은 모델 개발을 추진하고 프로젝트를 정상적으로 진행하는 데 도움이 됩니다.

Tableau에는 세그먼트를 반복적으로 빠르게 분석 및 비교할 수 있는 다양한 기능이 있습니다. 예를 들면 소수의 계산된 필드와 몇 번의 드래그 앤 드롭 작업만으로 날짜 및 시간별 환자 수로 세분화하는 대시보드를 만들 수 있고 진료과목에 따른 추세를 더 깊이 탐색할 수 있습니다(그림 1).

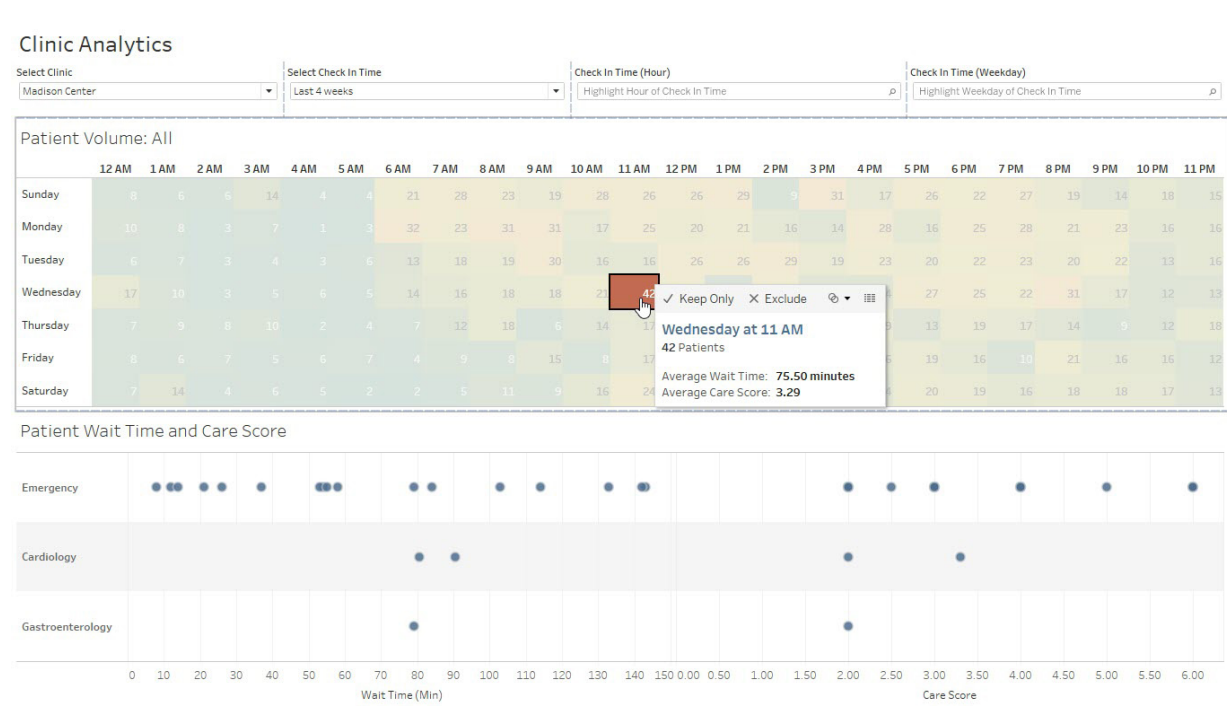


그림 1: 이 대화형의 병원 대시보드는 특정 일별, 시간별로 환자의 수, 방문 이유, 그리고 대기 시간을 보여줍니다.

클러스터 분석

Tableau의 **클러스터 분석**은 자율적인 기계 학습을 사용하여 데이터를 세분화하는 추가 옵션으로, 특히 다수의 변수를 고려해야 할 때 유용합니다.

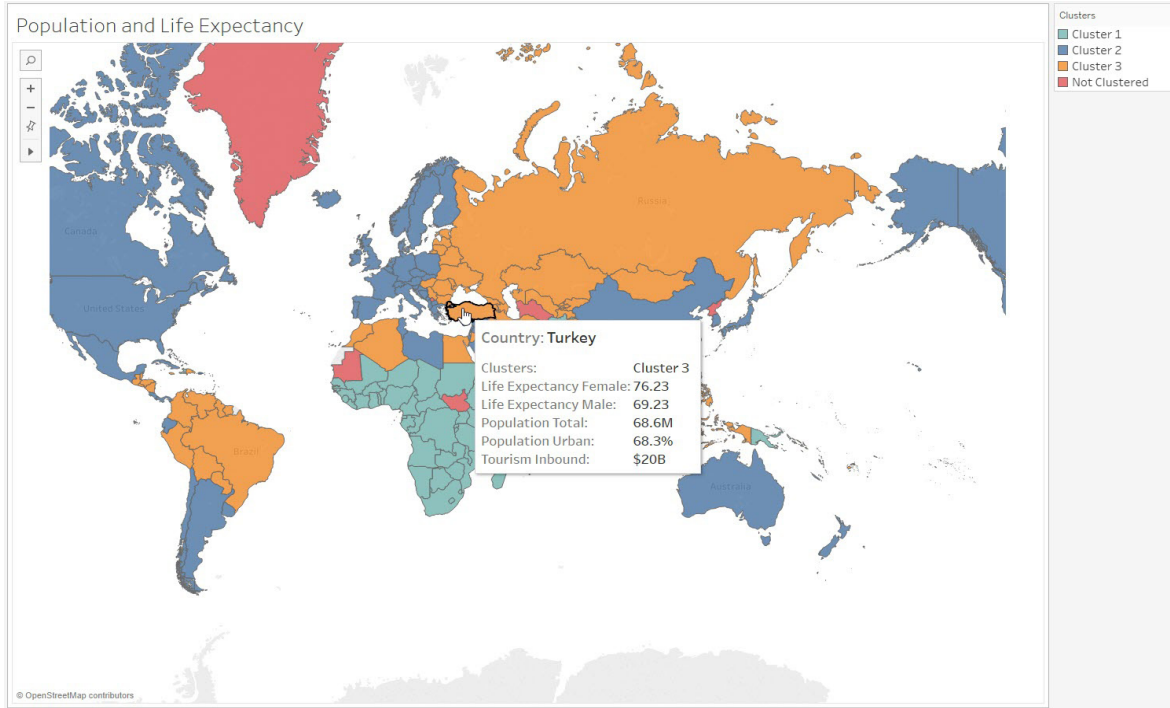


그림 2: 클러스터 분석을 사용하여 데이터를 자동으로 세분화합니다.

그림 2에서는, 미국의 개발 지표에 따라 국가를 클러스터 분석한 결과를 확인할 수 있습니다. 후진국, 개발도상국, 선진국에 해당하는 3가지 클러스터가 있는 솔루션을 원하는 알고리즘입니다.

또한 Tableau의 유연한 인터페이스를 통해 다양한 이론을 테스트하고 동질 집단 간 배포를 살펴볼 수 있습니다. 시각적으로 반복하는 기능을 사용하면 스크립트를 수정하고 다른 도구에서 시뮬레이션을 실행하는 데 드는 시간을 절약할 수 있습니다.

