

It is all about Visualizing Data

Remember to complete your evaluation for this session within the app!

Session ID:

10730

Prepared by:

Francisco Munoz Alvarez
Data Intensity
Director of Innovation
@fcomunoz

10/03/2019



Francisco Munoz Alvarez

- Oracle ACE Director
- 8/9/10g/11g/12c OCP, RAC OCE, AS OCA, E-Business OCP,
- SQL/PLSQL OCA, Oracle 7 OCM
- Oracle 7, 11GR2, 12cR1, 12cR2, ADWCS and OVM 3.1 and 3.2 and 3.3
 Beta Tester/Early Adopter
- IOUC LA and APAC Spokesperson, President of APACOUC, IAOUG, CLOUG and NZOUG
- ITIL Certified
- 2008 Top OTN Forum Contributor (All year #1)
- 2008 Oracle ACE Award Winner
- 2010 Oracle Excellence Award Winner
- 2010 Oracle Magazine Editors Choice Award Winner
- 2012 LAOUC Architect of the year Winner
- 2017 APAC Webinar Tour Best Session Winner
- Presented over 420 sessions at 48 Countries around the world

Blog: oraclenz.com - Email: fmunozalvarez@dataintensity.com

Twitter: fcomunoz

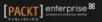
Data Intensity
Director of Innovation
www.dataintensity.com



Oracle Database 12c
Backup and Recovery
Survival Guide

A comprehensive guide for every DBA to learn recovery and backup solutions

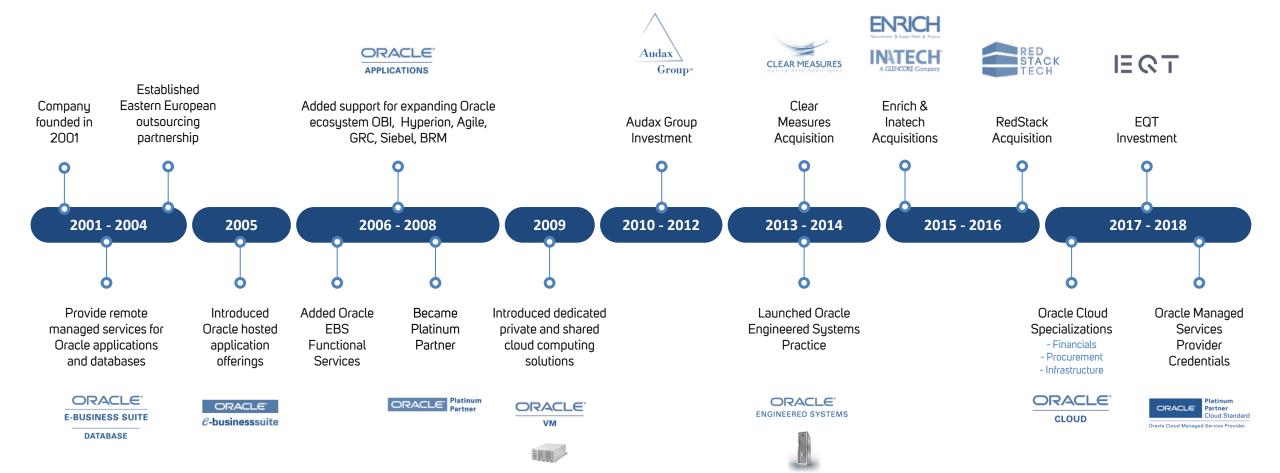
Francisco Munoz Alvarez Aman Sharma







Corporate Evolution & Strategy



Delivery & Service Capabilities That Span The Globe



Oracle Credentials & Deep Oracle Expertise











6,200+ Upgrades



+008

Professionals Worldwide





200+ **Cloud Migrations**



20+

Oracle Specializations



2,800+

Staff Certifications

SSAE 16 SOC 1 Type II ISO 9001 Certified **ITIL & ITSM Certified**

Oracle Cloud Applications

Financials Cloud **Procurement Cloud** Infrastructure Cloud

Oracle Applications

E-Business Suite R12 Financials E-Business Suite R12 Supply Chain

Oracle Engineered Systems

Exadata Database Machine X2 Exadata Database Machine Oracle Database Appliance

Oracle Platform

Linux 5/6 OVM 2/3

Oracle Technology

Oracle Cloud Managed Service Provider

ORACLE

Database 11q / 12c Database 11q / 12c Security Database 11q / 12c Data Warehousing Database 11q / 12c Performance Tuning

Platinum

Cloud Standard

Partner

Enterprise Manager 11g / 12c

Real Application Clusters 11q / 12c

Business Intelligence Foundation 10 / 11g

Data Integration 11q / 12c

WebLogic Server 12c Identity Management 12c

SOA Suite 12c

Big Data Appliance

Oracle Cloud Infrastructure

Associate Architect Certifications

Customer Advisory Board

Oracle EBS ATG Oracle Cloud Infrastructure

2017 Oracle Cloud Infrastructure

Partner of the Year – UK

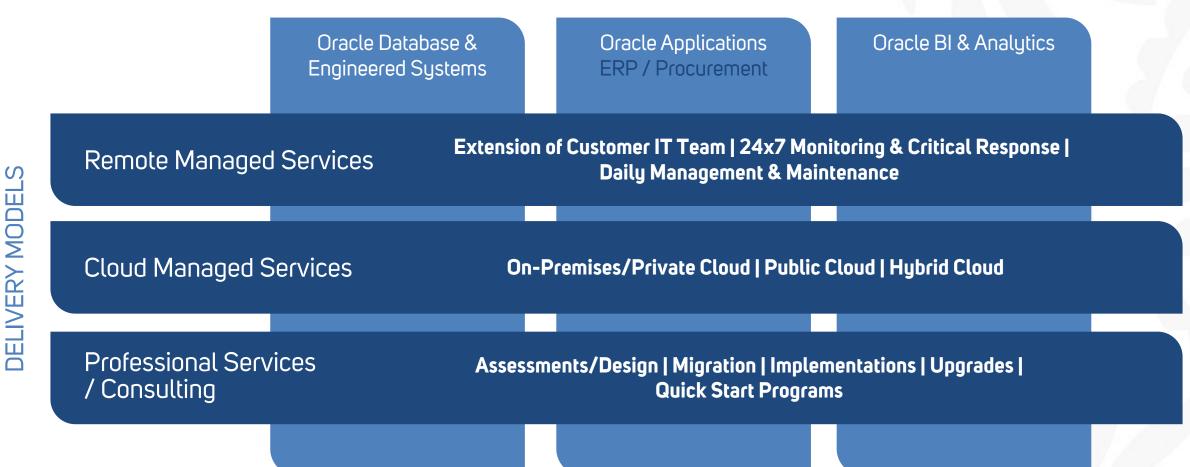
2015 Engineered **Systems**

Partner of the Year - Global

BETA Program Participant



Global Portfolio of Oracle Solutions and Services Delivered Your Way





The Data Intensity Oracle Advantage

We are an **Oracle First** service provider and believe in Oracle on Oracle for Oracle

Comprehensive portfolio of solutions from **Oracle Cloud** to **On-Premises Applications and technology.**

Oracle Platinum and Managed
Services Partner focused on
accelerated migration and
management services for Oracle
Cloud Infrastructure.



We have been selling enterprise cloud services for over 12 years and understand the sales process

We architected **OCI for our own EBS environment** and presented
it at Oracle Cloud Days

Proven with **hundreds of Oracle clients worldwide** who have entrusted us to handle the implementation, management, and support of their mission-critical Oracle applications and technologies

Oracle Capabilities & Alignment Partner for Growth

Delivery

SaaS

- Oracle ERP Cloud
 - Financials

HCM

Procurement

• PPM

OM

EPM

Java

PaaS

- Oracle Database Cloud
- BI
- Oracle Identity Cloud

Oracle SOA Cloud

- Oracle Data Integrator
- Oracle Management Cloud

laaS

Oracle Cloud Infrastructure

Capabilities

Applications (On-Premises)

- Oracle E-Business Suite Financials, SCM, Procurement
- Hyperion

PeopleSoft

ATG Web Commerce

Advanced Compression

OBIEE

BRM

- AgileGRC
- Siebel

Fusion Middleware

Database & Advanced Technology

- Oracle Database
- Advanced Security
 - •

- RAC Oracle StreamsActive
- WebLogic Suite
- Database Vault
- Advanced Compression
- WebCenterODI

Data GuardGoldenGate

• TDE

Infrastructure

- Oracle Exadata Machine
- Oracle Database Appliance
- Exalogic Elastic Cloud
- Oracle Private Cloud Appliance

Exalytics

- ZFS
- Oracle SuperCluster
- VM

OEL

Alignment to 3 Oracle Sales Plays

Move Workloads to the Cloud

Modernize Data Management

Connect & Extend Apps

Technical & Functional Integration



It is all about Visualizing Data

- Why Visualizing Data is so Important
- Introduction to Autonomous Data Warehouse Cloud Service
- Machine Learning
- Introduction to Oracle Data Visualization Desktop
- Gain insights from data in the ADWC
- Conclusion

It is all about Visualizing Data

- Why Visualizing Data is so Important
- Introduction to Autonomous Data Warehouse Cloud Service
- Machine Learning
- Introduction to Oracle Data Visualization Desktop
- Gain insights from data in the ADWC
- Conclusion

Why Visualizing Data is so Important

Data visualization is the presentation of data in a pictorial or graphical format. It enables decision makers to see analytics presented visually, so they can grasp difficult concepts or identify new patterns. With interactive visualization, you can take the concept a step further by using technology to drill down into charts and graphs for more detail, interactively changing what data you see and how it's processed.

