

Oracleで実現する クラウド・マシン・ラーニング

Oracle Advanced Analytics
in Oracle Database Cloud Service



Oracle DBA & Developer Day 2015

2015年12月9日(水) グランドプリンスホテル新高輪 国際館パミール

日本オラクル株式会社

クラウド・テクノロジー事業統括
データベースソリューション本部

小川 幹雄

ORACLE®

- 以下の事項は、弊社の一般的な製品の方向性に関する概要を説明するものです。また、情報提供を唯一の目的とするものであり、いかなる契約にも組み込むことはできません。以下の事項は、マテリアルやコード、機能を提供することをコミットメント（確約）するものではないため、購買決定を行う際の判断材料になさらないで下さい。オラクル製品に関して記載されている機能の開発、リリースおよび時期については、弊社の裁量により決定されます。

OracleとJavaは、Oracle Corporation 及びその子会社、関連会社の米国及びその他の国における登録商標です。文中の社名、商品名等は各社の商標または登録商標である場合があります。

アジェンダ

- 1 ▶ マシン・ラーニング（機械学習）概要
- 2 ▶ Oracle Advanced Analytics の紹介
- 3 ▶ クラウド・マシン・ラーニング
- 4 ▶ Oracle Data Miner で設計する予測分析
- 5 ▶ まとめ

アジェンダ

- 1 マシン・ラーニング（機械学習）概要
- 2 Oracle Advanced Analytics の紹介
- 3 クラウド・マシン・ラーニング
- 4 Oracle Data Miner で設計する予測分析
- 5 まとめ

マシン・ラーニング（機械学習）



明示的にプログラムしなくても学習する能力をコンピュータに与える

- 教師あり学習
- 教師なし学習

• 連想される言葉

– 人工知能、画像認識、自然言語処理、自動運転、
ディープラーニング、データマイニング、R、Python

顧客情報（属性値や、過去の利用状況など）から、その顧客が今後優良顧客になる可能性を算出したい



顧客データ



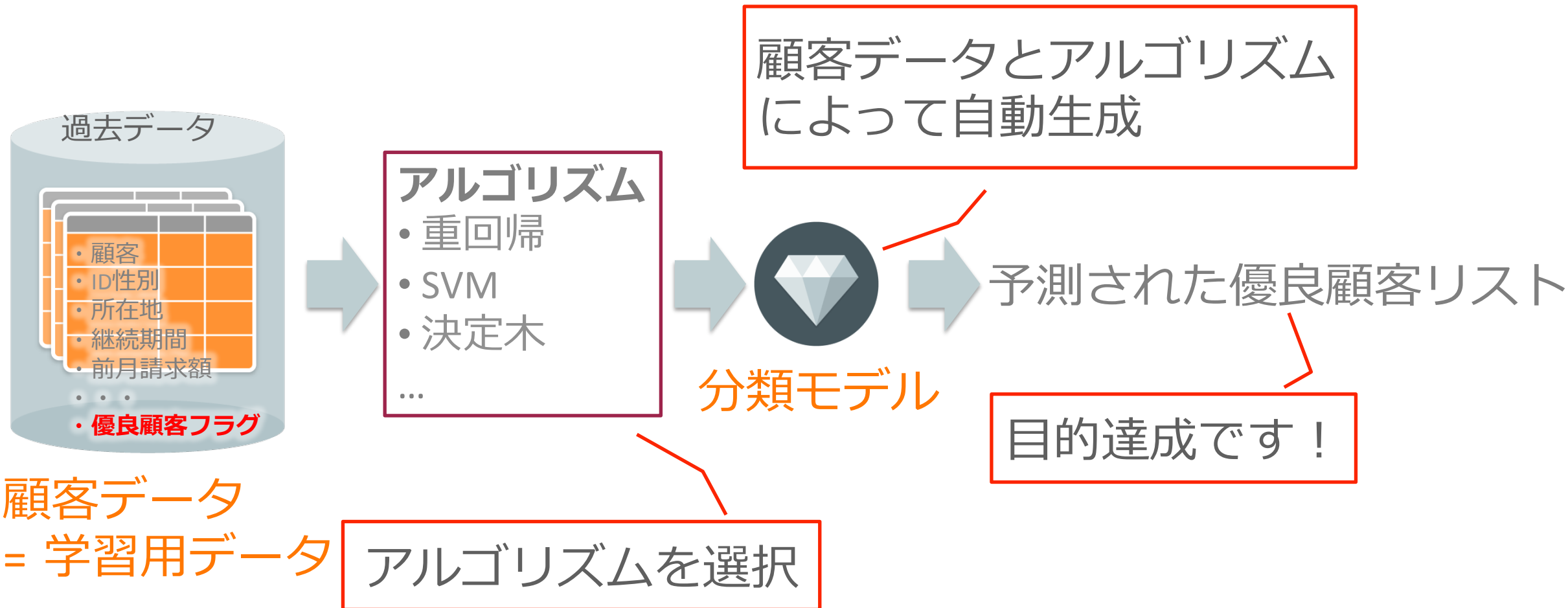
分類モデル



この顧客は優良顧客か？
その確率は？

このモデルをどうやって
作成するのか？

機械学習によるモデル作成



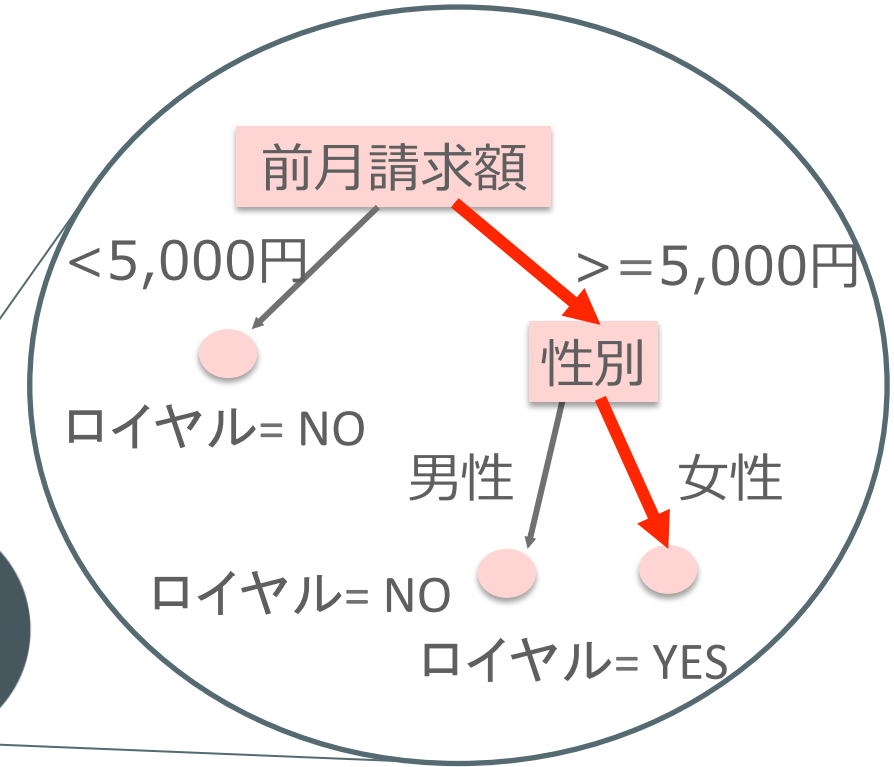
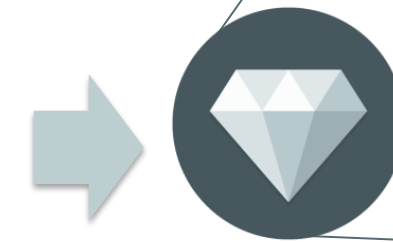
自動生成されるモデルのイメージ

過去顧客データ

| ID | 性別 | 年齢 | 職業 | 前月請求額 | ロイヤル |
|-----|----|----|-----|---------|------|
| 101 | 男性 | 31 | 会社員 | 10,000円 | NO |
| 102 | 女性 | 28 | 主婦 | 5,000円 | YES |
| 103 | 女性 | 36 | 主婦 | 200円 | NO |
| 104 | 男性 | 43 | 会社員 | 3,000円 | NO |
| 105 | 女性 | 22 | 会社員 | 7,000円 | YES |

新規顧客データ

| ID | 性別 | 年齢 | 職業 | 前月請求額 | ロイヤル |
|-----|----|----|-----|--------|------|
| 314 | 男性 | 40 | 会社員 | 4,000円 | NO |
| 315 | 女性 | 26 | 会社員 | 5,500円 | YES |



IF (前月請求額 >= 5000
AND 性別 = "女性")

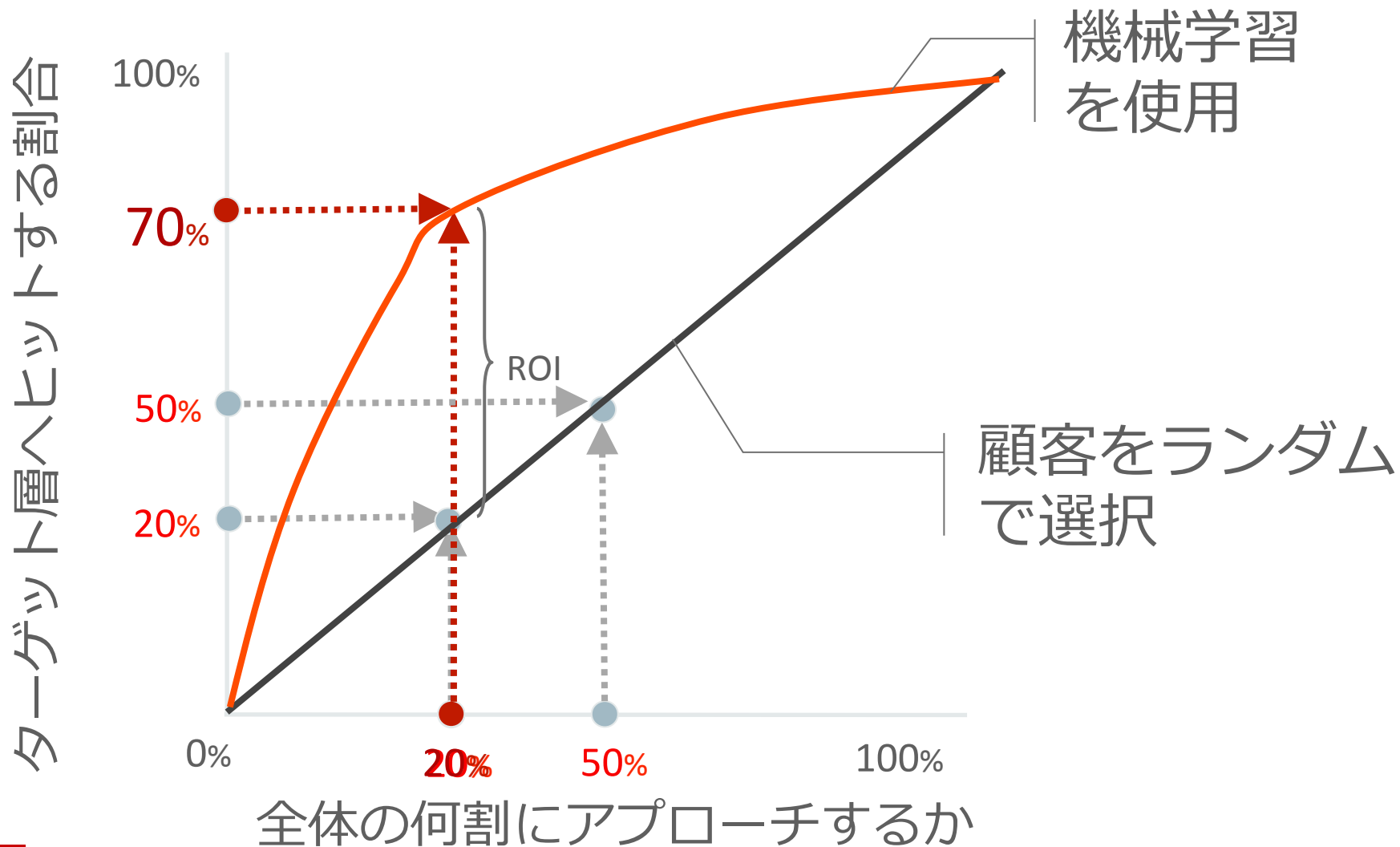
THEN

ロイヤル= YES

Probability = 0.77 (信頼度)

Support = 0.250 (組み合わせの出現率)

機械学習のビジネスへの有効性

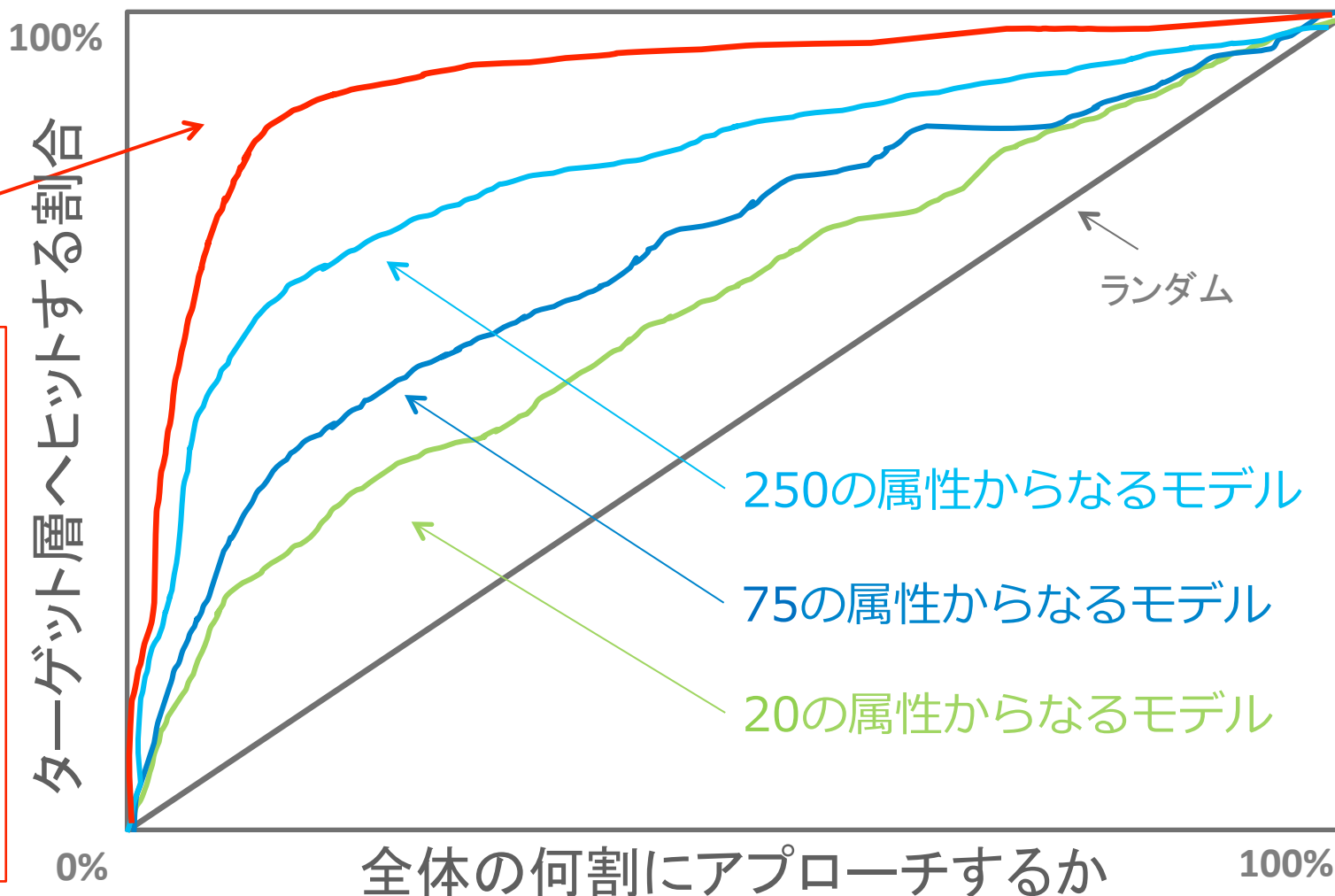


データ量を増やすことによって、予測精度を上げる



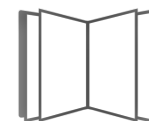
ビッグデータ（例えば数百の属性からなる入力）で生成されたモデル

- 統計データ
 - POSトランザクションデータ
 - テキストやコメント
 - 位置情報
 - 過去データと直近の行動データ
 - ウェブログデータ
 - センサーデータ
- etc.

