
ORACLE®

What's New in Oracle 18c for Data Warehousing

ORACLE OPEN WORLD

October 1–5, 2017
SAN FRANCISCO, CA

George Lumpkin
Vice President, Product Management
Oracle Database Server Technology
October 2, 2017

ORACLE

Safe Harbor Statement

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.

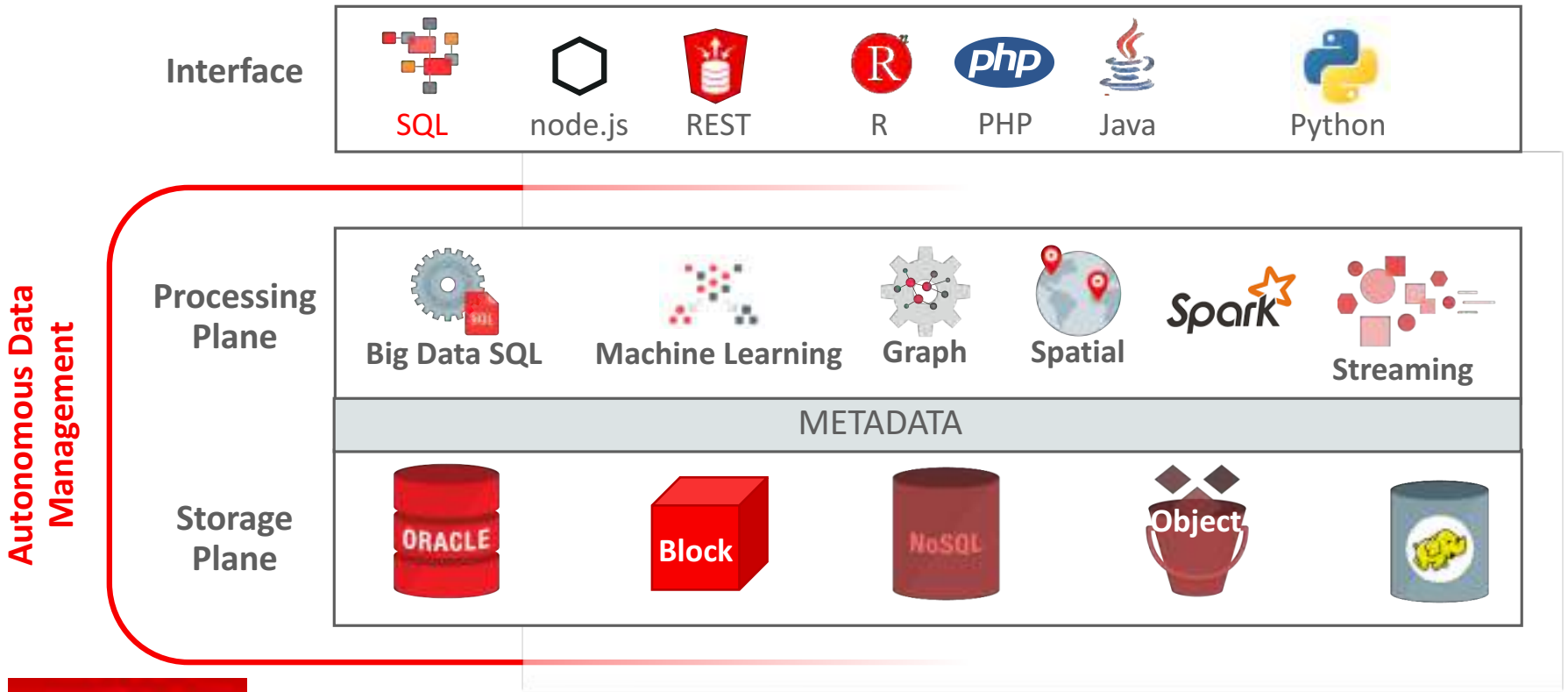
Agenda

- Vision for DW, Big Data and Analytics
- Enhancing the foundation: New DW features in 18c
- Building into Autonomous: Autonomous Data Warehouse Cloud

Data Management as a Service

- Philosophy: Data is a core asset. Focus on maximizing the value of data
- Data should be liquid:
  - easily accessed by any analytical processing engine
  - easily transformed into optimized formats
- Data architectures should be flexible, cost-effective and high-performant

Data Management as a Service



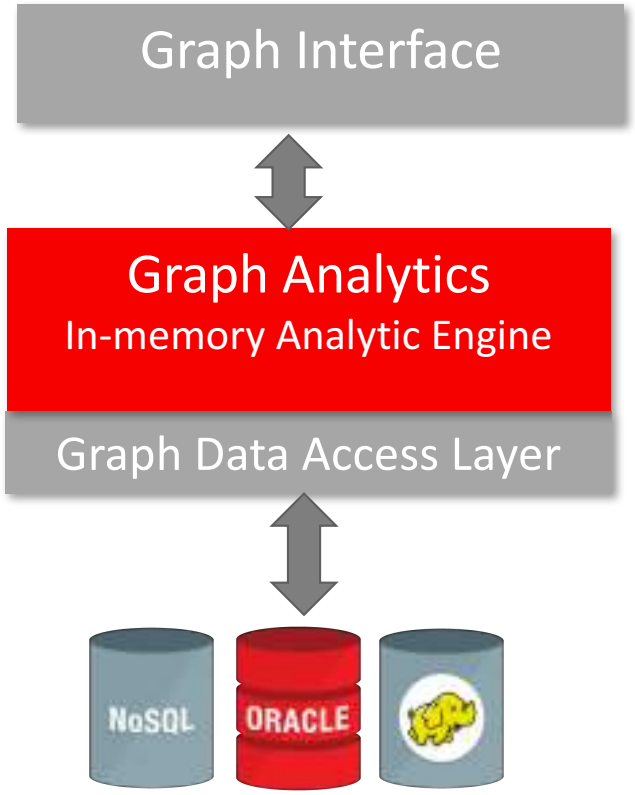
Data Management as a Service Vision

Key Concepts

1. Abstraction of storage and compute
   - Enables the logical data warehouse
   - Maximize value of data
2. The Cloud
   - Minimize costs without compromising performance
   - Accelerates the adoption of this architecture by customers

Oracle has made the necessary down-payments to deliver this architecture today

Abstraction of Storage and Compute - Property Graph



Ground-breaking Engineering

- Parallel in-memory graph engine developed by Oracle Labs
- 10-50x faster analytics than competitors' offerings
  - Analyze 20-30 Billion edge graph in memory on a single node
- Dozens of pre-built in-memory graph analysis algorithms

Domain-specific interfaces

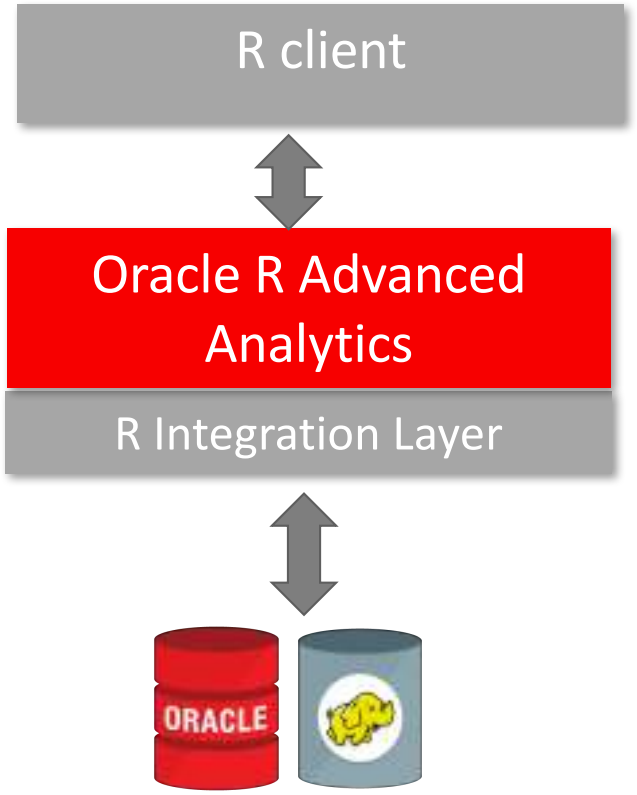
- Python, Groovy
- Java, Tinkerpop, Blueprints, Gremlin

Choice of Database Storage

- Oracle NoSQL, Hbase, Oracle Database

Commercial, supported software

Abstraction of Storage and Compute - Machine Learning



Ground-breaking Engineering

- 20+ integrated machine learning algorithms
  - Custom Spark and Map Reduce implementations
  - Wrapped Apache Spark MLlib algorithms
- Up to 10x plus faster model building than comparable open source offerings
  - Analyze billions of rows of data using Spark in the presence of limited memory