It is all about Visualizing Data

- Why Visualizing Data is so Important
- Introduction to Autonomous Data Warehouse Cloud Service
- Machine Learning
- Introduction to Oracle Data Visualization Desktop
- Gain insights from data in the ADWC
- Conclusion

Autonomous Data Warehouse Cloud

- Easy
 - Fully-managed, pre-configured and optimized for DW workloads
 - Simply load data and run
 - No need to define indexes, create partitions, etc.
- Fast
 - Based on Exadata technology
- Elastic
 - Instant scaling of compute or storage with no downtime



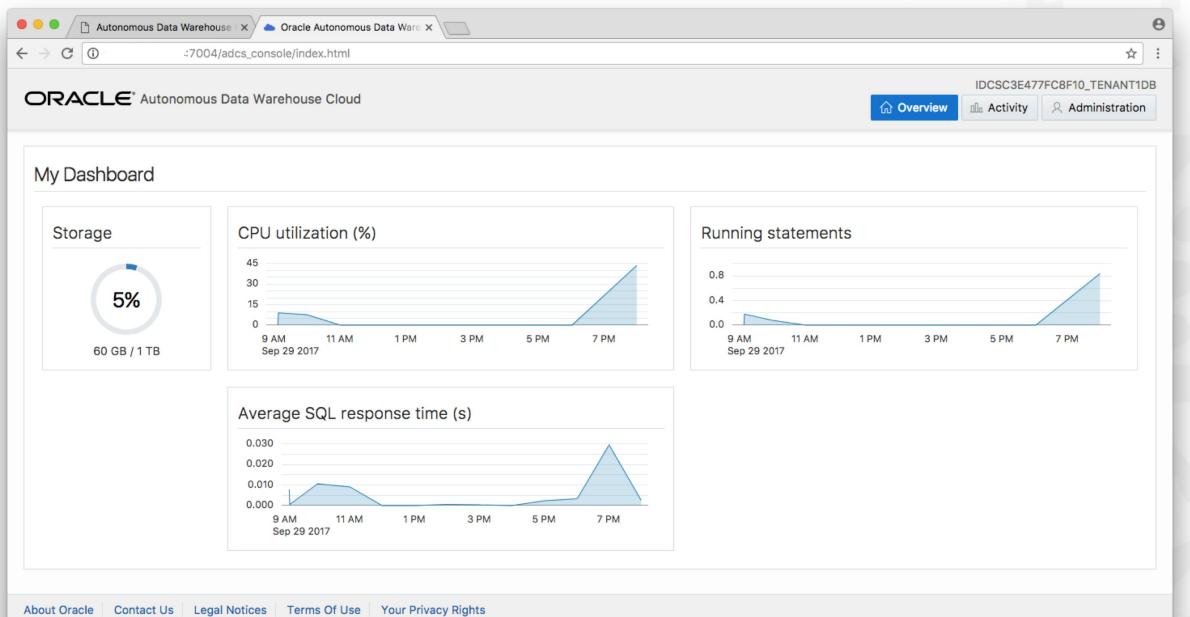


Fully Managed

- Automated end-to-end management of the DW
 - Provisioning new Databases
 - Growing/shrinking storage and/or compute
 - Patching and Upgrades
 - Backup and Recovery
- Fully lifecycle managed using the service console
 - Alternatively, can be managed via command line interface or REST API
- Load and Go
 - Fast Performance with zero tuning
 - Simple Web based monitoring console
 - Built-in resource-management plans

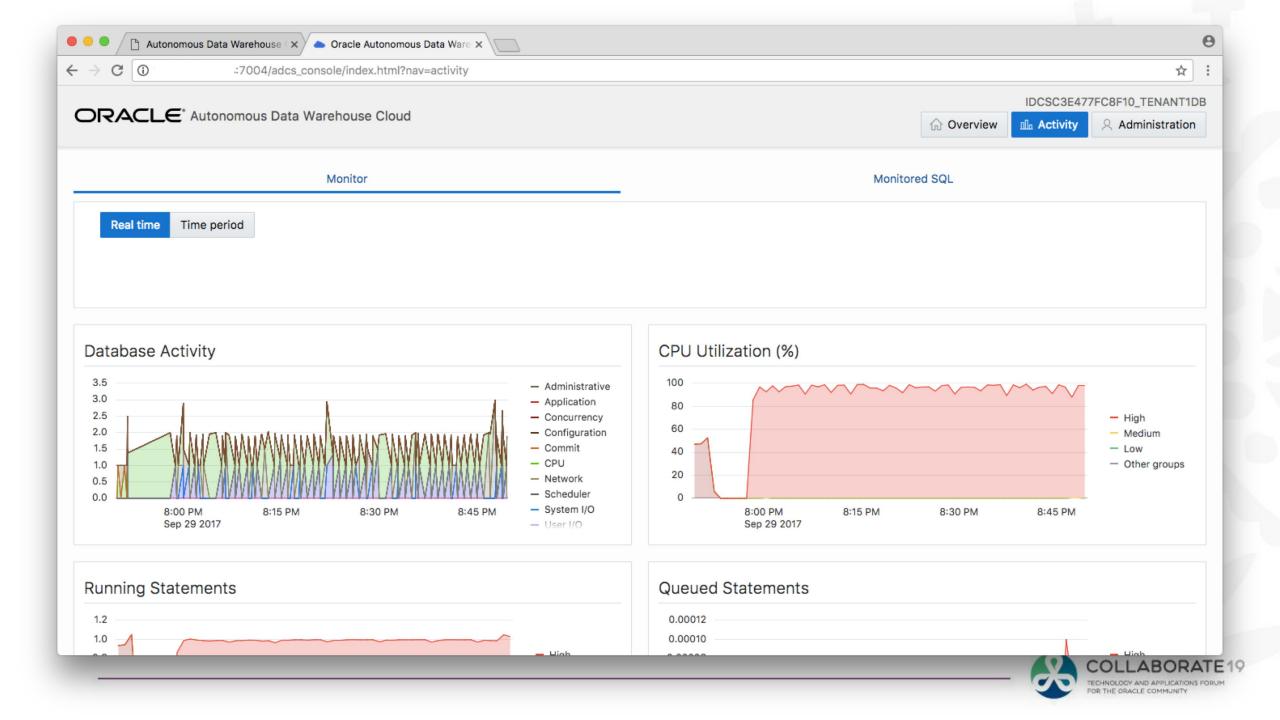




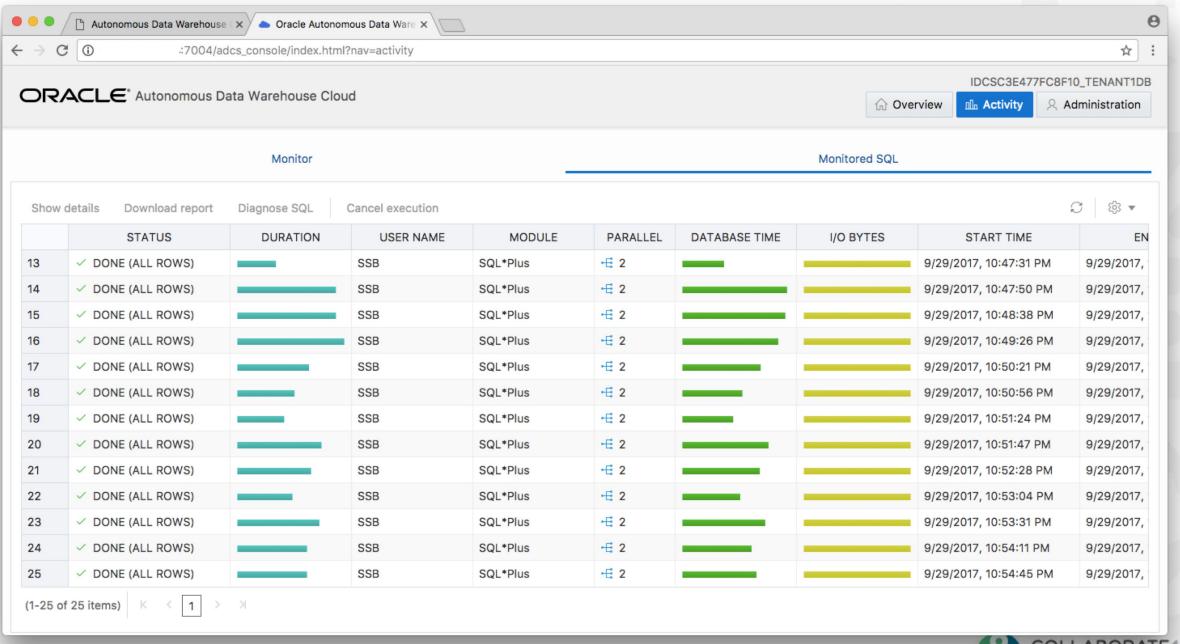


Copyright © 2017 Oracle and/or its affiliates. All rights reserved.

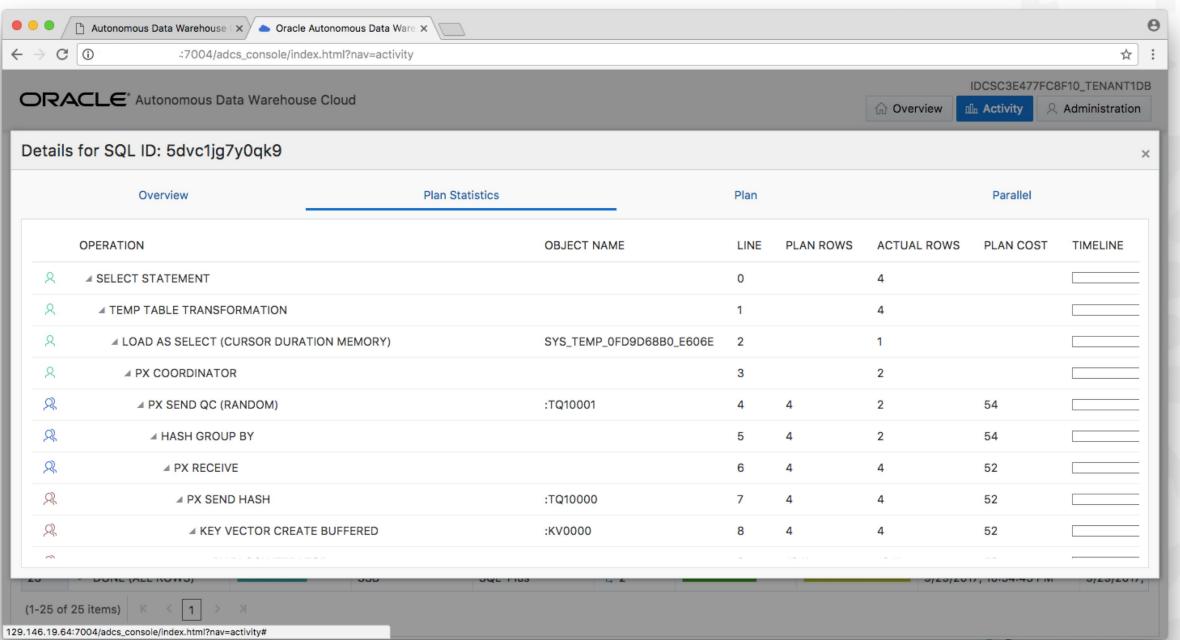












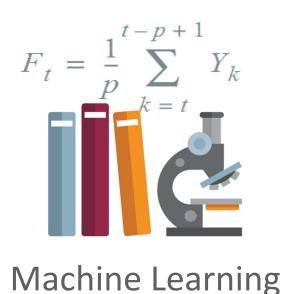
Key Use Cases for Autonomous Data Warehouse Cloud

Data Marts / Warehouses



Business Analytics

Sandboxes for Data Scientists



Data Lakes



Query Across All Data



ADW – In-Database SQL Analytics

Hierarchical Analytics

Window Functions

Forecasting

Statistics

Approximate Analytics

Pattern Matching SQL Modeling Advanced Aggregations

Ranking

Pivoting

Used-Defined PTFs

Text Analytics



ADW – In-Database SQL Analytics

Oracle Machine Learning SQL notebooks provide easy access to Oracle's parallelized, scalable in-database implementations of a library of <u>Oracle Advanced Analytics</u>' <u>machine learning algorithms</u> (classification, regression, anomaly detection, clustering, associations, attribute importance, feature extraction, times series, etc.), SQL, PL/SQL and Oracle's statistical and analytical SQL functions.