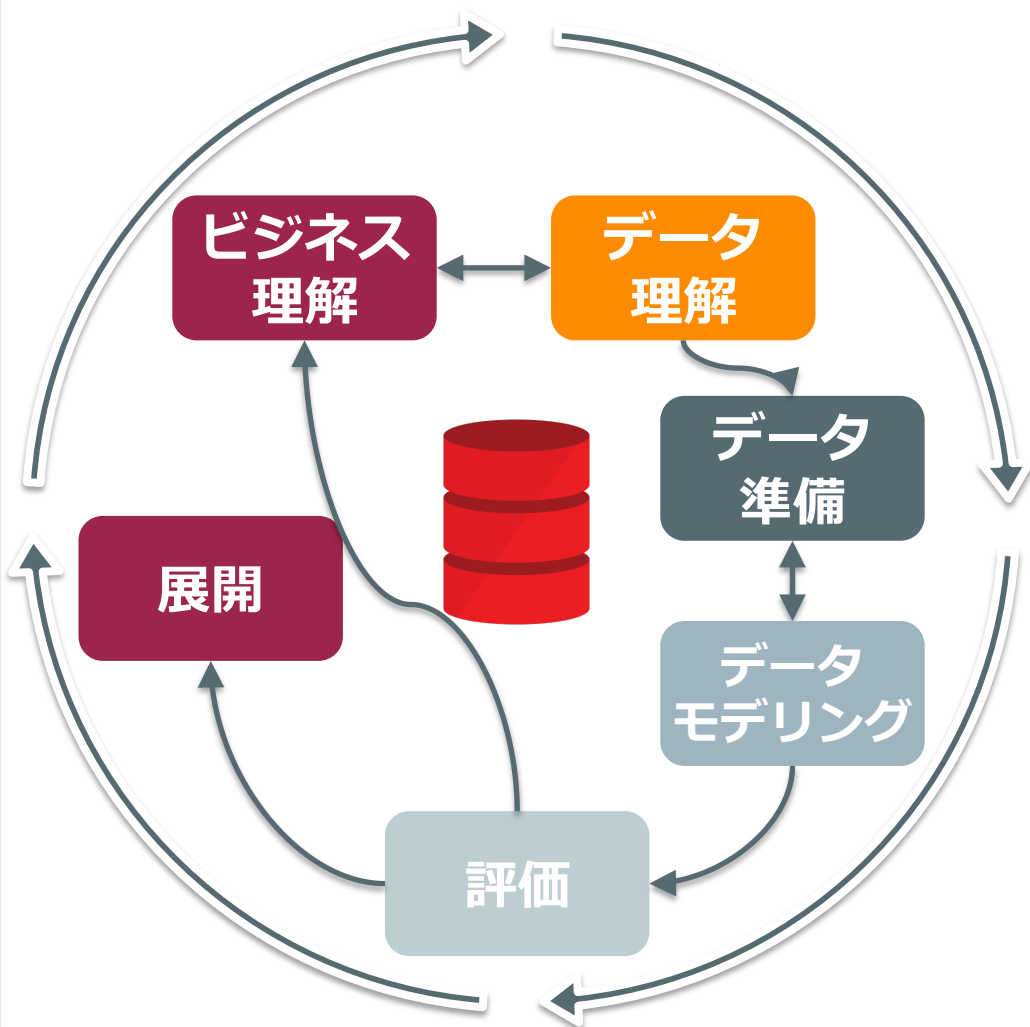


機械学習の様々な活用例

- ダイレクトメール送付 ⇒ 送付先アドレスリスト
- チャーン分析 ⇒ チャーン顧客のリスト
- クレジットリスク分析 ⇒ 個々人のクレジットスコア
- マーケット・バスケット分析 ⇒ アップセル・クロスセルのリコメンデーション
- 異常検出(Fraud 分析) ⇒ 異常トランザクションの検出
- 医学実験結果分析 ⇒ 実験データ要因抽出
- 国家安全保障 ⇒ テロリスト容疑者の検出



データ活用におけるPDCAサイクル (CRISP-DM)



| ビジネス理解 | データ理解 | データ準備 | データモデリング | 評価 | 展開 |
|--------------|-------------|----------|--------------|-------------|-------------|
| ビジネス目的の決定する | 初期データを収集する | データを選択する | モデリング手法を選択する | 結果を評価する | 展開を計画する |
| 状況を評価する | データを記述する | データを整理する | テスト結果を生成する | プロセスを審査する | 運用を計画する |
| 分析のゴールを決定する | データを調べる | データを構築する | モデルを構築する | 次のステップを決定する | 最終レポートを作成する |
| プロジェクト計画を建てる | データの品質を検証する | データを統合する | モデルを評価する | | プロジェクトを審査する |
| | | データを設定する | | | |

アジェンダ

- 1 マシン・ラーニング（機械学習）概要
- 2 Oracle Advanced Analytics の紹介
- 3 クラウド・マシン・ラーニング
- 4 Oracle Data Miner で設計する予測分析
- 5 まとめ

Oracleのマシン・ラーニングの進化

アナリティクスSQLをデータベースに



ORACLE® **12^c**
DATABASE



ORACLE® **11^g**
DATABASE



• シンキングマシンス社より“Darwin”データ・マイニングソフトウェアと開発チームを獲得

• 二つのアルゴリズム (NB, AR) をJava APIとして提供

• データ・マイニング SQL関数, 7つの新しいアルゴリズムを提供
• GUIウィザード型GUI Oracle Data Miner “Classic”を提供

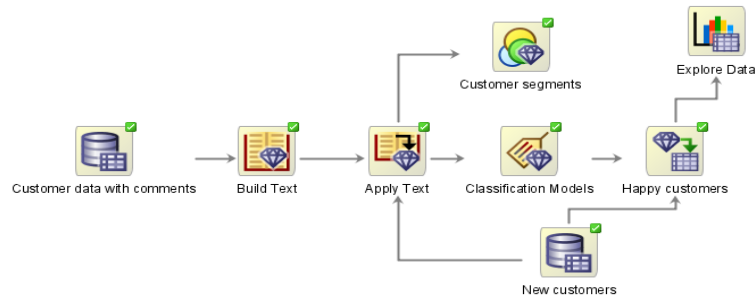
• 自動データ準備 (ADP), テキストマイニング機能追加
• SQL Developerプラグイン Oracle Data Miner 3.2 を提供
• Oracle R Enterpriseによる Oracle Databaseと“R”との関係
• “Oracle Advanced Analytics (ODM + ORE)に名称変更

• 3つの新しいアルゴリズム (EM, PCA, SVD)を追加
• SQLDEV/Oracle Data Miner 4.0 においてSQLスクリプト作成機能+SQLクエリーノード機能、R関係機能が提供
• OAA/ORE 1.3 + 1.4 スケールする新しいRアルゴリズムを提供
• Hadoopと連携するOracle R Advanced Analytics for Hadoopを提供

1998 → 1999 → 2002 → 2004 → 2005 → 2008 → 2011 → 2015

Oracle Advanced Analytics

高度な機械学習エンジンをデータベースに内包



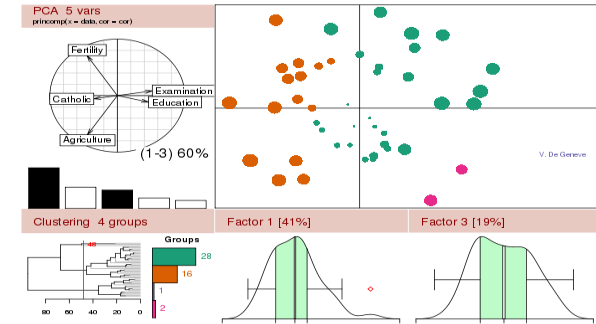
Oracle Data Mining

- データベース内部でマイニング処理
- 12のin-databaseデータマイニングアルゴリズム
- Predictive analytics アプリケーションを開発するための環境
- SQL Developer/Oracle Data MinerによるGUIによる開発
- PL/SQL APIとJava API
- Exadataのパワーを利用したモデルの適用



Oracle R Enterprise

- 利用者の多いOSSの統計言語/環境
- 拡張性を得るためにデータベースに統合
- 広範な統計、高度な分析のための関数を用意
- Rの機能をアプリケーションやOBIEEに統合
- 探索的なデータ分析
- すぐれたグラフ描画機能
- Open source R (CRAN) パッケージ



Statistics

Advanced Analytics

Data Mining

Predictive Analytics

Oracle Advanced Analytics Databaseアーキテクチャ

Users



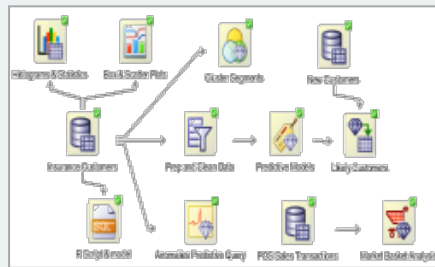
Platform

DBCS

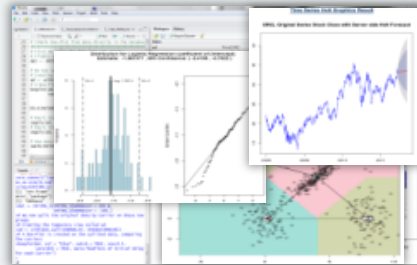
データサイエンティスト
+ ビジネスユーザー

Rプログラマー

SQL Developer



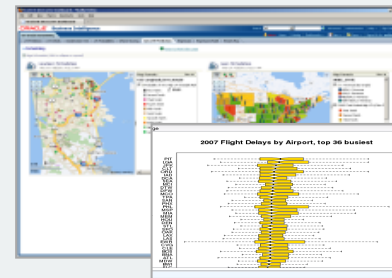
R Enterprise Client



BICS

ビジネスユーザー
(マネージャーレベル)

OBIEE



SaaS

エンドユーザー

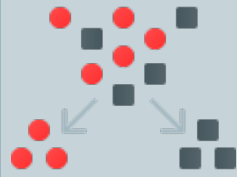
HCM, CRMなど



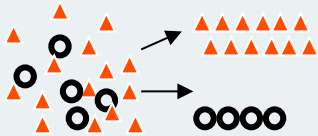
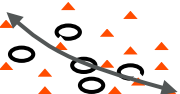



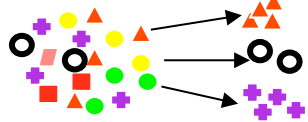
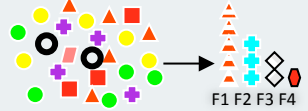
Oracle Database Enterprise Edition

Oracle Advanced Analytics

並列実行のためのSQLデータ・マイニング/分析関数
スケーラブルな分散、高性能のRとの統合



Oracle Advanced Analytics In-Database Data Mining アルゴリズム

| | アルゴリズム | 応用範囲の例 |
|---------|---|---|
| 分類 |  <p>Logistic Regression (GLM) Decision Trees Naïve Bayes Support Vector Machines (SVM)</p> | <p>ロイヤルカスタマーの予測 稼働会員から外れる会員の予測 優良商品へ成長する製品の予測 インバウンドニーズのある店舗の予測</p> |
| 回帰 |  <p>Linear Regression (GLM) Support Vector Machine (SVM)</p> | <p>特定商品の売り上げを予測 特定顧客の消費金額を予測</p> |
| 異常検出 |  <p>One Class SVM</p> | <p>突発的な商品需要の検知 機器の異常値検知</p> |
| 属性重要度 |  <p>Minimum Description Length (MDL) Principal Components Analysis (PCA)</p> | <p>属性の絞り込み、ノイズの低減</p> |
| 相関ルール |  <p>Apriori</p> | <p>バスケット分析/ NBO(Next Best Offer)分析</p> |
| クラスタリング |  <p>Hierarchical k-Means Hierarchical O-Cluster Expectation-Maximization Clustering (EM)</p> | <p>製品のグルーピング/ テキストマイニング 遺伝子-タンパク質分析</p> |
| 特徴抽出 |  <p>Nonnegative Matrix Factorization (NMF) Singular Value Decomposition (SVD)</p> | <p>テキスト分析 / 因子分析</p> |

アジェンダ

- 1 マシン・ラーニング（機械学習）概要
- 2 Oracle Advanced Analytics の紹介
- 3 クラウド・マシン・ラーニング**
- 4 Oracle Data Miner で設計する予測分析
- 5 まとめ

