

 REVISED 10-DAY SCHEDULE

Data Mining & Data Handling

Master analytics with **SQL**, **Python**, and **Snowflake**

SESSION DETAILS

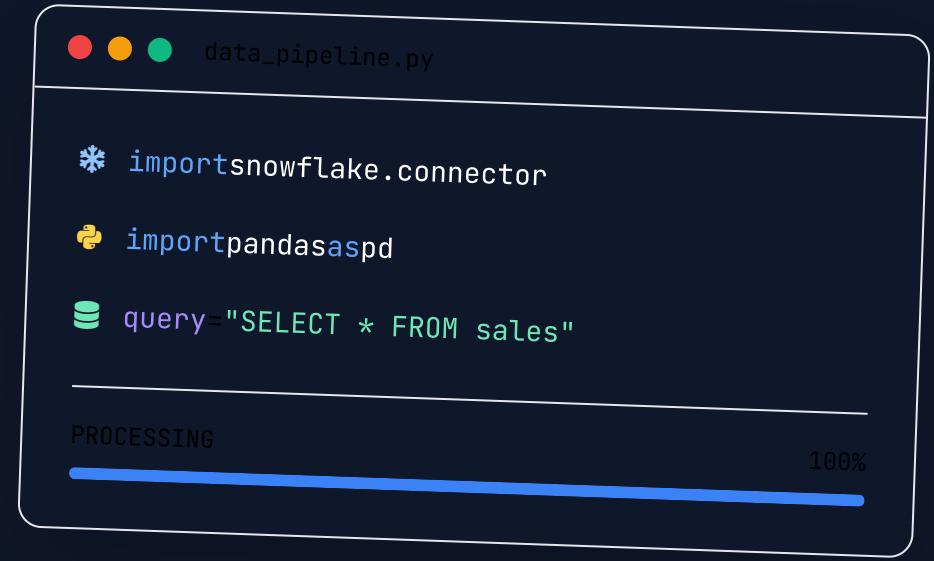
 2.30 Hrs / Day

Short lectures + Hands-on labs

IDEAL FOR

 Analysts & Engineers

Power users working with data



The screenshot shows a code editor window with a dark theme. At the top, there are three colored dots (red, yellow, green) followed by the file name `data_pipeline.py`. Below the header, the code is displayed in three lines:

```
* import snowflake.connector  
* import pandas as pd  
* query="SELECT * FROM sales"
```

Below the code, there is a progress bar labeled "PROCESSING" with a blue bar extending to the right, ending at the text "100%".

 Modeling

 Automation



Course Overview

Course Goals

- ✓ Build strong **SQL fundamentals** and Snowflake fluency.
- ✓ Apply **Python** for data wrangling, automation, and reporting.
- ✓ Deliver a **capstone project** that demonstrates end-to-end value.

Key Outcomes

- ★ Write performant SQL (joins, windows, CTEs)
- ★ Model and prepare data for analytics in Snowflake
- ★ Create automated Python reports and predictive insights

PREREQUISITES

Spreadsheet Skills

Familiarity with rows, columns, and basic formulas.

Scripting Basics

Prior exposure to coding logic is helpful but not required.

WHAT'S INCLUDED



Slides



Labs



Datasets



Solutions

Program Structure



Schedule (2.5 Hrs/Day)

SQL TRACK

- D1 SQL & Snowflake Foundations 1
- D2 SQL & Snowflake Foundations 2
- D3 Intermediate SQL (CTEs/Windows)
- D4 Intermediate SQL (Modeling)

PYTHON & PROJECT

- D6 Python Data Handling Part 1
- D7 Python Data Handling Part 2
- D8 Applied Modeling & Capstone



Modality: Live Walkthroughs Guided Labs Live Q&A

>_ TOOLS ACCESS

* Snowflake

Roles, warehouses, sample schemas

🐍 Python Environment

Conda/venv setup, connector creds

💽 Sample Data

Churn, Sales, Product datasets

DELIVERABLES



Daily Lab Check-ins

Progress validation on core concepts



Capstone Project

End-to-end solution & presentation



Learning Objectives

- ✓ Understand core **relational database concepts** and the basics of cloud data warehousing.
- ✓ Run basic `SELECT` queries with filters to retrieve specific data points.
- ✓ Navigate the **Snowflake UI** including roles, warehouses, and worksheets.



