

ORACLE

Vibe Data Engineering: Months to minutes with ADB and Select AI

Autonomous Database Learning Lounge

Hosted by Marcos Arancibia

Autonomous Database Product Management

AUTONOMOUS
DATABASE
LEARNING
LOUNGE

Agenda



**Javier de
la Torre
Medina**

Vibe Data Engineering, with Oracle Autonomous Database and Data Studio, **bridges the gap** between business ideas and engineering execution with unprecedented speed, backed by **enterprise-grade security** and **governance**.

Topics

- Explore Data with **Natural Language**: Ask business questions in plain English and get instant, actionable data using Select AI.
- **AI-Accelerate Your Data Lifecycle**: Generate & enrich Data effortlessly using simple prompts to **create synthetic data** or add **insights like sentiment** during data loading directly in Data Studio.
- **Transform Data with Ease**: Modify data structures using natural language prompts in **Table AI Assist**, which auto-generates and applies proper SQL.
- **Uncover Deeper Insights, Faster**: Let AI identify anomalies and patterns, empowering business users to build **intelligent data products**.

Q&A

- **Product Managers will answer any questions**

Before we begin....

This session is for you !!!

Ask your questions using Q&A

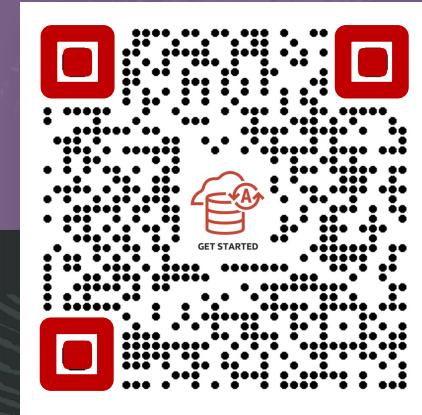
Product Managers are monitoring your questions

We will share links in Chat

**The recording will be made available in a few days at
oracle.com/goto/adb-learning-lounge**

Important links to bookmark

**Links to get you started
and to keep up to date with
Autonomous Database**



1 Get Started with ADB:
oracle.com/autonomous-database/get-started/

2 Join us: bit.ly/adb-linkedin-grp [@AutonomousDW](https://twitter.com/AutonomousDW)

2



Bluesky
autonomousdb.bsky.social

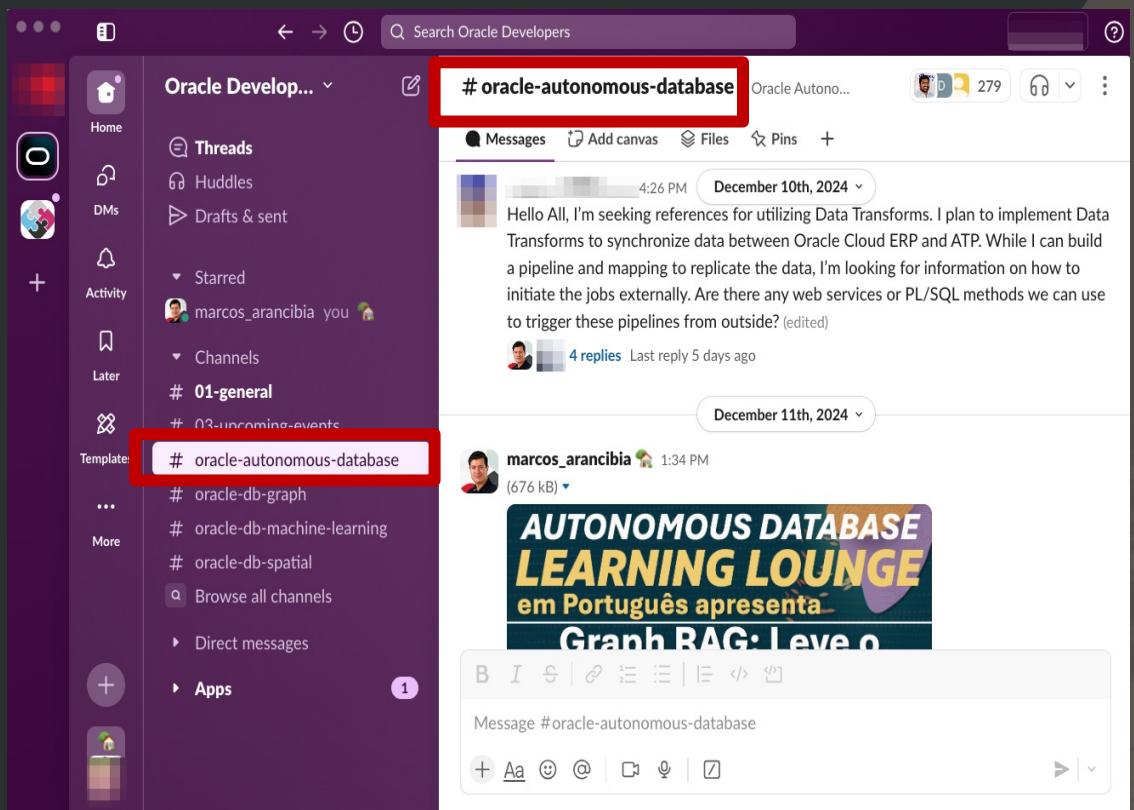
3 Got a question?
We are on stackoverflow
bit.ly/adb-stackoverflow