Oracle Database Cloud Service

No.1データベースをそのままクラウド上で利用可能



すぐに使える
すぐに拡張できる

データベースが使えるまで、わずか**5画面**
リソースが足りなくなったら、すぐに拡張できる

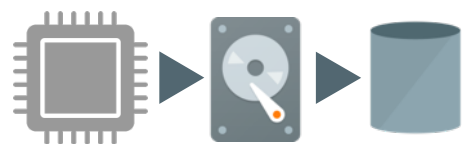
使った分だけ
お支払い

サーバー費用は**従量制**（1時間50円～）
ソフトウェアライセンス（オプション含）も**従量制**

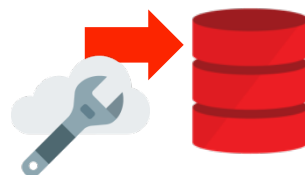
エンター
プライズ対応

定番機能から**機械学習**まですぐ使える形で提供
最高のデータベース基盤である **Exadata** を選択可能

潜在的な管理コストを含めたTCO削減



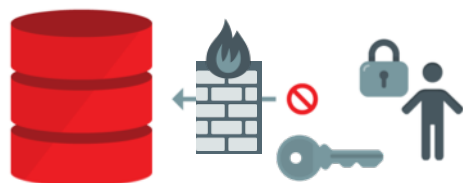
数クリックで
プロビジョニング



ワンクリックで
パッチ適用



容易な可用性対策
Backup/HA/DR



DB暗号化
Advanced Security

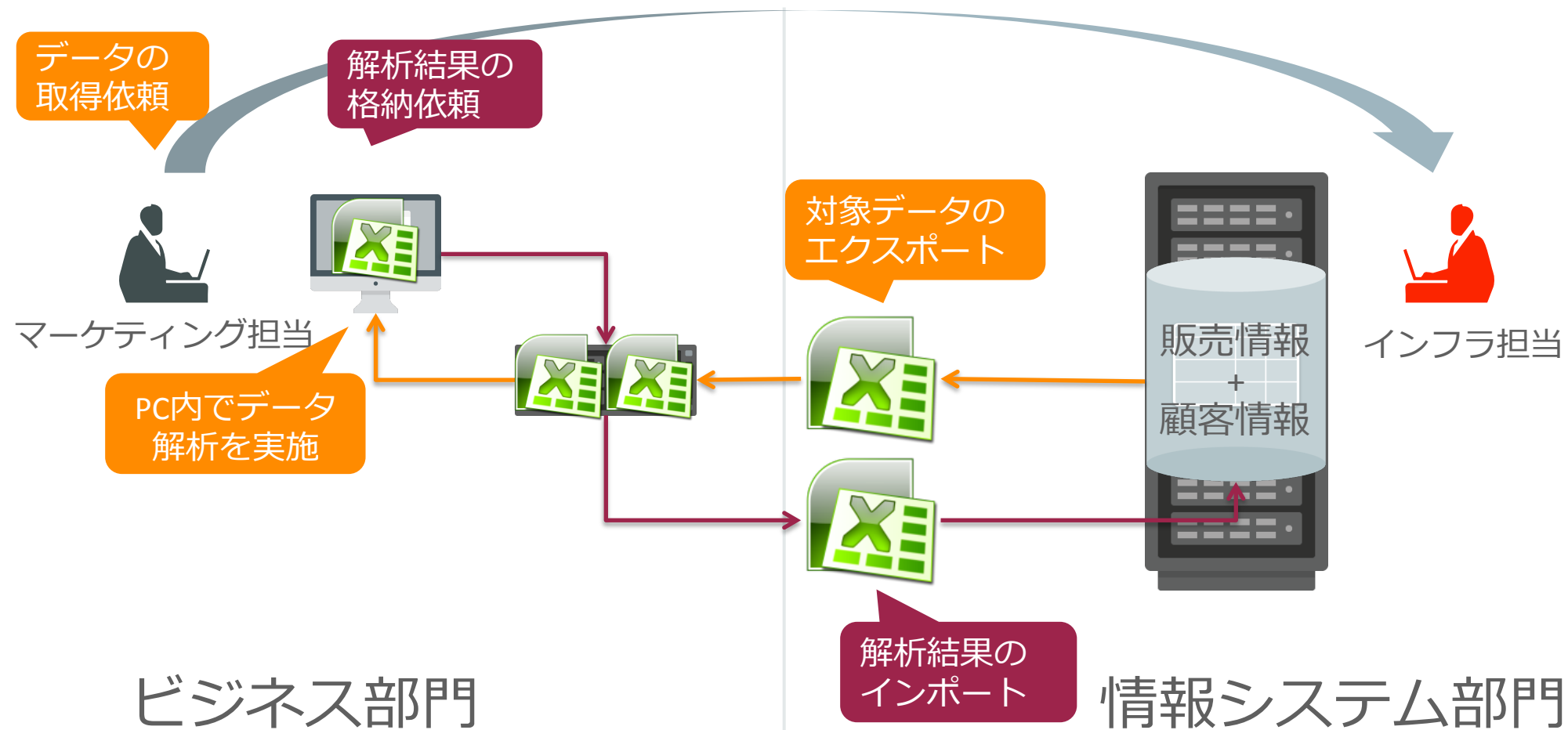


DB Monitor
クラウド専用
モニタリングツール

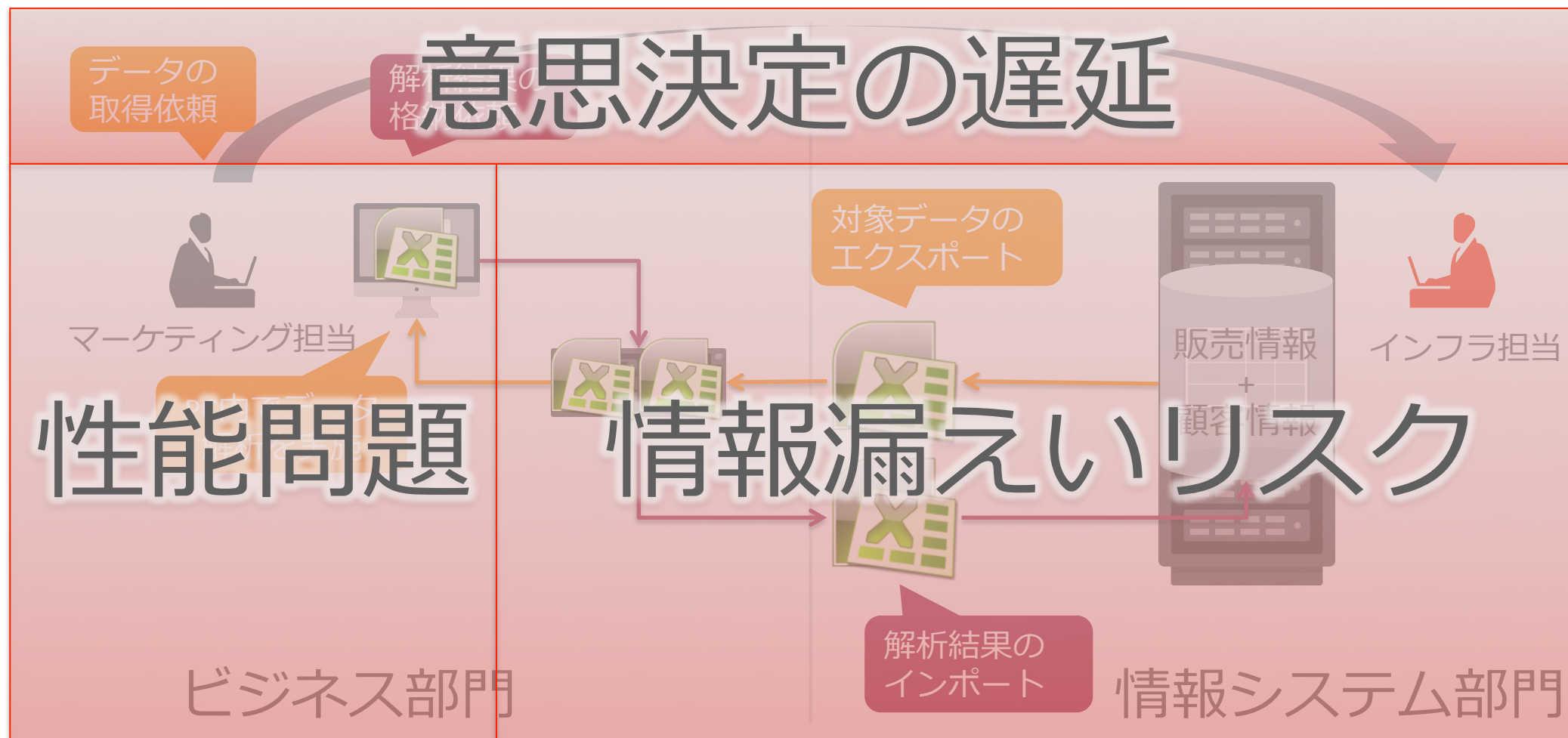


インフラ管理は
不要

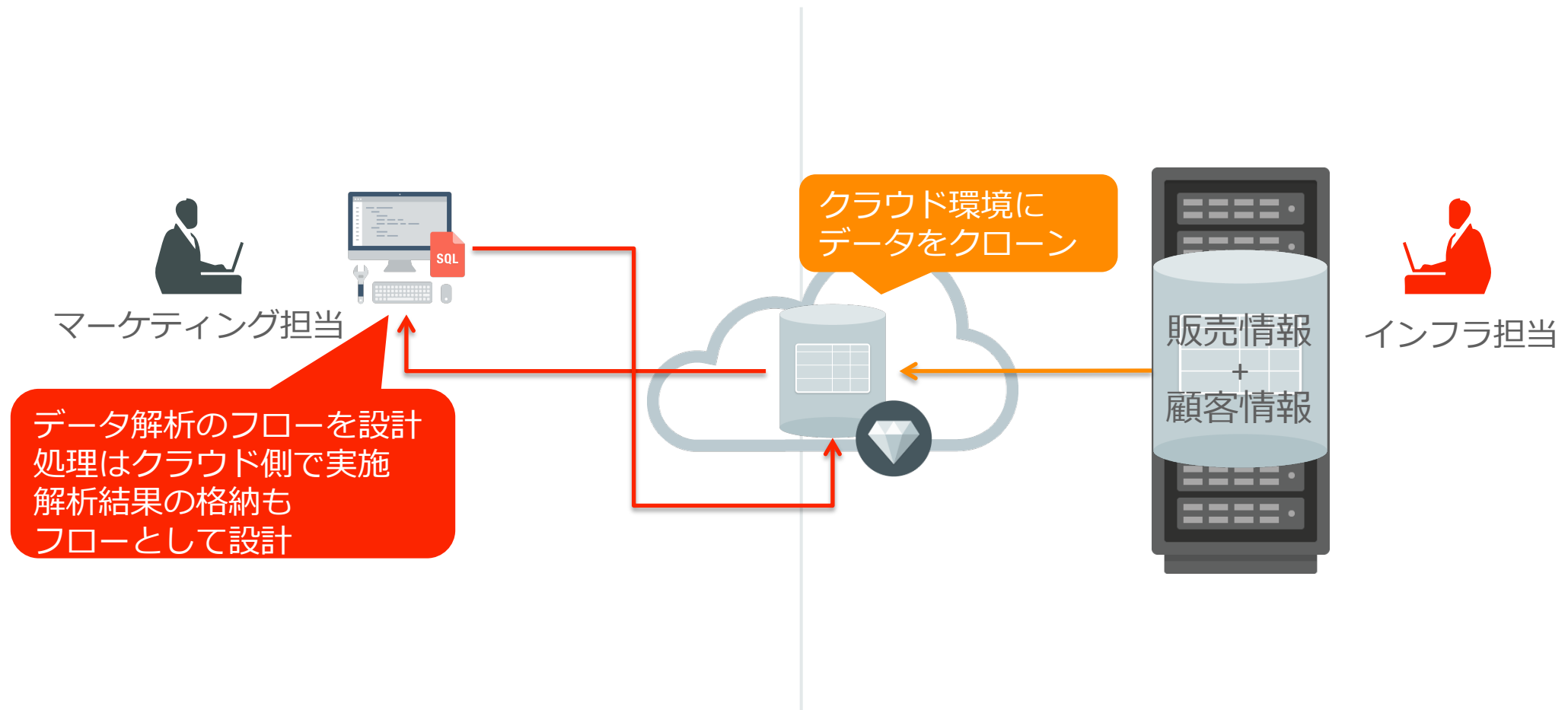
典型的なデータ解析基盤の課題



典型的なデータ解析基盤の課題



クラウド・マシン・ラーニング



クラウド・マシン・ラーニング

クイック・スモールスタート

オンプレミスにも移行可能な

マーケティング担当

ハイブリッドクラウド

インフラ担当

データ解析のフローを設計

処理はクラウド側で実施

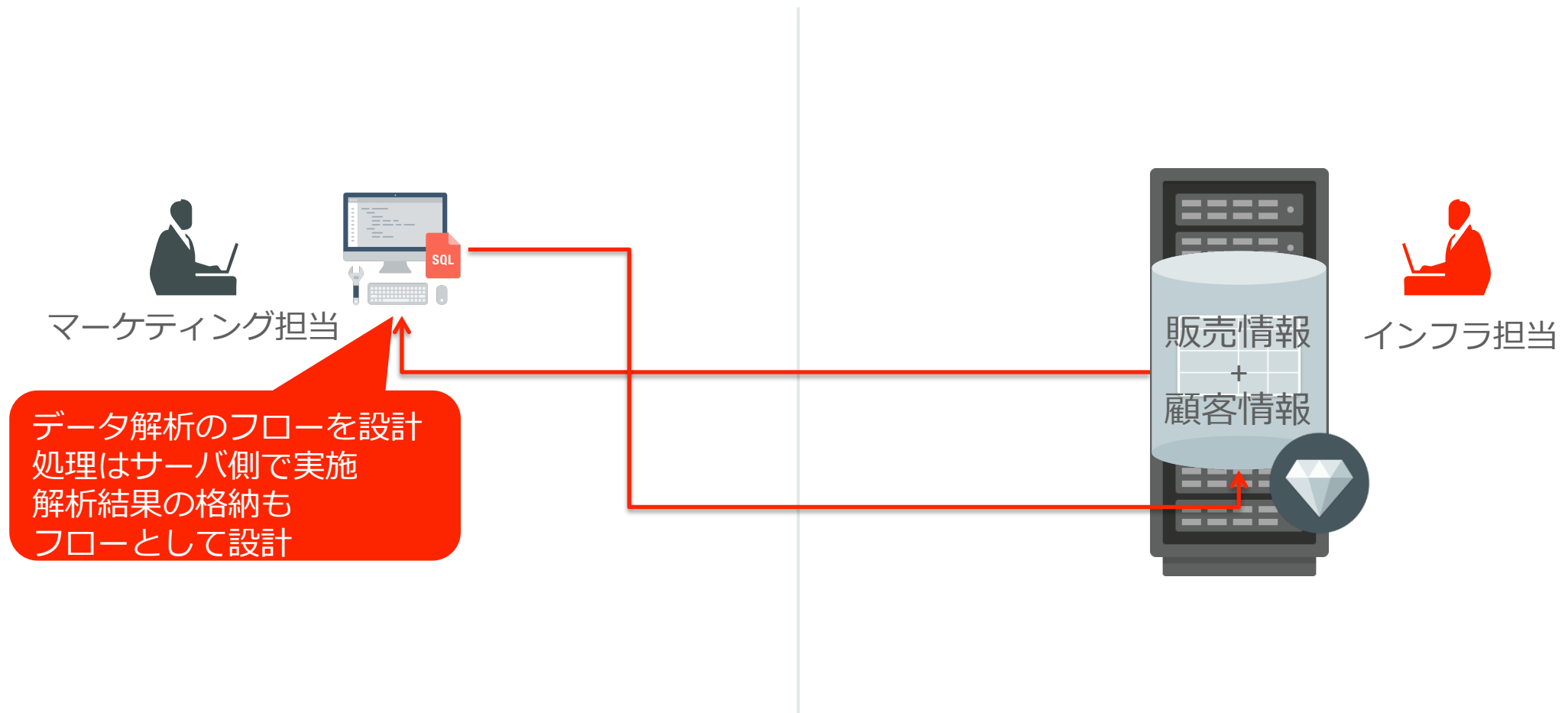
解析結果の

フローとして

暗号化・権限分掌による

高セキュリティレベル

参考：オンプレミス・マシン・ラーニング



参考：オンプレミス・マシン・ラーニング

単一部門で完結する意思決定

サーバ性能をフルに使える
高パフォーマンス解析基盤

データ解析のフローを設計

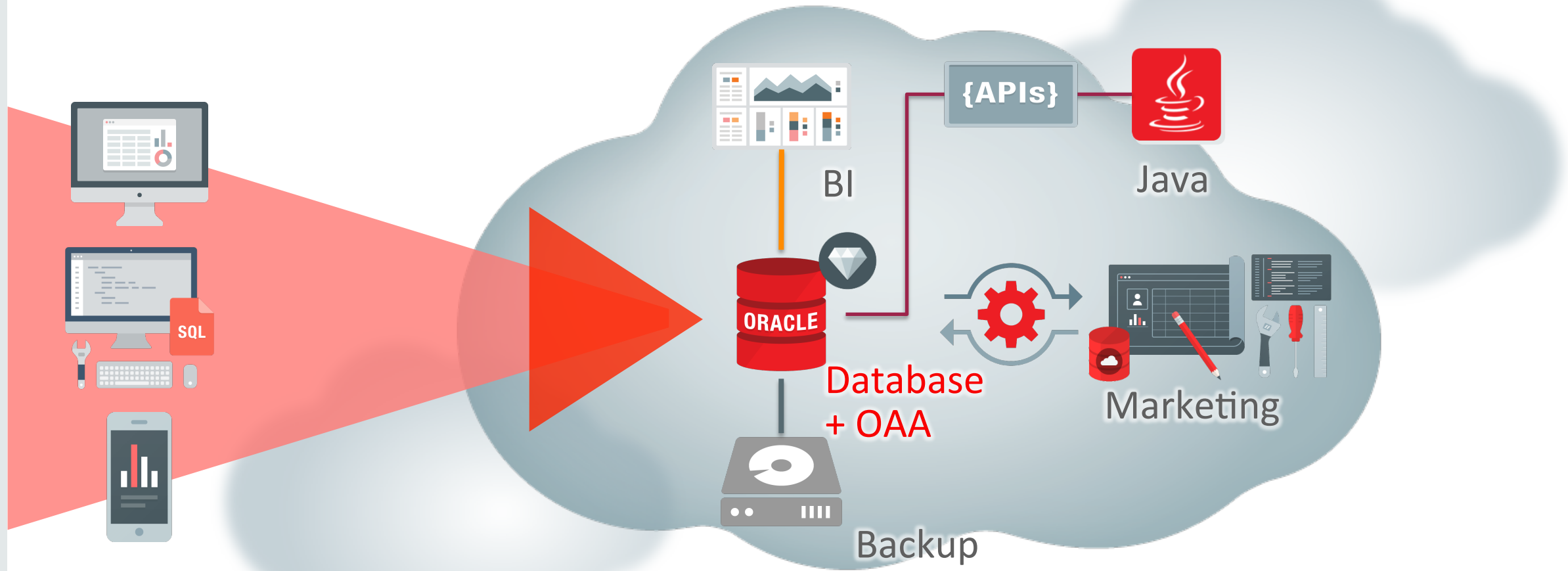
処理はサーバ側で実施

解析結果の共有

フローとして実行

暗号化・権限分掌による
高セキュリティレベル

クラウド・マシン・ラーニングと 他のOracle Cloudとの連携

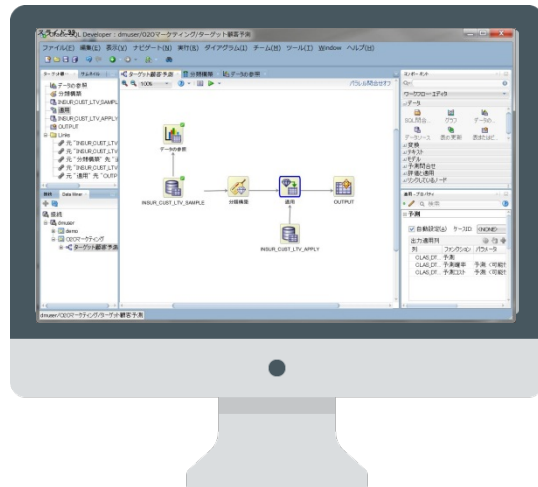


アジェンダ

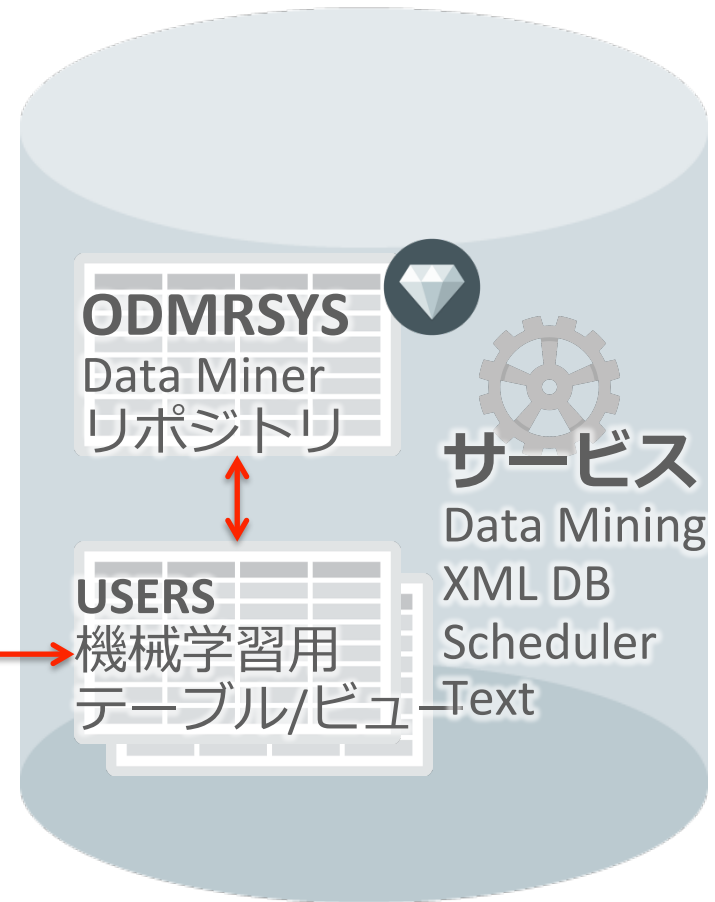
- 1 マシン・ラーニング（機械学習）概要
- 2 Oracle Advanced Analytics の紹介
- 3 クラウド・マシン・ラーニング
- 4 Oracle Data Miner で設計する予測分析**
- 5 まとめ

Oracle Data Minerアーキテクチャ

SQL Developer
+ Data Miner



SQL, PL/SQL



Oracle Advanced Analytics Oracle Data Miner GUI

- GUIによる操作で分析フローを設計
- フローをイメージで保存可能
- クライアントツールからデータベースを操作
- 複雑な機械学習の処理を自動で同時実行
- モデリングに適したデータに自動変換
- ETL処理も実行可能

The screenshot shows the Oracle Data Miner GUI within the Oracle SQL Developer environment. The main workspace displays a workflow diagram with several nodes: 'Explore Data 2', 'Colum filter and AI', '5 Response Mpdols', 'NEW CUST_INSUR_LTV1', 'Apply 14', and 'Clust Build 21'. A 'Component Palette' on the right side lists various models and data sources, including 'Anomaly Detection', 'Association', 'Classification', 'Clustering', 'Feature Extraction', 'Link', 'Model', 'Model Details', 'Regression', 'Evaluate and Apply', 'Apply', 'Link', 'Test', 'Data', 'Create Table', 'Data Source', 'Explore Data', 'Transforms', 'Aggregate', 'Filter Columns', 'Filter Columns Details', 'Filter Rows', 'Join', 'Link', 'Sample', 'Transform', 'Text', 'Apply Text', 'Build Text', 'Link', and 'Text Reference'. A 'Property Inspector' at the bottom shows the settings for the '5 Response Mpdols' model, including a table of model settings.

