



ORACLE



Oracle Code Innovate Hackfest @Motorola

Make.Break.Build

Derrick Cameron, Master Principal Solution Engineer

Aakash Desai, Staff Solution Engineer

Joe Hahn, Cloud Platform Specialist

NA Cloud Solution Engineering - Hub

Feb 21, 2020

Safe harbor statement

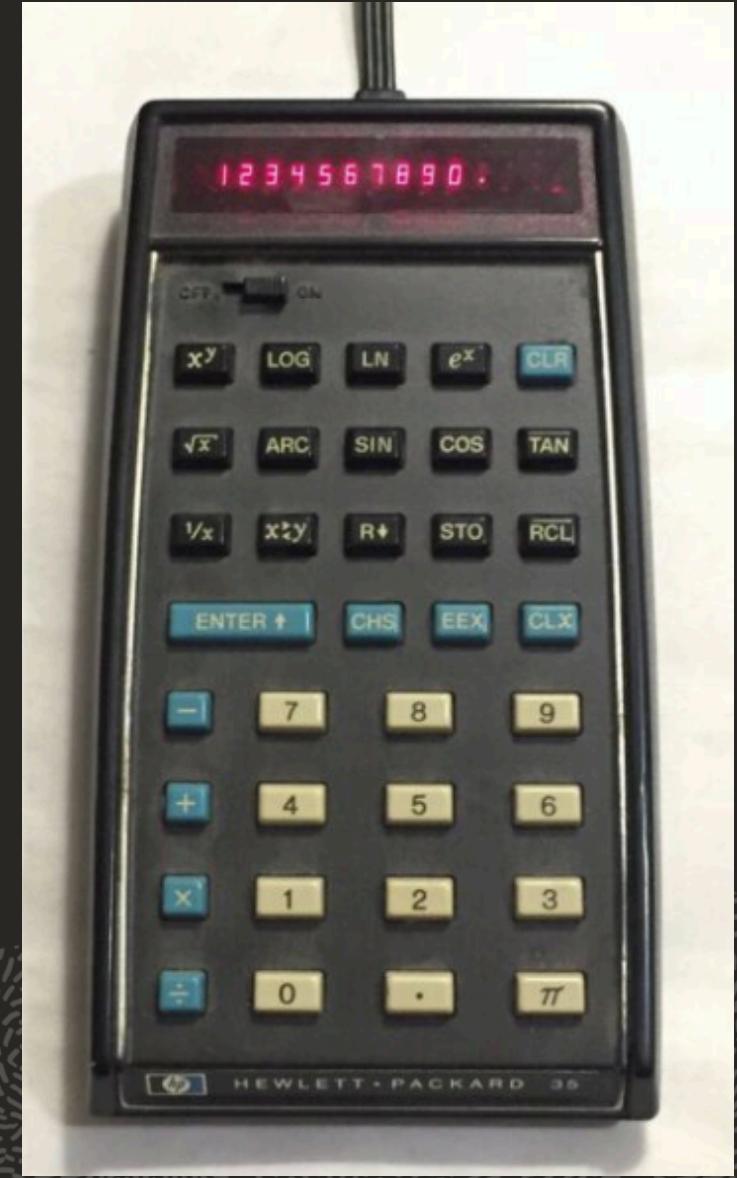
The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions.

The development, release, timing, and pricing of any features or functionality described for Oracle's products may change and remains at the sole discretion of Oracle Corporation.



—

My first exercise in exploring the science of outcome probabilities.
Use case: simulate a black jack game to determine the dealer advantage.



Decades of R&D...

- 1983 Doctoral work at MIT genesis of AI company Thinking Machines
- 1999 Oracle acquires Thinking Machines, including years of IP development
- • 2002 Oracle implements Machine Learning models in the database, and extends support to this day
- Next 20 years Oracle expands language, platform, tools, APIs, Application support

Benefits

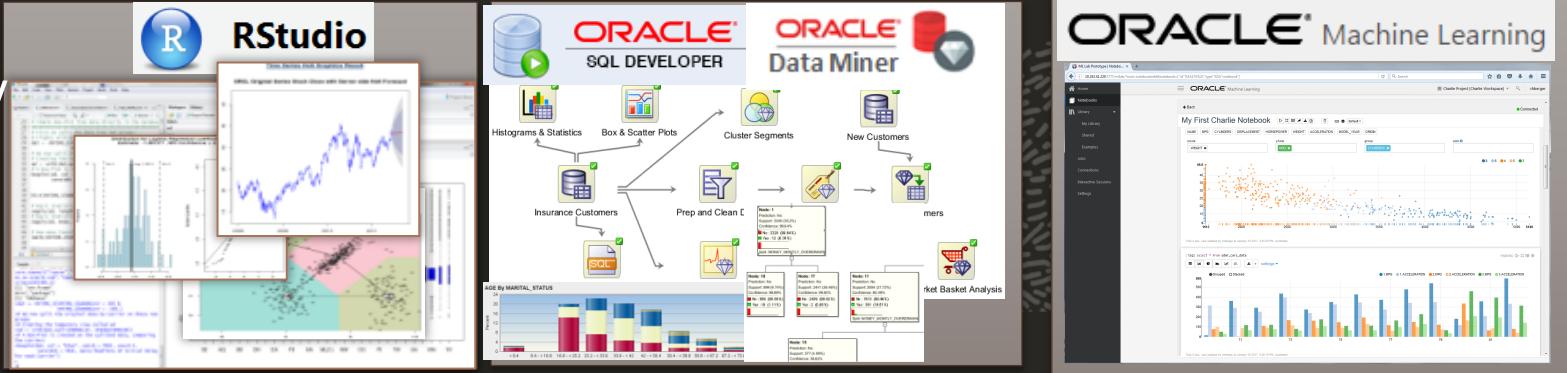
- A platform for everyone: citizen data scientists, career AI professionals, business analysts, anyone who accesses information.
- Choice of tools: Oracle Analytic Cloud, SQL Developer, Oracle Data Miner, Day by Day mobile, custom applications.
- Choice of language: SQL, Python, Java, Scala, R, Spark ...
- Choice of Data Store: Database, files, Hadoop, On-premise or cloud, Streaming data, No-SQL Databases,



Decades of R&D...