Join us on Developers Slack
(search #oracle-autonomous-database)
bit.ly/odevrel_slack

Join our External Slack

STEP 1: Join our Slack workspace at:
bit.ly/odevrel_slack

STEP 2: search for **#oracle-autonomous-database** at the top and click on the Channel



Speaker



Javier de la Torre Medina



Vibe Data Engineering

Months to minutes with ADB and Select AI

Javier de la Torre

Autonomous Data Warehouse Black Belt

All Started with Vibe Coding

Vibe coding (or **vibecoding**) is an approach **to producing software by using artificial intelligence** (AI), where a person describes a problem in a few natural language sentences as a **prompt** to a large language model (LLM) tuned for coding. The LLM generates software based on the description, shifting the programmer's role from manual coding to guiding, testing, and refining the AI-generated source code.

LLMs are getting better at Code Generation

Text Arena
View rankings across various LLMs on their versatility, linguistic precision, and cultural context across text.

Last Updated: Jun 5, 2025 | Total Votes: 597,826 | Total Models: 247

Coding | Search by model name... | Default

Rank (UB) ↑	Model ↑	Score ↑	95% CI (±) ↑	Votes ↑
1	gemini-2.5-pro-preview-06-05	1496	+18/-17	1,152
2	o3-2025-04-16	1460	+11/-9	3,642
2	chatgpt-4o-latest-20250326	1457	+10/-10	4,455
2	claude-opus-4-20250514	1456	+13/-13	2,497

OpenAI Code Model Accuracy

Model	Year	Accuracy %
Codex	2021	38
GPT-4.0	2023	33
GPT-4.1	2024	55
o3	2024	80
o4-mini-high	2025	70

And SQL is not an exception!

About BIRD

Page Views 232425

BIRD (Big Bench for LaRge-scale Database Grounded Text-to-SQL Evaluation) represents a pioneering, cross-domain dataset that examines the impact of extensive database contents on text-to-SQL parsing. BIRD contains over **12,751** unique question-SQL pairs, **95** big databases with a total size of **33.4 GB**. It also covers more than **37** professional domains, such as blockchain, hockey, healthcare and education, etc.

[Paper](#)

[Code](#)

[Mini Dev \(500\)](#)