Oracle Machine Learning SQL notebooks and Oracle Advanced Analytics' library of machine learning SQL functions combined with PL/SQL allow companies to automate their discovery of new insights, generate predictions and add "AI" to data viz dashboards and enterprise applications.



ADW: In-Database Machine Learning (Built in Algorithms Methods)

Anomaly Attribute Classification Regression **Detection Importance Association Feature Descriptive** Clustering Rules **Extraction** statistics **Hypothesis** Test Text **ANOVA Distribution Fit** Mining **Testing**

It is all about Visualizing Data

- Why Visualizing Data is so Important
- Introduction to Autonomous Data Warehouse Cloud Service
- Machine Learning
- Introduction to Oracle Data Visualization Desktop
- Gain insights from data in the ADWC
- Conclusion

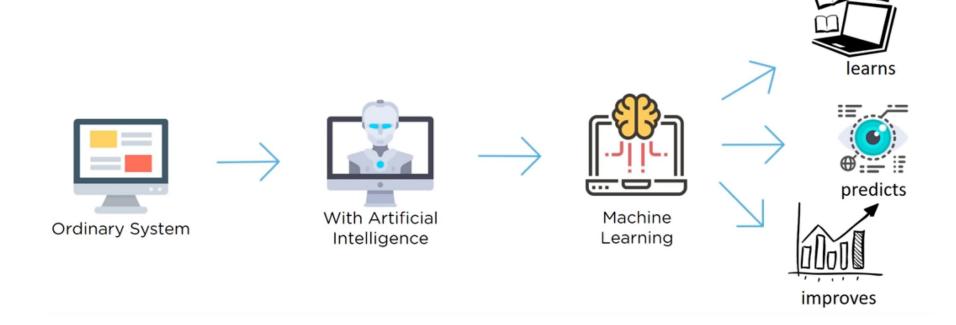
Data Science

First of all, let's review what machine learning is.

- Machine learning is an application of artificial intelligence (AI) that provides systems the ability to
 automatically learn and improve from experience without being explicitly programmed. Machine
 learning focuses on the development of computer programs that can access data and use it learn
 for themselves.
- The process of learning begins with observations or data, such as examples, direct experience, or instruction, in order to look for patterns in data and make better decisions in the future based on the examples that we provide. The primary aim is to allow the computers learn automatically without human intervention or assistance and adjust actions accordingly.
- Machine learning enables analysis of massive quantities of data. While it generally delivers faster,
 more accurate results in order to identify profitable opportunities or dangerous risks, it may also
 require additional time and resources to train it properly. Combining machine learning with AI and
 cognitive technologies can make it even more effective in processing large volumes of information.

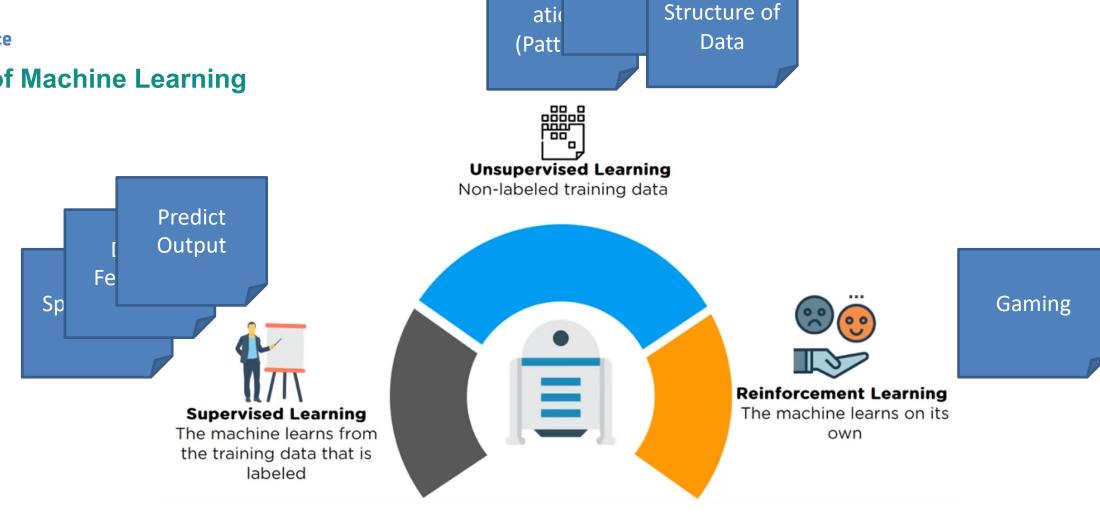
Data Science

First of all, let's review what machine learning is.



Data Science

Types of Machine Learning



Recon

No Fee

Find Hidden

Data Science

Algorithms (Methods to Solve Specific Problems)

- Supervised machine learning algorithms can apply what has been learned in the past to new data using labeled examples to predict future events. Starting from the analysis of a known training dataset, the learning algorithm produces an inferred function to make predictions about the output values. The system is able to provide targets for any new input after sufficient training. The learning algorithm can also compare its output with the correct, intended output and find errors in order to modify the model accordingly.
- In contrast, unsupervised machine learning algorithms are used when the information used to train is neither classified nor labeled. Unsupervised learning studies how systems can infer a function to describe a hidden structure from unlabeled data. The system doesn't figure out the right output, but it explores the data and can draw inferences from datasets to describe hidden structures from unlabeled data.
- Semi-supervised machine learning algorithms fall somewhere in between supervised and unsupervised learning, since they use both labeled and unlabeled data for training typically a small amount of labeled data and a large amount of unlabeled data. The systems that use this method are able to considerably improve learning accuracy. Usually, semi-supervised learning is chosen when the acquired labeled data requires skilled and relevant resources in order to train it / learn from it. Otherwise, acquiringunlabeled data generally doesn't require additional resources.
- Reinforcement machine learning algorithms is a learning method that interacts with its environment by producing actions and discovers errors or rewards. Trial and error search and delayed reward are the most relevant characteristics of reinforcement learning. This method allows machines and software agents to automatically determine the ideal behavior within a specific context in order to maximize its performance. Simple reward feedback is required for the agent to learn which action is best; this is known as the reinforcement signal.

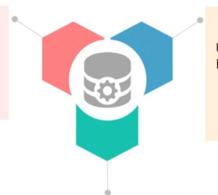
Data Science

Algorithms Methods (Methods to Solve Specific Problems)



Classification

Used when the output is categorical like 'YES' or 'NO'



Clustering

Used when the data needs to be organized to find patterns in the case of 'product recommendation'



Algorithms used

- · Decision Tree
- Naïve Bayes
- Random Forest
- · Logistic regression
- KNN



Regression

Used when a value needs to be predicted like the 'stock prices'

Algorithms used

· Linear Regression

Algorithms used

K Means

Data Science

Example:

The best way to explain what machine learning is would be to give you a simple example.

Let's say you want to develop a program that automatically detects what's in a picture, and we show a picture of a dog.

However, this kind of approach gets tricky pretty quickly. What if there's a white dog in the picture with no brown hair? What if the picture shows only the round parts of the table?

This is where machine learning comes in.

Machine learning typically implements an algorithm that automatically detects a pattern in the given input.

You can give, say, 1,000 pictures of a dog and 1,000 pictures of a table to a machine learning algorithm. Then, it will learn the difference between a dog and a table. When you give it a new picture of either a dog or a table, it will be able to recognize which one it is.

Data Science

Machine learning algorithms work much the same way.

You can apply the same idea to:

- Recommendation systems (think YouTube, Amazon, and Netflix)
- Face recognition
- Voice recognition
- Among other applications.

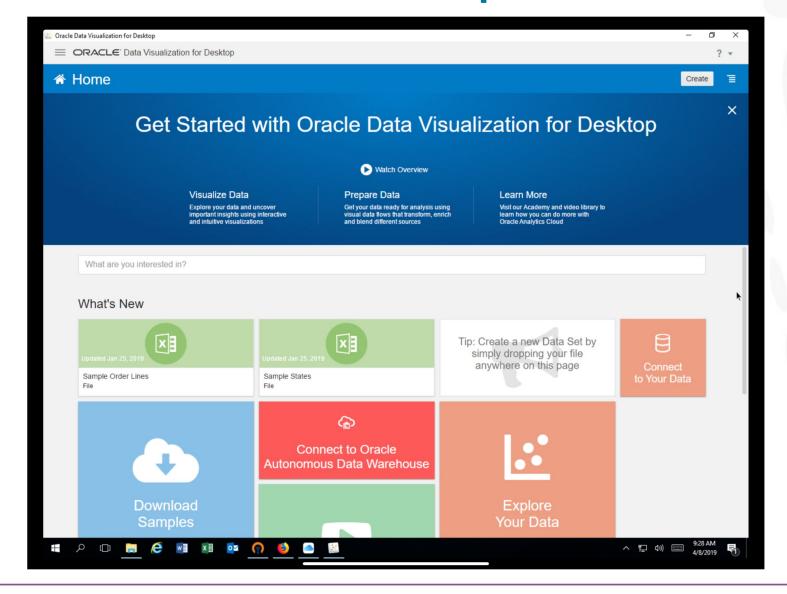
Popular machine learning algorithms you might have heard about include:

- Neural networks
- Deep learning
- Support vector machines
- Random forest

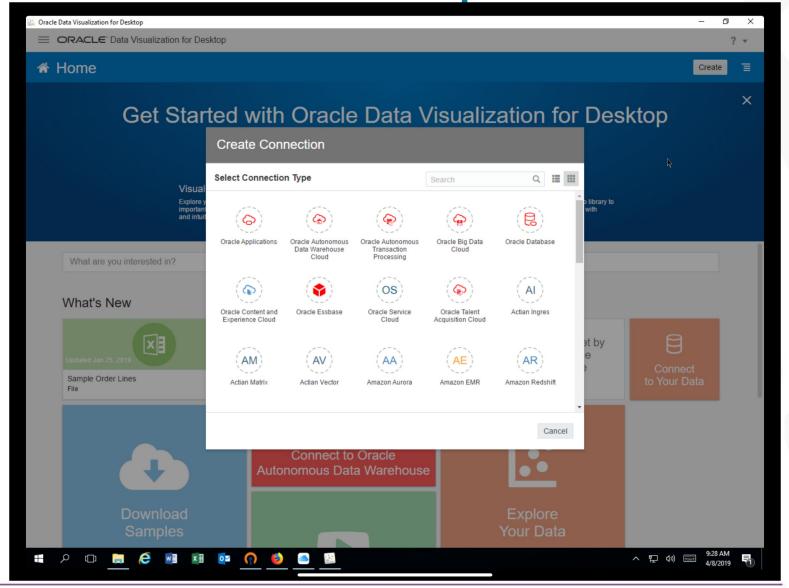
You can use any of the above algorithms to solve the picture-labeling problem I mentioned before.