Architecturally, Many Options and Flexibility

Data Scientists, R Users, Citizen Data Scientists



SQL / R / Python

Boil down the Data Lake

Big Data SQL / R



“Engineered Features”

- Derived attributes that reflect domain knowledge—key to best models e.g.:
- Counts
- Totals
- Changes
- over time



Machine Learning Platforms

- Oracle Advanced Analytics
 - Statistics and Machine Learning Algorithms
 - Graph Analytics
 - Spatial Analytics
- Oracle Machine Learning for Python (OML4PY)
- Oracle Machine Learning for Spark (OML4SPARK)
- Big Data: Oracle R Advanced Analytics for Hadoop (ORAAH)
- Oracle Analytics Cloud: OBIEE on – premise, Data Visualization Desktop, OAC for the cloud
- Oracle Machine Learning for ADW (Zeppelin Notebooks with in-database algorithms)
- Stream Analytics
- Oracle Data Science:
 - Jupyter Notebooks
 - Multi-Language
 - All the popular data platforms

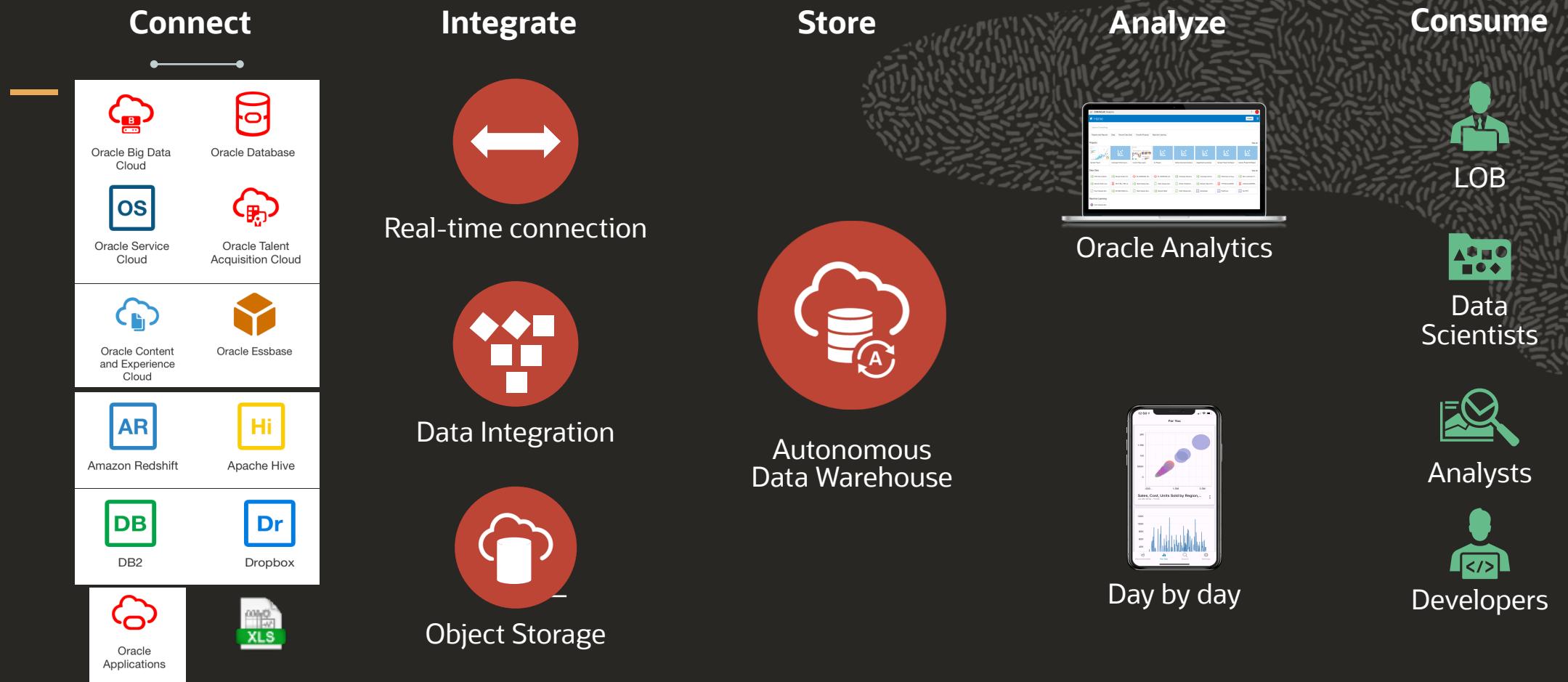


Oracle Analytics Cloud Service

An analytics platform that empowers business analysts and consumers with modern AI-powered, self-service analytics capabilities for data preparation, visualization, enterprise reporting, augmented analysis, and natural language processing/generation.



