



Informatica®
10.5.7

프로필 가이드

Informatica 프로펠 가이드

10.5.7

2024년 11월

© 저작권 Informatica LLC 2010, 2025

이 소프트웨어와 설명서는 사용 및 공개에 대한 제한 사항이 포함되어 있는 별도의 사용권 계약에 따라서만 제공됩니다. 본 문서의 어떤 부분도 Informatica LLC의 사전 통지 없이 어떠한 형태나 수단(전자적, 사진 복사, 녹음 등)으로 복제되거나 전송될 수 없습니다.

Informatica 및 Informatica 로고는 미국과 전 세계 여러 관할 국가에서 Informatica LLC의 상표 또는 등록 상표입니다. Informatica 상표의 현재 목록은 <https://www.informatica.com/trademarks.html>에서 확인할 수 있습니다. 다른 회사 및 제품명은 해당 소유자의 상표 또는 등록 상표일 수 있습니다.

수신 거부 권한에 따라 소프트웨어는 소프트웨어가 배포된 컴퓨팅 및 네트워크 환경 그리고 배포의 데이터 사용 및 시스템 통계에 대한 정보를 미국 내 Informatica에 자동으로 전송합니다. 이 전송은 Informatica 개인 정보 보호 정책에 의거하여 서비스의 일부로 간주되며 Informatica는 <https://www.informatica.com/in/privacy-policy.html>에서 제공되는 Informatica 개인 정보 보호 정책에 따라 이 정보를 사용하고 처리합니다. Administrator 도구에서 사용량 수집을 비활성화할 수 있습니다.

미국 정부 권한. 미국 정부 고객에게 제공되는 프로그램, 소프트웨어, 데이터베이스, 관련 문서 및 기술 데이터는 해당하는 연방 입수 규정 및 기관별 보안 규정에 따라 "상용 컴퓨터 소프트웨어" 또는 "상용 기술 데이터"입니다. 따라서 사용, 복제, 공개, 수정 및 조정은 해당하는 정부 계약에 규정된 제한 사항 및 라이선스 조건을 따르며, 정부 계약 조건에 의해 적용 가능한 한도 내에서, FAR 52.227-19, 상용 소프트웨어 라이선스에 규정된 추가 권한이 적용됩니다.

이 소프트웨어 및/또는 설명서의 일부에는 타사의 저작권이 적용될 수 있습니다. 필요한 타사 고지 사항은 제품에 포함되어 있습니다.

<https://www.informatica.com/legal/patents.html>에서 특허를 참조하십시오.

이 설명서의 정보는 예고 없이 변경될 수 있습니다. 이 문서에서 문제가 발견되는 경우 infa_documentation@informatica.com으로 보고해 주십시오.

Informatica 제품은 제품이 제공될 당시의 계약 조건에 따라 보증됩니다. Informatica는 상품성과 특정 목적에의 적합성에 대한 보증 그리고 비침해에 대한 보증 또는 조건을 포함하여 어떠한 종류의 명시적이거나 묵시적인 보증 없이 이 문서의 정보를 "있는 그대로" 제공합니다.

발행 날짜: 2025-02-23

목차

서문	8
Informatica 리소스	8
Informatica Network	8
Informatica 기술 자료	8
Informatica 설명서	8
Informatica Product Availability Matrix	9
Informatica Velocity	9
Informatica Marketplace	9
Informatica 글로벌 고객 지원 센터	9
파트 I: 프로파일 소개	10
장 1: 프로파일 소개	11
프로파일 소개 개요	11
프로파일링 프로세스	12
프로파일링 도구	12
프로파일 구성 요소	12
장 2: 열 프로파일 개념	14
열 프로파일 개념 개요	14
열 프로파일 옵션	15
리포지토리 프로파일 잠금 및 버전 지정 프로파일 관리	15
성과 기록표	15
장 3: 큐레이션 개념	17
큐레이션 개념 개요	17
큐레이션 태스크	17
파트 II: Informatica Analyst에서 프로파일링	19
장 4: Informatica Analyst의 열 프로파일	20
Informatica Analyst의 열 프로파일 개요	20
열 프로파일링 프로세스	21
프로파일 옵션	21
샘플링 옵션	22
드릴다운 옵션	23
런타임 환경	23
원시 환경	23
Hadoop 환경	23
Databricks 환경	24

Informatica Analyst의 운영 체제 프로필 개요.....	24
운영 체제 프로필 선택.....	25
리포지토리 자산 잠금 및 팀 기반 개발 개요.....	25
Informatica Analyst에서 열 프로필 생성.....	25
열 프로필 편집.....	26
프로필 실행.....	27
Spark 엔진에서 프로필 실행.....	28
Databricks 클러스터에서 프로필 실행.....	28
동기화 옵션.....	28
Informatica Analyst의 플랫폼 파일 데이터 개체 동기화.....	28
Informatica Analyst의 관계형 데이터 개체 동기화.....	30
장 5: Informatica Analyst의 규칙.....	31
Informatica Analyst의 규칙 개요.....	31
미리 정의된 규칙.....	31
미리 정의된 규칙 프로세스.....	32
미리 정의된 규칙 적용.....	32
식 규칙.....	32
식 규칙 작성.....	33
규칙 사양을 사용하여 식 규칙 생성.....	34
장 6: Informatica Analyst의 필터.....	36
Informatica Analyst의 필터 개요.....	36
필터 작성.....	36
단순 필터 작성.....	37
고급 필터 작성.....	38
SQL 필터 작성.....	39
필터 관리.....	39
장 7: Informatica Analyst의 열 프로필 결과.....	41
Informatica Analyst의 열 프로필 결과 개요.....	41
요약 보기.....	42
요약 보기 속성.....	43
요약 보기의 기본 필터.....	43
상세 보기.....	44
상세 보기 창.....	45
통계.....	46
데이터 미리보기.....	46
데이터 유형.....	47
이상값.....	47
패턴.....	48
값.....	49

프로필 실행의 유형.....	52
최근 프로필 실행.....	52
기록 프로필 실행.....	52
통합된 프로필 실행.....	52
프로필 실행 선택.....	53
여러 프로필 결과 비교 개요.....	54
여러 프로필 결과 비교.....	54
프로필 결과 비교의 요약 보기.....	54
프로필 결과 비교의 상세 보기.....	56
열 프로필 드릴다운.....	57
행 데이터 드릴다운.....	58
드릴다운 데이터에 필터 적용.....	58
Analyst 도구의 기간.....	58
데이터 유형 및 데이터 도메인 승인.....	58
데이터 유형 및 데이터 도메인 거부.....	59
Informatica Analyst의 열 프로필 내보내기 파일.....	59
CSV 파일의 프로필 내보내기 결과.....	59
Microsoft Excel의 프로필 내보내기 결과.....	60
Informatica Analyst에서 프로필 결과 내보내기.....	60
장 8: Informatica Analyst의 성과 기록표.....	62
Informatica Analyst의 성과 기록표 개요.....	62
Informatica Analyst 성과 기록표 프로세스.....	63
Informatica Analyst에서 성과 기록표 작성.....	64
기존 성과 기록표에 열 추가.....	65
성과 기록표 실행.....	66
성과 기록표 보기.....	66
성과 기록표 편집.....	66
메트릭.....	67
메트릭 가중치.....	67
데이터 품질의 값.....	67
임계값 정의.....	68
메트릭 그룹.....	68
메트릭 그룹 작성.....	69
메트릭 그룹으로 점수 이동.....	69
메트릭 그룹 편집.....	69
메트릭 그룹 삭제.....	70
열 드릴다운.....	70
경향 차트.....	70
점수 경향 차트.....	71
비용 경향 차트.....	71
경향 차트 보기.....	72

경향 차트 내보내기.	73
Informatica Analyst의 성과 기록표 내보내기 파일.	74
Microsoft Excel의 성과 기록표 내보내기 결과.	75
Informatica Analyst에서 성과 기록표 결과 내보내기.	75
성과 기록표 알림.	75
전자 메일 알림 메시지 템플릿.	76
성과 기록표 알림 설정.	77
성과 기록표 알림에 대한 전역 설정 구성.	77
성과 기록표 연계.	78
Informatica Analyst에서 성과 기록표 연계 보기.	78

파트 III: Informatica Developer에서 프로파일링..... 79

장 9: 데이터 개체 프로파일..... 80

Informatica Developer의 열 프로파일.	80
필터링 옵션.	80
샘플링 옵션.	81
Informatica Developer에서 단일 데이터 개체 프로파일 생성.	81
Informatica Developer에서 여러 데이터 개체 프로파일 생성.	82
Informatica Developer의 플랫폼 파일 데이터 개체 동기화.	83
Informatica Developer의 관계형 데이터 개체 동기화.	85

장 10: 반 구조화된 데이터 소스에 대한 열 프로파일..... 86

반 구조화된 데이터 소스에 대한 열 프로파일 개요.	86
JSON 및 XML 데이터 개체.	87
JSON 또는 XML 데이터 소스로부터 데이터 개체 생성.	87
HDFS에 있는 반 구조화된 데이터 소스에 대한 복잡한 파일 데이터 개체.	88
HDFS에 있는 JSON 또는 XML 데이터 소스의 복잡한 파일 데이터 개체.	88
HDFS에 있는 Avro 또는 Parquet 데이터 소스의 복잡한 파일 데이터 개체.	88
HDFS 연결 작성.	88
HDFS에 있는 JSON 또는 XML 파일로부터 복잡한 파일 데이터 개체 생성.	89
Avro 또는 Parquet 데이터 소스로부터 복잡한 파일 데이터 개체 생성.	90
반 구조화된 데이터 소스에 대한 열 프로파일 생성.	91

장 11: Informatica Developer의 규칙..... 93

Informatica Developer의 규칙 개요.	93
Informatica Developer에서 규칙 작성.	94
Informatica Developer에서 규칙 적용.	94

장 12: 맵렛 및 매핑 프로파일링..... 95

맵렛 및 매핑 프로파일링 개요.	95
맵렛 또는 매핑 개체에 대해 프로파일 실행.	95
매핑 또는 맵렛 개체용 프로파일 비교.	96

프로필에서 매핑 생성.	96
-------------------	----

장 13: Informatica Developer의 열 프로필 결과. 97

Informatica Developer의 열 프로필 결과.	97
---------------------------------------	----

열 값 속성.	98
--------------	----

열 패턴 속성.	98
---------------	----

열 통계 속성.	98
---------------	----

열 데이터 유형 속성.	99
-------------------	----

Informatica Developer의 큐레이션.	100
-----------------------------------	-----

데이터 유형 승인.	100
-----------------	-----

데이터 유형 거부.	100
-----------------	-----

Informatica Developer에서 프로필 결과 내보내기.	101
---	-----

장 14: Informatica Developer의 성과 기록표. 102

Informatica Developer의 성과 기록표 개요.	102
--	-----

성과 기록표 작성.	102
-----------------	-----

성과 기록표 연계의 리소스 파일 내보내기.	103
------------------------------	-----

Informatica Developer에서 성과 기록표 연계 보기.	103
--	-----

인덱스. 104

서문

Informatica 프로파일 가이드에서 프로파일 사용하여 데이터 소스의 콘텐츠 및 구조를 분석하는 방법을 알아보십시오. 값 빈도, 비율 및 패턴과 같은 열의 소스 데이터 특성을 확인할 수 있습니다. 이 가이드는 데이터 분석가 및 개발자를 대상으로 작성되었습니다.

Informatica 리소스

Informatica는 Informatica Network 및 기타 온라인 포털을 통해 다양한 범위의 제품 리소스를 제공합니다. 리소스를 통해 Informatica 제품 및 솔루션을 최대한 활용하고 다른 Informatica 사용자 및 주제별 전문가로부터 배울 수 있습니다.

Informatica Network

Informatica Network는 Informatica 기술 자료, Informatica 글로벌 고객 지원 센터 등 여러 리소스로 연결되는 관문입니다. Informatica Network를 시작하려면 <https://network.informatica.com>을 방문하십시오.

Informatica Network 멤버인 경우 다음 옵션이 가능합니다.

- 기술 자료에서 제품 리소스를 검색할 수 있습니다.
- 제품 사용 가능 여부에 대한 정보를 봅니다.
- 지원 사례를 생성하고 검토할 수 있습니다.
- 거주 지역의 Informatica 사용자 그룹 네트워크를 검색하고 동료와 협업 관계 유지

Informatica 기술 자료

Informatica 기술 자료를 사용하여 사용 방법 문서, 모범 사례, 비디오 자습서, 자주 묻는 질문에 대한 답변 등 제품 리소스를 확인할 수 있습니다.

기술 자료를 검색하려면 <https://search.informatica.com>을 방문하십시오. 기술 자료에 대한 질문, 의견 또는 아이디어가 있는 경우 KB_Feedback@informatica.com을 통해 Informatica 기술 자료 팀에 문의해 주시기 바랍니다.

Informatica 설명서

Informatica 설명서 포털에서 확장된 설명서 라이브러리를 탐색하여 현재 및 최근 제품 릴리스를 확인할 수 있습니다. 설명서 포털을 탐색하려면 <https://docs.informatica.com>을 방문하십시오.

제품 설명서에 대한 질문, 의견 또는 아이디어가 있는 경우 infa_documentation@informatica.com에서 Informatica 설명서 팀에 문의해 주시기 바랍니다.

Informatica Product Availability Matrix

PAM(Product Availability Matrix)은 제품 릴리스에서 지원하는 운영 체제 버전, 데이터베이스 및 데이터 소스 유형과 대상을 나타냅니다.

<https://network.informatica.com/community/informatica-network/product-availability-matrices>에서 Informatica PAM을 찾을 수 있습니다.

Informatica Velocity

Informatica Velocity는 수백 가지 데이터 관리 프로젝트의 실제 경험을 토대로 Informatica 전문 서비스업에서 개발한 팁과 모범 사례 모음입니다. Informatica Velocity는 전 세계의 조직과 협력하여 성공적인 데이터 관리 솔루션을 계획, 개발, 배포 및 유지 관리하는 Informatica 컨설턴트의 포괄적인 지식을 보여줍니다.

Informatica Velocity 리소스는 <http://velocity.informatica.com>에서 확인할 수 있습니다. Informatica Velocity에 대한 질문, 주석 또는 아이디어가 있으시면 Informatica 전문 서비스업(ips@informatica.com)에 문의하십시오.

Informatica Marketplace

Informatica Marketplace는 Informatica 구현을 확대 및 개선하기 위한 솔루션을 찾을 수 있는 포럼입니다. Marketplace에서 Informatica 개발자와 파트너가 제공하는 수백 개의 솔루션을 활용하여 생산성을 향상시키고 프로젝트의 구현에 걸리는 시간을 줄일 수 있습니다. <https://marketplace.informatica.com>에서 Informatica Marketplace를 찾을 수 있습니다.

Informatica 글로벌 고객 지원 센터

전화 또는 Informatica 네트워크를 통해 글로벌 지원 센터에 문의할 수 있습니다.

해당 지역의 Informatica 글로벌 고객 지원 전화 번호는 Informatica 웹 사이트 (<https://www.informatica.com/services-and-training/customer-success-services/contact-us.html>)를 방문하여 찾을 수 있습니다.

Informatica 네트워크에 대한 온라인 지원 리소스를 찾으려면 <https://network.informatica.com>으로 이동하고 eSupport 옵션을 선택하십시오.

파트 I: 프로필 소개

이 파트에 포함된 장:

- [프로필 소개, 11](#)
- [열 프로필 개념, 14](#)
- [큐레이션 개념, 17](#)

제 1 장

프로필 소개

이 장에 포함된 항목:

- [프로필 소개 개요, 11](#)
- [프로파일링 프로세스, 12](#)
- [프로파일링 도구, 12](#)
- [프로필 구성 요소, 12](#)

프로필 소개 개요

프로필을 작성하고 실행하여 응용 프로그램, 스키마 또는 엔터프라이즈의 데이터 소스 콘텐츠/품질/구조를 찾습니다. 데이터 소스 콘텐츠에는 값 빈도 및 데이터 유형이 포함됩니다. 데이터 소스 구조에는 키 및 함수 종속성이 포함됩니다.

검색 프로세스를 수행할 때 프로필을 작성하고 실행할 수 있습니다. 프로필은 엔터프라이즈 데이터 소스 전체의 모든 데이터 불규칙 및 데이터 프로젝트에 위험을 야기하는 숨겨진 데이터 문제를 찾고 분석하는 리포지토리 개체입니다. 엔터프라이즈의 모든 데이터 소스에 프로필을 실행하여 데이터 및 메타데이터의 강점과 약점을 제대로 파악할 수 있습니다.

Informatica Analyst 및 **Informatica Developer**를 사용하여 소스 데이터와 메타데이터를 분석할 수 있습니다. 분석가 및 개발자는 이러한 도구를 사용하여 공동 작업을 수행하고 데이터 품질 문제를 식별하고 데이터 관계를 분석할 수 있습니다. 사용자의 작업 역할에 따라 **Analyst** 도구 또는 **Developer tool**의 기능을 사용할 수 있습니다. 수행할 수 있는 프로파일링의 수준은 사용하는 도구에 따라 다릅니다.

Developer tool 및 **Analyst** 도구에서 다음 태스크를 수행할 수 있습니다.

- 열 프로파일링을 수행합니다. 이 프로세스에는 열의 고유한 값 수, null 값 수 및 데이터 패턴 수에 대한 검색이 포함됩니다.
- 열 프로필에 규칙을 추가합니다.
- 프로필 결과에서 유추된 데이터 유형을 큐레이션합니다.
- 성과 기록표를 사용하여 데이터 품질을 모니터링합니다.
- 프로필에서 매핑을 생성합니다.

프로파일링 프로세스

데이터 통합 프로젝트를 시작할 때의 첫 번째 단계는 주로 프로파일링입니다. 프로필을 생성하면 데이터 소스의 콘텐츠, 품질 및 구조를 분석할 수 있습니다. 프로파일링 프로세스에는 데이터 소스의 메타데이터를 검색하는 작업이 포함됩니다.

각 데이터 분석 유형에 대해 열 프로필 등의 서로 다른 프로필을 사용합니다. 데이터 품질 문제를 확인하고 문서화할 수도 있습니다. 프로파일링을 수행하려면 다음 태스크를 완료합니다.

1. 데이터 소스에서 데이터 콘텐츠를 찾고 분석합니다. 여기에는 데이터 유형, 값 빈도, 패턴 빈도 및 데이터 통계(예: 최소값 및 최대값)가 포함됩니다.
2. 프로파일링 결과를 검토합니다.
3. 참조 데이터를 생성합니다.
4. 프로필 결과를 드릴다운합니다.
5. 데이터 문제를 문서화합니다.
6. 규칙을 생성하고 실행합니다.
7. 데이터 품질을 모니터링할 성과 기록표를 생성합니다.

프로파일링 도구

여러 Informatica 도구를 사용하여 프로파일링 프로세스를 관리합니다.

다음 도구를 사용하여 프로파일링 프로세스를 관리할 수 있습니다.

Informatica Administrator

사용자, 그룹, 권한 및 역할을 관리합니다. 분석 서비스를 관리하고 Informatica Analyst의 프로젝트 및 개체에 대한 사용 권한을 관리할 수 있습니다. 이 도구를 사용하여 Informatica Developer의 액세스 사용 권한을 제어할 수 있습니다.

Informatica Developer

이 도구에서 프로필을 작성하고 실행하여 데이터 소스 하나 이상의 메타데이터를 찾고 분석합니다. 프로필은 마법사를 사용하여 생성합니다.

Informatica Analyst

Analyst 도구에서 열 프로필을 실행할 수 있습니다. 프로필을 실행한 후에는 데이터 소스의 데이터 행을 드릴다운할 수 있습니다.

프로필 구성 요소

프로필에는 데이터 소스의 콘텐츠와 구조를 효과적으로 분석하는 데 사용할 수 있는 여러 구성 요소가 포함됩니다.

프로필에 포함되는 구성 요소는 다음과 같습니다.

필터

원래 데이터 소스에서 특정 조건을 충족하는 하위 집합을 생성합니다. 그런 다음 샘플 데이터에 대해 프로필을 실행할 수 있습니다.

규칙

프로필을 실행할 때 데이터에 적용되는 조건을 정의하는 비즈니스 논리입니다. 프로필에 규칙을 추가하여 데이터 유효성을 검사합니다.

태그

비즈니스 사용법에 따라 모델 리포지토리의 개체를 정의하는 메타데이터입니다. 비즈니스 용도에 따라 개체를 그룹화하기 위해 태그를 작성합니다.

설명

프로필에 대한 설명입니다. 설명을 사용하면 다른 **Analyst** 및 **Developer** 도구 사용자와 프로필에 대한 정보를 공유할 수 있습니다.

성과 기록표

프로필 결과에서 열의 유효한 값 또는 규칙의 출력을 그래픽으로 표시한 것입니다. 성과 기록표는 데이터 품질의 진행 상황을 측정하는 데 사용됩니다.

제 2 장

열 프로필 개념

이 장에 포함된 항목:

- [열 프로필 개념 개요, 14](#)
- [열 프로필 옵션, 15](#)
- [리포지토리 프로필 잠금 및 버전 지정 프로필 관리, 15](#)
- [성과 기록표, 15](#)

열 프로필 개념 개요

열 프로필에 따라 값 빈도, 백분율, 패턴 등 데이터 소스에 있는 열의 특성이 결정됩니다.

열 프로파일링 시에는 데이터에 대한 다음과 같은 정보가 검색됩니다.

- 숫자 및 백분율로 표시되는 각 열의 **Null** 값, 고유한 값 및 고유하지 않은 값.
- 각 열의 데이터 패턴 및 이러한 값이 표시되는 빈도
- 각 열에 있는 값의 최대 길이 및 최소 길이와 각 열의 첫 번째 값과 마지막 값 등 열 값에 대한 통계를 제공합니다.
- 문서화된 데이터 유형, 유추된 데이터 유형 및 문서화된 데이터 유형과 유추된 데이터 유형 사이의 가능한 충돌.
- 패턴 및 값 빈도 이상값.

프로필을 작성하거나 편집할 경우 다음의 옵션을 구성할 수 있습니다.

- 열 프로필 옵션. 프로필을 실행할 열을 선택하고, 샘플링 옵션 및 드릴다운 옵션을 선택할 수 있습니다.
- 필터와 규칙 추가, 편집 또는 삭제.

프로필 결과에서는 프로필 및 프로필의 열에 설명과 태그를 추가할 수 있습니다. 열에 비즈니스 용어를 할당할 수 있습니다.

모델 리포지토리는 리포지토리 프로필 잠금을 사용하여 사용자가 작업을 덮어쓰지 못하도록 프로필을 잠급니다. 버전 제어 시스템은 여러 버전의 프로필을 저장하고 각 버전에 버전 번호를 할당합니다. 프로필을 체크 아웃한 다음 내용을 변경하고 다시 프로파일을 체크 인할 수 있습니다. 프로필을 다시 체크 인하기 전에 프로파일 체크 아웃 작업을 실행 취소할 수 있습니다.

성과 기록표를 작성하여 데이터 품질을 정기적으로 검토합니다. 프로필에 규칙을 적용하기 전과 적용한 후에 성과 기록표를 작성하여 열에 대해 유효한 값의 그래픽 표현을 확인할 수 있습니다.

스케줄러 서비스를 사용하면 특정 시간에 실행되거나 특정 간격으로 실행되도록 프로필 실행 및 성과 기록표 실행을 예약할 수 있습니다. 스케줄러 서비스는 프로필, 성과 기록표, 배포된 매핑 및 배포된 워크플로우의 일정을 관리합니다. Informatica Administrator에서 일정을 작성, 관리 및 실행할 수 있습니다.

열 프로필 옵션

프로필을 생성할 때 프로필 마법사를 사용하여 필터, 규칙, 드릴다운 옵션, 샘플링 옵션 및 연결을 정의할 수 있습니다. 이러한 옵션은 프로필이 소스 데이터에서 행을 읽는 방식을 결정합니다.

열 프로필, 데이터 도메인 검색 프로필 또는 엔터프라이즈 검색 프로필에 다음과 같은 옵션을 정의할 수 있습니다.

- 필터. 필터를 생성하여 프로필에 적용할 수 있습니다.
- 규칙. 프로필을 생성할 때 규칙을 추가할 수 있습니다. **Analyst** 도구나 **Developer tool**에서 생성하는 규칙을 재사용할 수 있습니다.
- 드릴다운 옵션. 데이터 소스의 현재 데이터를 읽거나, 프로파일링 웨어하우스의 준비된 프로필 데이터를 읽을지 선택할 수 있습니다.
- 샘플링 옵션. 샘플링 옵션 중 하나를 선택하여 프로필을 실행할 대상 행 수를 결정할 수 있습니다.
- 연결. 원시 또는 **Hadoop** 런타임 환경에서 프로필을 실행할 수 있습니다. **Hadoop** 런타임 환경에서 **Blaze** 또는 **Spark** 엔진을 선택할 수 있습니다.

리포지토리 프로필 잠금 및 버전 지정 프로필 관리

모델 리포지토리는 사용자가 작업을 덮어쓰지 못하도록 프로필을 잠급니다. 프로필 편집을 시작하면 다른 사용자가 변경 내용을 프로필에 저장하지 못하도록 프로필이 잠깁니다. 프로필을 저장하면 잠금이 해제됩니다. 버전 지정 프로필 관리를 통해 프로필의 버전이 작성되고 버전 기록을 볼 수 있습니다.

Developer tool 또는 **Analyst** 도구에서 프로필을 편집하는 경우 모델 리포지토리가 프로필을 잠급니다. 도구가 예기치 않게 중지되면 잠금이 유지됩니다. 모델 리포지토리에 다시 연결하면 잠긴 프로필을 볼 수 있습니다. 프로필 편집을 계속하거나 프로필의 잠금을 해제할 수 있습니다.

모델 리포지토리가 버전 제어 시스템과 통합된 경우 프로필의 버전을 관리할 수 있습니다. 예를 들어 프로필을 체크 아웃 및 체크 인하고, 체크 아웃을 실행 취소하고, 프로필의 특정 기록 버전을 보고, 체크 아웃한 프로필을 볼 수 있습니다. **Analyst** 도구의 리포지토리 자산 잠금 및 버전 지정 자산 관리에 대한 자세한 내용은 **Analyst** 도구 가이드를 참조하십시오. **Developer tool**의 리포지토리 개체 잠금 및 버전 지정 개체 관리에 대한 자세한 내용은 **Developer tool** 가이드를 참조하십시오.

성과 기록표

성과 기록표는 프로필 결과에서 열의 유효한 값 또는 규칙의 출력을 그래픽으로 표현한 것입니다. 성과 기록표는 데이터 품질의 진행 상황을 측정하는 데 사용됩니다. 프로필에서 성과 기록표를 작성하고 시간의 경과에 따른 데이터 품질 변화를 모니터링할 수 있습니다.

성과 기록표는 메트릭, 메트릭 그룹, 임계값 등의 여러 구성 요소를 포함합니다. 프로필을 실행한 후에는 소스 열을 성과 기록표에 메트릭으로 추가하고 메트릭에 대해 유효한 값을 구성할 수 있습니다. 조직은 성과 기록표를 통해 메트릭 및 성과 기록표 수준에서 잘못된 데이터의 비용을 추적하여 데이터 품질의 값을 측정할 수 있습니다. 각 메트릭에 대한 잘못된 데이터 비용을 측정하려면 비용 단위를 메트릭에 할당할 다음 고정 비용이나 가변 비용을 설정합니다. 성과 기록표를 실행할 때 성과 기록표 결과에는 각 메트릭에 대한 잘못된 데이터 비용과 모든 메트릭에 대한 총 비용 값이 포함됩니다.

메트릭 그룹을 사용하여 성과 기록표의 관련 메트릭을 집합으로 분류합니다. 임계값은 레코드의 열에 대해 허용되는 잘못된 데이터의 백분율 범위를 식별합니다. 데이터의 양호, 합격 또는 불합격 범위에 대한 임계값을 설정할 수 있습니다.

성과 기록표를 실행할 때 점수 메트릭을 드릴다운할 대상(라이브 데이터 또는 준비된 데이터)을 구성합니다. 성과 기록표를 실행하고 점수를 확인한 후 각 메트릭을 드릴다운하여 유효한 데이터 레코드와 유효하지 않은 레코드를 식별합니다. 성과 기록표의 각 메트릭 또는 메트릭 그룹에 대해 성과 기록표 연계를 확인할 수도 있습니다. 데이터 품질을 효율적으로 추적하려는 경우 점수 경향 차트와 비용 경향 차트를 사용할 수 있습니다. 이러한 차트에서는 일정 기간 동안 잘못된 데이터의 점수와 비용이 변경되는 방식을 모니터링합니다.

프로파일링 웨어하우스에는 성과 기록표 통계 및 구성 정보가 저장됩니다. 성과 기록표 결과를 가져오고 보고서를 실행하도록 타사 응용 프로그램을 구성할 수 있습니다. 웹 응용 프로그램, 포털 또는 보고서(예: 비즈니스 인텔리전스 보고서)에서 성과 기록표 결과를 표시할 수도 있습니다.

제 3 장

큐레이션 개념

이 장에 포함된 항목:

- [큐레이션 개념 개요, 17](#)
- [큐레이션 태스크, 17](#)

큐레이션 개념 개요

기간은 데이터 소스에서 검색된 메타데이터를 사용하고 보고할 수 있도록 해당 메타데이터의 유효성을 검사하고 관리하는 프로세스입니다.

다음과 같은 유추된 프로필 결과에 대해 기간을 지정할 수 있습니다.

- 데이터 유형
- 데이터 도메인
- 기본 키
- 외래 키

유추된 프로필 결과에 대한 기간을 지정하여 데이터베이스와 스키마의 열, 데이터 도메인 및 데이터 개체 관계에 대한 메타데이터를 정확하게 만들 수 있습니다. 그런 다음 검색을 사용하여 여러 리포지토리에서 정보를 검색할 때 가장 관련성이 높은 메타데이터를 찾을 수 있습니다. 엔터프라이즈 검색 결과에서 외래 키 관계 다이어그램을 볼 때 가장 관련성이 높은 메타데이터를 찾을 수도 있습니다.

프로필 실행의 일부분으로 프로필이 생성하는 특정 메타데이터 유추의 기간을 지정할 수 있습니다. 예를 들어 열 프로필 결과 및 데이터 도메인 검색 결과에서 유추된 데이터 유형을 승인하거나 거부할 수 있습니다. 엔터프라이즈 검색 결과에서 유추된 기본 키 및 외래 키를 승인하거나 거부할 수도 있습니다.

큐레이션 태스크

프로필 실행 후에 프로필 결과의 기간을 지정할 수 있습니다. 이전에 프로필을 실행할 때 내린 기간 결과를 반전할 수도 있습니다.

Analyst 도구에서 다음 기간 태스크를 수행할 수 있습니다.

- 여러 열 및 데이터 도메인에 대해 유추된 데이터 유형을 승인하거나 거부합니다.
- 승인되거나 거부된 데이터 유형을 유추된 상태로 복원합니다.

- 승인되거나 거부된 데이터 도메인을 유추된 상태로 복원합니다.
- 거부된 결과 행을 보거나 숨깁니다.
- 승인된 데이터 유형 및 데이터 도메인과 같은 특정 메타데이터 기본 설정을 기준으로 프로필 실행에서 열을 제외합니다.

Developer 도구에서 다음 기간 태스크를 수행할 수 있습니다.

- 여러 열에 대해 유추된 데이터 유형을 승인하거나 거부합니다.
- 승인되거나 거부된 데이터 유형을 유추된 상태로 복원합니다.
- 승인되거나 거부된 데이터 도메인을 유추된 상태로 복원합니다.
- 거부된 결과 행을 보거나 숨깁니다.
- 기본 키 검색 결과에서 데이터 개체를 승인하거나 거부합니다.
- 외래 키 검색 결과를 포함한 엔터프라이즈 검색 결과를 승인하거나 거부합니다.
- 승인된 데이터 유형 및 데이터 도메인과 같은 특정 메타데이터 기본 설정을 기준으로 프로필 실행에서 열을 제외합니다.

파트 II: Informatica Analyst에서 프로파일링

이 파트에 포함된 장:

- [Informatica Analyst의 열 프로필, 20](#)
- [Informatica Analyst의 규칙, 31](#)
- [Informatica Analyst의 필터, 36](#)
- [Informatica Analyst의 열 프로필 결과, 41](#)
- [Informatica Analyst의 성과 기록표, 62](#)

제 4 장

Informatica Analyst의 열 프로파일

이 장에 포함된 항목:

- [Informatica Analyst의 열 프로파일 개요, 20](#)
- [열 프로파일링 프로세스, 21](#)
- [프로파일 옵션, 21](#)
- [런타임 환경, 23](#)
- [Informatica Analyst의 운영 체제 프로파일 개요, 24](#)
- [리포지토리 자산 자금 및 팀 기반 개발 개요, 25](#)
- [Informatica Analyst에서 열 프로파일 생성, 25](#)
- [열 프로파일 편집, 26](#)
- [프로파일 실행, 27](#)
- [Spark 엔진에서 프로파일 실행, 28](#)
- [Databricks 클러스터에서 프로파일 실행, 28](#)
- [동기화 옵션, 28](#)

Informatica Analyst의 열 프로파일 개요

프로파일을 작성할 때는 프로파일을 실행할 데이터 개체의 열을 선택합니다. 빠른 프로파일링을 위한 샘플링 및 드릴다운 옵션을 구성할 수 있습니다. 런타임 환경을 선택할 수 있습니다. 프로파일을 작성할 때 프로파일에 규칙 및 필터를 추가할 수 있습니다. 프로파일을 실행한 후에 데이터를 파악할 수 있도록 프로파일링 통계를 검사할 수 있습니다.

열 수가 최대 1000개인 플랫폼 파일과 광범위한 테이블을 프로파일링할 수 있습니다. 프로파일을 작성하거나 실행할 때는 모든 열을 선택하거나 프로파일에 대한 각 열을 선택할 수 있습니다. 드릴다운할 모든 열을 선택하고 해당 열의 값 빈도를 볼 수 있습니다. 열 이름이 245자를 초과하는 경우 구분자로 분리된 파일에서 프로파일을 실행할 열을 선택할 수 없습니다.

Spark 엔진의 반 구조화된 데이터 소스에서 프로파일을 실행할 수 없습니다.

Informatica Analyst에서 다음과 같은 방법을 사용하여 열 프로파일을 작성할 수 있습니다.

