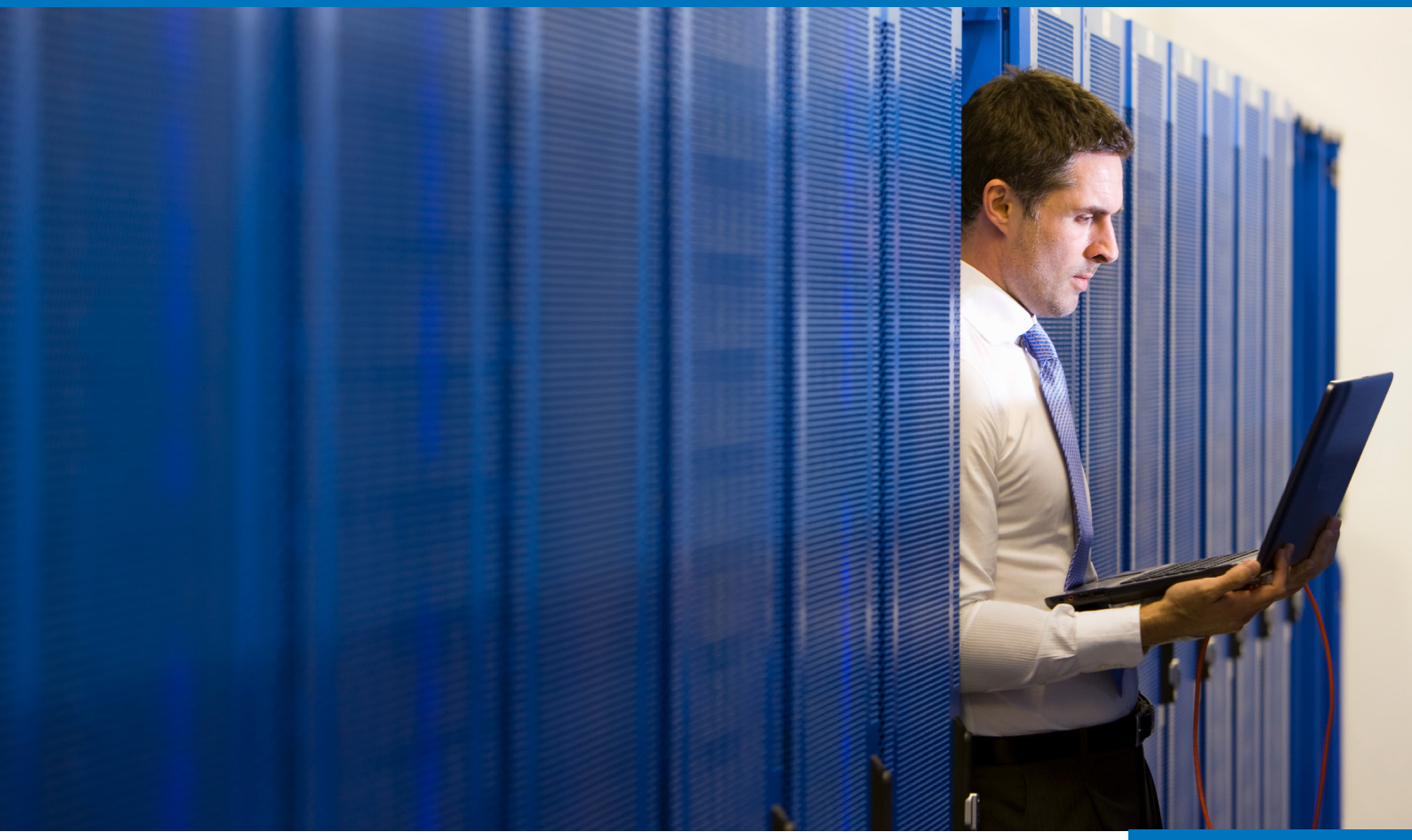
백서

IT 성공 비결: 선택과 통제 사이의 균형

분석 라이프사이클의 모든 단계를 포괄적으로 지원하는 통합 플랫폼으로서 분석 프로세스 조율



sas
THE POWER TO KNOW.

