

추세선

추세선에 대해 설명하는 동영상에 오신 것을 환영합니다.

예제 통합 문서를 다운로드하여 사용자의 시스템에 설치된 Tableau 에서 직접 따라 해보실 수 있습니다.

추세선 추가

추세는 분석에 대한 중요한 인사이트를 제공할 수 있습니다. 예를 들어 '풍속에 따라 전기 출력이 얼마나 상승하는지?'에 대한 질문에 답을 할 경우 풍속이 증가하면 전기 출력도 증가한다는 관계를 확인할 수 있습니다. 하지만 얼마나 정확할까요?

뷰에 추세선을 쉽게 추가할 수 있습니다.

- 분석 패널을 클릭하고 원하는 모델 유형으로 추세선을 드래그합니다.
- 추세선 삭제도 마찬가지로 간편합니다. 뷰 밖으로 드래그하면 됩니다. 하지만 이 예에서는 추세선이 필요하므로 실행을 취소하겠습니다.

추세선 옵션

기본적으로 추세선은 패널별 및 색상별로 표시됩니다.

- 데이터 탭으로 돌아와 위치 등과 같은 다른 차원을 뷰에 표시하려는 경우 추세선이 패널별 또는 분산형 차트별로 표시됩니다.
- 이와 유사하게 위치를 색상으로 이동하는 경우 추세선이 3 개로 표시됩니다.
- 위치를 색상에 표시하려는데 전체 추세가 하나만 있는 경우 추세선을 수정할 수 있습니다.

추세선을 편집하려면 마우스 오른쪽 버튼을 클릭하여 추세선을 선택하고 추세선 편집을 선택하면 됩니다.

- '색상별 추세선 허용'을 선택 취소하면 단일 전체 추세선으로 돌아올 수 있습니다.
- 또한, 신뢰 구간 표시를 선택 취소하여 뷰를 단순화할 수 있습니다.

이 대화 상자에서 여러 작업을 수행할 수 있습니다.

- 먼저 모델 유형입니다.
 - 이러한 옵션은 분석 패널에서 추세선을 처음 표시할 때의 옵션과 동일합니다.
 - 이러한 옵션을 사용하면 Tableau 에서 하나 또는 두 변수 모두의 전환 유형을 기반으로 선형 회귀 모델을 구축합니다.
 - 선형은 변수 관계가 아닌 계수를 말합니다.

- 모델 유형과 전환에 대한 자세한 내용은 [추세선 모델 유형](#)에 대한 온라인 도움말 문서를 참조하시기 바랍니다.
- 신뢰 구간은 모델의 95% 신뢰 구간을 표시합니다.
- Y 절편을 0 까지 강제로 만들 수도 있습니다.

추세선의 유의성

추세선이 가치 있는 정보를 제공하는지 여부를 평가하는 것은 중요합니다.

- 추세선에 마우스오버하면 추세선 수식과 p-값 및 R-제곱이 있는 도구 설명이 표시됩니다.
- 통계에서 p-값은 유의성이라는 개념에 할당되는 숫자입니다.
 - p-값이 일반적으로 0.05 인 한계값보다 적은 경우 결과가 유의하다고 해석됩니다.
 - 높은 p-값(0~1 의 범위 기준)은 데이터 추세가 모델의 요소로 인해 발생한 것이 아니라 우연이라는 것을 의미할 수 있습니다.
- 이 예에서 추세선에는 매우 작은 p-값이 있으며 이는 바람직한 상태입니다.

하지만 모델이 적합한지 정확하게 평가하려면 p-값 이외의 다른 값도 알아야 합니다.

- 여기에는 이 모델이 데이터에 얼마나 적합한지를 효과적으로 보여주는 R-제곱도 있습니다. R-제곱 값은 0-1 범위일 수 있으며 값이 클수록 더 좋게 평가됩니다.
 - 이 예에서는 R-제곱이 0.956 으로 매우 높습니다.
 - 이는 모델이 데이터에 매우 적합하다는 것을 의미합니다. R-제곱이 1 이면 완벽하게 적합한 상태이지만 R-제곱이 0.999 처럼 상당히 높은 경우 모델이 잘못된 것일 수도 있습니다. 일반적으로 R-제곱 값이 비정상적으로 너무 높으면 자유도가 낮거나 관찰값이 너무 많은 것입니다.

추세선의 나머지

추세선이 데이터를 정확하게 나타내는지 파악하려면 p-값이 작거나 R-제곱이 큰 것만으로는 부족합니다.

모든 요소가 예측된 추세선에 일치하는 것은 아닙니다. 특정 포인트와 예측값 사이의 거리를 오차 또는 나머지라고 합니다.

올바른 모델인 경우 나머지를 설명 변수를 기준으로 배치할 때 기준선 주위에서 임의의 정규 분포를 따르게 됩니다. 나머지 차트가 정규 분포가 아닌 경우 데이터가 예측 값과 일치하지 않는 추세가 존재한다는 것을 의미합니다. 따라서 해당 모델은 유용하지 않을 수 있습니다.

추세선이 있는 뷰의 나머지 값을 확인하려면

- 워크시트 > 내보내기 > 데이터로 이동합니다.
- 파일을 저장하라는 메시지가 표시됩니다. 유일한 형식 옵션은 Microsoft Access 입니다. 여기서 이름을 추세선의 나머지로 변경하고 저장을 눌러 보겠습니다.

- '내보낸 후 연결'을 선택하겠습니다.
- 이 데이터 원본에는 분산형 차트의 원래 데이터뿐만 아니라 추세선의 예측 값 및 나머지 값도 포함됩니다.

나머지 차트는 설명 변수가 가로 축(풍속은 열에 해당)으로 구성되며, 나머지가 세로 축(행)으로 구성됩니다. 풍차를 세부 수준으로 이동해 보겠습니다.

우수한 모델은 0 을 기준으로 한 정규 분포를 따릅니다.

현재 모델은 풍속 값을 기반으로 전기를 예측하는 데 유용하지 않습니다. 추세선의 p-값과 R-제곱은 맞지만 나머지 차트가 잘못되었습니다.

결론

지금까지 Tableau 의 추세선 및 관련 통계를 해석하는 방법에 대한 개요를 살펴보았습니다. 시청해 주셔서 감사합니다. 무료 교육 동영상을 통해 더 자세한 내용을 알아보시기 바랍니다.