Key Topics

ARCHITECTURE

Storage vs. Compute vs. Services layer separation.

ACCESS SETUP

Configuring Roles, Virtual Warehouses, Databases, and Schemas.

SQL ESSENTIALS

SELECT WHERE LIMIT

Hands-on Labs

01. Connect & Explore

Log in to Snowflake and query the `information_schema` to view objects.

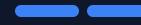
02. Data Extraction

Extract simple datasets applying basic filters (date ranges, categories).

03. Save & Share

Save your SQL worksheets and share result sets with other roles.

```
SELECT * FROM  
USERS WHERE  
ACTIVE = TRUE
```



Learning Objectives

- ✓ Aggregate and sort data effectively using **grouping** functions.
- ✓ Combine data from multiple tables using **INNER, LEFT, and RIGHT** joins with confidence.
- ✓ Compose **subqueries** and nested queries for modular data logic.



Key Topics

AGGREGATION & SORTING

GROUP BY HAVING ORDER BY

JOINS & RELATIONSHIPS

Connecting relational datasets.

INNER LEFT RIGHT

ADVANCED LOGIC

Subqueries, Nested Queries, and Derived Tables.

Hands-on Labs

01. Customer Analysis

`SELECT A.ID, SUM(B.VAL) FROM ORDERS B JOIN CUSTOMERS A ON A.ID = B.CUST_ID`

Analyze sales by segment and identify top-N products using GROUP BY .

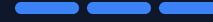
02. Multi-Table Joins

Join orders, customers, and products; apply filters and calculate rankings.

03. Reusable Logic

Build a reusable query snippet for common analytical questions.

Focus on understanding the join types for data integrity.



Learning Objectives

- ✓ Structure complex logic using **CTEs** (**Common Table Expressions**) for better readability.
- ✓ Master **window functions** for advanced analytic calculations like ranking and moving averages.
- ✓ Improve query **clarity and maintainability** by refactoring nested subqueries.



Key Topics

ADVANCED STRUCTURE

Common Table Expressions (WITH clause) for modular query building.

WINDOW FUNCTIONS

RANK ROW_NUMBER LEAD/LAG

ANALYTICAL FRAMES

Partitions, order frames, running totals, and period deltas.

Hands-on Labs

01. CTE Pipeline

WITH T1 AS (SELECT * FROM ...)
SELECT ...

Build a multi-step CTE pipeline: Clean → Enrich → Summarize.

02. Churn & Cohorts

Calculate churn flags and cohort metrics using window functions.

03. Compare Approaches

Refactor legacy subqueries into CTEs and compare performance/readability.



Learning Objectives

- ✓ Design **analytics-ready schemas** understanding the trade-offs between Star and Snowflake models.
- ✓ Implement efficient **multi-table joins** to unify fragmented data sources correctly.
- ✓ Translate business questions into a concrete **data model** in Snowflake.



Key Topics

SCHEMA DESIGN

Star vs. Snowflake schemas; Defining grain, dimensions, and facts.

RELATIONSHIPS

Handling constraints, foreign keys, and implementing surrogate keys.

PERFORMANCE BASICS

CLUSTERING PRUNING CACHING

Hands-on Labs

01. Multi-table Joins

Execute complex joins on a churn dataset (Customers + Usage + Support).

02. Design the Model

Identify entities and relationships; sketch dimension and fact tables.

03. Build in Snowflake

Create the physical tables and views to support the analytics model.

```
SELECT C.NAME  
FROM CHURN T  
JOIN CUSTOMERS U
```



Learning Objectives

- ✓ Leverage **Pandas & NumPy** for efficient tabular data manipulation and analysis.
- ✓ Establish secure connections between Python and Snowflake to **pull and push data**.
- ✓ Clean, reshape, and prepare raw datasets for downstream **predictive modeling**.