- **라이브러리** 작업 공간에서 데이터 개체를 마우스 오른쪽 단추로 클릭하여 프로파일을 작성합니다.
- 기본 옵션을 사용하여 기본 열 프로파일을 작성합니다.
- 프로파일에 대한 설정을 사용자 지정하여 사용자 지정 프로파일을 작성합니다.

참고: Avro, JSON, Parquet 및 XML 데이터 소스에 대해 프로필을 보고 실행할 수 있습니다. Informatica Developer에서 Avro, JSON, Parquet 및 XML 데이터 소스에 대해 열 프로필을 작성하고 편집할 수 있습니다.

열 프로파일링 프로세스

열 프로파일링 프로세스의 일부분으로 프로파일링을 위해 모든 소스 열을 선택하거나 특정 열을 선택할 수 있습니다. 기본 프로필 옵션을 그대로 사용하거나, 샘플링 옵션, 드릴다운 옵션 및 런타임 환경을 구성할 수 있습니다.

다음 단계에서는 열 프로파일링 프로세스에 대해 설명합니다.

1. 열 프로필의 이름, 설명 및 위치를 선택합니다.
2. 프로필이 실행될 외부 소스 또는 가져온 데이터 개체를 선택합니다.
3. 필요에 따라 소스 데이터를 미리 봅니다.
4. 프로필이 실행될 열을 선택합니다.
5. 기본 옵션으로 프로필을 작성할지 아니면 기본 옵션을 변경할지를 결정합니다. 샘플링 옵션, 드릴다운 옵션 및 런타임 환경 옵션을 구성할 수 있습니다.
6. 필요에 따라 프로필을 작성할 때 규칙 및 필터를 추가합니다.
7. 프로필을 실행합니다.

참고: 다국어/유니코드 데이터를 프로파일링하는 경우와 열 이름의 경우에는 다음 규칙과 지침을 고려하십시오.

- 다른 소스에서 다국어 데이터를 프로파일링한 다음 브라우저의 로캘 설정을 기준으로 프로필 결과를 볼 수 있습니다. Analyst 도구는 브라우저 로캘에 따라 날짜/시간, 숫자 및 10진수 데이터 유형을 변경합니다.
- 다국어 데이터 정렬. 다국어 데이터를 정렬할 수 있습니다. Analyst 도구는 브라우저 로캘에 따라 정렬 순서를 표시합니다.
- DB2 데이터베이스의 유니코드 데이터를 프로파일링하려면 데이터베이스에서 DB2CODEPAGE 데이터베이스 환경 변수를 설정한 다음 데이터 통합 서비스를 다시 시작합니다.

프로필 옵션

프로필 옵션에는 데이터 샘플링 옵션과 데이터 드릴다운 옵션이 포함됩니다. 데이터 개체에 대해 열 프로필을 작성하거나 편집할 때 이러한 옵션을 구성할 수 있습니다.

검색 작업 공간을 사용하여 프로필 옵션을 구성합니다. 열, 샘플링 및 드릴다운 옵션에 대한 기본 옵션을 사용하여 프로필을 작성할 수 있습니다. 드릴다운 옵션을 사용하여 라이브 데이터와 준비된 데이터 중에서 선택합니다.

샘플링 옵션

샘플링 옵션에 따라 **Analyst** 도구가 프로필을 실행하도록 선택하는 행의 수가 결정됩니다. 프로필을 정의하거나 실행할 때 샘플링 옵션을 구성할 수 있습니다.

다음 테이블에서는 프로필의 샘플링 옵션에 대해 설명합니다.

옵션	설명
모든 행	데이터 개체의 모든 행에서 프로필을 실행합니다. 원시, Blaze, Spark 및 Databricks 런타임 환경에서 지원됩니다.
처음 <수>개 행 샘플링	데이터 개체의 행 처음부터 샘플 행에서 프로필을 실행합니다. 최대 2,147,483,647개의 행을 선택할 수 있습니다. 원시 및 Blaze 런타임 환경에서 지원됩니다.
<수>개 행 무작위 샘플링	데이터 개체의 무작위로 선택한 행 수에서 프로필을 실행합니다. 최대 2,147,483,647개의 행을 선택할 수 있습니다. 원시 및 Blaze 런타임 환경에서 지원됩니다.
무작위 샘플링(자동)	데이터 개체의 행 수를 기초로 계산된 샘플 행에서 프로필을 실행합니다. 원시 및 Blaze 런타임 환경에서 지원됩니다.
n <수>개 행 제한	데이터 개체의 행 수를 기반으로 프로필을 실행합니다. Hadoop 유효성 검사 환경에서 프로필을 실행하도록 선택하면 Spark 엔진이 데이터 개체의 여러 파티션에서 샘플을 수집하고 샘플 크기를 계산할 단일 노드에 샘플을 푸시합니다. n 제한 샘플링 옵션은 Oracle, SQL Server 및 DB2 데이터베이스를 지원합니다. n 제한 샘플링 옵션과 함께 고급 필터를 적용할 수 없습니다. Spark 런타임 환경에서 지원됩니다.
무작위 비율	데이터 개체의 행 비율에서 프로필을 실행합니다. Spark 및 Databricks 런타임 환경에서 지원됩니다.
후속 프로필 실행 시 데이터 유형 및 데이터 도메인 유추에서 승인된 데이터 유형 및 데이터 도메인 제외	다음 프로필 실행부터 데이터 유형 및 데이터 도메인 유추에서 승인된 데이터 유형 또는 데이터 도메인을 제외합니다.

무작위 샘플링 행에서 프로필을 실행하도록 선택하면 무작위 샘플링 알고리즘이 데이터 개체에서 행을 무작위로 선택하여 프로필을 실행합니다. 열 프로필에 대한 무작위 샘플링 옵션을 선택하면 **Analyst** 도구가 준비된 데이터에 대해 드릴다운을 수행합니다. 이는 드릴다운 성능에 영향을 줄 수 있습니다. 데이터 도메인 검색 프로필에 대해 무작위 샘플링 옵션을 선택하면 **Analyst** 도구가 라이브 데이터에 대해 드릴다운을 수행합니다.

드릴다운 옵션

프로필을 정의하거나 편집할 때 드릴다운 옵션을 구성할 수 있습니다.

다음 테이블에서는 프로필의 드릴다운 옵션에 대해 설명합니다.

옵션	설명
라이브	라이브 데이터를 드릴다운하여 데이터 소스에서 현재 데이터를 읽습니다.
준비됨	준비된 데이터를 드릴다운하여 프로파일링 웨어하우스에 준비된 프로필 데이터를 읽습니다.
열 선택	프로파일링을 위해 선택하지 않은 열을 드릴다운용으로 식별합니다.

런타임 환경

열 프로필의 런타임 환경으로 원시, Hadoop 또는 Databricks를 선택할 수 있습니다. Hadoop 런타임 환경에서 Blaze 또는 Spark 엔진을 선택할 수 있습니다. Databricks 런타임 환경에서 Databricks Spark 옵션을 선택할 수 있습니다. Informatica Analyst에서는 사용자가 런타임 환경을 선택한 이후에 프로필 정의에 런타임 환경을 설정합니다.

원시 환경

원시 런타임 환경에서 프로필을 실행하면 Analyst 도구가 프로필 작업을 프로파일링 서비스 모듈에 제출합니다. 그러면 프로파일링 서비스 모듈이 프로필 작업을 매핑 집합으로 나눕니다. 데이터 통합 서비스는 데이터 통합 서비스가 실행되는 동일한 시스템에서 이러한 매핑을 실행하고 프로필 결과를 프로파일링 웨어하우스에 씁니다. 기본적으로 모든 프로필은 원시 런타임 환경에서 실행됩니다.

원시 소스를 사용하여 원시 환경에서 프로필을 생성 및 실행할 수 있습니다. 원시 데이터 소스는 플랫폼 파일, 관계형 소스 또는 메인프레임 소스와 같은 비Hadoop 소스입니다. 원시 환경에서 Hive 또는 HDFS 데이터 소스와 함께 논리적 데이터 소스 또는 매핑 사양에 대해 프로필을 실행할 수도 있습니다.

Hadoop 환경

Blaze 엔진 또는 Spark 엔진을 선택하여 Hadoop 런타임 환경에서 프로필을 실행할 수 있습니다.

Blaze 또는 Spark를 선택한 후 Hadoop 연결을 선택할 수 있습니다. 데이터 통합 서비스가 프로필을 실행하기 위해 프로필 논리를 Hadoop 클러스터에 있는 Blaze 또는 Spark 엔진에 푸시합니다.

Hadoop 환경에서 프로필을 실행하면 Analyst 도구가 프로필 작업을 프로파일링 서비스 모듈에 제출합니다. 그러면 프로파일링 서비스 모듈이 프로필 작업을 매핑 집합으로 나눕니다. 데이터 통합 서비스가 Hadoop 연결을 통해 매핑을 Hadoop 환경에 푸시합니다. Blaze 엔진 또는 Spark 엔진이 매핑을 처리하고 데이터 통합 서비스가 프로필 결과를 프로파일링 웨어하우스에 씁니다.

Sqoop 데이터 소스에 대한 열 프로필

Sqoop을 사용하는 데이터 개체에 대해 열 프로필을 실행할 수 있습니다. Hadoop을 유효성 검사 환경으로 선택한 후 열 프로필을 실행하기 위한 Hadoop 연결에서 Blaze 엔진 또는 Spark 엔진을 선택할 수 있습니다.

논리적 데이터 개체 또는 사용자 지정된 데이터 개체에 대해 열 프로필을 실행하는 경우 병렬 처리를 실현하고 성능을 최적화하도록 num-mappers 인수를 구성할 수 있습니다. 또한 split-by 인수를 구성하여 Sqoop이 작업 단위를 분할하는 기준이 되는 열을 지정해야 합니다.

다음 구문을 사용합니다.

```
--split-by <column_name>
```

기본 키의 값이 최소값 ~ 최대값 범위 내에 고르게 분포되지 않은 경우, split-by 인수를 구성해 균형 조정된 데이터 분포가 있는 다른 열을 지정하여 작업 단위를 분할할 수 있습니다.

split-by 열을 정의하지 않는 경우 Sqoop은 다음과 같은 조건을 기반으로 작업 단위를 분할합니다.

- 데이터 개체에 단일 기본 키가 포함되어 있는 경우 Sqoop은 기본 키를 split-by 열로 사용합니다.
- 데이터 개체에 복합 기본 키가 포함되어 있는 경우 Sqoop은 기본적으로 split-by 인수 없이 복합 기본 키를 처리하는 동작으로 설정됩니다. 자세한 내용은 Sqoop 설명서를 참조하십시오.
- 데이터 개체에 동일한 열이 있는 2개의 테이블이 포함되어 있는 경우 테이블 인가된 이름으로 split-by 열을 정의해야 합니다. 예를 들어 테이블 이름이 CUSTOMER이고 열 이름이 FULL_NAME인 경우 split-by 열을 다음과 같이 정의합니다.
--split-by CUSTOMER.FULL_NAME
- 데이터 개체에 기본 키가 포함되어 있지 않은 경우 m 인수 및 num-mappers 인수의 값이 기본적으로 1로 설정됩니다.

Cloudera Connector Powered by Teradata 또는 Hortonworks Connector for Teradata를 사용하고 Teradata 테이블에 기본 키가 포함되어 있지 않은 경우, split-by 인수가 필요합니다.

Databricks 환경

Databricks Spark 옵션을 선택하여 Databricks 런타임 환경에서 프로필을 실행할 수 있습니다.

Databricks Spark 옵션을 선택한 후 Databricks 연결을 선택할 수 있습니다. 데이터 통합 서비스가 프로필을 실행하기 위해 프로필 논리를 Databricks 클러스터에 있는 Spark 엔진에 푸시합니다.

Databricks 환경에서 프로필을 실행하면 Analyst 도구가 프로필 작업을 프로파일링 서비스 모듈에 제출합니다. 그러면 프로파일링 서비스 모듈이 프로필 작업을 매핑 집합으로 나눕니다. 데이터 통합 서비스가 Hadoop 연결을 통해 매핑을 Spark 엔진에 푸시합니다. Spark 엔진이 매핑을 처리하고 데이터 통합 서비스가 프로필 결과를 프로파일링 웨어하우스에 씁니다.

Informatica Analyst의 운영 체제 프로필 개요

Analyst 도구에서 운영 체제 프로필을 선택할 수 있습니다. 운영 체제 프로필을 선택하면 운영 체제 프로필 사용자의 사용 권한을 기반으로 데이터 통합 서비스가 열 프로필, 엔터프라이즈 검색 프로필 및 성과 기록표를 작성하고 실행합니다.

Analyst 도구에서는 기본 프로필을 사용하여 프로필 및 성과 기록표를 실행합니다. 운영 체제 프로필이 하나만 있는 경우에는 해당 운영 체제 프로필이 기본적으로 선택됩니다. 운영 체제 프로필이 여러 개 있는 경우에는 운영 체제 프로필 중에서 선택할 수 있습니다.

운영 체제 프로필 선택

Informatica Analyst에서 운영 체제 프로필을 선택할 수 있습니다. 데이터 통합 서비스는 운영 체제 프로필 사용자의 사용 권한을 이용하여 프로파일링 작업을 실행합니다.

1. Informatica Analyst 헤더 영역에서 <사용자 이름> > 설정을 클릭합니다.
설정 대화 상자가 나타납니다.
2. 운영 체제 프로필을 선택합니다. **저장**을 클릭합니다.

리포지토리 자산 잠금 및 팀 기반 개발 개요

모델 리포지토리는 사용자가 다른 사용자의 작업을 덮어쓰지 못하도록 프로필을 잠급니다. 모델 리포지토리가 버전 제어 시스템과 통합된 경우 여러 자산 버전이 저장되고 버전에 버전 번호가 할당됩니다. 프로필을 체크 아웃 및 체크 인하고 체크 아웃을 실행 취소할 수 있습니다. 체크 아웃한 프로필의 특정 버전을 볼 수 있습니다.

Analyst 도구에서 프로필 편집을 시작하면 다른 사용자가 프로필을 편집할 수 없도록 모델 리포지토리에서 프로필을 잠급니다. 프로필을 저장하면 잠금이 유지됩니다. 프로필을 닫으면 모델 리포지토리가 프로필을 잠금 해제합니다.

모델 리포지토리는 버전 지정 자산 관리를 통해 팀의 다른 멤버가 덮어쓰지 못하도록 프로필을 보호합니다. 다른 사용자가 체크 아웃한 프로필을 편집하려고 하면 프로필을 체크 아웃한 사용자를 나타내는 알림을 받게 됩니다. 체크 아웃된 프로필을 읽기 전용 모드에서 열거나 프로필을 다른 이름으로 저장할 수 있습니다.

프로필 속성 대화 상자에서 프로필의 버전을 선택하여 해당 버전에 대한 프로필 정의를 볼 수 있습니다. 작업 메뉴에서 프로필 속성 옵션에 액세스할 수 있습니다. 리포지토리 자산 잠금 및 버전 지정 자산 관리에 대한 자세한 내용은 *Analyst 도구 가이드*를 참조하십시오.

Informatica Analyst에서 열 프로필 생성

사용자 지정 프로필 또는 기본 프로필을 작성할 수 있습니다. 사용자 지정 프로필을 작성하는 경우 열, 샘플 행 및 드릴다운 옵션을 구성할 수 있습니다. 기본 프로필을 작성하는 경우 열 프로필 및 데이터 도메인 검색이 모든 데이터 도메인과 함께 전체 데이터 집합에서 실행됩니다.

1. 검색 작업 공간에서 **프로필**을 클릭하거나 헤더 영역에서 **새로 만들기 > 프로필**을 선택합니다.
참고: 라이브러리 작업 공간에서 데이터 개체를 마우스 오른쪽 단추로 클릭하고 프로필을 작성할 수 있습니다. 이 프로필에서 프로필 이름, 위치 이름, 데이터 개체가 데이터 개체 속성에서 추출됩니다. 기본 프로필을 작성하거나 설정을 사용자 지정하여 사용자 지정 프로필을 작성할 수 있습니다.
새 프로필 마법사가 표시됩니다.
2. **단일 소스** 옵션이 기본적으로 선택됩니다. **다음**을 클릭합니다.
3. **일반 속성 지정** 화면에서 프로필에 대한 이름과 선택적 설명을 입력합니다. 위치 필드에서 프로필을 작성할 프로젝트 또는 폴더를 선택합니다. **다음**을 클릭합니다.
4. **소스 선택** 화면에서 **선택**을 클릭하여 데이터 개체를 선택하거나 **새로 만들기**를 클릭하여 데이터 개체를 가져옵니다. **다음**을 클릭합니다.
 - **데이터 개체 선택** 대화 상자에서 데이터 개체를 선택합니다 **확인**을 클릭합니다.

속성 창은 선택된 데이터 개체의 속성을 표시합니다. 데이터 미리보기 창은 데이터 개체의 열을 표시합니다.

- **새 데이터 개체** 대화 상자에서 연결, 스키마, 테이블 또는 프로필을 작성할 보기를 선택하고 위치를 선택하고 데이터 개체를 가져올 폴더를 작성할 수 있습니다. **확인**을 클릭합니다.
5. **소스 선택** 화면에서 프로필을 실행할 열을 선택합니다. 필요에 따라 **이름**을 선택하여 모든 열을 선택합니다. **다음**을 클릭합니다.

기본적으로 모든 열이 선택됩니다. **Analyst** 도구는 각 열에 대한 기본 키의 이름, 데이터 유형, 전체 자릿수, 소수 자릿수, **null** 가능 및 참여와 같은 열 속성을 나열합니다.

6. **설정 지정** 화면에서 열 프로필, 데이터 도메인 검색 또는 열 프로필과 데이터 도메인 검색을 실행하도록 선택합니다. 기본적으로 열 프로필 옵션이 선택됩니다.

- **열 프로필 실행**을 선택하여 열 프로필을 실행합니다.
 - **데이터 도메인 검색 실행**을 선택하여 데이터 도메인 검색을 수행합니다. **데이터 도메인** 창에서 검색할 데이터 도메인을 선택하고 준수 조건을 선택한 후 **데이터 도메인 검색을 위한 열 선택 편집** 대화 상자에서 데이터 도메인 검색에 사용할 열을 선택합니다.
 - **열 프로필 실행 및 데이터 도메인 검색 실행**을 선택하여 열 프로필과 데이터 도메인 검색을 실행합니다. **데이터 도메인** 창의 데이터 도메인 옵션을 선택합니다.
- 참고:** 기본적으로 선택하는 열은 열 프로필과 데이터 도메인 검색을 위한 것입니다. 데이터 도메인 검색을 위한 열을 선택하거나 선택 취소하려면 **편집**을 클릭하십시오.
- 데이터 도메인 검색을 실행할 데이터, 열 또는 데이터와 열을 선택합니다.
 - 샘플링 옵션을 선택합니다. **프로필 실행** 창에서 **모든 행(완전한 분석)**, **샘플 먼저**, **무작위 샘플링**, **무작위 샘플링(자동)**, **n 제한** 또는 **무작위 비율**을 샘플링 옵션으로 선택할 수 있습니다. 이 샘플링 옵션은 열 프로필과 데이터 도메인 검색에 적용됩니다.
 - 드릴다운 옵션을 선택합니다. **라이브** 또는 **준비됨** 드릴다운 옵션을 선택하거나 **꺼짐**을 선택하여 **드릴다운** 창에서 드릴다운을 비활성화할 수 있습니다. 필요에 따라 **열 선택**을 클릭하여 드릴다운할 열을 선택합니다. 승인된 데이터 유형 또는 데이터 도메인이 포함된 열에 대한 데이터 유형 및 데이터 도메인 유추를 생략하도록 선택할 수 있습니다.
 - 런타임 환경으로 **원시**, **Blaze**, **Spark** 또는 **Databricks**를 선택합니다. **Blaze** 또는 **Spark**를 선택하는 경우 **선택**을 클릭하여 **Hadoop 연결 선택** 대화 상자에서 Hadoop 연결을 선택합니다. **Databricks**를 선택하는 경우 **선택**을 클릭하고 Databricks 연결을 선택합니다.

7. **다음**을 클릭합니다.

규칙 및 필터 지정 화면이 열립니다.

8. **규칙 및 필터 지정** 화면에서 다음 태스크를 수행할 수 있습니다.

- 규칙을 작성, 편집 또는 삭제합니다. 프로필에 기존 규칙을 적용할 수 있습니다.
- 필터를 작성, 편집 또는 삭제합니다.

참고: 이 프로필에 대한 성과 기록표를 작성하는 경우 프로필에 대해 작성된 필터를 재사용할 수 있습니다.

9. **저장 및 마침**을 클릭하여 프로필을 작성하거나 **저장 및 실행**을 클릭하여 프로필을 작성 및 실행합니다.

열 프로필 편집

열 프로필을 실행한 후에 변경할 수 있습니다.

1. **라이브러리** 작업 공간에서 프로필이 포함된 프로젝트를 선택하거나 **자산** 창의 프로필을 선택합니다.

2. 프로필 이름을 클릭합니다.
요약 보기가 **검색** 작업 공간에 표시됩니다.
3. 버전 제어 시스템이 활성화된 경우 **작업 > 체크 아웃**을 클릭하여 프로필을 체크 아웃합니다.
4. **작업 > 프로필 편집**을 클릭합니다.
프로필 마법사가 나타납니다.
5. 변경하려는 항목에 따라 다음 페이지 옵션 중 하나를 선택합니다.
 - **일반 속성 지정.** 이름, 설명 및 위치와 같은 기본 속성을 변경합니다.
 - **소스 선택.** 일치하는 다른 데이터 소스와 프로필을 실행할 열을 선택합니다.
 - **설정 지정.** 열 프로필 또는 열 프로필과 데이터 도메인 검색을 실행하도록 선택합니다. 검색할 데이터 도메인을 선택하고 데이터 도메인 검색, 샘플링 및 드릴다운 옵션을 수정합니다.
 - **규칙 및 필터 지정.** 규칙 및 필터를 작성, 편집 또는 삭제합니다.
6. **저장 및 마침**을 클릭하여 프로필 편집을 완료하거나 **저장 및 실행**을 클릭하여 프로필을 편집 및 실행합니다.
7. 버전 제어 시스템이 활성화된 경우 다음 태스크를 수행해야 합니다.
 - **저장 및 마침**을 클릭하여 프로필 편집을 완료합니다.
 - 요약 보기에서 **체크 인**을 클릭하여 프로필을 체크 인합니다.
 - **작업 > 프로필 실행**을 클릭하여 프로필을 실행합니다.

프로필 실행

프로필을 실행하여 데이터 소스의 콘텐츠와 구조를 분석하고 드릴다운할 열과 규칙을 선택합니다. 열 및 규칙에 대해 라이브 데이터 또는 준비된 데이터를 드릴다운할 수 있습니다. 초기 프로필 실행 후 모든 소스 열에서 프로필을 실행하지 않고 하나의 열 또는 규칙에서만 프로필을 실행할 수 있습니다.

1. **라이브러리** 작업 공간의 프로젝트 창에서 프로필이 포함된 프로젝트 또는 폴더를 선택하거나, **자산** 창에서 프로필을 선택합니다.
2. **작업 > 열기**를 클릭합니다.
요약 보기가 **검색** 작업 공간에 나타납니다.
3. **작업 > 프로필 실행**을 클릭합니다.

Analyst 도구는 프로필 실행을 수행하고 요약 보기에서 프로필 결과를 표시합니다.

프로필 요약과 매핑 로그 파일을 확인하면 **Analyst** 도구에서 수행한 태스크에 대한 자세한 정보를 볼 수 있습니다.

참고: 운영 체제 프로필을 사용하여 프로필을 실행한 경우 요약 로그는 데이터 통합 서비스에 대해 구성된 로그 디렉터리에 나타나고 매핑 로그는 운영 체제 프로파일에 대해 구성된 로그 디렉터리에 나타납니다.

Spark 엔진에서 프로필 실행

Spark 엔진에서 JDBC 연결과 함께 프로필을 실행하는 경우 프로필 실행이 실패합니다.

Spark 엔진에서 프로필을 실행하기 전에 다음 단계를 수행합니다.

1. JDBC 웨어하우스 연결을 생성합니다.
2. 데이터를 추출하는 데 사용하는 데이터베이스의 Data Direct JAR 파일을 얻습니다.
3. 파일을 다음 위치에 복사합니다. <INFA_HOME>/externaljdbcjars

Databricks 클러스터에서 프로필 실행

Databricks 클러스터를 사용하여 Spark 엔진에서 JDBC 연결과 함께 프로필을 실행하는 경우 프로필 실행이 실패합니다.

Spark 엔진에서 프로필을 실행하기 전에 다음 단계를 수행합니다.

1. JDBC 웨어하우스 연결을 생성합니다.
2. 데이터를 추출하는 데 사용하는 데이터베이스의 Data Direct JAR 파일을 얻습니다.
3. 파일을 다음 위치에 복사합니다. <INFA_HOME>/externaljdbcjars

동기화 옵션

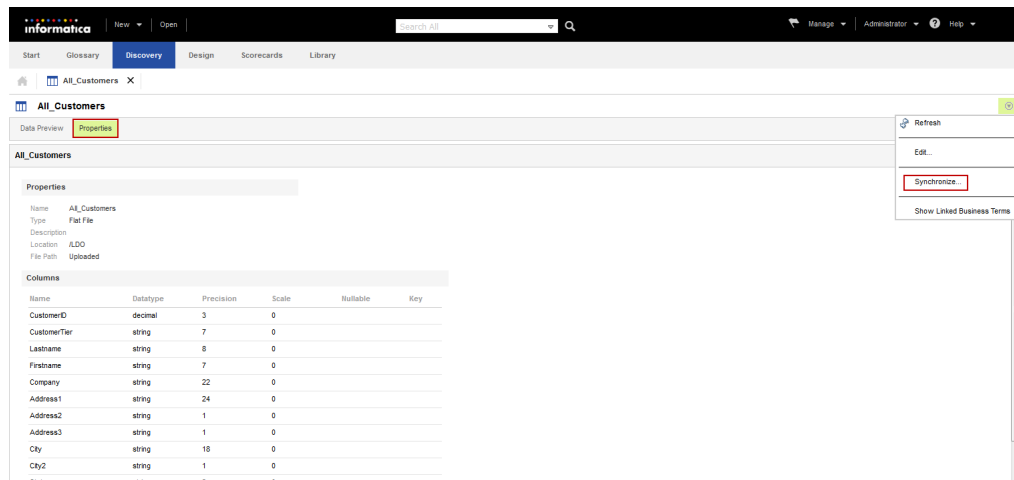
외부 데이터 소스의 메타데이터를 변경하면 모델 리포지토리의 데이터 개체 메타데이터가 기본적으로 업데이트되지 않습니다. 데이터 개체 메타데이터를 데이터 소스 메타데이터에 동기화하려면 동기화 옵션을 사용하십시오. 오. 열 프로필, 엔터프라이즈 검색 프로필 및 성과 기록표에 대해 동기화 옵션을 사용할 수 있습니다. 외부 데이터 소스는 관계형 데이터 소스나 플랫폼 파일 데이터 소스일 수 있습니다.

Informatica Analyst의 플랫폼 파일 데이터 개체 동기화

Analyst 도구에서 외부 플랫폼 파일 데이터 소스에 대한 변경 사항을 해당하는 데이터 개체와 동기화할 수 있습니다. 플랫폼 파일 동기화 마법사를 사용하여 데이터 개체를 동기화합니다.

1. 라이브러리 작업 공간을 엽니다.
2. 프로젝트 섹션에서 프로젝트의 플랫폼 파일 데이터 개체를 선택합니다.
Analyst 도구가 속성 탭에 테이블의 속성을 표시합니다.
3. 작업 메뉴에서 동기화를 클릭합니다.

다음 이미지는 작업 메뉴의 속성 탭과 동기화 옵션을 보여 줍니다.



플랫 파일 동기화 마법사가 나타납니다.

4. 위치를 찾으도록 선택하거나 플랫 파일을 가져올 네트워크 경로를 입력합니다.
 - 위치를 찾으려면 **파일 선택**을 클릭하여 시스템이 액세스할 수 있는 디렉터리에서 플랫 파일을 선택합니다.
 - 네트워크 경로를 입력하려면 **네트워크 경로 입력**을 선택하고 파일 경로 및 파일 이름을 구성합니다.
- 다음 이미지는 플랫 파일 동기화 마법사를 보여 줍니다.

Synchronize Flat File: Step 1 of 5

Specify a location to import the flat file from and specify how to import the flat file.

☒ **Browse and Upload:** Browse... No file selected.

☐ **Enter a Network Path:**

☐ **Hadoop File System**

Description

Upload files from a local machine. Recommended for smaller files up to 10 MB. The Analyst tool uploads a copy of the file to the node on which the Analyst Service runs. Upload the file again if you modify the file.

Back
Next
Finish
Cancel

5. 다음을 클릭합니다.
6. 구분자로 분리된 플랫 파일 또는 고정 너비 플랫 파일을 가져올지 선택합니다.
 - 구분자로 분리된 플랫 파일을 가져오려면 **구분자로 분리됨** 옵션을 선택합니다.

- 고정 너비 플랫폼 파일을 가져오려면 **고정 너비** 옵션을 선택합니다.
- 7. **다음**을 클릭합니다.
- 8. 구분자로 분리된 플랫폼 파일 또는 고정 너비 플랫폼 파일에 대한 플랫폼 파일 옵션을 구성합니다.
- 9. **다음**을 클릭합니다.
- 10. 필요에 따라 열 특성을 변경합니다.
- 11. **다음**을 클릭합니다.
- 12. 플랫폼 파일의 기본 이름을 허용하거나 다른 이름을 입력합니다.
- 13. 필요에 따라 설명을 입력합니다.
- 14. **마침**을 클릭합니다.
동기화 메시지에서 작업을 확인하라는 메시지를 표시합니다.
- 15. 플랫폼 파일을 동기화하려면 **예**를 클릭합니다.
동기화가 완료되었음을 나타내는 메시지가 표시됩니다. 메타데이터 변경에 대한 세부 정보를 보려면 **세부 정보 표시**를 클릭합니다.
- 16. **확인**을 클릭합니다.

Informatica Analyst의 관계형 데이터 개체 동기화

외부 관계형 데이터 소스에 대한 변경 사항을 해당하는 테이블 데이터 개체와 동기화할 수 있습니다. 외부 데이터 소스 변경 사항에는 소스 열 및 규칙 열의 추가, 변경 및 제거가 포함됩니다.

1. **라이브러리** 작업 공간을 엽니다.
2. **프로젝트** 섹션에서 프로젝트의 테이블 데이터 개체를 선택합니다.
Analyst 도구가 **속성** 탭에 테이블의 속성을 표시합니다.
3. 작업 메뉴에서 **동기화**를 클릭합니다.
작업을 확인하라는 메시지가 표시됩니다.
4. 동기화 프로세스를 완료하려면 **예**를 클릭합니다.
동기화 상태 메시지가 표시됩니다.
5. 동기화가 완료되었음을 나타내는 메시지가 표시됩니다.
메타데이터 변경에 대한 세부 정보를 보려면 **세부 정보 표시**를 클릭합니다.
6. **확인**을 클릭합니다.

제 5 장

Informatica Analyst의 규칙

이 장에 포함된 항목:

- [Informatica Analyst의 규칙 개요, 31](#)
- [미리 정의된 규칙, 31](#)
- [식 규칙, 32](#)

Informatica Analyst의 규칙 개요

규칙은 열 프로필을 실행할 때 소스 데이터에 적용되는 조건을 정의하는 비즈니스 논리입니다. 프로필에 규칙을 추가하여 데이터 유효성을 검사할 수 있습니다.

다른 상황에서 규칙을 사용할 수도 있습니다. 하나 이상의 데이터 열을 정리하는 규칙을 추가할 수 있습니다. 소스 데이터가 제공하지 않는 정보를 제공하는 조회 규칙을 추가할 수 있습니다. 데이터 품질 또는 데이터 통합 프로젝트에 대한 정리 규칙 유효성을 검사하는 규칙을 추가할 수 있습니다.

열 프로필을 생성하거나 편집할 때 규칙을 생성하고 프로필에 규칙을 추가하거나, 기존 규칙을 프로필에 적용할 수 있습니다. 열 프로필에서 식 규칙 또는 미리 정의된 규칙을 사용할 수 있습니다.

프로필을 실행한 후 **Analyst** 도구는 요약 보기에 규칙 열의 프로필 결과를 표시합니다. 상세 보기에서 규칙에 대한 열 결과를 볼 수 있습니다. 규칙의 출력은 하나 이상의 가상 열일 수 있습니다. 가상 열은 프로필 결과에 있습니다. **Analyst** 도구는 가상 열에서 프로필을 실행합니다. 이름과 성이 포함된 열을 **FIRST_NAME** 및 **LAST_NAME** 가상 열로 분할하는 미리 정의된 규칙을 사용하는 경우를 예로 들어 보겠습니다. **Analyst** 도구는 **FIRST_NAME** 및 **LAST_NAME** 열에서 프로필을 실행합니다.

참고: 다른 개체 유형이 참조하는 규칙 개체를 삭제하면 **Analyst** 도구는 해당 개체 유형을 나열하는 메시지를 표시합니다. 규칙을 삭제하기 전에 규칙 삭제의 영향을 확인합니다.

미리 정의된 규칙

미리 정의된 규칙은 **Developer tool**에서 작성되거나 **Developer tool** 및 **Analyst** 도구에서 제공되는 규칙입니다. 미리 정의된 규칙을 열 프로필에 적용하여 소스 데이터를 수정하거나 유효성을 검사합니다.

미리 정의된 규칙은 변환을 사용하여 규칙 논리를 정의합니다. 여러 프로필에서 미리 정의된 규칙을 사용할 수 있습니다. 모델 리포지토리의 미리 정의된 규칙은 규칙 논리를 정의하는 변환, 출력 그룹 및 입력 그룹이 포함된 맵셋입니다.

미리 정의된 규칙 프로세스

새 규칙 마법사를 사용하여 미리 정의된 규칙을 프로필에 적용합니다.

다음 단계를 수행하여 미리 정의된 규칙을 적용할 수 있습니다.

1. 프로필을 엽니다.
2. 미리 정의된 규칙을 선택합니다.
3. 규칙 매개 변수를 검토합니다.
4. 입력 열을 선택합니다. 둘 이상의 열에 규칙을 적용하려는 경우 여러 열을 선택하면 됩니다.
5. 프로파일링 옵션을 구성합니다.