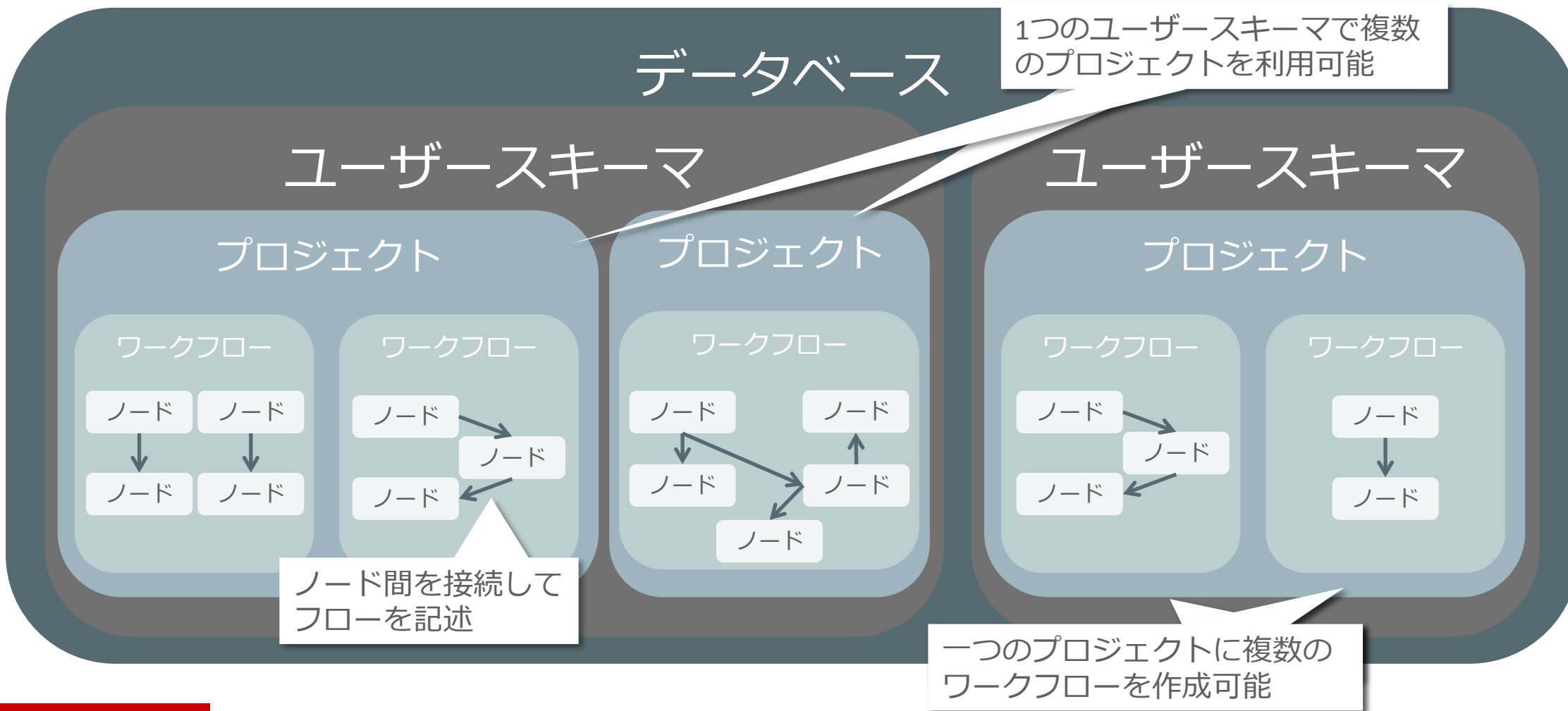
| Name | Build | Test | Tune | Algorithm | Comment |
|--------------|-----------------|-----------------|-----------|---------------------|---------|
| CLAS_GLM_... | 7/13/10 6:07... | 7/13/10 6:07... | Automatic | Generalized Line... | |
| CLAS_SVM_... | 7/13/10 6:06... | 7/13/10 6:06... | Automatic | Support Vector ... | |
| CLAS_SVM_... | 7/13/10 6:06... | 7/13/10 6:06... | Automatic | Support Vector ... | |
| CLAS_DT_3_3 | 7/13/10 6:06... | 7/13/10 6:06... | Automatic | Decision Tree | |
| CLAS_NB_3_3 | 7/13/10 6:06... | 7/13/10 6:06... | Automatic | Naive Bayes | |

A red arrow points to the workflow editor with the text 'ドラッグ&ドロップ' (Drag & Drop).

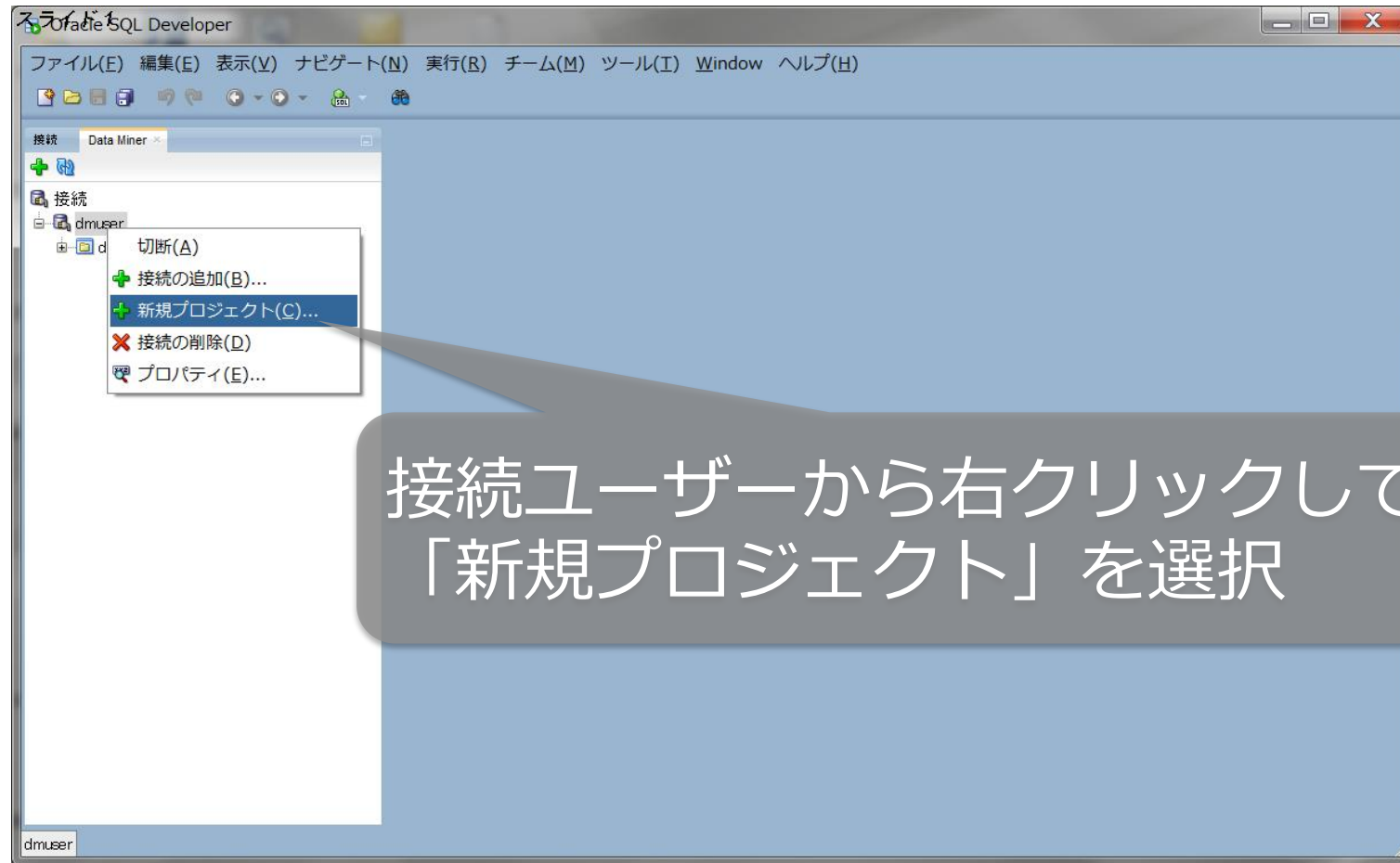
Oracle Data Minerの主な概念

| 用語 | 概要 | 補足 |
|----------|--|---|
| ユーザースキーマ | データベース接続スキーマ データマイニング実施アカウント | 複数作成可能 |
| プロジェクト | 分析プロジェクト | ワークフローの集合体 プロジェクト名とコメントを付与可能 |
| ワークフロー | データマイニングオペレーション(モデリング、テスト、スコアリング)の設計領域 | ノードが結びついたもの ノードの組み合わせによって、 データマイニングオペレーションを設計可能 |
| ノード | データノード 変換ノード テキストノード モデルノード 予測問合せノード 評価と適用ノード リンクノード | ノード間を接続して、フローを実現 名前とコメントを付与可能 詳細パラメータを設定できる ノード単位で実行を行える ※実行に必要な他のノードがあれば、 そちらも自動的に実行される ノードのステータスが表示 |

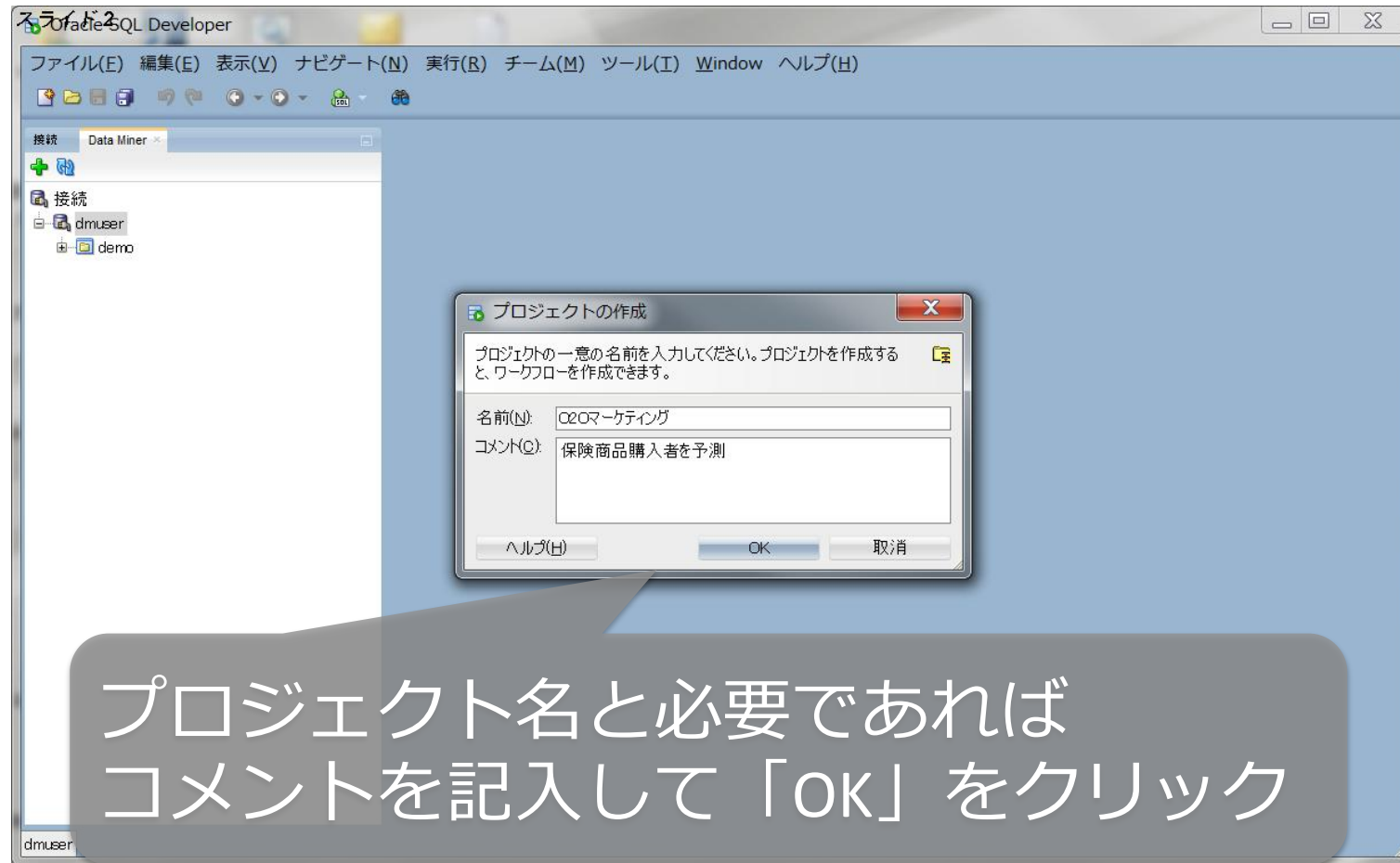
Oracle Data Minerの各概念の関係



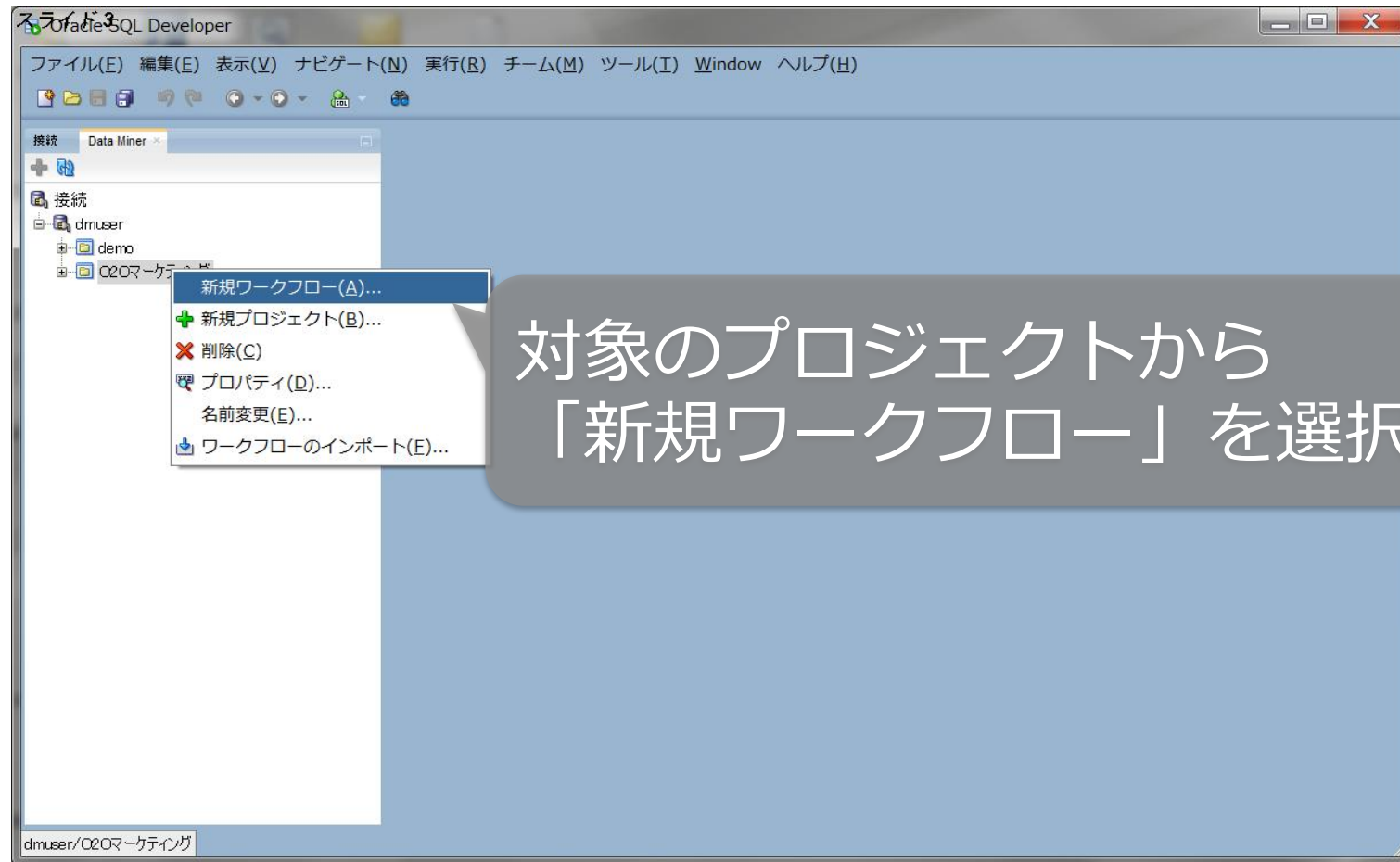
Step1 新規プロジェクト作成(1/2)