Oracle Analytics Platform



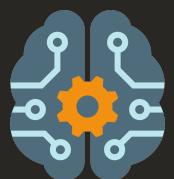
Oracle Analytics Platform



- Natural language processing and chat bot integrations



- Personalized insights and knowledge graphs



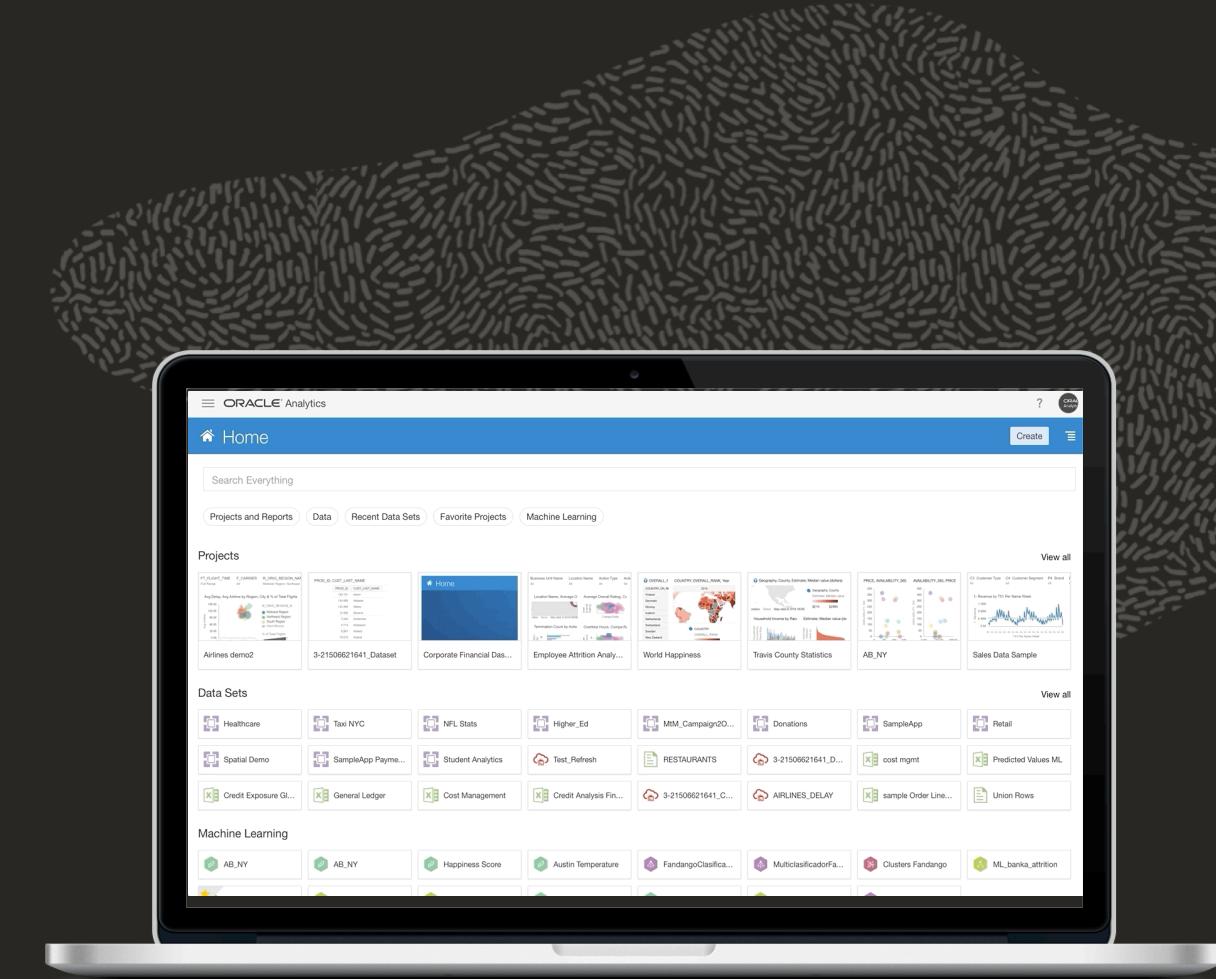
- Automated and custom machine learning



- Deep data profiling and intelligent enrichment

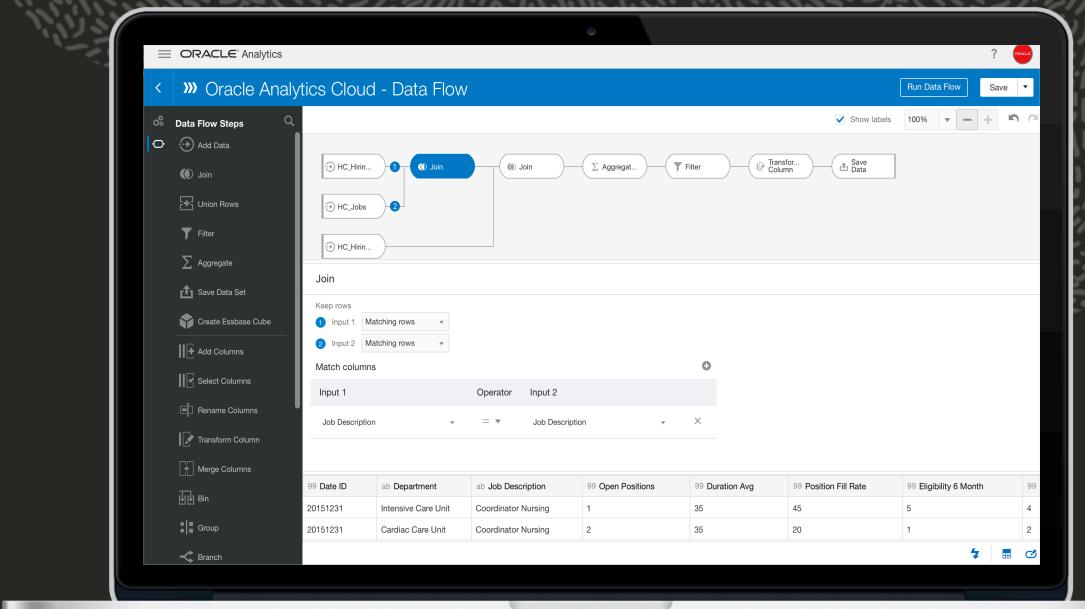
Open connectivity

- Fully integrated with Autonomous Database
- Connectors to various data sources including: Amazon Redshift, Fusion Apps, EPM Cloud, Microsoft SQL Server, Snowflake, Oracle Database and more!