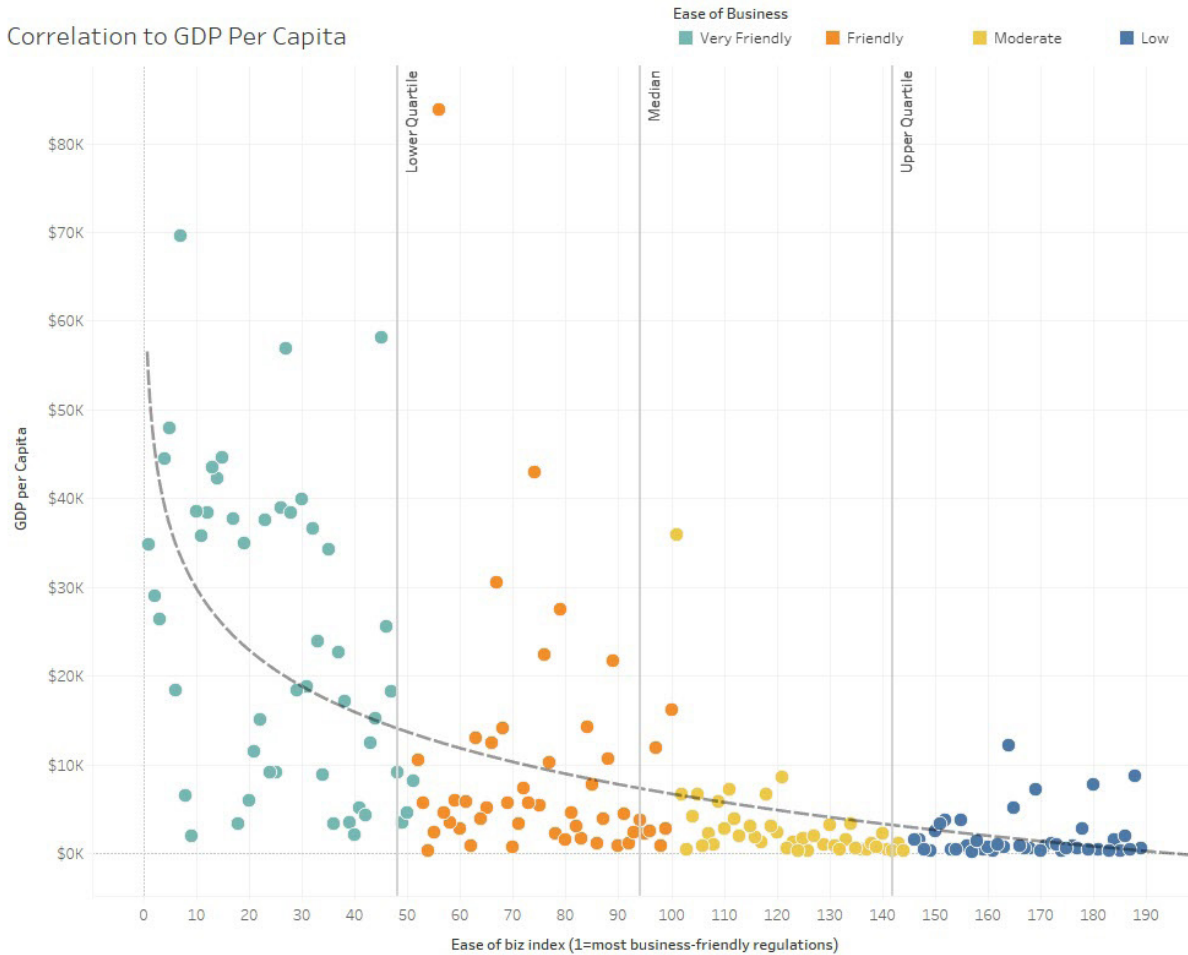


그림 3: 몇 초 만에 데이터를 세분화하고 탐색합니다. 이 예에서, '기업 친화적 규제'와 1인당 GDP로 분류된 국가들을 봅니다.

그림 3에서 보는 대로, 클러스터 분석에 사용된 세분화 필드와 측정값을 캔버스로 끌면 데이터를 보는 새로운 방식이 생성되며 네 개의 국가 클러스터 간의 차이가 하이라이트됩니다.

집합 및 집합 작업

집합을 사용하면 수동 선택 또는 프로그래밍 논리를 통해 데이터 객체의 모음을 정의할 수 있습니다. 집합은 필터링, 하이라이팅, 동질 집단 계산, 이상값 분석을 비롯한 다양한 시나리오에 유용할 수 있습니다. 또한 여러 집합을 결합하여 다양한 시나리오를 테스트하거나 시뮬레이션을 위해 여러 동질 집단을 생성할 수 있습니다(예: 보유 분석을 위해 독립적으로 생성된 다양한 고객 그룹 결합 또는 여러 개의 연속 기준 적용).

집합 작업은 비주얼리제이션 내에서 선택한 데이터 요소들을 하나의 집합으로 저장할 수 있는 기능입니다. 워크시트에서 사용자가 선택한 사항은 집합을 업데이트하여 조율된 단일 선택 사항으로 모든 대상 워크시트를 수정할 수 있습니다. 이로써 비례 선긋기와 같은 다양한 사용 사례가 가능합니다(그림 4).

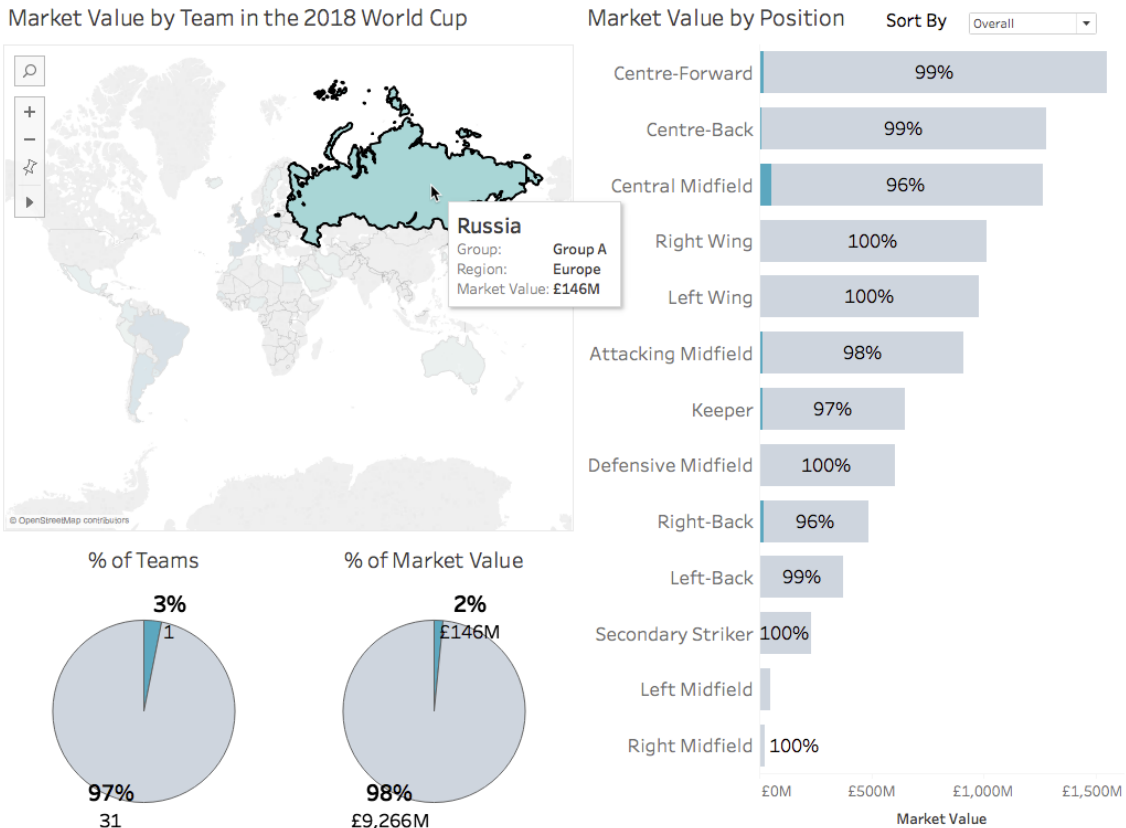


그림 4: 집합 작업은 전체 규모에 비례하여 선택 사항의 규모를 표시하는 대화형 분석 기술인 비례 선긋기를 가능하게 합니다. '러시아'를 선택하면 월드컵 참가 선수의 포지션별로 그 나라의 시장 가치가 나타납니다.

그룹

애드혹 범주를 만들고 계층 구조를 설정하는 것을 지원하기 위해 Tableau에는 그룹이라는 기능이 있습니다. 그룹은 기본적인 데이터 정리 요구 사항에도 도움이 됩니다.

그룹을 사용하여 필드 내에서 관련된 멤버를 결합함으로써, 사용자는 많은 분석 임무를 위해 직관적인 방식으로 데이터를 구조화할 수 있습니다. 예를 들어 전공별 평균 시험 점수를 표시하는 뷰로 작업하는 경우 특정 전공을 그룹화하여 범주를 만들 수 있습니다. 영어와 역사를 교양 전공 그룹으로 결합하고 생물학과 물리학을 과학 전공으로 그룹화할 수 있습니다. 이렇게 하여 분석가는 프레젠테이션을 사용자 지정하고, 분석 과정 전체에서 데이터 집계를 제어할 수 있습니다.