It is all about Visualizing Data

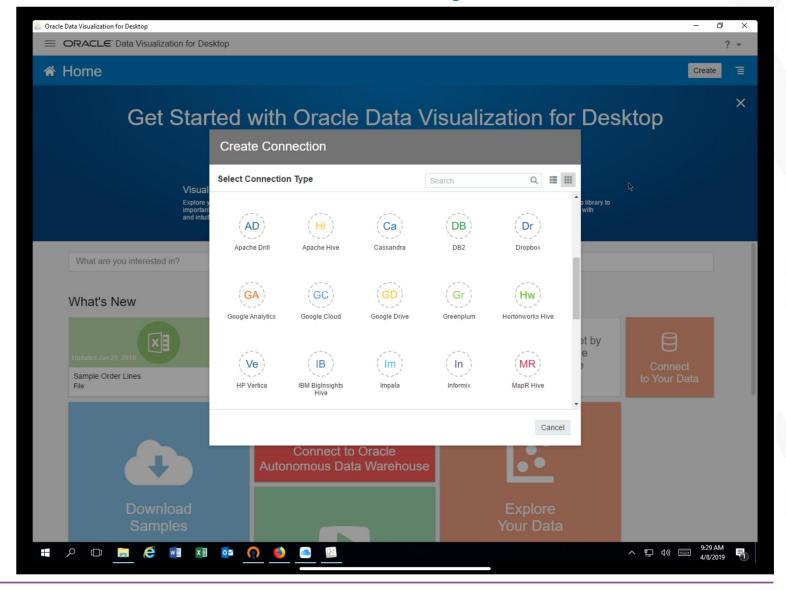
- Why Visualizing Data is so Important
- Introduction to Autonomous Data Warehouse Cloud Service
- Introduction to Oracle Data Visualization Desktop
- Gain insights from data in the ADWC
- Conclusion



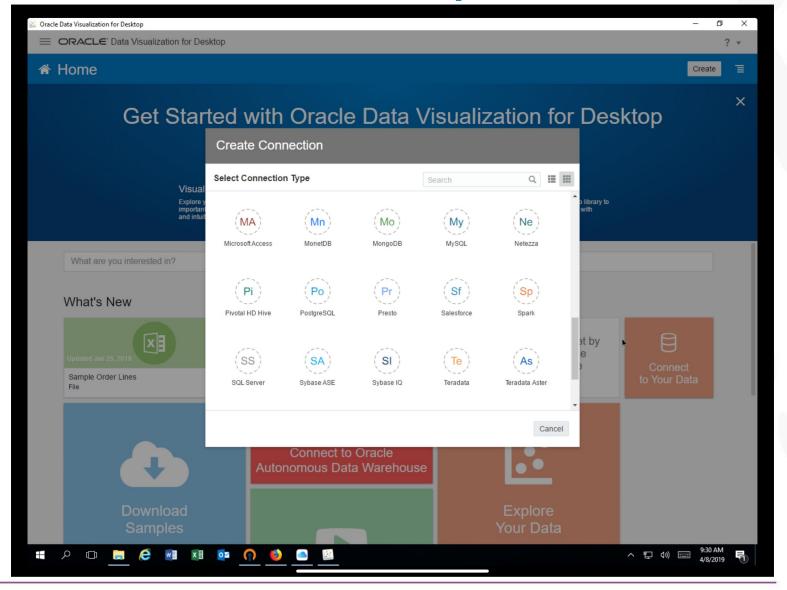






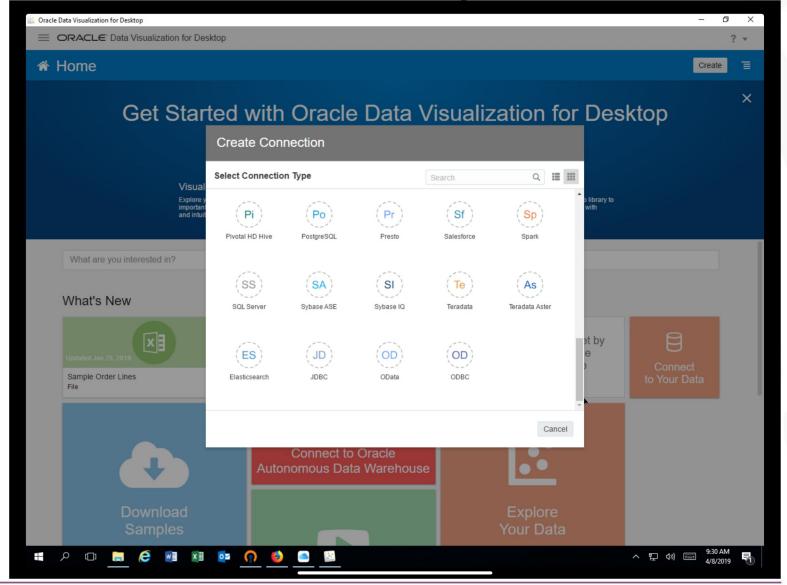






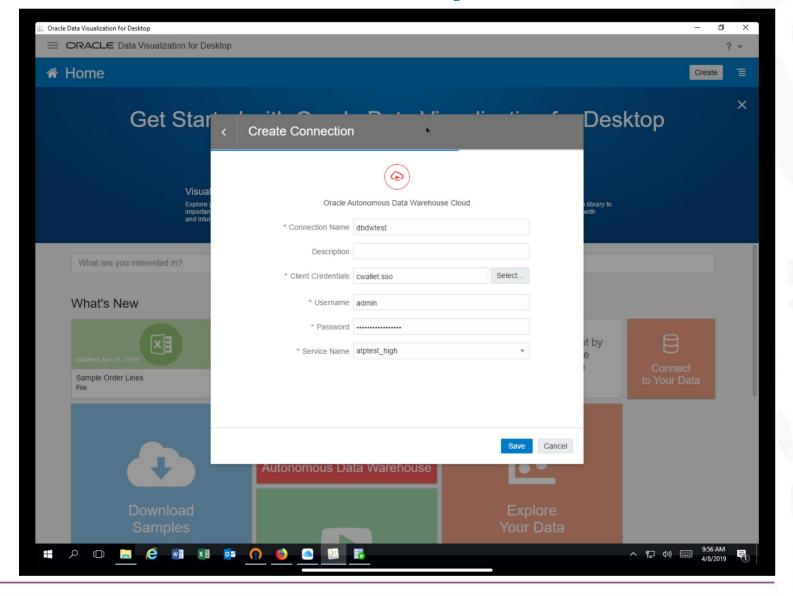


Oracle Data Visualization Desktop



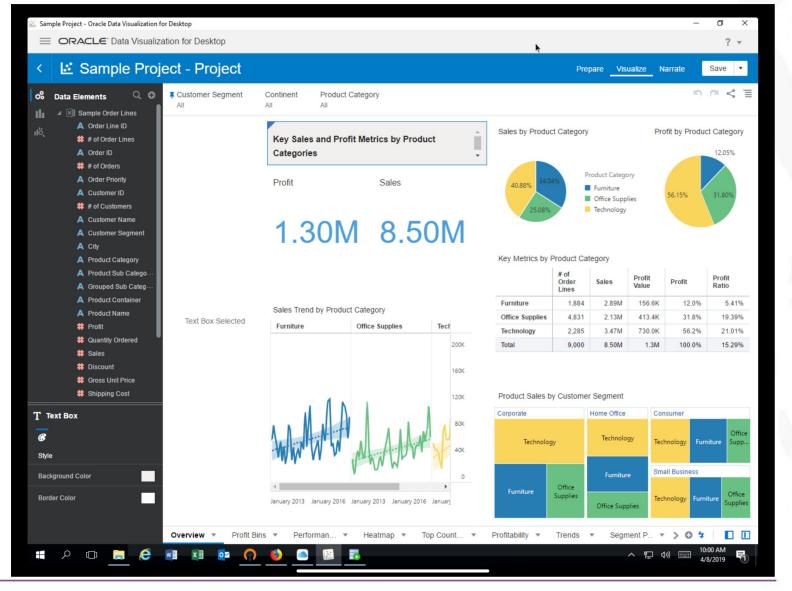


Oracle Data Visualization Desktop





Oracle Data Visualization Desktop





It is all about Visualizing Data

- Why Visualizing Data is so Important
- Introduction to Autonomous Data Warehouse Cloud Service
- Introduction to Oracle Data Visualization Desktop
- Gain insights from data in the ADWC
- Conclusion

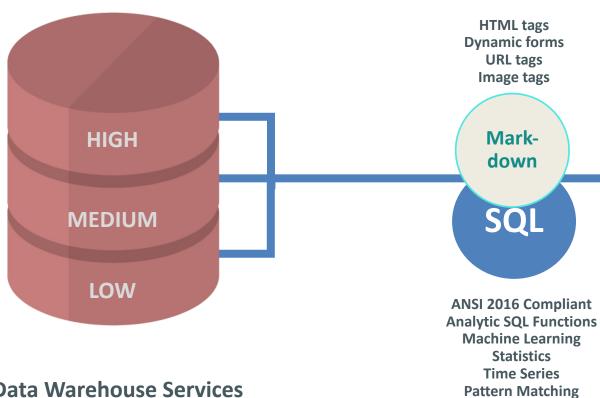
What is Oracle ML

- Built-in, web-based SQL Notebook
- Bundled with Autonomous Data Warehouse
- Derived from open-source Apache Zeppelin
 - Extensive additions that are being fed back into the Apache Zeppelin project
- Provides web-based SQL access to ADW
- Simple but powerful set of data visualizations
- Includes sharing and collaboration framework
- All workspaces, projects, notebooks etc saved inside database
 - Automatically (i.e. autonomously) managed and backed up by database



