

2025년 제2회 H-iDEA

HERO Guide book

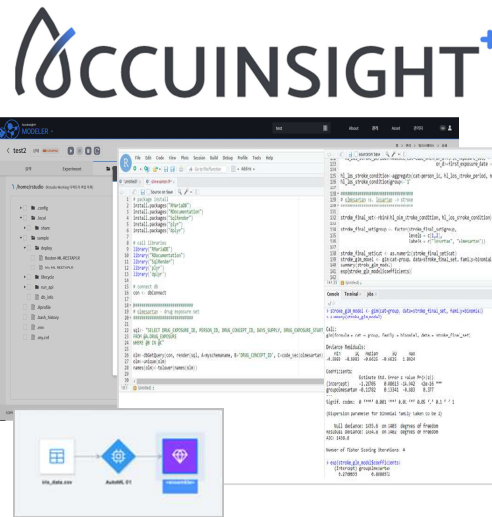
HERO & 분석솔루션

HERO



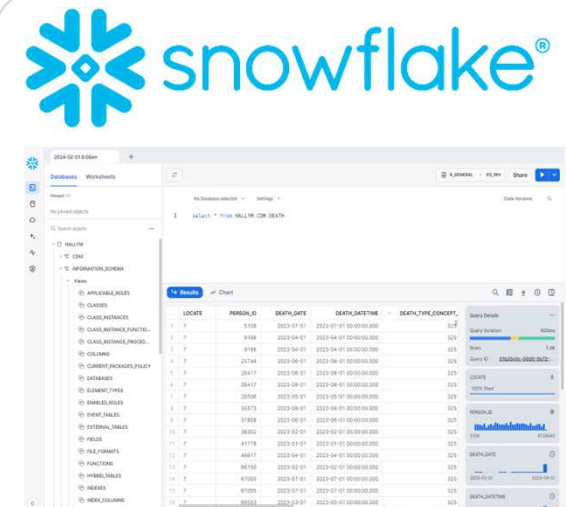
- 한림대학교의료원 차세대 연구 분석 포털 (2022.03.02 v1.0 오픈)
- Harmonic intEgrated Research platfOrm
- Data Lake : OCS/EMR, CDM, K-CURE, 임상특화 DB, PACS 등의 Data 탑재
- Hybrid Cloud 환경 (Private & Public)
- 개인정보삭제 및 비식별화 적용
- 경진대회용 가명DB 심의/개방

AccuInsight+3.5



- SK(주)의 클라우드 기반 다기능 분석솔루션
- 워크플로우를 통한 객체형 Task 코딩
- HDFS, Local file, DB 소스 연결 가능
- AutoML 기능으로 초보자 범용 AI 가능
- Sandbox를 통해 R, Python 코딩 환경
- 다기관 연계 분석 가능

Snowflake

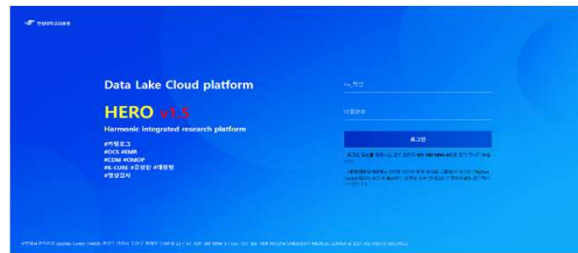


- AWS Public SaaS 설치 Public RDBMS
- 컴퓨팅 리소스 관리 용이
- DB/Schema/Table/Column Role 정의 가능(부분 공유, 읽기/쓰기 등)
- On demand 형식의 유연한 과금 정책
- Data를 안전하게 로드 및 사용자 간 공유
- 금융/의료/교육/경영 등 7200여개 Reference
- 현재 가장 Hot한 솔루션 중 하나

