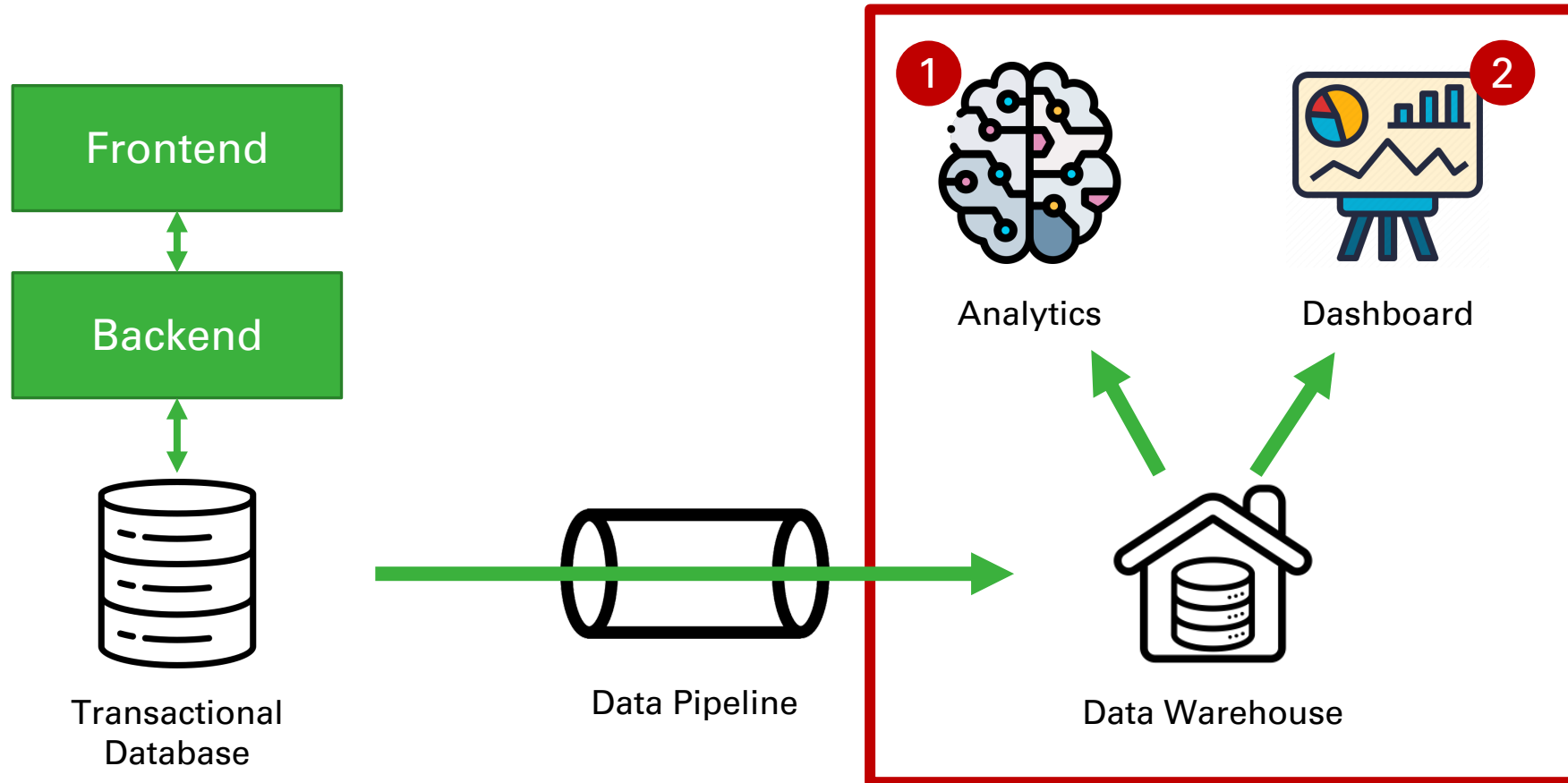


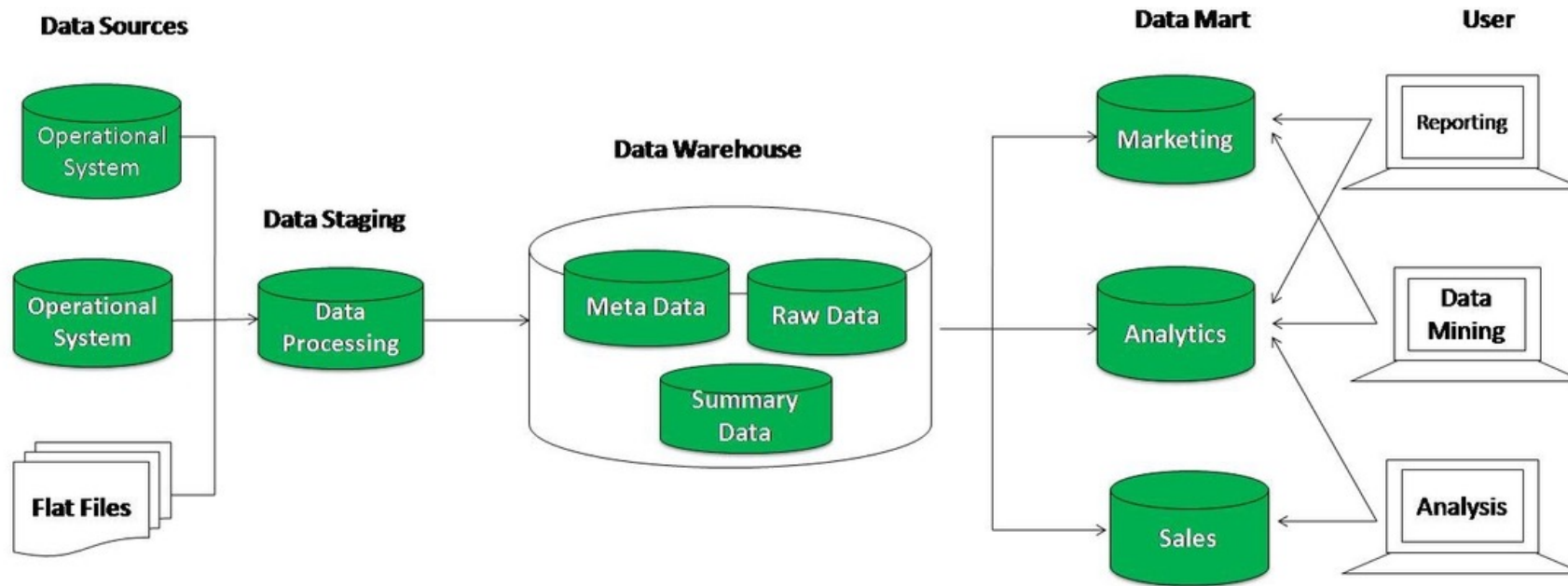
Database Demo: Data Science & Analytics

DATABASE SYSTEM 2110322
(2022/1)

Architecture



Architecture



<https://www.geeksforgeeks.org/top-15-popular-data-warehouse-tools/>

OnLine Transaction Processing vs OnLine Analytical Processing

	OLTP	OLAP
Function	Day to day operation	Decision support
Database Design	Application oriented	Subject oriented
Data	Current, up-to-date detailed, flat relational, isolated	Historical, summarized multi-dimensional, consolidated
Usage	Repetitive	Ad-hoc
Access	Read/Write	Lots of scans
Unit of Work	Short, simple transaction	Complex query
Database Size	Gigabytes	Terabytes
Metric	Transaction throughput	Query throughput, response

<https://medium.com/t%C3%BCrk-telekom-bulut-teknolojileri/whats-the-difference-between-oltp-and-olap-bdcafdffb1c3>

Data Warehouse

Top Data Warehouse Tools & Software



Snowflake



Microsoft Azure



SQL Server



Google BigQuery



Amazon Redshift

<https://www.oneclickitsolution.com/blog/what-is-data-warehouse/>

Data Science Demo - Prerequisites

- Database

- Mysql / MariaDB / Docker & Docker-compose

- Database Driver

- mysql-client / mariadb-connector-c
 - (M1/M2: Make sure the architecture of your Python and the driver matches)
 - <https://github.com/PyMySQL/mysqlclient/issues/496>

- Jupyter Notebook Editor of your choices

- Jupyter / JupyterLab / Spyder / PyCharm / DataSpell
 - VSCode with Jupyter Extension
 - Or even Colab (the db must be accessible from colab)

- Python Libraries

- Data Science:

- numpy
 - pandas, pandasql
 - matplotlib, seaborn
 - sklearn, xgboost

- Stats:

- scipy
 - researchpy

- Database Clients:

- sqlalchemy
 - mysql-connector-python / mariadb

Code & Dataset

Github: <https://github.com/BlueTogepi/dbsys2110322-demo-datasci>

Dataset from: <https://www.kaggle.com/datasets/aungpyaeap/supermarket-sales>

(Already converted to .sql script which is inside db_dump/dump-warehouse.sql)

Start the Database Server by: "docker-compose up"

Or import db_dump/dump-warehouse.sql to your MySQL/MariaDB Database

Data Science Demo

A thin vertical line is positioned to the right of the text, extending from the top of the word 'Science' down to the bottom of the word 'Demo'.

Data Dashboard



Figure 1: Magic Quadrant for Analytics and Business Intelligence Platforms



Source: Gartner (March 2022)

Power BI

<https://app.powerbi.com>

<https://powerbi.microsoft.com/en-us/>

Power BI Pro

Per user

\$13.70

Per user/month

License individual users with modern, self-service analytics to visualise data with live dashboards and reports, and share insights across your organisation.

- Power BI Pro is included in [Microsoft 365 E5](#).
- Available to buy now with a credit card.¹

Buy now >

Power BI Premium

Per user

\$27.50

Per user/month ²

License individual users to accelerate access to insights with advanced AI, unlock self-service data prep for big data, and simplify data management and access at enterprise scale.

- Includes all the [features](#) available with Power BI Pro.
- Available to buy now with a credit card.¹

Buy now >

Tableau

<https://www.tableau.com/academic/students>

<https://www.tableau.com/support/releases>

Driver Support:

<https://www.tableau.com/support/drivers>



Hi XXX,

Please retain this email for your records. You will need the product key and instructions enclosed.

Welcome to Tableau for Students! Your academic license now includes Tableau Desktop, Tableau Prep, and eLearning for free.

The product key below can be used to activate both **Tableau Desktop** and **Tableau Prep** on two separate computers, Windows or Mac. This key will expire in one year.

- [Download Tableau Desktop](#)
- [Download Tableau Prep](#)
- Activate with your product key: XXXX-XXXX-XXXX-XXXX-XXXX
- If you're receiving the error "product key is invalid" visit the [knowledge base page](#) to resolve your issue.

Get started with free eLearning online self-paced courses:

1. Go to <https://elearning.tableau.com>
2. Create (or login to) your TableauID account, and confirm email address via the TableauID confirmation email
3. Go to <https://elearning.tableau.com> and log on using your TableauID
4. Register at the Tableau Learning Center with this Access Key: xxxxxxxxxxxxxxxxxxxxxx

These keys are for your personal use only. Please do not share them. Need additional help? check out the [FAQ's](#) or [submit a case](#) for installation support.

Best,

Tableau Academic Team
Tableau Software

e: tft@tableau.com

Data Dashboard Demo - Prerequisites

- Database
 - Mysql / MariaDB / Docker & Docker-compose
- Database Driver
 - MySQL ODBC / MariaDB ODBC
(M1/M2: Use x86 driver as Tableau uses x86 architecture running under Rosetta 2)
 - (Mac) iODBC
- Tableau Desktop
 - This demo uses Tableau 2022.3
<https://www.tableau.com/support/releases/desktop/2022.3>

Data Dashboard Demo

What to go from here

- We've only touched Tabular Data and Classical Machine Learning
 - More in-depth Data Science / Engineering / Visualization: 2110446 Data Science and Data Engineering
 - Time-series Data Analysis: 2110430 Time Series Mining and Knowledge Discovery
- Other types of Data & AI Models:
 - Image (Computer Vision): 2110431 Introduction to Digital Imaging, 2110433 Computer Vision
 - Text (Natural Language Processing): 2110572 Natural Language Processing System
 - Audio (Speech Recognition): 2110432 Automatic Speech Recognition
 - Biological Data (Bioinformatics): 2110581 Bioinformatics I
 - Artificial Intelligence & Machine Learning & Deep Learning:
2110477 Artificial Intelligence II, 2110571 Neural Network, 2110573 Pattern Recognition
- Self-Leaning Roadmap: <https://github.com/Moataz-Elmesmary/Data-Science-Roadmap>
- Try Competitive Data Science!: <https://www.kaggle.com/>