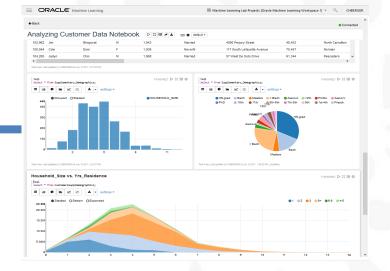
Autonomous Data Warehouse Cloud: Architecture Overview

Analytic Views



Data Warehouse Services

(EDWs, DW, departmental marts and sandboxes)



Oracle Machine Learning SQL Notebook



Smart Way to Use Oracle ML

- 1. Write **SQL** using 18c in-database **analytics**
- 2. Document and annotate using the HTML, URL, image markdown tags
- 3. Build **interactive** visualizations using markdown tags
- 4. Collaborate and share
 - Version notebooks for team work
 - Share entire notebook with report consumers
 - Export specific paragraphs or visualizations



Key Concepts

- Workspace
 - Folder for organizing projects
 - Controls user access to projects
- **Project**
 - Folder for organizing and storing SQL notebooks, SQL scripts and jobs

Notebook

- Contains one or more paragraphs
- Includes data visualizations (tables and graphs)

Run SQL

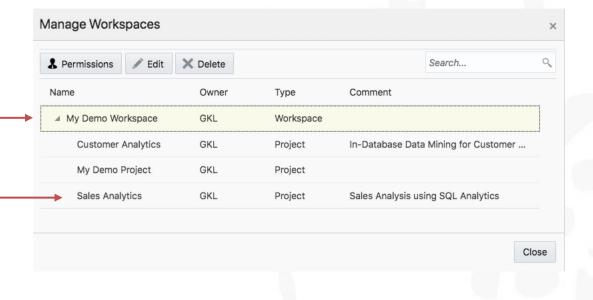
Statement

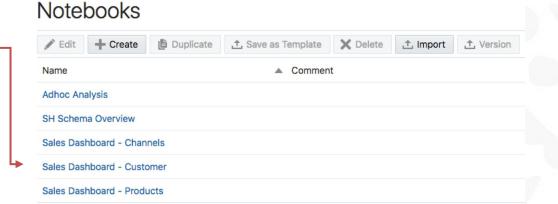
Enter and run SQL statements

- Two built-in shortcuts
 - SQL Query Scratchpad
 - SQL Script Scratchp











Key Concepts for Notebooks

Interpreters

- Bindings that connect a notebook to ADW resource group (high, medium, low)
- Markdown is automatically enabled

Paragraph

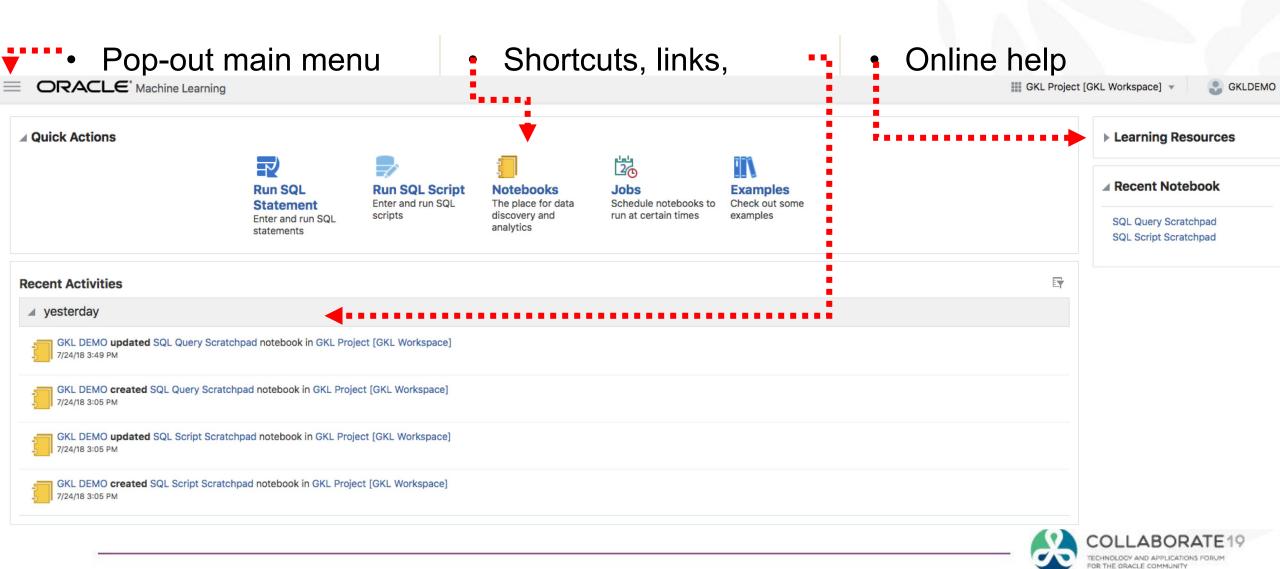
- Contains a single SQL statement (default) or
 - SQL Query offers multiple data visualizations
- Contains a SQL script (%script)
- Contains markdown tags + text (%md)







Oracle ML Home Page



SQL Query Scratchpad DX 10 1 1 1





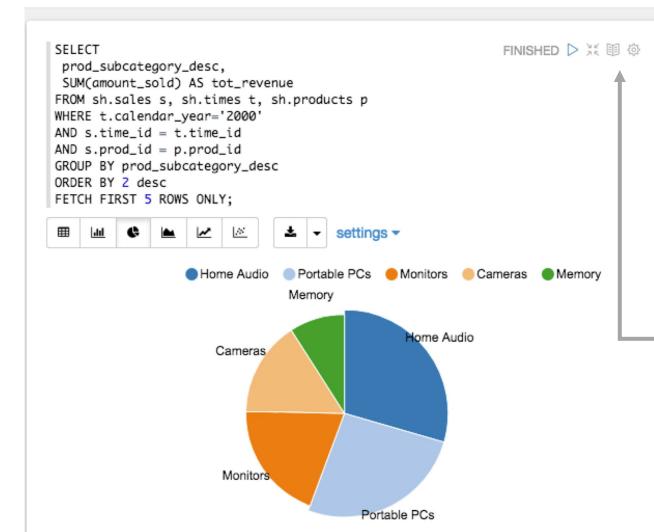
Editor Area (Show/Hide)

SQL Query Scratchpad









Notebook menu area

- Run/refresh all paragraphs
- Show/hide code
- Show/output
- Clear output
- Clear notebook
- Export notebook
- List of Shortcuts
- Interpreter bindings
- Report type

Paragraph menu area

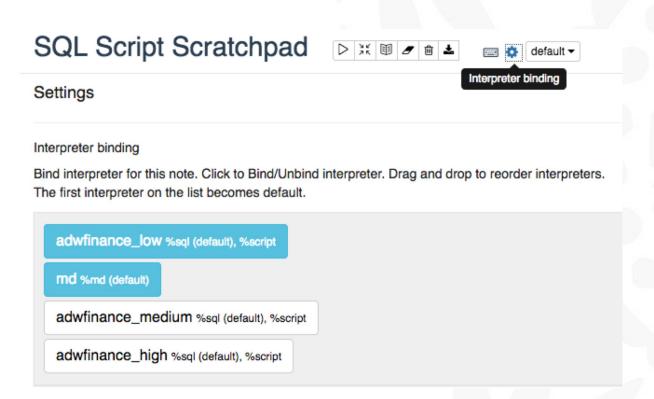
- Run/refresh paragraphs
- Show/hide code
- Show/output
- Settings

Took 1 sec. Last updated by GKLDEMO at July 24 2018, 3:16:30 PM. (outdated)



ADW Resource Interpreters

- Three interpreters linked to three ADW resource groups:
 - High
 - Medium
 - Low
- Normally only need to select two interpreters:
 - md to access markdown tags
 - ADW resource group interpreter





Interpreters Linked to Database Resource Groups

- 3 pre-defined database services
 - Choice of performance and concurrency
- HIGH
 - Highest resources, lowest concurrency
 - Queries run in parallel
- MEDIUM
 - Less resources, higher concurrency
 - Queries run in parallel
- LOW
 - Least resources, highest concurrency
 - Queries run serially

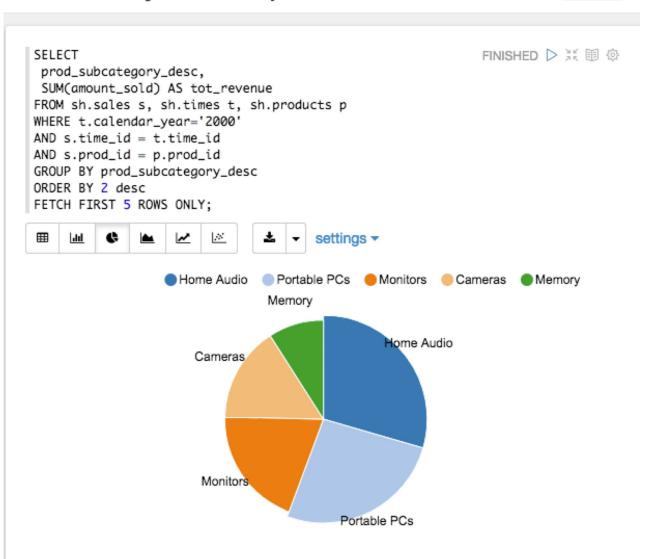
Example for a database with 16 OCPUs

	No of concurrent queries	Max idle time	CPU shares
HIGH	3	5 mins	4
MEDIUM	20	5 mins	2
LOW	32	1 hour	1



SQL Query Scratchpad





SQL Editor Area (Show/Hide)

Output Area (Show/Hide)

Paragraph

Took 1 sec. Last updated by GKLDEMO at July 24 2018, 3:16:30 PM. (outdated)



OML Markdown Key Concepts

- Set of predefined markdown tags
 - Includes basic HTML tags, e.g. <H1>, <H2>…
- Markdown tags can be used in:
 - Column selection
 - Setting aggregation methods
 - WHERE, GROUP BY, ORDER BY clause
 - FROM clause
- Note: Markdown is always parsed before SQL is executed
 - Can't create dynamic input to populate structure of markdown tag