그룹은 또한 데이터에 일관성 및 품질 문제가 있는 경우에도 유용합니다. 예를 들면 캘리포니아를 전체 이름으로 부를 수도 있지만, CA 또는 Calif라고 할 수도 있습니다. 대부분의 경우 분석가와 비즈니스 사용자는 문제를 해결하기 위해 원본 시스템을 직접 변경할 수 있는 권한이 없기 때문에 데이터에 작은 오류라도 있으면 조사 분석에 큰 지장을 초래할 수 있습니다. 데이터 변경을 요청하기 위해 질문을 중단하면 프로젝트가 지연되고 아이디어의 신속한 개발이 저해됩니다. 그룹을 사용하면 흐름을 방해하지 않고도 분석용 대체 이름이 모두 포함된 새 세그먼트를 신속하게 정의하고 계속 질문할 수 있습니다.

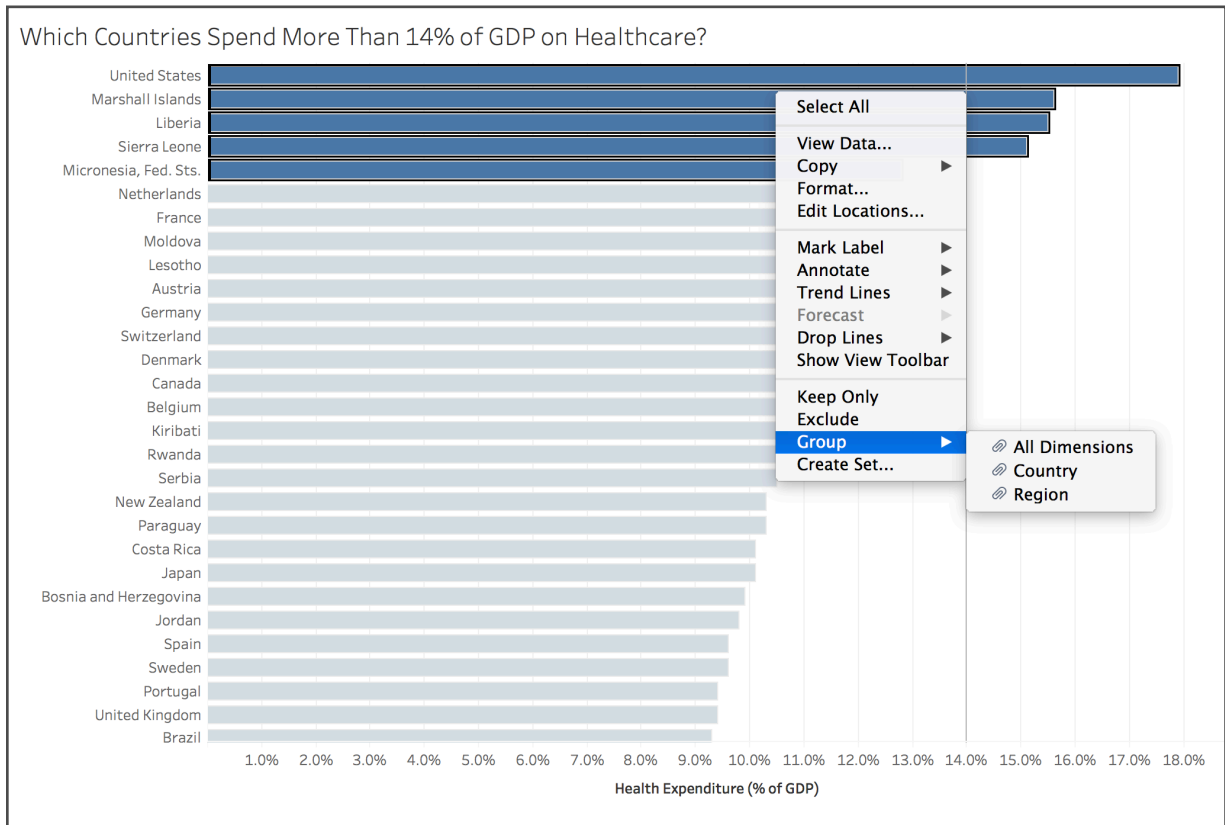


그림 5: 필드에서 관련된 멤버를 결합하는 그룹을 생성합니다. 이 예에서 의료 서비스에 지출하는 GDP의 비율이 높은 순서대로 상위 5개 국가를 그룹으로 만들고자 합니다.

이러한 모든 기능에는 단순 업데이트 기능이 포함되어 있습니다. Tableau에서 라이브 연결을 선택하고 데이터를 업데이트하면 분석은 물론 집합 및 그룹과 같은 모든 기본 구성 요소도 업데이트됩니다. 즉, 보고서 또는 종속 스크립트를 수동으로 다시 실행하지 않아도 동질 집단 구성원이 자동으로 업데이트됩니다. 단순 업데이트는 보고 업무의 부담을 줄여주는 것은 물론 시나리오를 테스트할 수 있는 또 다른 방법이기도 합니다. 이를 통해 분석 스택을 업데이트하지 않고도 초기 조건에 대한 민감도를 조사하기 위해 기초 데이터를 교체할 수 있습니다.

Tableau를 사용하면 사용자가 데이터를 빠르게 세분화하고 범주화할 수 있어 비즈니스 사용자와 분석가 모두 강력한 동질 집단 분석을 비교적 손쉽게 수행할 수 있습니다. 이러한 기능은 데이터 과학자가 초기 가설을 조사하고 시나리오를 테스트하는 데도 도움이 됩니다.

2. 가정 및 시나리오 분석

기능: 매개 변수, Story Points

사용자는 때때로 특정 값이나 값 집합을 변경하는 것이 분석 결과에 어떤 영향을 미치는지 알아보려고 합니다. 이러한 결과는 다양한 이론을 테스트하거나 동료에게 중요한 시나리오를 하이라이트하거나 새로운 비즈니스 가능성을 조사하는 데 사용될 수 있습니다. Tableau를 사용하면 분석 입력 내용을 실험할 수 있고, 단순 제어로 시나리오를 쉽게 플래그 지정하고 공유하는 한편 데이터를 최신으로 유지할 수 있습니다.

매개 변수

가정 분석을 수행할 때 사용자는 계산의 기준 값을 변경하거나 할당량을 다시 정의하거나 초기 조건을 설정하고자 할 수 있습니다. Tableau의 **매개 변수**를 사용하면 이러한 작업을 쉽게 수행할 수 있습니다. 사용자는 매개 변수를 정의하여 모델 또는 대시보드에 대한 입력 값을 변경할 수 있는 방법을 제공합니다. 매개 변수를 사용하면 계산을 주도하고 필터 임계값을 변경하는 것은 물론 대시보드에 입력되는 데이터를 선택할 수도 있습니다. 기술 전문가가 아닌 사용자도 매개 변수를 활용하여 다양한 입력을 실험하고 복잡한 모델에서 나올 수 있는 결과를 탐색할 수 있습니다.

Tableau의 매개 변수 기능을 사용하면 가설을 테스트할 수 있을 뿐만 아니라 대화형 보고서로 가정 분석 결과를 보여줄 수 있습니다. 그림 6에서 매개 변수는 영업 커미션에 관한 가정 분석을 이끌어냅니다. 영업 관리자는 커미션 비율, 기본 급여, 할당량을 실험하는 동시에 이러한 변경 내용이 주요 메트릭에 어떻게 영향을 미치는지에 대해 실시간 피드백을 받을 수 있습니다.