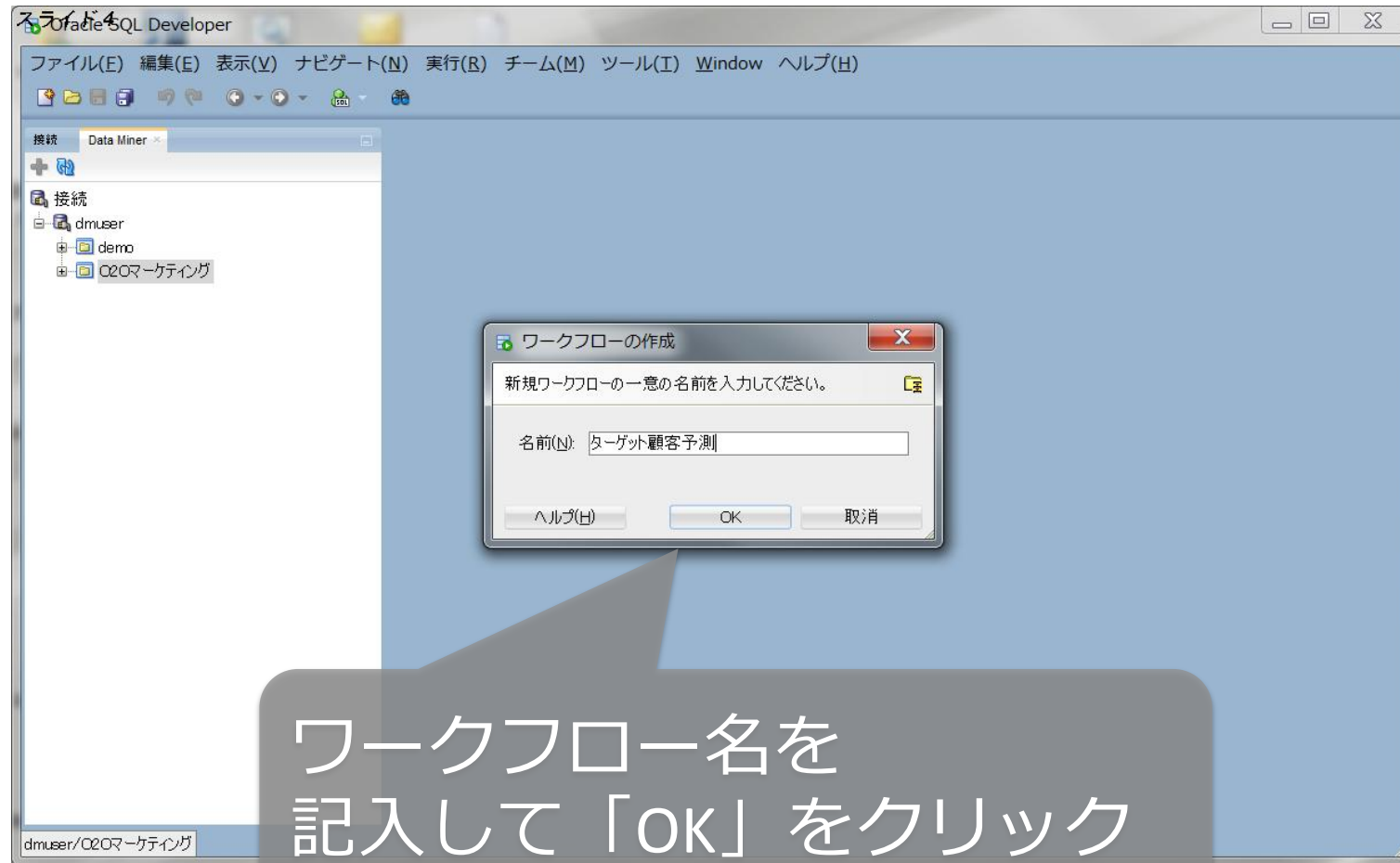
Step1 新規プロジェクト作成(2/2)



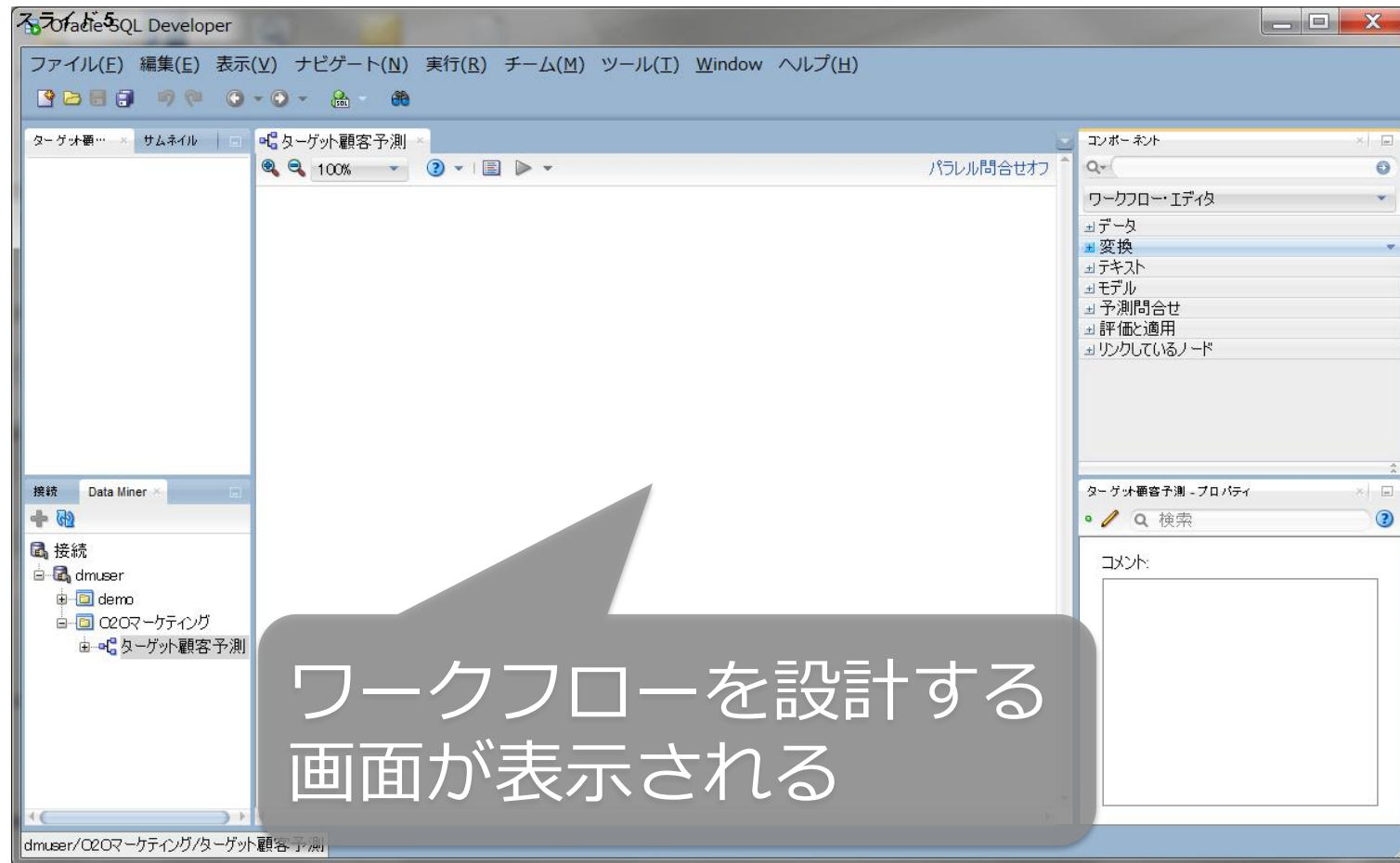
Step2 新規ワークフロー作成(1/3)



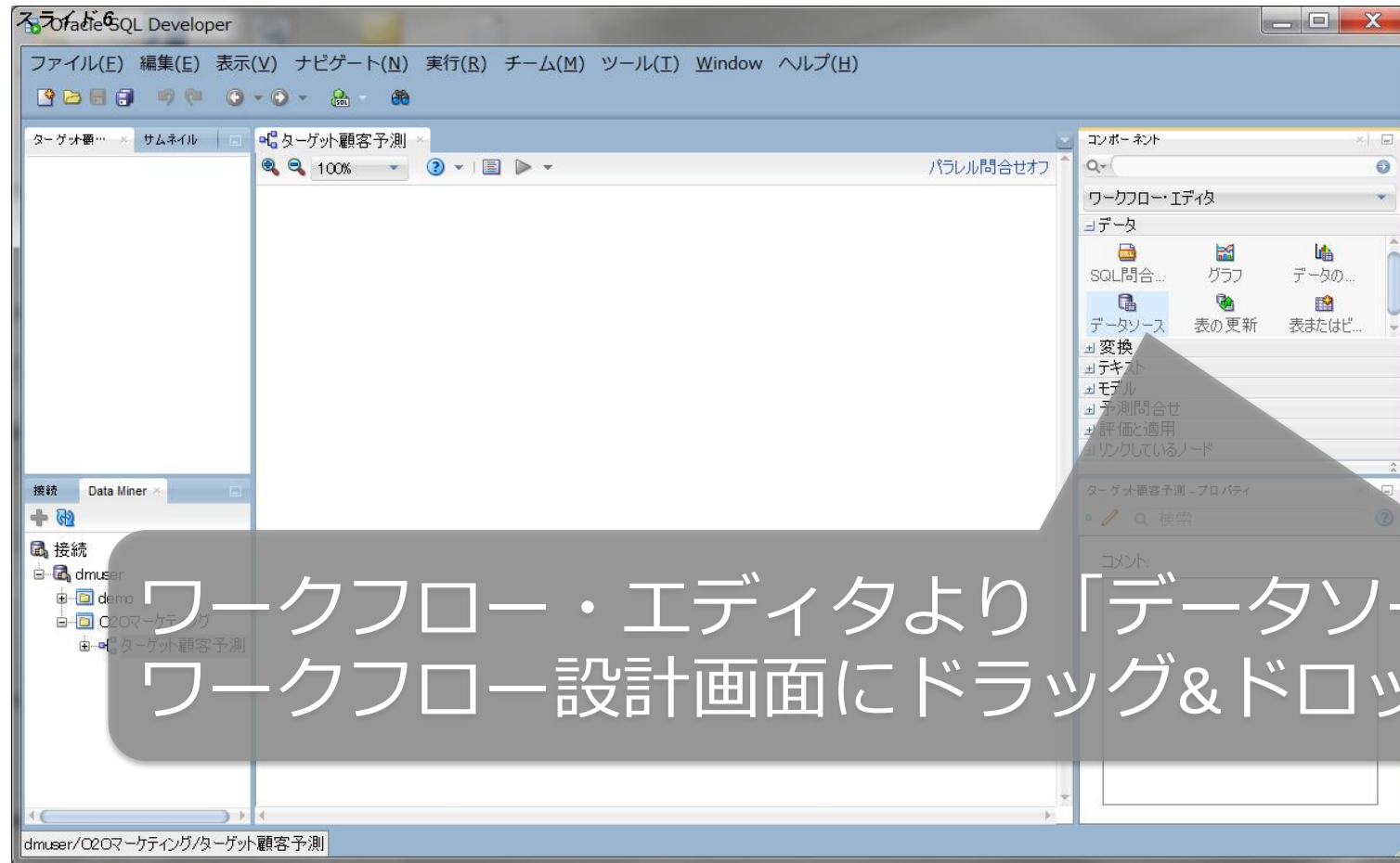
Step2 新規ワークフロー作成(2/3)



Step2 新規ワークフロー作成(3/3)



Step3 データソースを選択(1/3)



Step3 データソースを選択(2/3)

データの定義 - ステップ 1 / 2

表の選択

使用可能な表/ビュー:

| 名前 | タイプ |
|-------------------------------------|--------------|
| DMUSER.CHURN_CT_DATA | TABLE |
| DMUSER.CLUS_ID | TABLE |
| DMUSER.INSUR_CUST_LTV_SAMPLE | TABLE |
| DMUSER.KEIRIN_CUST_LTV_SAMPLE | TABLE |
| DMUSER.MINING_DATA_APPLY_V | VIEW |
| DMUSER.MINING_DATA_BUILD_V | VIEW |
| DMUSER.MINING_DATA_TEST_V | VIEW |
| DMUSER.MINING_DATA_TEXT_APPLY_V | VIEW |
| DMUSER.MINING_DATA_TEXT_BUILD_V | VIEW |
| DMUSER.MINING_DATA_TEXT_TEST_V | VIEW |
| DMUSER.ODMR_CARS_DATA | TABLE |
| DMUSER.ODMR_SALES_JSON_DATA | TABLE |

スキーマの編集(E)

対象の表を選択して「終了」をクリック
※一部の列を選択可能

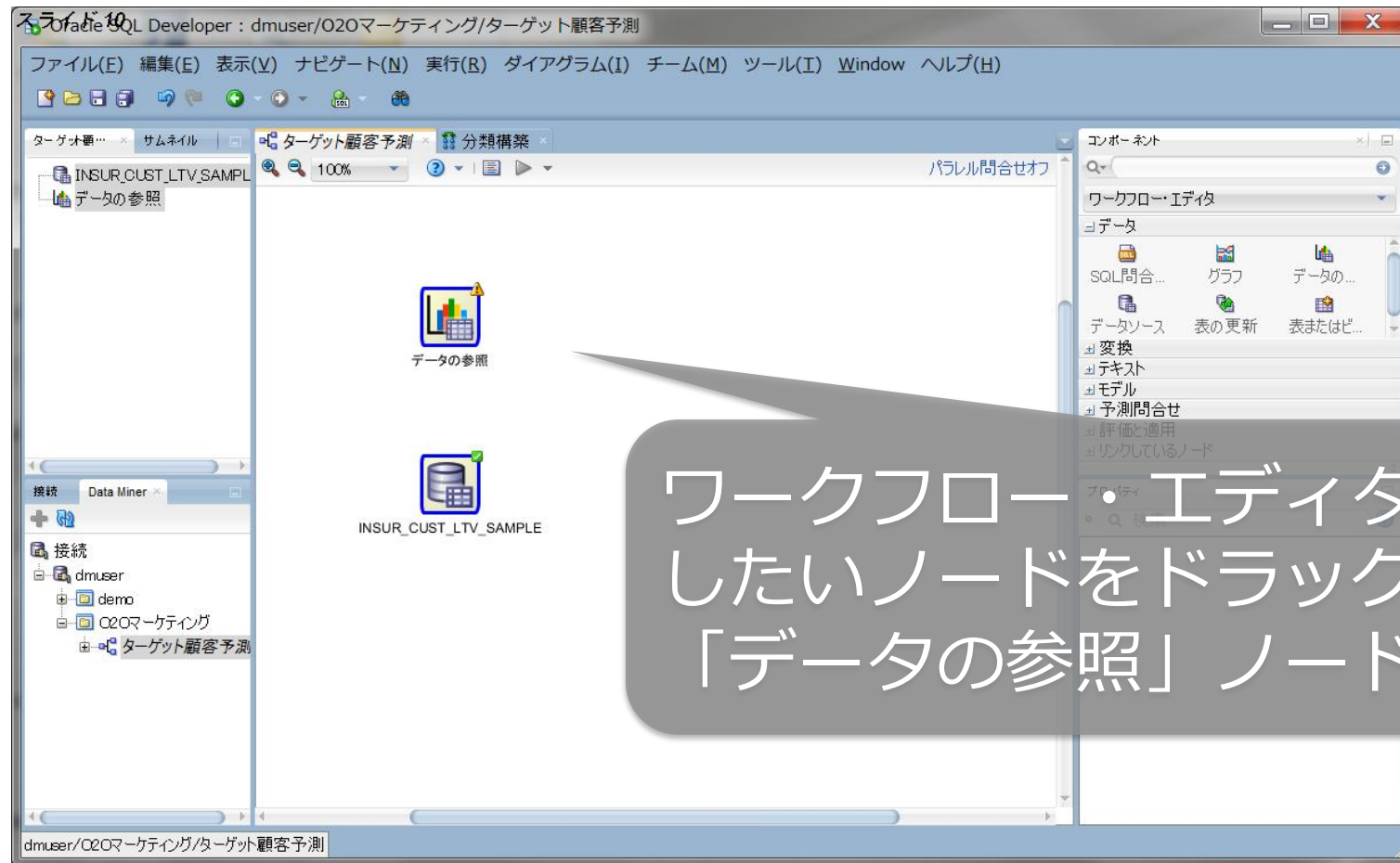
| データ型 | マイニング型 | 長さ | 列ID |
|----------|--------|-----|-----|
| NUMBER | 数値型 | 22 | 9 |
| NUMBER | 数値型 | 22 | 18 |
| VARCHAR2 | カテゴリ型 | 100 | 8 |
| NUMBER | 数値型 | 22 | 13 |
| NUMBER | 数値型 | 22 | 19 |
| NUMBER | 数値型 | 22 | 17 |
| NUMBER | 数値型 | 22 | 27 |
| VARCHAR2 | カテゴリ型 | 100 | 1 |
| VARCHAR2 | カテゴリ型 | 100 | 3 |