Overall Leaderboard Single-Model Leaderboard

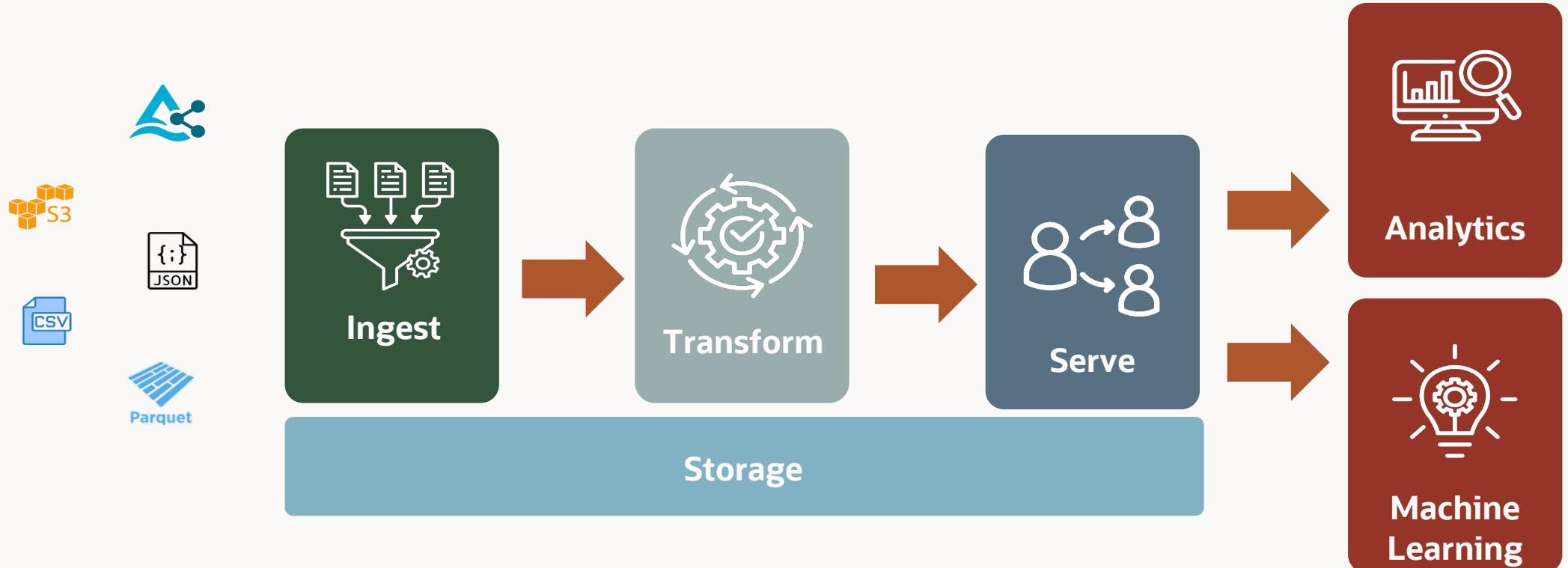
Leaderboard - Execution Accuracy (EX)

Model	Code Size	Oracle Knowledge	Dev (%)	Test (%)
Human Performance <i>Data Engineers + DB Students</i>		✓		92.96
AskData + GPT-4o AT&T CDO - DSAIR [Shkpenyuk et al. '25]	UNK	✓	75.36	77.14
CHASE-SQL + Gemini Google Cloud [Pourreza et al. '24]	UNK	✓	74.90	76.02



Data Engineer Lifecycle

SQL is Everywhere!! Let's use AI to boost productivity!



Security



Data Management



DataOps



Data Architecture



Orchestration

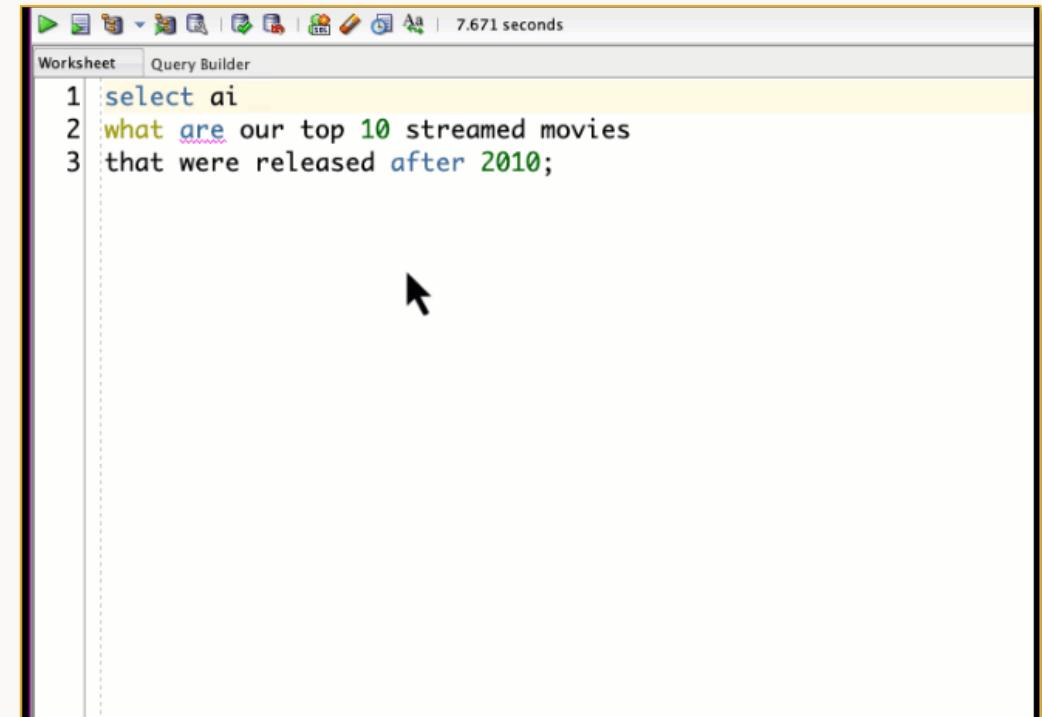
Which is the Secret Sauce?