Domain-specific interfaces

- Standard R Interfaces

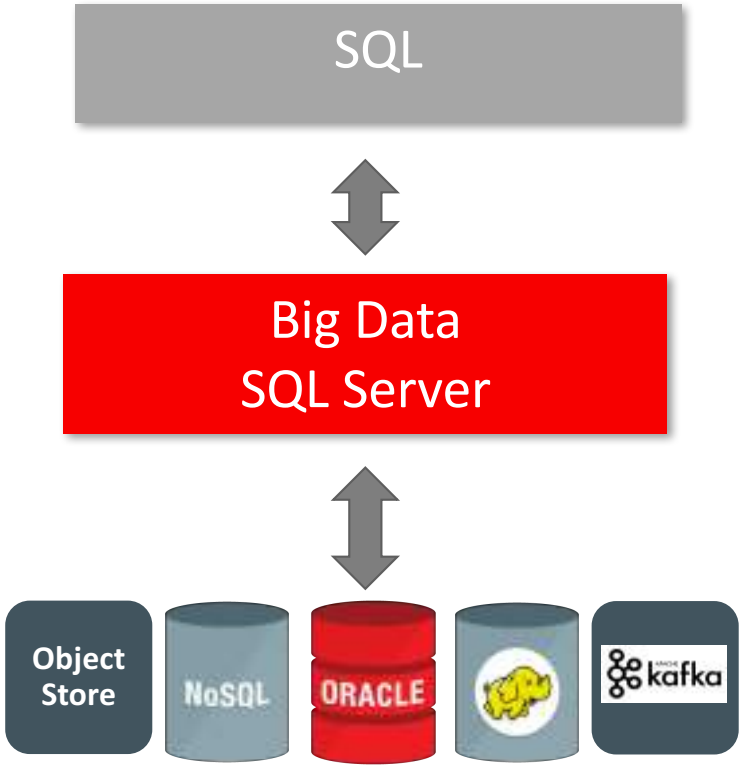
Choice of Database Storage

- Hadoop, Oracle Database

Commercial, supported software

- Including Oracle R Enterprise, Oracle's R distro

Abstraction of Storage and Compute - Big Data SQL



Ground-breaking Engineering

- Based on innovative Exadata Storage technology
  - Oracle query processing directly on storage nodes
- Data virtualization via SQL
  - Join across different data stores
- Tight integration with Hadoop ecosystem