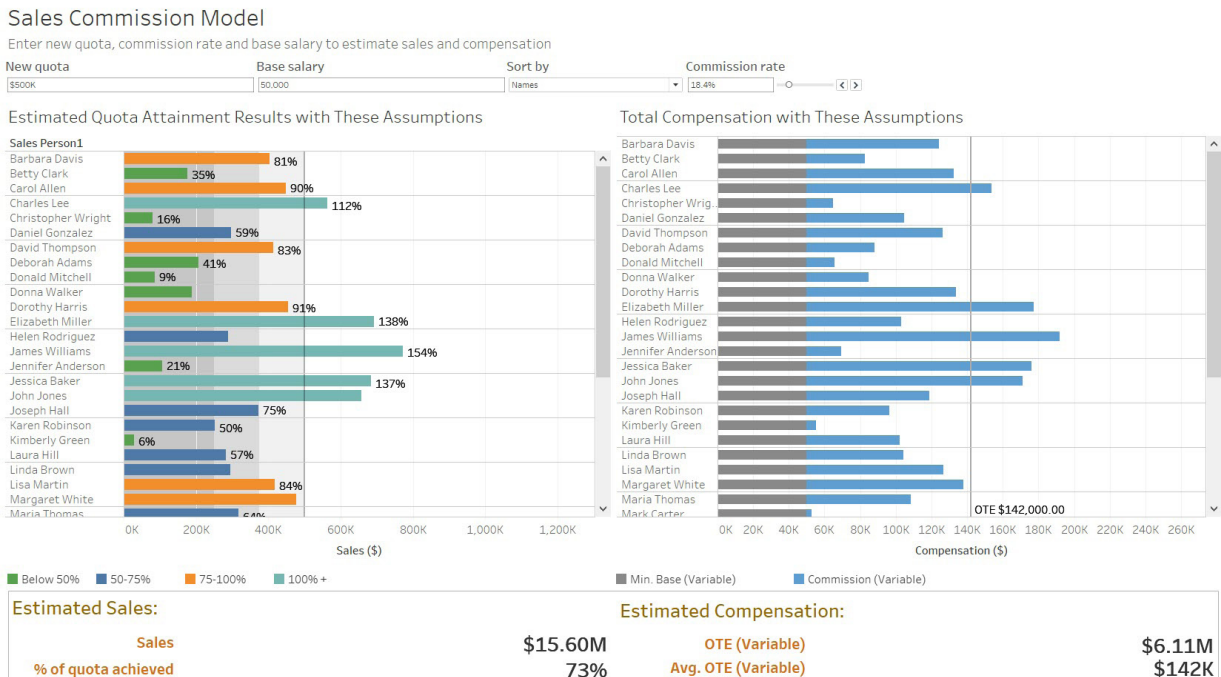


그림 6: 이러한 매개 변수 중심의 영업 보고서를 통해 인터랙터는 조직 내 할당량, 커미션, 급여에 미치는 영향을 알아볼 수 있습니다.

Story Points

매개 변수를 **Story Points**(데이터로 이야기를 만드는 Tableau의 방식)와 결합하면 흥미로운 결과를 스냅샷으로 만들고 탐색을 계속할 수 있습니다. Story Points를 사용하면 데이터 변경 및 비주얼리제이션 수정을 통해 프레젠테이션을 계속 업데이트하도록 구성할 수 있습니다. 다른 한편, Story Points는 매개 변수 값을 유지할 수 있을 정도로 스마트하기 때문에, 시나리오에 플래그를 지정할 수 있으며 분석 흐름을 중단하지 않은 채 안심하고 시나리오로 돌아갈 수 있습니다. 또한 스크린샷이 오래되어 걱정하거나 시뮬레이션을 다시 실행하지 않고도 다양한 입력 집합에서 나오는 결과를 비교할 수 있습니다.

Tableau에서는 집합, 그룹, 끌어 놓기 방식의 세분화, 매개 변수를 통해 이론과 질문으로부터 전문적인 대시보드로 이동하여 비전문가도 질문을 통해 자신의 시나리오를 테스트할 수 있습니다. 가정 분석의 간소화 기능을 통해 데이터 전문가가 분석의 더 복잡한 부분에 집중하여 더 효과적인 인사이트를 제공할 수 있으며, 직관적인 비주얼리제이션을 통해 최종 사용자는 데이터를 편리하게 활용할 수 있습니다. 이처럼 데이터 활용이 늘어나면 변화를 이끌고 조직의 의사 결정을 개선하는 데 도움이 됩니다.

3. 정교한 계산

기능: 계산된 필드, 세부 수준 식, 테이블 계산

일반적으로 원본 데이터에는 포괄적인 분석에 필요한 모든 필드가 포함되어 있지 않습니다. 분석가에게는 데이터를 변환하고 복잡한 논리를 정의하기 위한 간단하면서도 강력한 언어가 필요합니다.

Tableau는 사용하기 쉬우면서도 복잡한 논리를 지원하는 강력한 기능 또한 제공합니다. Tableau의 **계산된 필드**를 사용하면 데이터 원본의 기존 데이터로부터 새로운 데이터를 만들 수 있습니다. 계산된 필드를 사용하면 산술 연산을 수행하거나 조건부 논리를 표현하고 특정 데이터 유형에 대한 특수 연산을 수행하는 작업이 쉬워집니다. 고급 분석을 가능하게 하는 두 가지 유형의 계산된 필드는 세부 수준(LOD) 식 및 테이블 계산입니다.

세부 수준 식

세부 수준 식은 Tableau 계산 언어의 확장 프로그램으로, 하나의 비주얼리제이션에서 여러 세부 수준에 관련된 질문에 답할 수 있는 강력한 방법입니다. 세부 수준 식을 통해 이전에는 불가능했거나 어려웠던 시나리오를 매우 간단하고 간결한 식으로 처리할 수 있습니다.

이전 섹션에서 설명한 대로 세부 수준 식 덕분에 동질 집단 분석과 다중 패스 집계가 크게 간소화되었습니다. 그림 7에서는 '신규'와 '기존'으로 분류된 고객 동질 집단에 대한 지역 시장별 고객 유치의 누계를 보여줍니다. 시계열 분석에 대한 다음 섹션에서는 이러한 분석을 가능하게 하는 계산 언어의 다른 측면에 대해 살펴보겠습니다. 차트를 보면 북미의 고객 유치가 1월부터 그해 내내 감소했음을 알 수 있습니다. 복잡한 그룹별 SQL 문이 필요할 수 있는 세분화가 세부 수준 식을 통해 Tableau의 프론트엔드에서 편집하고 조정할 수 있는 단순하고 직관적인 식으로 전환되었습니다.

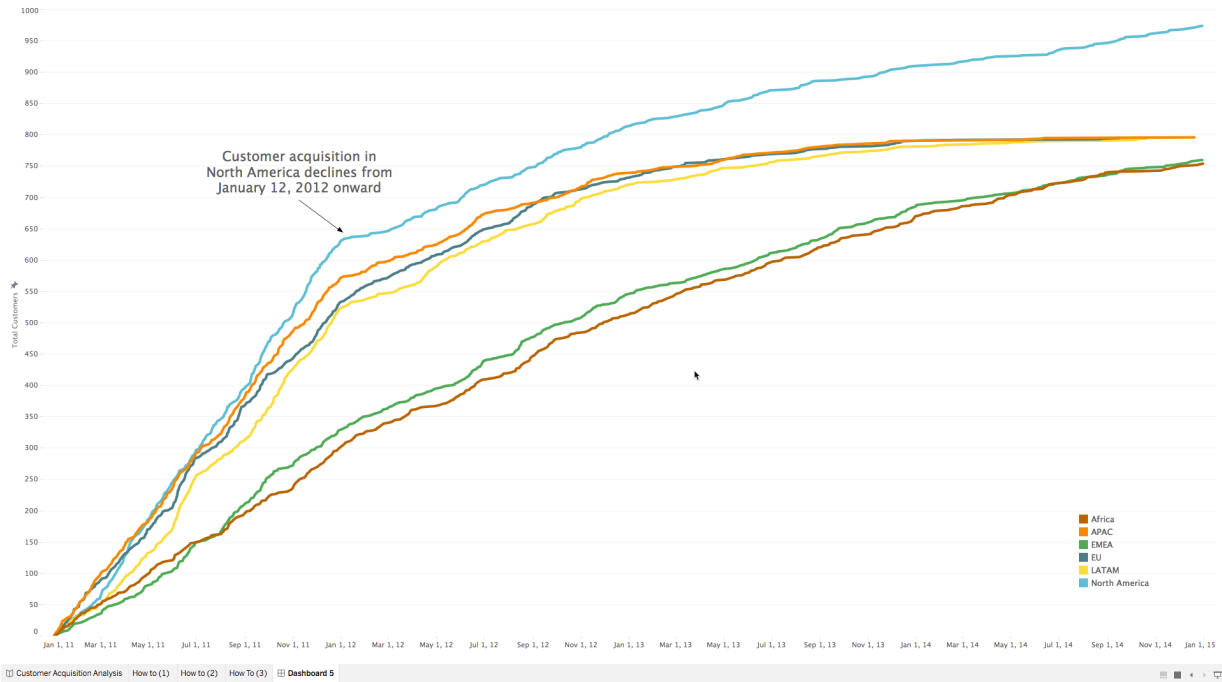


그림 7: 이 비주얼리제이션에서는 시장별 총 고객 유지 수의 일별 추세를 보는 데 세부 수준 식을 사용합니다. 세부 수준 식은 반복된 고객을 신규 고객으로 잘못 세지 않게 해줍니다.