ヘルプ(H) < 戻る(B) 次(N) > 終了(E) 取消

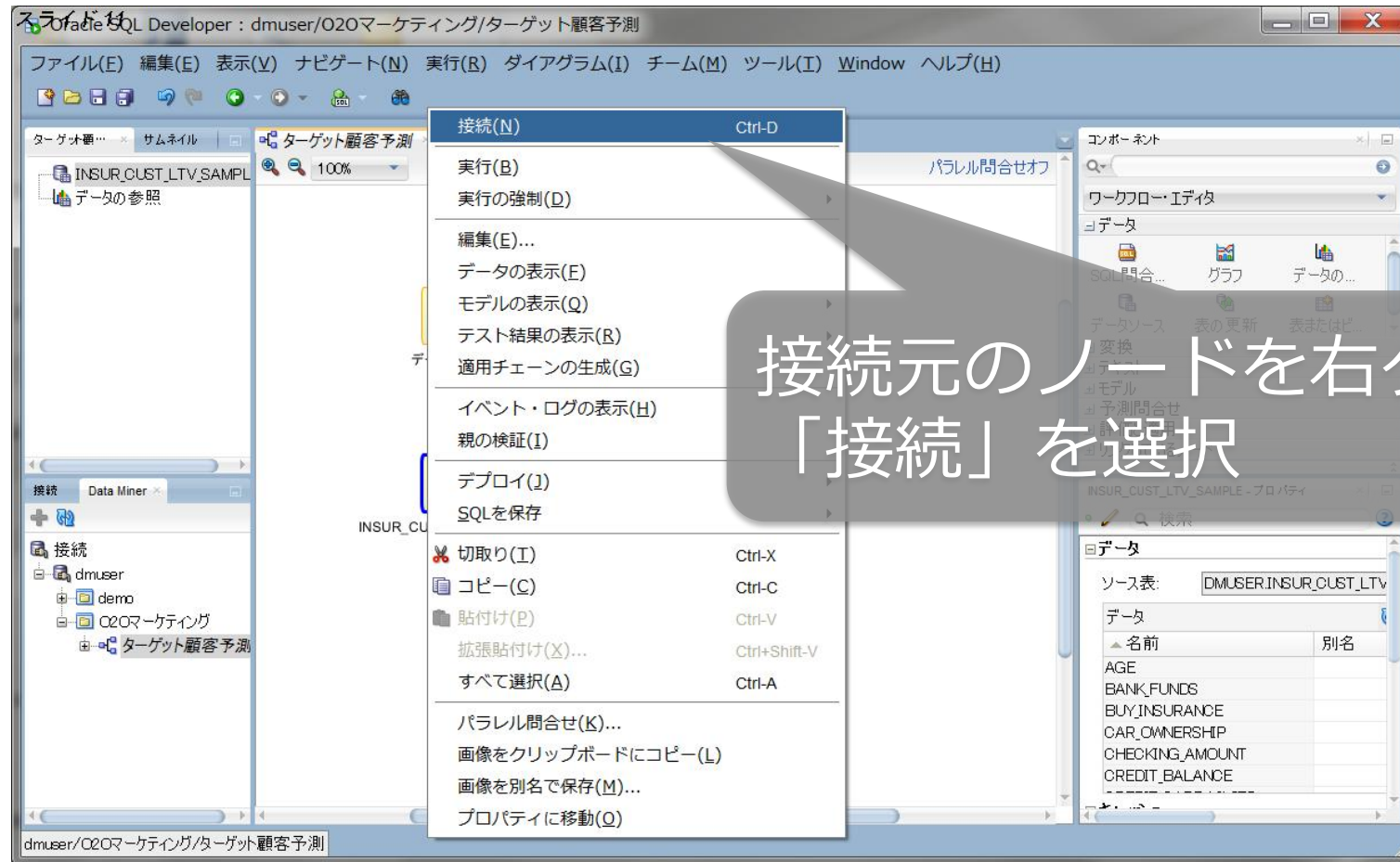
Step3 データソースを選択(3/3)

The screenshot shows the Oracle SQL Developer interface. The main workspace displays a diagram with a node labeled 'INSUR_CUST_LTV_SAMPLE'. A callout box points to this node with the text: 対象の表名を持ったノードが作成される. The right-hand pane shows the 'INSUR_CUST_LTV_SAMPLE - プロパティ' window, which lists the source table as 'DMUSER.INSUR_CUST_LTV' and displays a list of columns including AGE, BANK_FUNDS, BUY_INSURANCE, CAR_OWNERSHIP, CHECKING_AMOUNT, and CREDIT_BALANCE.

Step4 ノードの接続とデータ統計の確認(1/5)



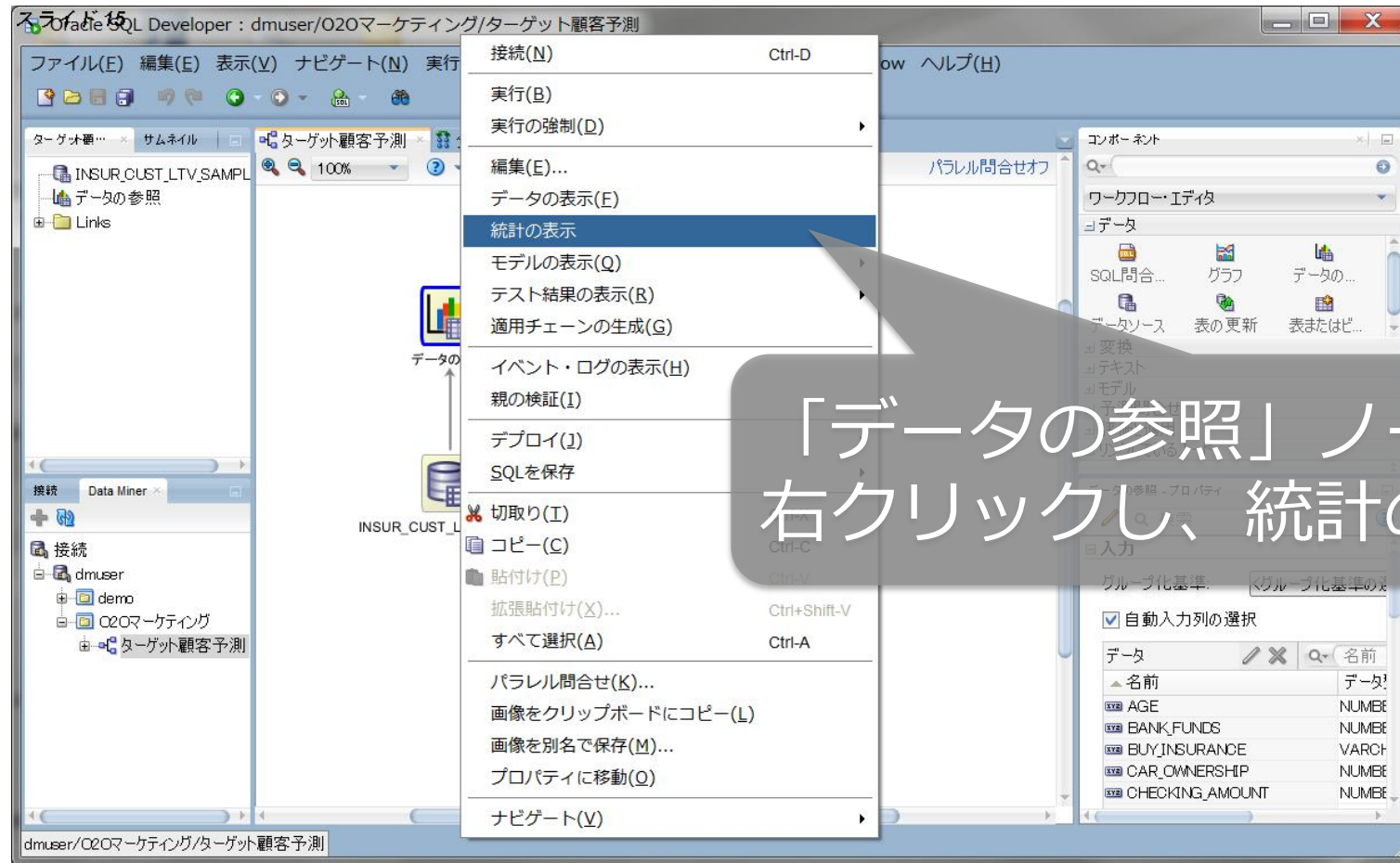
Step4 ノードの接続とデータ統計の確認(2/5)



Step4 ノードの接続とデータ統計の確認(3/5)

接続先のノードをクリックすると矢印で結ばれる

Step4 ノードの接続とデータ統計の確認(4/5)



Step4 ノードの接続とデータ統計の確認(5/5)

Oracle SQL Developer : dmuser/O2Oマーケティング/ターゲット顧客予測/データの参照

ファイル(E) 編集(E) 表示(V) ナビゲート(N) 実行(R) チーム(M) ツール(I) Window ヘルプ(H)

データの参照 - 構造

ターゲット顧客予測 分類構築 データの参照

統計 データ参照

dmuser/O2Oマーケティング/ターゲット顧客予測

統計: 101,015からの列 行(サンプル) NULLを表示 フィルタ(E): 名前

| 名前 | ヒストグラム | データ型 | パーセ... | 個別値 | 個別... | モード | 平均 | 中間 | 最小... | 最大... | 標準偏... | 分散 |
|--------------------|--------|------------|--------|---------|-------|-----|------------|-------|-------|---------|--------------|-------------|
| AGE | | NUMBER 0 | 66 | 6.5025 | | | 38.1901 | 36 | 0 | 84 | 14.9184 | 222.5585 |
| BANK_FUNDS | | NUMBER 0 | 280 | 27.5862 | | | 2,639.8... | 500 | 0 | 36,000 | 4,996.1007 | 24,961,0... |
| BUY_INSURANCE | | VARCH... 0 | 2 | 0.197 | No | | | | | | | |
| CAR_OWNERSHIP | | NUMBER 0 | 2 | 0.197 | | | 0.9468 | 1 | 0 | 1 | 0.2245 | 0.0504 |
| CHECKING_AMOUNT | | NUMBER 0 | 327 | 32.2167 | | | 1,065.8... | 25 | 25 | 23,476 | 3,124.406 | 9,761,91... |
| CREDIT_BALANCE | | NUMBER 0 | 93 | 9.1626 | | | 2,234.1... | 0 | 0 | 170,498 | 11,727.87... | 137,543,... |
| CREDIT_CARD_LIMITS | | NUMBER 0 | 23 | 2.266 | | | 1,285.6... | 1,000 | 500 | 5,000 | 858.2067 | 736,518... |
| CUSTOMER_ID | | VARCH... 0 | 1,015 | 100 | <その他> | | | | | | | |
| FIRST | | VARCH... 0 | 847 | 83.4483 | <その他> | | | | | | | |
| HAS_CHILDREN | | NUMBER 0 | 2 | 0.197 | | | | | | | | |
| HOUSE_OWNERSHIP | | NUMBER 0 | 3 | 0.2956 | | | | | | | | |

接続 Data Miner

接続

- dmuser
 - demo
 - O2Oマーケティング
 - ターゲット顧客予測

dmuser/O2Oマーケティング/ターゲット顧客予測

接続元ノードの各列の統計情報をグラフで表示

| 年齢範囲 | パーセント |
|---------------|-------|
| < 8.4 | ~2% |
| 8.4 - < 16.8 | ~18% |
| 16.8 - < 25.2 | ~25% |
| 25.2 - < 33.6 | ~18% |
| 33.6 - < 42 | ~22% |
| 42 - < 50.4 | ~10% |
| 50.4 - < 58.8 | ~8% |
| 58.8 - < 67.2 | ~5% |
| 67.2 - < 75.6 | ~3% |
| >= 75.6 | ~1% |

Step5 予測モデル(教師あり学習)を作成(1/2)

ワークフロー・エディタより「分類構築」ノードをドラッグ&ドロップし、データソースと接続する

Step5 予測モデル(教師あり学習)を作成(2/2)

スライド 24

Oracle SQL Developer : dmuser/O2Oマーケティング/ターゲット顧客予測

ファイル(E) 編集(E) 表示(V) ナビゲート(N) 実行(R) ダイアグラム(I) チーム(M) ツール(I) Window ヘルプ(H)

ターゲット... サムネイル

データの参照
INSUR_CUST_LTV_SAMPL
分類構築
Links

分類ビルド・ノードの編集

ビルド 入力 テキスト

ターゲット: BUY_INSURANCE

ケースID: CUSTOMER_ID

ターゲットフラグを選択

モデル設定

| 名前 | アルゴリズム | 日付 | データの使用方法 |
|-------------|------------------------|----|-------------------------------------|
| CLAS_GLM1_5 | 一般化線形モデル | | <input checked="" type="checkbox"/> |
| CLAS_SVM1_5 | Support Vector Machine | | <input checked="" type="checkbox"/> |
| CLAS_DT1_5 | ディビジョン・ツリー | | <input checked="" type="checkbox"/> |
| CLAS_NB1_5 | Naive Bayes | | <input checked="" type="checkbox"/> |

デフォルトで分類モデルでは、4つのアルゴリズムが実行される

ヘルプ(H) OK 取消

dmuser/O2Oマーケティング/ターゲット顧客予測

Step5 予測モデルの予測精度を確認(1/4)

Oracle SQL Developer : dmuser/O2Oマーケティング/ターゲット顧客予測

スライド 22

ファイル(E) 編集(E) 表示(V) ナビゲート(N) 実行(R) ダイアグラム(I) チーム(M) ツール(T) Window ヘルプ(H)

ターゲット顧客予測 100%

データの参照

INSUR_CUST_LTV_SAMPL

分類構築

Links

データの参照

INSUR_CUST_LTV_SAMPL

接続 Data Miner

接続

dmuser

demo

O2Oマーケティング

ターゲット顧客予測

接続(N) Ctrl-D

実行(B)

実行の強制(D)

編集(E)...

拡張設定(E)...

モデルの表示(Q)

テスト結果の表示(B)

テスト結果の比較(S)

適用チェーンの生成(G)

イベント・ログの表示(H)

親の検証(I)

すべて選択(A) Ctrl-A

パラレル問合せ(K)...

画像をクリップボードにコピー(L)

画像を別名で保存(M)...

プロパティに移動(Q)

ナビゲート(V)

コンポーネント

ワークフロー エディタ

データ

変換

テキスト

モデル

アシエン... クラスタリ... モデル

モデル詳細 分類 帰属

予測問合せ

評価と適用

名前 ビルド テスト チ...

