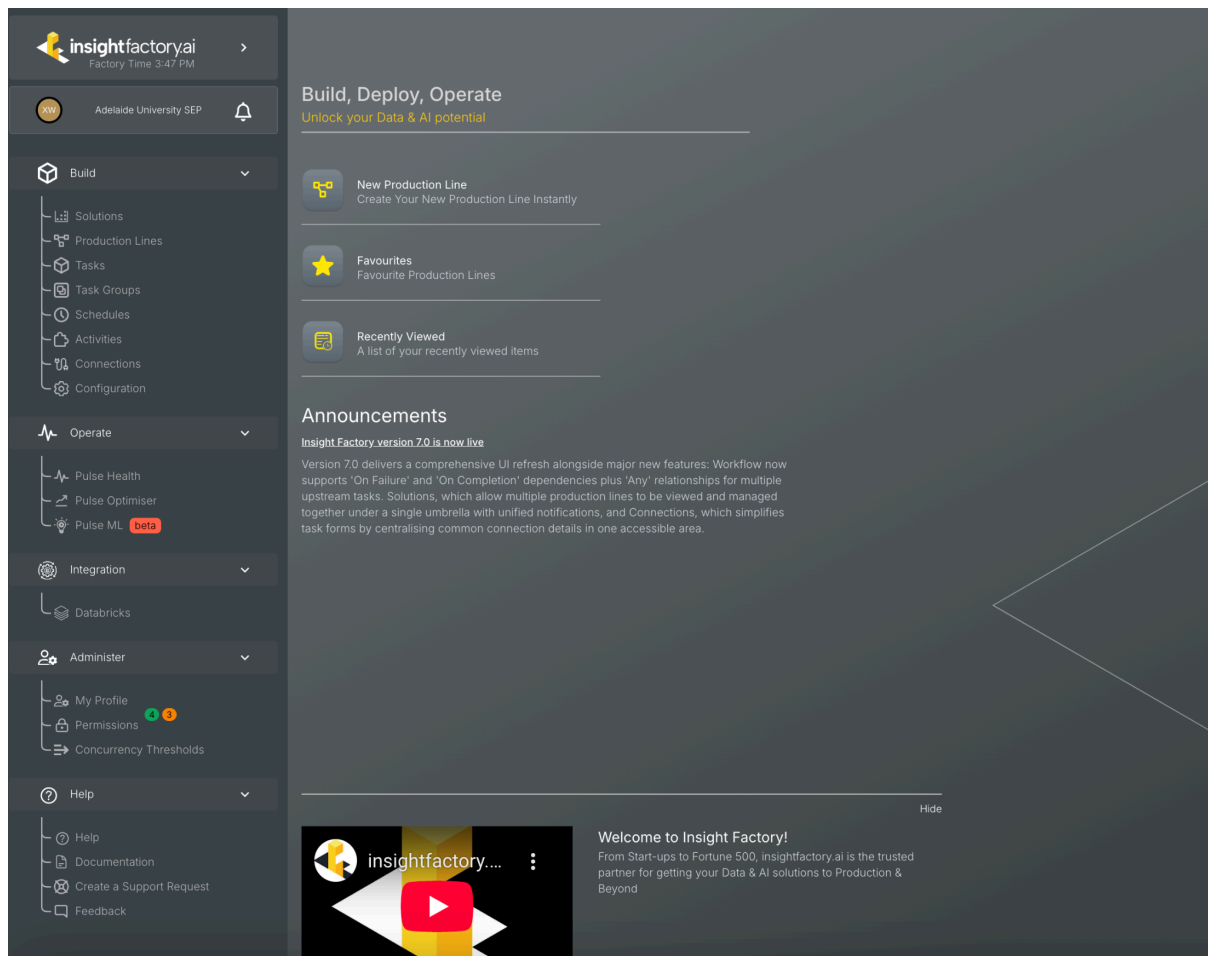


Home Page of Insight Factory.ai platform



The home page mainly consists of five modules: Build, Operate, Integration, Administer, and Help.

Build

This is the core area used to build and configure the AI production process.

- **Solutions**
A collection of solutions that enables bundling multiple Production Lines for unified management.
- **Production Lines**
The specific implementation of AI production lines. For example, a production line may consist of the workflow: Data Preprocessing → Feature Extraction → Model Training → Deployment.

- **Tasks**
Individual execution steps, such as data cleaning, training tasks, and inference tasks.
- **Task Groups**
A logical group formed by combining multiple related tasks to facilitate management.
- **Schedules**
Scheduling tasks that set up automatic execution times.
- **Activities**
Activity records used to view the operation status of tasks/production lines in the system (similar to execution logs).
- **Connections**
Data connectors that centrally manage external connections such as databases, cloud storage, and third-party APIs.
- **Configuration**
The configuration center for the platform and tasks, including settings like runtime environments and parameter templates.