테이블 계산

Tableau의 **테이블 계산**을 통해 본질적으로 상대적인 계산을 수행할 수 있습니다. 좀 더 구체적으로 말하면 테이블 계산은 테이블의 모든 값에 적용되는 계산이며 대부분 테이블 구조 자체에 종속적입니다. 이러한 유형의 계산에는 시간 간격 또는 누계와 같은 다양한 시계열 연산이 포함되며 순위 지정 및 가중 평균과 같은 계산도 포함됩니다.

Tableau에서는 두 가지 방법으로 테이블 계산을 사용할 수 있습니다. 첫 번째는 일반적으로 사용되는 테이블 계산을 모아놓은 **퀵 테이블 계산**입니다. 이 방법을 사용하면 한 번의 클릭으로 테이블 계산을 정의할 수 있어 시작할 때 사용하기 좋은 방법입니다. 또한 계산 언어에서 **테이블 계산 함수**를 사용하여 고유의 테이블 계산을 만들 수도 있습니다. 통합 문서 작성자는 이러한 함수를 사용하여 결과 집합을 정밀하게 처리할 수 있습니다. 또한 모든 테이블 계산을 같은 계산 언어로 표현할 수 있으므로 더 복잡한 식이 필요한 경우 **퀵 테이블 계산** 중 하나를 시작점으로 사용하여 수동으로 편집할 수 있습니다.

테이블 계산을 사용하면 집계된 데이터의 능숙한 처리, 데이터 구조에 따른 집계 생성과 같이 까다로운 데이터베이스 작업을 몇 번의 클릭이나 간단한 식만으로 수행할 수 있습니다. 따라서 기술 전문가가 아닌 사용자도 이러한 작업을 수행하고 전문가는 많은 시간을 절약하며 힘든 SQL 코드 작업에서 벗어날 수 있습니다.

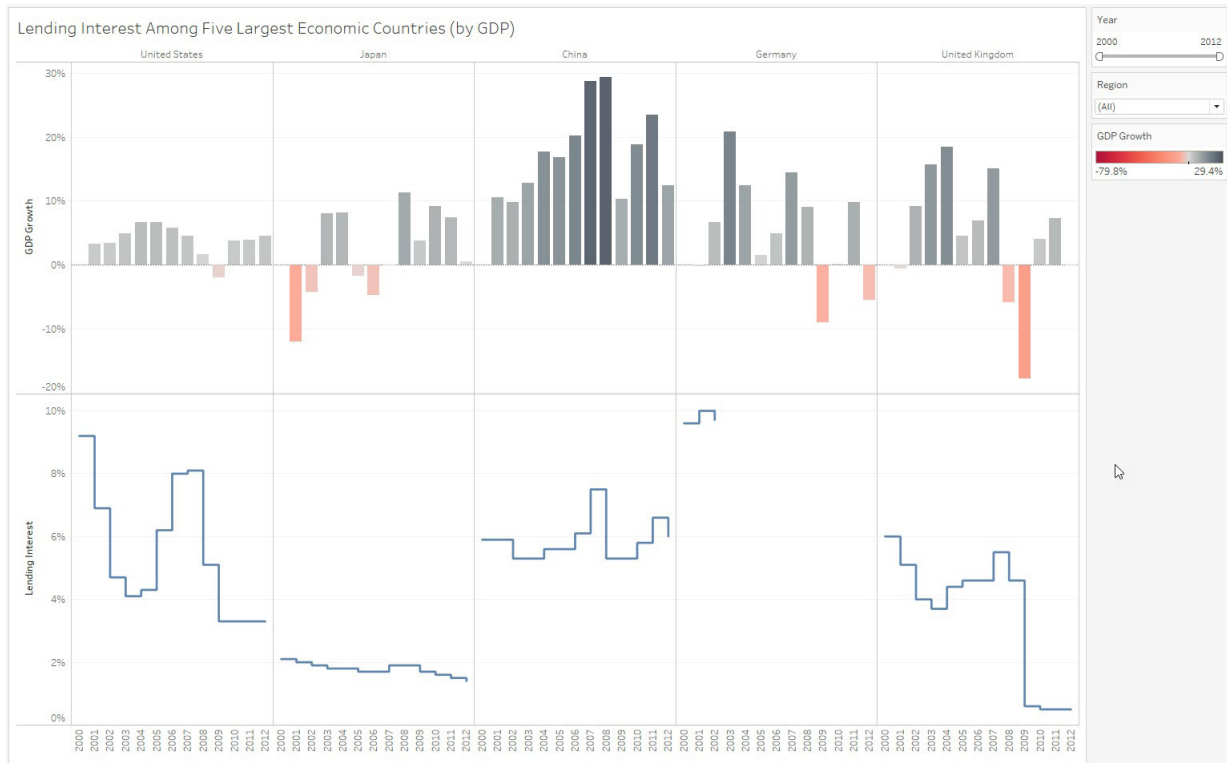


그림 8: Tableau가 기본 제공하는 킷 테이블 계산의 하나(비율 차이)를 사용하여 일부 국가에서 이자율의 인상이 GDP에 어떻게 영향을 미치는지 신속하게 볼 수 있습니다.

4. 시계열 및 예측 분석

기능: 시계열 분석, 예측

센서 판독에서 주식 시장 가격, 졸업률에 이르기까지 전 세계 데이터의 대부분은 시계열을 사용하여 효과적으로 모델링할 수 있습니다. 시간은 그 자체로 분석 프로젝트에 사용되는 가장 일반적인 독립 변수 중 하나입니다. 더 많은 인사이트를 찾으려면, 예측 기능을 활용하는 것이 좋습니다. Tableau를 사용하면, 많은 노력 없이도 예측 분석을 추가하여 다양한 시나리오를 빠르게 탐색할 수 있습니다.

시계열 분석

Tableau의 유연한 프론트엔드 및 강력한 백엔드 덕분에 올바른 질문을 할 수 있도록 시계열 분석이 간단해집니다. 원하는 필드를 해당 뷰로 끌어 놓고 질문 프로세스를 시작하기만 하면 분석이 시작됩니다.

이중 축 기능과 차원화된 집계계를 활용하면 다양한 시계열을 살펴볼 수 있습니다. 시계열을 다양한 세부 수준 및 집계 수준에서 신속하게 검사할 수 있는 기능이 없으면 질문을 생성하지 못할 수도 있습니다.