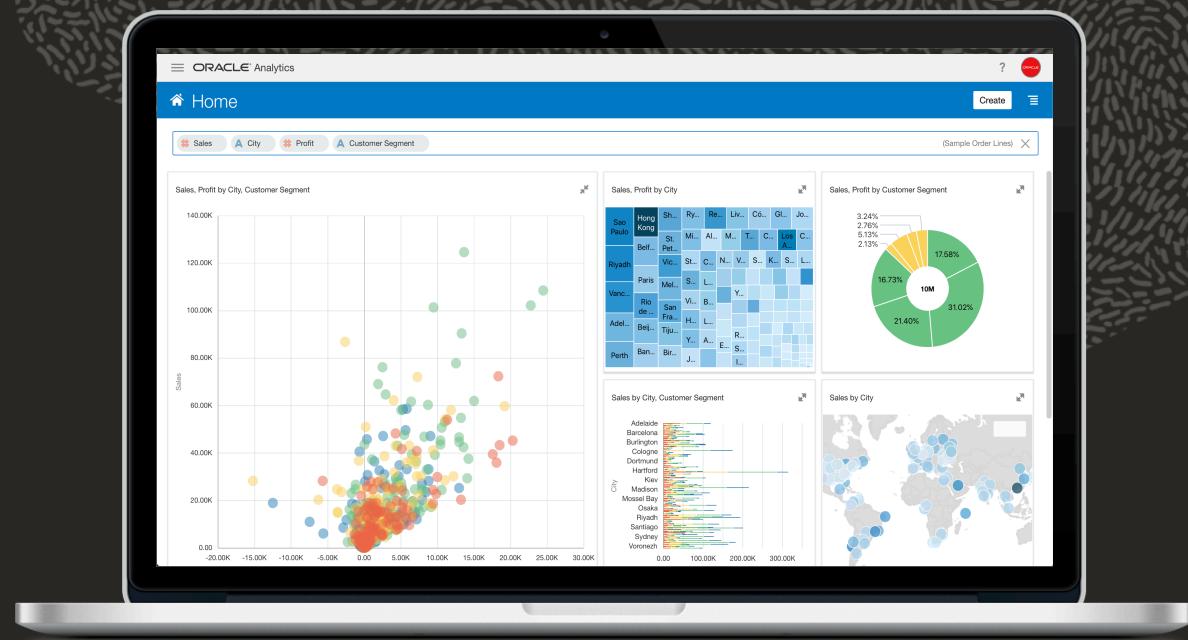
Data flow

- Use Data Flow to ingest, clean and transform easily your data from various sources and cloud!
- Group, split, join, union, merge and save data into a database or a cube



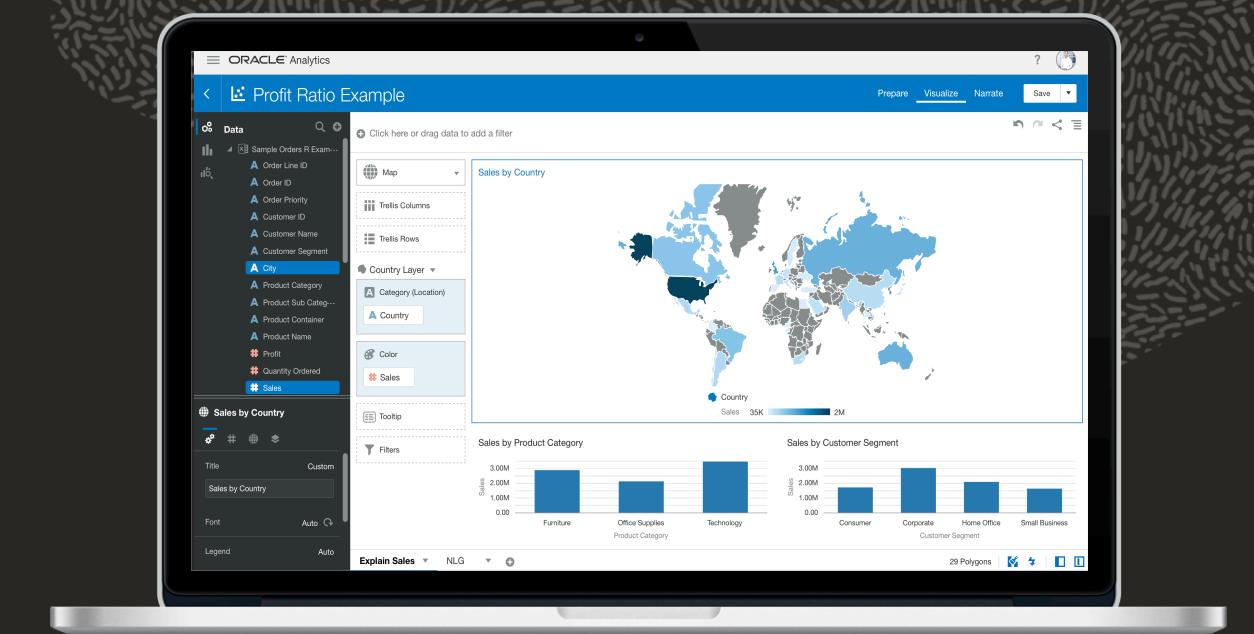
Natural language processing

- Query data directly using Natural language processing
- No Query or SQL needed
- Augmented analytics
- Voice enabled



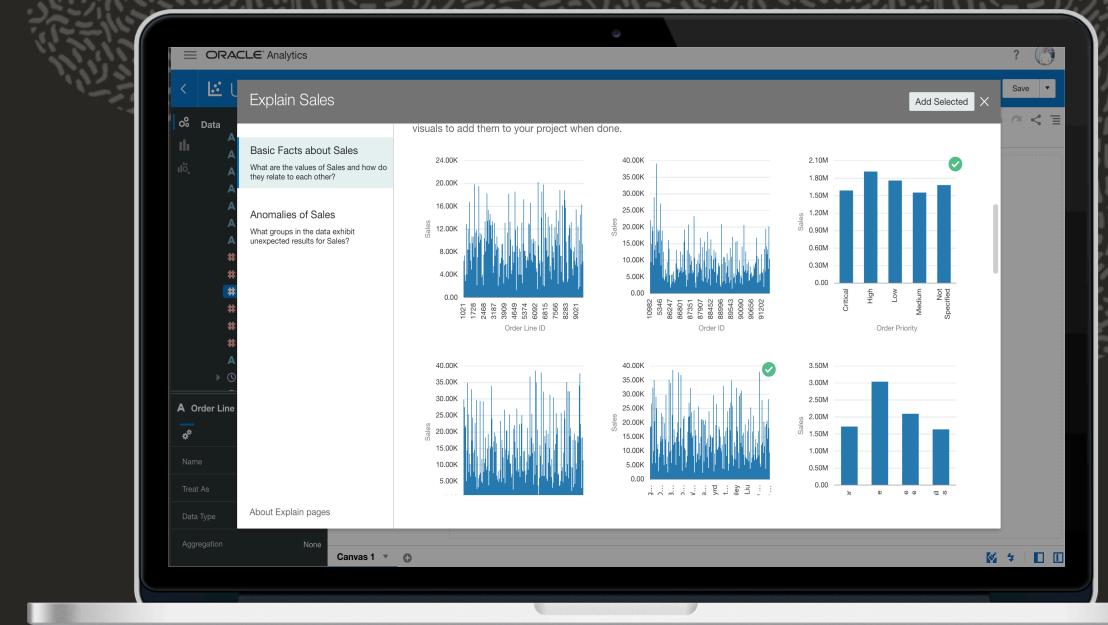
Self-service analytics

- Self-service analytics
- 1-click to forecast for predictive analytics
- What-if scenarios
- Multiple data sources join
- Data story telling