Operate

Used for health monitoring and optimization of production lines after they are put into operation.

- **Pulse Health**
Monitors the operation status and health of production lines (e.g., whether a task has failed, task duration, etc.).
- **Pulse Optimiser**
Analyzes and optimizes task performance to help identify bottlenecks.
- **Pulse ML (beta)**
Machine learning operation monitoring (beta version), which may include model performance tracking, data drift detection, etc.

Integration

- **Databricks**
Integration with the external big data/analytics platform (Databricks) to facilitate the connection between data processing and AI training.

Administer

Interface Settings and permission management, etc.

Help

Channels for problem feedback.

Data ingestion test

Production Lines

New production line

Production Lines

Filter

Search

Columns

Supported

Favourites

New Production Line

Current Status

Last Schedule Run

Product Line Permission Management

X

Properties

Status

Runs

Permissions

...

Protection

Remove Protection

This Production Line is currently protected.

Permissions

Email or Group Name

Permission Level

Select a user by email or a group

View

Add

Save

New Task - Data Ingestion

Task Production Line

Get started by adding/creating your **first task**

Create/Add Task

+ Create New ▶

+ Add Existing ▶

Activity Category

Ingest to Lakehouse ▶

Enrich Lakehouse Table ▶

Machine Learning ▶

Ingest Azure SQL DB as Parquet

Task settings

Code *

12 <= 30

M94AG872N5L0

Unique code (containing only letters, numbers, underscores and '-')

Name *

19 <= 100

ingest tonnage data

Unique name for this task.

Activity *

Ingest Azure SQL DB as Parquet

Task Group

Optionally select a task group

☒ Is Active?

Source

T Connection *

ReferenceData Database

Source connection to use.

T Extract Query *

```
1 SELECT * FROM
2   predictive_maintenance_uofa_2025.
3   tonnagedata
```

SQL query to extract data from the source database.

Data Lake

T Data Lake System Folder *

demo_xw

Name of the folder in the Data Lake that acts as the parent folder for all datasets belonging to this System.



T Data Lake Dataset Folder *

demo_xw_datasetfolder

Name of the folder in the Data Lake that the dataset will be stored under. Used with 'Data Lake System Folder' to form the fully qualified path to the dataset within the data Container in the Data Lake.



Lakehouse Table

T Catalog Name *

09ad024f-822f-48e4-9d9e-b5e03c1839a2

The name of the Catalog this transformation lives in. Required if Copy to Lakehouse Table is enabled.

A Copy to Lakehouse Table



Ingest directly to Lakehouse Table

T Schema Name

predictive_maintenance_uofa_2025

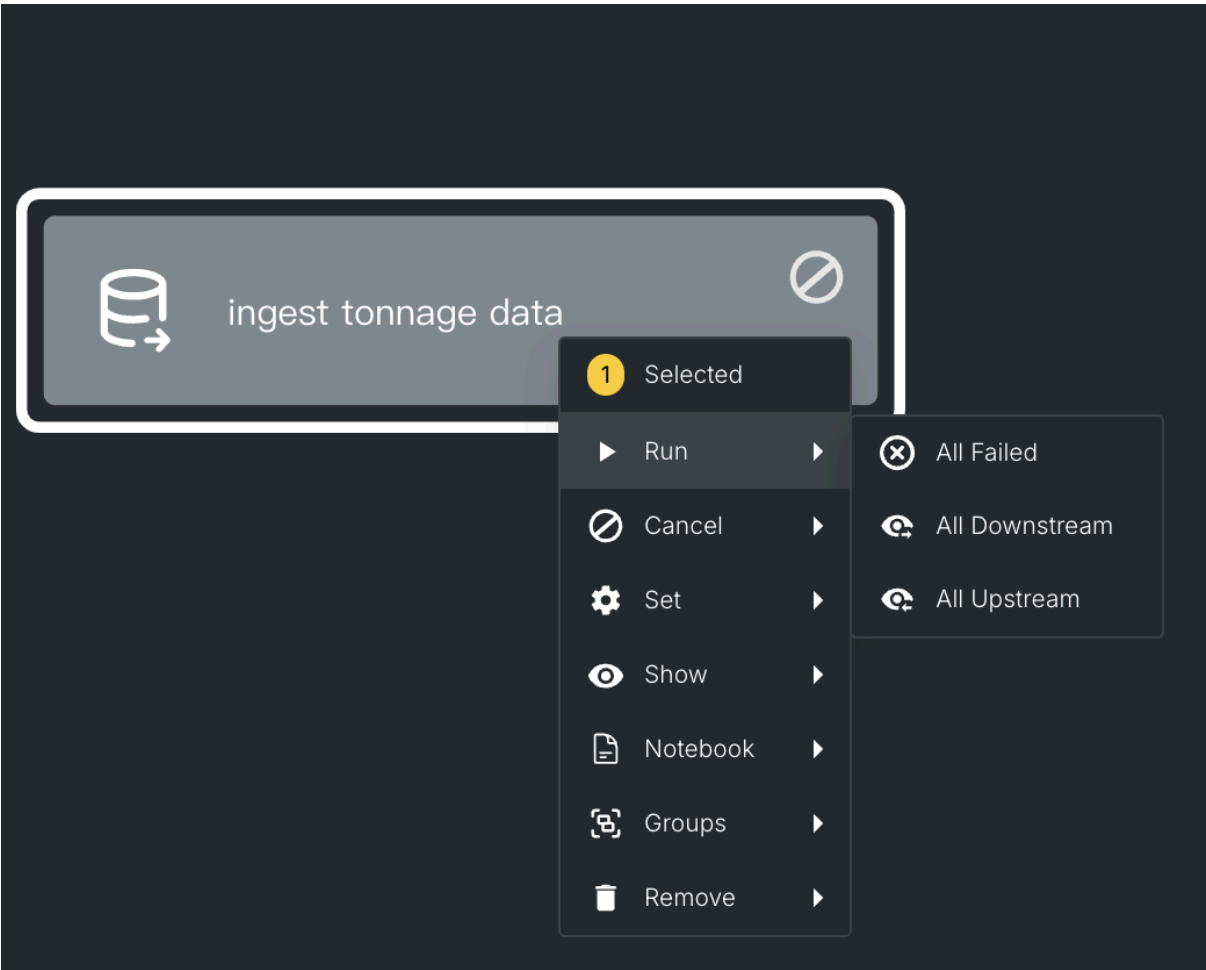
The name of the Schema this transformation lives in.