HERO 기본 사용법

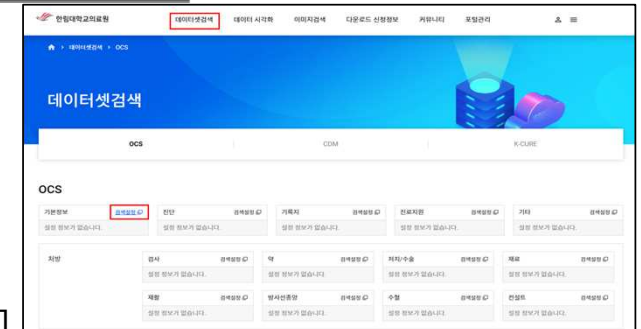
※ **가산포인트** 과정을 완수할 시 **가산점** 부여

1. 접속하기



- www.hero.re.kr로 접속
- ID: hu_학번/ PW: 생년월일(6자리)
- 패스워드 반드시 변경 후 재로그인
- Chrome 또는 Edge 브라우저 이용

2. 데이터셋 검색하기



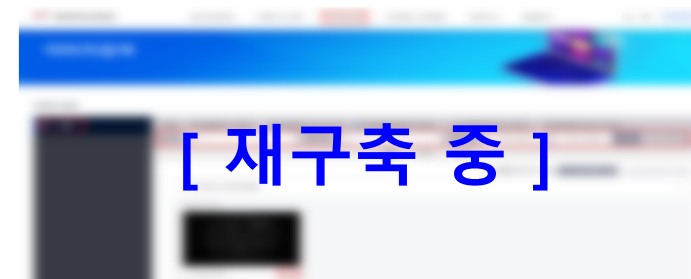
- 상단메뉴 {데이터셋검색} 클릭
- [검색설정] 클릭하여 기간 [검색]
- <미리보기> 확인 및 [전체 다운로드 신청] 버튼 선택

3. 데이터셋 다운로드 신청



- [사용 용도] 입력
- [Snowflake 체크박스]를 체크 시 Snowflake로 전송 가능 옵션(선택 요망)
- 위 옵션 체크 하지 않을 시 .csv파일로만 다운로드
- 접수/승인 후 다운로드 가능

4. 이미지 검색하기



HERO to Snowflake Pipeline 방법

1. 데이터셋 다운로드 신청

The screenshot shows the '데이터셋 신청' (Dataset Request) form. Under the '신청서 내용' (Request Content) section, there is a checkbox labeled 'Snowflake Upload'. This checkbox is highlighted with a red box, indicating it should be selected for the upload process.

- 데이터셋 검색 후 다운로드 신청에서 [Snowflake] 체크박스를 선택

2. 다운로드 승인

The screenshot shows the '다운로드 신청정보' (Download Request Information) page. At the top, there is a navigation bar with a button labeled '승인' (Approve), which is highlighted with a red box. Below the header, there is a table listing the download requests, with the '승인' column highlighted.

- 상단메뉴 [다운로드 신청정보]
- 데이터셋 신청 목록 중 승인 목록 선택

3. Snowflake Upload

The screenshot shows the 'Snowflake Upload' form. At the bottom of the form, there are two buttons: 'Snowflake Upload' and '확인' (Confirm). The 'Snowflake Upload' button is highlighted with a red box, indicating it should be clicked to proceed with the upload.

- [Snowflake Upload] 버튼 선택

4. 데이터 업로드

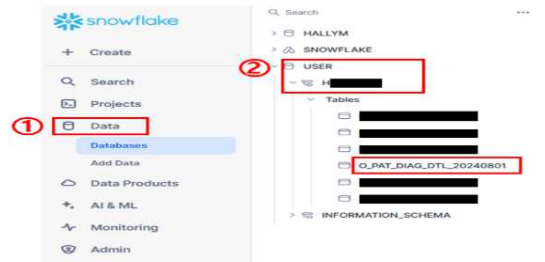
가산포인트

The screenshot shows the 'Snowflake Upload' form. At the bottom of the form, there is a button labeled '확인' (Confirm), which is highlighted with a red box. Above the button, there is a table with columns for '데이터명' (Data Name), '데이터형식' (Data Format), and '데이터크기' (Data Size). The table contains three rows of data, with the first row highlighted.