미리 정의된 규칙 적용

미리 정의된 규칙을 적용할 때는 규칙을 선택하고 규칙에 대한 입력 열과 출력 열을 구성합니다. 재사용 가능 규칙으로 승격된 규칙 또는 개발자가 작성한 규칙을 사용하려면 미리 정의된 규칙을 적용합니다.

1. 라이브러리 작업 공간에서 프로필이 포함된 프로젝트를 선택하거나 **자산** 창의 프로필을 선택합니다.
2. 작업 > 열기를 클릭하여 프로필을 엽니다.
요약 보기가 검색 작업 공간에 표시됩니다.
3. 작업 > 프로필 편집을 클릭합니다.
프로필 마법사가 나타납니다.
4. 규칙 및 필터 지정을 클릭합니다.
5. 규칙 및 필터 지정 화면의 규칙 패널에서 작업 > 기존 규칙 적용을 클릭합니다.
규칙 적용 마법사 대화 상자가 나타납니다.
6. 규칙을 선택하고 다음을 클릭합니다.
7. 추가를 클릭합니다.
입력 포트에 대한 열 선택 대화 상자가 나타납니다.
8. 필드 및 입력 열을 선택합니다. 확인을 클릭합니다.
입력 열 및 출력 열이 규칙 적용 마법사 대화 상자에 나타납니다.
9. 규칙 적용 마법사 대화 상자에서 확인을 클릭합니다.
규칙이 규칙 및 필터 지정 화면에 나타납니다.

식 규칙

식 규칙은 식 함수와 열을 사용하여 규칙 논리를 정의합니다. Analyst 도구에서 식 규칙을 생성한 다음 열 프로필에 추가합니다.

식 규칙을 사용하여 열 프로필에서 열의 값 유효성을 검사하거나 값을 변경합니다. 프로필에서 사용할 식 규칙을 하나 이상 작성할 수 있습니다. 식 함수는 소스 데이터를 변환하는 데 사용되는 SQL 형식의 함수입니다. 다음과 같은 유형의 함수로 식 규칙 논리를 작성할 수 있습니다.

- 문자
- 변환

- 데이터 정리
- 날짜
- 인코딩
- 재무
- 숫자
- 지수
- 특수
- 테스트

다음 방법을 사용하여 식 규칙을 생성할 수 있습니다.

- **프로필 마법사.** 열 프로필을 생성하거나 편집할 때 프로필 마법사에서 식 규칙을 생성하고 적용할 수 있습니다. 규칙을 재사용 가능 규칙으로 승격하고 여러 프로필에서 사용할 수 있습니다.
- **규칙 사양.** **Analyst** 도구에서 규칙 사양을 구성하고 열 프로필에서 규칙 사양을 사용할 수 있습니다. 규칙 사양을 구성하는 경우 비즈니스 규칙의 요구 사항을 하나 이상의 규칙 문으로 변환합니다. 규칙 문은 데이터 집합이 비즈니스 규칙을 준수하는지 여부를 결정하는 논리를 나타냅니다. 규칙 사양에서 맷을 생성하고 **Developer tool**에서 생성한 열 프로필에서 맷을 사용합니다.

식 편집기를 사용하여 식 함수를 추가하고, 열을 함수에 대한 입력으로 구성하고, 식 유효성을 검사하고, 반환 유형, 전체 자릿수, 소수 자릿수를 구성할 수 있습니다. 식 규칙을 생성 및 유효성 검사한 후에는 출력 규칙 열의 전체 자릿수 값을 편집할 수 있습니다. 기본적으로 출력 규칙 열의 전체 자릿수 값은 **10**으로 설정되어 있습니다. 출력 규칙 열이 설정된 전체 자릿수 값을 초과하는 경우 전체 자릿수 값이 잘립니다.

식 규칙의 출력은 규칙 이름을 열 이름으로 사용하는 가상 열입니다. **Analyst** 도구는 가상 열에서 열 프로필을 실행합니다. 식 규칙을 사용하여 우편 번호 유효성을 검사하는 경우를 예로 들어 보겠습니다. 이 규칙은 우편 번호가 유효하면 **1**을 반환하고 유효하지 않으면 **0**을 반환합니다. **Informatica Analyst**는 규칙의 출력 값 **1** 및 **0**에 대해 열 프로필을 실행합니다.

식 규칙 작성

프로필 마법사를 사용하여 식 규칙을 작성한 다음 프로필에 추가합니다. 식 규칙을 작성하여 프로필에서 열의 값 유효성을 검사합니다.

1. 프로필을 엽니다.
2. 요약 보기에서 **작업 > 프로필 편집**을 클릭하여 **프로필** 마법사를 엽니다.
3. **규칙 및 필드 지정**을 클릭합니다.
4. 규칙 창에서 **작업 > 규칙 추가**를 클릭합니다.



새 규칙 대화 상자가 나타납니다.

5. 새 규칙 대화 상자에서 규칙의 이름 및 선택적 설명을 입력합니다. 함수 패널 또는 열 패널에서 규칙을 작성할 수 있습니다.
 - 함수 패널에서 함수 범주를 선택하고 오른쪽 화살표(>>) 단추를 클릭합니다. 대화 상자에서 매개 변수를 지정하고 **확인**을 클릭합니다.
함수가 열 및 값과 함께 식 패널에 나타납니다.
 - 열 패널에서 열을 선택하고 오른쪽 화살표(>>) 단추를 클릭합니다. 열이 식 패널에 나타납니다. 함수, 식 및 값을 추가하여 규칙을 작성합니다.
6. 규칙을 확인하려면 **유효성 검사**를 클릭합니다.
7. 필요에 따라 규칙을 재사용 가능 규칙으로 승격하도록 선택하고 프로젝트 및 폴더 위치를 구성합니다. 규칙을 재사용 가능 규칙으로 승격하는 경우 자신이나 다른 사용자가 다른 프로필의 규칙을 미리 정의된 규칙으로 사용할 수 있습니다.
8. **확인**을 클릭합니다.
규칙 및 필터 지정 화면이 규칙 창에 규칙과 함께 나타납니다.

규칙 사양을 사용하여 식 규칙 생성

규칙 사양을 사용하여 Informatica Analyst에서 식 규칙을 생성할 수 있습니다. 열 프로필에 규칙을 추가하여 데이터 유효성을 검사할 수 있습니다.

1. 헤더 영역에서 **새로 만들기 > 규칙 사양**을 클릭합니다.
새로운 규칙 사양 마법사가 나타납니다.

2. **새로운 규칙 사양** 마법사에서 규칙의 이름 및 선택적 설명을 입력합니다.
3. **위치 필드**에서 **찾아보기**를 클릭하여 규칙을 저장할 프로젝트나 폴더를 선택합니다.
4. **계속**을 클릭합니다.
규칙 사양이 **디자인** 작업 공간에 나타납니다.
5. 규칙의 속성을 입력하려면 규칙에서 최상위 수준의 8각형 모양을 선택한 후 **속성**을 클릭합니다.
6. 기본 규칙 집합을 구성하려면 규칙에서 그 다음 수준의 사각형 모양을 클릭합니다.
7. 규칙 집합에 입력을 지정하려면 **속성 > 입력**을 클릭합니다.
입력 관리 대화 상자가 나타납니다.
8. **입력 관리** 대화 상자에서 **입력 추가**를 클릭하고 입력에 대한 이름, 데이터 유형, 최대 길이 및 설명을 입력합니다. 선택적으로 여러 입력을 지정할 수 있습니다.
9. **확인**을 클릭합니다.
입력은 **속성** 섹션에 나타납니다.
10. 규칙 논리를 정의하려면 **규칙 논리**를 클릭하고 연산자, 조건을 입력한 다음 **작업** 목록에서 작업을 선택합니다.
11. 필요에 따라 여러 규칙 집합을 입력합니다(선택 사항).
12. 규칙의 유효성을 검사하려면 **유효성 검사**() 아이콘을 클릭합니다.
13. 열 프로필에서 규칙 사양을 저장하고 사용하려면 **저장 및 마침**을 클릭합니다.
14. 규칙을 저장하고 계속 작업하려면 **저장 및 계속**을 클릭합니다.
15. **Developer tool**에서 규칙 사양을 사용하려면 **규칙 생성**() 아이콘을 클릭하여 맵셋을 생성합니다.
Analyst 도구가 모델 리포지토리에 맵셋을 생성합니다. 규칙으로 맵셋의 유효성을 검사한 후 **Developer tool**에서 생성하는 맵셋을 열 프로필에서 사용합니다.

제 6 장

Informatica Analyst의 필터

이 장에 포함된 항목:

- [Informatica Analyst의 필터 개요, 36](#)
- [필터 작성, 36](#)
- [필터 관리, 39](#)

Informatica Analyst의 필터 개요

필터 조건을 충족하는 원래 데이터 소스의 하위 집합을 작성할 수 있도록 필터를 작성할 수 있습니다. 그런 다음 필터링된 데이터에 대해 프로필을 실행할 수 있습니다.

필터를 작성하여 필터 조건을 충족하는 프로필 결과를 볼 수 있습니다. 요약 보기에서 사용 가능한 기본 필터가 포함된 프로필 결과를 볼 수 있습니다.

필터 작성

필터 조건을 충족하는 원래 데이터 소스의 하위 집합을 작성할 수 있도록 필터를 작성할 수 있습니다.

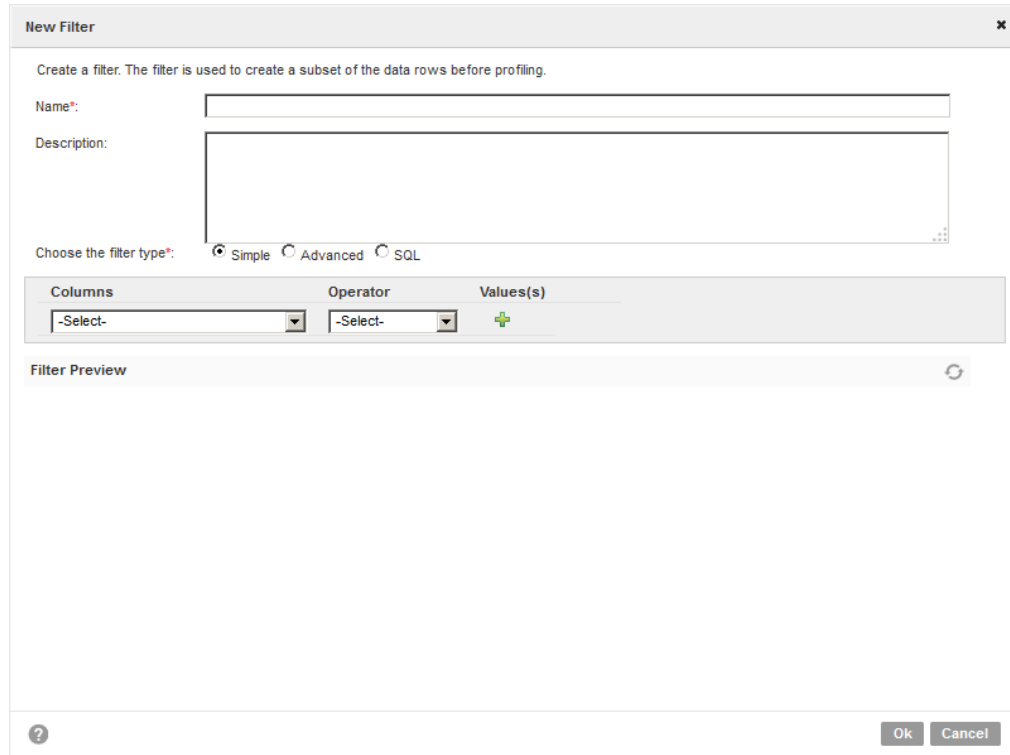
1. 프로필을 엽니다.
2. 요약 보기에서 **작업 > 프로필 편집**을 클릭합니다.
프로필 마법사가 나타납니다.
3. **규칙 및 필터 지정**을 클릭합니다.
4. **필터** 창에서 **작업 > 필터 추가**를 클릭합니다.
새 필터 대화 상자가 나타납니다.
5. 단순, 고급 또는 SQL 필터를 작성합니다.
참고: 날짜 열에 대한 단순 또는 고급 필터의 경우 YYYY/MM/DD HH:MM:SS 형식으로 조건을 입력합니다.
데이터 미리보기 창은 필터 조건을 충족하는 원래 데이터 소스의 하위 집합을 표시합니다.
6. **확인**을 클릭합니다.
규칙 및 필터 지정 화면이 **필터** 창에 필터와 함께 나타납니다.

단순 필터 작성

=, !=, >, < 같은 조건 연산자를 사용하여 단순 필터를 작성할 수 있습니다. 필터를 사용하여 원래 데이터 소스의 하위 집합을 작성합니다.

1. 새 필터 대화 상자에서 단순을 클릭합니다.

다음 이미지는 새 필터 대화 상자에서 단순 필터를 작성하는 데 사용할 수 있는 옵션을 보여 줍니다.



The 'New Filter' dialog box is shown. It has a title bar with a close button. The main area contains a text box for 'Name*', a larger text box for 'Description', and three radio buttons for 'Simple' (selected), 'Advanced', and 'SQL'. Below these is a table with three columns: 'Columns', 'Operator', and 'Values(s)'. The 'Columns' column has a dropdown menu with '-Select-' selected. The 'Operator' column has a dropdown menu with '-Select-' selected. The 'Values(s)' column has a green plus icon. At the bottom of the dialog is a 'Filter Preview' section with a refresh icon, and 'Ok' and 'Cancel' buttons.

Columns	Operator	Values(s)
-Select-	-Select-	+

2. 이름 및 선택적 설명을 입력합니다.
3. 열을 선택합니다.
4. 조건 연산자를 선택합니다.
5. 값을 입력합니다.
6. 필요에 따라 더하기(+) 아이콘을 클릭하여 필터를 추가합니다.
7. 확인을 클릭합니다.

규칙 및 필터 지정 페이지가 필터 창에 필터와 함께 나타납니다.

고급 필터 작성

AND, OR 및 NOT과 같은 식으로 고급 필터를 작성하여 원래 데이터 소스의 하위 집합을 작성할 수 있습니다.

1. 새 필터 대화 상자에서 고급을 클릭합니다.

다음 이미지는 새 필터 대화 상자의 고급 필터 옵션을 보여 줍니다.

The image shows the 'New Filter' dialog box with the 'Advanced' tab selected. The dialog has a title bar 'New Filter' and a close button. Below the title bar is a text area for 'Name*' and a larger text area for 'Description'. Below these is a section 'Choose the filter type*' with three radio buttons: 'Simple', 'Advanced' (selected), and 'SQL'. Below the radio buttons are two tabs: 'Functions' (selected) and 'Columns'. The 'Functions' tab shows a list of function categories: All, Character, Consolidation, Conversion, Data Cleansing, Date, Decision, and Encoding. To the right of this list is a '>>' button. Below the list is a large text area for the 'Expression'. To the right of the expression area is a 'Validate' button. At the bottom of the dialog is a toolbar with logical operators: AND, OR, NOT, (,), =, !=, <, <=, >=, >. At the very bottom are 'Ok' and 'Cancel' buttons.

2. 고급 필터의 이름과 선택적 설명을 입력합니다.
3. 함수 패널 또는 열 패널로 고급 필터를 작성할 수 있습니다.
 - 함수 패널에서 함수 범주를 선택하고 오른쪽 화살표(>>) 단추를 클릭합니다.
대화 상자에서 매개 변수를 지정하고 **확인**을 클릭합니다. 함수가 열 및 값과 함께 식 패널에 나타납니다.
 - 열 패널에서 열을 선택하고 오른쪽 화살표(>>) 단추를 클릭합니다. 열이 식 패널에 나타납니다.
함수, 식 및 값을 추가하여 고급 필터를 작성합니다.
4. 고급 필터를 확인하려면 **유효성 검사**를 클릭합니다.
5. **확인**을 클릭합니다.
규칙 및 필터 지정 화면이 필터 창에 필터와 함께 나타납니다.

SQL 필터 작성

SQL 쿼리로 SQL 필터를 작성할 수 있습니다. 관계형 데이터 소스용으로 SQL 필터를 작성할 수 있습니다.

1. 새 필터 대화 상자에서 SQL을 클릭합니다.

다음 이미지는 새 필터 대화 상자의 SQL 필터 옵션을 보여 줍니다.

The image shows a 'New Filter' dialog box. At the top, it says 'Create a filter. The filter is used to create a subset of the data rows before profiling.' Below this are fields for 'Name*' and 'Description:'. Underneath, there are radio buttons for 'Simple', 'Advanced', and 'SQL', with 'SQL' being selected. Below the radio buttons is a section titled 'Type or paste a SQL statement' with a 'Validate' button. At the bottom, there is a 'Filter Preview' section with a refresh icon. The dialog box has 'Ok' and 'Cancel' buttons at the bottom right.

2. SQL 필터의 이름과 선택적 설명을 입력합니다.
3. 텍스트 상자에서 SQL 쿼리를 입력하거나 붙여 넣습니다.
4. 유효성 검사를 클릭하여 SQL 쿼리를 확인합니다.
5. 확인을 클릭합니다.

규칙 및 필터 지정 페이지가 필터 창에 SQL 필터와 함께 나타납니다.

필터 관리

필터를 편집 및 삭제할 수 있습니다.

1. 라이브러리 작업 공간에서 프로필이 포함된 프로젝트를 선택하거나 필터링할 자산 창의 프로필을 선택합니다.
2. 프로필을 엽니다.
3. 요약 보기에서 작업 > 프로필 편집을 클릭하여 프로필 마법사를 엽니다.
4. 규칙 및 필터 지정을 클릭합니다.
5. 필터 창에서 필터를 선택하고 작업 > 필터 편집을 클릭합니다.

필터 편집 대화 상자가 나타납니다.

6. 필터 설정을 편집하고 **확인**을 클릭합니다.
7. 필터를 삭제하려면 필터를 선택하고 **작업 > 필터 삭제**를 클릭합니다.

제 7 장

Informatica Analyst의 열 프로필 결과

이 장에 포함된 항목:

- [Informatica Analyst의 열 프로필 결과 개요, 41](#)
- [요약 보기, 42](#)
- [상세 보기, 44](#)
- [통계, 46](#)
- [프로필 실행의 유형, 52](#)
- [여러 프로필 결과 비교 개요, 54](#)
- [열 프로필 드릴다운, 57](#)
- [Analyst 도구의 기간, 58](#)
- [Informatica Analyst의 열 프로필 내보내기 파일, 59](#)

Informatica Analyst의 열 프로필 결과 개요

프로필 결과를 보고 데이터의 콘텐츠, 구조 및 품질을 이해 및 분석합니다. 요약 보기에서 프로필의 모든 열 및 규칙을 볼 수 있습니다. 상세 보기에서 열 또는 규칙의 속성을 자세히 볼 수 있습니다.

검색 작업 공간에서 프로필 결과를 볼 수 있습니다. 보기 헤더는 프로필의 유형, 프로필의 열 수, 프로필의 규칙 수, 샘플링 데이터, 작성 날짜 및 시간을 표시합니다.

요약 보기에서 값, 가로 막대 차트 또는 백분율로 각 열의 속성을 볼 수 있습니다. Null 값, 고유한 값, 고유하지 않은 값, 패턴, 데이터 유형 및 데이터 도메인 같은 열 속성을 볼 수 있습니다. 기본 필터를 기반으로 요약 보기에서 프로필 결과를 볼 수 있습니다.

상세 보기에서 Null 값, 고유한 값, 고유하지 않은 값, 유추된 데이터 유형, 유추된 데이터 도메인, 유추된 패턴, 값 및 비즈니스 용어를 보고 창에서 데이터를 미리 볼 수 있습니다.

최근 실행, 기록 실행 및 통합된 실행에 대한 프로필 결과를 볼 수 있습니다. 2개의 프로필 실행에 대한 프로필 결과를 비교하고 요약 보기 및 상세 보기에서 결과를 볼 수 있습니다. 프로필 통계를 보고 데이터를 쿼리할 수 있습니다. 프로필 통계에는 열 및 규칙에 대한 값, 패턴, 데이터 유형, 이상값 및 통계가 포함됩니다. 데이터 검색을 수행하고 데이터를 드릴다운할 수 있습니다.

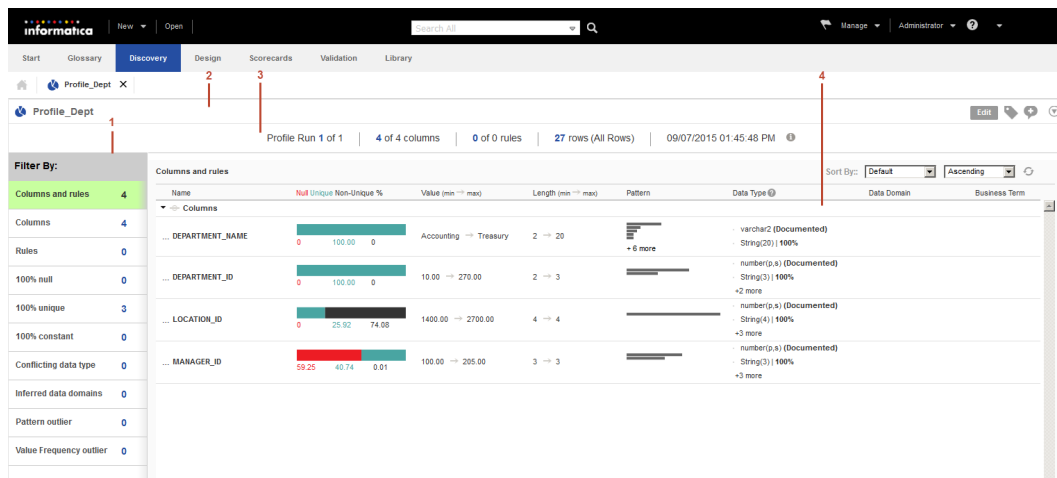
참고: Avro, JSON, Parquet 및 XML 데이터 소스에서 프로필을 보고 실행할 수 있습니다. 최근 실행, 기록 실행 및 통합된 실행에 대한 프로필 결과를 보고 2개의 프로필 실행에 대한 프로필 결과를 비교할 수 있습니다.

값 빈도, 패턴 빈도, 드릴다운 데이터, 설명, 태그 및 비즈니스 용어를 CSV 파일로 내보낼 수 있습니다. 추가 분석을 위해 모든 데이터를 파일에서 확인할 수 있도록 프로파일 요약 정보를 Microsoft Excel 파일로 내보낼 수 있습니다. 프로파일 결과의 규칙 정보를 볼 수 있습니다. 표시되는 프로파일 결과는 프로파일 구성 및 샘플링 옵션에 따라 달라집니다.

요약 보기

프로파일 결과의 요약은 요약 보기에서 그리드 형식으로 나타납니다. 요약 보기에서 기본 필터를 사용하여 특정 통계를 볼 수 있습니다. 예를 들어 규칙을 선택하는 경우 요약 보기가 프로파일의 모든 규칙을 표시합니다.

다음 이미지는 요약 보기의 샘플 그래픽 형식 보기를 보여 줍니다.



1. 기본 필터. 기본 필터를 기반으로 요약 보기에서 프로파일 결과를 볼 수 있습니다.
2. 프로파일 헤더. 헤더에서 프로파일 이름을 볼 수 있습니다. 편집 단추를 사용하여 프로파일을 편집하고, 태그 및 설명 아이콘을 사용하여 태그와 설명을 추가 또는 편집하고, 작업 메뉴에서 옵션을 선택합니다.
3. 요약 보기 헤더. 요약 보기 헤더에서 프로파일 관련 정보를 볼 수 있습니다. 프로파일 실행 수, 총 프로파일 실행 수, 열 및 규칙 수, 프로파일의 행 수를 볼 수 있습니다.
4. 요약 보기. 프로파일의 모든 열 및 규칙에 대한 속성을 볼 수 있습니다.

요약 보기에서 프로파일을 실행 또는 편집하고, 패턴 또는 값 빈도 이상값을 검색하고, 성과 기록표에 열을 추가하고, 프로파일 실행을 선택하고, 2개의 프로파일 실행을 비교하고, 프로파일 결과 또는 데이터 도메인 검색 결과를 Microsoft Excel 스프레드시트로 내보내고, 여러 열의 유추 결과를 확인하고, 설명 및 태그를 추가 또는 삭제하고, 프로파일 속성을 볼 수 있습니다.

요약 보기 속성

요약 보기는 프로필의 모든 열 및 규칙에 대한 속성을 표시합니다. 요약 보기에는 속성의 시각적 표현이 포함됩니다. 각 요약 속성을 클릭하여 속성의 값을 정렬할 수 있습니다.

다음 테이블에는 프로필 결과 요약 속성이 설명되어 있습니다.

속성	설명
이름	프로필의 열 또는 규칙의 이름을 표시합니다.
Null 고유 비고유 %	열 또는 규칙 출력에 대해 Null 값, 고유한 값 및 고유하지 않은 값을 백분율로 표시합니다. 가로 막대 차트에서 값을 볼 수 있습니다.
패턴	열의 여러 패턴을 가로 막대 차트로 표시합니다. 마우스를 막대 차트로 가져가면 열의 유사한 패턴 수와 패턴 문자를 백분율로 볼 수 있습니다.
값	열 또는 규칙 출력에서 최소값 및 최대값을 표시합니다.
길이	열 또는 규칙 출력에서 값의 최소 및 최대 길이를 표시합니다.
데이터 유형	<p>열 또는 규칙의 문서화된 데이터 유형을 표시합니다. 필드로 마우스를 가져가면 유추된 데이터 유형을 표시합니다. Analyst 도구는 다음 데이터 유형을 유추할 수 있습니다.</p> <ul style="list-style-type: none">- 문자열- Varchar- 10진수- 정수- 날짜 <p>유추된 데이터 유형을 기준으로 준수율을 볼 수도 있습니다.</p> <p>참고: Analyst 도구는 전체 자릿수가 38자리보다 긴 숫자 열의 값에서 데이터 유형을 파생시킬 수 없습니다. Analyst 도구는 전체 자릿수가 255자리보다 긴 문자열 열의 값에서 데이터 유형을 파생시킬 수 없습니다. 연도 값이 1800년 이전인 날짜 열에 대해 열 프로필을 작성하는 경우 유추된 데이터 유형이 고정 길이 문자열로 표시될 수 있습니다. 필요에 따라 InferDateTimeConfig.xml의 year-minimum 매개 변수 기본값을 변경합니다.</p>
데이터 도메인	열과 연결된 데이터 도메인의 이름과 함께 준수율 및 준수 행 수를 표시합니다.
비즈니스 용어	열에 할당된 비즈니스 용어를 표시합니다.

요약 보기의 기본 필터

기본 필터를 기반으로 요약 보기에서 프로필 결과를 볼 수 있습니다.

요약 보기는 기본적으로 모든 소스 열, 가상 열 및 규칙 열에 대한 프로필 결과를 표시합니다. 필터 기준 창은 기본 필터를 적용할 수 있는 열의 수를 표시합니다.

요약 보기에서 다음 기본 필터 옵션을 사용하여 프로필 결과를 볼 수 있습니다.

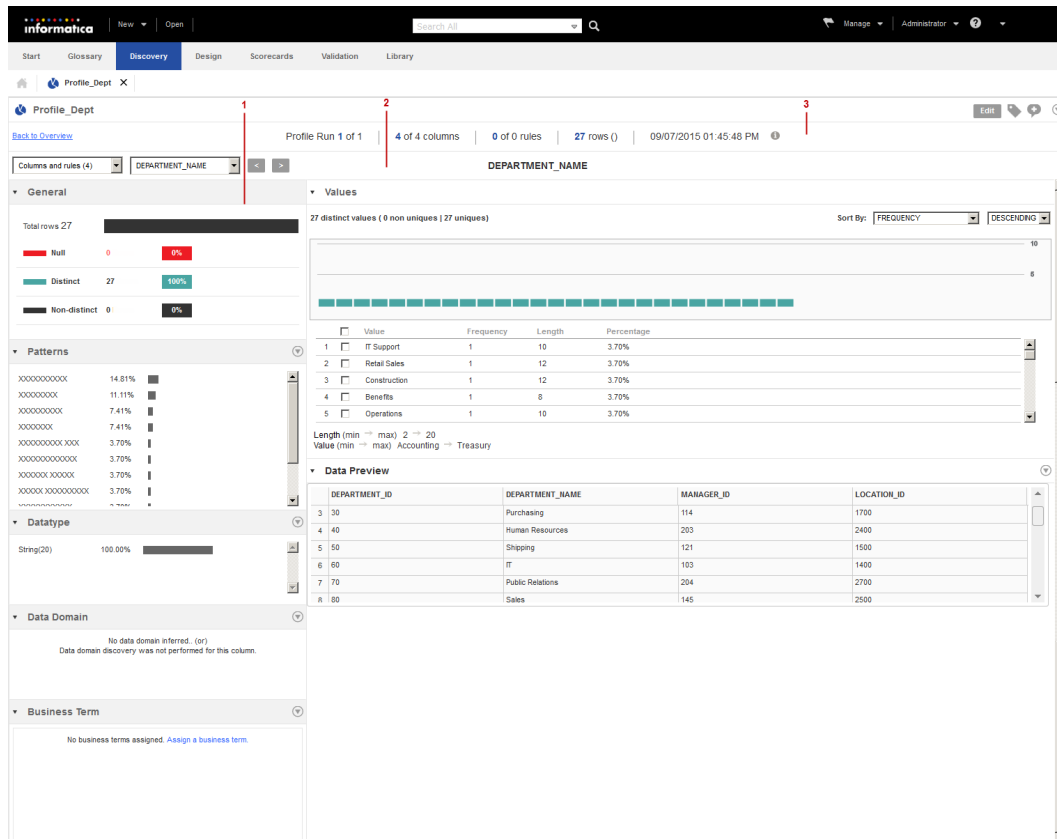
기본 필터 옵션	설명
열 및 규칙	소스 열 및 규칙 열에 대한 프로필 결과를 표시합니다. 소스 열 및 규칙 열을 확장 및 축소하여 결과를 볼 수 있습니다.
열	소스 열에 대한 프로필 결과를 표시합니다.
규칙	규칙 열에 대한 프로필 결과를 표시합니다.
100% null	100% null 값을 가진 열에 대한 프로필 결과를 표시합니다.
100% 고유	100% 고유한 값을 포함하는 열의 프로필 결과를 표시합니다.
100% 상수	모든 레코드에 대해 동일한 값을 가진 열에 대한 프로필 결과를 표시합니다. 예를 들어 100% 상수 필터에는 "미국" 값만 포함된 경우 국가 열의 프로필 결과가 포함되어 있습니다.
충돌하는 데이터 유형	문서화된 데이터 유형과 유추된 데이터 유형이 일치하지 않는 열에 대한 프로필 결과를 표시합니다. 예를 들어 이 필터는 열에 대한 문서화된 데이터 유형이 정수(2)이고 유추된 데이터 유형이 문자열이기 때문에 CustomerTier 열을 표시합니다.
유추된 데이터 도메인	유추된 데이터 도메인이 구성된 데이터 도메인과 동일한 열에 대한 프로필 결과를 표시합니다.
패턴 이상값	패턴 이상값을 가진 열에 대한 프로필 결과를 표시합니다.
값 빈도 이상값	값 또는 빈도 이상값을 가진 열에 대한 프로필 결과를 표시합니다.

상세 보기

열 결과는 상세 보기에 나타납니다. 열 속성을 자세하게 볼 수 있습니다.

요약 보기에서 열을 클릭하면 열에 대한 상세 보기가 나타납니다.

다음 이미지는 상세 보기에 표시된 열 속성의 샘플 그래픽 형식 보기를 보여 줍니다.



1. 창. 창에서 일반 속성, 열의 값, 데이터 미리보기, 유추된 패턴, 유추된 데이터 유형, 유추된 데이터 도메인 및 비즈니스 용어를 볼 수 있습니다.
2. 열 세부 정보 헤더. 드롭다운 목록에서 열을 선택하거나 탐색 단추를 사용하여 열 결과를 볼 수 있습니다.
3. 요약 보기 헤더. 요약 보기 헤더에서 프로파일 관련 정보를 볼 수 있습니다. 프로파일 실행, 열의 수, 규칙, 프로파일 실행의 행, 프로파일 실행의 시간과 날짜를 볼 수 있습니다.

상세 보기에서는 프로파일을 실행 또는 편집하고, 성과 기록표에 열을 추가하고, 프로파일 실행을 선택하고, 2개의 프로파일 실행을 비교하고, 프로파일 결과를 **Microsoft Excel** 스프레드시트로 내보내고, 값 빈도, 패턴 빈도, 데이터 유형, 선택한 값에 대한 드릴다운 데이터 또는 선택한 패턴에 대한 드릴다운 데이터를 **csv** 파일로 내보내고, 설명 및 태그를 열에 추가하거나 열에서 삭제하고, 프로파일 속성을 볼 수 있습니다.

열 속성에 대한 추가 작업을 수행하려면 각 창에서 작업 메뉴를 사용합니다. 창을 축소하거나 확장할 수 있습니다.

상세 보기 창

상세 보기에는 고유한 값, 고유하지 않은 값 및 **Null** 값의 개수와 백분율, 패턴, 유추된 데이터 유형, 유추된 데이터 도메인, 값, 데이터 미리보기 및 연결된 비즈니스 용어가 여러 창에 표시됩니다.

열 또는 규칙을 클릭하면 열 또는 규칙에 대한 상세 보기가 열립니다.

다음 테이블에는 상세 보기의 창이 설명되어 있습니다.