분석은 많은 기업의 운영 모델에서 중요한 비중을 차지합니다. 종사하는 산업에 관계없이 모든 기업은 데이터와 분석에서빠른 시간 내에 인사이트를 도출해야 경쟁 우위를 점할 수 있습니다. 많은 분석 솔루션 제공업체가 자사 제품의 장점이라고 강조하는 부분이기도 합니다.

그러나 전례 없는 경쟁을 경험하고 있는 오늘날의 분석 시장에서 그 말을 곧이곧대로 믿기는 어렵습니다. 기업 지도자는 데이터에서 진정한 가치를 창출하려면 무엇이 필요하고, 진화하는 기술을 활용해 기업을 성공적으로 이끌려면 무엇이 필요한지 고민해볼 필요가 있습니다.

기업이 분석 기술을 시험해보고 궁극적으로 도입하는 경우 그 결과에 대한 책임은 고스란히 IT 부서가 떠안아야 합니다. IT 부서는 많은 데이터에 안전하게 액세스할 수 있는 환경, 분석하고 테스트하기에 안정적인 환경, 모델을 신속하게 구현하고 관리할 수 있는 환경을 제공해야 합니다.

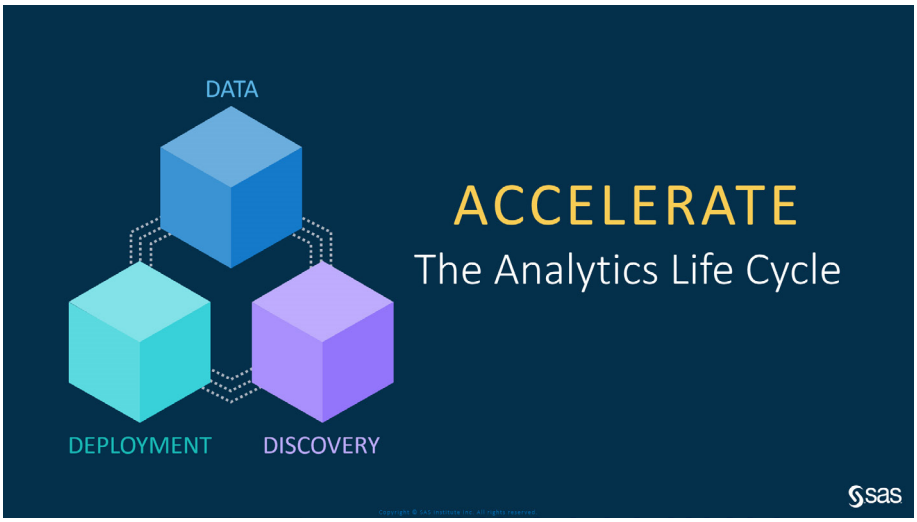
IT 부서가 비즈니스를 제대로 지원하려면 데이터, 기술, 분석, 프로그래밍 언어에 대한 다양한 옵션을 제공해야 합니다. 즉, 사용자가 이러한 기술을 자신에게 익숙한 방식으로 사용할 수 있는 환경을 조성해야 합니다. 또한 분석 결과를 신뢰하고 신속하게 사용할 수 있도록 보안 및 거버넌스 기술을 구현해야 합니다. 확장성과 탄력성을 갖춘 환경 역시 구현되어야 합니다. 분석 에코시스템이 복잡해지고 혼란스러울수록 많은 조율, 연결, 감독이 필요합니다. 이런 혼란을 바로잡는 것이 진정한 분석 플랫폼의 역할이기도 합니다. 최고의 플랫폼을 사용하면 분석 에코시스템의 모든 요소를 조율하고 연결하여 보다 빠른 시간 내에 결과를 얻을 수 있습니다.

분석에 고급 분석과 AI를 활용할 경우 고성능 분석 플랫폼이 훨씬 더 뛰어난 성능을 발휘합니다. AI 시스템이 제 기능을 발휘하려면 대량의 데이터, 다양한 고급 분석 기술, 다수의 모델을 구현 및 관리 능력이 뒷받침돼야 합니다. SAS® 플랫폼을 도입하면 분석 라이프사이클의 모든 측면에서 복잡성을 손쉽게 해소할 수 있습니다. 복잡한 분석일수록 더 많은 공을 들여서 더 짧은 시간 내에 결과를 도출할 수 있어야 합니다. SAS 플랫폼은 바로 이 부분에서 진가를 발휘합니다.