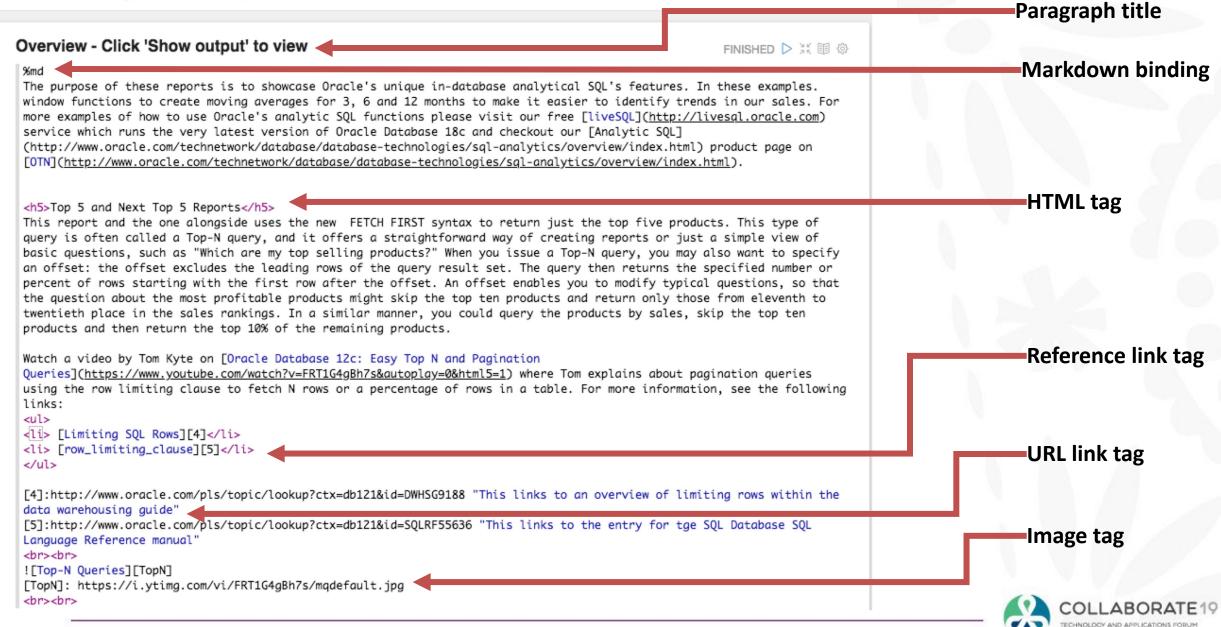
Simple Documentation Markdown Tags

- Useful tags for dynamic content generation
 - https://sourceforge.net/p/zeppelin/wiki/markdown_syntax/
 - Links
 - [click here for more information](http://your-specific-url)
 - Reference Links
 - Use references to [link first URL][1] and like to another URL like [this][2] with hover-text
 - [1]: http://url-link-1
 - [2]: http://url-link-2 "My hover-over text message"
 - Images
 - ![alternate-text-for-image](ImageURL)
 - Basic formatting tags includes HTML heading tags and font-format tags:
 - *use this for italic*
 - **use this for bold**
 - ***use this for bold and italic***



SQL Query Scratchpad





SQL Query Scratchpad DX 10 - 1 - 1



Overview - Click 'Show output' to view

FINISHED ▷ 🂢 🗐 🕸

The purpose of these reports is to showcase Oracle's unique in-database analytical SQL's features. In these examples, window functions to create moving averages for 3, 6 and 12 months to make it easier to identify trends in our sales. For more examples of how to use Oracle's analytic SQL functions please visit our free liveSQL service which runs the very latest version of Oracle Database 18c and checkout our Analytic SQL product page on OTN.

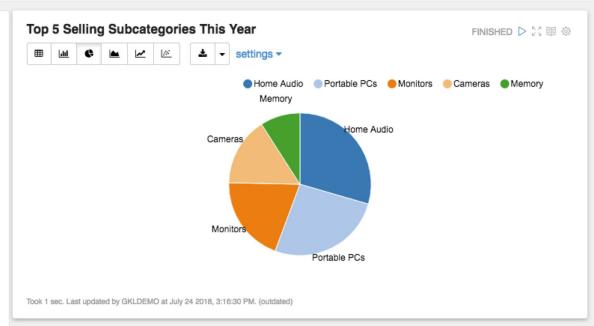
Top 5 and Next Top 5 Reports

This report and the one alongside uses the new FETCH FIRST syntax to return just the top five products. This type of query is often called a Top-N query, and it offers a straightforward way of creating reports or just a simple view of basic questions, such as "Which are my top selling products?" When you issue a Top-N query, you may also want to specify an offset: the offset excludes the leading rows of the query result set. The query then returns the specified number or percent of rows starting with the first row after the offset. An offset enables you to modify typical questions, so that the question about the most profitable products might skip the top ten products and return only those from eleventh to twentieth place in the sales rankings. In a similar manner, you could query the products by sales, skip the top ten products and then return the top 10% of the remaining products.

Watch a video by Tom Kyte on Oracle Database 12c: Easy Top N and Pagination Queries where Tom explains about pagination queries using the row limiting clause to fetch N rows or a percentage of rows in a table. For more information, see the following links:

- · Limiting SQL Rows
- row_limiting_clause





Took 0 sec. Last updated by GKLDEMO at July 24 2018, 3:28:01 PM.



OML Markdown Additional Key Concepts

- Set of predefined markdown tags
 - Includes basic HTML tags, e.g. <H1>, <H2>...
- Markdown tags can be used in:
 - Column selection
 - Setting aggregation methods
 - WHERE, GROUP BY, ORDER BY clause
 - FROM clause
- Note: Markdown is always parsed before SQL is executed
 - Can't create dynamic input to populate structure of markdown tag

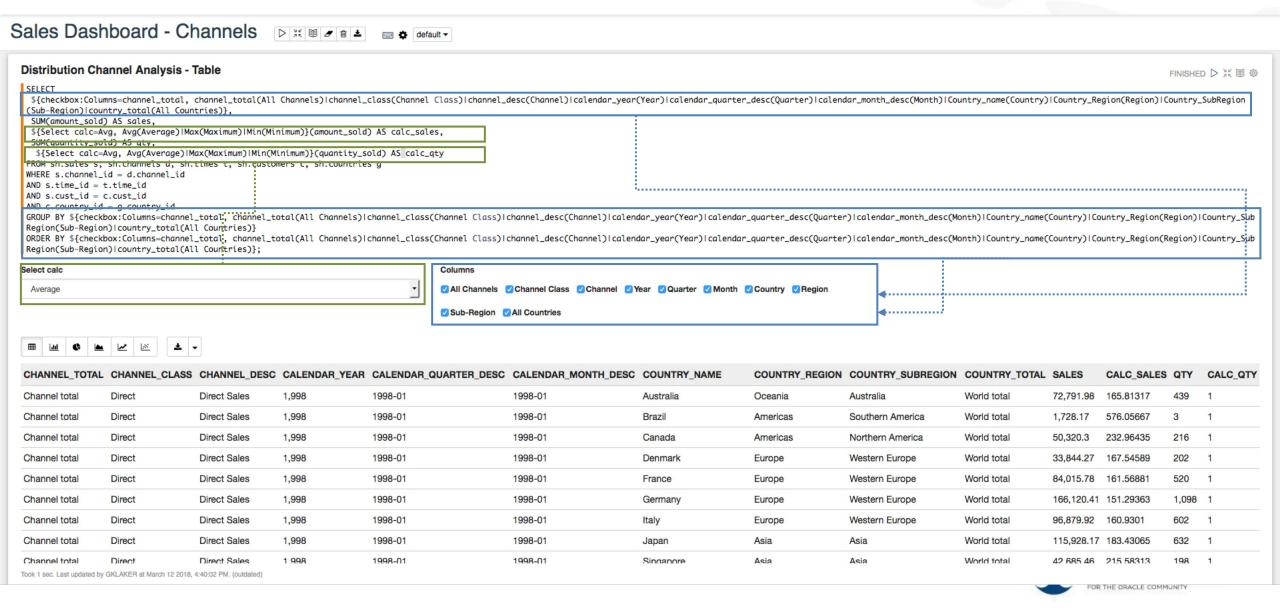


More Sophisticated Markdown Tags

- Useful tags for dynamic content generation
 - https://zeppelin.apache.org/docs/latest/manual/dynamicform.html#select-form
 - Simple input box
 - \${formName} templates.
 - Select form/pulldown selection
 - \${formName=defaultValue,option1(DisplayName)|option2(DisplayName)...}
 - Checkbox
 - **\${checkbox**:formName=defaultValue1|defaultValue2...,option1|option2...}
 - URL Links
 - [text-for-link] < http://url-to-content>



Building More Sophisticated Interactive Reports



Documenting Interactive Reports

Sales Dashboard - Channels



Briefing Notes - Click 'Show output' to view

FINISHED ▷ ※ III @

Wmd

The purpose of these reports is to showcase Oracle's unique in-database analytical SQL's features. In these examples we are using a range of different analytical techniques and functions...

<h5>52-Week Revenue Trend Analysis</h5>

The initial data set from the join of the sales fact table to the channel dimension and times dimension tables results in a sparse data set for 1998 in the channel class Direct. To create a fully dense report where all weeks are shown, even where there is no data we use a partitioned outer join with a dense set of time data. In the this query, you alias the original query as v and you select data from the times table, which you alias as t. You will find additional rows are added to the output for the missing time values (i.e. where no sales have been made) and these have 0 as their Sales value set to 0 by using the NVL function.

<h5>Ouarterly Analysis of Sales by Channel</h5>

Aggregation is a fundamental part of data warehousing. In this report we are using the ROLLUP extension to the GROUP BY clause along with one of the GROUPING functions.

ROLLUP enables a SELECT statement to calculate multiple levels of subtotals across a specified group of dimensions. It also calculates a grand total. ROLLUP is a simple extension to the GROUP BY clause, so its syntax is extremely easy to use. The ROLLUP extension is highly efficient, adding minimal overhead to a query. The action of ROLLUP is straightforward: it creates subtotals that roll up from the most detailed level to a grand total, following a grouping list specified in the ROLLUP clause.

The GROUPING_ID function is used to determine to programmatically determine which result set rows are subtotals, and the exact level of aggregation for a given subtotal. This information is then used to substitute new descriptions for each dimension member, i.e. "All Quarters" and "All Channels".

For more information about the ROLLUP function [go here][6] and for more information Grouping Functions [go here][7].