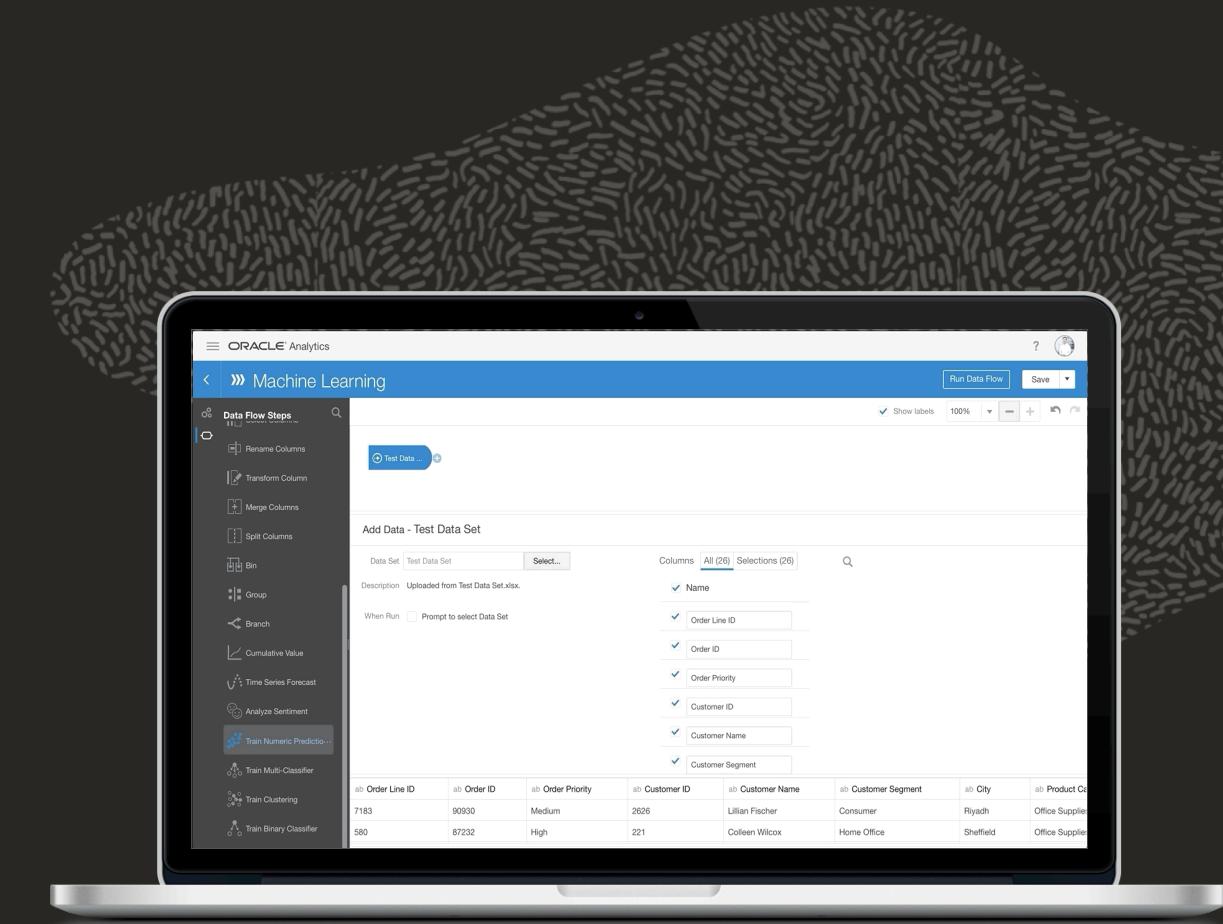
AI data exploration

- AI data exploration and insights
- Machine Learning models to provide top insights
- AI data profiling and data enrichment
- One click forecast, clustering and more



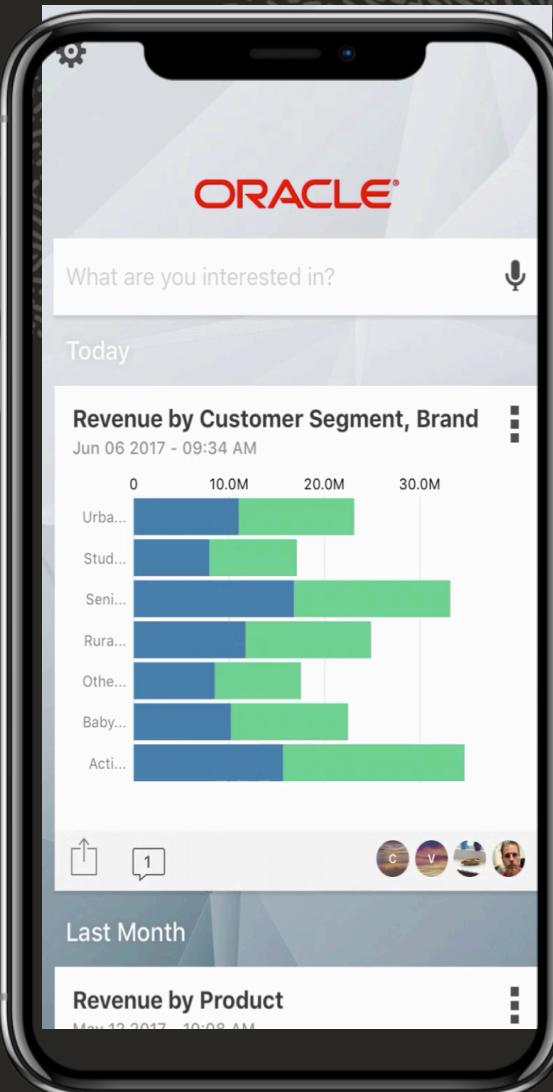
Machine learning

- Data flow to train machine learning models
- Linear Regression, Random Forest, CART, k-means clustering and more algorithms
- Machine Learning models for predictive analytics