Domain-specific interfaces

- Industry-leading, standards-compliant SQL

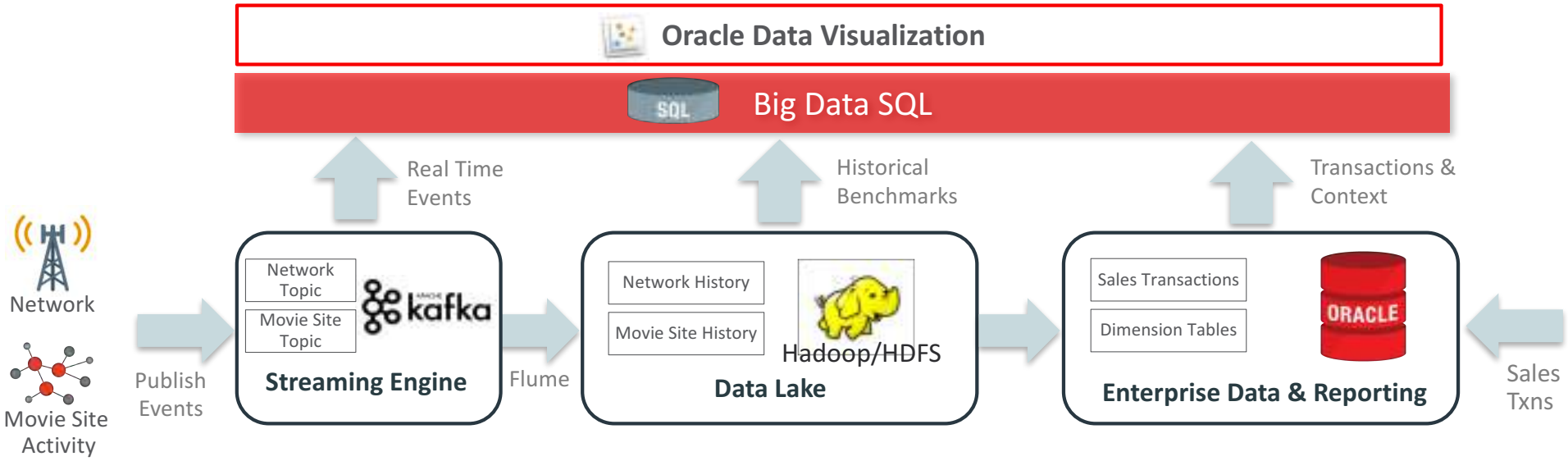
Choice of Database Storage

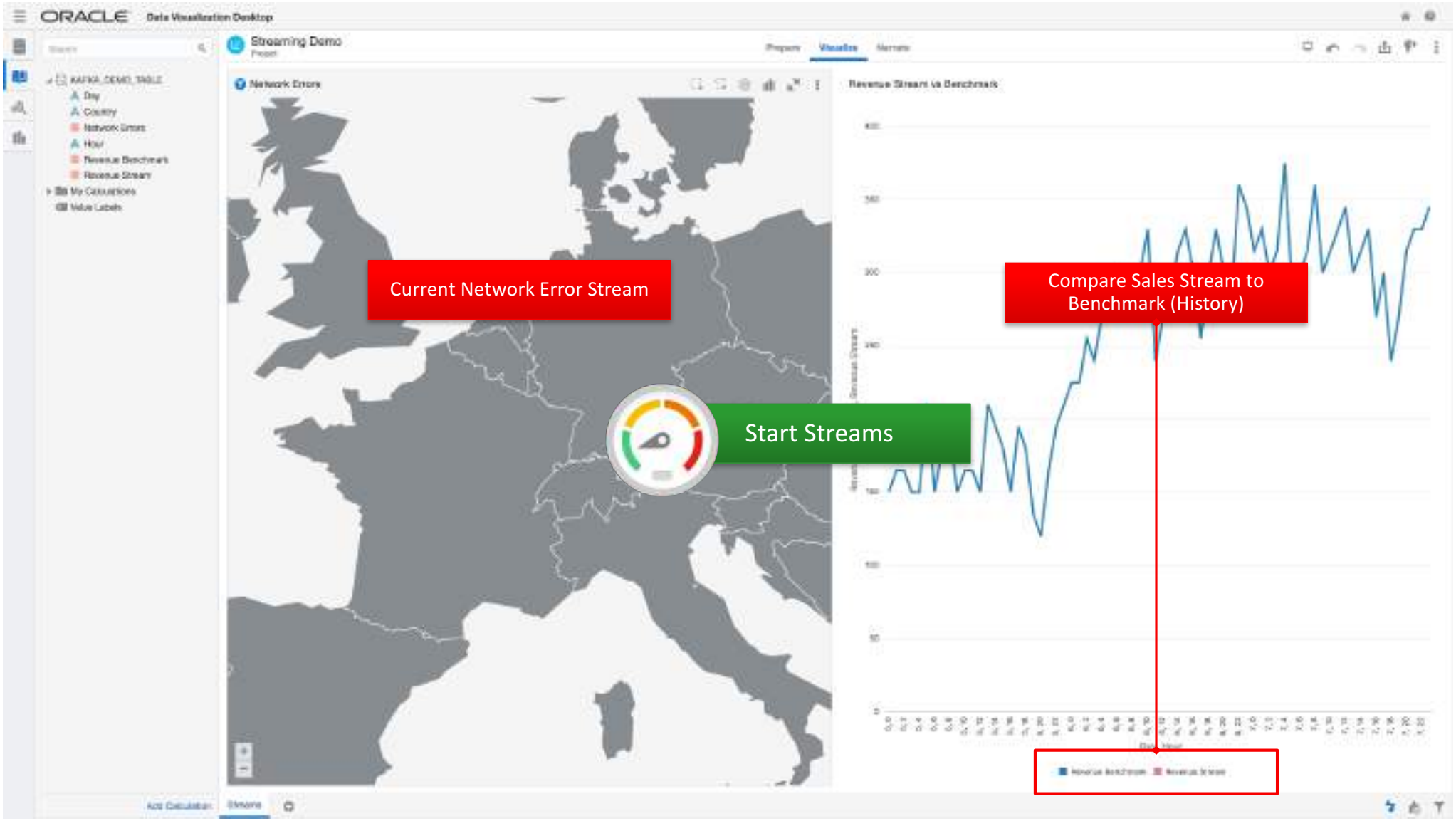
- NoSQL, Hadoop, Kafka, Object Storage

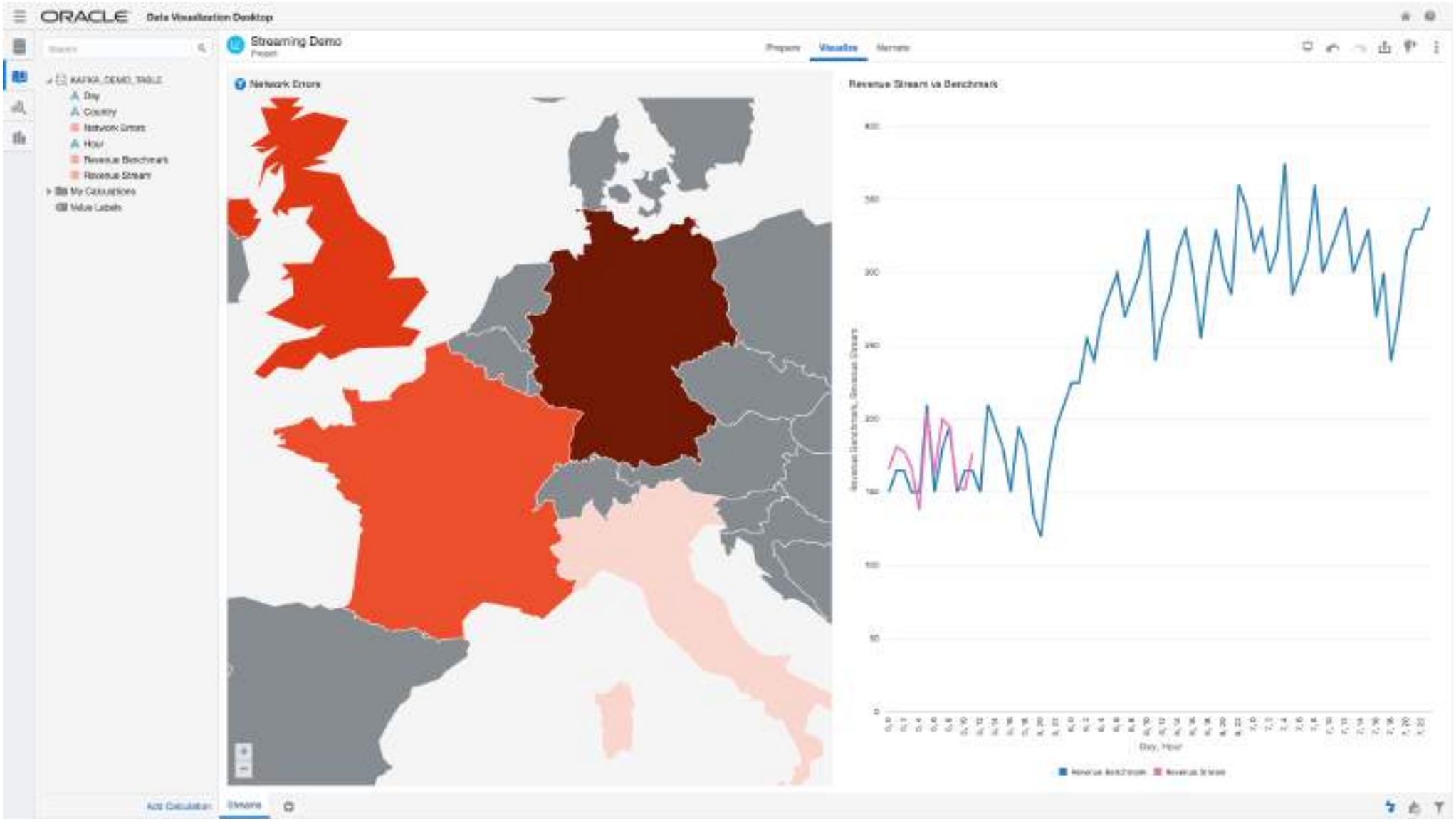
Commercial, supported software

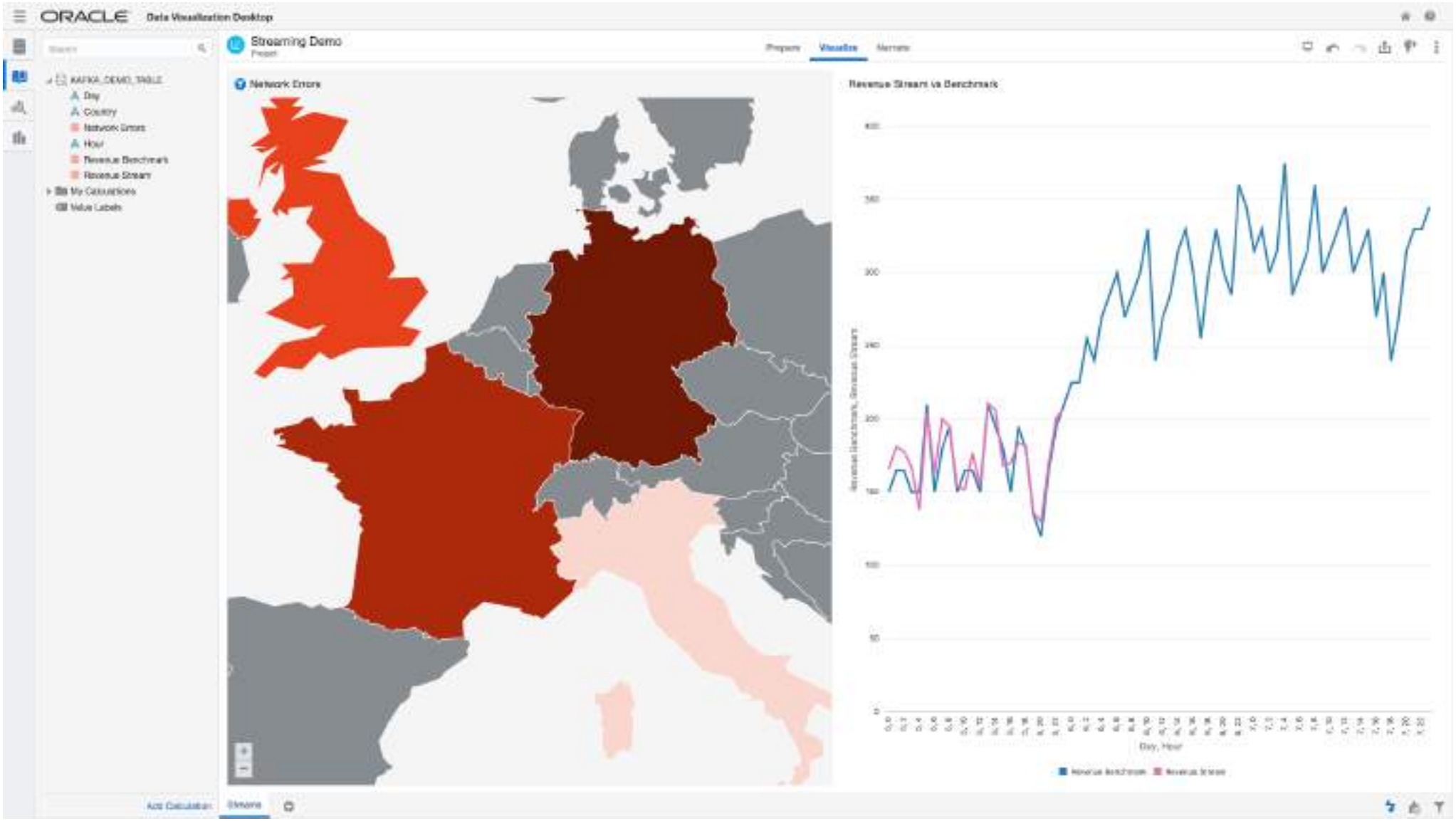