<h5>Distribution Channel Analysis - Table<h5>

This report is designed to highlight how to use some of the Oracle ML markdown constructs such as CHECKBOX and SELECT. There is more information about the markdown tags:

- [Apache Zeppelin Documentation][8]
- [Zeppelin Wiki Page][9]

[8]:https://zeppelin.apache.org/docs/0.7.0/manual/dynamicform.html "This is the Zeppelin doumentation page for dynamic form markdown taas"

<h5>Using the Report Menu</h5>

All the graphs are interactive so you can drill down by clicking on a product label such as Electronics or Hardware. The trend analysis graph allows you to select the time period and the product category as filters for the graph data set. If you want to reset a grand simply press the triangle icon in the top-right menu ![ML Menu Icons][MenuIcons]

[MenuIcons]:https://lh3.googleusercontent.com/--tHD5Mo8n8M/WRXRE7wC5GI/AAAAAAAADRs/VEvZDVwNVeo41QgveLQ1yJmtxseJeTo8wCHM/fullsizeoutput_1079.jpeg



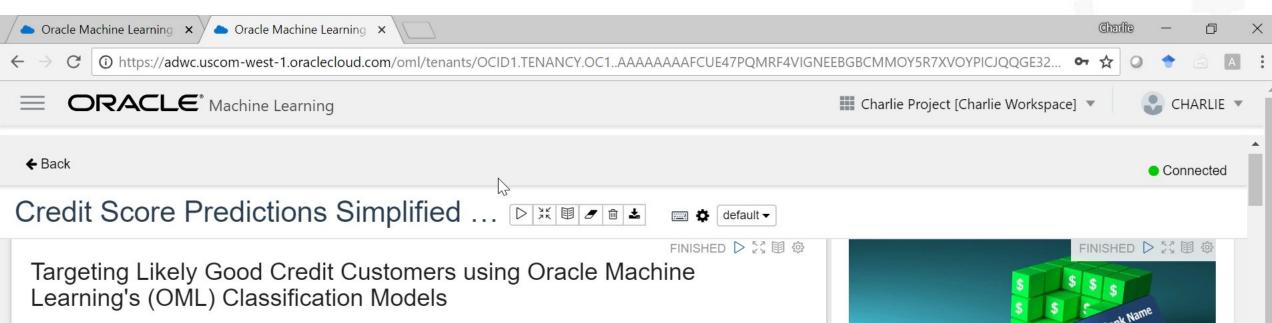
Download Sample Notebooks

- Sales Analysis series of notebooks
 - Based on sales history schema (SH)
 - Includes 5 notebooks:
 - Overview, Channel Analysis Dashboard, Customer Analysis Dashboard, Product Analysis Dashboard, adhoc analysis
 - Download links https://www.dropbox.com/s/2if07dja88y8s1u/Sales-Analysis.zip?dl=0
- Customer Insight Analytics series of notebooks
 - Machine learning using sales history schema (SH)
 - Includes 8 notebooks covering machine learning examples
 - Download links https://www.dropbox.com/s/47oj06yw4g0e654/Deep-Customer-Analytics.zip?dl=0



It is all about Visualizing Data

- Why Visualizing Data is so Important
- Introduction to Autonomous Data Warehouse Cloud Service
- Introduction to Oracle Data Visualization Desktop
- Gain insights from data in the ADWC
- Conclusion



Heather has spent most of her time over the past couple of years extracting and preparing data for analysis. The large volumes of data need extracting and processing mean she spends most of her time waiting for jobs to finish and very little of her time analyzing the data. Demands from marketing are forcing a new approach whereby the data remains in the data warehouse and is processed there. The alternative cloud solution is more complex, and has no direct out of the box processes to analyze the data in place. She started taking a look at Oracle, and found the simple SQL commands in ADWC are familiar, and execute extremely fast, leveraging all the performance features of the platform. Further once she is done can can apply the learning models to incoming data on the fly, and allow end user analysts to immediately see mining results. This drastically reduces the cycle of data preparation, analysis, and publishing. It also means there is no change to analysis/reporting Data Visualization toolset that users are familiar with.

Scroll down this notebook and learn how to use OML to create predictive perspectives on data in ADWC, WITHOUT moving it. We will process a small 100k data set, but could use a 100M or billion row data set without worrying about processing time.

This is an extract of Alphaoffice customer information. We will first get acclimated to Apache Zepplin, the open source interface for interactive collaboration in a team environment.

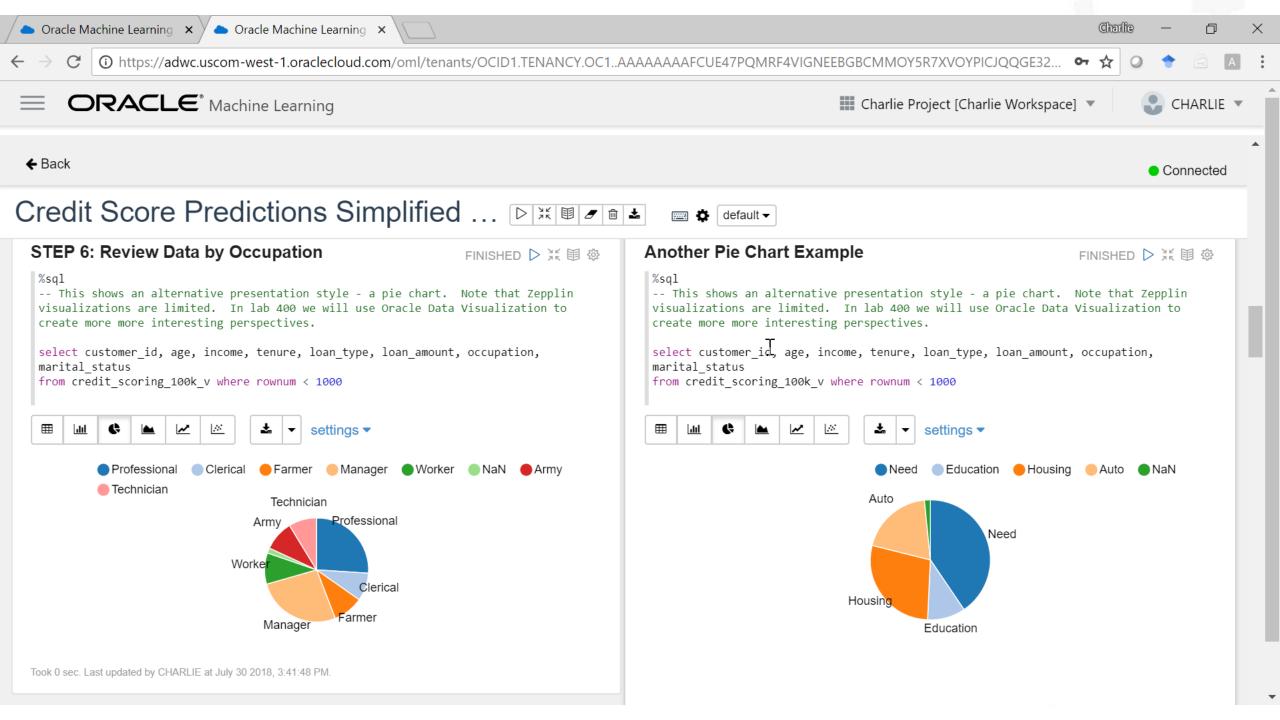
The Business Problem:

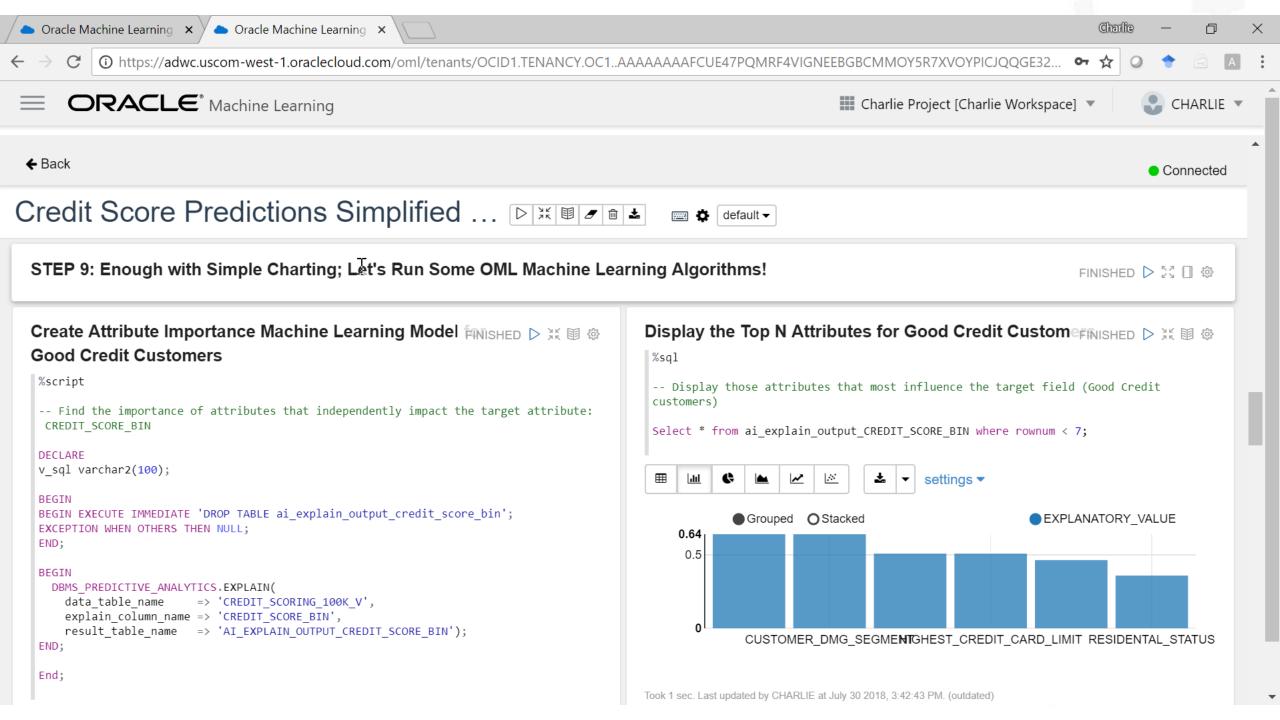
Increase Sales by Targeting our Best Customers; Good Credit Customers!

Heather has a hunch that weakening sales may be due to the company selling to non-optimal customers; customers who



Took 0 sec. Last updated by CHARLIE at July 16 2018, 12:58:40 PM. (outdated)







FINISHED ▷ 🂢 国 墩

Credit Score Predictions Simplified ... ▷※ ■ 🗗 📼 🕏 default 🕶

STEP 11: Create Predictive Model to Target Good Credit Customers

Now that Heather has found the key attributes that most influence findig more Good Credit customers and also making better Maximum Credit Card Amount decisions, she wants to leverage Oracle Machine Learning's powerful in-Database, parallelized algorithms to build predictive models that help her company to better target "the right customers" with the "right offers".