True mobile analytics

- True mobile augmented analytics
- All data sources available to query
- Intelligent recommendations collaborative



FIFA-Worldcup-2019-Women

[Prepare](#) [Visualize](#) [Narrate](#)

Save

▼

[+ Click here or drag data to add a filter](#)

TOTAL GOALS SCORED

138

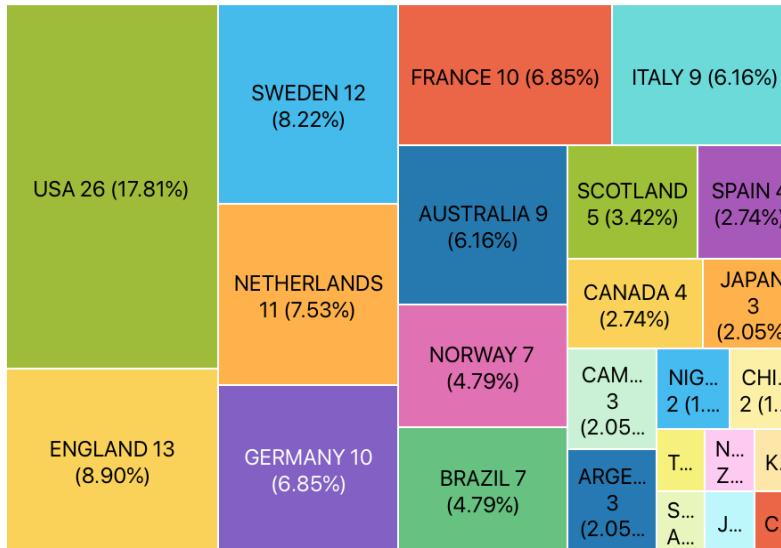
TOTAL MATCHES PLAYED

52

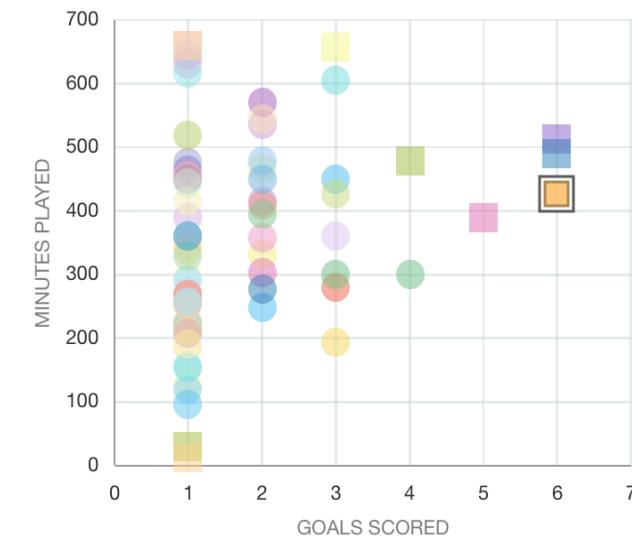
TOP PLAYER GOALS: Megan Rapinoe



GOALS BY COUNTRIES

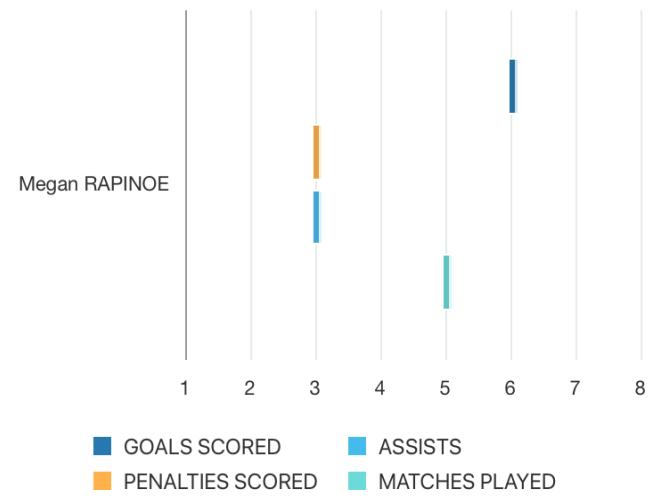


GOALS SCORED BY MINUTES PLAYED



STATISTICS

PLAYER: Megan RAPINOE



Oracle Machine Learning

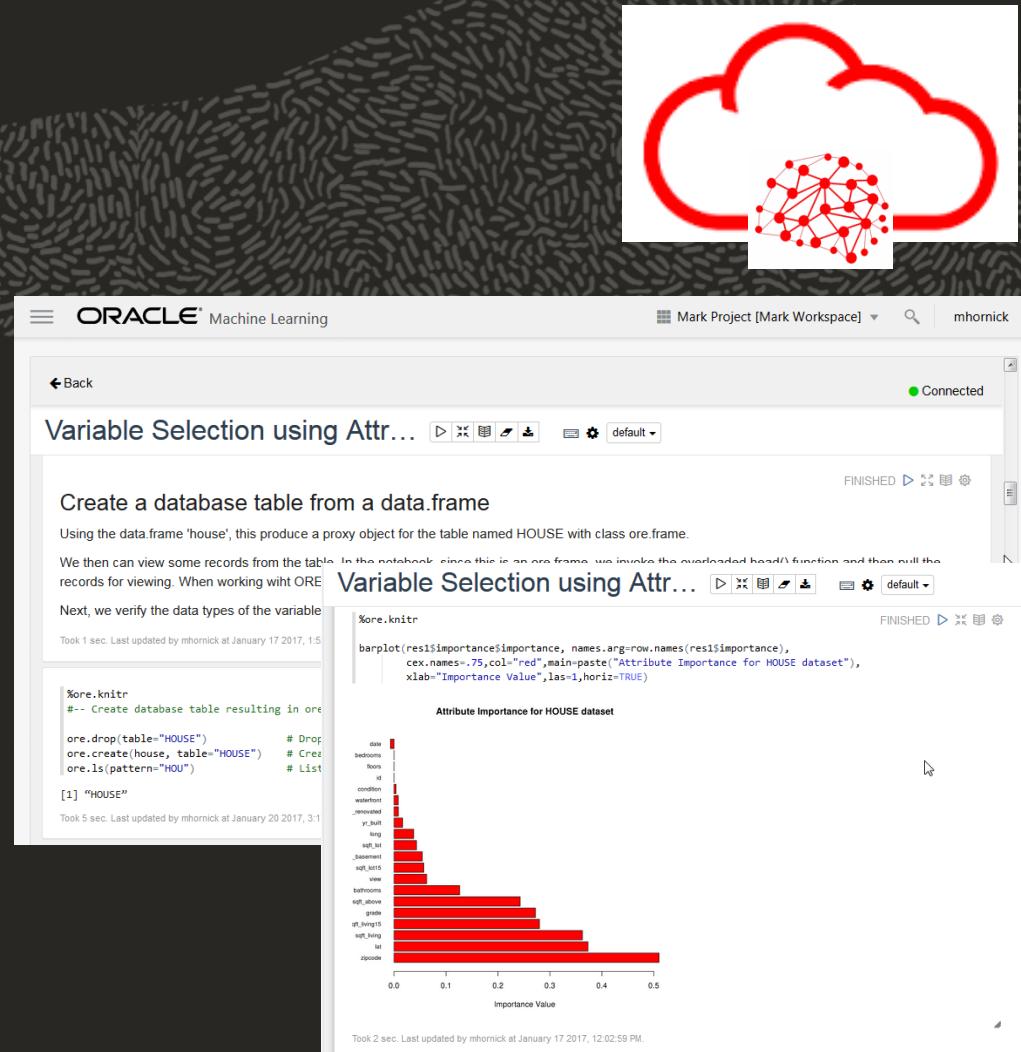
—

A collaborative Zeppelin notebook for
Oracle's machine learning capabilities



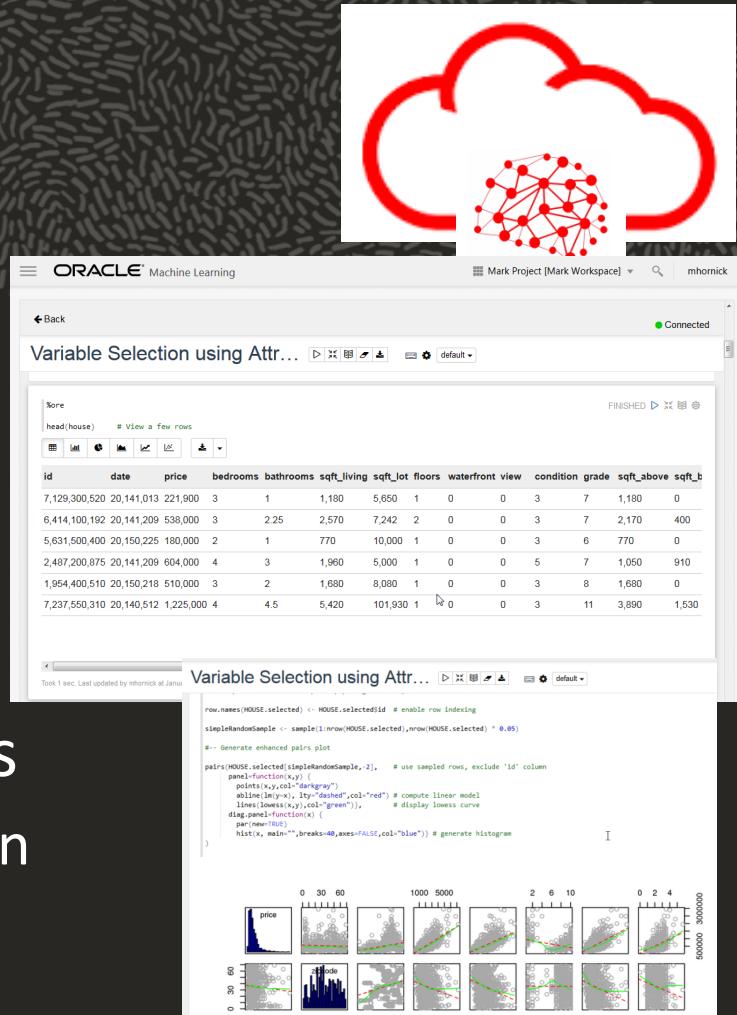
Introduction

- A collaborative unified notebook user interface for Oracle's machine learning capabilities