Select AI

Use Natural Language to query data and get responses using Gen AI

Easily access LLMs from multiple providers

- OCI GenAI Service
- OpenAI
- Azure OpenAI Service
- Cohere
- Google
- Anthropic
- Hugging Face
- OpenAWS Bedrock
- AI API-compatible providers



The screenshot shows a software interface for querying data using natural language. The title bar indicates '7.671 seconds' for the query execution. The 'Worksheet' tab is active, and the 'Query Builder' tab is visible. The query text is as follows:

```
1 select ai
2 what are our top 10 streamed movies
3 that were released after 2010;
```

A cursor arrow is positioned in the middle of the third line of the query.

Actions

- **runsql** - (default) return SQL result set for NL2SQL
return vector result set for RAG
- **showsqli** – return the generated query
- **explainsqli** – explain generated SQL query
- **narrate** – return a conversational result for NL2SQL/RAG
- **chat** – return LLM response to prompt – a general AI chat

Select AI

Fully integrated in Autonomous Database

Use cases

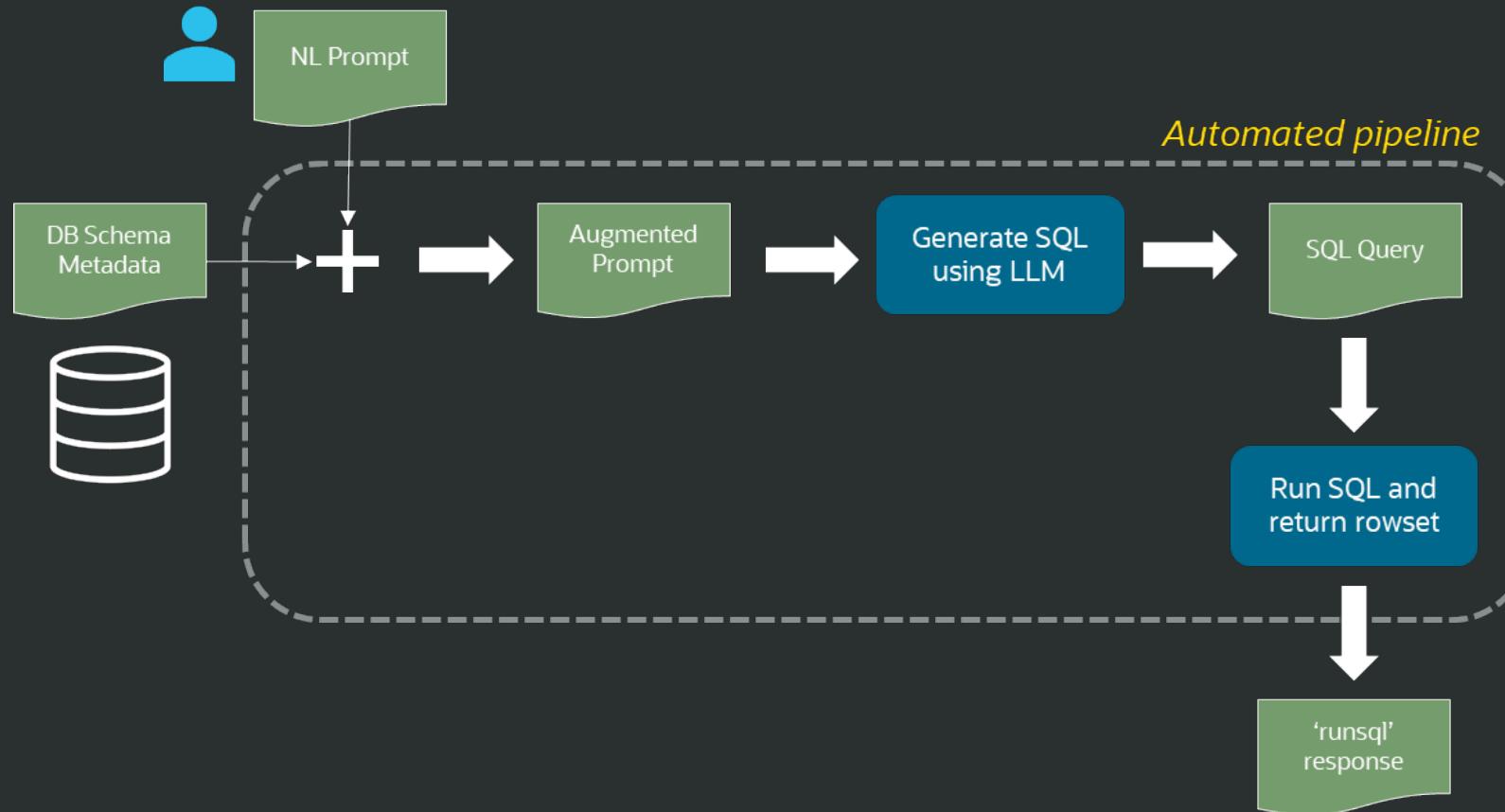
- Get help writing SQL to support your applications using **your** database
- Build applications that enable users to interact with their data *without* SQL
- Address business problems using generative AI:
customer email generation, recommendations, sentiment analysis, document summarization, etc.