Lets' quickly review the machine learning process:

Problem Definition: Target Good Credit Customers

Data Gathering and Preparation: We've assembled 100K records with 100+ variable about each customer and have created a target field (Good Customer/Other Customer) so we can use OML's Supervised Algorithms, specifically let's start by using a decision tree algorithm.

Model Building and Evaluation: We'll create a randomly selected sample from our Credit_Scoring_100k historical data and use 60% as training data for the machine learning model building phase. Then, we'll use the remaining 40% as a holdout sampple to test our model's accuracy using various model evaluation tools such as a "lift chart".

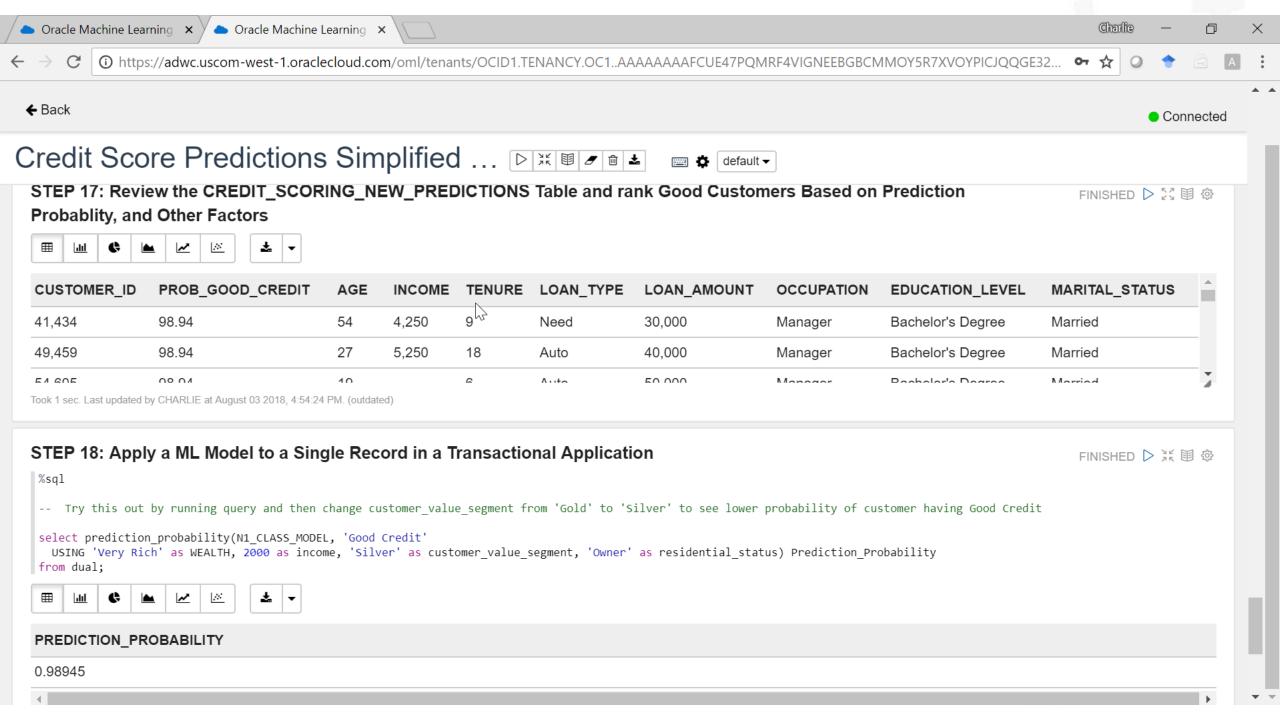
Knowledge Deployment: Once we're satisfied that we have a useful ML model that can predict with some accuracy which customers we should target (Good Credit customers), we want to apply our OML model to new customer data inside ADWC and then take a deeper look a them. Lastly, we'll jump over to Oracle Analytics Cloud for a more interactive, exploratory data analysis experience but now focusing on our customers of interest (Good Credit customers).

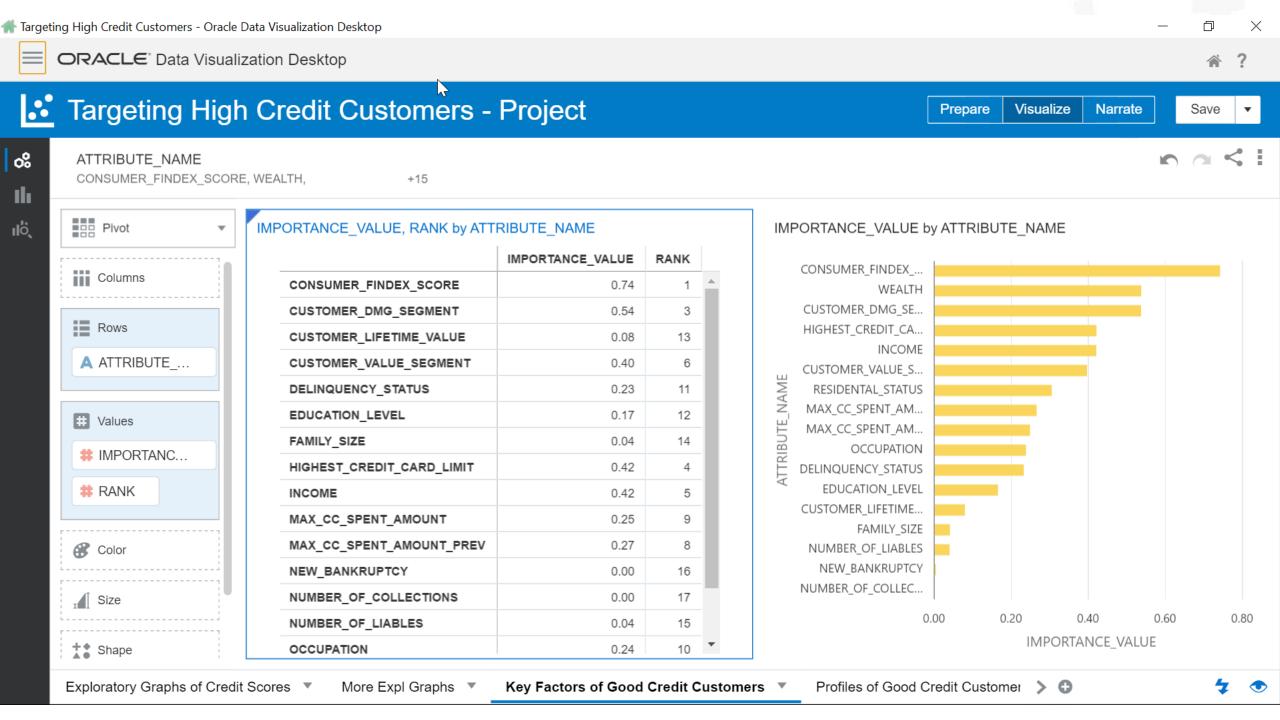
Data Mining and Machine Leanished ▷ 🂢 🗏 🕸 Process Problem Definition Data Gathering and Preparation Data access Data sampling Model Building Data Transformation and Evaluation Create model Knowledge Deployment Test model Evaluate and interpret model Model apply Custom reports External applications

Connected

Took 0 sec. Last updated by CHARLIE at July 16 2018, 5:31:12 PM. (outdated)

Took 0 sec. Last updated by CHARLIE at July 19 2018, 3:06:35 PM. (outdated)





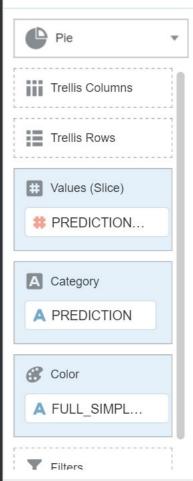
Targeting High Credit Customers - Project o

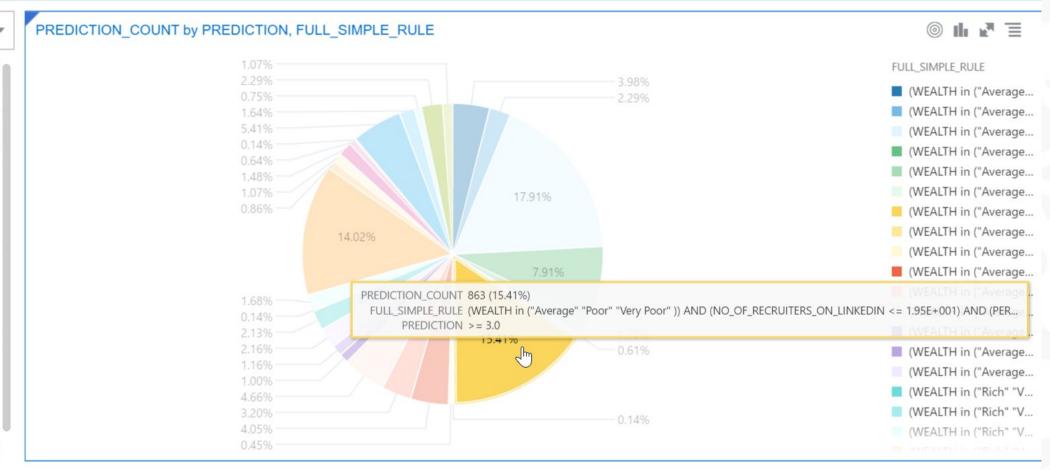
th

ıβ

Click here or drag data to add a filter













Where to get more information

- Product information: <u>cloud.oracle.com/datawarehouse</u>
- Documentation
 - ADW: https://docs.oracle.com/en/cloud/paas/autonomous-data-warehouse-cloud/index.html
 - Oracle ML: https://docs.oracle.com/en/cloud/paas/autonomous-data-warehouse-cloud/omlug/getting-started-oracle-machine-learning1.html
- Hands-on Workshop
 - https://oracle.github.io/learning-library/workshops/journey4-adwc



Where to get more information

- New Q&A Forum on Cloud Customer Connect
 - https://cloudcustomerconnect.oracle.com/resources/32a53f8587/summary
- Forbes: Autonomous Capabilities Will Make DBAs More Valuable
 - https://www.forbes.com/sites/oracle/2018/03/21/autonomous-capabilities-will-make-datawarehouses-and-dbas-more-valuable/#73b134c6624e
- Oracle Analytics Library (Example and Tutorials)
 - https://www.oracle.com/solutions/business-analytics/data-visualization/library-overview.html
- Oracle Data Visualization Desktop (Download)
 - https://www.oracle.com/middleware/technologies/oracle-data-visualization-desktop.html





fmunozalvarez@dataintensity.com

Session ID:

10730

Remember to complete your evaluation for this session within the app!