창	설명
일반	Null 값, 고유한 값 및 고유하지 않은 값이 포함된 행 수를 다른 색상으로 표시합니다. 값을 백분율로 볼 수 있습니다. 모든 연속 프로필 실행의 일반 값의 증가 및 감소를 스파크라인으로 볼 수 있습니다. 5개의 최근 연속 프로필 실행에 걸쳐 Null 값, 고유한 값 또는 고유하지 않은 값의 변형이 스파크라인으로 표시됩니다. 각 프로필 실행에 해당하는 스파크라인 위로 포인터를 이동하면 값의 개수와 값의 백분율을 볼 수 있습니다. 열에 태그와 설명을 추가할 수 있습니다.
패턴	열 값에 대한 패턴을 표시합니다. 패턴이 열에 나타나는 빈도가 가로 막대 차트 및 백분율로 나타납니다. 패턴에서 드릴다운하고, 패턴을 참조 테이블에 추가하거나, 선택한 패턴으로 데이터 도메인을 작성할 수 있습니다.
데이터 유형	열에 대해 유추된 데이터 유형을 표시합니다. 열의 데이터 유형의 빈도가 가로 막대 차트 및 백분율로 나타납니다. 데이터 유형에서 드릴다운하고, 선택한 유추 데이터 유형을 승인, 거부 또는 재설정할 수 있습니다. 거부된 항목 표시 옵션은 거부된 유추된 데이터 유형을 표시합니다.
데이터 도메인	열에 대해 유추된 데이터 도메인을 표시합니다. 데이터 도메인에서 드릴다운하여 준수 행, 비준수 행 또는 Null 값이 포함된 행을 볼 수 있습니다. 데이터 도메인 값을 승인, 거부 또는 재설정할 수 있습니다. 거부된 항목 표시 옵션은 거부된 데이터 도메인을 표시합니다. 데이터 도메인 값을 확인할 수 있습니다.
비즈니스 용어	열에 대해 할당된 비즈니스 용어를 표시합니다. 열에 비즈니스 용어를 할당하거나 할당 해제할 수 있습니다.
값	빈도, 길이 및 백분율과 함께 그래픽 표현으로 열의 모든 값을 표시합니다. 각 값에서 드릴다운할 수 있습니다. 참조 테이블에 값을 추가하고 값 빈도 규칙을 작성하고 데이터 도메인을 작성할 수 있습니다.
데이터 미리보기	선택한 패턴, 데이터 유형, 데이터 도메인 또는 값에 대한 드릴다운 데이터를 표시합니다.

통계

프로필의 열 및 규칙에 대한 값, 패턴, 데이터 유형, 데이터 도메인 및 이상값과 같은 통계를 볼 수 있습니다.

요약 보기에서 프로필 통계를 볼 수 있으며 요약 보기 및 상세 보기에서 열 통계를 볼 수 있습니다. 최근 프로필 실행, 기록 프로필 실행 및 통합된 프로필 실행에 대한 통계를 볼 수 있습니다. 2개의 프로필 실행에 대한 프로필 결과를 비교하고 요약 보기 및 상세 보기에서 프로필 및 열에 대한 통계를 볼 수 있습니다.

데이터 미리보기

데이터 미리보기 창에서는 선택한 패턴, 데이터 유형, 데이터 도메인 또는 값에 대한 드릴다운 데이터를 볼 수 있습니다.

데이터 미리보기 창은 상세 보기에서 볼 수 있습니다. 요약 보기에서 열을 클릭하면 상세 보기가 나타나고, 데이터 미리보기 창은 기본적으로 축소되어 있습니다. **작업 > 미리보기 표시**를 클릭하여 열 데이터를 볼 수 있습니다.

다음 테이블에는 데이터 미리보기 창의 **작업** 메뉴에 있는 옵션이 설명되어 있습니다.

옵션	설명
필터에 추가	프로필 결과의 하위 집합에서 데이터 불규칙을 분석할 수 있도록 드릴다운 데이터를 필터링할 드릴다운 필터를 작성합니다.
필터 저장	드릴다운 필터를 저장합니다.
미리보기 표시	소스 행을 표시합니다.
데이터 내보내기	드릴다운 결과를 CSV 파일 또는 Microsoft Excel 파일로 내보냅니다.

데이터 유형

데이터 유형에는 프로필 결과의 각 열에 대해 유추된 모든 데이터 유형이 포함됩니다.

요약 보기 및 상세 보기에서 데이터 유형을 볼 수 있습니다. 요약 보기에서 문서화된 데이터 유형 및 유추된 데이터 유형을 볼 수 있습니다. **충돌하는 데이터 유형** 필터는 문서화된 데이터 유형과 유추된 데이터 유형 간에 충돌이 존재하는 열을 표시합니다. 상세 보기에서는 열에 대해 유추된 데이터 유형을 볼 수 있습니다. 열의 데이터 유형의 빈도가 가로 막대 차트 및 백분율로 나타납니다. 선택한 유추된 데이터 유형을 드릴다운, 승인, 거부 또는 재설정할 수 있습니다. 거부된 항목 표시 옵션은 거부된 유추된 데이터 유형을 표시합니다.

다음 테이블에는 데이터 유형에 대한 속성이 설명되어 있습니다.

속성	설명
데이터 유형	프로필의 열에 대해 문서화된 데이터 유형 및 유추된 데이터 유형의 목록을 표시합니다.
빈도	데이터 유형이 열에 대해 나타나는 숫자로 표시된 횟수를 표시합니다.
백분율	데이터 유형이 열에 대해 나타나는 백분율을 표시합니다.
드릴다운	열 데이터 유형에 따라 특정 소스 행으로 드릴다운합니다. 참고: 여러 유추된 데이터 유형을 선택하는 경우에는 드릴다운 작업을 수행할 수 없습니다.
상태	데이터 유형의 상태를 나타냅니다. 상태는 유추됨, 승인됨 또는 거부됨입니다. 유추됨 Analyst 도구가 유추한 열의 데이터 유형을 나타냅니다. 승인됨 열에 대해 승인된 데이터 유형을 나타냅니다. 데이터 유형을 승인하면 해당 데이터 유형이 모델 리포지토리에 커밋됩니다. 거부됨 열에 대해 거부된 데이터 유형을 나타냅니다.

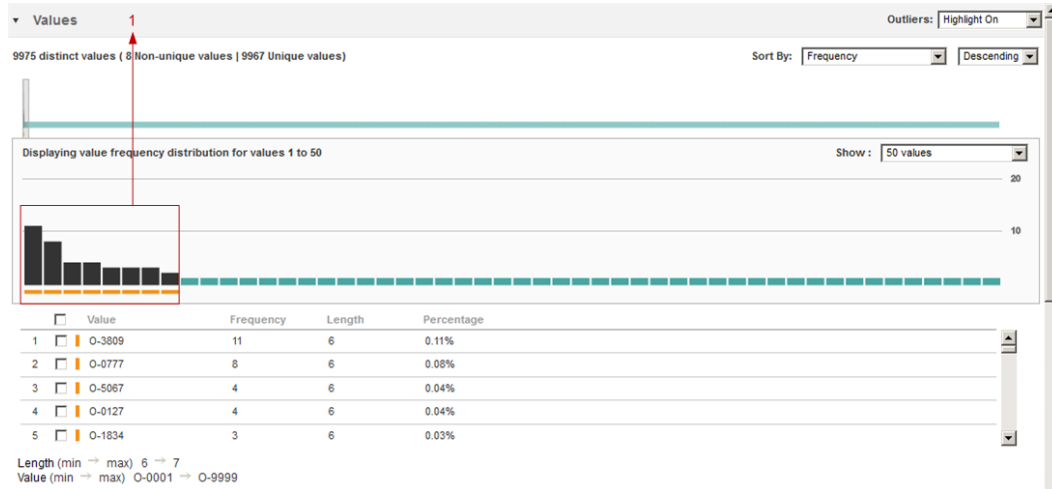
이상값

이상값은 값의 예상 범위에 속하지 않는 프로필 결과의 열에 대한 패턴, 값 또는 빈도입니다.

데이터 통합 서비스에 포함된 프로파일링 플러그인은 열의 대다수 값의 범위에 속하지 않는 값을 식별하는 알고리즘을 실행합니다. 열의 이러한 대다수 값의 예상 범위에 속하지 않는 모든 패턴, 값 또는 빈도가 이상값입니다.

기본적으로 **Analyst** 도구는 프로필 결과의 이상값을 결정하지 않습니다. 요약 보기에서 이상값을 실행하여 이상값 결과를 볼 수 있습니다. 패턴 이상값 필터는 열의 패턴을 기준으로 이상값을 표시합니다. 값 빈도 이상값 필터는 열의 값 또는 빈도를 기준으로 이상값을 표시합니다. 요약 보기에서 다른 작업을 수행할 수 있도록 백그라운드에서 이상값 검색이 발생합니다.

상세 보기에서 값 창에 있는 목록의 **강조 표시 켜짐** 옵션을 선택하면 이상값을 볼 수 있습니다. 이상값이 주황색 밑줄이 포함된 세로 막대로 나타납니다. 이상값만 보려면 드롭다운 목록에서 **필터** 옵션을 선택해야 합니다.



1. 이상값. 이상값이 주황색 밑줄이 포함된 세로 막대로 나타납니다.

이상값 실행

이상값을 실행하여 값의 예상 범위에 속하지 않는 열의 패턴, 값 또는 빈도를 식별합니다.

1. 요약 보기에서 **작업 > 이상값 검색**을 클릭합니다.
필터 기준 창에서 패턴 이상값 및 값 빈도 이상값이 N/A에서 검색된 이상값의 수로 변경됩니다.
2. 필터 기준 창에서 **패턴 이상값**을 클릭합니다.
패턴 이상값이 포함된 열이 요약 보기에 나타납니다.
3. 필터 기준 창에서 **값 빈도 이상값**을 클릭합니다.
값 또는 빈도 이상값이 포함된 열이 요약 보기에 나타납니다.
4. 상세 보기의 이상값 드롭다운 목록에서 **강조 표시 켜짐**을 선택합니다.
값 창에서 이상값이 주황색 밑줄이 포함된 세로 막대로 나타납니다.
5. 이상값만 보려면 이상값 드롭다운 목록에서 **필터**를 클릭합니다.

패턴

패턴이 요약 보기 및 상세 보기에서 나타나는 열 값 및 빈도에 대한 패턴을 볼 수 있습니다.

요약 보기에서 열의 여러 패턴을 가로 막대 차트로 볼 수 있습니다. 마우스를 막대 차트로 가져가면 열의 유사한 패턴 수와 패턴 문자를 백분율로 볼 수 있습니다. 상세 보기에서 패턴이 열에서 가로 막대 차트 및 백분율로 나타나는 빈도를 볼 수 있습니다. 패턴을 드릴다운하고, 참조 테이블에 추가하거나, 선택한 패턴으로 데이터 도메인을 작성할 수 있습니다.

프로파일링 웨어하우스에는 기본적으로 프로필 결과에 대한 Null 값을 포함하여 고유한 최대 빈도 값 16,000개가 저장됩니다. 프로필 결과에 Null 값이 하나 이상 있으면 Analyst 도구는 Null 값을 패턴으로 표시할 수 있습니다.

참고: Analyst 도구는 전체 자릿수가 38자리보다 긴 숫자 열에 대해 패턴을 과생시킬 수 없습니다. Analyst 도구는 전체 자릿수가 255자리보다 긴 문자열 열에 대해 패턴을 과생시킬 수 없습니다.

다음 테이블에는 열 패턴에 대한 속성이 설명되어 있습니다.

속성	설명
패턴	프로필의 열에 대한 패턴을 표시합니다.
빈도	패턴이 열에 대해 나타나는 숫자로 표시된 횟수를 표시합니다.
백분율	패턴이 열에 대해 나타나는 백분율을 표시합니다.

다음 테이블에서는 패턴 문자 및 해당 문자가 나타내는 내용을 설명합니다.

문자	설명
'B' 또는 'b' 또는 ''	빈 공간을 나타냅니다.
'C' 또는 'c'	임의의 문자를 나타냅니다.
'L' 또는 'l'	임의의 소문자 알파벳 문자를 나타냅니다.
'T' 또는 't'	탭을 나타냅니다.
'U' 또는 'u'	임의의 대문자 알파벳 문자를 나타냅니다.
9	임의의 숫자 문자를 나타냅니다. Informatica Analyst에서는 최대 3개의 문자가 "9" 형식으로 개별 표시됩니다. 문자가 3개보다 많으면 괄호로 묶여 값으로 표시됩니다. 예를 들어 "9(8)" 형식은 8자리 숫자 값을 나타냅니다.
'X' 또는 'x'	임의의 알파벳 문자를 나타냅니다. Informatica Analyst에서는 최대 3개의 문자가 "X" 형식으로 개별 표시됩니다. 문자가 3개보다 많으면 괄호로 묶여 값으로 표시됩니다. 예를 들어 "X(6)" 형식은 "Boston"이라는 값을 나타낼 수 있습니다. 참고: 패턴 문자 X는 대/소문자를 구분하지 않으며 소스 데이터에서 대문자 또는 소문자를 나타낼 수 있습니다.
'P' 또는 'p'	여는 괄호인 "("를 나타냅니다.
'Q' 또는 'q'	닫는 괄호인 ")"를 나타냅니다.

참고: 열 패턴에는 특수 문자가 포함될 수도 있습니다. 예를 들어 ~, [,], =, -, ?, =, {, *, -, >, < 및 \$입니다.

값

값이 열에 나타나는 열 및 빈도에 대한 값을 볼 수 있습니다.

요약 보기에서 열의 최소값 및 최대값을 봅니다. 상세 보기에서 열에 대한 값 속성을 볼 수 있습니다.

요약 보기의 값

요약 보기에서 최근 프로필 실행, 기록 프로필 실행 및 통합된 프로필 실행의 모든 열 및 규칙에 대한 최소값 및 최대값을 볼 수 있습니다.

예

소매점 데이터베이스에는 100~250 범위의 직원 ID로 채워진 직원 테이블의 직원 ID로 이름이 지정된 열이 있으며 Bob, Robert와 같은 이름이 있습니다. 직원 테이블에서 열 프로필을 실행하면 요약 보기에서 직원 ID에 대한 값 열이 100 --> Robert를 표시합니다.

상세 보기의 값

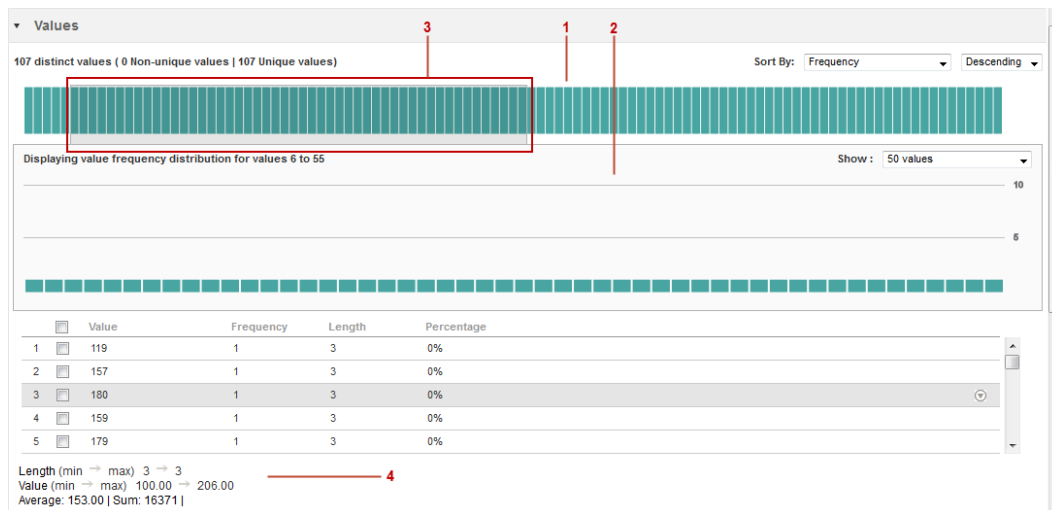
상세 보기의 열 값에는 값이 열에 나타나는 열 및 빈도에 대한 값이 포함됩니다.

값 창에는 열 값이 그래픽 표현으로 표시됩니다. 각 값의 빈도, 길이 및 백분율을 볼 수 있습니다. 값을 빈도를 기준으로 값을 정렬할 수 있습니다. 데이터를 드릴다운하거나, 참조 테이블에 값을 추가하거나, 값 빈도 규칙을 작성하거나 데이터 도메인을 작성할 수 있습니다. Null 값은 빨강색 세로 막대로 표시되고 값 빈도는 검정색 세로 막대로 표시되며 이상값은 주황색으로 강조 표시된 세로 막대로 표시됩니다. 이상값을 강조 표시하거나, 이상값을 비활성화하거나, 열에 이상값만 표시되도록 결과를 필터링할 수 있습니다.

값 창에는 그래픽 레이아웃 및 값 섹션이 포함됩니다.

그래픽 레이아웃은 2개의 패널로 나뉩니다.

다음 이미지는 상세 보기의 값 창을 보여 줍니다.



1. 상부 패널. 값을 세로 막대 차트로 볼 수 있습니다. 값을 빈도 및 값을 기준으로 정렬할 수 있습니다. 값을 오름차순 또는 내림차순으로 정렬할 수 있습니다. 이상값을 주황색으로 강조 표시된 세로 막대로 볼 수 있습니다.
2. 하부 패널. 하부 패널에서는 각 값이 세로 막대로 표시되는 슬라이더에서 값을 볼 수 있습니다. 값을 드릴다운하고, 참조 테이블에 값을 추가하고, 값 빈도 규칙을 작성하고, 값에 대해 데이터 도메인을 작성할 수 있습니다. 한 번에 50개, 75개 또는 100개의 값을 볼 수 있습니다.
3. 슬라이더. 상부 패널의 값 위에서 슬라이더를 이동할 수 있습니다. 그러면 슬라이더 내의 값이 하부 패널에 표시됩니다.
4. 값 속성. 값 속성 섹션에는 값과 속성이 표시됩니다.

다음 테이블에는 그래픽 레이아웃의 패널이 설명되어 있습니다.

패널	설명
상부 패널	모든 값을 세로 막대 차트로 표시합니다. 상부 패널에서 최대 16,000개의 값을 볼 수 있습니다. 슬라이더를 사용하여 일괄 값을 볼 수 있습니다.
하부 패널	상부 패널에서 선택한 일괄에 대한 값을 표시합니다. 기본적으로 Analyst 도구는 50개의 값을 표시합니다. 한 번에 75개 또는 100개의 값을 볼 수 있습니다.

다음 테이블에는 값 섹션의 열 값에 대한 속성이 설명되어 있습니다.

속성	설명
값	상부 패널에서 선택한 일괄에 대한 값 목록을 표시합니다. 참고: Analyst 도구는 열 값에서 CLOB, BLOB, Raw 및 이진 데이터 유형을 제외합니다.
빈도	값이 열에 나타나는 숫자로 표시된 횟수를 표시합니다.
길이	열 값의 길이를 표시합니다.
백분율	값이 열에 나타나는 백분율을 표시합니다.

다음 테이블에는 선택된 열에 대한 통계가 설명되어 있습니다.

통계	설명
길이(최소 - 최대)	열에 대한 가장 짧은 값과 가장 긴 값의 길이를 표시합니다.
값(최소 - 최대)	열의 최소값과 최대값을 표시합니다.
평균	열에 대한 값의 평균을 표시합니다.
합계	열의 모든 값의 합계를 표시합니다.

프로필 결과 비교를 위한 상세 보기의 값

프로필 결과 비교를 위한 상세 보기의 값 창은 고유 값 수, 최소값, 최대값, 최대 및 최소 길이, 평균, 표준 편차 및 값의 합계와 같은 값 속성을 표시합니다.

프로필 결과 비교를 위한 열의 상세 보기는 가로 막대 차트와 함께 값 속성, 값 및 값의 빈도를 표시합니다. 다음 테이블에서는 2개 프로필 실행의 결과를 비교하는 경우 상세 보기에서 열 값에 대한 속성을 설명합니다.:

속성	설명
고유 값 수	열의 고유 값 수를 표시합니다.
최소값	열의 최소값을 표시합니다.
최대값	열의 최대값을 표시합니다.
길이(최소 - 최대)	열에 대한 가장 짧은 값과 가장 긴 값의 길이를 표시합니다.

속성	설명
평균	열에 대한 값의 평균을 표시합니다.
표준 편차	열의 모든 값에 대한 열 값 간의 표준 편차 또는 가변성을 표시합니다.
합계	열의 모든 값의 합계를 표시합니다.

프로필 실행의 유형

최근 프로필 실행, 기록 프로필 실행 및 통합된 프로필 실행에 대한 프로필 결과를 볼 수 있습니다. 요약 보기에서 프로필 실행 결과를 볼 수 있습니다.

최근 프로필 실행

요약 보기에서 프로필의 최근 프로필 실행에 대한 프로필 결과를 봅니다.

다음과 같은 경우 요약 보기에서 최근 프로필 실행에 대한 프로필 결과를 볼 수 있습니다.

- 프로필을 작성, 저장 및 실행합니다.
- **라이브러리** 작업 공간에서 이전에 실행한 프로필을 엽니다.
- 통합된 프로필 실행에 대한 요약 보기 또는 상세 보기에서 **최근 프로필 실행으로 돌아가기** 링크를 클릭합니다.
- 기록 프로필 실행에 대한 요약 보기 또는 상세 보기에서 **최근 프로필 실행으로 돌아가기** 링크를 클릭합니다.
- **프로필 실행 선택** 대화 상자에서 최근 프로필 실행을 선택하고 **확인**을 클릭합니다.

기록 프로필 실행

요약 보기에서 이전 프로필 실행에 대한 프로필 결과를 봅니다.

프로파일링 웨어하우스는 프로필의 모든 프로필 실행의 프로필 결과를 저장합니다. 프로필 실행 선택 대화 상자에서 프로필 실행을 선택하여 프로필 실행의 이전 버전에서 결과를 볼 수 있습니다.

통합된 프로필 실행

요약 보기에서 프로필의 각 열에 대한 최근 프로필 결과를 봅니다.

통합된 프로필 실행에서 프로필의 각 열에 대한 최근 결과를 볼 수 있습니다. **프로필 실행 선택** 대화 상자에서 통합된 프로필 실행을 선택하면 프로파일링 웨어하우스가 프로필의 모든 프로필 실행에서 최근 열 결과를 검색합니다. 요약 보기에서 결과를 볼 수 있으며 요약 보기 헤더가 증분 프로필 실행을 표시합니다.

예제

데이터 분석가는 프로필의 각 열에 대한 최근 결과를 볼 수 있습니다. 예를 들어 프로필 실행 A를 수행하기 위한 열 1, 2 및 3을 선택하고 프로필 실행 B에 대한 열 3, 4 및 5를 선택할 수 있습니다. 모든 열에 대한 최근 결과를 보려면 프로필 실행 선택 대화 상자에서 통합된 프로필 실행을 선택합니다. 요약 보기는 실행 A에서 열 1 및 2에 대한 결과를 표시하고 실행 B에서 열 3, 4 및 5에 대한 결과를 표시합니다.

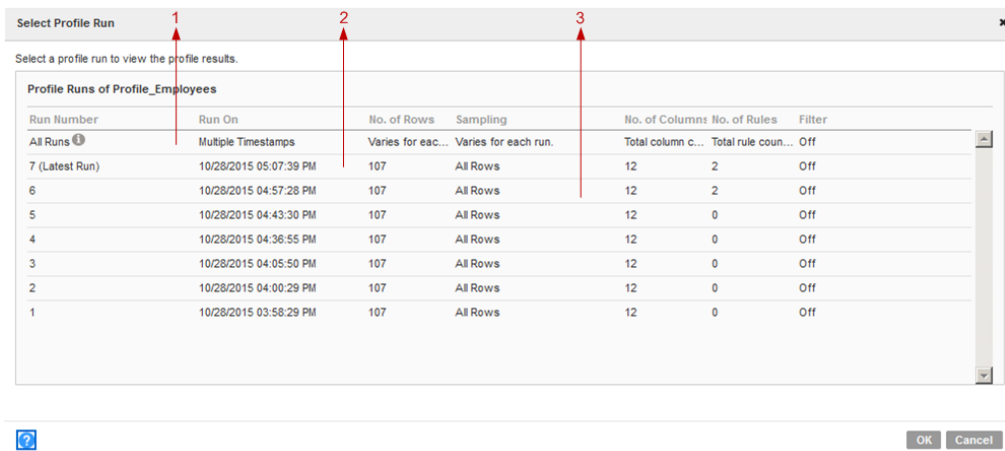
프로필 실행 선택

프로필 결과를 볼 기록 프로필 실행, 최근 프로필 실행 또는 통합된 프로필 실행을 선택할 수 있습니다. 요약 보기에서 프로필 결과를 보고 상세 보기에서 열 결과를 볼 수 있습니다.

1. 라이브러리 작업 공간에서 프로필을 포함하는 프로젝트 또는 폴더를 선택하거나 자산 창에서 프로필을 선택합니다.
2. 작업 > 열기를 클릭하여 프로필을 엽니다.
요약 보기가 검색 작업 공간에 표시됩니다.
3. 요약 보기에서 작업 > 프로필 실행 선택을 클릭합니다.

프로필 실행 선택 대화 상자가 표시됩니다.

다음 이미지는 프로필 실행 선택 대화 상자를 보여 줍니다.



1. 통합된 프로필 실행. 이 프로필 실행을 선택하면 요약 보기에서 각 열에 대한 최근 프로필 결과를 볼 수 있습니다.
2. 최근 프로필 실행. 이 프로필 실행을 선택하면 요약 보기에서 프로필에 대한 최근 프로필 결과를 볼 수 있습니다.
3. 기록 프로필 실행. 이 프로필 실행을 선택하면 요약 보기에서 이전 프로필 실행에 대한 기록 프로필 결과를 볼 수 있습니다.
4. 프로필 실행 선택 대화 상자에서 해당 프로필 결과를 볼 프로필 실행 중 하나를 선택합니다.
 - 최근 프로필 실행에 대한 프로필 결과를 보려면 최근 프로필 실행을 선택하고 **확인**을 클릭합니다.
 - 기록 프로필 실행에 대한 프로필 결과를 보려면 최근이 아닌 프로필 실행을 선택하고 **확인**을 클릭합니다.
 - 통합 프로필 실행에 대한 프로필 결과를 보려면 **모든 실행**을 선택하고 **확인**을 클릭합니다. 각 열에 대한 최근 프로필 결과가 요약 보기에 표시됩니다.

Analyst 도구는 프로필 실행을 수행하고 요약 보기에서 프로필 결과를 표시합니다.

5. 요약 보기에서 열을 클릭하여 열 결과를 봅니다.
상세 보기가 나타납니다.

여러 프로필 결과 비교 개요

2개의 프로필 실행에 대한 프로필 결과를 비교할 수 있습니다. 요약 보기에서 비교 결과를 보고 상세 보기에서 열 결과를 볼 수 있습니다.

요약 보기에서는 두 프로필 실행의 모든 열에 대한 비교 결과를 볼 수 있습니다.

여러 프로필 결과 비교

2개의 프로필 실행을 비교할 때 요약 보기에서 프로필 결과 비교를 볼 수 있습니다.

1. 요약 보기에서 **작업 > 프로필 실행 비교**를 클릭합니다.

다음 이미지는 **프로필 실행 비교** 대화 상자를 보여 줍니다.

The dialog box 'Compare Profile Runs' contains two sections for selecting profile runs. The first section, 'Select the profile runs that you want to compare.', shows a table with three runs. Run 3 is selected. The second section, 'Select another profile run', shows a table with three runs. Run 1 is selected.

Run Number	Run On	No. of Rows Pr	Sampling	No. of Columns	No. of Rules Pr	Filter
3 (Latest Run)	09/09/2015 02:47:10 PM	27	No Sampling	4	0	Off
2	09/09/2015 02:45:24 PM	27	No Sampling	4	0	Off
1	09/07/2015 01:45:48 PM	27	No Sampling	4	0	Off

Run Number	Run On	No. of Rows Pr	Sampling	No. of Columns	No. of Rules Pr	Filter
3 (Latest Run)	09/09/2015 02:47:10 PM	27	No Sampling	4	0	Off
2	09/09/2015 02:45:24 PM	27	No Sampling	4	0	Off
1	09/07/2015 01:45:48 PM	27	No Sampling	4	0	Off

1. 실행 A. 프로필 실행을 실행 A로 선택합니다.
2. 실행 B. 프로필 실행을 실행 B로 선택합니다.

프로필 실행 비교 대화 상자가 나타납니다.

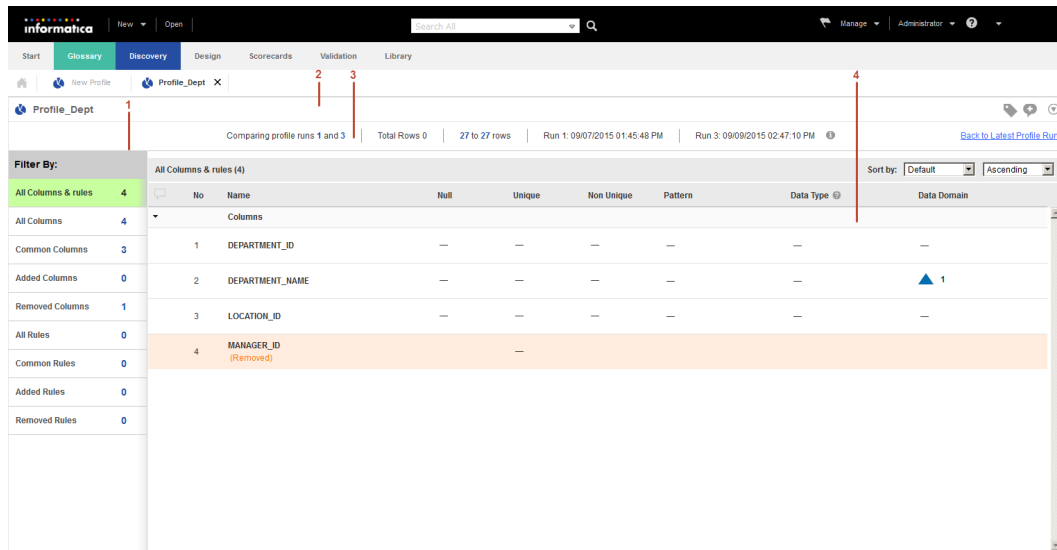
2. 실행 A 창에서 프로필을 선택하고 실행 B 창에서 다른 프로필을 선택합니다.
3. **확인**을 클릭합니다.

요약 보기는 프로필 결과의 통합 보기를 표시합니다.

프로필 결과 비교의 요약 보기

두 개의 프로필 실행을 비교할 경우 요약 보기에서 그리드 형식으로 결과를 볼 수 있습니다. 요약 보기에서 기본 필터를 사용하여 특정 통계를 볼 수 있습니다.

다음 이미지는 두 개의 프로필 실행에 대한 프로필 비교 결과를 요약 보기에서 보여 줍니다.



1. 기본 필터. 기본 필터를 기반으로 요약 보기에서 프로파일 비교 결과를 볼 수 있습니다.
2. 프로파일 헤더. 헤더에서 프로파일 이름을 볼 수 있습니다.
3. 요약 보기 헤더. 요약 보기 헤더에서 프로파일 관련 정보를 볼 수 있습니다. 비교할 프로파일 실행, 프로파일 실행 간 행의 증가 또는 감소, 프로파일 행의 수, 프로파일 실행의 시간과 날짜를 볼 수 있습니다.
4. 요약 보기. 두 프로파일 실행에서 열 사이의 비교를 볼 수 있습니다.

프로파일 결과 비교에 대한 요약 보기 속성

프로파일 결과 비교에 대한 요약 보기에는 고유한 값, 고유하지 않은 값 및 Null 값의 개수와 백분율, 패턴, 유추된 데이터 유형, 유추된 데이터 도메인 및 연결된 비즈니스 용어가 포함됩니다. 요약 보기에는 속성의 시각적 표현이 포함됩니다. 각 요약 속성을 클릭하여 속성의 값을 정렬할 수 있습니다.

요약 보기에서는 데이터 통합 서비스가 모든 열 및 규칙에 오름차순으로 번호를 할당합니다.

참고: 숫자 개수가 포함된 위쪽 화살표는 특정 프로파일 실행에서 다른 프로파일 실행으로 속성의 값 증가를 표시합니다. 숫자 개수가 포함된 아래쪽 화살표는 특정 프로파일 실행에서 다른 프로파일 실행으로 속성의 값 감소를 표시합니다.

다음 테이블에는 프로파일 결과 비교에 대한 요약 속성이 설명되어 있습니다.

속성	설명
아니요	열 또는 규칙의 번호를 표시합니다.
이름	프로파일의 열 또는 규칙의 이름을 표시합니다.
Null	null 값의 증가 또는 감소를 표시합니다.
고유	고유한 값의 증가 또는 감소를 표시합니다.
비고유	고유하지 않은 값의 증가 또는 감소를 표시합니다.
패턴	프로파일 실행 사이의 패턴의 변형을 표시합니다.

속성	설명
데이터 유형	2개의 프로필 실행의 열 또는 규칙에 대한 유추된 데이터 유형 사이의 변형을 표시합니다.
데이터 도메인	2개의 프로필 실행의 열 또는 규칙과 연결된 유추된 데이터 도메인 사이의 변형을 표시합니다.

요약 보기의 프로필 결과 비교에 대한 기본 필터

요약 보기의 기본 필터를 기준으로 프로필 결과를 볼 수 있습니다.

요약 보기에서 소스 열 및 가상 열을 볼 수 있습니다. 요약 보기에서 규칙에 대한 출력은 가상 열로 표시됩니다. 규칙에 대한 출력 포트를 변경하고 프로필 실행을 기록 실행과 비교하는 경우 기록 규칙 출력 열이 **제거된 규칙** 필터에 표시되고 새로운 규칙 출력 열이 **추가된 규칙** 필터에 표시됩니다. 프로필 실행에서 단일 출력 규칙에 대한 규칙 논리를 변경하거나 여러 규칙 출력에 대한 입력을 변경하고 기록 실행과 비교하는 경우 **추가된 규칙** 및 **제거된 규칙** 필터 출력이 변경되지 않습니다. 필터가 열에 대한 변경 내용만 필터에 대한 유효한 입력으로 간주하기 때문에 필터 출력은 변경되지 않습니다.

다음과 같은 기본 필터 옵션을 사용하여 특정 조건을 충족하는 프로필 결과를 볼 수 있습니다.

기본 필터 옵션	설명
모든 열 및 규칙	소스 열, 가상 열 및 규칙 열에 대한 프로필 결과를 표시합니다. 소스 열 및 규칙 열을 확장 및 축소하여 결과를 볼 수 있습니다.
모든 열	소스 열 및 가상 열에 대한 프로필 결과를 표시합니다.
공통 열	두 프로필 실행 결과에서 사용 가능한 열을 표시합니다.
추가된 열	최근 프로필 실행에서 사용 가능한 열을 표시합니다. 예를 들어 실행 5를 실행 3과 비교하는 경우 추가된 열은 실행 3이 아닌 실행 5에서 사용 가능한 열을 표시합니다.
제거된 열	기록 프로필 실행에서 사용 가능한 열을 표시합니다. 예를 들어 실행 5를 실행 3과 비교하는 경우 제거된 열은 실행 5가 아닌 실행 3에서 사용 가능한 열을 표시합니다.
모든 규칙	모든 규칙 열에 대한 프로필 결과를 표시합니다.
추가된 규칙	최근 프로필 실행에서 사용 가능한 규칙을 표시합니다. 예를 들어 실행 5를 실행 3과 비교하는 경우 추가된 규칙은 실행 3이 아닌 실행 5에서 사용 가능한 규칙을 표시합니다.
제거된 규칙	기록 프로필 실행에서 사용 가능한 규칙을 표시합니다. 예를 들어 실행 5를 실행 3과 비교하는 경우 제거된 규칙은 실행 5가 아닌 실행 3에서 사용 가능한 규칙을 표시합니다.