Leverage the data security and user authentication of the database

Easily select and switch between large language models (LLMs)



SQL Query generation pipeline with 'runsql' action



Synthetic Data Generation (SDG)

Realistic test data generated by LLMs

Support multiple use cases

- Jumpstart a new project
- Validate the user experience
- Integrate with Autonomous Database
Metadata Clone for test-dev scenarios

Simple and Scalable



Example of synthetic data generation for multiple tables

```
-- Create new/use existing tables with
-- optional referential integrity constraints

BEGIN
  DBMS_CLOUD_AI.generate_synthetic_data(
    profile_name => 'GENAI',
    object_list =>
      '[{"owner": "SELECT_AI_USER",
        "name": "Director", "record_count":5},
       {"owner": "SELECT_AI_USER",
        "name": "Movie_Actor", "record_count":5},
       {"owner": "SELECT_AI_USER",
        "name": "Actor", "record_count":10},
       {"owner": "SELECT_AI_USER",
        "name": "Movie", "record_count":5,
        "additional_prompt":
          "all movies are released in 2009"}]'
  );
END;
```

```
-- Query the table to see results
```

```
SQL> SELECT * FROM SELECT_AI_USER.Movie
```

MOVIE_ID	TITLE	RELEASE_D	GENRE	DIR_ID
1	The Dark Knight	15-JUL-09	Action	8
2	Inglourious Basterds	21-AUG-09	War	3
3	Up in the Air	04-SEP-09	Drama	6
4	The Hangover	05-JUN-09	Comedy	1
5	District 9	14-AUG-09	Sci fi	10

```
-- Or ask select ai to show the results
```

```
SQL> SELECT AI how many actors are there
```

```
Number of Actors
```

```
-----
```