- 전송 컬럼 아래 체크박스 [체크] 후 [실행]버튼 클릭
- 전송 실행 완료 후 화면 종료 가능
- 테이블(신규,기존) 명 변경 가능 중복에 Y체크 시 필수 변경
- 테이블명은 영문 대문자,숫자,언더바(_)만 사용 가능, 첫 글자 영문 대문자

HERO to Snowflake Pipeline 방법

5. 데이터 확인하기



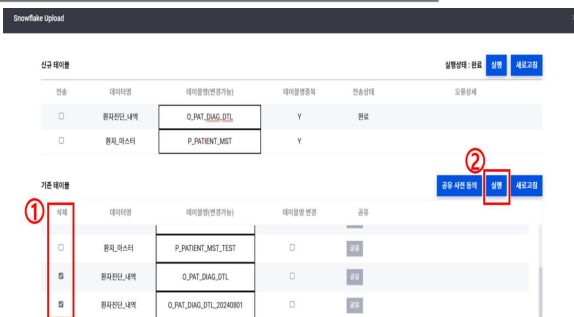
- 상단 메뉴 {데이터 분석 솔루션} > Snowflake 접속 후
- Data>Databases>USER>schema (H+hu_학번)>입력한 테이블 명

6. 기존 테이블 명 변경



- 기존 테이블 목록에서 변경할 테이블명 변경 체크 박스 선택 후
- 테이블 명 변경 후 실행(여러 테이블 동시 선택 가능)
- 테이블 명은 영문 대문자,숫자,언더바(_)만 사용 가능, 첫 글자 영문 대문자

7. 기존 테이블 삭제



- 기존 테이블 목록에서 [삭제] 체크 박스 선택 후 [실행] (여러 테이블 동시 선택-실행 가능)

8-1. 테이블 공유

가산포인트

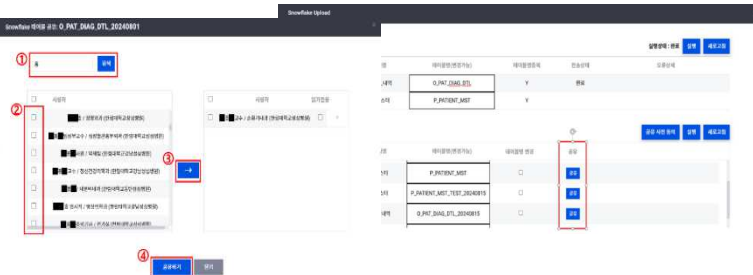


- 자신의 Snowflake 테이블을 다른 사용자가 접근 할 수 있도록 하는 기능
- [공유 사전 동의] 버튼 선택 후 공유 버튼 활성화

HERO to Snowflake Pipeline 방법

8-2. 테이블 공유

가산포인트



- 공유 사전 등의 후 [공유] 버튼 클릭(공유는 개별 테이블 별로 가능)
- 사용자 성명으로 [검색] 후 해당 사용자를 [체크] 후 우측 화살표로 이동

8-3. 테이블 공유

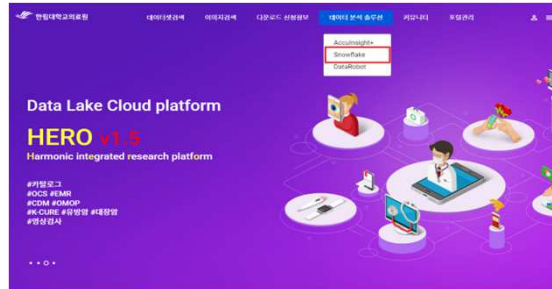
가산포인트



- [읽기전용] 체크박스 선택 시 select(읽기) 권한만 부여
- 체크하지 않으면 읽기, 쓰기, 삭제(select, insert, update, delete) 권한이 부여
- 공유 권한을 회수하려면 오른쪽 사용자 목록에서 [X] 버튼 클릭 후 [공유하기] 버튼 클릭하여 반영