SAS는 고급 분석 및 예측 분석 분야에서 오랫동안 1위 자리를 고수해왔으며 IDC에 따르면 30%의 고급 분석 및 예측 분석 제품 시장 점유율을 차지하고 있습니다.

(Gopal, Chandana. *Worldwide Advanced and Predictive Analytics Software Market Shares, 2017: Open Sources Continues to Threaten* . IDC 보고서, 2018년 6월).

분석 라이프사이클: 개선과 촉진



분석의 가시적인 가치는 모델의 분석 결과를 실행에 옮길 때 발생합니다. 분석 라이프사이클(데이터 수집, 탐색, 모델 구현)이 완료되는 데 짧게는 몇 주에서 길게는 몇 개월이 걸리는 경우가 많습니다. 그러나 그 무렵이면 데이터가 오래되어 비즈니스 가치를 상실합니다. 데이터 준비, 모델 제작 및 비교, 모델 준비 및 구현에 너무 많은 시간이 걸립니다. 모델의 예측 정확도를 모니터링하는 데 더 많은 시간이 낭비됩니다. 이에 SAS는 기업이 분석에 대한 투자 효과를 거둘 수 있도록 분석 라이프사이클을 촉진하는 데 힘쓰고 있습니다. 이를 위해서는 기업이 아이디어 착안 단계에서 모델 제작 및 구현 단계로 신속하게 넘어가되, 신뢰할 수 있는 분석 결과를 확보하고 이를 제대로 이해할 수 있어야 합니다.

데이터 수집, 데이터 관리, 프로그래밍 언어, 분석 기술에 존재하는 선택지가 많다 보니 분석 라이프사이클을 효율적으로 진행할 방법을 찾기가 어려울 수 있습니다. 거버넌스, 보안, 안정성을 확보해야 하는 경우 특히 그렇습니다. 분석 결과를 신속하게 실행에 옮기려면 모든 요소를 조율해야 합니다. 그래야 데이터 처리 및 사용 과정에서 흔히 발생하는 실수를 줄이고 일선 부서 및 IT 부서 관계자 간의 협업과 거버넌스를 지원함으로써 모델링 프로세스에 수반되는 위험과 수고를 덜 수 있습니다.

분석 플랫폼은 데이터 수집에서 시작하여 가시적인 결과로 귀결되는 과정을 조율하는 데 유용합니다.

분석 라이프사이클: 데이터

분석가는 분석을 위한 데이터를 최대한 빠르고 쉽게 확보할 수 있기를 원합니다. 또한 분석가는 분석하기 적합하게 데이터를 준비하는 데만 약 80%의 작업 시간을 소모합니다. 데이터 수집 및 준비 작업을 간소화하는 분석 플랫폼을 사용하면 적절한 통제력을 유지하고 데이터 품질을 개선함과 동시에 분석 인력의 생산성과 작업 속도를 크게 개선할 수 있습니다. 다음과 같은 사항을 고려해야 합니다.

- 데이터 이동: 네트워크를 통해 이동해야 하는 데이터가 많을수록 분석 프로세스의 속도가 느려집니다. 극단적인 경우로 너무 많은 데이터를 전송하게 되면 네트워크가 마비될 수도 있습니다. 한 가지 선택지는 데이터가 저장되어 있는 곳으로 분석 프로세스를 이동하여 분석 작업을 수행하는 것입니다. 또 다른 선택지는 여러 프로세스가 효율적으로 데이터를 공유할 수 있도록 데이터를 메모리로 이동하는 것입니다. IoT 데이터의 경우 센서를 통해 스트리밍되는 시점에 데이터를 처리하는 것도 좋은 방법입니다. SAS 플랫폼은 이 모든 옵션을 지원하므로 고객이 가장 적합한 방법을 선택할 수 있습니다.
- 새로운 유형의 데이터: 데이터는 끊임없이 변화하며 진화하고 있습니다. 새로운 유형의 데이터로는 센서 데이터, 비디오, 오디오, 텍스트, 이미지 등이 있습니다. 그러나 내일이면 또 달라질지도 모릅니다. 새로운 데이터 유형을 지속적으로 수용하고 활용할 수 있는 분석 플랫폼이 필요합니다.
- 대량의 고품질 데이터 확보: 고급 분석 기술은 모델 학습에 대량의 데이터를 필요로 합니다. AI는 다른 어떤 기술보다 더 많은 양의 매우 복잡한 데이터를 사용하므로 가장 많은 데이터를 필요로 하는 기술로 손꼽힙니다. AI 모델 제작 프로세스를 자동화할 수 있지만 정확한 결과를 확보하려면 데이터의 품질이 우수해야 합니다.