- Enables teams to...
 - Leverage the languages, platforms and engines of their choice
 - Explore and prepare data
 - Build, access, and use machine learning models
 - Deploy machine learning solutions



Features

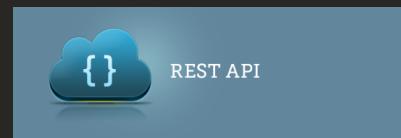
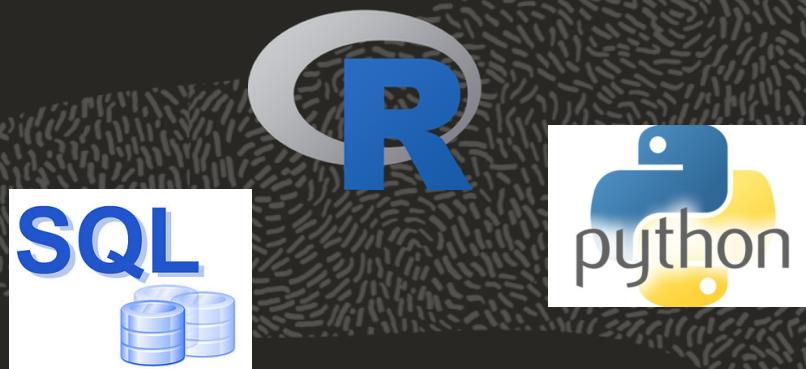
- Zeppelin-based Unifying Notebook UI
 - Multiple backend servers, analytical engines and languages
 - Oracle and open source algorithms, platforms, tools, and data sources (SQL, R, Spark, Python)
 - Provides a flexible toolbox for the data scientist
- Supports collaborative development
 - Shared notebooks and templates, with access permissions
- Supports deployment of predictive analytics solutions
 - Enables publishing libraries, templates, examples of common use cases
- Dynamic and expandable platform