Key Topics

PYTHON DATA STACK

Pandas DataFrames, NumPy arrays, I/O operations, Merges & Joins.

CONNECTIVITY

Snowflake Connector for Python, SQLAlchemy, credential management.

WRANGLING & MODELING

NULL
Handling

Type
Conversion

Baselines

Hands-on Labs

01. Connect & Query

import pandas as pd
df = pd.read_sql(
Write your first Python script to connect to Snowflake and load data into a Pandas DataFrame.

02. Clean & Prepare

Wrangle a raw sales dataset: handle missing values, parse dates, and calculate KPIs.

03. Derived Views

Push processed data back to Snowflake or create view definitions for reporting.



Learning Objectives

- ✓ Engineer features and create **summary statistics** for downstream reporting.
- ✓ Build reliable end-to-end **SQL + Python workflows** that are production-ready.
- ✓ Implement **error handling** and optimization techniques for data tasks.



Key Topics

FEATURE ENGINEERING

Time windows, rolling averages, categorical encodings.

REPORTING LOGIC

Grouped summaries, basic charts, and file exports.

ROBUST WORKFLOWS

try/except logging optimization



Hands-on Labs

```
df['rolling_avg'] = df.sales.rolling(7).mean()  
if df['sales'].isna().any():  
    df['sales'].fillna(0, inplace=True)
```

01. Summary Reports

Build scripts to generate weekly/monthly KPI summaries from raw data.

02. Automated Reporting

Automate a weekly reporting job (via CLI script or scheduled notebook).

03. Shared Utilities

Package common functions into shared utilities for team reuse.



Learning Objectives

- ✓ Apply **end-to-end skills** on a real-world dataset, integrating SQL and Python.
- ✓ Collaborate effectively on **modeling, querying, and reporting** tasks within a team.
- ✓ Communicate data findings clearly and make actionable **recommendations**.



Key Topics

PROJECT SCOPING

Defining objectives, scope, and success criteria for analysis.

MODEL DESIGN

Creating an appropriate data model for the specific use case.

STAKEHOLDER ALIGNMENT

TEMPLATES KPIs VISUALS



Group Project

01. Choose Dataset

Select a real-world scenario (e.g., churn reduction, process optimization).

02. Prepare Assets

Develop SQL transformation scripts, Python notebooks, and dashboards.

03. Draft Initiatives

Create 6+ concrete improvement initiatives with estimated impact.

PROJECT_START

TEAM_ASSIGN

EXECUTE

FINAL

Capstone Project

Goal: End-to-End Analytics Solution



Group or Individual

Project Brief

Design and implement a complete data workflow starting from raw data ingestion in Snowflake to actionable insights in Python.

TOPIC OPTIONS

Churn Reduction

Process Optimization

Growth Insights

PROJECT TIMELINE

01

Kickoff & Scoping

Day 8: Select dataset & define KPIs.

02

Midpoint Check

Validate data model & initial transforms.

🏁

Final Presentation

Demo workflow & present insights.

Core Deliverables



SQL Layer

- Schema DDL
- Transform Scripts
- Analytical Views



Python Layer

- Clean/Wrangle NB
- Feature Eng.
- Auto-Report Script



Reporting

- KPI Summary
- Visualizations
- README.md

EVALUATION CRITERIA

Technical Quality

40%

Efficient queries, clean Python code, error handling.

Business Relevance

40%

Clarity of insights and actionable recommendations.

Documentation

20%

Reproducibility, comments, and README.

COMPLETE

Closing & Next Steps

Course Summary & Future Learning Path

Congratulations! 

★ Key Takeaways



Analytics Foundation

Mastered SQL fundamentals and Snowflake architecture for performant data warehousing.



Python Workflows

Built automated pipelines using Pandas and connectors to bridge data and insights.



End-to-End Delivery

Executed a complete analytics project from raw data to final report delivery.



Your Next Steps

LEARNING RESOURCES



Course Repository

Slides, Labs, Solutions



Snowflake Docs

Official Documentation



Pandas Guide

User Guide & API Ref



Q & A

Open discussion. Feel free to ask about specific challenges or implementation details.

Feedback Survey Link Available in Chat