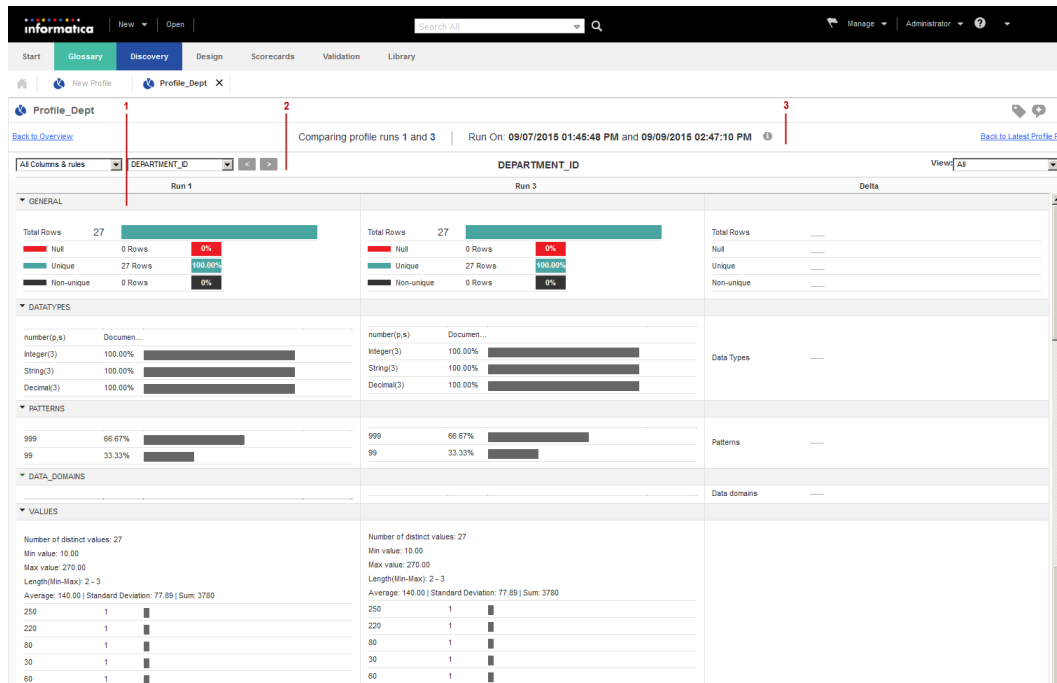
요약 보기는 기본적으로 모든 소스 열 및 가상 열에 대한 프로필 결과를 표시합니다.

프로필 결과 비교의 상세 보기

열 결과가 상세 보기에서 그리드 형식으로 나타납니다. 열 세부 정보에는 고유한 값, 고유하지 않은 값, Null 값, 패턴, 데이터 유형, 데이터 도메인, 비즈니스 용어, 값 및 데이터 미리보기와 같은 일반 정보가 포함됩니다.

열 이름을 클릭하면 열에 대한 상세 보기가 나타납니다. 실행 A 및 실행 B에서 개별 열로 열 결과를 볼 수 있으며 델타 열에서 데이터의 비교를 사용할 수 있습니다.

다음 이미지는 상세 보기에서 열에 대한 프로파일 결과 비교를 보여 줍니다.



1. 창. 2개의 프로파일 실행에서 프로파일 결과와 열에 대한 통계를 보고, 2개의 프로파일 실행에서 열에 대한 델타 정보를 창에서 볼 수 있습니다.
2. 프로파일 헤더. 드롭다운 목록에서 열을 선택하거나 탐색 단추를 사용하여 열 결과를 볼 수 있습니다. 열 이름을 보고, 보기 드롭다운 목록에서 옵션을 사용하여 특정 결과를 볼 수 있습니다.
3. 요약 보기 헤더. 요약 보기 헤더에서 프로파일 관련 정보를 볼 수 있습니다. 비교할 프로파일 실행을 보고 프로파일 실행의 시간과 날짜를 볼 수 있습니다.

프로파일 결과 비교에 대한 상세 보기 창

상세 보기는 2개 프로파일 실행의 열에 대한 프로파일 결과 및 비교 결과를 자세히 표시합니다.

상세 보기는 실행 A 및 실행 B에 대한 열 결과를 표시하며 델타 열에서 데이터의 비교를 사용할 수 있습니다. 다른 열 결과를 보려면 필터 드롭다운 목록에서 필터를 선택하거나 열 드롭다운 목록에서 열을 선택합니다.

열 프로파일 드릴다운

열 프로파일의 드릴다운 옵션을 사용하여 열 값을 기반으로 데이터 소스의 특정 행으로 드릴다운합니다. 드릴다운을 위해 데이터 소스의 현재 데이터를 읽을 수도 있고 프로파일 웨어하우스의 준비된 프로파일 데이터를 읽을 수도 있습니다. 준비된 프로파일 데이터의 특정 행으로 드릴다운하면 **Analyst** 도구가 일치하는 열 값에 대한 드릴다운 필터를 작성합니다. 드릴다운한 후에는 드릴다운 필터를 편집, 회수, 재설정 및 저장할 수 있습니다.

프로파일링을 위해 선택하지 않은 열도 드릴다운용으로 선택할 수 있습니다. 드릴다운을 위해 데이터 소스의 현재 데이터를 읽을 수도 있고 프로파일링 웨어하우스의 준비된 프로파일 데이터를 읽을 수도 있습니다. 열 값에 대한 드릴다운을 수행한 후에는 선택한 값이나 패턴에 대한 드릴다운 데이터를 선택한 위치에 **CSV** 파일로 내보낼 수 있습니다. **Informatica Analyst**는 드릴다운 데이터에 대해 처음 200개 값을 표시하지만 **CSV** 파일로는 모든 값을 내보냅니다.

행 데이터 드릴다운

프로필을 실행한 후에는 열 값, 데이터 유형 또는 패턴과 일치하는 특정 행으로 드릴다운할 수 있습니다.

1. 프로필을 실행합니다.
프로필 결과가 요약 보기에 나타납니다.
2. 요약 보기에서 열 이름을 클릭합니다.
열 결과가 상세 보기에 나타납니다.
3. 상세 보기에서 값 창의 값을 마우스 오른쪽 단추로 클릭하고 **드릴다운**을 선택합니다.
데이터 미리보기 창이 드릴다운 데이터를 표시합니다.

드릴다운 데이터에 필터 적용

드릴다운 데이터를 반복적으로 필터링하여 프로필 결과 하위 집합의 데이터 불규칙성을 분석할 수 있습니다.

1. 값 탭에서 열 값을 선택합니다.
2. 마우스 오른쪽 단추를 클릭하고 **드릴다운**을 선택합니다.
드릴다운 결과가 **데이터 미리보기** 창에 나타납니다.
3. 필터 조건을 추가하려면 **데이터 미리보기** 창에서 열 값을 마우스 오른쪽 단추로 클릭한 다음 **필터에 추가**를 선택합니다.
필터 조건이 포함된 **드릴다운 필터** 대화 상자가 표시됩니다.
4. 필요한 필터 조건을 추가하고 **확인**을 클릭합니다.
유추된 데이터 유형에는 드릴다운 필터를 적용할 수 없습니다.
5. 필터를 저장하려면 **작업 > 필터 저장**을 클릭합니다.
6. 드릴다운 필터를 지우려면 **작업 > 새로 고침**을 클릭합니다.
7. 드릴다운 데이터를 Microsoft Excel 스프레드시트로 내보내려면 **작업 > 데이터 내보내기**를 클릭합니다.

Analyst 도구의 기간

기간은 데이터 소스에서 검색된 메타데이터를 사용하고 보고할 수 있도록 해당 메타데이터의 유효성을 검사하고 관리하는 프로세스입니다. **Analyst** 도구에서 메타데이터 기간을 지정할 때는 프로필 결과의 유추된 데이터 유형 또는 데이터 도메인을 승인, 거부 및 재설정할 수 있습니다.

열에 대해 하나의 데이터 유형 및 하나의 데이터 도메인을 승인할 수 있습니다. 열에 대해 거부된 데이터 유형 또는 데이터 도메인을 숨길 수 있습니다. 유추된 데이터 유형 또는 데이터 도메인을 승인하거나 거부한 후에 데이터 유형 또는 데이터 도메인을 재설정하여 유추된 상태를 복원할 수 있습니다.

데이터 유형 및 데이터 도메인 승인

프로필 결과에는 데이터 소스의 각 열에 대한 유추된 데이터 유형 및 데이터 도메인이 포함됩니다. **Analyst** 도구의 각 열에 대해 단일 데이터 유형 및 단일 데이터 도메인을 선택 및 승인할 수 있습니다.

1. 프로필을 실행합니다.
프로필 결과가 요약 보기에 표시됩니다.

2. 요약 보기에서 열 이름을 클릭합니다.
열 결과가 상세 보기에 나타납니다.
3. 상세 보기에서 **데이터 유형** 창의 데이터 유형 또는 **데이터 도메인** 창의 데이터 도메인을 선택합니다.
4. **작업 > 승인**을 클릭합니다.
5. 데이터 유형 또는 데이터 도메인의 유추된 상태를 복원하려면 데이터 유형 또는 데이터 도메인을 선택하고 **작업 > 재설정**을 클릭합니다.

데이터 유형 및 데이터 도메인 거부

상세 보기에서 데이터 유형 또는 데이터 도메인을 거부할 수 있습니다. 거부된 데이터 유형 및 데이터 도메인을 표시하거나 숨길 수 있습니다.

1. 프로필을 실행합니다.
프로필 결과가 요약 보기에 표시됩니다.
2. 요약 보기에서 열 이름을 클릭합니다.
열 결과가 상세 보기에 나타납니다.
3. 상세 보기에서 **데이터 유형** 창의 데이터 유형 또는 **데이터 도메인** 창의 데이터 도메인을 선택합니다.
4. **작업 > 거부**를 클릭합니다.
Analyst 도구는 데이터 유형 목록에서 거부된 데이터 유형을 제거합니다.
5. 거부된 데이터 유형을 보려면 **작업 > 거부된 항목 표시**를 클릭합니다.

Informatica Analyst의 열 프로필 내보내기 파일

프로필 결과의 일부분을 선택하는지 아니면 전체 결과 요약을 선택하는지에 따라 **CSV** 파일 또는 **Microsoft Excel** 파일로 열 프로필 결과를 내보낼 수 있습니다.

선택한 값과 패턴에 대해 값 빈도, 패턴 빈도, 데이터 유형 또는 드릴다운 데이터를 **CSV** 파일로 내보낼 수 있습니다. 모든 열에 대한 프로파일링 결과 요약을 **Microsoft Excel** 파일로 내보낼 수 있습니다. 데이터 통합 서비스 권한인 **결과 드릴다운 및 내보내기**를 사용하여 프로필 결과를 내보내는 사람(사용자 또는 그룹)을 결정합니다.

CSV 파일의 프로필 내보내기 결과

값 빈도, 패턴 빈도, 데이터 유형 또는 드릴다운 데이터를 내보내 파일에서 데이터를 볼 수 있습니다. **Analyst** 도구는 정보를 **CSV** 파일에 저장합니다.

유추된 열 패턴을 내보내면 **Analyst** 도구가 열 패턴의 다른 형식을 내보냅니다. 예를 들어 유추된 열 패턴 X(5)를 내보내면 **Analyst** 도구는 열 패턴의 XXXXX 형식을 **CSV** 파일에 표시합니다.

Microsoft Excel의 프로필 내보내기 결과

전체 프로필 결과 요약을 내보내면 **Analyst** 도구가 정보를 **Microsoft Excel** 파일의 여러 워크시트에 저장합니다. **Analyst** 도구는 파일을 ".xlsx" 형식으로 저장합니다.

다음 테이블에는 내보내기 파일의 각 워크시트에 표시되는 정보에 대한 설명이 나와 있습니다.

탭	설명
열 프로필	프로필이 실행된 후 요약 보기에서 내보낸 요약 정보입니다. 열 이름, 규칙 이름, 고유한 값/Null 값의 수, 유추된 데이터 유형 및 마지막 프로필 실행 날짜/시간 등을 예로 들 수 있습니다.
값	열과 규칙의 값 및 각 열에 대해 값이 표시되는 빈도입니다.
패턴	프로필을 실행한 열과 규칙에 대한 값 패턴 및 패턴이 표시되는 빈도입니다.
데이터 유형	열에 대한 모든 데이터 유형, 각 데이터 유형의 빈도, 백분율 값과 데이터 유형의 상태(예: 유추됨, 승인됨, 거부됨)입니다.
통계	각 열과 규칙에 대한 통계입니다. 평균, 길이, 상위 값, 하위 값, 표준 편차 등이 있습니다.
속성	프로필 이름, 유형, 샘플링 정책, 행 수 등의 속성 보기 정보입니다.

Informatica Analyst에서 프로필 결과 내보내기

프로필의 결과를 ".csv" 또는 ".xlsx" 파일로 내보내 파일에서 데이터를 확인할 수 있습니다.

1. **라이브러리** 작업 공간에서 프로필을 포함하는 프로젝트 또는 폴더를 선택합니다.
2. 프로필을 클릭하여 엽니다.
프로필 결과가 요약 보기에 나타납니다.
3. 요약 보기에서 **작업 > 데이터 내보내기**를 클릭합니다.
파일로 데이터 내보내기 대화 상자가 표시됩니다.
4. **파일로 데이터 내보내기** 대화 상자에서 파일 이름을 입력합니다. 필요한 경우 기본 파일 이름을 사용합니다.
5. **모두(요약, 값, 패턴, 통계, 속성)** 또는 **데이터 도메인 검색 결과**를 선택하고 **코드 페이지**를 선택합니다. **확인**을 클릭합니다.
데이터가 **Microsoft Excel** 스프레드시트로 내보내집니다.
6. 요약 보기에서 열을 클릭합니다.
열 결과가 상세 보기에 나타납니다.
7. 상세 보기에서 **작업 > 데이터 내보내기**를 클릭합니다.
파일로 데이터 내보내기 대화 상자가 표시됩니다.
8. **파일로 데이터 내보내기** 대화 상자에서 파일 이름을 입력합니다. 필요한 경우 기본 파일 이름을 사용합니다.
9. 다음 옵션 중 하나를 선택합니다.
 - 모두(요약, 값, 패턴, 통계, 속성)
 - 선택한 열에 대한 값 빈도
 - 선택한 열에 대한 패턴 빈도
 - 선택한 열에 대한 데이터 유형
 - 선택한 값에 대한 드릴다운 데이터

- 선택한 패턴에 대한 드릴다운 데이터
 - 선택한 데이터 유형에 대한 드릴다운 데이터
10. 파일 형식을 입력합니다. 형식은 **모두** 옵션의 경우 **Excel**이고 나머지 옵션의 경우에는 **CSV**입니다. 필드 이름을 파일의 첫 번째 행으로 내보내도록 선택할 수 있습니다.
 11. 파일의 코드 페이지를 선택합니다.
 12. **확인**을 클릭합니다.
데이터가 파일로 내보내집니다.

제 8 장

Informatica Analyst의 성과 기록표

이 장에 포함된 항목:

- [Informatica Analyst의 성과 기록표 개요, 62](#)
- [Informatica Analyst 성과 기록표 프로세스, 63](#)
- [Informatica Analyst에서 성과 기록표 작성, 64](#)
- [기존 성과 기록표에 열 추가, 65](#)
- [성과 기록표 실행, 66](#)
- [성과 기록표 보기, 66](#)
- [성과 기록표 편집, 66](#)
- [메트릭, 67](#)
- [메트릭 그룹, 68](#)
- [열 드릴다운, 70](#)
- [경향 차트, 70](#)
- [Informatica Analyst의 성과 기록표 내보내기 파일, 74](#)
- [성과 기록표 알림, 75](#)
- [성과 기록표 연계, 78](#)

Informatica Analyst의 성과 기록표 개요

성과 기록표는 프로필에서 열에 대한 유효한 값을 그래픽으로 표현한 것입니다. 성과 기록표를 생성하여 라이브 데이터 또는 준비된 데이터를 드릴다운할 수 있습니다.

성과 기록표는 데이터 품질의 진행 상황을 측정하는 데 사용됩니다. 예를 들어 성과 기록표를 생성해 데이터 품질 규칙을 적용하기 전에 데이터 품질을 측정할 수 있습니다. 데이터 품질 규칙을 적용한 후에는 다른 성과 기록표를 생성하여 데이터 품질에 대한 규칙의 영향을 비교할 수 있습니다.

성과 기록표에는 열에 대한 값 빈도가 점수로 표시됩니다. 점수는 열의 유효한 값의 백분율을 반영합니다. 프로필을 실행한 후에는 프로필의 열을 메트릭으로 성과 기록표에 추가할 수 있습니다. 메트릭 그룹을 생성하여 관련 메트릭을 단일 항목으로 그룹화할 수 있습니다. 레코드의 열에 대해 허용되는 잘못된 데이터 범위를 지정하는 임계값을 정의하고 각 메트릭에 대해 메트릭 가중치를 할당할 수 있습니다. 성과 기록표를 실행하면 **Analyst** 도구가 각 메트릭 그룹에 대해 가중치 평균 값을 생성합니다. 데이터 품질을 추가로 평가하려는 경우 각 메트릭에 고정 비용 또는 가변 비용을 할당할 수도 있습니다. 성과 기록표를 실행할 때 **Analyst** 도구는 각 메트릭에 대해 잘못된 데이터 비용의 합계를 계산하여 총 비용을 표시합니다.

성과 기록표를 생성하거나 편집할 때 소스 데이터를 기반으로 성과 기록표 필터를 생성할 수 있습니다. 성과 기록표 필터를 사용하면 필터 조건을 기반으로 메트릭 점수를 다시 계산할 수 있습니다. 유효한 데이터 레코드 및 올바르지 않은 레코드를 식별하려는 경우 각 메트릭을 드릴다운할 수 있습니다. 경향 차트를 사용하여 메트릭 점수 및 메트릭의 잘못된 데이터 비용이 일정 기간 동안 변경되는 방식을 추적할 수 있습니다. 성과 기록표의 프로필 필터를 재사용할 수 있습니다.

Analyst 도구에서 버전 제어 시스템이 활성화된 경우 여러 버전의 성과 기록표를 생성하고 특정 성과 기록표에 대한 버전 기록을 볼 수 있습니다. 기본적으로 성과 기록표를 생성하고 나면 성과 기록표가 체크 아웃됩니다. 다른 사용자가 성과 기록표를 편집할 수 있도록 성과 기록표를 체크 인해야 합니다.

성과 기록표 작업 공간에서 성과 기록표 대시보드를 볼 수 있습니다. 성과 기록표 대시보드에서는 성과 기록표가 있는 데이터 개체, 프로젝트의 성과 기록표, 지난 여섯 달 동안의 성과 기록표 실행 경향 및 한 달 내의 모든 성과 기록표 실행에 대한 양호, 합격 및 불합격 메트릭의 집계를 볼 수 있습니다.

Informatica Analyst에서 성과 기록표에 대한 전자 메일 알림을 구성하고 관리할 수 있습니다. 전자 메일 알림은 전자 메일 서비스를 사용하여 관리합니다. 전자 메일 서비스는 **Informatica Administrator**에서 구성할 수 있는 시스템 서비스입니다.

Informatica Analyst 성과 기록표 프로세스

Developer tool 및 **Analyst** 도구에서 성과 기록표를 작성하고 편집할 수 있습니다. **Analyst** 도구에서 성과 기록표를 실행할 수 있습니다. 프로파일링 웨어하우스의 준비된 데이터 또는 데이터 개체의 현재 데이터에 대해 성과 기록표를 실행할 수 있습니다.

성과 기록표 작업 공간에서 성과 기록표를 볼 수 있습니다. 성과 기록표를 실행한 후에는 **성과 기록표** 패널에서 점수를 볼 수 있습니다. 데이터 개체를 선택하여 성과 기록표 내의 점수에서 데이터 개체를 탐색할 수 있습니다. **Analyst** 도구는 다른 탭의 데이터 개체를 엽니다.

성과 기록표로 작업할 때는 다음 태스크를 수행할 수 있습니다.

1. **Developer tool** 또는 **Analyst** 도구에서 성과 기록표를 작성하고 프로필에서 열을 추가합니다.
2. **Analyst** 도구에서 성과 기록표를 엽니다.
3. 프로필을 실행한 후에는 프로필 열을 메트릭으로 성과 기록표에 추가합니다.
4. 필요에 따라 소스 데이터를 기반으로 성과 기록표 필터를 작성합니다.
5. 필요에 따라 각 메트릭에 대해 올바르지 않은 데이터 비용을 구성합니다.
6. 성과 기록표를 실행하여 열에 대한 점수를 생성합니다.
7. 성과 기록표를 표시하여 레코드의 각 열에 대해 점수를 확인합니다.
8. 점수에 대한 열을 드릴다운합니다.
9. 성과 기록표를 편집합니다.
10. 성과 기록표의 각 메트릭에 대한 임계값을 설정합니다.
11. 그룹을 작성하여 성과 기록표의 관련 메트릭을 추가하거나 이동합니다.
12. 필요에 따라 그룹을 편집하거나 삭제합니다.
13. 각 점수의 점수 경향 차트를 확인하여 시간 경과에 따른 점수 변경 방식을 모니터링합니다.
14. 필요에 따라 각 메트릭의 비용 경향 차트를 확인하여 데이터 품질 값을 모니터링합니다.
15. 각 메트릭 또는 메트릭 그룹에 대한 성과 기록표 연계를 확인합니다.
16. 읽기 액세스 권한이 있는 성과 기록표에 대한 통합된 정보를 봅니다.

Informatica Analyst에서 성과 기록표 작성

성과 기록표를 작성하고 프로필의 열을 성과 기록표에 추가합니다. 먼저 프로필을 실행한 다음 성과 기록표에 열을 추가해야 합니다.

1. **라이브러리** 작업 공간에서 프로필을 포함하는 프로젝트 또는 폴더를 선택합니다.

2. 프로필을 클릭하여 엽니다.

프로필 결과가 **검색** 작업 공간의 요약 보기에 표시됩니다.

3. **작업 > 성과 기록표에 추가**를 클릭합니다.

성과 기록표에 추가 마법사가 표시됩니다.

4. **성과 기록표에 추가** 화면에서 새 성과 기록표를 작성할지 아니면 기존 성과 기록표를 편집하여 미리 정의된 성과 기록표에 열을 추가할지 선택할 수 있습니다. **새 성과 기록표** 옵션이 기본적으로 선택됩니다. **다음**을 클릭합니다.

5. **2/8단계** 화면에서 성과 기록표의 이름을 입력합니다. 필요한 경우 성과 기록표에 대한 설명을 입력할 수 있습니다. 성과 기록표를 저장할 프로젝트 및 폴더를 선택합니다. **다음**을 클릭합니다.

기본적으로 성과 기록표 마법사는 프로필에 정의된 열과 규칙을 선택합니다. 프로필에 포함되지 않은 열은 추가할 수 없습니다.

6. **3/8단계** 화면에서 성과 기록표에 매트릭스로 추가할 열과 규칙을 선택합니다. 필요에 따라 왼쪽 열 헤더의 확인란을 클릭하여 모든 열을 선택합니다. 필요에 따라 **열 이름**을 선택하여 열 이름을 정렬합니다. **다음**을 클릭합니다.

7. **4/8단계** 화면에서 매트릭에 필터를 추가할 수 있습니다.

프로필에 대해 작성한 필터를 매트릭에 적용하거나 새 필터를 작성할 수 있습니다. **메트릭 필터** 창에서 매트릭을 선택하고 **필터 관리** 아이콘을 클릭하여 **필터 편집: 열 이름** 대화 상자를 엽니다. **필터 편집: 열 이름** 대화 상자에서 다음 태스크 중 하나를 수행하도록 선택할 수 있습니다.

- 프로필에 대해 생성한 필터를 선택합니다. **다음**을 클릭합니다.
- 기존 필터를 선택합니다. 편집 아이콘을 클릭하여 **필터 편집** 대화 상자에서 필터를 편집합니다. **다음**을 클릭합니다.
- 더하기(+) 아이콘을 클릭하여 **새 필터** 대화 상자에서 필터를 작성합니다. **다음**을 클릭합니다.

필요에 따라 선택한 필터를 성과 기록표의 모든 매트릭에 적용하도록 선택할 수 있습니다.

필터가 **메트릭 필터** 창에 표시됩니다.

8. **4/8단계** 화면에서 **다음**을 클릭합니다.

9. **5/8단계** 화면의 **메트릭** 창에서 각 매트릭을 선택하고 다음 태스크를 수행합니다.

- 유효한 값을 구성합니다. **사용 중인 점수: 값** 창의 **사용 가능한 값** 창에서 값을 하나 이상 선택한 후 오른쪽 화살표 단추를 클릭하여 값을 **유효한 값** 창으로 이동합니다. 매트릭의 유효한 값 총 수는 **사용 가능한 값** 창 위쪽에 표시됩니다.
- 매트릭 임계값을 구성합니다. **메트릭 임계값** 창에서 **양호, 합격 및 불합격** 점수에 대해 임계값을 설정합니다.
- 올바르지 않은 데이터에 대한 비용을 구성합니다. 매트릭 비용에 상수 값을 할당하려면 **고정 비용**을 선택합니다. 숫자 열을 매트릭에 가변 비용으로 연결하려면 **가변 비용**을 선택하고 **열 선택**을 클릭하여 숫자 열을 선택합니다. 필요에 따라 **비용 단위 변경**을 클릭하여 비용의 단위를 변경합니다. 매트릭에 대해 올바르지 않은 데이터에 대한 비용을 구성하지 않으려면 **없음**을 선택합니다.

10. **다음**을 클릭합니다.

11. **6/8단계** 화면에서 매트릭스를 추가할 수 있는 매트릭 그룹을 선택하거나 새 매트릭 그룹을 작성할 수 있습니다. 새 매트릭 그룹을 작성하려면 그룹 아이콘을 클릭합니다. **다음**을 클릭합니다.

12. **7/8단계** 화면에서 그룹의 매트릭에 대한 가중치와 그룹의 임계값을 지정합니다.

13. **8/8단계** 화면에서 성과 기록표를 실행할 런타임 환경으로 **원시**, **Hadoop** 또는 **Databricks**을 선택합니다. Hadoop 런타임 환경에서 **Blaze** 또는 **Spark** 엔진을 선택할 수 있습니다. **Databricks**를 선택하는 경우 **찾아보기**를 클릭하여 프로필을 실행하기 위한 **Databricks** 연결을 선택합니다. **Blaze** 또는 **Spark** 엔진을 선택하는 경우 **찾아보기**를 클릭하여 프로필을 실행하기 위한 **Hadoop** 연결을 선택합니다.
14. **저장**을 클릭하여 성과 기록표를 저장하거나 **저장 및 실행**을 클릭하여 성과 기록표를 저장하고 실행합니다. 성과 기록표가 **성과 기록표** 작업 공간에 표시됩니다.

기존 성과 기록표에 열 추가

프로필을 실행한 후 기존 성과 기록표에 열을 추가할 수 있습니다.

1. 프로필을 클릭하여 엽니다.
프로필 결과가 요약 보기에 표시됩니다.
2. 열을 선택합니다. **작업 > 성과 기록표에 추가**를 클릭합니다.
성과 기록표에 추가 마법사가 표시됩니다.
참고: 성과 기록표에 열을 추가하기 전에 다음 규칙과 지침을 사용합니다.
 - 열 이름과 성과 기록표 이름이 일치하면 성과 기록표에 열을 추가할 수 없습니다.
 - 열 이름을 변경하더라도 같은 열을 성과 기록표에 두 번 추가할 수는 없습니다.
3. **기존 성과 기록표**를 선택하여 미리 정의된 성과 기록표에 열을 추가합니다. **다음**을 클릭합니다.
4. **2/7단계** 화면에서 열을 추가할 성과 기록표를 선택합니다. **다음**을 클릭합니다.
성과 기록표와 연결된 기존 메트릭 및 메트릭 그룹을 볼 수 있습니다.
5. **3/7단계** 화면에서 성과 기록표에 메트릭으로 추가할 열과 규칙을 선택합니다. 필요에 따라 왼쪽 열 헤더의 확인란을 클릭하여 모든 열을 선택합니다. **열 이름**을 선택하여 열 이름을 정렬합니다. **다음**을 클릭합니다.
6. **4/7단계** 화면에서는 메트릭스에 대한 필터를 작성할 수 있습니다. 또한 프로필에 대해 작성한 필터를 메트릭에 적용할 수도 있습니다.
7. **5/7단계** 화면에서는 다음 태스크를 수행할 수 있습니다.
 - **메트릭** 창에서 각 메트릭을 선택하고 다른 창에서 메트릭 값을 구성합니다.
 - **사용 중인 점수:** 값 창의 **사용 가능한 값** 창에서 여러 값을 선택하고 오른쪽 화살표 단추를 클릭하여 값을 **유효한 값** 창으로 이동합니다.
메트릭의 유효한 값 총 수는 **사용 가능한 값** 창 위쪽에 표시됩니다.
 - **메트릭 임계값:** 창에서 **양호**, **합격** 및 **불합격** 점수에 대해 임계값을 설정할 수 있습니다.
 - **올바르지 않은 데이터에 대한 비용**에서 다음을 할 수 있습니다.
 - 각 메트릭을 선택하고 메트릭에 대해 올바르게 않은 데이터에 대한 비용을 구성합니다.
 - **고정 비용** 옵션을 선택하여 메트릭 비용에 상수 값을 할당합니다. **비용 단위 변경**을 클릭하여 비용의 단위를 변경할 수 있습니다.
 - **가변 비용** 옵션을 선택하여 숫자 열을 메트릭에 가변 비용으로 연결합니다. **열 선택**을 클릭하여 숫자 열을 선택할 수 있습니다.
8. **다음**을 클릭합니다.
9. **6/7단계** 화면에서는 다음 태스크를 수행할 수 있습니다.
 - 메트릭을 추가할 메트릭 그룹을 선택합니다.

- **기본 - 메트릭** 창에서 기본 메트릭 가중치인 0을 두 번 클릭하여 값을 변경할 수 있습니다.
 - **메트릭 임계값:** 창에서 **양호, 합격 및 불합격** 점수에 대해 임계값을 설정할 수 있습니다.
10. 다음을 클릭합니다.
 11. **7/7단계** 화면에서 런타임 환경을 선택합니다.
 12. **저장**을 클릭하여 성과 기록표를 저장하거나 **저장 및 실행**을 클릭하여 성과 기록표를 저장하고 실행합니다.

성과 기록표 실행

성과 기록표를 실행하여 열에 대한 점수를 생성합니다.

1. **자산** 패널에서 실행할 성과 기록표를 선택합니다.
2. 성과 기록표를 클릭하여 엽니다.
성과 기록표가 **성과 기록표** 작업 공간에 표시됩니다.
3. **작업 > 성과 기록표 실행**을 클릭합니다.
4. **메트릭** 창에서 점수를 선택하고 **열** 창에서 드릴다운할 열을 선택합니다.
5. **드릴다운** 옵션에서 드릴다운할 데이터(라이브 데이터 또는 준비된 데이터)를 선택합니다.
최적의 성능을 위해 라이브 데이터를 드릴다운하십시오.
6. **실행**을 클릭합니다.

성과 기록표 보기

성과 기록표를 실행하여 각 메트릭에 대한 점수를 확인합니다. 성과 기록표에는 점수가 백분율과 표시줄로 표시됩니다. 유효하거나 올바르지 않은 데이터를 확인합니다. 메트릭 가중치, 메트릭 그룹 점수, 점수 경향 및 데이터 개체의 이름과 같은 성과 기록표 정보도 확인할 수 있습니다.

1. 성과 기록표를 실행하여 점수를 확인합니다.
2. 보려는 점수가 있는 메트릭을 선택합니다.
3. **작업 > 드릴다운**을 클릭하여 열에 대해 유효한 데이터의 행 또는 유효하지 않은 데이터의 행을 확인합니다.
Analyst 도구의 **드릴다운** 섹션에는 기본적으로 유효하지 않은 데이터 행이 표시됩니다.

성과 기록표 편집

성과 기록표의 각 메트릭에 대해 유효한 값을 편집합니다. 성과 기록표는 실행해야 편집할 수 있습니다.

1. **라이브러리** 작업 공간에서 **자산** 창에서 편집할 성과 기록표를 클릭합니다.
성과 기록표가 **성과 기록표** 작업 공간에 표시됩니다.
2. 버전 제어 시스템이 활성화된 경우 **작업 > 체크 아웃**을 클릭합니다.

3. 작업 > 편집 > 일반을 클릭합니다.
성과 기록표 편집 대화 상자가 표시됩니다.
4. 일반 탭에서 성과 기록표의 이름과 설명을 필요에 따라 편집할 수 있습니다.
5. 메트릭 탭을 클릭합니다.
6. 메트릭스 창에서 점수를 선택한 다음 **사용 중인 점수: 값** 창의 모든 값 목록에서 유효한 값을 구성합니다.
7. 메트릭 임계값 창에서 점수 임계값을 필요에 따라 변경할 수 있습니다.
8. 각 메트릭에 대해 올바르게 않은 데이터의 비용을 검토한 후 필요에 따라 변경합니다.
9. 성과 기록표 필터 탭을 클릭합니다.
10. 필터를 추가, 편집 또는 삭제할 수 있습니다.
11. 메트릭 그룹 탭을 클릭합니다.
12. 메트릭 그룹을 작성, 편집 또는 제거할 수 있습니다.
메트릭 그룹 탭에서 메트릭 가중치 및 메트릭 임계값을 편집할 수도 있습니다.
13. 알림 탭을 클릭합니다.
14. 성과 기록표 알림 설정을 필요에 따라 변경할 수 있습니다.
메트릭 및 메트릭 그룹에 대해 전역 설정과 사용자 지정 설정을 지정할 수 있습니다.
15. 런타임 환경으로 **원시**, **Blaze**, **Spark** 또는 **Databricks**를 선택합니다. 런타임 환경으로 **Blaze** 또는 **Spark**를 선택하는 경우 **찾아보기**를 클릭하여 Hadoop 연결을 선택합니다. **Databricks**를 선택하는 경우 **선택**을 클릭하고 **Databricks** 연결을 선택합니다.
16. **저장**을 클릭하여 성과 기록표의 변경 내용을 저장하거나 **저장 및 실행**을 클릭하여 변경 내용을 저장하고 성과 기록표를 실행합니다.
17. **체크 인**을 클릭합니다.

