**Azure Data Warehouse ETL Pipeline Documentation  
Data Factory Pipeline – Copy Activities**

**A screenshot of a computer

AI-generated content may be incorrect.**

**Description**:  
This diagram illustrates the data movement pipeline within Azure Data Factory. Each "Copy data" activity extracts data from operational sources and loads it into the staging or destination layer of the Data Warehouse.

**Components**:

* VDimUser: Dimension table for user information.
* DimCourse: Dimension table for course metadata.
* FactEnrollment: Fact table capturing user-course enrollment data.
* SupDimSection, SupDimReq, SupDimQuiz: Supporting dimension tables containing course section, prerequisites, and quiz data.
* FactOrder: Fact table representing course purchases or subscriptions.

**Purpose**:  
To extract and load raw data from operational sources into the Data Warehouse for further transformation and analysis.

**Project Stage**:  
ETL (Extract-Load) phase of the Data Warehouse pipeline.

**Data Flow – Transformation Logic**

**A screenshot of a computer

AI-generated content may be incorrect.**

**Description**:  
This diagram shows a Data Flow in Azure Data Factory where different dimension and fact tables are joined together using transformation logic. This step prepares clean, enriched, and linked data for insertion into the final Data Warehouse tables.

**Flow Sequence**:

* Begins with DimCourse
* Linked through surrogate keys to SupDimQuizKey, SupDimSectionKey, and SupDimReqKey
* Then joined with EnrollmentKey and DimUserKey
* Final output is loaded into the DWH table

**Purpose**:  
To apply business logic, perform necessary joins and transformations, and ensure referential integrity before data is inserted into the final warehouse.

**Project Stage**:  
ETL (Transform) phase – data integration and transformation.

**Azure Resources Overview**

**A screenshot of a computer

AI-generated content may be incorrect.**

**Description**:  
This screenshot presents the Azure resources used throughout the project as seen in the Azure Portal.

**Resources**:

* UdemyDB: SQL operational database containing raw source data (OLTP).
* Udemy\_DWH: SQL Data Warehouse where final integrated data is stored.
* udemydwh: Azure Data Factory instance used for orchestrating ETL processes.
* ITI: Azure Resource Group housing all project-related services.

**Purpose**:  
To manage and organize the platform’s data infrastructure in Azure. These components represent the foundational architecture for the data flow from source to warehouse.

**Project Stage**:  
All phases – development, orchestration, deployment, and monitoring.