Demonstration Scenario

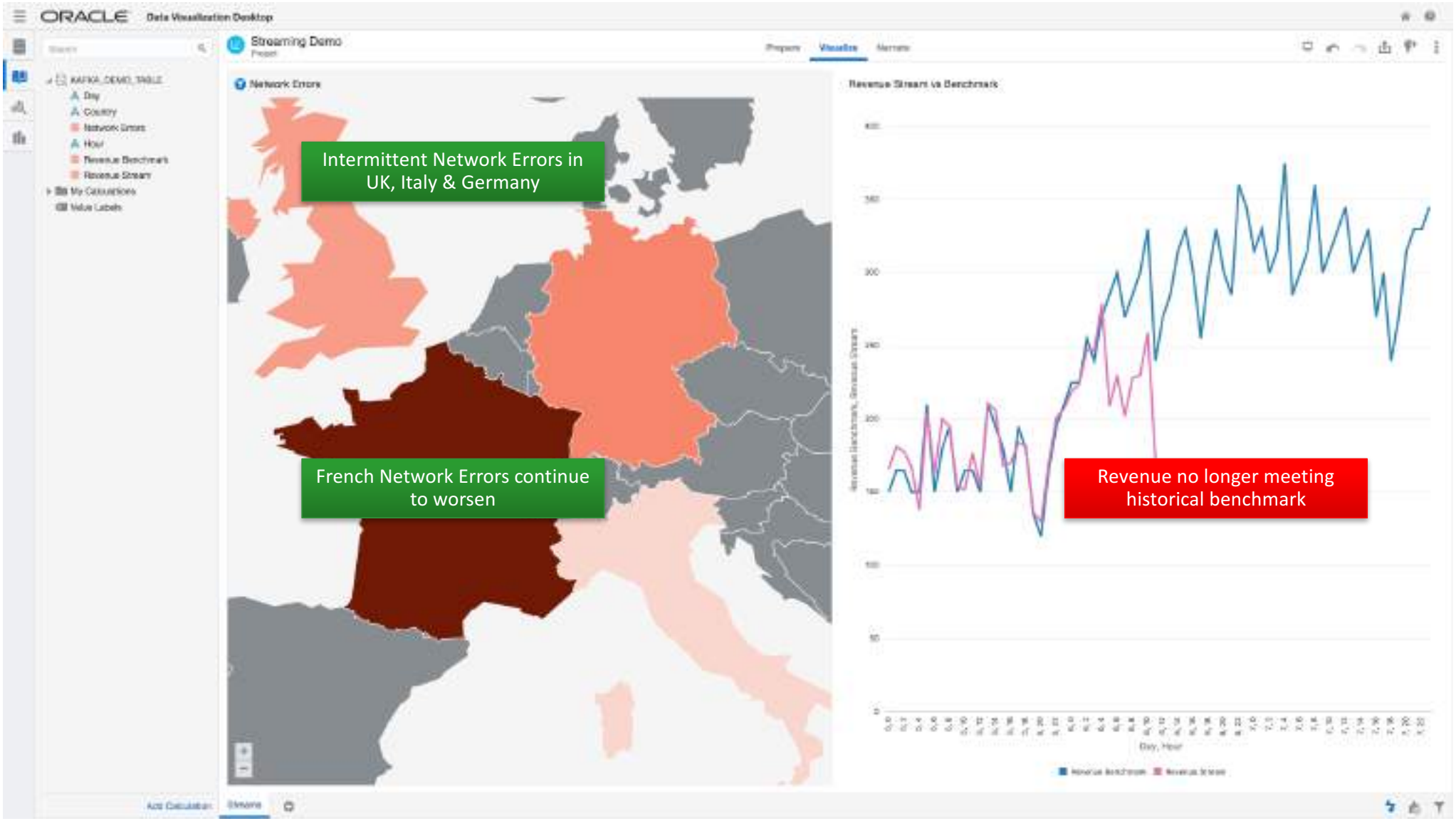
Analyzing Real-time Streams

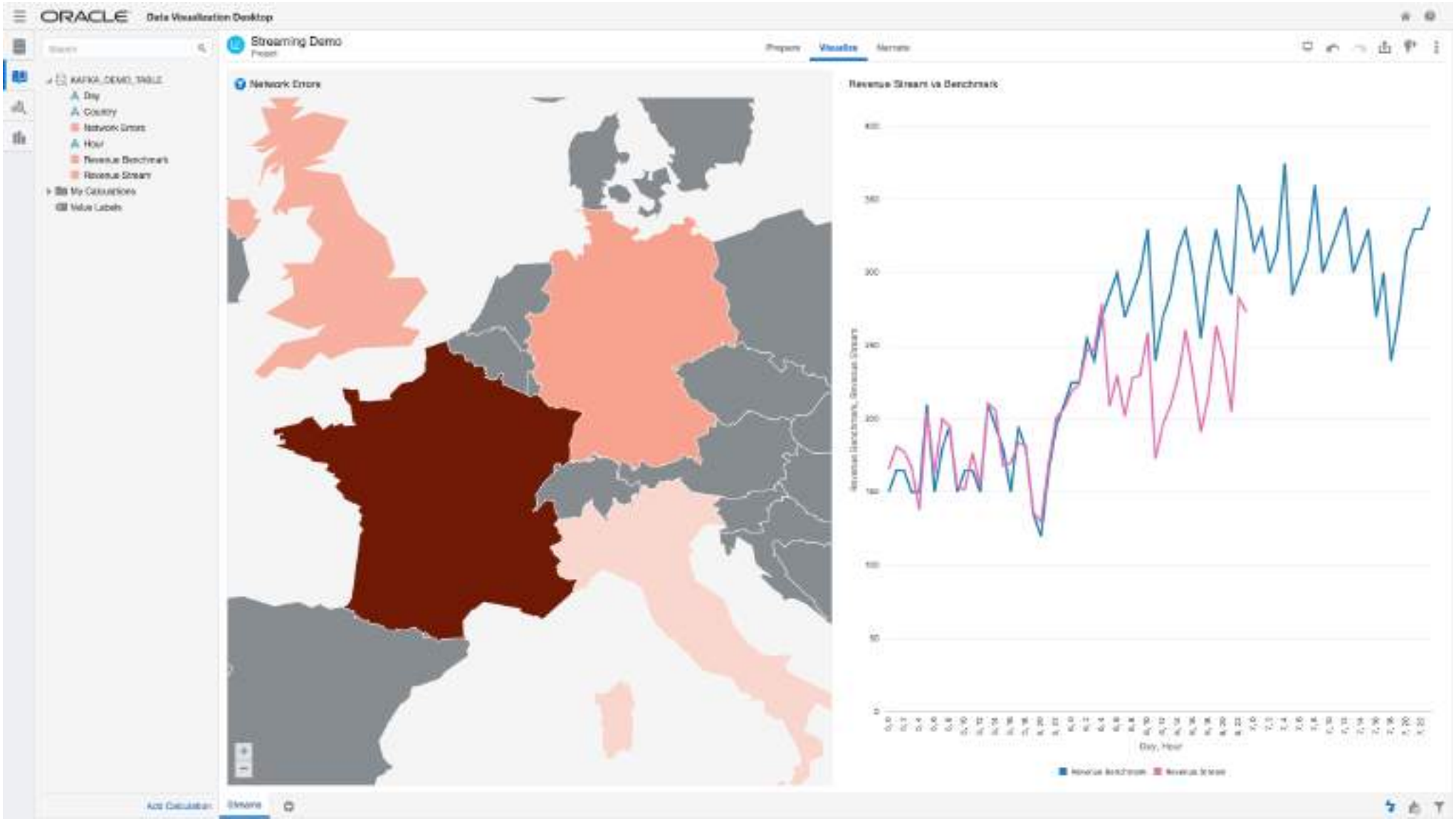


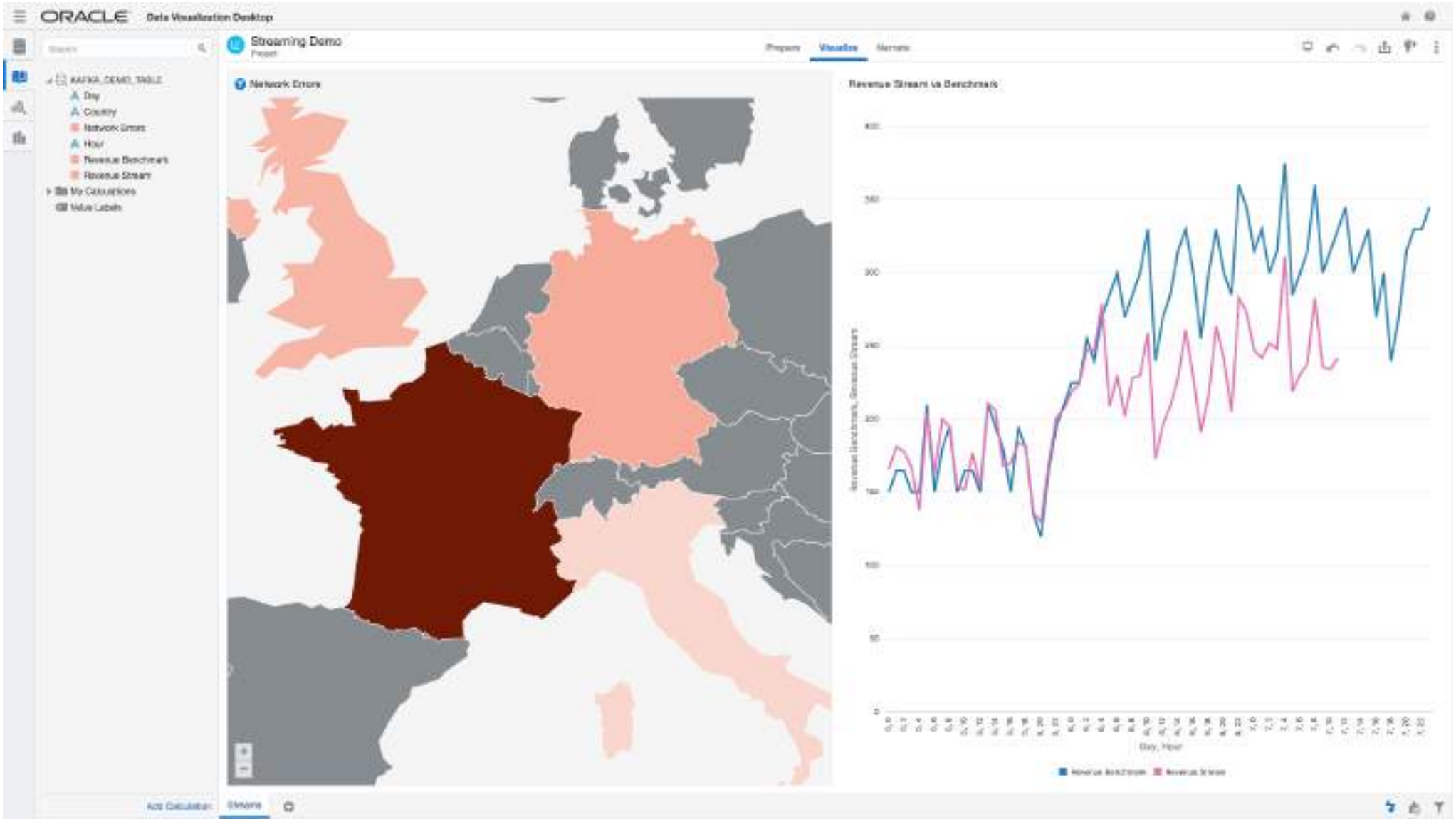

ORACLE Data Visualization Desktop
Streaming Demo
Network Errors
Revenue Stream vs Benchmark
Current Network Error Stream
Start Streams
Compare Sales Stream to Benchmark (History)