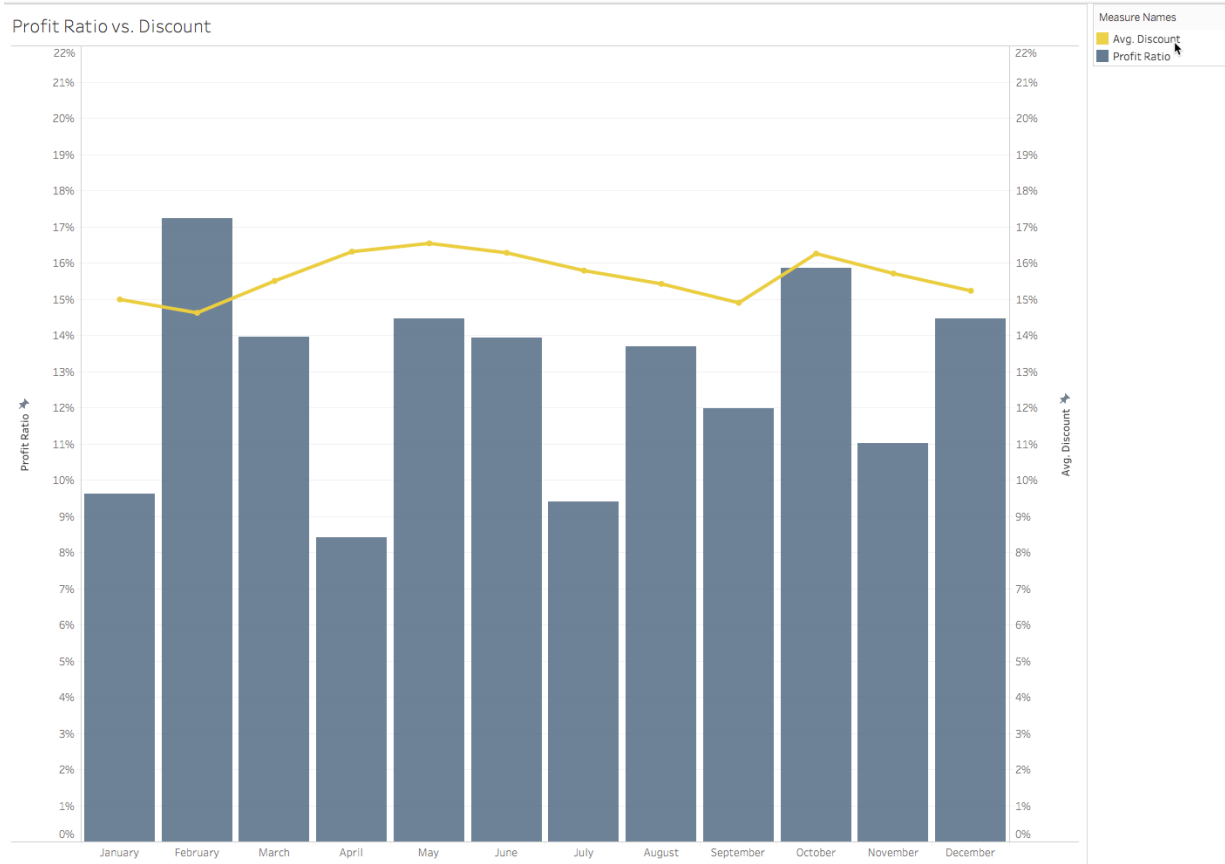


그림 9: 이중 축 플롯은 수익률과 평균 할인 간의 상관 관계를 보여줍니다.

특정 기간으로 살펴보려면 데이터를 정확한 날짜 집합으로 필터링하거나 Tableau의 상대 날짜 필터를 활용하면 됩니다. 상대 날짜 필터를 사용하면 '지난주' 또는 '지난달'과 같이 상대적인 기간으로 살펴볼 수 있습니다. 이러한 기간은 새로운 데이터를 통합하기 위해 뷰를 열 때마다 업데이트되어 강력한 보고 도구의 역할을 합니다.

시계열로 작업할 때에는 종종 다른 시간 계산을 평활화 또는 수행해야 합니다. Tableau에는 이동 평균, 전년 대비 계산 및 누계와 같은 일반적인 시계열 연산을 단순화하도록 설계된 다양한 기능이 있습니다.

이전에 설명한 것처럼 Tableau의 테이블 계산을 사용하면 시계열 조작의 공통 집합에서 선택하거나(쿼리 테이블 계산) 사용자 지정 계산을 작성할 수 있습니다.

Stock Tracking

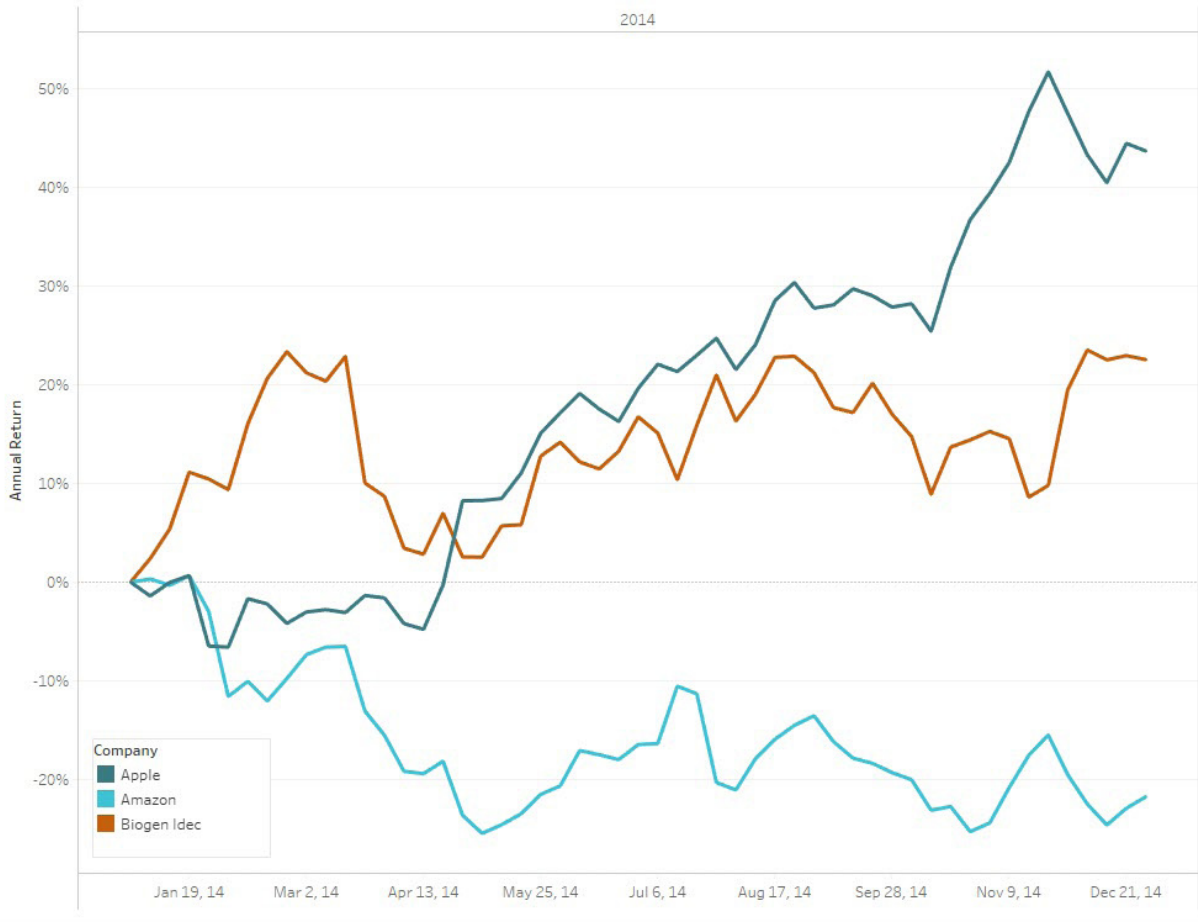


그림 10: 이 시계열 분석은 3개의 주요 회사의 월별 주식 가격을 보여줍니다.

시계열 분석은 매우 일반적이므로 Tableau의 기능은 프로젝트를 더 빨리 완료하고 조직에 더 많은 가치를 제공합니다. 데이터 전문가와 비즈니스 분석가는 직관적인 기능을 통해 자신의 데이터에 대해 더 많은 질문을 효과적으로 할 수 있습니다.

예측

미래를 예측해야 하는 사용자를 위해 Tableau에는 추세 파악 및 예측 등의 몇 가지 기본 모델링 기능이 있습니다.

모든 차트에 빠르게 추세선을 추가하고 마우스 오른쪽 버튼으로 추세선을 클릭하기만 하면 해당값(예: p-값 및 R-제곱)에 대한 세부 사항을 확인할 수 있습니다. 추세선이 프론트엔드에 완전히 통합되어 쉽게 세분화할 수 있으므로 Tableau의 끌어 놓기 기능을 사용하면 한 번의 클릭으로 여러 그룹을 모델링할 수 있습니다. 그림 11에서 볼 수 있듯이 Tableau는 코드 없이도 서로 다른 세그먼트에 대해 세 개의 추세선을 자동으로 만듭니다. Tableau는 대수, 다항식, 지수 등 기타 여러 가지 유형의 적합성을 지원합니다.