메트릭

메트릭은 성과 기록표에 포함되는 데이터 소스의 열 또는 규칙의 출력입니다. 성과 기록표를 작성할 때는 각 메트릭에 가중치를 할당할 수 있습니다. 메트릭 그룹을 작성하여 성과 기록표의 관련 메트릭을 집합으로 분류합니다.

메트릭 가중치

성과 기록표를 작성할 때는 각 메트릭에 가중치를 할당할 수 있습니다. 가중치의 기본값은 0입니다.

성과 기록표를 실행하면 **Analyst** 도구가 메트릭 점수와 각 메트릭에 할당된 가중치를 바탕으로 각 메트릭 그룹에 대한 가중치 평균을 계산합니다.

예를 들어 가중치 **W1**을 메트릭 **M1**에 할당하고 가중치 **W2**를 메트릭 **M2**에 할당합니다. **Analyst** 도구는 다음 수식을 사용하여 가중치 평균을 계산합니다.

$$(M1 \times W1 + M2 \times W2) / (W1 + W2)$$

데이터 품질의 값

소스 데이터의 데이터 품질 측정값은 조직의 데이터 자산을 관리하는 데 중요한 정보입니다. 성과 기록표에 표현된 메트릭의 올바르게 않은 데이터 비용을 통해 조직은 소스 데이터의 데이터 품질 모니터링 과정에서 값을 파생

시킬 수 있습니다. 데이터 분석가는 통화 단위 또는 사용자 지정 단위와 같은 값을 메트릭 및 메트릭 그룹에 연결할 수 있습니다. 그런 다음 성과 기록표를 실행하여 소스 데이터에서 올바르게 읽은 데이터의 총 비용을 확인할 수 있습니다.

비즈니스 요구 사항에 따라 메트릭에 대해 비용 단위를 정의할 수 있습니다. 성과 기록표를 작성하거나 편집할 때 각 메트릭에 대해 가변 비용이나 고정 비용을 구성할 수도 있습니다.

고정 비용

고정 비용은 성과 기록표의 메트릭에 할당할 수 있는 상수 값입니다. 미리 정의된 비용 단위를 선택하거나 비즈니스 요구 사항을 충족하는 사용자 지정 비용 단위를 작성할 수 있습니다.

가변 비용

가변 비용은 데이터 소스의 숫자 열 값을 기준으로 메트릭에 할당하는 값입니다. 데이터 통합 서비스는 비용에 할당하는 열이나 가상 열을 기준으로 메트릭에 대해 가변 비용을 계산합니다.

예

주택 담보 대출 담당자가 고객이 주택 담보 대출 불입액을 전송할 수 있도록 고객에게 지급장을 제공해야 합니다. 성과 기록표를 사용하면 지급장이 올바르게 배달되도록 고객 주소 정확도를 측정할 수 있습니다. 주소 정확도 메트릭의 월별 불입액 열에 대해 가변 비용을 설정할 수 있습니다. 성과 기록표를 실행하여 고객이 월별 불입액을 제때 불입하지 않는 경우 주택 담보 대출 조직에서 손해를 입는 총 비용을 계산합니다.

임계값 정의

성과 기록표의 각 점수에 대한 임계값을 설정할 수 있습니다. 임계값은 레코드의 열에 대해 허용되는 잘못된 데이터의 백분율 범위를 지정합니다. 데이터의 양호, 합격 또는 불합격 범위에 대한 임계값을 설정할 수 있습니다. 성과 기록표에 열을 추가하거나 성과 기록표를 편집할 때 각 열에 대해 임계값을 정의할 수 있습니다.

성과 기록표의 열에 대해 임계값을 정의할 때 다음과 같은 선행 조건 태스크 중 하나를 완료합니다.

- 프로필을 열고 **성과 기록표에 추가** 대화 상자에서 프로필의 열을 성과 기록표에 추가합니다.
 - 필요에 따라 **라이브러리** 작업 공간에서 성과 기록표를 클릭하고 **작업 > 편집**을 선택하여 **성과 기록표 편집** 대화 상자에서 성과 기록표를 편집합니다.
1. **성과 기록표에 추가** 대화 상자 또는 **성과 기록표 편집** 대화 상자의 **메트릭** 창에서 각 메트릭을 선택합니다.
 2. **메트릭 임계값** 창에서 불합격 범위의 상한과 양호 범위의 하한을 나타내는 임계값을 입력합니다.
임계값에 대해 10진수 자릿수를 최대 두 자리까지 설정할 수 있습니다.
 3. **다음** 또는 **저장**을 클릭합니다.

메트릭 그룹

메트릭 그룹을 작성하여 성과 기록표의 관련 점수를 집합으로 분류합니다. 기본적으로 **Analyst** 도구는 기본 메트릭 그룹의 모든 점수를 분류합니다.

메트릭 그룹을 작성한 후에는 점수를 기본 메트릭 그룹에서 다른 메트릭 그룹으로 이동할 수 있습니다. 기본 메트릭 그룹을 비롯한 메트릭 그룹을 편집하여 해당 이름과 설명을 변경할 수 있습니다. 더 이상 사용하지 않는 메트릭 그룹은 삭제할 수 있습니다. 기본 메트릭 그룹은 삭제할 수 없습니다.

메트릭 그룹 작성

메트릭 그룹을 작성하여 성과 기록표의 관련 점수를 그룹에 추가합니다.

1. 라이브러리 작업 공간에서 **자산** 창에서 편집할 성과 기록표를 클릭합니다.
성과 기록표가 **성과 기록표** 작업 공간에 표시됩니다.
2. **작업 > 편집**을 클릭합니다.
성과 기록표 편집 창이 표시됩니다.
3. **메트릭 그룹** 탭을 클릭합니다.
기본 그룹이 **메트릭 그룹** 패널에 표시되며 기본 그룹의 점수가 **메트릭** 패널에 표시됩니다.
4. **새 그룹** 아이콘을 클릭하여 메트릭 그룹을 작성합니다.
메트릭 그룹 대화 상자가 표시됩니다.
5. 이름 및 선택적 설명을 입력합니다.
6. **확인**을 클릭합니다.
7. **저장**을 클릭하여 성과 기록표의 변경 내용을 저장합니다.

메트릭 그룹으로 점수 이동

메트릭 그룹을 작성한 후에는 관련 점수를 메트릭 그룹으로 이동할 수 있습니다.

1. 라이브러리 작업 공간에서 **자산** 창에서 편집할 성과 기록표를 클릭합니다.
성과 기록표가 **성과 기록표** 작업 공간에 표시됩니다.
2. **작업 > 편집**을 클릭합니다.
성과 기록표 편집 창이 표시됩니다.
3. **메트릭 그룹** 탭을 클릭합니다.
기본 그룹이 **메트릭 그룹** 패널에 표시되며 기본 그룹의 점수가 **메트릭** 패널에 표시됩니다.
4. **메트릭** 패널에서 메트릭을 선택하고 **메트릭 이동** 아이콘을 클릭합니다.
메트릭 이동 대화 상자가 표시됩니다.
참고: 여러 개의 점수를 선택하려면 Shift 키를 길게 누릅니다.
5. 점수를 이동시킬 대상 메트릭 그룹을 선택합니다.
6. **확인**을 클릭합니다.

메트릭 그룹 편집

메트릭 그룹을 편집하여 이름과 설명을 변경합니다. 기본 메트릭 그룹의 이름을 변경할 수 있습니다.

1. 라이브러리 작업 공간에서 **자산** 창에서 편집할 성과 기록표를 클릭합니다.
성과 기록표가 **성과 기록표** 작업 공간에 표시됩니다.
2. **작업 > 편집**을 클릭합니다.
성과 기록표 편집 창이 표시됩니다.
3. **메트릭 그룹** 탭을 클릭합니다.
기본 메트릭 그룹이 **메트릭 그룹** 패널에 표시되며 기본 메트릭 그룹의 메트릭이 **메트릭** 패널에 표시됩니다.
4. **메트릭 그룹** 패널에서 **그룹 편집** 아이콘을 클릭합니다.

편집 대화 상자가 표시됩니다.

- 이름 및 선택적 설명을 입력합니다.
- 확인**을 클릭합니다.

메트릭 그룹 삭제

더 이상 유효하지 않은 메트릭 그룹은 삭제할 수 있습니다. 메트릭 그룹을 삭제할 때는 메트릭 그룹의 점수를 기본 메트릭 그룹으로 이동하도록 선택할 수 있습니다. 기본 메트릭 그룹은 삭제할 수 없습니다.

- 라이브러리** 작업 공간에서 **자산** 창에서 편집할 성과 기록표를 클릭합니다.
성과 기록표가 **성과 기록표** 작업 공간에 표시됩니다.
- 작업 > 편집**을 클릭합니다.
성과 기록표 편집 창이 표시됩니다.
- 메트릭 그룹** 탭을 클릭합니다.
기본 메트릭 그룹이 **메트릭 그룹** 패널에 표시되며 기본 메트릭 그룹의 메트릭이 **메트릭** 패널에 표시됩니다.
- 메트릭 그룹** 패널에서 메트릭 그룹을 선택하고 **그룹 삭제** 아이콘을 클릭합니다.
그룹 삭제 대화 상자가 표시됩니다.
- 메트릭 그룹을 삭제하기 전에 메트릭 그룹에서 메트릭을 삭제하는 옵션이나 메트릭을 기본 메트릭 그룹으로 이동하는 옵션을 선택합니다.
- 확인**을 클릭합니다.

열 드릴다운

점수의 열을 드릴다운하여 올바른 데이터 행 또는 올바르지 않은 데이터 행을 볼 때 표시되는 열을 선택합니다. 드릴다운하도록 선택한 열이 **드릴다운** 패널에 표시됩니다.

- 성과 기록표를 실행하여 점수를 확인합니다.
- 보려는 점수가 있는 열을 선택합니다.
- 작업 > 드릴다운**을 클릭하여 열에 대해 유효한 데이터의 행 또는 올바르지 않은 데이터의 행을 확인합니다.
- 작업 > 열 드릴다운**을 클릭합니다.
선택한 점수에 대해 열이 **드릴다운** 패널에 표시됩니다. **Analyst** 도구에는 기본적으로 열에 대해 유효한 데이터 행이 표시됩니다. 필요에 따라 **잘못됨**을 클릭하여 유효하지 않은 데이터 행을 확인합니다.

경향 차트

경향 차트를 사용하여 메트릭 점수 및 메트릭의 올바르지 않은 데이터 비용이 일정 기간 동안 변경되는 방식을 모니터링합니다.

경향 차트에는 수평 축의 모든 성과 기록표 실행을 기준으로 수직 축에 점수 또는 비용 값을 그리는 점수 및 비용 그래프가 모두 포함됩니다. 기본적으로 경향 차트에는 지난 **10개** 성과 기록표 실행의 데이터가 표시됩니다. 경향 차트의 메트릭에 대해 전체 행과 올바르지 않은 행의 수를 확인할 수 있습니다. 경향 차트에는 지난 성과 기록표

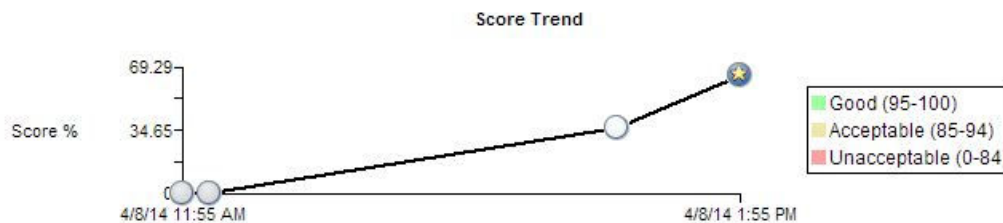
실행을 기준으로 하여 점수 및 비용 경향이 일정하게 유지되었는지 아니면 위나 아래로 이동되었는지도 표시됩니다.

Analyst 도구는 각 날짜의 기록 성과 기록표 실행 데이터와 최신 유효 점수 값을 사용하여 점수를 계산합니다. **Analyst** 도구는 차트의 최신 임계값 설정을 사용하여 점수 지점의 색상을 표시합니다. 점수에 대해 양호, 합격 및 불합격 임계값을 확인할 수 있습니다. 성과 기록표의 점수 값을 편집한 후에 성과 기록표를 실행할 때마다 이 임계값이 변경됩니다. 성과 기록표를 내보낼 때 **Analyst** 도구는 내보낸 파일의 점수 및 비용 정보를 비롯한 경향 차트 정보를 포함합니다.

점수 경향 차트

점수 경향 차트는 여러 프로파일 실행에서 메트릭 점수가 변경되는 방식을 그래픽으로 표현한 것입니다. 점수 경향 차트에는 수평 축의 모든 성과 기록표 실행을 기준으로 수직 축에 메트릭 점수 값이 그려집니다.

다음 이미지는 점수 경향 차트 샘플을 보여 줍니다.



예제

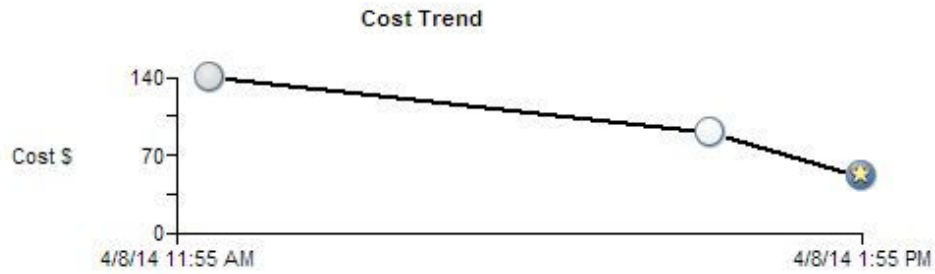
데이터 분석가는 데이터 품질을 모니터링하여 매핑 및 기타 프로세스 변경 내용으로 인해 데이터 품질 점수가 높아지는지 여부를 분석할 수 있습니다. 데이터 품질 변경을 측정 한 후에는 조직에서 분석하고 사용할 수 있도록 데이터 품질 변경을 다시 보고할 수 있습니다. 예를 들어 여러 성과 기록표 실행이 끝나면 주민등록번호 열의 유효한 값 백분율이 84에서 90으로 변경될 수 있습니다. 빠른 분석을 위해 이 데이터 품질 변경을 시각적 차트로 보고할 수 있습니다.

비용 경향 차트

비용 경향 차트는 여러 프로파일 실행에서 메트릭의 올라르지 않은 데이터 비용이 변경되는 방식을 그래픽으로 표현한 것입니다. 비용 경향 차트에서는 조직의 데이터 품질 영향을 측정할 수 있습니다. 비용 경향 차트에는 수평 축의 모든 성과 기록표 실행을 기준으로 수직 축에 비용 값이 그려집니다. 비용 경향 차트 아래의 그리드에서 올라르지 않은 데이터의 총 비용과 메트릭의 유효한 값도 확인할 수 있습니다.

비용 경향 차트를 통해 가치가 높은 레코드에 대한 올라르지 않은 데이터의 영향을 추적할 수 있습니다. 고정 비용을 사용하여 올라르지 않은 데이터를 계산할 때 가치가 높은 레코드에 대한 올라르지 않은 데이터의 영향을 간과하는 경우가 있습니다. 경향 차트에서 여러 성과 기록표 실행에 대해 점수는 높아졌는데 전체 비용은 낮아진 것으로 표시되는 경우 이 문제가 발생합니다. 그러나 가치가 높은 레코드에서는 성과 기록표에 표현되는 소수의 데이터 품질 문제가 있을 수 있습니다.

다음 이미지는 비용 경향 차트 샘플을 보여 줍니다.



예제

금융 기관에 미화 1000만 달러와 같이 상당한 은행 잔고와 투자 금액을 갖춘 고객이 여러 명 있습니다. 또한 잔고가 적은 고객도 많습니다. 점수 경향 차트에는 일정 기간 동안의 점수 상승이 표시될 수 있습니다. 그러나 잔고가 많은 소수 고객 계정의 주소나 성별이 잘못된 경우 조직에서 가장 중요한 고객과의 관계에 영향을 줄 수 있습니다. 올바르게 표시되지 않은 데이터 계산을 위한 가변 비용 열로 계좌 잔고 열을 설정할 수 있습니다. 열 값으로 인해 올바르게 표시되지 않은 데이터의 비용이 높으면 전체 값이 위험한 것으로 간주하여 수정 작업을 즉시 수행할 수 있습니다.

경향 차트 보기

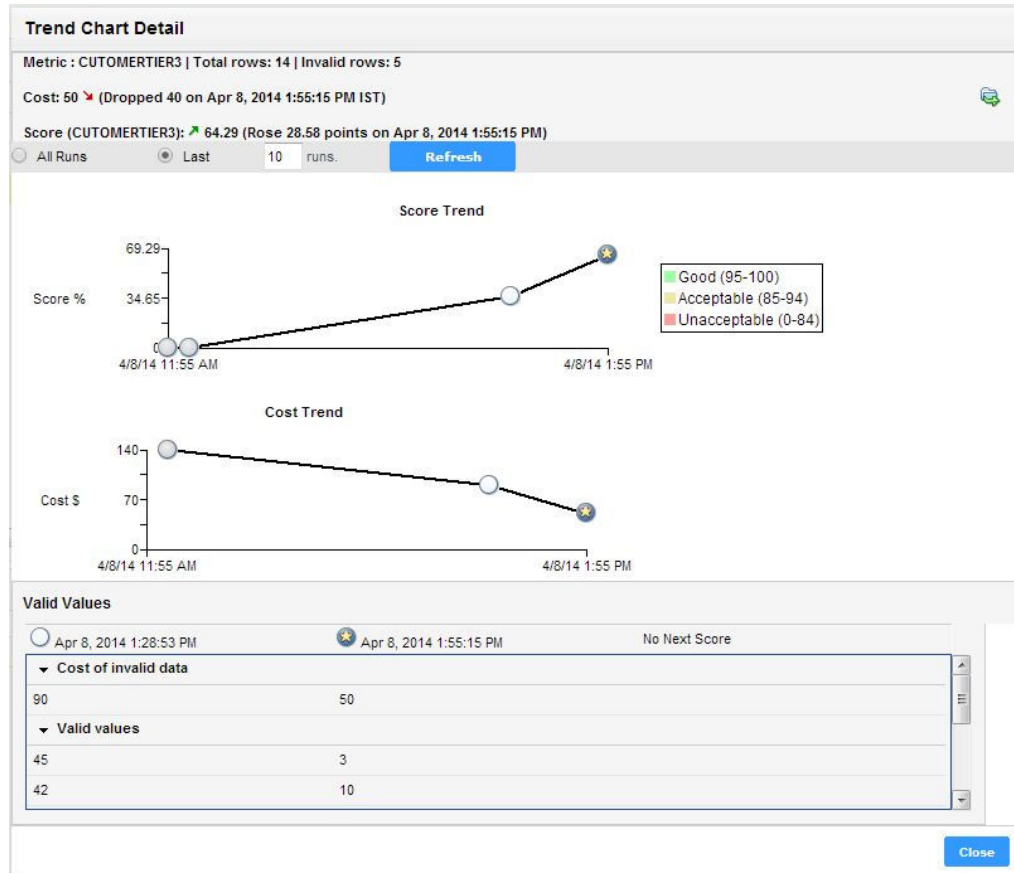
각 메트릭의 경향 차트를 확인하여 올바르게 표시되지 않은 데이터의 점수나 비용이 시간이 지남에 따라 변경되는 방식을 모니터링할 수 있습니다.

1. **라이브러리** 작업 공간에서 성과 기록표를 포함하는 프로젝트 또는 폴더를 선택합니다.
2. 성과 기록표를 클릭하여 엽니다.
성과 기록표가 **성과 기록표** 작업 공간에 표시됩니다.
3. **성과 기록표** 보기에서 메트릭을 선택합니다.

4. 작업 > 경향 차트 표시를 클릭합니다.

경향 차트 세부 정보 대화 상자가 표시됩니다.

다음 이미지는 경향 차트 세부 정보 대화 상자를 보여 줍니다.



시간의 경과에 따라 변경된 점수 및 비용 값을 볼 수 있습니다. 대화 상자 위쪽에서 행의 총 수와 올바르게 않은 행의 수를 확인할 수 있습니다. Analyst 도구는 각 날짜의 기록 성과 기록표 실행 데이터와 최신 유효 점수 값을 사용하여 점수를 계산합니다. 점수 및 비용 경향 차트에서 메트릭에 대한 유효한 값과 올바르게 않은 데이터의 비용을 확인할 수 있습니다.

경향 차트 내보내기

점수 및 비용 경향 차트를 ".xlsx" 파일로 내보내 파일에서 데이터를 확인할 수 있습니다.

- 성과 기록표를 엽니다.
- 메트릭을 선택하고 작업 > 경향 차트 표시를 클릭합니다.
경향 차트 세부 정보 대화 상자가 나타납니다.
- 데이터 내보내기 아이콘을 클릭합니다.
파일로 데이터 내보내기 대화 상자가 표시됩니다.
- 파일 이름 필드에 파일 이름을 입력합니다. 필요한 경우 기본 파일 이름을 사용할 수 있습니다.
기본 파일 형식은 Microsoft Excel입니다.
- 내보내려는 올바르게 않은 행 수를 올바르게 않은 행 수 필드에 입력합니다. 최대 100,000의 값을 필드에 입력할 수 있습니다.

6. **코드 페이지** 필드에서 파일의 코드 페이지를 선택할 수 있습니다.
7. **확인**을 클릭합니다.

Informatica Analyst의 성과 기록표 내보내기 파일

성과 기록표 결과를 Microsoft Excel 파일로 내보낼 수 있습니다. Analyst 도구는 파일을 XLSX 형식으로 저장합니다. 성과 기록표를 내보내면 Microsoft Excel 파일의 여러 워크시트에 성과 기록표 요약, 경향 차트, 올바르게 않은 행 수 및 성과 기록표 속성이 표시됩니다.

성과 기록표를 내보내면 다음과 같은 옵션을 구성할 수 있습니다.

데이터

다음 옵션 중 하나를 선택하여 데이터를 내보냅니다.

- 모두. 성과 기록표 요약, 경향 차트, 올바르게 않은 행 수 및 성과 기록표 속성을 Microsoft Excel 파일로 내보냅니다.
- 요약 보기. 성과 기록표 요약, 경향 차트 및 성과 기록표 속성을 Microsoft Excel 파일로 내보냅니다.

올바르지 않은 행 수

내보내려는 올바르게 않은 행 수를 입력합니다. 최대 100,000의 값을 필드에 입력할 수 있습니다. 기본적으로 필드에는 100이 표시됩니다. Analyst 도구는 각 메트릭에 대해 최대 100개의 행을 워크시트로 내보냅니다.

메트릭에 대해 100개를 초과하는 올바르게 않은 행 수를 내보내도록 선택하는 경우 데이터 통합 서비스는 다음 단계를 수행하여 나머지 데이터를 내보냅니다.

1. <INFA_HOME>/tomcat/bin/reject 위치에 성과 기록표에 대한 폴더를 생성합니다. 데이터 통합 서비스는 성과 기록표 이름 및 파일 생성 시간에 따라 <scorecard_name>_HH_MM_SS 형식으로 폴더 이름을 지정합니다. 예를 들어 성과 기록표 이름이 SD1이고 파일 생성 시간이 2:23:15인 경우 데이터 통합 서비스는 폴더 이름을 SD1_02_23_15로 지정합니다.
2. 성과 기록표의 각 메트릭에 대한 하위 폴더를 생성합니다. 예를 들어 성과 기록표 SD1에 M1, M2 및 M3이라는 이름의 메트릭이 있는 경우 데이터 통합 서비스는 M1, M2 및 M3이라는 이름의 하위 폴더 3개를 생성합니다.
3. 메트릭 하위 폴더에 Microsoft Excel 파일을 생성합니다. 파일 이름은 메트릭 이름에 따라 증분 숫자를 사용하여 <metric_name>_<IncrementalNumberStartingWith0> 형식으로 지정됩니다. 마지막으로 생성된 파일에는 <metric_name>_Remaining 형식이 사용됩니다. 각 Excel 파일에는 최대 10,000개의 올바르게 않은 행이 포함될 수 있습니다. 예를 들어 메트릭 M1에 30,000개의 올바르게 않은 행이 있는 경우 데이터 통합 서비스는 M1_0, M1_1 및 M1_Remaining이라는 이름의 Microsoft Excel 파일 3개를 생성하고 각 파일에 10,000개의 올바르게 않은 행을 저장합니다.
4. 데이터 통합 서비스는 성과 기록표의 다른 모든 메트릭에 대해 3단계를 반복합니다.

코드 페이지

코드 페이지를 선택하십시오.

Microsoft Excel의 성과 기록표 내보내기 결과

성과 기록표 결과를 내보내면 **Analyst** 도구가 정보를 **Microsoft Excel** 파일의 여러 워크시트에 저장합니다. 성과 기록표 요약, 경향 차트, 올바르지 않은 행 및 성과 기록표 속성이 파일의 워크시트로 표시됩니다. **Analyst** 도구는 파일을 "xlsx" 형식으로 저장합니다.

다음 테이블에는 내보내기 파일의 각 워크시트에 표시되는 정보에 대한 설명이 나와 있습니다.

탭	설명
성과 기록표 요약	내보낸 성과 기록표 결과의 요약 정보입니다. 이러한 정보로는 성과 기록표 이름, 각 열의 전체 행 수, 유효하지 않은 행 수, 점수 및 메트릭 가중치가 포함됩니다.
경향 차트	점수의 경향 차트입니다.
올바르지 않은 행 수	각 열에 대해 유효하지 않은 행의 세부 정보입니다. Analyst 도구는 최대 100개의 행을 워크시트로 내보냅니다. 올바르지 않은 행 워크시트는 파일로 데이터 내보내기 대화 상자에서 데이터 > 모두 옵션을 선택한 경우에 표시됩니다.
속성	이름, 유형, 설명, 위치 등의 성과 기록표 속성입니다.

Informatica Analyst에서 성과 기록표 결과 내보내기

성과 기록표 결과를 ".xlsx" 파일로 내보내 파일에서 데이터를 확인할 수 있습니다.

1. 성과 기록표를 엽니다.
2. **작업 > 데이터 내보내기**를 클릭합니다.
파일로 데이터 내보내기 대화 상자가 표시됩니다.
3. 파일 이름을 입력하십시오. 필요한 경우 기본 파일 이름을 사용합니다.
기본 파일 형식은 **Microsoft Excel**입니다.
4. 파일의 코드 페이지를 선택합니다.
5. **확인**을 클릭합니다.

성과 기록표 알림

특정 메트릭 점수, 메트릭 그룹 점수 또는 메트릭 비용이 여러 임계값 간을 이동할 때 **Analyst** 도구가 전자 메일을 보내도록 성과 기록표 알림 설정을 구성합니다. 메트릭 점수 또는 메트릭 그룹 점수는 여러 임계값 간을 이동할 수도 있고 불합격, 합격 및 양호와 같은 특정 점수 범위에서 유지될 수도 있습니다. 메트릭 비용 값은 설정된 상한 및 하한 비용 임계값을 초과하여 이동할 수 있습니다.

개별 메트릭 점수, 메트릭 그룹 및 메트릭 비용에 대해 전자 메일 알림을 구성할 수 있습니다. 점수에 대해 전역 설정을 사용하는 경우 **Analyst** 도구는 특정 메트릭 점수가 임계값을 넘어 양호 점수 범위에서 합격으로/합격 점수 범위에서 잘못됨으로 변경되면 전자 메일 알림을 보냅니다. 성과 기록표를 연속 실행했는데 점수가 불합격 점수 범위에서 유지되는 경우에도 각 성과 기록표 실행에 대해 전자 메일 알림을 받습니다. 메트릭 비용에 대해 전역 설정을 사용하는 경우 **Analyst** 도구는 선택한 메트릭의 올바르지 않은 데이터 비용이 상한 및 하한 임계값을 초과하면 전자 메일 알림을 보냅니다.

점수가 불합격에서 합격 점수 범위로/합격에서 양호 점수 범위로 이동할 때 성과 기록표 사용자가 전자 메일 알림을 받도록 알림 설정을 사용자 지정할 수 있습니다. 모든 성과 기록표 실행에 대해 메트릭 점수 또는 메트릭 비용이 특정 범위 내에서 유지되면 전자 메일 알림을 보내도록 선택할 수 있습니다. 비용 임계값 설정 기준으로 사용할 수 있는 알림 설정의 각 메트릭에 대해 올바르게 않은 데이터의 현재 비용을 확인할 수 있습니다.

전자 메일 알림을 보내도록 성과 기록표를 구성하려면 먼저 관리자가 **Administrator** 도구에서 전자 메일 서비스를 구성해야 합니다.

전자 메일 알림 메시지 템플릿

Analyst 도구가 성과 기록표 알림의 일부분으로 수신자에게 보내는 전자 메일 메시지의 구조와 메시지 텍스트를 설정할 수 있습니다. 전자 메일 템플릿에는 선택적인 소개 텍스트 섹션, 읽기 전용 메시지 본문 섹션 및 선택적인 마무리 텍스트 섹션이 있습니다.

다음 테이블에서는 전자 메일 템플릿의 태그에 대해 설명합니다.

태그	설명
ScorecardName	성과 기록표의 이름입니다.
ObjectURL	성과 기록표에 대한 링크입니다. 사용자 이름과 암호를 입력해야 합니다.
MetricGroupName	메트릭이 속하는 메트릭 그룹의 이름입니다.
CurrentWeightedAverage	현재 성과 기록표 실행의 메트릭 그룹에 대한 가중치 평균 값입니다.
CurrentRange	현재 성과 기록표 실행의 메트릭 그룹에 대한 양호/합격/불합격 등의 점수 범위입니다.
PreviousWeightedAverage	이전 성과 기록표 실행의 메트릭 그룹에 대한 가중치 평균 값입니다.
PreviousRange	이전 성과 기록표 실행의 메트릭 그룹에 대한 양호/합격/불합격 등의 점수 범위입니다.
MetricName	메트릭의 이름입니다.
MetricGroupName	메트릭 그룹의 이름입니다.
CurrentScore	최신 성과 기록표 실행을 기준으로 한 점수입니다.
CurrentRange	최신 성과 기록표 실행을 기준으로 하여 현재 점수가 유지되는 점수 범위입니다.
PreviousScore	이전 성과 기록표 실행을 기준으로 한 점수입니다.
PreviousRange	이전 성과 기록표 실행을 기준으로 한 점수 범위입니다.
CurrentCost	최신 성과 기록표 실행을 기준으로 한 메트릭의 올바르게 않은 데이터 비용입니다.
PreviousCost	이전 성과 기록표 실행을 기준으로 한 메트릭의 올바르게 않은 데이터 비용입니다.
ColumnName	메트릭이 할당된 소스 열의 이름입니다.
ColumnType	소스 열의 유형입니다.

태그	설명
RuleName	규칙의 이름입니다.
RuleType	규칙의 유형입니다.
DataObjectName	소스 데이터 개체의 이름입니다.

성과 기록표 알림 설정

메트릭 및 메트릭 그룹 수준 둘 다에서 성과 기록표 알림을 설정할 수 있습니다. 전역 알림 설정은 개별 알림 설정이 없는 메트릭 및 메트릭 그룹에 적용됩니다.

1. **Analyst** 도구에서 성과 기록표를 실행합니다.
2. **작업 > 편집**을 클릭합니다.
3. **알림** 탭을 클릭합니다.
4. **알림 활성화**를 선택하여 성과 기록표 알림 구성을 시작합니다.
5. 메트릭 또는 메트릭 그룹을 선택합니다.
6. **알림** 확인란을 클릭하여 메트릭 또는 메트릭 그룹에 대한 전역 설정을 활성화합니다.
7. **사용자 지정 설정 사용**을 선택하여 메트릭 또는 메트릭 그룹에 대한 설정을 변경합니다.

점수가 **불합격**, **합격** 및 **양호** 범위이며 임계값 간을 이동할 때 전자 메일 알림을 보내도록 선택할 수 있습니다. 메트릭 비용이 상한 또는 하한 임계값을 초과할 때도 전자 메일 알림을 보낼 수 있습니다.

8. 성과 기록표 알림에 대한 전역 설정을 편집하려면 **전역 설정 편집** 아이콘을 클릭합니다.
전자 메일 템플릿을 비롯한 설정을 편집할 수 있는 **전역 설정 편집** 대화 상자가 표시됩니다.

성과 기록표 알림에 대한 전역 설정 구성

전역 성과 기록표 알림 설정을 선택하면 점수가 **불합격** 범위일 때 **Analyst** 도구가 대상 사용자에게 전자 메일을 보냅니다. 메트릭 점수 또는 메트릭 비용이 임계값 간을 이동할 때 전자 메일을 보내도록 알림 설정을 구성할 수도 있습니다. 성과 기록표에 대해 전자 메일 주소 및 메시지 텍스트를 포함한 전자 메일 템플릿을 구성할 수 있습니다.

1. **Analyst** 도구에서 성과 기록표를 실행합니다.
2. **작업 > 편집 > 알림**을 클릭하여 **성과 기록표 편집** 대화 상자를 엽니다.
3. **알림 활성화**를 선택하여 성과 기록표 알림 구성을 시작합니다.
4. **전역 설정 편집** 아이콘을 클릭합니다.
전자 메일 템플릿을 비롯한 설정을 편집할 수 있는 **전역 설정 편집** 대화 상자가 표시됩니다.
5. **점수 및 점수 이동 시** 확인란을 사용하여 메트릭 점수에 대해 전자 메일 알림을 보내려는 경우를 선택합니다.
6. **비용 위치** 확인란을 사용하여 메트릭 비용에 대해 전자 메일 알림을 보내려는 경우를 선택합니다.
7. **전자 메일 수신인** 필드에 수신자의 전자 메일 ID를 입력합니다. 세미콜론을 사용하여 여러 전자 메일 ID를 구분합니다.
기본 보내는 사람 전자 메일 ID는 도메인 SMTP 속성에서 구성된 **보내는 사람 전자 메일 주소**입니다.
8. 전자 메일 제목의 텍스트를 입력합니다.