CLAS_GLM1_5 → 15... 15... 自動

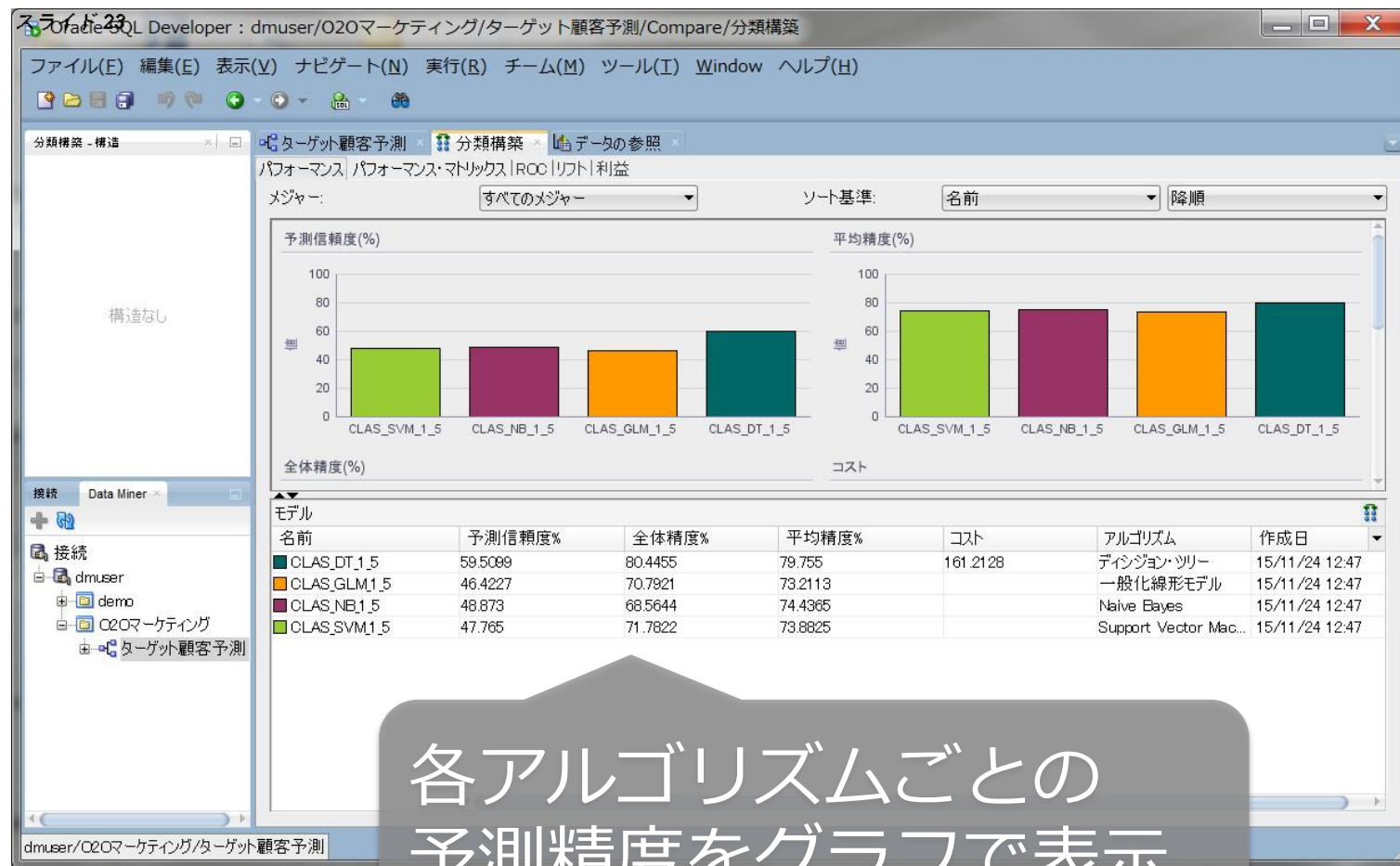
CLAS_SVM1_5 → 15... 15... 自動

CLAS_DT_1_5 → 15... 15... 自動

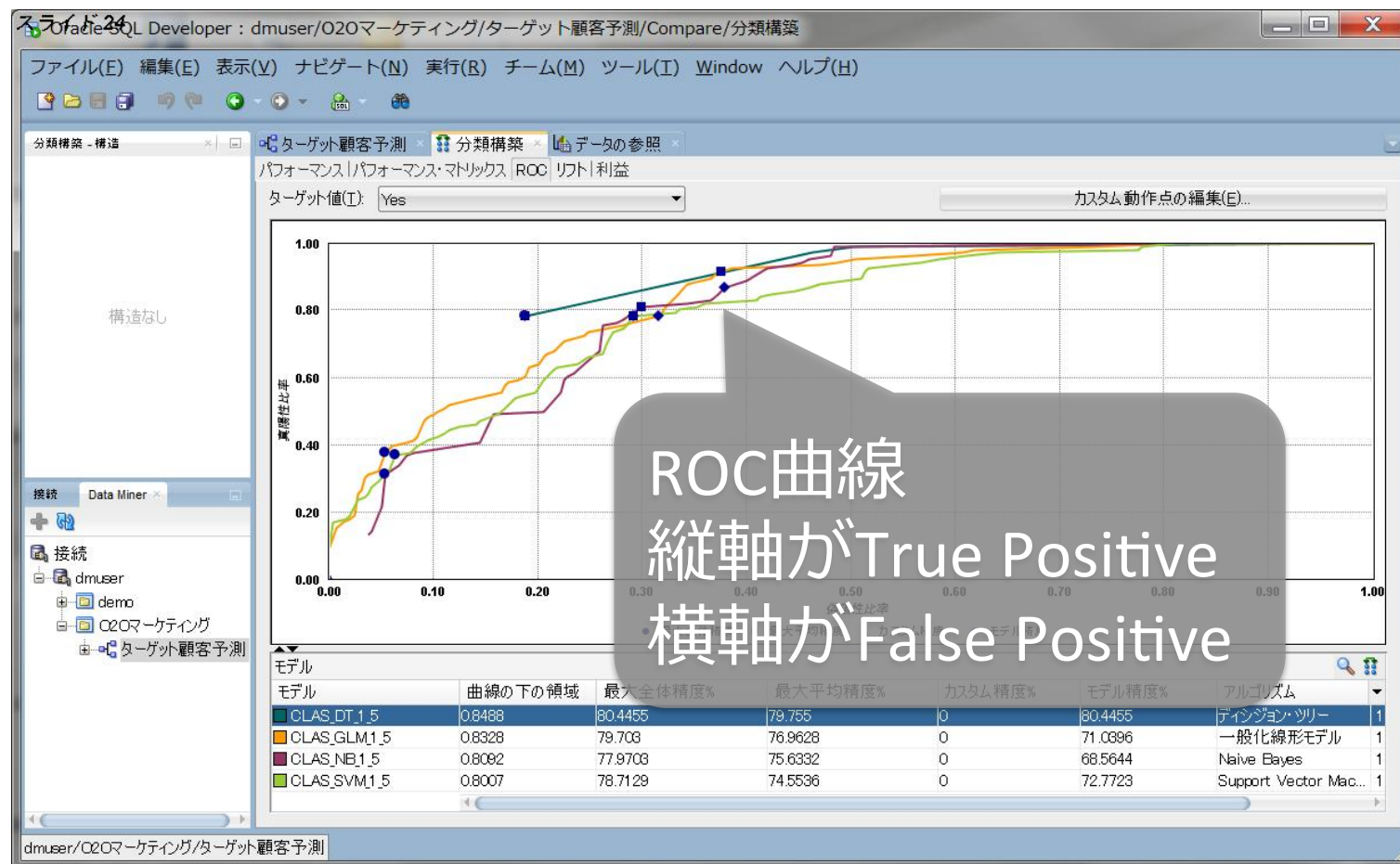
CLAS_NB1_5 → 15... 15... 自動

「分類構築」ノードを右クリックし、「テスト結果の比較」を選択

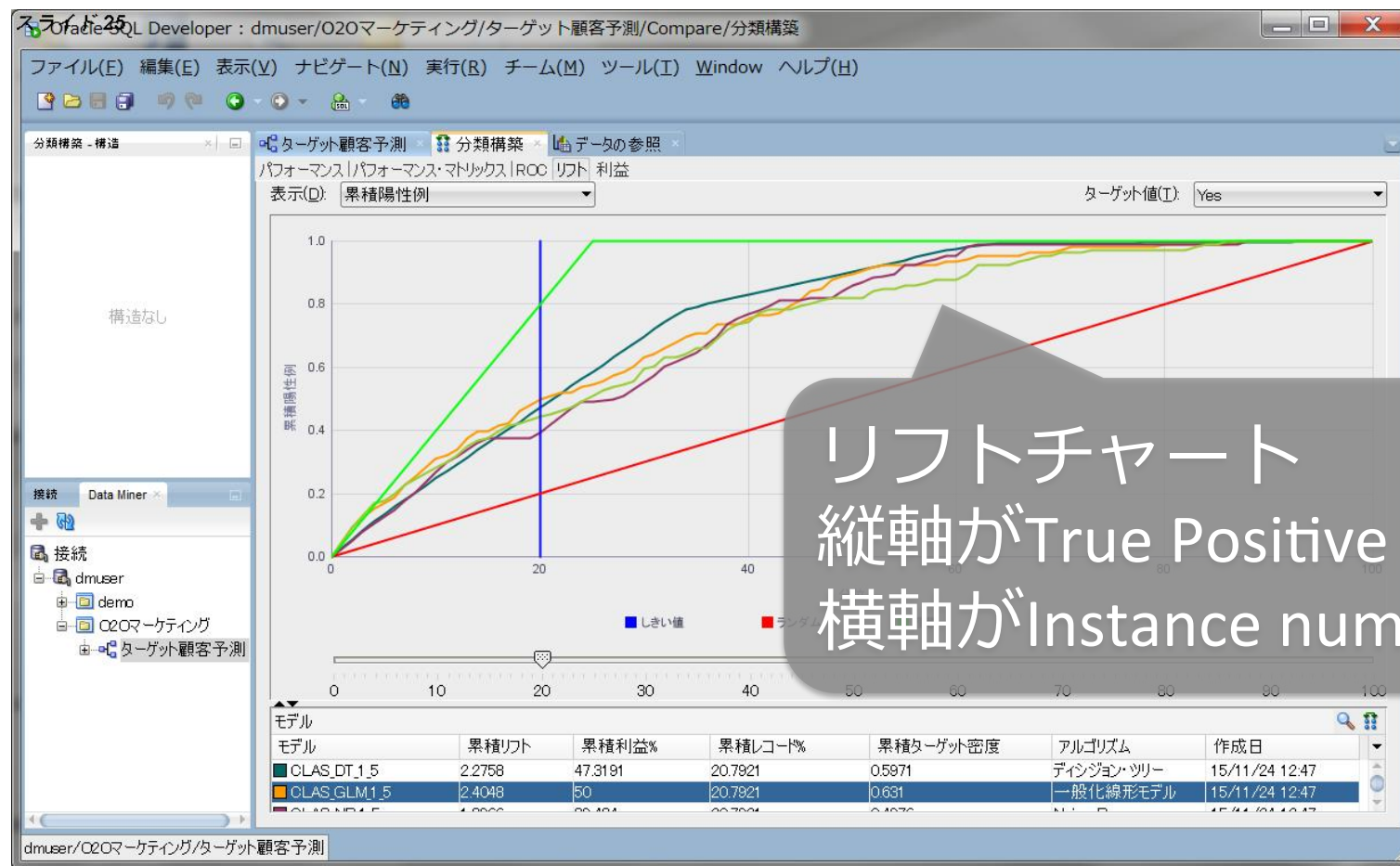
Step5 予測モデルの予測精度を確認(2/4)



Step5 予測モデルの予測精度を確認(3/4)



Step5 予測モデルの予測精度を確認(4/4)



Step6 予測モデルのアルゴリズムを選択

The screenshot shows the Oracle Data Miner interface. The central workspace displays a workflow diagram with three nodes: 'データの参照' (Data Reference) at the top, 'INSUR_CUST_LTV_SAMPLE' (Data Source) at the bottom left, and '分類構築' (Classification Model) at the bottom right. Arrows indicate the flow from the data source to the data reference, and then to the classification model.

On the right side, the 'コンポーネント' (Components) pane is open, showing a list of models under the 'モデル' (Models) section. The following table represents the data shown in this pane:

| 名前 | ビルド | テスト | チ... |
|-------------|-------|-------|------|
| CLAS_GLM1_5 | 15... | 15... | 自動 |
| CLAS_SVM1_5 | 15... | 15... | 自動 |
| CLAS_DT_1_5 | 15... | 15... | 自動 |
| CLAS_NB1_5 | 15... | 15... | 自動 |

A semi-transparent grey box is overlaid at the bottom of the screenshot, containing the following text in Japanese:

スコアリングに使用しないモデルを
スコアリング対象から外す

参考 予測モデル(教師あり学習)の確認

スライド 43

Oracle SQL Developer : dmuser/O2Oマーケティング/ターゲット顧客予測/分類構築/CLAS_DT_1_5

ファイル(E) 編集(E) 表示(V) ナビゲート(N) 実行(R) ダイアグラム(I) チーム(M) ツール(I) Window ヘルプ(H)

CLAS_DT_1_5... サムネイル ターゲット顧客予測 分類構築 データの参照 OUTPUT CLAS_DT_1_5 コンポーネント

Tree 設定

100% 最大ターゲット値: 2 ルールの保存...

ノード: 1
予測: No
サポート: 358 (58.59%)
信頼度: 53.35%
No: 191 (53.35%)
Yes: 167 (46.65%)
分割: CHECKING_AMOUNT

ノード: 2
予測: Yes
サポート: 211 (34.53%)
信頼度: 64.93%
No: 74 (35.07%)
Yes: 137 (64.93%)
分割: CREDIT_BALANCE

ノード: 3
予測: No
サポート: 253 (41.41%)
信頼度: 100%
No: 253 (100.00%)
Yes: 0 (0.00%)

ノード: 6
予測: No
サポート: 147 (24.06%)
信頼度: 79.59%
No: 117 (79.59%)
Yes: 30 (20.41%)

ルール サロゲート ターゲット値

ノード ルール: 折返し

If BANK_FUNDS > 225.5
Then No

dmuser/O2Oマーケティング/ターゲット顧客予測

作成されたモデルの詳細を確認
※ディジションツリー

Step7 スコアリング結果を表に格納

Oracle Data Miner: dmuser/O2Oマーケティング/ターゲット顧客予測

メニュー: ファイル(E) 編集(E) 表示(V) ナビゲート(N) 実行(R) ダイアグラム(I) チーム(M) ツール(I) Window ヘルプ(H)

ワークスペース: 100%

ワークフロー:

- データの参照 (Data Reference)
- INSUR_CUST_LTV_SAMPLE (Source Table)
- 分類構築 (Classification)
- 適用 (Apply) - 設定済み
- OUTPUT (Output Table)
- INSUR_CUST_LTV_APPLY (Target Table)

Callouts:

- モデルとスコアリング対象の表を結びつける「適用」ノード
- 格納先の表を設定
- スコアリング対象の表を設定

出力適用列 (Output Apply Columns):

| 列 | ファンクション | パラメータ |
|------------|---------|----------|
| CLAS_DT... | 予測 | |
| CLAS_DT... | 予測確率 | 予測: <可能↑ |
| CLAS_DT... | 予測コスト | 予測: <可能↑ |

Step8 出力結果の編集(1/2)

The screenshot shows the Oracle SQL Developer interface with a workflow diagram. The diagram consists of three nodes: 'データの参照' (Data Reference), '分類構築' (Classification Structure), and '適用' (Apply). The '適用' node is selected, and a context menu is open over it. The menu options include '接続(N)', '実行(B)', '実行の強制(D)', '編集(E)...', 'データの表示(E)', 'モデルの表示(Q)', 'テスト結果の表示(R)', '適用チェーンの生成(G)', 'コピー(C)', '貼り付け(P)', '拡張貼り付け(X)...', 'すべて選択(A)', 'パラレル問合せ(K)...', '画像をクリップボードにコピー(L)', '画像を別名で保存(M)...', 'プロパティに移動(O)', and 'ナビゲート(V)'. The '適用' node is highlighted with a blue border, and the '編集(E)...' option is highlighted in the menu.

「適用」ノードを右クリックし、「編集」を選択

Step8 出力結果の編集(2/2)

スライド 37

Oracle SQL Developer : dmuser/O2Oマーケティング/ターゲット顧客予測

ファイル(E) 編集(E) 表示(V) ナビゲート(N) 実行(R) ダイアグラム(I) チーム(M) ツール(I) Window ヘルプ(H)

適用ノードの編集

予測 追加出力

出力データ列

列

「出力データ列の編集」ダイアログ

| 使用可能な属性 | | 選択された属性 | |
|------------------|----------|-------------|----------|
| 名前 | データ型 | 名前 | データ型 |
| N_TRANS_ATM | NUMBER | CUSTOMER_ID | VARCHAR2 |
| N_TRANS_KIOSK | NUMBER | STATE | VARCHAR2 |
| N_TRANS_TELL... | NUMBER | | |
| N_TRANS_WEB... | NUMBER | | |
| PROFESSION | VARCHAR2 | | |
| REGION | VARCHAR2 | | |
| SALARY | NUMBER | | |
| SEX | VARCHAR2 | | |
| TIME_AS_CUSTO... | NUMBER | | |
| T_AMOUNT_AUT... | NUMBER | | |

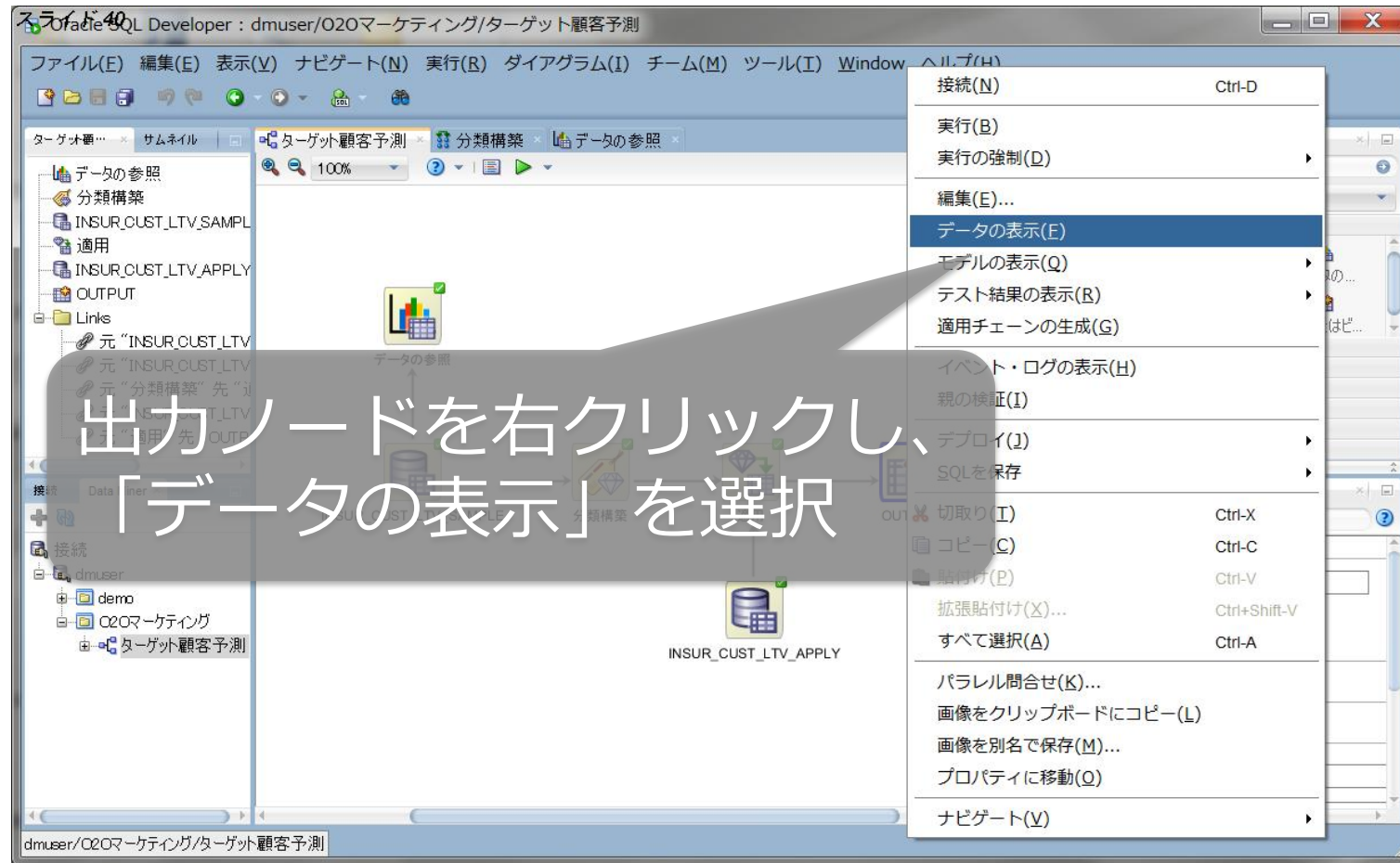
追加で出力したい列を選択

OK 取消

OK 取消

dmuser/O2Oマーケティング/ターゲット顧客予測

Step9 出力結果の表示(1/2)