Target Personas

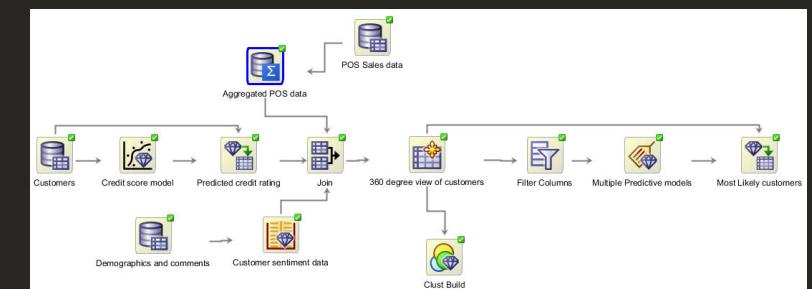
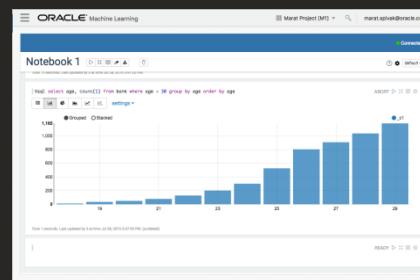
“Coders”

- Data Scientists (v1)
 - R/ORE/ORAAH, SQL/ODM
 - Python, Scala
- Database and Application Developers
 - SQL and Big Data SQL (v1)
 - REST API (v2)



“Clickers”

- Citizen Data Scientists (v3)
 - Visual Pipelines / Workflows
- Business Users (v2)
 - Lightweight Notebook Viewers



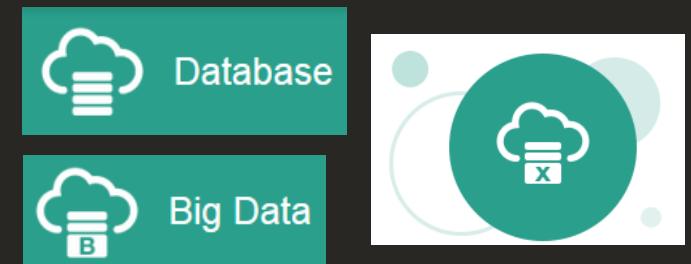
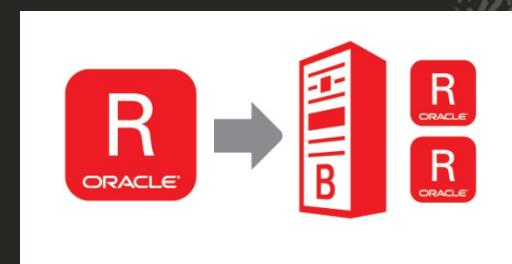
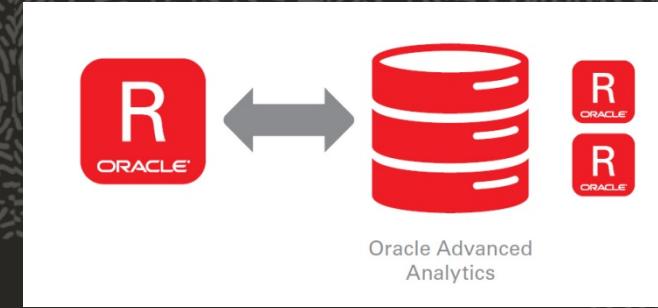
Machine Learning Models



- **Classification**
 - Logistic Regression
 - Decision Tree
 - Random Forest
 - Neural Network
 - Support Vector Machine
 - Naïve Bayes
 - Explicit Semantic Analysis
- **Clustering**
 - Hierarchical K-Means
 - Hierarchical O-Cluster
 - Expectation Maximization
- **Anomaly Detection**
 - One-Class Support Vector Machine
- **Regression**
 - Generalized Linear Model
 - Support Vector Machine
 - Random Forest
 - Linear Model
 - Stepwise Linear regression
 - LASSO
- **Feature Extraction**
 - Non-negative Matrix Factorization
 - Principal Component Analysis
 - Singular Value Decomposition
- **Algorithm Text Support**
 - Tokenization and theme extraction
 - Document similarity
 - Algorithms support text type
- **Association Rules**
 - A priori
- **Attribute Importance**
 - Minimum Description Length
 - Principal Component Analysis
 - Unsupervised Pairwise KL Divergence
- **Time Series**
 - Single Exponential Smoothing
 - Double Exponential Smoothing
- **SQL Predictive Queries**
 - Open Source CRAN R Algorithm Packages through Embedded R Execution
 - Spark MLlib algorithm integration

Apache Zeppelin

- Support for Oracle-specific and other interpreters
 - Oracle Database (SQL, PL/SQL, with Oracle Data Mining)
 - Oracle R Enterprise
 - Oracle R Advanced Analytics for Hadoop
 - Oracle Graph Server (PGX)
 - MySQL
 - Livy
- Workspaces / projects / versioning
- Advanced privilege concept for collaboration
- Deploy/schedule notebooks
 - Dynamic startup/shutdown of Zeppelin servers
- Zeppelin Docker Management
 - Spawn Zeppelin Server Instance per user per connection group
 - Auto shut down



Demonstration

Oracle Analytic Services

Thank you

**Derrick Cameron
Aakash Desai
Joe Hahn**

Oracle Cloud Solution Hub