Obesity's correlation with diet, exercise, and smoking (% of pop.) - All

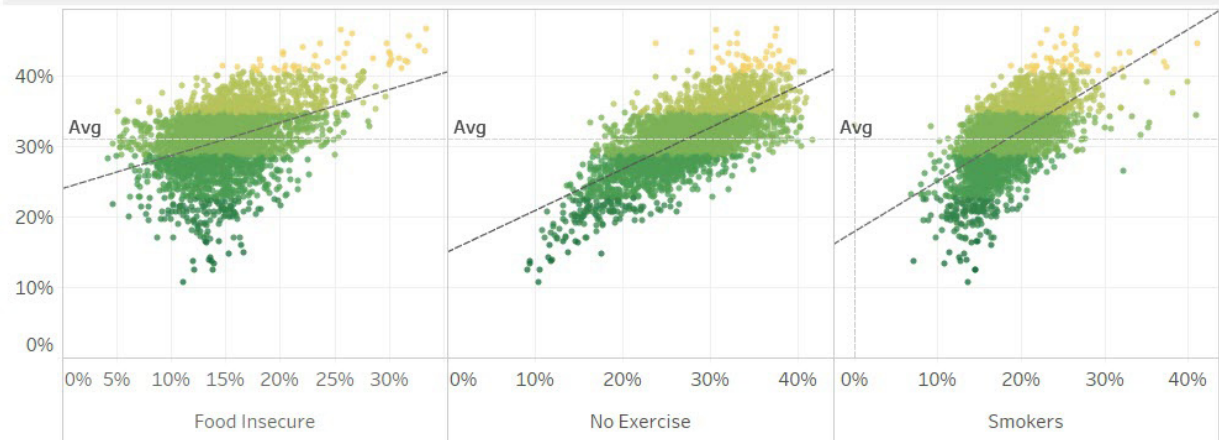


그림 11: 추세선이 비만과 식량 공급 불안성, 운동 및 흡연의 관계를 하이라이트합니다.

기본적으로 Tableau의 예측 기능은 백그라운드에서 여러 모델을 실행하고 최상의 모델을 선택하여 계절성과 같은 데이터 문제를 자동으로 보고합니다. Tableau의 예측에서는 **지수평활법 모델**이라는 기술을 사용합니다. 지수평활법 모델은 과거 값의 가중치 평균을 사용하여 시계열의 미래 값을 반복적으로 예측합니다. 앞에서 살펴보았듯이, 예측에 관한 거의 모든 것을 구성할 수 있습니다. 예를 들어, 예측 기간, 모델의 유형(덧셈 또는 곱셈), 그리고 계절적 변동 고려 여부를 선택할 수 있습니다.

이 기능은 또한 매우 쉽게 사용할 수 있기 때문에 초보 사용자도 기본 설정을 사용하여 몇 번의 클릭만으로 예측을 만들 수 있으며 고급 사용자는 모델의 거의 모든 측면을 구성할 수 있습니다. Tableau는 통계 요소 외에도 신뢰 구간을 표시하여 초보 사용자에게 예측 품질에 대한 추정치를 제공합니다. 또한 예측은 나머지 Tableau 기능과 원활하게 결합되므로, 사용자 인터페이스의 다른 분석 객체와 마찬가지로 예측을 쉽게 세분화하고 능숙하게 처리할 수 있습니다.

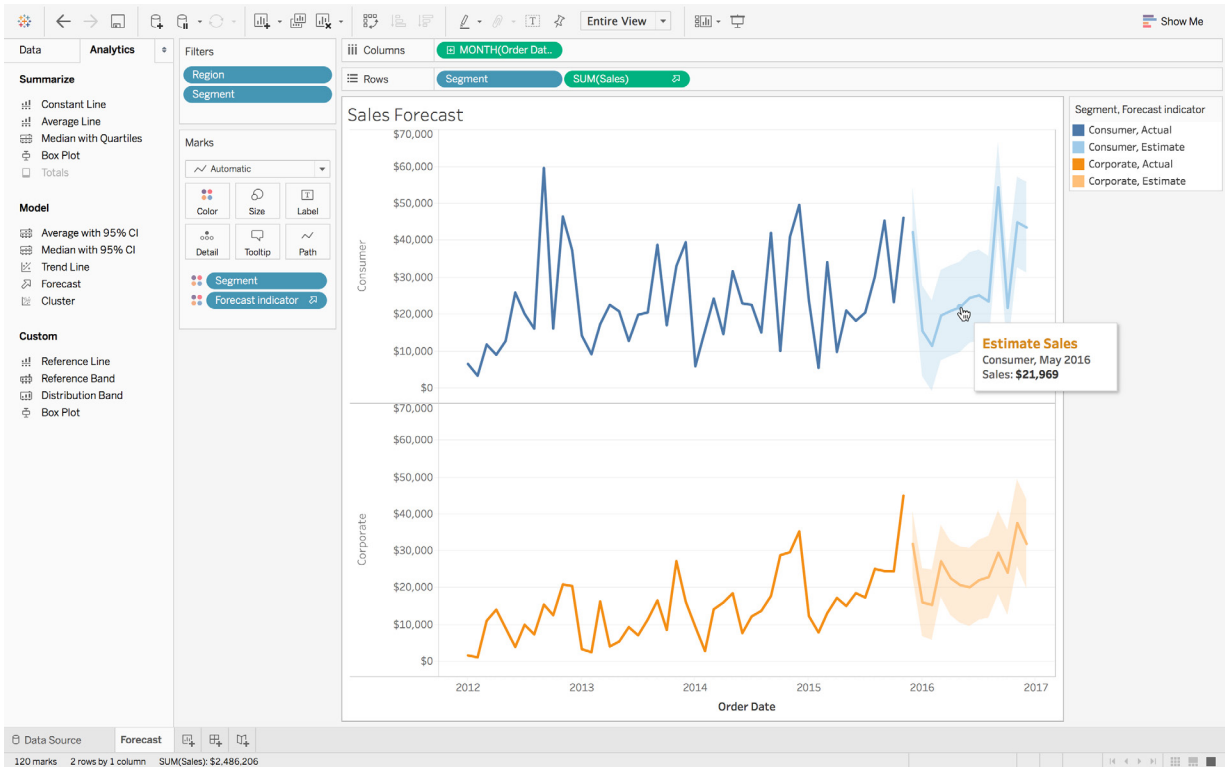


그림 12: 예측 기능은 세그먼트별 판매를 자동으로 예측합니다.

손쉬운 예측 분석을 통해 모든 데이터 프로젝트에 엄청난 가치를 제공합니다. 복잡한 구성과 단순한 대화형 모델링을 모두 지원하므로 데이터 과학자와 최종 사용자 모두 플랫폼을 원활히 사용할 수 있습니다.

5. 외부 서비스 통합

기능: Python, R, MATLAB 통합

많은 조직에서 분석 플랫폼과 제도적 지식에 투자하고 있으므로, 여러분의 조직에도 매우 구체적인 요구 사항과 기존 작업에서 얻은 유용한 자료가 있을 것입니다.

포괄적인 분석 플랫폼이라면 다른 고급 분석 기술을 통합할 수 있는 기능을 지원하여 조직에서 수행 가능한 기능을 확장하고 다른 솔루션에서 기존 투자를 활용할 수 있도록 해야 합니다. Tableau를 Python, R, MATLAB과 같은 외부 서비스와 통합할 수 있습니다. Tableau는 사용 용이성 덕분에 조직 내 모든 사용자가 고급 알고리즘을 간단히 끌어 놓기 방식으로 사용하게 함으로써 데이터 과학을 대중화할 수 있습니다.

Tableau는 Python, R 및 MATLAB을 직접 통합하여 기존 모델을 사용하는 사용자를 지원하고 전 세계 통계 커뮤니티를 활용합니다. Tableau는 RServe 및 TabPy와 같은 외부 예측 서비스로 스크립트 및 데이터를 전송합니다. 그 결과가 Tableau로 반환되면 Tableau 시각화 엔진에서 사용할 수 있습니다. 이를 통해 사용자는 Tableau의 데이터에 대해 R 또는 Python에 사용할 수 있는 모든 함수를 호출할 수 있으며, Tableau를 사용하여 이러한 환경에서 만든 모델을 효과적으로 처리할 수 있습니다.