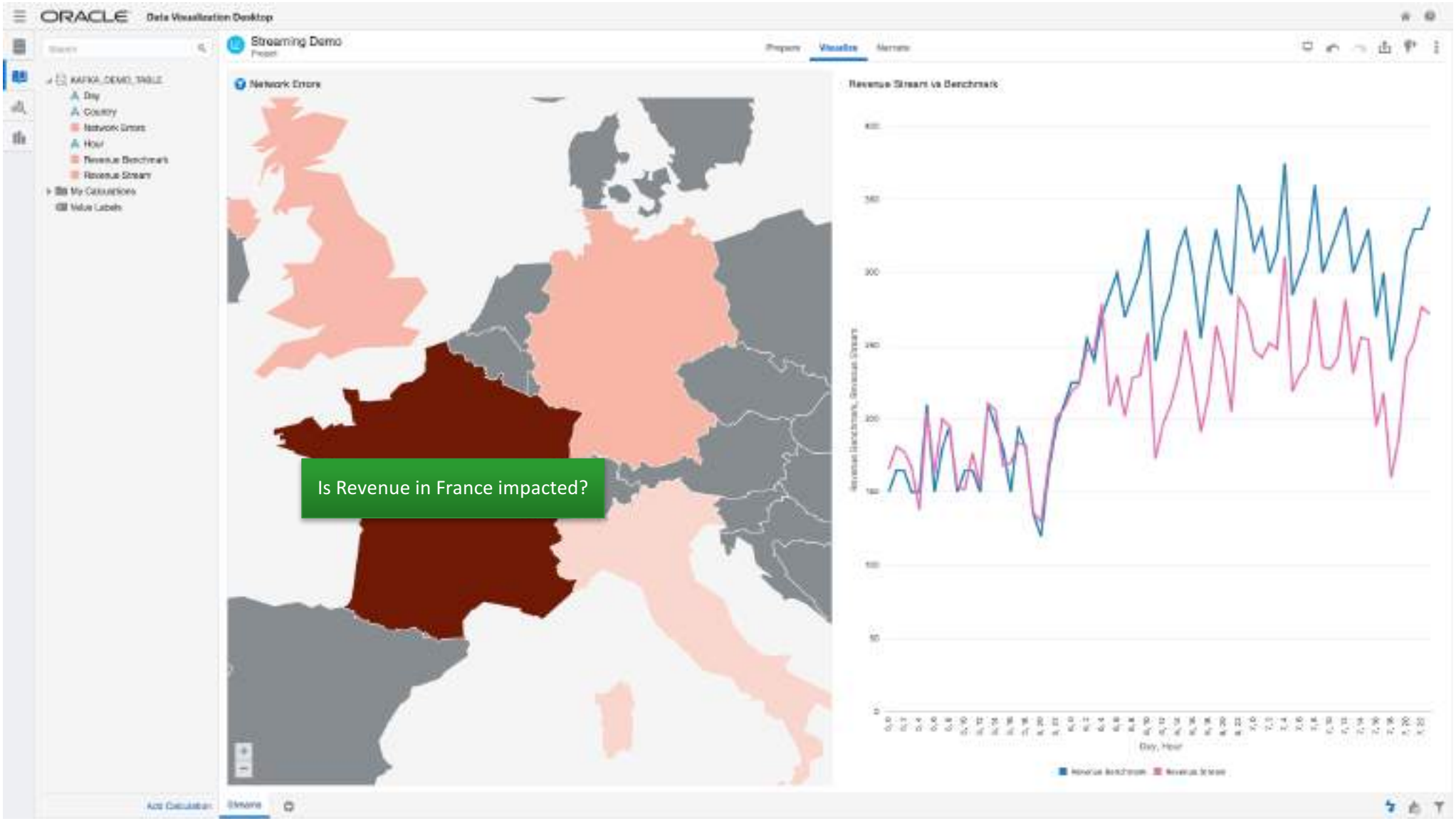

ORACLE
Data Visualization Desktop
Streaming Demo
Network Errors
Revenue Stream vs Benchmark
Day, Hour

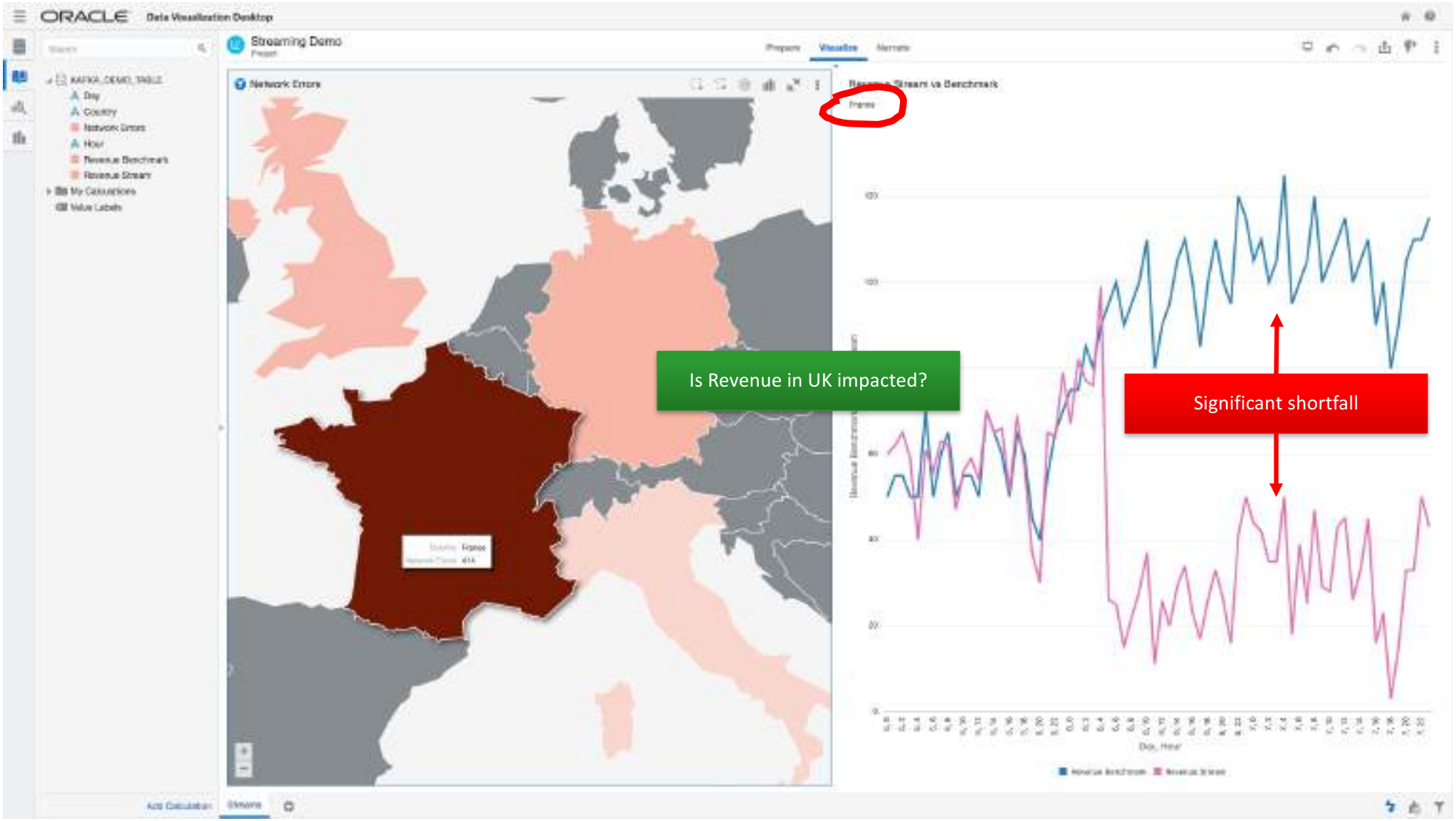

ORACLE
Data Visualization Desktop
Streaming Demo
Prepare
Visualize
Narrate
KAFKA_DEMO_TABLE
Day
Country
Network Errors
Hour
Revenue Benchmark
Revenue Stream
My Calculations
Value Labels
Network Errors
Revenue Stream vs Benchmark
Day, Hour
Add Calculation
Errors