10

Innovation with AI

Using Data Studio



Select AI: Accelerate application development

Natural
Language
to SQL

AI SQL
Functions

Retrieval
Augmented
Generation

Synthetic
Data
Generation

APEX AI
Assistant

AI Assistants for Data Management

AI Assist
for
Transformations

AI Assist
for
Data Loading

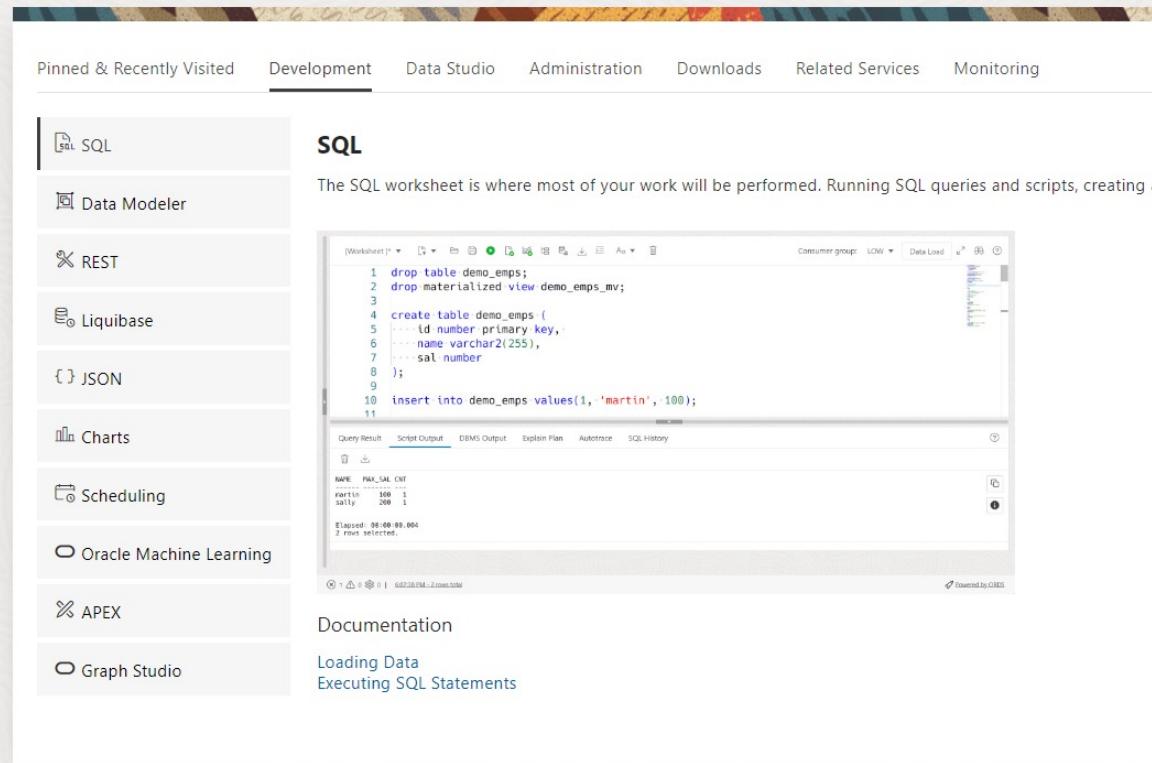
AI Assist
for
Table Definitions

Multicloud-Native: Leverage AI services in every cloud



Powerful built-in self-service tool

For Data Engineers and Data Analysts



The screenshot shows the Oracle Cloud Infrastructure Data Studio interface. The top navigation bar includes 'Pinned & Recently Visited', 'Development' (which is selected and highlighted in blue), 'Data Studio', 'Administration', 'Downloads', 'Related Services', and 'Monitoring'. The left sidebar contains links for 'SQL', 'Data Modeler', 'REST', 'Liquibase', 'JSON', 'Charts', 'Scheduling', 'Oracle Machine Learning', 'APEX', and 'Graph Studio'. The main workspace is titled 'SQL' and contains a SQL worksheet with the following code:

```
1 drop table demo_emps;
2 drop materialized view demo_emps_mv;
3
4 create table demo_emps (
5     id number primary key,
6     name varchar2(255),
7     sal number
8 );
9
10 insert into demo_emps values(1, 'martin', 100);
11
```

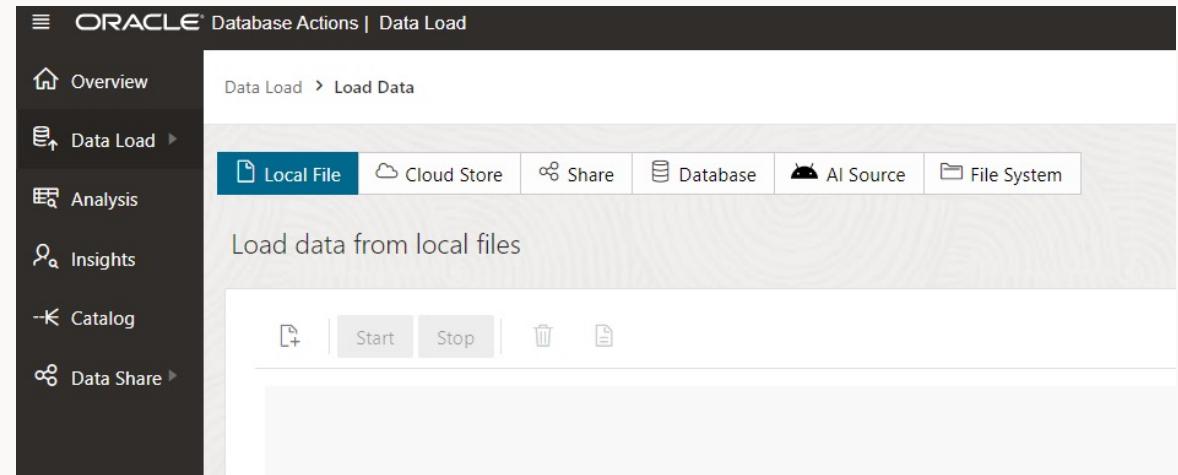
Below the code, the 'Query Result' tab shows the output:

NAME	MAX_SAL_CNT
sally	100
sally	200

Elapsed: 00:00:00.004
2 rows selected.

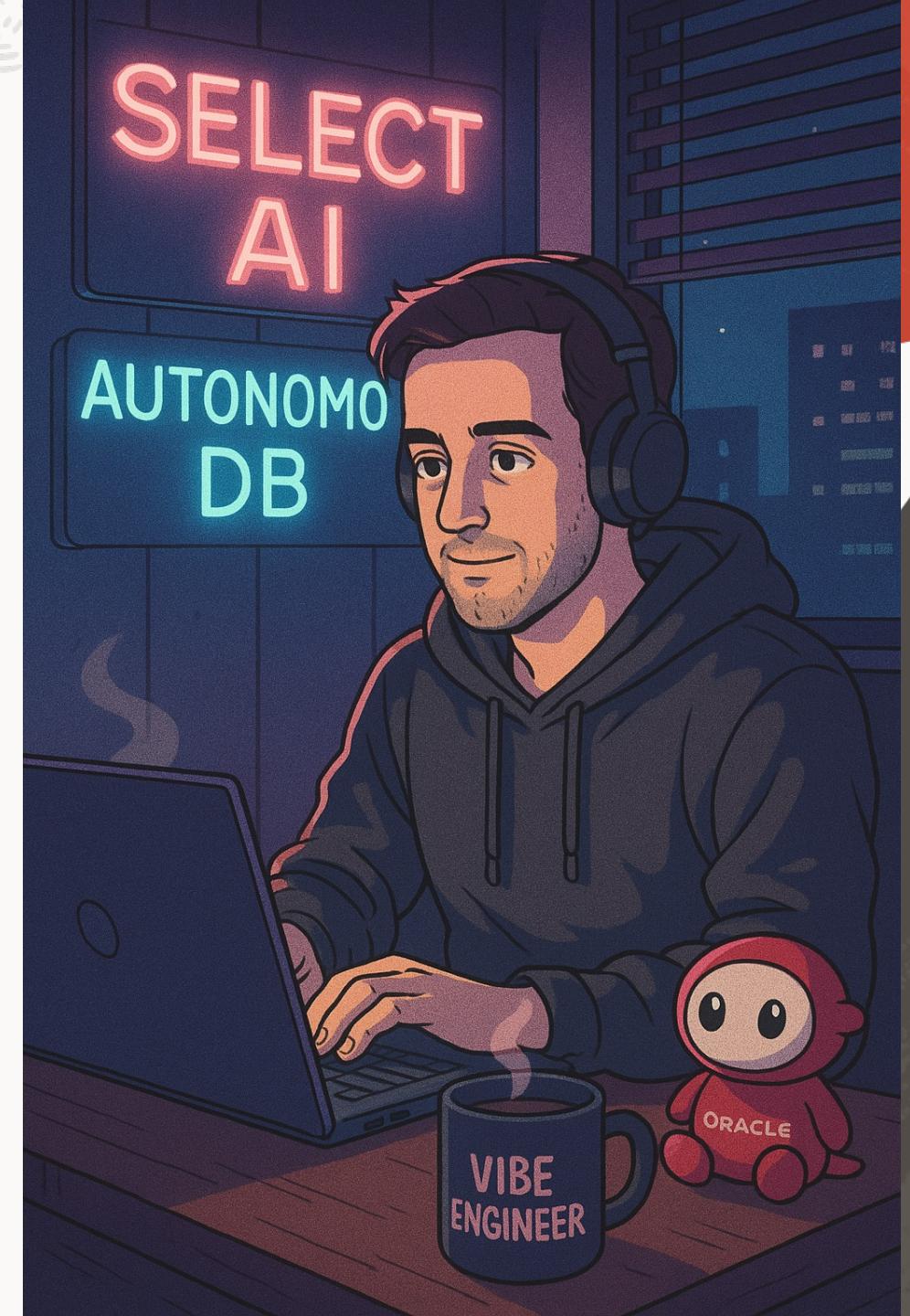
At the bottom of the SQL worksheet, there are tabs for 'Query Result', 'Script Output', 'DBMS Output', 'Explain Plan', 'Autotrace', and 'SQL History'. The 'SQL History' tab is currently selected.