그림 13: Tableau에서 시각화한 R을 사용한 k-평균 클러스터 분석 알고리즘의 예입니다. 클러스터가 세 종류의 아이리스 꽃에 대한 매개 변수(꽃받침과 꽃잎의 길이와 폭)의 차이를 보여줍니다.

모델링은 기본 통계보다 훨씬 심도 있게 수행할 수 있습니다. R 및 Python 통합을 사용하면 텍스트를 분석하여 감성을 추출하거나(그림 14) 공급망 또는 주식 포트폴리오 최적화와 같은 규범적 분석 결과를 표시할 수 있습니다. 통합을 통해 Tableau 내부에서 R, MATLAB 또는 Python 코드를 직접 실행할 수 있으며 Tableau 외부에 정의된 모델 및 스크립트도 지원됩니다.

Comment	Sentiment
I had a pleasant experience. Responsive customer service, great product.	Positive
This is an awesome product. I enjoy using it every day.	Positive
Crisp and bright screen, well worth the money.	Positive
What a joy to use.	Positive
It is OK for what it costs. Not great but works as expected.	Negative
Much better than the competing products for a much more reasonable price.	Positive
I would recommend this product to anyone. I will buy another one myself.	Positive
Don't waste your time on this product. One word. Bad.	Negative
Do not buy this product. It only lasts 2 weeks. It is garbage. And they won't accept returns.	Positive
Customer service was unresponsive. They wouldn't allow me to return the defective item.	Negative
Took 4 weeks to receive it even though I paid for 2 day delivery. What a scam.	Negative
Product description is misleading. It is much smaller than what's shown in the picture and ugly. Not worth the money.	Negative
My experience was terrible. The product was dead on arrival.	Negative

그림 14: Tableau와 R 통합으로 작성한 정서 분석의 예입니다.

그림 15에서는 그라디언트 부스팅이라는 앙상블 기계 학습 기술을 사용하여 과거 암 발병 사례를 바탕으로 유방 종양에 대한 진단 의견을 제공하는 대시보드가 표시되어 있습니다. 데이터 과학자는 Python 프로그래밍 환경에서 이 모델을 훈련하고 정확성에 대해 통계적으로 평가한 다음 TabPy(Tableau Python Server)에 게시합니다. 대시보드에는 사용자가 다양한 의료 지표 값을 입력할 수 있는 양식 입력란이 있는데, 그 후에 해당 값은 이 호스팅된 기계 학습 모델로 전달되어 실시간으로 진단 의견으로 전환됩니다. 이를 통해 다수의 분석가 및 대시보드가 공유 기계 학습 자산을 구축하고 데이터 과학자가 자신의 선택 도구를 사용하여 단일 위치에서 해당 자산을 쉽게 관리할 수 있습니다.

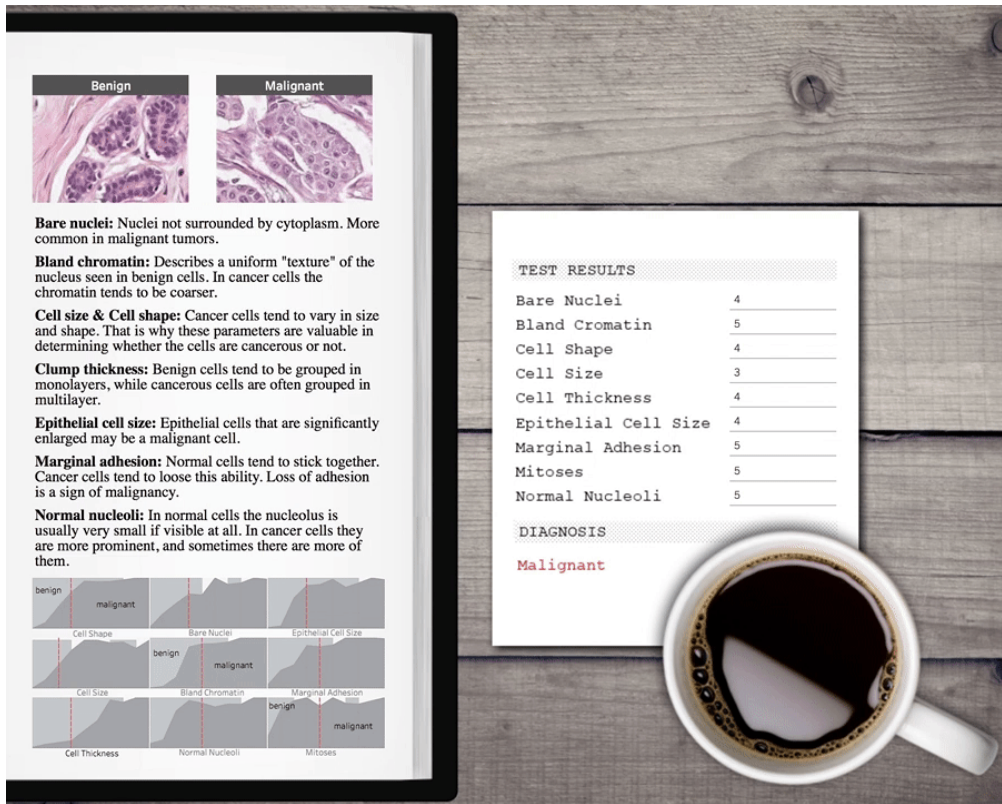


그림 15: 입력 매개 변수가 있는 앙상블 기계 학습 모델입니다.

Tableau가 Python, R 또는 MATLAB과 결합하면 아주 강력해집니다. 이를 통해 기술 수준에 관계없이 모든 사용자가 사용할 수 있도록 복잡한 모델링 결과를 아름답고 매력적인 비주얼로 표현할 수 있으며, 사용자는 대시보드에 포함된 제어 기능으로 상호 작용하고 가정 질문을 하고 가설 시나리오를 평가할 수 있습니다.

결론

Tableau는 여러 면에서 독보적인 분석 플랫폼입니다. Tableau는 인간의 지능을 높이려는 사명감을 바탕으로 비즈니스 사용자와 데이터 과학자 모두를 위해 설계되었습니다. 사용자가 가능한 한 빨리 데이터에 대해 흥미로운 질문을 할 수 있도록 집중적으로 지원함으로써 모든 수준의 사용자에게 유용한 기능을 제공하는 플랫폼을 구축했습니다.

Tableau의 유연한 프론트엔드를 통해 비즈니스 사용자는 데이터베이스를 코드화하거나 이해할 필요 없이 질문을 할 수 있습니다. 또한 Tableau에는 데이터 과학자에게 강력한 무기가 되어주는 심도 깊은 분석이 있습니다. 데이터 과학자는 정교한 계산, R 및 Python 통합, 신속한 동질 집단 분석, 예측 기능을 활용하여 Tableau에서 복잡한 분석을 완료하고 비주얼로 표현된 결과를 쉽게 공유할 수 있습니다. Tableau를 데이터 탐색 및 품질 관리에 사용하든 모델 설계 및 테스트에 사용하든, 어떤 경우에도 대화형 플랫폼의 특성 덕분에 프로젝트 수명 전반에 걸쳐 시간이 크게 단축됩니다. Tableau는 모든 수준에서 분석을 더 쉽게 이용하고 신속하게 완료하도록 함으로써 기업 전체에서 중요한 협업과 개선된 의사 결정이 가능하도록 지원합니다.

Tableau 정보

Tableau는 정보에서 의미를 추출하도록 지원합니다. Tableau는 분석 주기를 지원하고, 시각적 피드백을 제공하며, 진화하는 복잡성 속에서 질문의 답을 찾도록 돕는 분석 플랫폼입니다. 데이터에서 혁신을 이루려면 계속해서 새로운 질문을 던지며 관점을 바꾸도록 지원하는 응용 프로그램이 필요합니다. 데이터로 영향력을 발휘할 준비가 되셨다면, 지금 Tableau Desktop [무료 평가판](#)을 다운로드하십시오.

관련 백서

[R과 Tableau 사용](#)

[세부 수준 식의 이해](#)

[분석의 정의](#)

[모든 백서 보기](#)

추가 리소스 탐색

[제품 데모](#)

[교육 및 자습서](#)

[커뮤니티 및 지원](#)

[고객 스토리](#)

[솔루션](#)