분석가는 모든 유형의 데이터(정형, 비정형, 스트리밍)를 확보하여 모델을 학습시킬 수 있어야 합니다. 한 가지 방법은 분석 프로세스와 연결된 중앙 리포지토리를 이용하는 것입니다. 이 방법은 다양한 데이터 출처를 셀프 서비스 방식으로 확보하는 데 효과적입니다. 리포지토리는 주기적으로 업데이트되며, 데이터 창구가 일원화되므로 분석 결과를 신뢰할 수 있을 뿐더러 재사용할 수도 있습니다. 이를 위해선 누가 데이터에 액세스하고 데이터를 어떻게 사용하는지 명확하게 알 수 있도록 리포지토리에 보안 체제가 구현되어 있어야 합니다.

SAS 플랫폼을 사용하면 데이터를 적절하게 보호하고 관리하는 데 필요한 통제력과 거버넌스 체제를 유지하면서 분석가의 업무 부담을 완화하고 분석가에게 사용할 수 있는 데이터에 대한 선택지가 주어지는 환경을 구현하기 용이합니다. SAS의 데이터 관리 솔루션은 분석하기 적합한 데이터를 생성하는 데 대단히 효과적이며 관계도, 통합, 품질 개선, 메타데이터, 즉석 데이터 랭글링을 지원합니다.

SAS는 관계형 및 비관계형 데이터베이스, PC 파일, Hadoop, Amazon Redshift, 데이터 웨어하우스 어플라이언스를 비롯한 60가지 이상의 출처에 저장된 데이터를 손쉽게 읽거나, 쓰거나, 업데이트할 수 있는 솔루션을 제공합니다. 일선 부서 직원은 최소한의 교육만 받고도 SAS의 솔루션을 손쉽게 사용할 수 있습니다. IT 부서 직원은 더욱 다양한 기능들을 SQL 같은 기본 인터페이스를 사용하여 활용할 수 있습니다. SAS 소프트웨어는 데이터베이스 쿼리, 조인 및 기타 함수를 분석에 사용하는 데이터 출처에 전송하므로 네트워크 트래픽이 감소하고 데이터 처리 시간이 단축됩니다. 데이터 병렬 전송을 통해 로드 시간을 크게 단축하고 클라우드로 전송되는 맞춤형 데이터의 보안과 형식을 최적화할 수 있습니다.

LDAP이 내장된 SAS 플랫폼은 기업의 기존 보안 인프라와 통합하여 사용할 수 있으며 기존의 사용자와 그룹이 사용하는 데이터를 한결 수월하게 모니터링, 감사, 보호할 수 있습니다. 인데이터베이스 및 인메모리 처리 옵션을 사용하면 데이터 이동을 줄이고 분석 라이프사이클의 첫 단계인 데이터 수집 속도를 개선할 수 있습니다.

분석 라이프사이클: 탐색

이 단계에서 분석가는 데이터를 탐구 및 분석하여 특정 문제에 관한 인사이트를 발견합니다. 예측 정확도가 가장 우수한 모델을 확보하기 위해 많은 알고리즘이 반복해서 데이터를 처리할 것을 요구하기 때문에 이 작업은 컴퓨팅 집약적으로 이뤄집니다. AI 모델에서 이런 현상이 특히 두드러지는데, 그 이유는 모델 학습 및 스코어링 과정에 사용되는 정교한 반복 기법과 데이터 복잡성 및 크기 때문입니다.

또한 분석 워크로드 유형도 고려해야 합니다. 예를 들어, 분석 워크로드에 따라 가용성이 상시 유지되는 환경이 필요하거나, 신속하게 프로비저닝할 수 있는 환경이 필요하거나 또는 특정 프로젝트에 맞게 확장할 수 있는 환경을 구현해야 합니다.