Below the SQL worksheet, there are links for 'Documentation', 'Loading Data', and 'Executing SQL Statements'.



The screenshot shows the Oracle Database Actions | Data Load interface. The top navigation bar is 'ORACLE Database Actions | Data Load'. The left sidebar includes 'Overview', 'Data Load' (which is selected and highlighted in blue), 'Analysis', 'Insights', 'Catalog', and 'Data Share'. The main workspace is titled 'Data Load > Load Data' and contains a sub-menu for 'Local File' (which is selected and highlighted in blue), 'Cloud Store', 'Share', 'Database', 'AI Source', and 'File System'. Below this, the text 'Load data from local files' is displayed, followed by a set of buttons for 'Start', 'Stop', and other actions.

Let's start!



Demo Objective

We are going to start collecting customer feedback from our application.
Does it have any impact in our sales?

- Generate **synthetic data** to prepare pipelines
- **Sentiment Analysis** using AI
- **Data Quality** using AI
- **Understand** data
- Prepare Analytical View for **AI insights**



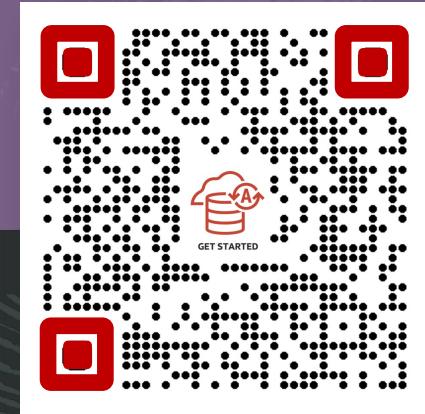
Thank you



Q&A is open

Important links to bookmark

**Links to get you started
and to keep up to date with
Autonomous Database**



1 Get Started with ADB:
oracle.com/autonomous-database/get-started/

2 Join us: bit.ly/adb-linkedin-grp [@AutonomousDW](https://twitter.com/autonomousDW)



3 Got a question?
We are on stackoverflow
bit.ly/adb-stackoverflow

Join us on Developers Slack
(search #oracle-autonomous-database)
bit.ly/odevrel_slack

Final Thoughts

oracle.com/goto/adb-learning-lounge

The screenshot shows the Oracle Autonomous Database Learning Lounge website. The top navigation bar includes 'ASK TOM', a search bar, 'Sign In', and tabs for 'Questions', 'Office Hours', 'Videos', 'Resources', and 'Classes'. Below the navigation is a menu with 'Sessions' (selected), 'Series' (highlighted in green), and 'My Dashboard'. A large section titled 'Autonomous Database Learning Lounge' is displayed, with a sub-section for 'Upcoming' sessions and a 'Replays' section. The 'Upcoming' section lists three events: 'Build AI-powered apps: A Step-by-Step Guide to Autonomous Database and GenAI' (March 11, 2025), 'Developer's nirvana with Autonomous Database: JSON-Relational Duality in Oracle Database 23ai' (March 18, 2025), and 'Autonomous Database: SQL Firewall, because hackers deserve 404s' (March 25, 2025). The 'Replays' section lists six recorded sessions: 'Migration to ADB Part III: OCI Database Management, the Swiss Army knife for databases' (November 21, 2024), 'Graph RAG: Bring the Power of Graphs to Generative AI' (November 21, 2024), 'Migration to ADB Part II: Easily migrate from previous database releases with DMS' (November 19, 2024), 'Migration to ADB Part I: Visualize and Evaluate your entire database estate with Oracle Estate Explorer' (November 12, 2024), and two other sessions by Marcos Arancibia and Jorge Martinez. Red arrows point from the text 'Links' to the 'Autonomous Database Learning Lounge' title, and from 'Upcoming' and 'Replays' to their respective sections on the website.

Links

Upcoming
Replays

Thank you for joining !!!

**AUTONOMOUS
DATABASE**
**LEARNING
LOUNGE**