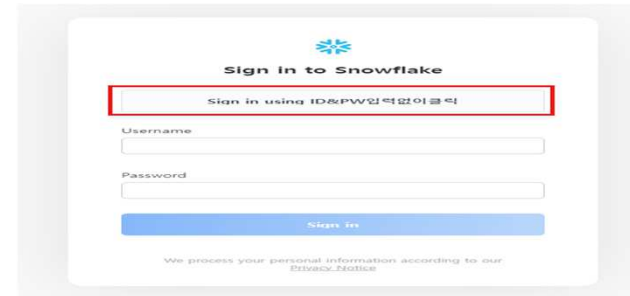
Snowflake 기본 사용법

1. 접속하기



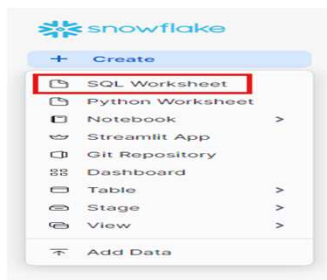
- www.hero.re.kr로 접속 & 로그인 후
- 상단메뉴 {데이터 분석 솔루션} > Snowflake

2. SSO 로그인



- [Sign In Using ID&PW입력없이클릭] Click

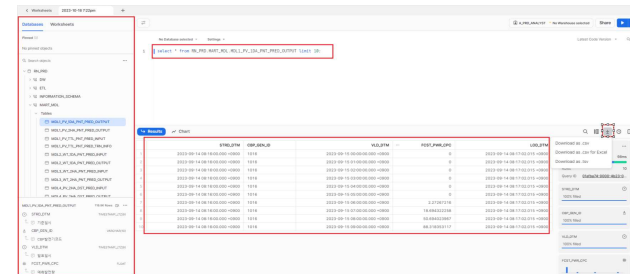
3. SQL Worksheet 생성



- 왼쪽 상단 +Create
- SQL Worksheet 선택

4. 데이터 조회

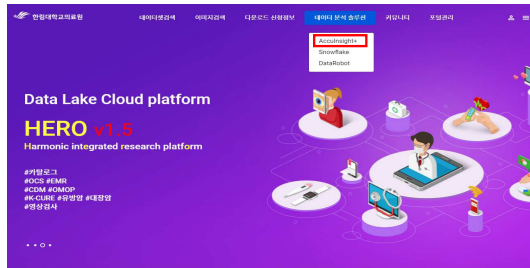
가산포인트



- 왼쪽 메뉴에 데이터베이스, 스키마, 테이블 구조 확인 가능
- SQL로 데이터 핸들링
- 결과 다운로드 가능

Acculnsight+ 기본 사용법

1. 접속하기



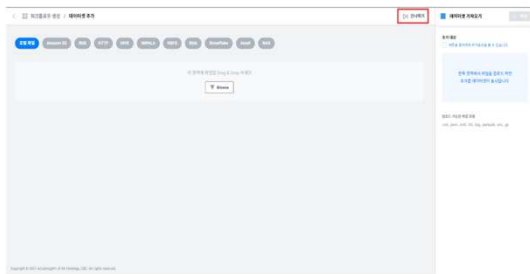
- www.hero.re.kr로 접속 & 로그인 후
- 상단메뉴 {데이터 분석 솔루션} > Acculnsight+

2. 워크플로우 생성



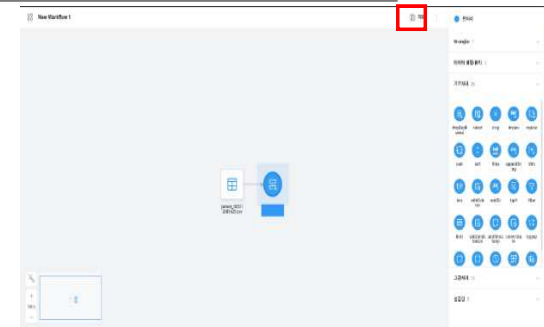
- 상단 {프로젝트} 메뉴 or 중앙화면 [프로젝트 생성] 후
- [워크플로우] 생성 후 [다음]