9. 본문 필드에 전자 메일 메시지의 소개 및 마무리 텍스트를 추가합니다.
10. 전역 설정을 적용하려면 **모든 메트릭 및 메트릭 그룹에 설정 적용**을 선택합니다.
11. **확인**을 클릭합니다.

성과 기록표 연계

성과 기록표 연계에는 데이터의 출처를 보여 주고, 경로를 설명하고, 메트릭 또는 메트릭 그룹의 데이터 흐름을 표시합니다. 성과 기록표 연계를 사용하면 메트릭 또는 메트릭 그룹 내 불합격 점수 분산의 근본 원인을 분석할 수 있습니다. **Analyst** 도구에서 성과 기록표 연계를 확인합니다.

성과 기록표 연계를 보려면 다음 태스크를 완료합니다.

1. **Informatica Administrator**에서 **Metadata Manager** 서비스를 분석 서비스와 연결합니다.
2. 프로젝트를 선택한 다음, **Developer tool**에 있는 **Metadata Manager**의 리소스 파일 내보내기 옵션을 사용하거나 **infacmd** 도구인 **exportResources** 명령을 사용하여 해당 프로젝트 내의 성과 기록표 개체를 XML 파일로 내보냅니다.
3. **Metadata Manager**에서 내보낸 XML 파일을 사용하여 리소스를 작성한 다음 로드합니다.
참고: **Metadata Manager**에서 작성하여 로드하는 리소스 파일의 이름은 **<MRS name>_<project name>** 명명 규칙을 사용해야 합니다. 리소스 파일을 작성 및 로드하는 방법에 대한 자세한 내용은 *Informatica PowerCenter Metadata Manager 사용자 가이드*를 참조하십시오.
4. **Analyst** 도구에서 성과 기록표를 열고 메트릭 또는 메트릭 그룹을 선택합니다.
5. 성과 기록표 연계를 확인합니다.

Informatica Analyst에서 성과 기록표 연계 보기

메트릭 또는 메트릭 그룹에 대한 성과 기록표 연계 다이어그램을 확인할 수 있습니다. **Analyst** 도구에서 성과 기록표 연계 다이어그램을 확인하려면 **Metadata Manager**에서 성과 기록표 연계 및 메타데이터를 로드해야 합니다.

1. **라이브러리** 작업 공간에서 **자산** 창에서 볼 성과 기록표를 클릭합니다.
성과 기록표가 **성과 기록표** 작업 공간에 표시됩니다.
2. **성과 기록표** 보기에서 메트릭 또는 메트릭 그룹을 선택합니다.
3. 마우스 오른쪽 단추를 클릭하고 **연계 표시**를 선택합니다.
성과 기록표 연계 다이어그램이 새 창에 표시됩니다.

중요: 성과 기록표 개체의 내보낸 XML 파일을 사용하여 **Metadata Manager**에서 리소스를 작성 및 로드하지 않으면 카탈로그에서 리소스를 사용할 수 없다는 오류 메시지가 표시될 수 있습니다. 성과 기록표 연계 확인을 위해 XML 파일을 내보내는 방법에 대한 자세한 내용은 [“성과 기록표 연계의 리소스 파일 내보내기” 페이지 103](#)를 참조하십시오.

파트 III: Informatica Developer에서 프로파일링

이 파트에 포함된 장:

- [데이터 개체 프로필, 80](#)
- [반 구조화된 데이터 소스에 대한 열 프로필, 86](#)
- [Informatica Developer의 규칙, 93](#)
- [맵렛 및 매핑 프로파일링, 95](#)
- [Informatica Developer의 열 프로필 결과, 97](#)
- [Informatica Developer의 성과 기록표, 102](#)

제 9 장

데이터 개체 프로필

이 장에 포함된 항목:

- [Informatica Developer의 열 프로필, 80](#)
- [Informatica Developer에서 단일 데이터 개체 프로필 생성, 81](#)
- [Informatica Developer에서 여러 데이터 개체 프로필 생성, 82](#)
- [Informatica Developer의 플랫폼 파일 데이터 개체 동기화, 83](#)
- [Informatica Developer의 관계형 데이터 개체 동기화, 85](#)

Informatica Developer의 열 프로필

열 프로필을 사용하면 값 백분율, 값 패턴 등 데이터 소스의 열 특성을 분석할 수 있습니다. 필터를 추가하여 프로필이 런타임 시 읽는 행을 결정할 수 있습니다. 프로필에서 필터 조건을 충족하지 않는 행은 처리되지 않습니다.

프로필을 실행하는 열에 대해 다음과 같은 유형의 정보를 검색할 수 있습니다.

- 값이 열에 표시되는 횟수입니다.
- 열의 각 값이 발생하는 빈도이며, 백분율 또는 행 수로 표시됩니다.
- 열에 포함된 값의 문자 패턴입니다.
- 열의 값 최대/최소 길이와 같은 통계와 첫 번째 값 및 마지막 값입니다.
- 유추된 데이터 유형, 빈도, 데이터 도메인 검색의 준수 조건 및 데이터 유형 유추 상태입니다.

모델 리포지토리의 개체, 맵렛 또는 매핑에서 데이터 개체에 대한 열 프로필을 정의할 수 있습니다. 리포지토리의 개체는 단일 데이터 개체 프로필, 여러 데이터 개체 프로필 또는 엔터프라이즈 검색 프로필에 포함되어 있을 수 있습니다.

열 프로필에 대해 샘플링 옵션, 드릴다운 옵션 및 런타임 환경을 선택할 수 있습니다. 열 프로필에 규칙 및 필터를 추가할 수 있습니다.

필터링 옵션

고급 필터 또는 SQL 필터를 추가하여 프로필을 실행할 때 열 프로필이 사용하는 행을 결정할 수 있습니다. 프로필에서 필터 조건을 충족하지 않는 행은 처리되지 않습니다.

샘플링 옵션

샘플링 옵션은 **Developer tool**이 프로필을 실행하는 행 수를 결정합니다. 프로필을 정의하거나 실행할 때 샘플링 옵션을 구성할 수 있습니다.

다음 테이블에서는 프로필의 샘플링 옵션에 대해 설명합니다.

속성	설명
모든 행	데이터 개체의 모든 행에서 프로필을 실행합니다. 원시, Blaze, Spark 및 Databricks 런타임 환경에서 지원됩니다.
처음 <수>개 행 샘플링	데이터 개체의 행 처음부터 샘플 행에서 프로필을 실행합니다. 최대 2,147,483,647개의 행을 선택할 수 있습니다. 원시 및 Blaze 런타임 환경에서 지원됩니다.
<수>개 행 무작위 샘플링	데이터 개체의 무작위로 선택한 행 수에서 프로필을 실행합니다. 최대 2,147,483,647개의 행을 선택할 수 있습니다. 원시 및 Blaze 런타임 환경에서 지원됩니다.
무작위 샘플링(자동)	데이터 개체의 행 수를 기초로 계산된 샘플 행에서 프로필을 실행합니다. 원시 및 Blaze 런타임 환경에서 지원됩니다.
n <수>개 행 제한	데이터 개체의 행 수를 기반으로 프로필을 실행합니다. Hadoop 유효성 검사 환경에서 프로필을 실행하도록 선택하면 Spark 엔진이 데이터 개체의 여러 파티션에서 샘플을 수집하고 샘플 크기를 계산할 단일 노드에 샘플을 푸시합니다. n 제한 샘플링 옵션은 Oracle, SQL Server 및 DB2 데이터베이스를 지원합니다. n 제한 샘플링 옵션과 함께 고급 필터를 적용할 수 없습니다. Spark 런타임 환경에서 지원됩니다.
무작위 비율	데이터 개체의 행 비율에서 프로필을 실행합니다. Spark 및 Databricks 런타임 환경에서 지원됩니다.
후속 프로필 실행 시 데이터 유형 및 데이터 도메인 유추에서 승인된 데이터 유형 및 데이터 도메인 제외	다음 프로필 실행부터 데이터 유형 및 데이터 도메인 유추에서 승인된 데이터 유형 또는 데이터 도메인을 제외합니다.

무작위 샘플링 행에서 프로필을 실행하도록 선택하면 무작위 샘플링 알고리즘이 데이터 개체에서 행을 무작위로 선택하여 프로필을 실행합니다. 열 프로필에 대한 무작위 샘플링 옵션을 선택하면 **Developer tool**이 준비된 데이터에 대해 드릴다운을 수행합니다. 이는 드릴다운 성능에 영향을 줄 수 있습니다. 데이터 도메인 검색 프로필에 대해 무작위 샘플링 옵션을 선택하면 **Developer tool**이 라이브 데이터에 대해 드릴다운을 수행합니다.

Informatica Developer에서 단일 데이터 개체 프로필 생성

데이터 개체에서 하나 이상의 열에 대해 단일 데이터 개체 프로필을 생성하고 모델 리포지토리에 해당 프로필 개체를 저장할 수 있습니다.

1. **개체 탐색기** 보기에서 프로파일링할 데이터 개체를 선택합니다.
2. **파일 > 새로 만들기 > 프로필**을 클릭하여 프로필 마법사를 엽니다.

3. **프로필**을 선택하고 **다음**을 클릭합니다.
4. 프로필의 이름을 입력하고 프로젝트 위치를 확인합니다. 필요한 경우 새 위치를 찾아봅니다.
5. 필요한 경우 프로필의 텍스트 설명을 입력합니다.
6. 선택한 데이터 개체의 이름이 **데이터 개체** 섹션에 표시되는지 확인합니다.
7. **다음**을 클릭합니다.
8. 수행할 프로필 작업을 구성합니다. 다음 작업을 구성할 수 있습니다.
 - 열 프로파일링
 - 기본 키 검색
 - 함수 종속성 검색
 - 데이터 도메인 검색

참고: 프로필 작업을 활성화하려면 해당 작업에 대해 "**프로필 실행**" 작업의 일부로 **활성화됨**을 선택합니다. 기본적으로 열 프로파일링이 활성화됩니다.
9. 프로필에 대한 옵션을 검토합니다.

모든 프로필 유형에 대한 열 선택을 편집할 수 있습니다. 열 프로필에 대한 필터 및 샘플링 옵션을 검토합니다. 기본 키, 함수 종속성 및 데이터 도메인 검색에 대한 유추 옵션을 검토할 수 있습니다. 데이터 도메인 검색을 위한 데이터 도메인 선택도 검토할 수 있습니다.
10. 드릴다운 옵션을 검토하고 필요하면 편집합니다. 기본적으로 **행 드릴다운 활성화** 옵션이 선택됩니다. 열 프로필의 드릴다운 옵션을 편집할 수 있습니다. 이 옵션은 드릴다운 작업이 데이터 소스에서 읽는지 아니면 준비된 데이터에서 읽는지와 프로필이 이전 프로필 실행의 결과 데이터를 저장하는지도 결정합니다.
11. **실행 설정** 섹션에서 유효성 검사 환경을 선택합니다. 유효성 검사 환경으로 **원시**, **Hadoop** 또는 **Databricks**를 선택합니다. 런타임 환경으로 **원시**, **Blaze** 또는 **Spark**를 선택할 수 있습니다. **Blaze** 또는 **Spark**를 선택한 후 **Hadoop** 연결을 선택할 수 있습니다. **Databricks**를 선택하는 경우 **Databricks** 연결을 선택할 수 있습니다.
12. **마침**을 클릭합니다.

Informatica Developer에서 여러 데이터 개체 프로필 생성

여러 데이터 개체에 대해 여러 데이터 개체 프로필을 실행할 경우, **Developer tool**은 기본 열 프로파일링 옵션을 사용하여 하나 이상의 데이터 개체에 대해 열 프로필을 생성합니다. 필요에 따라 여러 데이터 개체에 대해 프로필을 실행하기 위해 엔터프라이즈 검색 프로필을 생성할 수 있습니다.

1. **개체 탐색기** 보기에서 프로파일링할 데이터 개체를 선택합니다.
2. **파일 > 새로 만들기 > 프로필**을 클릭하여 새 **프로필** 마법사를 엽니다.
3. **새로 만들기** 마법사에서 **여러 프로필** 옵션을 선택하고 **다음**을 클릭합니다.
4. **여러 프로필** 창에서 프로필을 생성할 위치를 선택합니다. 프로파일링된 개체와 같은 위치에 각 프로필을 작성할 수도 있고 프로필에 대해 공통 위치를 지정할 수도 있습니다.
5. 선택한 데이터 개체의 이름이 **데이터 개체** 섹션에 표시되는지 확인합니다.

필요에 따라 **추가**를 클릭하여 다른 데이터 개체를 추가합니다.
6. 필요에 따라 프로파일링할 행 수를 지정하고 마법사 완료 시 프로필을 실행할지 여부를 선택합니다.
7. **다음**을 클릭합니다.

8. **유효성 검사 환경** 섹션에서 **원시**를 선택합니다.

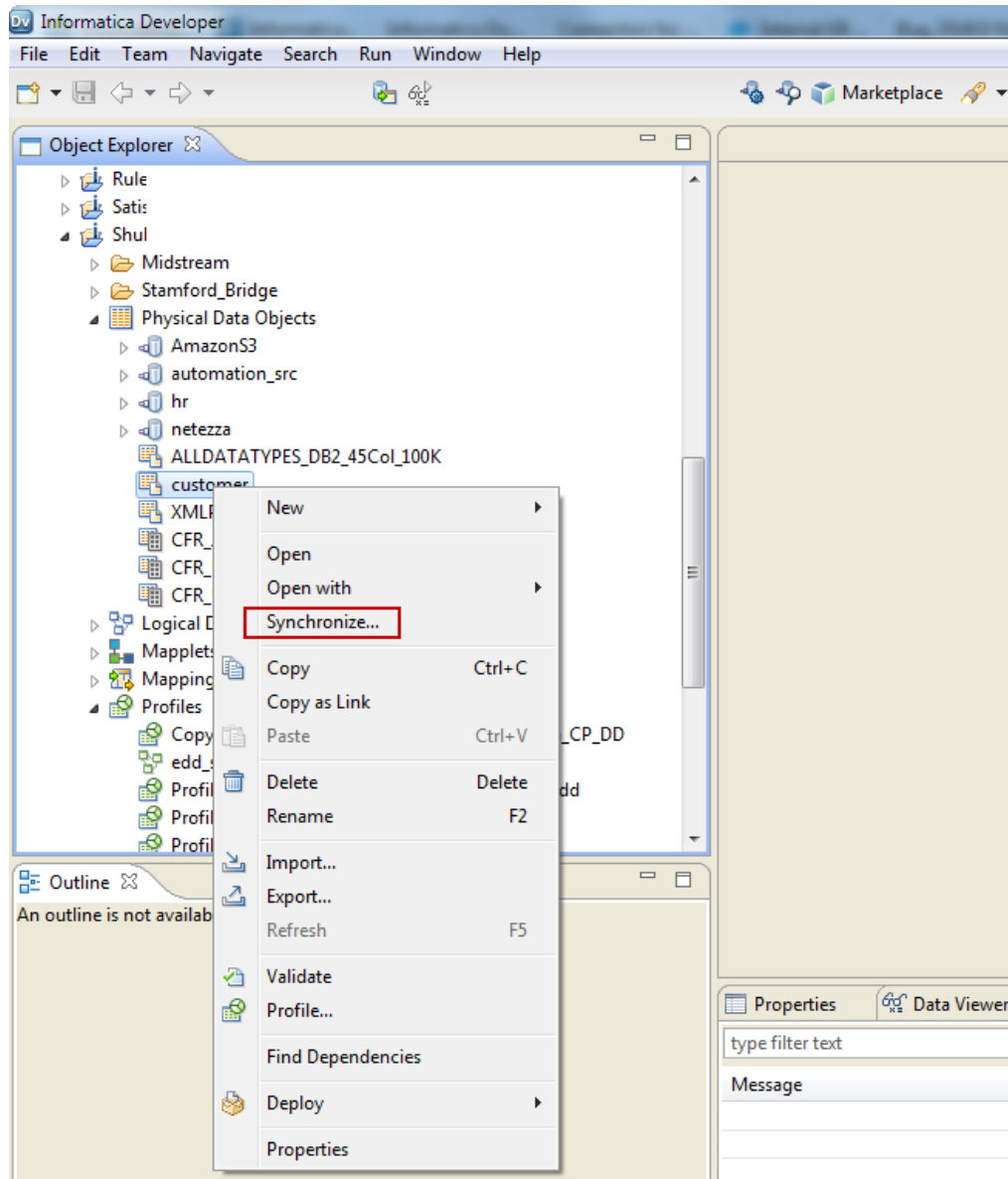
참고: 여러 데이터 개체 프로필을 실행하려면 원시 옵션만 선택해야 합니다. **Hadoop** 런타임 환경의 **Blaze** 또는 **Spark** 엔진에서 여러 데이터 개체를 실행하려는 경우에는 엔터프라이즈 데이터 검색 프로필을 선택할 수 있습니다. **Databricks** 런타임 환경의 **Databricks Spark**에서 여러 데이터 개체를 실행하려는 경우에는 엔터프라이즈 데이터 검색 프로필을 선택할 수 있습니다.

9. **마침**을 클릭합니다.
10. 필요에 따라 프로필 이름에 추가할 접두사 및 접미사 문자열을 입력합니다.
11. **확인**을 클릭합니다.

Informatica Developer의 플랫폼 파일 데이터 개체 동기화

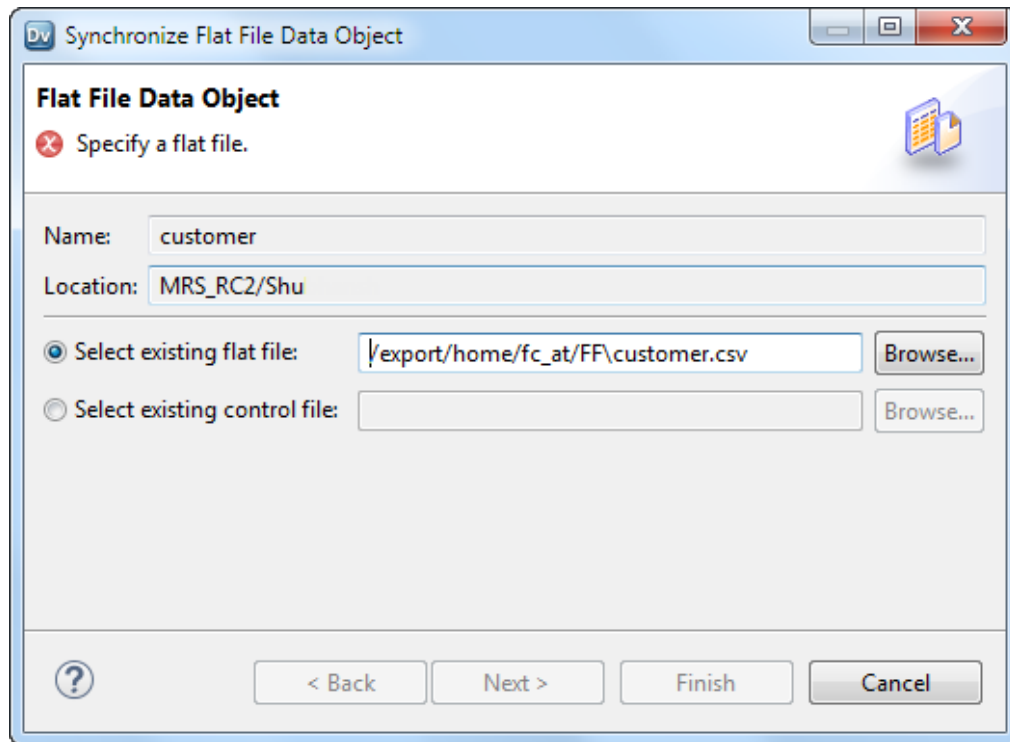
Informatica Developer에서 외부 플랫폼 파일 데이터 소스에 대한 변경 내용을 해당 데이터 개체와 동기화할 수 있습니다. **플랫폼 파일 동기화** 마법사를 사용하여 데이터 개체를 동기화합니다.

1. **Object Explorer** 보기에서 플랫폼 파일 데이터 개체를 선택합니다.
2. 마우스 오른쪽 단추를 클릭하고 **동기화**를 선택합니다.
다음 이미지는 데이터 개체에 대한 동기화 옵션을 보여 줍니다.



플랫 파일 데이터 개체 동기화 마법사가 나타납니다.

3. 플랫 파일 데이터 개체 동기화 마법사에서 기존 플랫 파일 선택 필드에 있는 플랫 파일 경로를 확인합니다.
다음 이미지는 플랫 파일 데이터 개체 동기화 마법사를 보여 줍니다.



4. 다음을 클릭합니다.
5. 필요에 따라 코드 페이지, 형식, 구분자로 분리된 형식 속성 및 열 속성을 선택합니다.
6. 마침을 클릭한 다음 확인을 클릭합니다.

Informatica Developer의 관계형 데이터 개체 동기화

Informatica Developer에서 관계형 데이터 소스의 외부 데이터 소스 변경 내용을 해당 데이터 개체와 동기화할 수 있습니다. 외부 데이터 소스 변경 내용으로는 열 추가/변경/제거, 규칙 변경 등이 있습니다.

1. 개체 탐색기 보기에서 관계형 데이터 개체를 선택합니다.
2. 마우스 오른쪽 단추를 클릭하고 동기화를 선택합니다.
작업을 확인하라는 메시지가 표시됩니다.
3. 동기화 프로세스를 완료하려면 확인을 클릭합니다.
동기화 프로세스 상태 메시지가 나타납니다.
4. 동기화가 완료되었습니다. 메시지가 표시되면 확인을 클릭합니다.
메시지에는 데이터 개체에 대한 메타데이터 변경 내용의 요약이 표시됩니다.

제 10 장

반 구조화된 데이터 소스에 대한 열 프로파일

이 장에 포함된 항목:

- [반 구조화된 데이터 소스에 대한 열 프로파일 개요, 86](#)
- [JSON 및 XML 데이터 개체, 87](#)
- [HDFS에 있는 반 구조화된 데이터 소스에 대한 복잡한 파일 데이터 개체, 88](#)
- [HDFS 연결 작성, 88](#)
- [HDFS에 있는 JSON 또는 XML 파일로부터 복잡한 파일 데이터 개체 생성, 89](#)
- [Avro 또는 Parquet 데이터 소스로부터 복잡한 파일 데이터 개체 생성, 90](#)
- [반 구조화된 데이터 소스에 대한 열 프로파일 생성, 91](#)

반 구조화된 데이터 소스에 대한 열 프로파일 개요

Avro, JSON, Parquet 및 XML 데이터 소스에서 데이터 개체를 작성한 다음 데이터 개체에 대해 열 프로 파일을 작성할 수 있습니다.

Avro, JSON, Parquet 및 XML 형식은 반 구조화된 데이터 소스입니다. 반 구조화된 데이터 소스를 사용하여 열 프로 파일을 작성하려면 다음과 같은 작업을 수행할 수 있습니다.

1. 반 구조화된 데이터 소스에 대해 실제 데이터 개체를 작성합니다.
2. 실제 데이터 개체에 대해 열 프로 파일을 작성하고 실행합니다.

JSON 또는 XML 데이터 소스에 대해 플랫폼 파일 데이터 개체를 작성할 수 있습니다. HDFS(Hadoop Distributed File System)에 있는 Avro, JSON, Parquet 및 XML 데이터 소스에 대해 복잡한 파일 데이터 개체를 작성할 수 있습니다.

JSON 및 XML 데이터 개체

JSON 또는 XML 데이터 소스로부터 플랫폼 파일 데이터 개체 또는 복잡한 파일 데이터 개체를 생성할 수 있습니다. 데이터 개체에 대해 열 프로필을 작성하고 실행할 수 있습니다.

JSON 또는 XML 데이터 소스의 경로가 포함된 텍스트 파일을 작성한 후 이 텍스트 파일을 플랫폼 파일 데이터 개체를 작성할 때 데이터 소스로 사용합니다. 여러 JSON 또는 여러 XML 데이터 소스의 파일 경로를 텍스트 파일에 추가할 수도 있습니다.

복잡한 파일 관독기를 사용하면 JSON 또는 XML 데이터 소스에서 복잡한 파일 데이터 개체를 작성할 수 있습니다. 복잡한 파일 관독기는 데이터 프로세서 변환에 입력을 제공하고, 데이터 프로세서 변환은 이 파일을 구문 분석하여 소스 데이터를 값이 쉼표로 구분된 플랫폼 레코드로 변환합니다.

참고: Developer tool은 UTF-8 인코딩을 사용하는 JSON 데이터 소스를 지원하지 않습니다.

JSON 또는 XML 데이터 소스로부터 데이터 개체 생성

JSON 또는 XML 데이터 소스로부터 플랫폼 파일 데이터 개체 또는 복잡한 파일 데이터 개체를 생성할 수 있습니다.

1. Developer tool의 **Object Explorer** 보기에서 데이터 개체 및 열 프로필을 작성할 프로젝트를 선택합니다.
2. **파일 > 새로 만들기 > 데이터 개체**를 클릭합니다.

새로 만들기 대화 상자가 나타납니다.

3. 플랫폼 파일 데이터 개체 또는 복잡한 파일 데이터 개체를 작성하도록 선택할 수 있습니다.
 - 플랫폼 파일 데이터 개체를 작성하려면 다음 태스크를 수행합니다.
 1. **실제 데이터 개체 > 플랫폼 파일 데이터 개체**를 선택하고 **다음**을 클릭합니다.
새 플랫폼 파일 데이터 개체 대화 상자가 나타납니다.
 2. **기존 플랫폼 파일에서 작성**을 선택하고 **찾아보기**를 클릭하여 텍스트 파일을 선택합니다. **다음**을 클릭합니다.
 3. 코드 페이지가 **MS Windows Latin 1 (ANSI), superset of Latin 1**이고 형식이 구분자로 분리되었는지 확인합니다. **다음**을 클릭합니다.
 4. 구분자가 **쉼표**로 설정되어 있는지 확인합니다. **마침**을 클릭합니다.
 - 복잡한 파일 데이터 개체를 작성하려면 다음 태스크를 수행합니다.
 1. **실제 데이터 개체 > 복잡한 파일 데이터 개체**를 선택하고 **다음**을 클릭합니다.
새 복잡한 파일 데이터 개체 대화 상자가 나타납니다.
 2. 데이터 개체의 이름을 입력합니다. 액세스 유형을 **파일**로 선택합니다.
 3. **찾아보기**를 클릭하여 JSON 또는 XML 파일을 선택합니다. **마침**을 클릭합니다.
개발자 서버가 Linux인 경우 데이터 소스의 파일 경로를 서버의 위치로 업데이트해야 합니다. 파일 경로를 업데이트하려면 복잡한 파일 데이터 개체를 선택하고 **데이터 개체 작업** 탭에서 **읽기**를 클릭한 다음 **데이터 개체 작업 세부 정보** 창의 **고급** 탭에 파일 경로를 추가합니다.

데이터 개체가 프로젝트 폴더에 나타납니다.

HDFS에 있는 반 구조화된 데이터 소스에 대한 복잡한 파일 데이터 개체

HDFS를 사용하는 Avro, JSON, Parquet 또는 XML 파일에 대해 열 프로필을 작성하고 실행할 수 있습니다. HDFS에 있는 JSON 또는 XML 파일을 읽으려면 JSON 또는 XML 입력을 데이터 프로세서 변환으로 전달하는 복잡한 파일 관독기가 필요합니다.

HDFS에 있는 JSON 또는 XML 데이터 소스의 복잡한 파일 데이터 개체

JSON 또는 XML 파일에서 복잡한 파일 데이터 개체를 작성할 수 있습니다. 데이터 개체에 대해 열 프로필을 작성하고 실행할 수 있습니다.

HDFS에 있는 JSON 또는 XML 파일에 대해 데이터 개체를 작성하기 전에 HDFS에 대한 연결을 작성합니다.

다음 방법 중 하나로 HDFS에 있는 JSON 또는 XML 파일에서 데이터 개체를 작성할 수 있습니다.

- JSON 또는 XML 파일에 대해 복잡한 파일 데이터 개체를 작성합니다.
- 여러 개의 JSON 또는 여러 개의 XML 파일이 포함된 폴더에 대해 복잡한 파일 데이터 개체를 작성합니다.

데이터 개체를 작성한 후에는 데이터 개체에 대해 열 프로필을 작성하여 실행할 수 있습니다.

HDFS에 있는 Avro 또는 Parquet 데이터 소스의 복잡한 파일 데이터 개체

HDFS에 있는 Avro 또는 Parquet 데이터 소스에서 복잡한 파일 데이터 개체를 작성할 수 있습니다. 데이터 개체를 사용하여 열 프로필을 작성하고 실행할 수 있습니다.

Avro 파일이나 Parquet 파일 또는 여러 개의 Avro 파일이나 Parquet 파일이 들어 있는 폴더로부터 복잡한 파일 데이터 개체를 생성할 수 있습니다. 파일 또는 연결 액세스 유형을 사용하고 이진, Avro 또는 Parquet을 리소스 형식으로 하여 Avro 및 Parquet 데이터 소스로부터 복잡한 파일 데이터 개체를 생성할 수 있습니다. Avro 및 Parquet 데이터 소스로부터 복잡한 파일 데이터 개체를 생성하려면 먼저 HDFS 연결을 생성해야 합니다.

참고: Avro 및 Parquet 데이터 소스가 플랫 형식으로 구조화된 경우에만 리소스 형식으로 Avro 또는 Parquet을 선택할 수 있습니다.

HDFS에 있는 Avro 및 Parquet 파일에서 데이터 개체를 작성할 때 다음의 옵션 중 하나를 선택할 수 있습니다.

- 액세스 유형으로 파일을 선택하고 리소스 형식으로 이진을 선택합니다.
- 액세스 유형으로 파일을 선택하고 리소스 형식으로 Avro 또는 Parquet을 선택합니다.
- 액세스 유형으로 연결을 선택하고 리소스 형식으로 Avro 또는 Parquet을 선택합니다.

HDFS 연결 작성

HDFS에 있는 Avro, JSON, Parquet 및 XML 데이터 소스에 대해 열 프로필을 작성하려면 Informatica Developer에서 HDFS 연결을 구성합니다. 복잡한 파일 데이터 개체는 HDFS 연결을 작성한 후에 작성할 수 있습니다.

1. **창 > 기본 설정**을 클릭합니다.
2. **Informatica > 연결**을 선택합니다.

3. 도메인을 확장합니다.
4. 연결 유형으로 **파일 시스템 > Hadoop 파일 시스템**을 선택하고 **추가**를 클릭합니다.
5. 연결 이름을 입력합니다.
6. 필요에 따라 연결 설명을 입력합니다.
7. **다음**을 클릭합니다.
8. 연결 속성을 구성합니다.
9. **연결 테스트**를 클릭하여 HDFS에 대한 연결을 확인합니다.
10. **마침**을 클릭합니다.

HDFS에 있는 JSON 또는 XML 파일로부터 복잡한 파일 데이터 개체 생성

HDFS를 사용하는 JSON 또는 XML 소스 파일로부터 복잡한 파일 데이터 개체를 생성하고 데이터 개체에 열 프로필을 생성할 수 있습니다.

1. Developer tool의 **Object Explorer** 보기에서 실제 데이터 개체 및 열 프로필을 생성할 프로젝트를 선택합니다.
2. **파일 > 새로 만들기 > 데이터 개체**를 클릭합니다.
새로 만들기 대화 상자가 나타납니다.
3. **실제 데이터 개체 > 복잡한 파일 데이터 개체**를 선택하고 **다음**을 클릭합니다.
새 복잡한 파일 데이터 개체 대화 상자가 나타납니다.
4. 데이터 개체의 이름을 입력합니다. 액세스 유형을 **연결**로 선택합니다.
5. JSON 파일이나 XML 파일로부터 데이터 개체를 생성하거나, 여러 개의 JSON 파일 또는 여러 개의 XML 파일이 포함된 폴더에 대해 생성할 수 있습니다.
 - JSON 또는 XML 파일로부터 복잡한 파일 데이터 개체를 생성하려면 다음 단계를 수행합니다.
 1. **찾아보기**를 클릭하여 연결을 선택합니다.
 2. **리소스 추가** 대화 상자에서 **추가**를 클릭하여 JSON 또는 XML 파일을 선택합니다.
 3. **마침**을 클릭합니다.
데이터 개체가 프로젝트 폴더에 나타납니다.
 - 여러 개의 JSON 파일 또는 여러 개의 XML 파일이 포함된 폴더에 대해 복잡한 파일 데이터 개체를 생성하려면 다음 단계를 수행합니다.
 1. **찾아보기**를 클릭하여 연결을 선택합니다.
 2. **리소스 추가** 대화 상자에서 **추가**를 클릭하여 폴더의 JSON 또는 XML 파일을 선택합니다.
 3. **마침**을 클릭합니다.
데이터 개체가 프로젝트 폴더에 나타납니다.
 4. 프로젝트 폴더에서 데이터 개체를 선택하고 **고급 > 런타임: 읽기 > 소스 파일 디렉터리**를 클릭합니다.
 5. 파일 경로에서 파일 이름을 제거하고 폴더 이름은 유지합니다.

Avro 또는 Parquet 데이터 소스로부터 복잡한 파일 데이터 개체 생성

액세스 유형을 **파일** 또는 **연결**로 지정하여 Avro 또는 Parquet 데이터 소스로부터 복잡한 파일 데이터 개체를 생성할 수 있습니다. 데이터 개체에 대해 열 프로필을 생성할 수 있습니다.