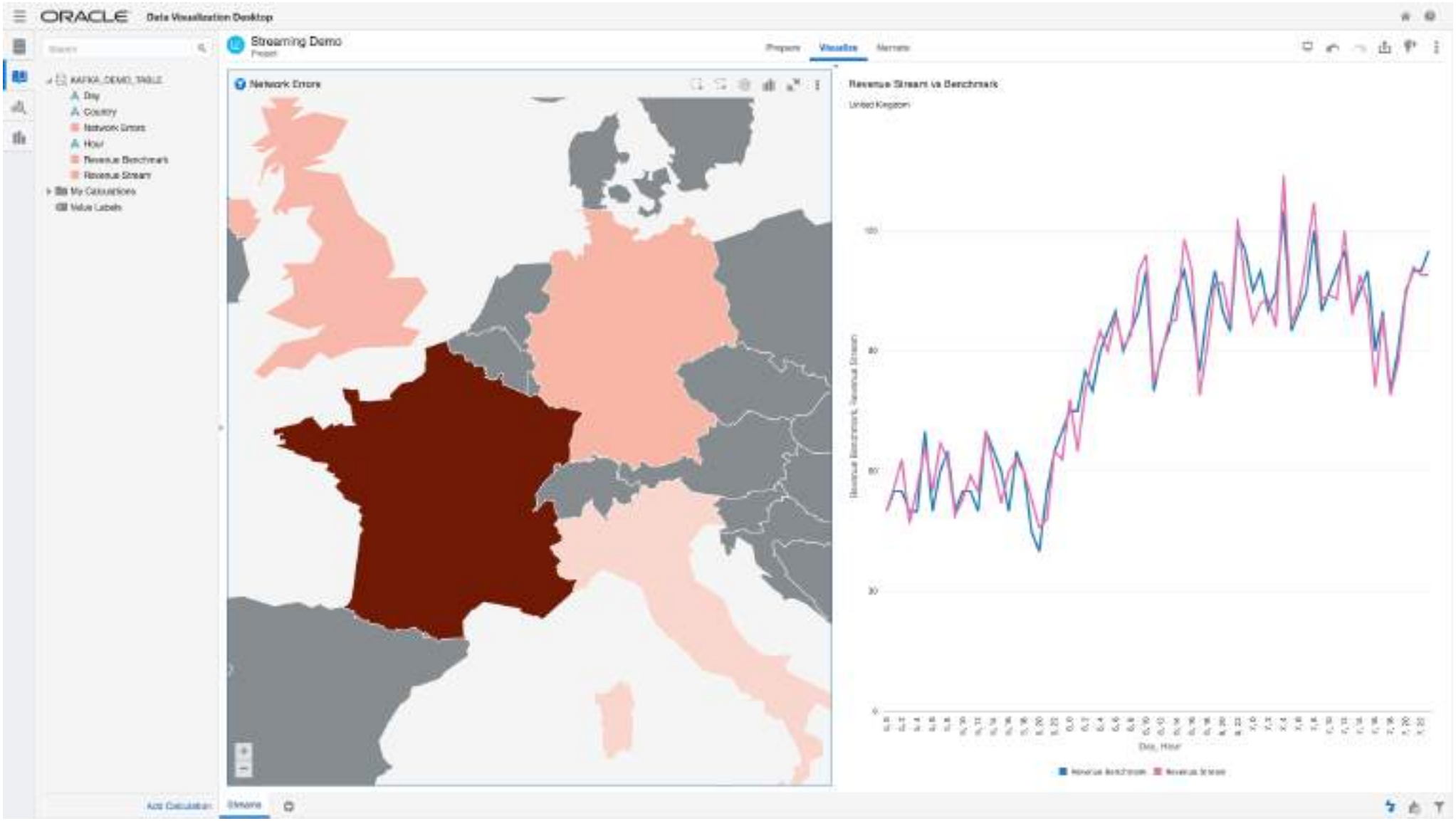

ORACLE Data Visualization Desktop
Streaming Demo
Network Errors
Revenue Stream vs Benchmark
Intermittent Network Errors in UK, Italy & Germany
French Network Errors continue to worsen
Revenue no longer meeting historical benchmark


ORACLE
Data Visualization Desktop
Streaming Demo
Network Errors
Revenue Stream vs Benchmark
Day, Hour


ORACLE
Data Visualization Desktop
Streaming Demo
Network Errors
Revenue Stream vs Benchmark
Day, Hour


ORACLE
Data Visualization Desktop
Streaming Demo
Network Errors
Revenue Stream vs Benchmark
Is Revenue in France impacted?


ORACLE
Data Visualization Desktop
Streaming Demo
Network Errors
Is Revenue in UK impacted?
Significant shortfall


ORACLE
Data Visualization Desktop
Streaming Demo
Network Errors
Revenue Stream vs Benchmark
United Kingdom

Agenda

- Vision for DW, Big Data and Analytics
- Enhancing the foundation: New DW features in 18c
- Building into Autonomous: Autonomous Data Warehouse Cloud

Key 18c Features to Enhance Data Warehouses

NEW IN 18c

DATABASE INFRASTRUCTURE

- Autonomous health framework
- Standby support for nologging ops

DATA MANAGEMENT

- Online partition merge
- Partitioning scheme online conversion of a table
- Private temporary tables

DATA PROCESSING

- SQL language
  - Enhanced **approximate query** processing
  - Enhanced **analytic views**, incl. MDX interface

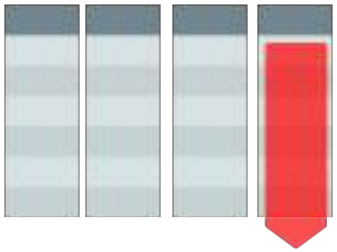
DATA PROCESSING, cont.

- **In-Memory**
  - In-memory external tables
  - In-memory for NVRAM
  - Automatic in-memory management
- Optimizer and SQL processing
  - **Inline external tables**
  - **Polymorphic table functions**
  - Better fine-grained cursor Invalidation
  - Enhanced SQL Plan Management
- And much more …

ORACLE

In-Memory Key Features

NEW IN 18c



- Automatic In-Memory Management
- In-Memory For External Tables
  - Support for all storage tiers
- In-Memory for Extreme Capacity NVRAM Memory
  - Dramatic increase of in-memory capacity
- Warehouse-specific performance improvements:
  - In-Memory Dynamic Scans
  - In-Memory Optimized Arithmetic

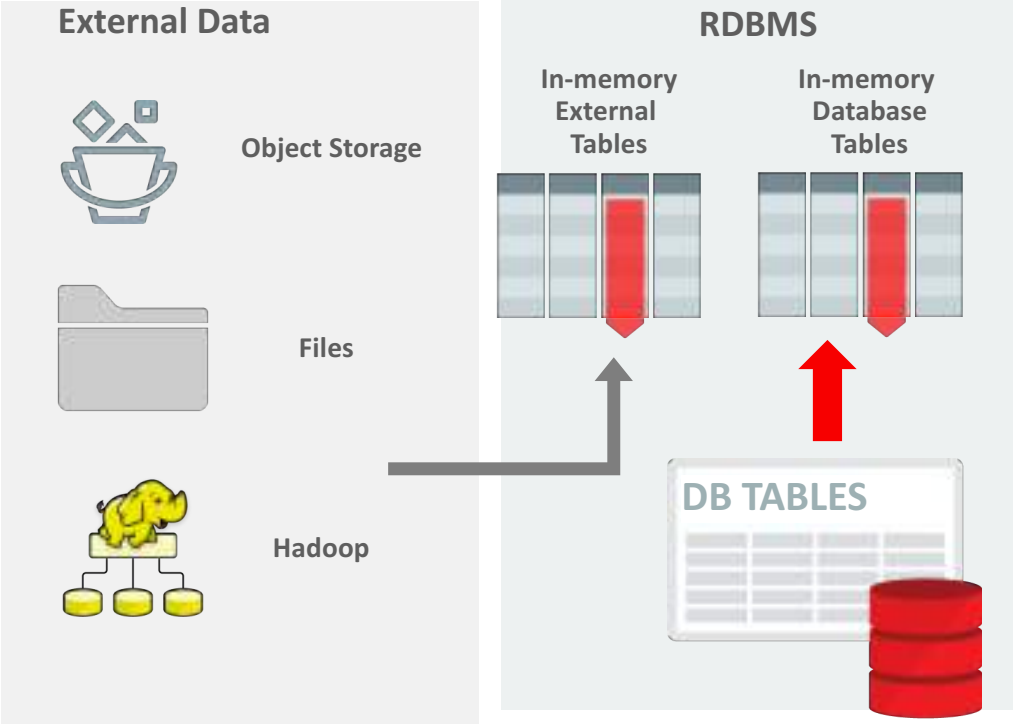
For more details, see “Oracle Database In-Memory Deep Dive: Past, Present, and Future” (Tues, 11:30am)

ORACLE

In-Memory For External Tables

Fast Analytics on External Data

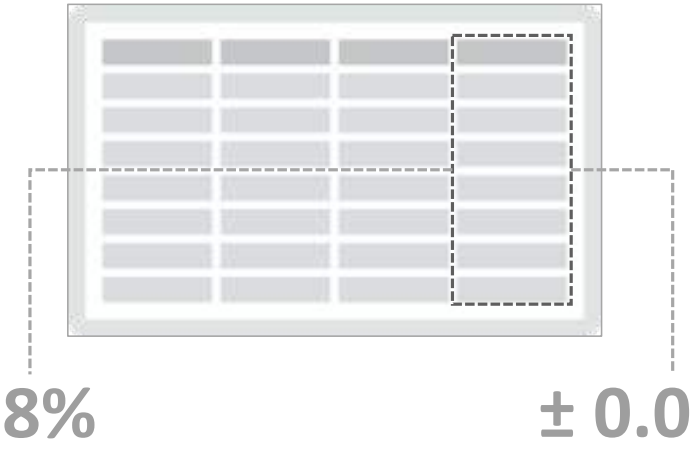
NEW IN 18c



- External Tables allow transparent access to data outside the DB
- In-Memory For External Tables builds in-memory column cache of data outside the DB for ultra-fast analytics on external data
- All In-Memory Optimizations apply
  - Vector processing, JSON expressions extend transparently to external data
- Up to 100X faster

Approximate Query Processing

Delivers significantly **faster** analysis for **interactive** and highly **iterative** data exploration



- Approximations for expensive aggregate calculations:
  APPROX_COUNT_DISTINCT
  APPROX_PERCENTILE
  APPROX_MEDIAN
  - 6-13X faster, accuracy typically within < 1%
- Use with ZERO code changes
  - approx_for_aggregation = TRUE
- Accuracy and error rate provided

Top-N approximate aggregation

NEW IN 18c

- Approximate results for common top n queries
  - Approximately how many page views did the top five blog posts get last week?
  - What were the top 50 customers in each region and their approximate spending?
- Order of magnitudes faster processing with high accuracy (error rate < 0.5%)
- New approximate functions APPROX_COUNT(), APPROX_SUM(), APPROX_RANK()

Top 5 blogs with approximate hits

```
SELECT blog_post, APPROX_COUNT(*)
FROM weblog
GROUP BY blog_post
FETCH FIRST 5 ROWS ONLY;
```