한 가지 분석 언어만 사용하던 시대는 지났습니다. Python, R, Lua 같은 오픈 소스는 분석 툴킷의 핵심 구성 요소가 됐습니다. SAS 플랫폼은 다양한 프로그래밍 언어를 지원합니다. 이러한 기능을 통해 사용할 수 있는 알고리즘과 채용할 인재에 대한 선택의 폭이 넓어집니다. 그러나 다양한 언어를 작업에 사용하는 경우 조율하기 어려울 수 있습니다. 예를 들어, 다음과 같은 고민을 해결해야 합니다. 어떻게 해야 서로 다른 언어가 사용된 알고리즘을 효율적으로 비교하여 예측 정확도가 가장 우수한 모델을 파악할 수 있는가? 어떻게 해야 모델링 프로젝트를 확장할 수 있는가? 어떻게 해야 서로 다른 언어로 제작된 모델을 구현하고 모니터링할 수 있는가? 어떻게 해야 분석가들이 협업할 수 있는가? 투자 효과를 극대화하려는 기업은 모든 모델링 프로세스를 통제하고 효율적으로 처리할 방법을 강구해야 합니다.

SAS 플랫폼을 도입하면 원하는 언어로 분석 모델을 제작하고 SAS 공통 컴퓨팅 엔진을 사용하여 모델링 프로세스를 처리할 수 있습니다. 다시 말해서, 서로 다른 언어로 제작된 모델을 비교하기 용이하므로 예측 정확도가 가장 우수한 모델을 선택하여 구현할 수 있습니다.

분석 라이프사이클: 모델 구현

운영 프로세스 범위 내에 분석 모델을 구현하면 투자 가치가 실현됩니다. 요컨대, 고객 만족도 및 운영 민첩성이 향상되고, 프로세스가 더욱 간결해지며, 자산을 보다 효율적으로 활용할 수 있습니다. 분석 모델을 구현하는 데 소요되는 시간이 짧을수록 기업은 더 빠르게 데이터에서 가치를 창출할 수 있습니다.

그러나 많은 기업이 구현 단계에서 분석 모델링 프로세스가 크게 둔화되는 어려움을 겪고 있습니다. 탐색 단계와 구현 단계가 명확하게 구분되어 있지 않은 탓일 수 있습니다. 또는 자동화가 미흡하거나, 모델 개발자와 IT 아키텍트 간의 협업이 제대로 이뤄지지 않는 탓일 수도 있습니다. 특히 운영상의 의사결정에 예측 모델을 사용하고 있는 경우, 대다수 기업의 구현 환경은 탐색 환경과 판이합니다. 종종 IT 부서는 비즈니스와 관련된 계약조건을 이행하기 위해 엄격한 거버넌스 정책을 구현 환경에 적용해야 합니다.

여러 지역으로 모델의 사용 범위를 넓히기 위해 코드를 수정하거나 보완(Refactoring)하는 경우가 많습니다. 운영 시스템이 지원하는 언어로 모델의 코드를 수정하는 작업은 비효율적이며 많은 시간을 요합니다. 자동화되어 있지 않은 경우 전담 인력을 동원해야 하기 때문에 IT 직원의 부담이 가중될 수 있습니다.

또한 투명하고 통제된 프로세스는 모든 관계자, 특히 감사 기관에 중요한 영향을 미칩니다. 모델을 구현하고 나면 모델을 테스트하고 운영 시스템에 저장된 실제 데이터를 모델에 사용할 준비가 완료되었다고 선언해야 합니다.

모델의 예측 정확도를 유지하려면 모델을 지속적으로 모니터링하고 관리해야 합니다. 고급 예측 분석 모델은 구현만 하면 끝나는 기술이 아닙니다. 모델을 지속적으로 모니터링하여 기대한 수준의 예측 정확도를 유지하는지 확인하고 필요한 경우 다시 학습시켜야 합니다. AI를 사용하는 경우 재학습이 더욱 중요합니다. 모델을 지속적으로 학습시키고 진화할 수 있도록 데이터를 제공해야 합니다.