T Table Name


tonnagedata


The name of the Table representing this transformation.

Run




Result in Databricks

 Integration

 Databricks

Microsoft Azure



Search data, notebooks, recent, and more...

dbw-adunisecc-dev

New

Workspace

Recents

Catalog

Jobs & Pipelines

Compute

Marketplace

SQL

SQL Editor

Queries

Dashboards

Genie

Alerts

Query History

SQL Warehouses

Data Engineering

Job Runs

Data Ingestion

AI/ML

Playground

Experiments

Features

Models

Serving

Table of contents

Workspace/Users/a1912958@adelaide.edu.au/Test Notebook 2025-08-23 10:33:14

Last edit was 1 minute ago

A table of contents will be added here when a notebook has Markdown headings.

Heading 1

Test Notebook 2025-08-23 10:33:14

Run all

Student Cluster 4

Schedule

Share

1 minute ago (3s)

1

SQL

%sql

select * from `09ad024f-822f-48e4-9d9e-b5e03c1839a2`.`predictive_maintenance_uofa_2025`.`tonnagedata`

(1) Spark Jobs

_sqlidf: pyspark.sql.connect.dataframe.DataFrame = [BaseCode: string, SectionBreakStartKM: double ... 5 more fields]

Table

+

	BaseCode	SectionBreakStartKM	SectionBreakFinishKM	FromDate	ToDate	Tonnage	load_date
1	ARTC-12	198.06	198.08	01/07/2020	30/06/2021	52.67	20250
2	ARTC-12	198.66	198.68	01/07/2018	30/06/2019	55.2	20250
3	ARTC-12	199.04	199.06	01/07/2019	30/06/2020	55.9	20250
4	ARTC-12	199.32	199.34	01/07/2017	30/06/2018	53	20250
5	ARTC-12	200.44	200.46	01/07/2017	30/06/2018	53	20250
6	ARTC-12	201.74	201.76	01/07/2019	30/06/2020	55.9	20250
7	ARTC-12	201.78	201.8	01/07/2016	30/06/2017	54	20250
8	ARTC-12	208.74	208.76	01/07/2022	30/06/2023	45.40062023	20250
9	ARTC-12	208.82	208.84	01/07/2015	30/06/2016	53.8	20250
10	ARTC-12	209.7	209.72	01/07/2015	30/06/2016	53.8	20250
11	ARTC-12	212.32	212.34	01/07/2020	30/06/2021	175.58	20250
12	ARTC-12	215.64	215.66	01/07/2021	30/06/2022	173.85	20250
13	ARTC-12	216.06	216.08	01/07/2016	30/06/2017	180	20250
14	ARTC-12	218.24	218.26	01/07/2017	30/06/2018	53	20250
15	ARTC-12	219.8	219.82	01/07/2016	30/06/2017	54	20250

10,000+ rows

Truncated data

2.96s runtime

Refreshed 1 minute ago

This result is stored as _sqlidf and can be used in other Python and SQL cells.