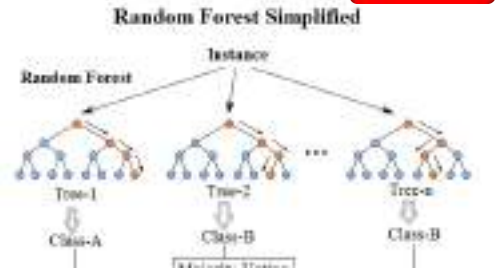
Top 50 customers per region with approximate spending

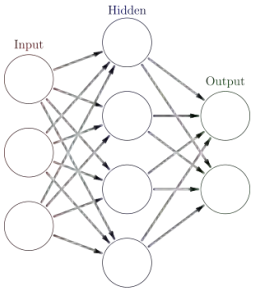
```
SELECT region, customer_name,
       APPROX_RANK(PARTITION BY region
                   ORDER BY APPROX_SUM(sales) DESC) appr_rank,
       APPROX_SUM(sales) appr_sales
FROM sales_transactions
GROUP BY region, customer_name
HAVING APPROX_RANK(..) <=50;
```

Oracle Advanced Analytics in 18.1

NEW IN 18c

- New Scalable Machine Learning Algorithms (SQL API)
  - Random Forests for Classification
  - Neural Networks for both classification and regression
  - Explicit Semantic Analysis ML algorithm extended to support classification
  - Time Series via Exponential Smoothing
  - CUR decomposition-based algorithm for attribute and row importance
- Ability to export ML models to C and Java for applications deployment





Analytic Views

- Moves business logic (Aggregations, Hierarchies, Calculations) back into database
- Simple SQL for complex analytic queries
  - no joins or GROUP-BY clauses necessary
- Works on top of pre-existing tables or views
  - no persistent storage
- Built-in data visualization via APEX

For more details, see “Using Analytic Views to Enhance BI Applications and Simplify Development” (Wed, 11:00am)

Analytic Views

Easier Access To You Data

Your Data

Analytic View

Organized & Enhanced

SQL

Simple SQL

Your Applications

Demo

“Standard” and Analytic Views

	“Standard” View	Analytic View
Data Sources (FROM)	Yes	Yes
Joins	Yes	Yes
Business Model-Based Calculations	No	Yes
Automatic Hierarchical Columns	No	Yes
Automatic Multi-Level Aggregation	No	Yes
Automatic Filter Expansion	No	Yes
Automatic Outer Join	No	Yes
Automatic Order of Calculation	No	Yes
Presentation Metadata	No	Yes

Analytic Views enhancements in 18.1

NEW IN 18c

- More calculations within Analytic Views:
  - Ranking and statistical functions
    - RANK_*, PERCENTILE_*, STATS_*, COVAR_*
  - Hierarchical expressions
    - HIER_DEPTH, HIER_LEVEL, HIER_MEMBER_NAME, etc
- Broader schema support for Analytic Views:
  - Snowflake schemas; flat/denormalized fact tables (in addition to star schemas)
- More powerful SQL over Analytic Views:
  - Dynamic definition of calculations within SQL queries

ORACLE

MDX Query Language with 18.1 Analytic Views

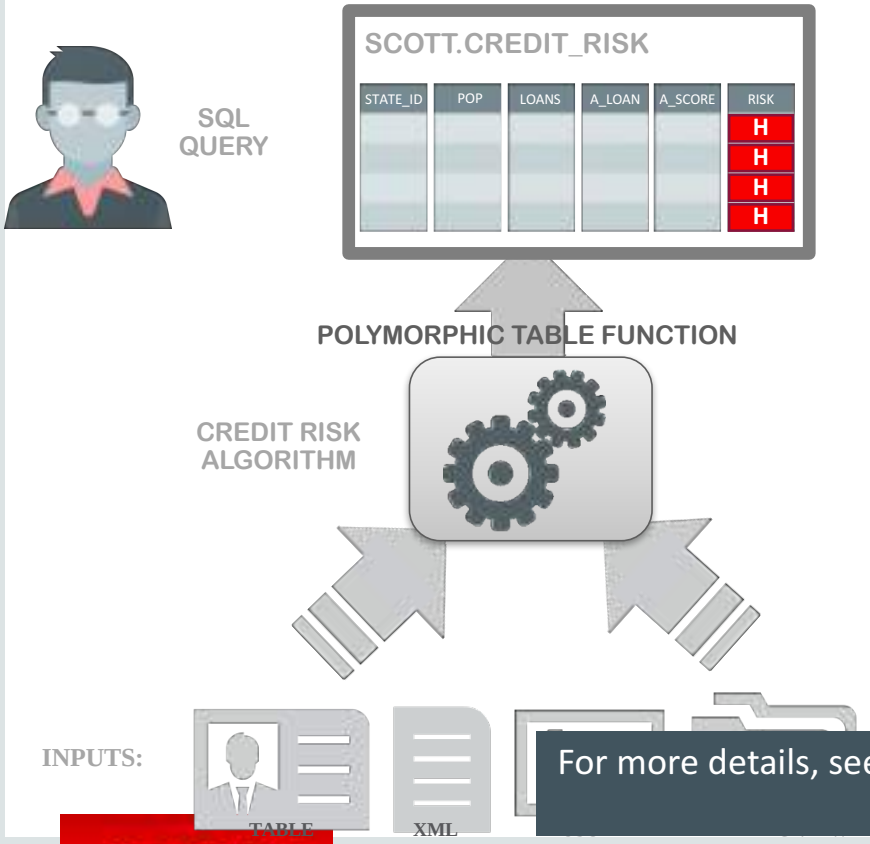
NEW IN 18c

- Support for MDX (Multi-Dimensional Expression) query language
  - Initially certified for use by Microsoft Excel Pivot Tables
    - Support/certification for other applications to follow
  - Includes a multi-dimensional query cache
    - Similar to the SQL Result Cache


```
SELECT
    {[Measures].[Sales],
      [Measures].[Units_Sold]} ON COLUMNS,
    {[Time].[Calendar].[Year].&[2014],
      [Time].[Calendar].[Year].&[2015]} ON ROWS
FROM [Sales_View]
WHERE ([Customer].[Region].[North America],
      [Product].[Departments].[Category].&[Cameras])
```

Polymorphic Tables: Self-Describing, Fully Dynamic SQL

NEW IN 18c



- Part of ANSI 2016
- Encapsulate **sophisticated algorithms**
  - Hides implementation of algorithms
  - Leverage powerful, dynamic capabilities of SQL
  - Pass in any table-columns for processing
  - Returns SQL rowset (table, JSON, XML doc etc)
    - E.g. return credit score and associated risk level

```
SELECT
 state_id, . . .,  AVG(credit_score), risk
FROM CREDIT_RISK(
  tab  => scott.customers,
  cols => columns(dob, zip,loan_default),
```

For more details, see "Building Agile, Self-Describing In-Database Analytical Algorithms for Big Data" (Tue, 4:45pm)

Inline external tables

NEW IN 18c

- External table definition provided at runtime
  - Similar to inline view
- No need to pre-create external tables that are used one time only
  - Increased developer productivity

```
CREATE TABLE sales_xt
   (prod_id number, … )
    TYPE ORACLE_LOADER
    …
    LOCATION 'new_sales_kw13')
    REJECT LIMIT UNLIMITED );
INSERT INTO sales SELECT * FROM
sales_xt;
DROP TABLE sales_xt;
```

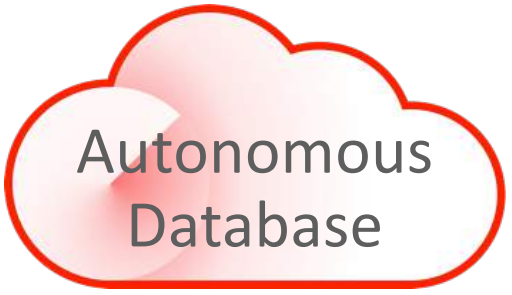
```
INSERT INTO sales
SELECT sales_xt.*
FROM EXTERNAL(
     (prod_id number, … )
      TYPE ORACLE_LOADER
      …
      LOCATION 'new_sales_kw13')
      REJECT LIMIT UNLIMITED );
```