1. **Object Explorer** 보기에서 프로젝트를 선택합니다.
2. **파일 > 새로 만들기 > 데이터 개체**를 클릭합니다.
새로 만들기 대화 상자가 나타납니다.
3. **실제 데이터 개체 > 복잡한 파일 데이터 개체**를 선택하고 **다음**을 클릭합니다.
새 복잡한 파일 데이터 개체 대화 상자가 나타납니다.
4. 데이터 개체의 이름을 입력합니다.
5. 액세스 유형으로 **연결** 또는 **파일**을 선택할 수 있습니다.
 - 액세스 유형으로 **연결**을 선택한 경우, 다음 단계를 수행합니다.
 1. **찾아보기**를 클릭하고 HDFS 연결을 선택합니다.
 2. **연결 선택** 대화 상자에서 데이터 소스를 선택하고 **확인**을 클릭합니다.
 3. **새 복잡한 파일 데이터 개체** 대화 상자에서 **마침**을 클릭합니다.
데이터 개체가 프로젝트 폴더에 나타납니다.
 - 액세스 유형으로 **파일**을 선택하고 리소스 형식으로 **이진**을 선택한 경우, 다음 단계를 수행합니다.
 1. **찾아보기**를 클릭하고 로컬 시스템에서 Avro 또는 Parquet 파일을 선택합니다.
 2. **새 복잡한 파일 데이터 개체** 대화 상자에서 **마침**을 클릭합니다.
데이터 개체가 프로젝트 폴더에 나타납니다.
 3. 프로젝트 폴더에서 데이터 개체를 선택하고 **데이터 개체 작업** 보기를 클릭합니다.
 4. **데이터 개체 작업** 보기에서 **읽기 > 고급** 탭을 클릭합니다.
 5. **고급** 탭의 **파일 경로** 필드에 Linux 또는 Windows 시스템에 있는 데이터 소스의 파일 경로를 입력합니다.
 6. 파일 형식으로 **사용자 지정** 입력을 입력합니다.
 7. Avro 데이터 소스의 경우 **입력 형식** 필드에 **com.informatica.avro.AvroToXML**을 입력하고, Parquet 데이터 소스의 경우 **입력 형식** 필드에 **com.informatica.parquet.ParquetToXML**을 입력합니다. 입력 형식을 추가하면 데이터 프로세서 변환이 런타임 시 Avro 또는 Parquet 형식의 데이터 소스를 XML 형식의 데이터 소스로 변환합니다.
 - 액세스 유형으로 **파일**을 선택하고 리소스 형식으로 **Avro** 또는 **Parquet**을 선택한 경우, 다음 단계를 수행합니다.
 1. **찾아보기**를 클릭하고 로컬 시스템에서 Avro 또는 Parquet 파일을 선택합니다.
 2. **새 복잡한 파일 데이터 개체** 대화 상자에서 **마침**을 클릭합니다.
데이터 개체가 프로젝트 폴더에 나타납니다.
 3. 데이터 개체를 생성한 후에는 **데이터 개체 작업 > 읽기 > 고급** 탭으로 이동한 후 **파일 경로** 필드의 파일 경로가 Linux 또는 Windows 시스템에 있는 데이터 소스의 경로인지 확인합니다.

참고: Avro 및 Parquet 데이터 소스가 플랫폼 형식으로 구조화된 경우에만 리소스 형식으로 **Avro** 또는 **Parquet**을 선택할 수 있습니다.

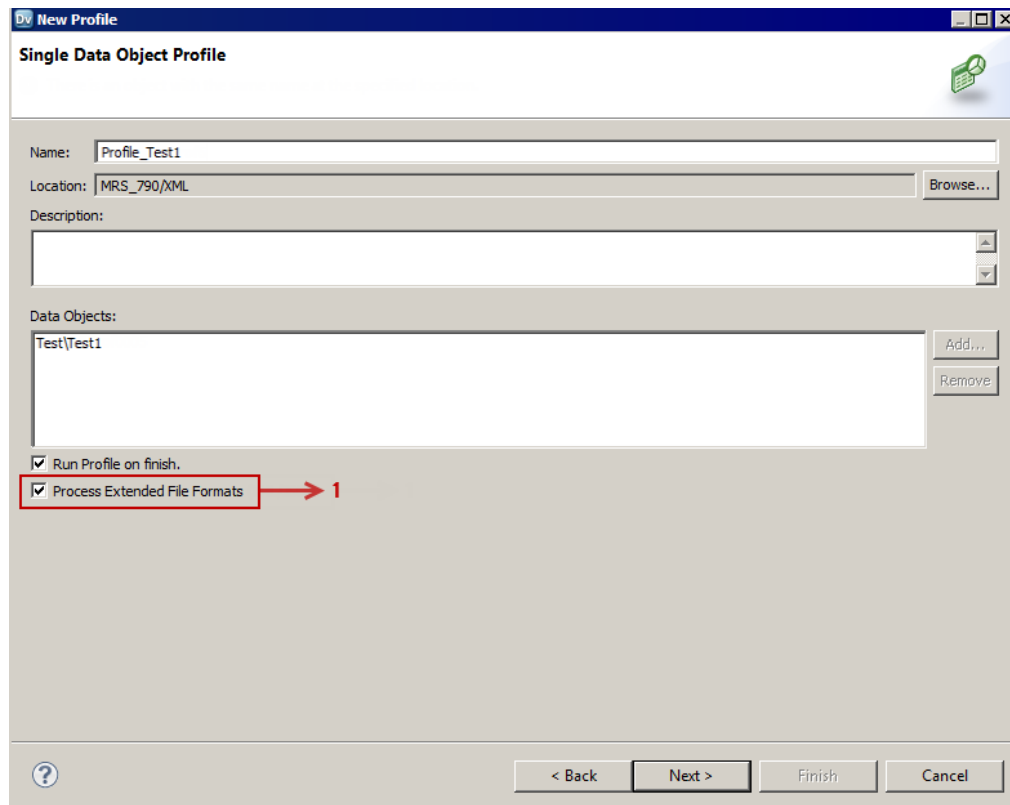
데이터 개체를 생성하기 위해 여러 개의 Avro 파일 또는 여러 개의 Parquet 파일이 있는 폴더를 선택할 수 있습니다. 데이터 개체를 생성한 후에는 **데이터 개체 작업 > 읽기 > 고급** 탭으로 이동한 후 **파일 경로** 필드의 파일 경로가 Linux 또는 Windows 시스템에 있는 데이터 소스 폴더의 경로를 가리키는 지 확인합니다.

반 구조화된 데이터 소스에 대한 열 프로필 생성

Avro, JSON, Parquet 또는 XML 데이터 소스에서 플랫폼 파일 데이터 개체 또는 복잡한 파일 데이터 개체를 생성한 후에는 데이터 개체에 대해 열 프로필을 생성하고 실행할 수 있습니다.

1. **Object Explorer** 보기에서 **Avro, JSON, Parquet** 또는 **XML** 파일의 데이터 개체를 선택합니다.
2. **파일 > 새로 만들기 > 프로필**을 클릭합니다.
새로 만들기 대화 상자가 나타납니다.
3. **프로필**을 선택합니다. **다음**을 클릭합니다.
새 **프로필** 대화 상자가 나타납니다.
4. 새 **프로필** 대화 상자에서 프로필에 대한 이름과 선택적 설명을 추가합니다.
5. **확장 파일 형식 처리** 옵션을 선택합니다. **다음**을 클릭합니다.

다음 이미지는 새 **프로필** 마법사의 **확장 파일 형식 처리** 옵션을 보여 줍니다.



1. 확장 파일 형식 처리. 반 구조화된 데이터 소스를 처리하려면 이 옵션을 선택합니다.

참고: 리소스 형식을 **Avro** 또는 **Parquet**으로 선택한 경우에는 Avro 및 Parquet 데이터 소스에 대해 **확장 파일 형식 처리** 옵션이 나타나지 않습니다.

6. **단일 데이터 개체 프로필** 페이지에서 필요에 따라 **열 선택** 및 **데이터 도메인 검색**에서 열 및 옵션을 선택합니다. **마침**을 클릭합니다.

참고: Developer tool이 Linux 시스템에 설치되어 있고 JSON 또는 XML 실제 데이터 개체가 텍스트 파일이 포함된 플랫폼 파일 데이터 개체인 경우에는 다음 태스크를 수행하십시오.

1. 서버의 데이터 소스 파일 경로의 문자 수가 포함되도록 **개요** 탭에서 **전체 자릿수** 값을 업데이트합니다.

2. 플랫폼 파일 데이터 개체에 대한 프로필을 생성한 후 데이터 소스의 파일 경로를 서버의 위치로 업데이트합니다. 파일 경로를 업데이트하려면 **고급** 탭에서 **런타임: 읽기 > 소스 파일 디렉터리**를 클릭하고 파일 경로를 추가합니다.
7. 마우스 오른쪽 단추로 프로필을 클릭하고 **프로필 실행**을 선택합니다.
프로필 결과가 나타납니다.

제 11 장

Informatica Developer의 규칙

이 장에 포함된 항목:

- [Informatica Developer의 규칙 개요, 93](#)
- [Informatica Developer에서 규칙 작성, 94](#)
- [Informatica Developer에서 규칙 적용, 94](#)

Informatica Developer의 규칙 개요

규칙은 열 프로필을 실행할 때 소스 데이터에 적용되는 조건을 정의하는 비즈니스 논리입니다. 프로필에 규칙을 추가하여 데이터 유효성을 검사할 수 있습니다. 열 프로필에서 규칙, 미리 정의된 규칙 또는 재사용 가능한 규칙으로 유효성이 검사되는 맵셋을 사용할 수 있습니다.

다음 방법을 사용하여 열 프로필에서 규칙을 사용할 수 있습니다.

- **Developer tool**에서 맵셋을 생성하고 규칙으로 유효성을 확인합니다. 규칙은 **Analyst** 도구에서 재사용 가능 규칙으로 나타납니다. **Analyst** 도구 및 **Developer tool**에서 열 프로필에 규칙을 적용할 수 있습니다.
- 열 프로필에서 미리 정의된 규칙을 사용할 수 있습니다. Informatica는 **Developer tool** 및 **Analyst** 도구에 미리 정의된 규칙을 제공합니다.
- **Analyst** 도구에서 규칙 사양을 생성하고 맵셋을 생성합니다. **Analyst** 도구의 열 프로필에 규칙 사양을 적용할 수 있습니다. **Developer tool**에서 규칙으로 맵셋의 유효성을 확인합니다. 규칙은 열 프로필에서 사용할 수 있는 재사용 가능 규칙으로 나타납니다.

참고: **Developer tool**에서는 열 프로필의 규칙 사양을 추가, 편집 또는 삭제할 수 없습니다.

규칙은 다음 요구 사항을 충족해야 합니다.

- 규칙은 입력 및 출력 변환을 포함해야 합니다. 규칙에서는 데이터 소스를 사용할 수 없습니다.
- 규칙은 식 변환, 조회 변환 및 수동 데이터 품질 변환을 포함할 수 있습니다. 규칙은 다른 유형의 변환을 포함할 수 없습니다. 예를 들어 규칙은 활성 변환인 일치 변환을 포함할 수 없습니다.
- 이는 입력 그룹 간의 카디널리티를 지정하지 않습니다.

Informatica Developer에서 규칙 작성

Developer 도구에서 규칙을 작성하려면 맵렛을 규칙으로 유효성을 검사해야 합니다.

Developer 도구에서 맵렛을 작성합니다.

1. 맵렛 편집기를 마우스 오른쪽 단추로 클릭합니다.
2. 다음으로 유효성 검사 > 규칙을 선택합니다.

Informatica Developer에서 규칙 적용

저장된 열 프로필에 규칙을 추가할 수 있습니다. 조인 분석에 대해 구성된 프로필에는 규칙을 추가할 수 없습니다.

1. **개체 탐색기** 보기로 이동하여 필요한 프로필을 찾습니다.
2. 프로필을 마우스 오른쪽 단추로 클릭하고 **열기**를 선택합니다.
편집기에서 프로필이 열립니다.
3. **정의** 탭을 클릭하고 규칙을 선택합니다.
4. **추가**를 클릭합니다.
규칙 적용 대화 상자가 열립니다.
5. **찾아보기**를 클릭하여 적용할 규칙을 찾습니다.
리포지토리 프로젝트에서 규칙을 선택하고 **확인**을 클릭합니다.
6. **입력 값** 아래의 **값** 열을 클릭하여 규칙의 입력 포트를 선택합니다.
7. 필요에 따라 **출력 값** 아래의 **값** 열을 클릭하여 규칙 출력 포트의 이름을 편집합니다.
규칙이 **정의** 탭에 표시됩니다.

제 12 장

맵렛 및 매핑 프로파일링

이 장에 포함된 항목:

- [맵렛 및 매핑 프로파일링 개요, 95](#)
- [맵렛 또는 매핑 개체에 대해 프로파일 실행, 95](#)
- [매핑 또는 맵렛 개체용 프로파일 비교, 96](#)
- [프로파일에서 매핑 생성, 96](#)

맵렛 및 매핑 프로파일링 개요

맵렛 또는 매핑의 개체에 대해 열 프로파일을 정의할 수 있습니다. 프로파일 결과를 저장하지 않고 매핑 또는 맵렛의 디자인을 확인하려면 맵렛 또는 매핑 개체에 대해 프로파일을 실행합니다. 프로파일에서 매핑을 생성할 수도 있습니다.

맵렛 또는 매핑 개체에 대해 프로파일 실행

맵렛 또는 매핑 개체에 대해 실행하는 프로파일은 모든 데이터 열에서 실행되며 데이터 개체용으로 준비된 데이터에 대한 드릴다운 작업이 활성화됩니다. 여러 출력 포트를 사용하여 맵렛 또는 매핑 개체에 대해 프로파일을 실행할 수 있습니다.

프로파일은 선택한 개체의 출력 포트에 대한 매핑을 통해 소스 데이터를 추적합니다. 프로파일은 매핑을 실행한 경우 해당 포트에 표시되는 데이터를 분석합니다.

1. 맵렛 또는 매핑을 엽니다.
2. 맵렛 또는 매핑이 유효한지 확인합니다.
3. 데이터 개체 또는 변환을 마우스 오른쪽 단추로 클릭하고 **지금 프로파일링**을 선택합니다.
변환에 출력 그룹이 여러 개인 경우 **출력 그룹 선택** 대화 상자가 표시됩니다.
4. 변환에 여러 출력 그룹이 있는 경우 필요에 따라 출력 그룹을 선택합니다.
5. **확인**을 클릭합니다.

프로파일 결과가 프로파일의 **결과** 탭에 표시됩니다. 프로파일 요약과 매핑 로그 파일을 확인하면 Developer tool에서 수행한 태스크에 대한 자세한 정보를 볼 수 있습니다.

참고: 운영 체제 프로파일을 사용하여 프로파일을 실행한 경우 요약 로그는 데이터 통합 서비스에 대해 구성된 로그 디렉터리에 나타나고 매핑 로그는 운영 체제 프로파일에 대해 구성된 로그 디렉터리에 나타납니다.

매핑 또는 맵렛 개체용 프로필 비교

맵렛이나 매핑의 두 개체를 분석하고 해당 개체에 대해 열 프로필 결과를 비교할 수 있는 프로필을 생성할 수 있습니다.

단일 매핑 또는 맵렛 개체의 열 프로필과 마찬가지로 프로필 비교도 모든 데이터 열에서 실행되며 데이터 개체용으로 준비된 데이터에 대한 드릴다운 작업이 활성화됩니다. 소스 테이블의 데이터를 대상 테이블로 이동한 후 프로필을 비교하여 데이터 마이그레이션을 확인할 수 있습니다. 또한 시간이 지남에 따라 변경되는 데이터 소스의 프로필을 비교할 수도 있습니다.

1. 맵렛 또는 매핑을 엽니다.
2. 맵렛 또는 매핑이 유효한지 확인합니다.
3. **Ctrl** 키를 누르고 편집기에서 두 개체를 클릭합니다.
4. 개체 중 하나를 마우스 오른쪽 단추로 클릭하고 **프로필 비교**를 선택합니다.
5. 필요에 따라 개체 간에 열이 일치하는지를 확인하도록 프로필 비교를 구성합니다.
6. 필요에 따라 개체 하나에서 열을 클릭하여 다른 개체의 열로 끄는 방법으로 열 일치를 확인합니다.
7. 필요에 따라 프로필이 모든 열을 분석하는지 아니면 일치하는 열만 분석하는지를 선택합니다.
8. **확인**을 클릭합니다.

프로필에서 매핑 생성

프로필에서 매핑 개체를 작성할 수 있습니다. 작성한 매핑 개체를 사용하여 유효한 매핑을 개발합니다. 작성한 매핑은 프로파일링된 개체를 기반으로 하는 데이터 소스를 포함하며 프로필 규칙 논리를 기반으로 하는 변환을 포함할 수 있습니다. 매핑을 작성한 후 개체를 추가하여 매핑을 완성합니다.

1. **개체 탐색기** 보기에서 매핑을 작성할 프로필을 찾습니다.
 2. 프로필 이름을 마우스 오른쪽 단추로 클릭하고 **매핑 생성**을 선택합니다.
매핑 생성 대화 상자가 표시됩니다.
 3. 매핑 이름을 입력합니다. 필요에 따라 매핑의 설명을 입력합니다.
 4. 매핑의 폴더 위치를 확인합니다.
기본적으로 **Developer** 도구는 프로필과 같은 프로젝트의 **매핑** 폴더에 매핑을 작성합니다. 매핑에 대해 다른 위치를 선택하려면 **찾아보기**를 클릭합니다.
 5. **Developer** 도구가 매핑을 작성하는 데 사용하는 프로필 정의를 확인합니다. 다른 프로필을 사용하려면 **프로필 선택**을 클릭합니다.
 6. **마침**을 클릭합니다.
매핑이 **개체 탐색기**에 표시됩니다.
- 매핑에 개체를 추가하여 매핑을 완성합니다.

제 13 장

Informatica Developer의 열 프로파일 결과

이 장에 포함된 항목:

- [Informatica Developer의 열 프로파일 결과, 97](#)
- [열 값 속성, 98](#)
- [열 패턴 속성, 98](#)
- [열 통계 속성, 98](#)
- [열 데이터 유형 속성, 99](#)
- [Informatica Developer의 큐레이션, 100](#)
- [Informatica Developer에서 프로파일 결과 내보내기, 101](#)

Informatica Developer의 열 프로파일 결과

열 프로파일 분석은 값 빈도, 패턴 및 데이터 통계를 강조 표시하여 데이터 품질에 대한 정보를 제공합니다.

다음 테이블에서는 각 분석 유형에 대한 프로파일 결과를 설명합니다.

프로파일 유형	프로파일 결과
열 프로파일	<ul style="list-style-type: none">- 고유한 값과 Null 값에 대한 백분율 및 수 통계- 유추된 데이터 유형- 데이터 소스가 데이터에 대해 선언하는 데이터 유형- 최대값 및 최소값- 가장 최근의 프로파일 실행 날짜 및 시간- 열의 각 고유 데이터 요소에 대한 백분율 및 수 통계- 열의 각 고유 문자 패턴에 대한 백분율 및 수 통계
기본 키 프로파일	<ul style="list-style-type: none">- 유추된 기본 키- 키 위반
함수 종속성 프로파일	<ul style="list-style-type: none">- 유추된 함수 종속성- 함수 종속성 위반

열 값 속성

열 값 속성에는 프로파일링된 열의 값과 각 값이 각 열에 나타나는 빈도가 표시됩니다. 빈도는 숫자, 백분율 및 막대 차트로 표시됩니다.

열 값 속성을 보려면 **표시** 목록에서 값을 선택합니다. 열 값을 두 번 클릭하여 값이 포함된 행으로 드릴다운합니다.

다음 테이블에는 열 값에 대한 속성이 설명되어 있습니다.

속성	설명
값	프로필에 있는 열에 대한 모든 값의 목록입니다.
빈도	값이 열에 표시되는 횟수입니다.
백분율	열에서 값이 표시되는 횟수로, 열에 포함된 모든 값의 백분율로 표현됩니다.
차트	백분율을 나타내는 막대 차트입니다.

열 패턴 속성

열 패턴 속성에는 프로파일링된 열의 데이터 패턴과 패턴이 각 열에 나타나는 빈도가 표시됩니다. 패턴은 숫자, 백분율 및 막대 차트로 표시됩니다.

패턴 정보를 보려면 **표시** 목록에서 패턴을 선택합니다. 패턴을 두 번 클릭하여 패턴이 포함된 행으로 드릴다운합니다.

다음 테이블에는 열 값 패턴에 대한 속성이 설명되어 있습니다.

속성	설명
패턴	선택한 열에 대한 패턴입니다.
빈도	패턴이 열에 표시되는 횟수입니다.
백분율	열에서 패턴이 표시되는 횟수가 열에 포함된 모든 값의 백분율로 표현됩니다.
차트	백분율에 대한 막대 차트입니다.

열 통계 속성

열 통계에는 값의 최대 길이/최소 길이와 첫 번째 값/마지막 값 등의 속성이 포함됩니다.

통계 정보를 보려면 **표시** 목록에서 통계를 선택합니다.

다음 테이블에는 열 통계 속성이 설명되어 있습니다.

속성	설명
최대 길이	열에서 가장 긴 값의 길이입니다.
최소 길이	열에서 가장 짧은 값의 길이입니다.
하위	열의 마지막 5개 값입니다.
상위	열의 처음 5개 값입니다.
합계	숫자 데이터 유형인 열에서 모든 값의 합계입니다.

참고: 프로필에는 정수 유형 열에 대한 평균 및 표준 편차 통계도 표시됩니다.

열 데이터 유형 속성

열 데이터 유형에는 프로필 결과의 각 열에 대해 유추된 모든 데이터 유형이 포함됩니다.

데이터 유형 정보를 보려면 **표시** 목록에서 **데이터 유형**을 선택합니다. 데이터 유형을 두 번 클릭하여 데이터 유형이 포함된 행으로 드릴다운합니다.

다음 테이블에는 열 데이터 유형에 대한 속성이 설명되어 있습니다.

속성	설명
데이터 유형	프로필에 있는 열에 대해 모든 유추된 데이터 유형의 목록입니다.
빈도	열에 대해 데이터 유형이 표시되는 횟수가 숫자로 표시됩니다.
준수율(%)	열에 대해 데이터 유형이 표시되는 백분율입니다.
상태	<p>데이터 유형의 상태를 나타냅니다. 상태는 유추됨, 승인됨 또는 거부됨입니다.</p> <p>유추됨</p> <p>Developer tool이 유추한 열의 데이터 유형을 나타냅니다.</p> <p>승인됨</p> <p>열에 대해 승인된 데이터 유형을 나타냅니다. 데이터 유형을 승인하면 해당 데이터 유형이 모델 리포지토리에 커밋됩니다.</p> <p>거부됨</p> <p>열에 대해 거부된 데이터 유형을 나타냅니다.</p>

Informatica Developer의 큐레이션

기간은 데이터 소스에서 검색된 메타데이터를 사용하고 보고할 수 있도록 해당 메타데이터의 유효성을 검사하고 관리하는 프로세스입니다. Informatica Developer에서 메타데이터를 큐레이션할 때 프로필 결과의 유추된 데이터 유형 또는 데이터 도메인을 승인, 거부 및 재설정할 수 있습니다.

열에 대해 하나의 데이터 유형 또는 데이터 도메인을 승인할 수 있습니다. 열에 대해 거부된 데이터 유형 또는 데이터 도메인을 숨길 수 있습니다. 유추된 데이터 유형 또는 데이터 도메인을 승인하거나 거부한 후에 데이터 유형 또는 데이터 도메인을 재설정하여 유추된 상태를 복원할 수 있습니다.

데이터 유형 승인

프로필 결과에는 데이터 소스의 각 열에 대한 유추 상태, 준수 백분율, 빈도 및 유추된 데이터 유형이 포함됩니다. 각 열에 대해 단일 데이터 유형을 선택 및 승인할 수 있습니다.

1. **Object Explorer** 보기에서 프로필을 선택하여 엽니다.
2. **결과** 탭에 있는지 확인합니다.
3. **열 프로파일링** 보기에서 열을 선택하여 오른쪽 패널에서 값 빈도, 패턴, 데이터 유형 및 통계를 확인합니다.
4. **세부 정보** 패널의 **표시** 목록에서 **데이터 유형**을 선택합니다.
열의 유추된 데이터 유형이 표시됩니다.
5. 승인할 열을 마우스 오른쪽 단추로 클릭하고 **승인**을 클릭합니다.
데이터 유형의 상태가 **승인됨**으로 변경됩니다.
6. 데이터 유형의 유추된 상태를 복원하려면 데이터 유형을 마우스 오른쪽 단추로 클릭하고 **재설정**을 클릭합니다.

데이터 유형 거부

Informatica Developer에서는 유추된 데이터 유형이 프로필 결과에 기본적으로 표시됩니다. 유추되거나 승인된 데이터 유형을 거부할 수 있습니다. 거부된 데이터 유형을 표시하거나 숨기도록 선택할 수도 있습니다.

1. **Object Explorer** 보기에서 프로필을 선택합니다.
2. 프로필을 두 번 클릭하여 엽니다.
프로필이 탭에 열립니다.
3. **열 프로파일링** 보기에서 행을 선택합니다.
4. 유추된 열 데이터 유형을 거부하려면 오른쪽 패널에서 **데이터 유형** 보기를 선택합니다. 거부하려는 유추된 데이터 유형을 선택하고 행을 마우스 오른쪽 단추로 클릭한 후 **거부**를 선택합니다.
Informatica Developer에서 데이터 유형 목록에는 거부된 데이터 유형이 회색으로 표시됩니다.
5. 거부된 데이터 유형을 숨기려면 행을 마우스 오른쪽 단추로 클릭하고 **거부된 항목 숨기기**를 선택합니다.
6. 거부된 데이터 유형을 보려면 행 중 하나를 마우스 오른쪽 단추로 클릭하고 **거부된 항목 표시**를 선택합니다.

Informatica Developer에서 프로필 결과 내보내기

열 프로필 결과를 .csv 파일 또는 Microsoft Excel 파일로 내보낼 수 있습니다. 프로필 결과를 Microsoft Excel 파일로 내보내면 Developer tool이 정보를 .xlsx 파일에 저장합니다.

1. **Object Explorer** 보기에서 프로필을 엽니다.
2. 필요에 따라 프로필을 실행하여 프로필 결과를 업데이트합니다.
3. **결과** 보기를 선택합니다.
4. 열을 선택합니다.
5. **세부 정보**에서 **값**, **패턴** 또는 **데이터 유형**을 선택하고 **내보내기** 아이콘을 클릭합니다.
파일로 데이터 내보내기 대화 상자가 열립니다.
6. 기본 파일 이름을 적용하거나 변경합니다.
7. 내보낼 데이터의 유형을 선택합니다. **선택한 열에 대한 값**, **선택한 열에 대한 패턴**, **선택한 열에 대한 데이터 유형** 또는 **모두(요약, 값, 패턴, 데이터 유형, 통계, 속성)**를 선택할 수 있습니다.
8. **찾아보기**를 클릭하여 위치를 선택하고 파일을 컴퓨터에 로컬로 저장합니다.
9. 필드 이름을 첫 번째 행으로 내보내지 않으려면 **필드 이름을 첫 번째 행으로 내보내기** 확인란의 선택을 취소합니다.
10. **확인**을 클릭합니다.

제 14 장

Informatica Developer의 성과 기록표

이 장에 포함된 항목:

- [Informatica Developer의 성과 기록표 개요, 102](#)
- [성과 기록표 작성, 102](#)
- [성과 기록표 연계의 리소스 파일 내보내기, 103](#)
- [Informatica Developer에서 성과 기록표 연계 보기, 103](#)

Informatica Developer의 성과 기록표 개요

성과 기록표는 프로필의 품질 측정을 그래픽으로 표현합니다. Developer 도구에서 성과 기록표를 볼 수 있습니다. Developer 도구에서 성과 기록표를 작성한 후 Analyst 도구에 연결하여 편집을 위해 성과 기록표를 열 수 있습니다. 프로파일링 웨어하우스의 준비된 데이터 또는 데이터 개체의 현재 데이터에서 성과 기록표를 실행합니다.

Analyst 도구에서 성과 기록표를 편집/실행하고 메트릭 또는 메트릭 그룹에 대한 성과 기록표 연계를 확인할 수 있습니다.

성과 기록표 작성

성과 기록표를 작성하고 프로필의 열을 성과 기록표에 추가합니다. 먼저 프로필을 실행한 다음 성과 기록표에 열을 추가해야 합니다.

1. **개체 탐색기** 보기에서 성과 기록표를 작성할 프로젝트 또는 폴더를 선택합니다.
2. **파일 > 새로 만들기 > 성과 기록표**를 클릭합니다.
새 **성과 기록표** 대화 상자가 나타납니다.
3. **추가**를 클릭합니다.
프로필 선택 대화 상자가 표시됩니다. 추가할 열이 있는 프로필을 선택합니다.
4. **확인**과 **다음**을 차례로 클릭합니다.
5. 성과 기록표에 추가할 열을 선택합니다.

기본적으로 성과 기록표 마법사는 프로필에 정의된 열과 규칙을 선택합니다. 프로필에 포함되지 않은 열은 추가할 수 없습니다.

6. **마침**을 클릭합니다.

Developer 도구에서 성과 기록표를 작성합니다.

7. 필요에 따라 **Informatica Analyst로 열기**를 클릭하여 Analyst 도구에 연결한 다음 Analyst 도구에서 성과 기록표를 엽니다.

성과 기록표 연계의 리소스 파일 내보내기

성과 기록표 및 종속 개체가 포함된 프로젝트를 Metadata Manager의 리소스 파일로 내보낼 수 있습니다. XML 형식으로 내보낸 리소스 파일을 사용하여 Metadata Manager에서 성과 기록표 연계의 리소스를 작성하고 로드합니다.

1. **내보내기** 마법사를 열려면 **파일 > 내보내기**를 클릭합니다.
2. **Informatica > Metadata Manager**의 리소스 파일을 선택합니다.
3. **다음**을 클릭합니다.
4. **찾아보기**를 클릭하여 내보내야 하는 성과 기록표 개체 및 연계가 포함된 프로젝트를 선택합니다.
5. **다음**을 클릭합니다.
6. 내보낼 성과 기록표 개체를 선택합니다.
7. 내보내기 파일 이름 및 파일 위치를 입력합니다.
8. 선택한 개체와 함께 **내보내기** 마법사가 내보내는 종속 개체를 보려면 **다음**을 클릭합니다.
내보내기 마법사에 종속 개체가 표시됩니다.
9. **마침**을 클릭합니다.
Developer 도구가 개체를 XML 파일로 내보냅니다.

Informatica Developer에서 성과 기록표 연계 보기

Developer 도구에서 메트릭 또는 메트릭 그룹에 대한 성과 기록표 연계를 보려면 Analyst 도구를 실행합니다.

1. **개체 탐색기** 보기에서 성과 기록표가 포함된 프로젝트 또는 폴더를 선택합니다.
2. 성과 기록표를 두 번 클릭하여 엽니다.
성과 기록표가 탭에 표시됩니다.
3. **Informatica Analyst로 열기**를 클릭합니다.
Analyst 도구가 브라우저 창에서 열립니다.
4. Analyst 도구의 **성과 기록표** 보기에서 메트릭 또는 메트릭 그룹을 선택합니다.
5. 마우스 오른쪽 단추를 클릭하고 **연계 표시**를 선택합니다.
성과 기록표 연계 다이어그램이 대화 상자에 표시됩니다.

인덱스

A

Analyst 도구의 열 프로필 결과
열 세부 정보 [44, 56](#)
요약 [42](#)
인터페이스 [43, 54, 56](#)

B

프로파일링
도구 [12](#)
잠금 및 버전 관리 [15](#)
프로세스 [12](#)
프로필
Avro 또는 Parquet 형식 [88](#)
XML 및 JSON 형식 [87, 88](#)
구성 요소 [12](#)
소개 [11](#)
실행 [27, 53, 54](#)
열 프로필 작성 [25](#)
열 프로필 편집 [26](#)
필터 작성 [36](#)
필터 추가 [39](#)
프로필 결과
Developer tool의 데이터 유형 거부 [100](#)
Excel [60](#)
Informatica Analyst에서 내보내기 [60](#)
Informatica Developer에서 내보내기 [101](#)
Informatica Developer에서 데이터 유형 승인 [100](#)
내보내기 [59](#)
데이터 유형 거부 [59](#)
데이터 유형 승인 [58](#)
드릴다운 [58](#)
상세 보기 [45](#)
열 값 [49](#)
열 데이터 유형 [47, 99](#)
열 패턴 [48](#)
요약 [55, 57](#)
요약 보기 [43](#)
플랫 파일 데이터 개체
동기화 [28](#)
필터
개요 [36](#)

I

Informatica Analyst
규칙 [31](#)
열 프로필 개요 [20, 54](#)
열 프로필 결과 [41, 52](#)
잠금 및 버전 관리 [25](#)
Informatica Developer
규칙 [93](#)

N

내보내기
XML에 대한 성과 기록표 연계 [103](#)
데이터 개체 프로필
단일 프로필 작성 [81](#)
여러 프로필 작성 [82](#)
런타임 환경
Analyst 도구 [23](#)
Hadoop [23, 24](#)
매핑 개체
프로필 실행 [95](#)
맷 및 매핑 프로파일링
개요 [95](#)
미리 정의된 규칙
프로세스 [32](#)
성과 기록표
Informatica Analyst [62](#)
Informatica Analyst 프로세스 [63](#)
Informatica Developer [102](#)
가변 비용 [68](#)
개요 [15](#)
경향 차트 [70](#)
고정 비용 [68](#)
드릴다운 [70](#)
메트릭 [67](#)
메트릭 가중치 [67](#)
메트릭 그룹 [68](#)
메트릭 그룹 삭제 [70](#)
메트릭 그룹 작성 [69](#)
메트릭 그룹 편집 [69](#)
보기 [66](#)
성과 기록표에 열 추가 [65](#)
실행 [66](#)
알림 [75](#)
알림 구성 [77](#)
올바르지 않은 데이터 비용 [68](#)
임계값 정의 [68](#)
전역 알림 설정 구성 [77](#)
점수 이동 [69](#)
편집 [66](#)
성과 기록표 결과
Excel로 내보내기 [75](#)
Informatica Analyst에서 내보내기 [75](#)
내보내기 [74](#)
성과 기록표 연계
Informatica Analyst에서 보기 [78](#)
Informatica Developer에서 보기 [103](#)
식 규칙 작성
규칙 [33](#)
열 프로필
Informatica Developer [80](#)
개요 [14](#)
드릴다운 [57](#)
옵션 [15](#)

열 프로필 (계속)
 운영 체제 프로필 [25](#)
 프로세스 [21](#)
열 프로필 결과
 Informatica Developer [97](#)
열 프로필 작성
 프로필 [25](#)
이상값
 감지 [48](#)
큐레이션
 Informatica Analyst [58](#)
 Informatica Developer [100](#)
 개념 [17](#)
 태스크 [17](#)
테이블 데이터 개체
 동기화 [30](#)

S

Sqoop 구성
 프로파일링 [24](#)

T

경향 차트
 Informatica Analyst에서 내보내기 [73](#)
 보기 [72](#)
 비용 [71](#)
 점수 [71](#)
규칙
 Informatica Developer에서 생성 [94](#)
 Informatica Developer에서 적용 [94](#)
 규칙 사양을 사용하여 식 규칙 생성 [34](#)
 미리 정의된 규칙 적용 [32](#)
 미리 정의됨 [31](#)
 식 [32](#)
 식 규칙 작성 [33](#)