이 프로세스는 자동화되어 있어야 합니다. 모델이 수백 개 또는 수천 개로 늘어난 경우에 특히 그렇습니다. SAS 플랫폼의 통합 모델 관리 기능을 사용하면 자동화되고 지속 가능한 방식으로 모델을 제작, 등록, 비교, 구현, 모니터링할 수 있습니다. 코드를 수정하는 데만 몇 주나 낭비할 필요 없이 단 몇 번의 클릭만으로 간단하게 모델을 구현할 수 있습니다. SAS 플랫폼은 코드를 수정할 필요 없이 애플리케이션, 데이터베이스, API 또는 웹 서비스를 통해 모델을 구현할 수 있는 통합 코드 기반을 제공합니다. SAS 플랫폼의 이벤트 스트림 프로세싱 기능을 사용하는 기업은 모델을 스트리밍 데이터에 구현하여 변화하는 상황에 즉시 대처할 수 있습니다.

엔터프라이즈 아키텍처에 유연한 분석 플랫폼 추가

오늘날의 IT 환경은 끊임없이 변화하고 있습니다. 온프레미스에서 클라우드 아키텍처나 하이브리드 환경으로 이전하는 기업이 있는가 하면, 컨테이너를 사용하는 기업이 있고, 인프라를 직접 관리하기를 원하는 기업이나 외주를 선호하는 기업도 있습니다. 어떤 선택지를 택하든 SAS 플랫폼은 기업의 아키텍처와 필요사항에 맞춰 함께 진화할 수 있습니다.

SAS 플랫폼은 한 가지 아키텍처를 고집하지 않습니다. SAS는 운영 체제, 구현 방법, 관리 면에서 유연한 선택지를 제시하는 전략을 추구합니다. 고객이 클라우드에 구현하기로 결정한 경우 SAS는 업계 최고의 클라우드 구현 기술을 동원합니다. Docker 컨테이너, 조율용 Kubernetes, Cloud Foundry 지원, AWS Quick Start 및 Azure Resource Manager 같은 퍼블릭 클라우드 자동화 기술을 단적인 예로 들 수 있습니다.

또한 SAS 플랫폼에는 시스템 다운타임과 그로 인한 사용자 불만을 줄이기 위해 장애 조치 보호, 인프라 풀링, 우수한 가용성으로 무장하고 있습니다. 이와 같은 탄력성 덕분에 기업이 비즈니스 차질을 걱정할 필요 없이 분석 혁신을 촉진하는 데 집중할 수 있습니다.

SAS와 함께 하면 그리드 아키텍처, 인메모리 프로세싱, 분석 모델 병렬 처리를 통해 컴퓨팅 리소스의 활용도를 극대할 수 있습니다.

SAS 플랫폼을 도입하면 IT 부서가 다양한 분석 워크로드에 대한 늘어나는 요구에 부응할 수 있으며, 데이터 과학자와 비즈니스 사용자는 원하는 작업 방식을 선택할 수 있습니다. 또한 SAS 플랫폼을 도입하면 선택의 폭을 유지하거나 넓히기도 용이합니다.

자세히 보기

SAS는 데이터 및 분석 중심의 기업으로 거듭나는 데 필요한 지원을 아끼지 않습니다. SAS와 함께 하는 기업은 자사의 비즈니스에 적합한 선택지를 택하여 가치를 실현할 수 있습니다. 또한 SAS는 기업이 다음과 같은 문제를 해결하여 기술 투자 효과를 극대화할 방법을 제시합니다.

- 급변하는 복잡한 기술 환경/ 기술 변화 추세
- 기존 기술을 폐기할 시기/방법
- 오픈 소스 채택, 사용, 통합 방법
- 개인정보 보호 및 보안
- IT 부서가 혁신과 가치를 창출하는 전략적 지원자로 거듭나는 방법
- 비용 절감과 예산 감축 압박 증가
- 거버넌스와 민첩성 간의 균형
- 디지털 변환 프로젝트

www.sas.com/korea/platform를 방문하여 SAS 플랫폼으로 선택권과 통제력의 완벽한 균형을 유지하면서 분석 프로세스를 조율하는 방법을 확인하십시오.

더 자세한 내용은 sas.com/korea/offices에서 확인하실 수 있습니다.

SAS 및 기타 모든 SAS Institute Inc.의 제품 또는 서비스 명은 미국 및 다른 국가에서 SAS Institute Inc.의 등록상표 또는 상표입니다.
®은 미국 등록을 나타냅니다. 그 밖의 브랜드 및 제품 명칭은 각 기업의 상표입니다. Copyright © 2019, SAS Institute Inc.
All rights reserved. 0315_110368