3. 데이터셋 가져오기



- 워크플로우의 [데이터 셋 추가] (건너뛰기 가능)
- [로컬 파일] Browse로 csv 로드 후 우측 상단 [추가] → 로드 완료

4. 전처리

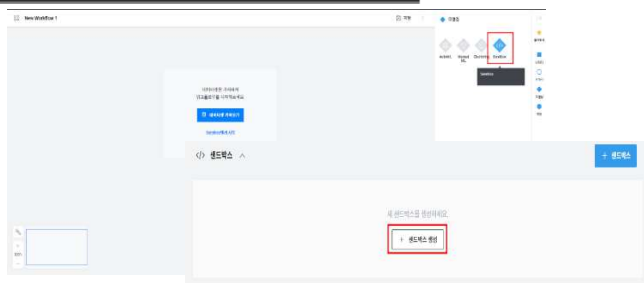


- 우측 패널의 전처리 항목 설정 후 [저장]

Acculnsight+ 기본 사용법

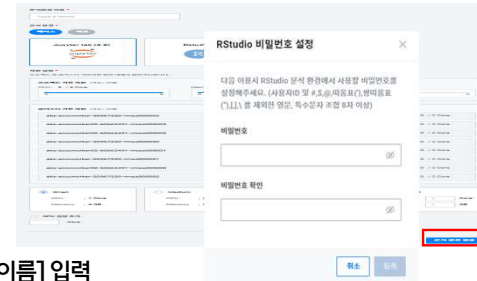
5-1. 모델링

가산포인트



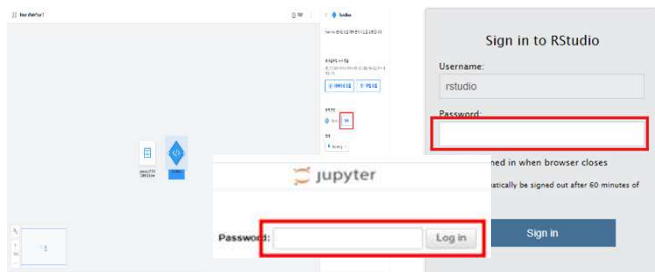
- 전처리 후 모델링에서 Sandbox 선택
- 또는 프로젝트에서 [샌드박스] 생성 가능

5-2. Sandbox 생성



- [분석환경 이름] 입력
- 분석환경(Jupyter lab 또는 Rstudio) 선택 후 [분석 환경 생성] 클릭
- 비밀번호 설정

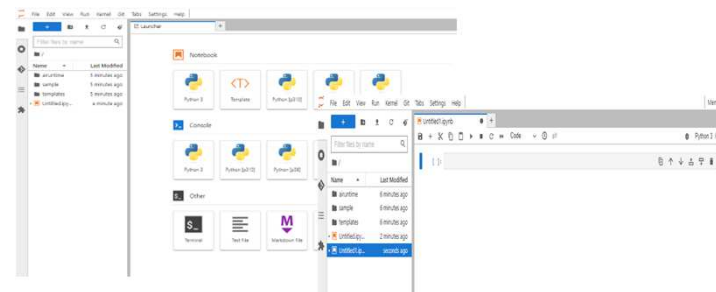
6. Sandbox 접속



- 본인 샌드박스 선택 후 [접속] 클릭
- 본인 설정한 비밀번호 [입력]

7. 분석 환경 활용

가산포인트



- 노트(코딩) 환경 선택
- 클라우드 환경 코딩

문의처 (한림대학교의료원 BigData Center)

[경진대회 전반]

장수연 031-380-1805 suyeon@hallym.or.kr

[HERO, Acculnisght+, Snowflake 사용]

김윤희 031-380-1947 yoonhee@hallym.or.kr