Agenda

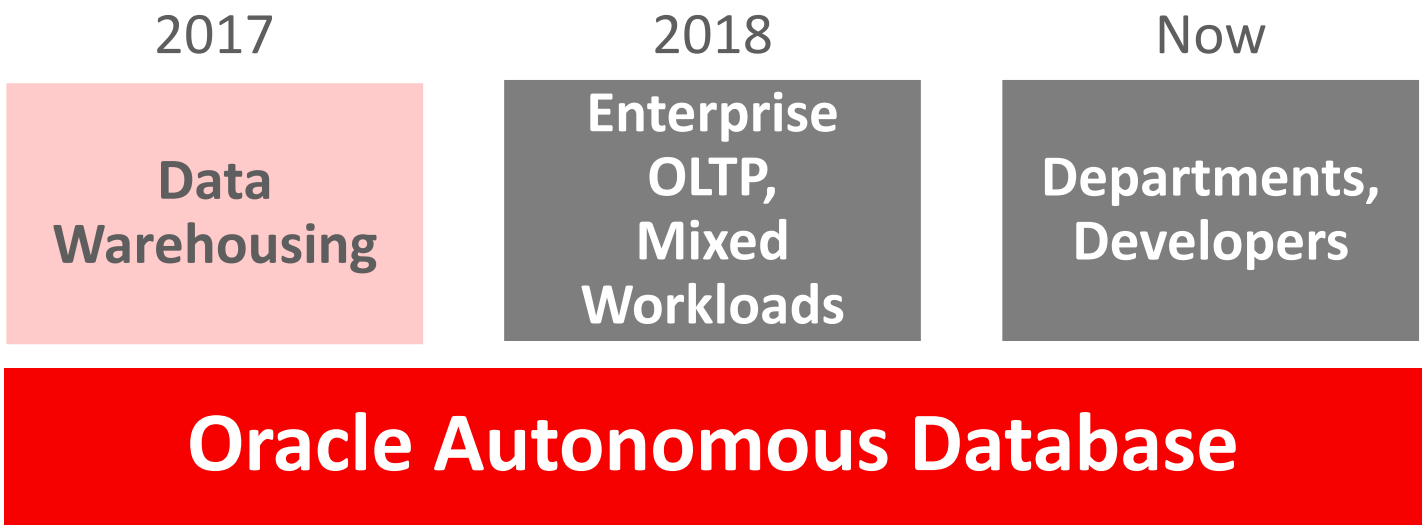
- Vision for DW, Big Data and Analytics
- Enhancing the foundation: New DW features in 18c
- Building into Autonomous: Autonomous Data Warehouse Cloud

Oracle's Vision for Autonomous Database

- Self-Driving
  - User defines service levels, database makes them happen
- Self-Securing
  - Protection from both external attacks and malicious internal users
- Self-Repairing
  - Automated protection from all downtime



One Autonomous Database – Optimized by Use Case



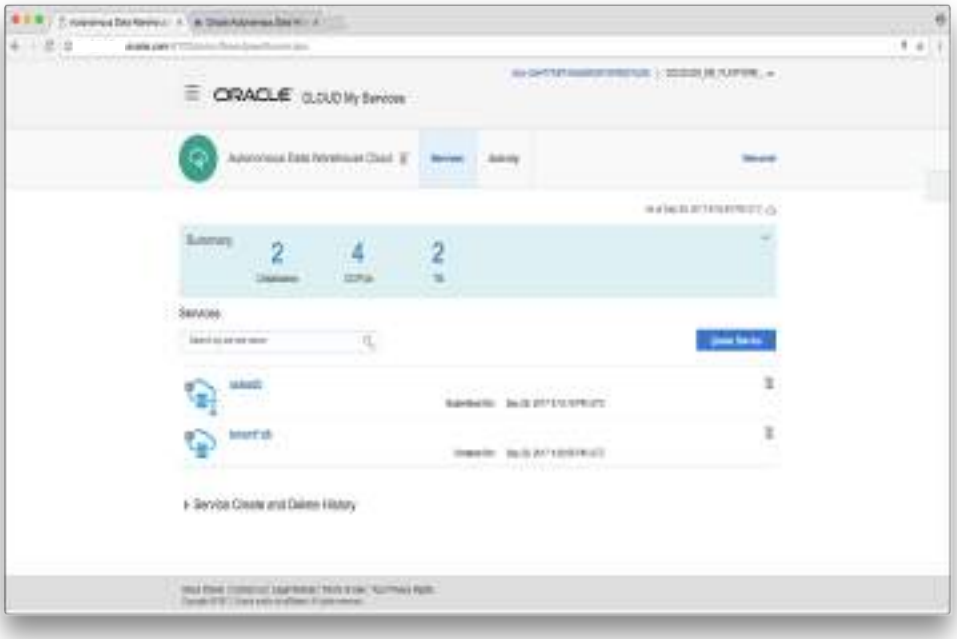
Introducing: Autonomous Data Warehouse Cloud

- Easy
  - Automated management
  - Automated tuning: Simply load data and run
- Fast
  - Based on Exadata technology
- Elastic
  - Instant scaling of compute or storage with no downtime

For more details, see “Data Warehousing for Everybody: Oracle Database Cloud Service for Data Warehousing” (Mon, 4:45pm)

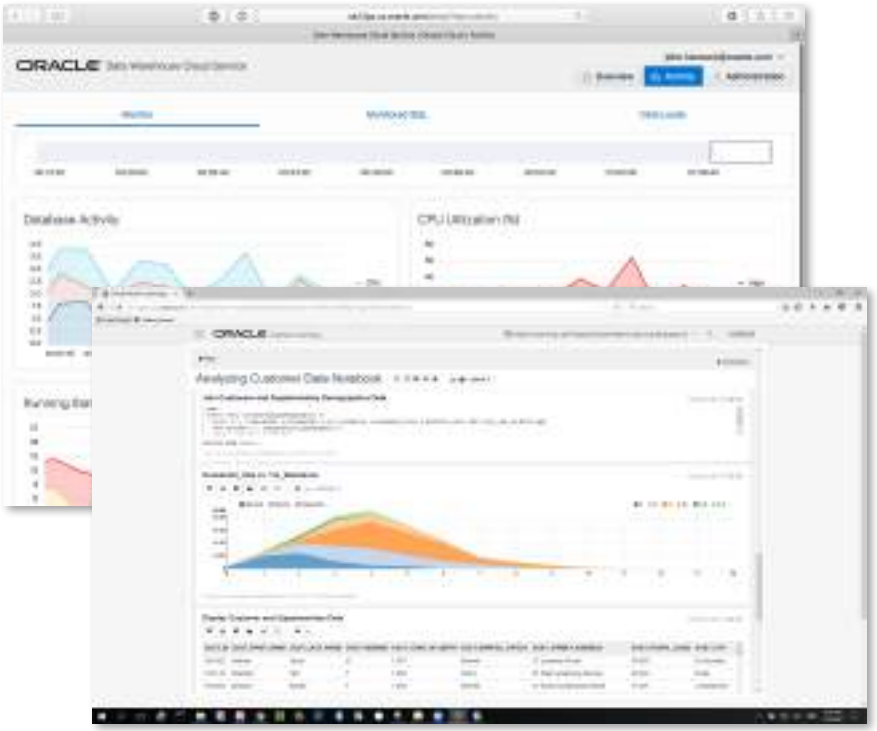
Automated management

- Oracle automates end-to-end management of data warehouse
  - Provisioning new database instances
  - Growing/shrinking storage and/or compute
  - Patching and upgrades
  - Backup and recovery
- Full lifecycle managed using ADWC Service Console



Automated Tuning

- "Load and go"
  - Define tables, load data, run queries
    - No tuning
    - No special database expertise required
  - Good performance out of the box
- Query using any business analytics tool or cloud service
  - Built-in SQL notebook also included



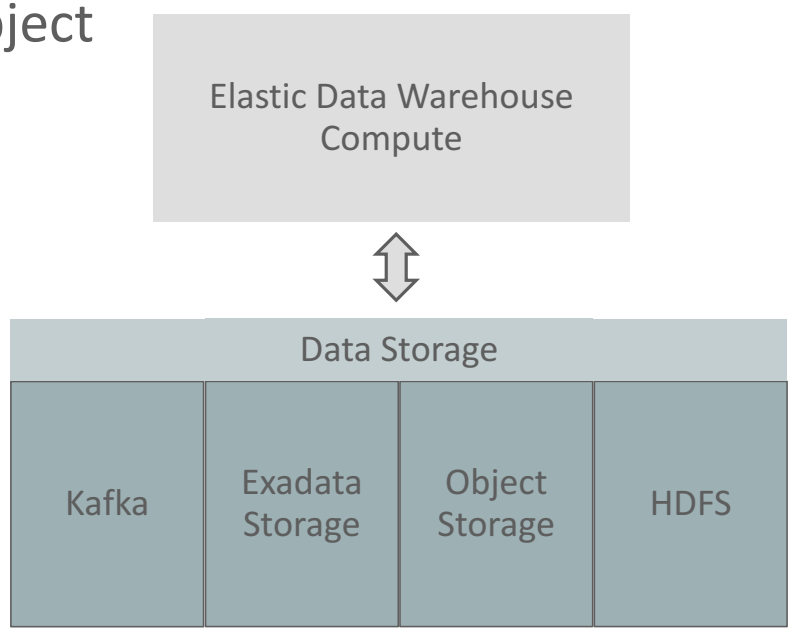
Instant Elasticity

- Customer specifies number of database cores (in OCPU) and database size (in TB)
  - CPU's and storage are independent
  - Databases can grow and shrink in either dimension
  - Pricing based upon CPU ($/CPU/hour) and Storage ($/TB/Month)
- Examples:
  - Adding new storage or compute capacity is just a click on the cloud console
  - CPU capacity can be added on-demand for end-of-quarter processing, then reduced after the processing is completed
  - CPU capacity can be set to zero, while storage remains in place (for example, to reduce costs for databases which are not used during weekends / nights)

Compute and storage processing across all dimensions

Transparent access of object store data with Autonomous Data Warehouse Cloud

- External tables can be created on data in object stores
  - Oracle Object Store or AWS S3
  - Any supported Oracle-loader file format
    - Or any Hadoop file format (e.g. Parquet) – coming soon
  - Big Data SQL integration – coming soon
- Seamless queries over object store
- Seamless queries across object store and database



Oracle 18c for Data Warehousing

1. Autonomous Database Cloud services
2. Latest Database optimizations
3. Next-generation data management architecture
   – With full integration of Big Data technologies

All of the benefits of modern architectures
… at the same time
... without compromises


ORACLE®