Step9 出力結果の表示(2/2)

Oracle SQL Developer : dmuser/O2Oマーケティング/ターゲット顧客予測/OUTPUT

ファイル(E) 編集(E) 表示(V) ナビゲート(N) 実行(R) チーム(M) ツール(I) Window ヘルプ(H)

OUTPUT - 構... サムネイル

ターゲット顧客予測 分類構築 データの参照 OUTPUT

コンポーネント

データ列 | SQL

表示(Y) [実際のデータ] ソート(S) 並列問合せオフ... フィルタ(F) WHERE句を入力

| | CLAS_DT_1_5_PRED | CLAS_DT_1_5_PROB | CLAS_DT_1_5_PCST | CUSTOMER_ID | STATE |
|----|------------------|--------------------|--------------------|-------------|-------|
| 1 | Nb | 0.7959183673469388 | 0.746669925455212 | CU12406 | NY |
| 2 | Nb | 0.7959183673469388 | 0.746669925455212 | CU1141 | NY |
| 3 | Nb | 0.7959183673469388 | 0.746669925455212 | CU6623 | NY |
| 4 | Nb | 1.0 | 0.0 | CU10020 | NY |
| 5 | Nb | 0.7959183673469388 | 0.746669925455212 | CU2708 | MI |
| 6 | Nb | 1.0 | 0.0 | CU885 | CA |
| 7 | Yes | 0.6938775510204082 | 0.4212630998345284 | CU8239 | CA |
| 8 | Nb | 0.7959183673469388 | 0.746669925455212 | CU8441 | CA |
| 9 | Yes | 0.6938775510204082 | 0.4212630998345284 | CU7412 | CA |
| 10 | Yes | 0.6938775510204082 | 0.4212630998345284 | CU8147 | MI |
| 11 | Yes | 0.6938775510204082 | 0.4212630998345284 | CU14708 | NY |
| 12 | Nb | 0.7959183673469388 | 0.746669925455212 | CU8461 | NV |
| 13 | Nb | 0.9333333333333333 | 0.2439121756487026 | CU8520 | CA |
| 14 | Nb | 1.0 | 0.0 | CU4009 | CA |
| 15 | Nb | 1.0 | 0.0 | CU8587 | MI |
| 16 | Nb | 0.7959183673469388 | 0.746669925455212 | CU6758 | NY |
| 17 | Yes | 0.6938775510204082 | 0.4212630998345284 | CU14746 | NY |
| 18 | Nb | 1.0 | 0.0 | CU13188 | NC |
| 19 | Nb | 0.7959183673469388 | 0.746669925455212 | CU14965 | CA |
| 20 | Nb | 1.0 | 0.0 | CU15852 | CA |
| 21 | Nb | 1.0 | 0.0 | CU6275 | MI |
| 22 | Nb | 1.0 | 0.0 | CU12404 | CA |
| 23 | Nb | 0.7959183673469388 | 0.746669925455212 | CU14487 | CA |
| 24 | Nb | 1.0 | 0.0 | CU9791 | DC |
| 25 | Yes | 0.6938775510204082 | 0.4212630998345284 | CU18696 | CA |

接続 Data Miner

接続 dmuser demo O2Oマーケティング ターゲット顧客予測

dmuser/O2Oマーケティング/ターゲット顧客予測

予測値、予測精度、追加で出力設定した、CUSTOMER_IDとSTATE列が確認できる

Step10 出力結果のグラフ表示(1/3)

The screenshot shows the Oracle Data Miner interface with a workflow diagram. The workflow consists of the following nodes: 'INSUR_CUST_LTV_SAMPLE' (Data Source) -> '分類構築' (Classification) -> '適用' (Apply) -> 'OUTPUT' (Output). A 'データの参照' (Data Reference) node is connected to the '適用' node. A 'グラフ' (Graph) node is connected to the 'OUTPUT' node. The 'グラフ' node is highlighted with a blue border. A callout box points to the 'グラフ' node with the text: 「グラフ」ノードを出カノードと結びつける (Connect the 'Graph' node to the output node).

Step10 出力結果のグラフ表示(2/3)

Oracle SQL Developer : dmuser/O2Oマーケティング/ターゲット顧客予測/グラフ

ファイル(E) 編集(E) 表示(V) ナビゲート(N) 実行(R) チーム(M) ツール(I) Window ヘルプ(H)

グラフ - 構造

ターゲット顧客予測 分類構築 データの参照 OUTPUT CLAS_DT_1_5 グラフ

表示 実際の

新規グラフ

Line Scat... Bar Hist... Box

Title: グラフ1

Comment:

棒グラフの設定

X軸: STATE 設定...

Y軸: CLAS_DT_1_5_PCST

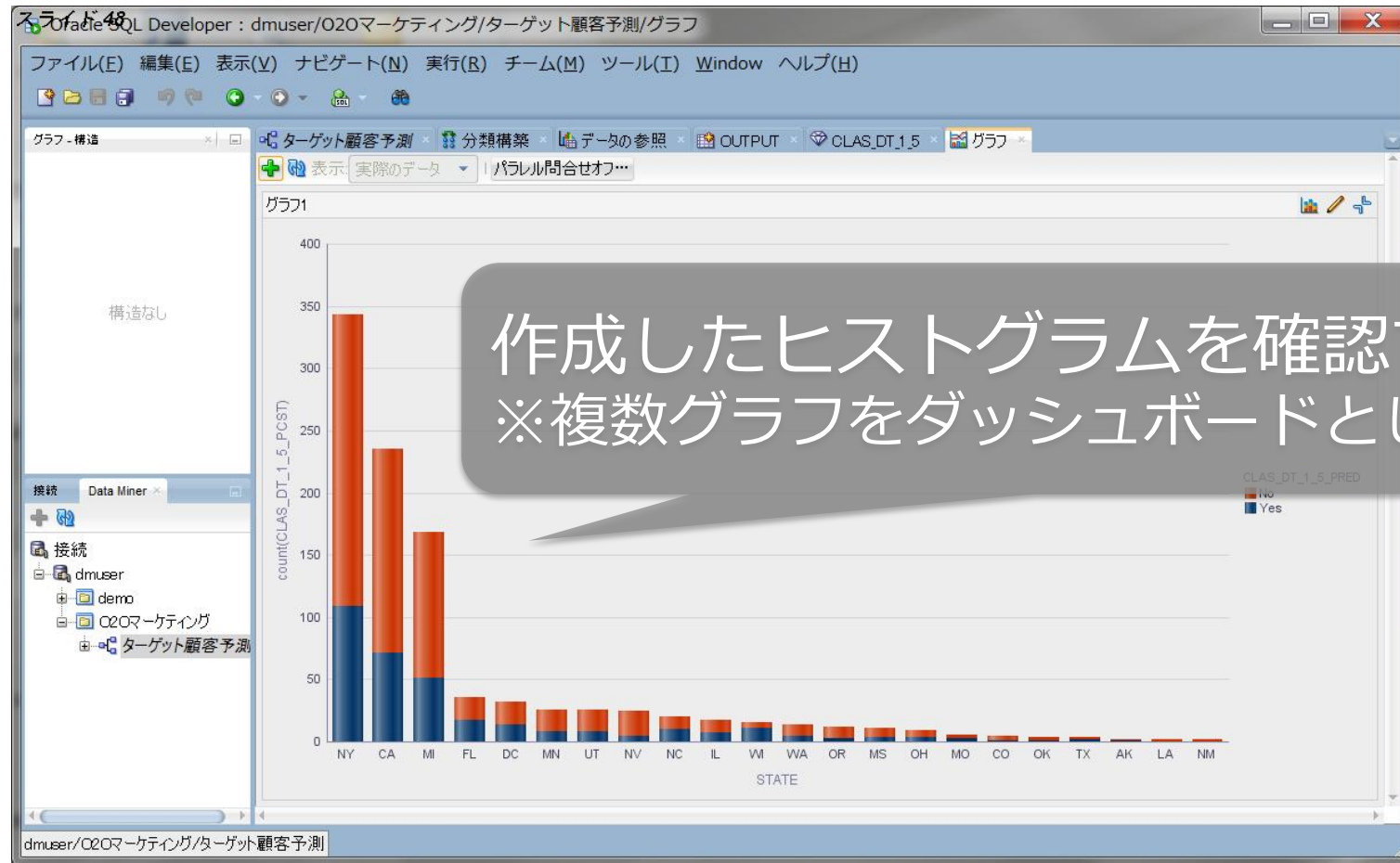
統計: 件数

グループ化基準

属性: CLAS_DT_1_5_PCST 設定

折れ線、散布図、棒グラフ、ヒストグラム、ひげ付き棒グラフからグラフを選択し、属性を埋めていく

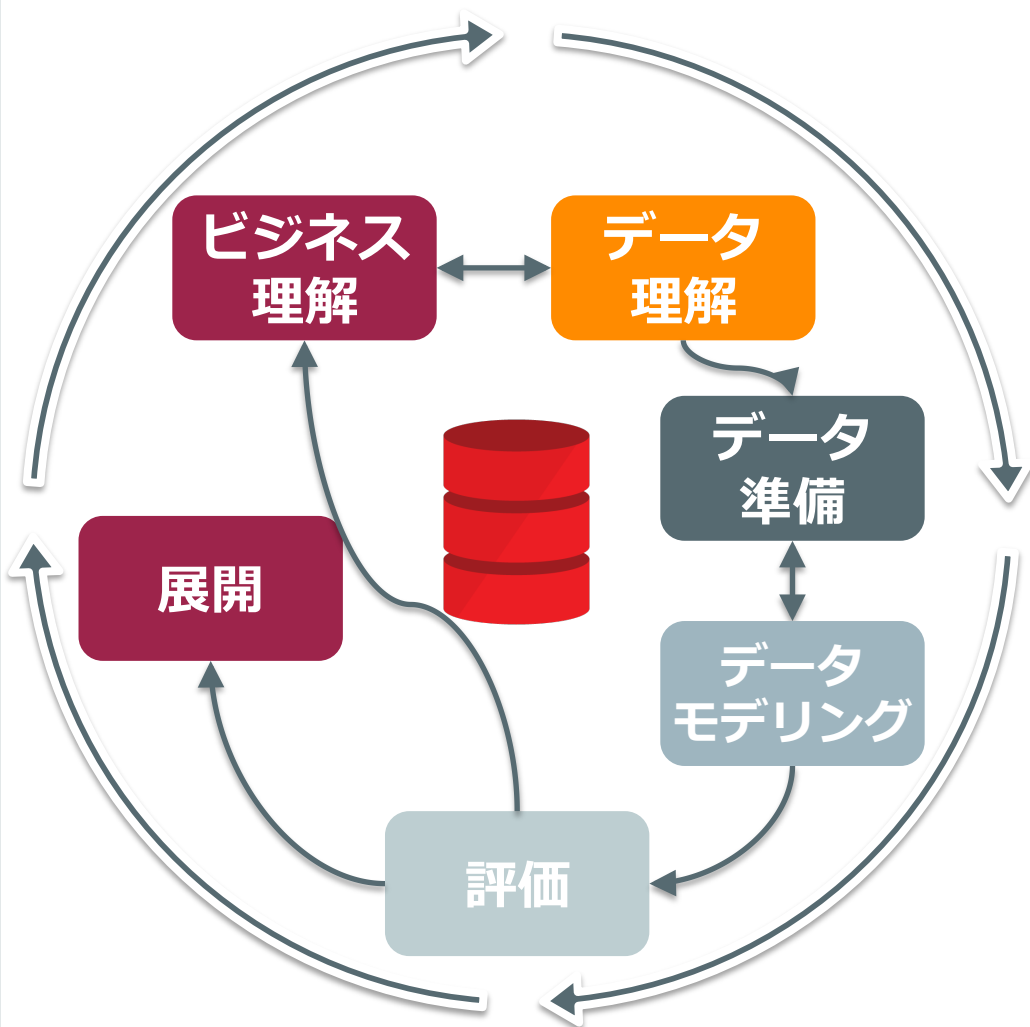
Step10 出力結果のグラフ表示(3/3)



アジェンダ

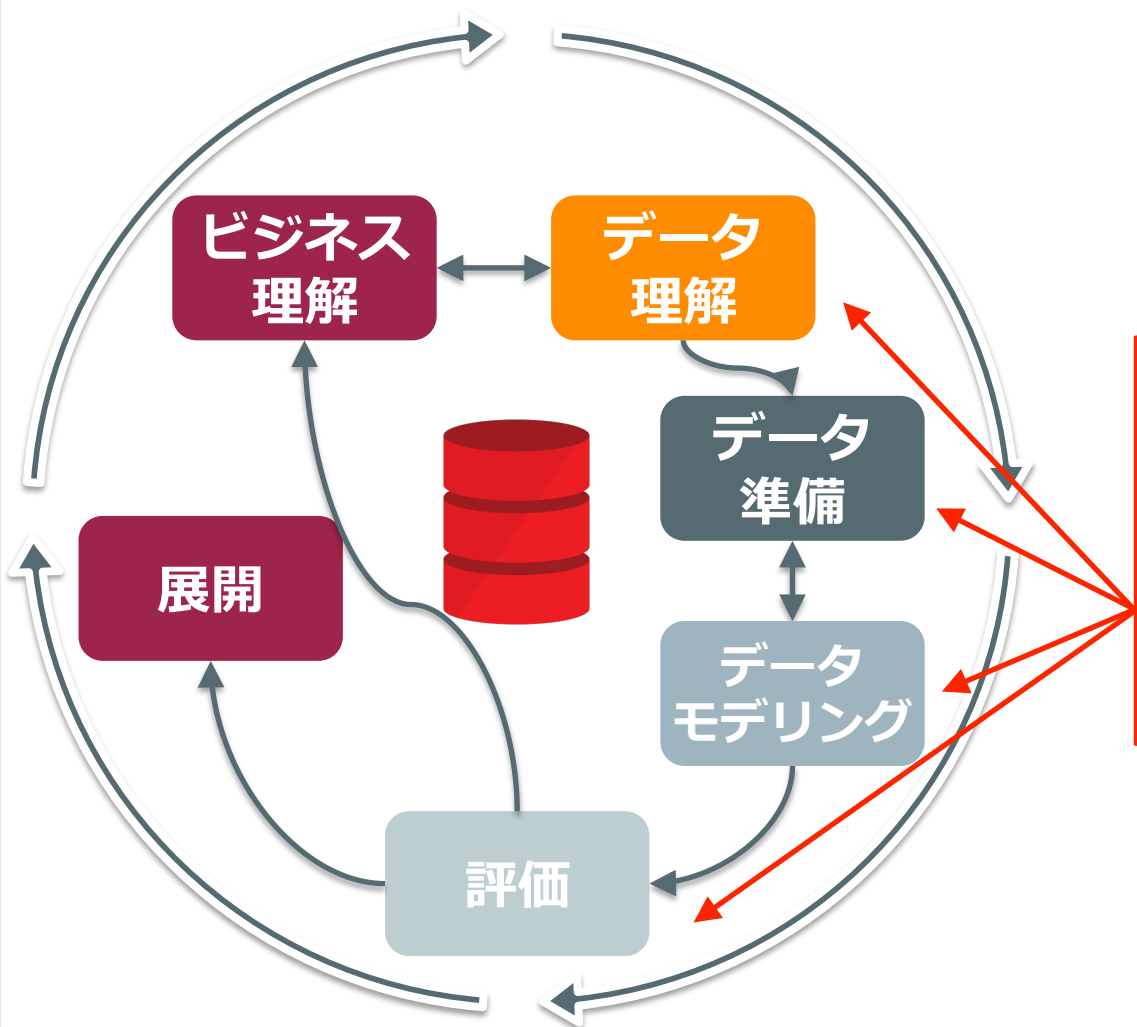
- 1 マシン・ラーニング（機械学習）概要
- 2 Oracle Advanced Analytics の紹介
- 3 クラウド・マシン・ラーニング構築方法
- 4 Oracle Data Miner で設計する予測分析
- 5 まとめ

データ活用におけるPDCAサイクル (CRISP-DM)



| ビジネス理解 | データ理解 | データ準備 | データモデリング | 評価 | 展開 |
|--------------|-------------|----------|--------------|-------------|-------------|
| ビジネス目的の決定する | 初期データを収集する | データを選択する | モデリング手法を選択する | 結果を評価する | 展開を計画する |
| 状況を評価する | データを記述する | データを整理する | テスト結果を生成する | プロセスを審査する | 運用を計画する |
| 分析のゴールを決定する | データを調べる | データを構築する | モデルを構築する | 次のステップを決定する | 最終レポートを作成する |
| プロジェクト計画を建てる | データの品質を検証する | データを統合する | モデルを評価する | | プロジェクトを審査する |
| | | データを設定する | | | |

データ活用におけるPDCAサイクル (CRISP-DM)



データ分析におけるPDCAサイクルを
クラウド・マシン・ラーニング
によって自動化・サポート



クラウド・テクノロジーを語ろう
Oracle Cloud Developers

第1回 Meetup 2016年1月29日 19:00~
@オラクル青山センター

参加登録はこちら : <http://ora.cl/l8s>

```
var community = React.createClass({
  init : function(){
    return {
      date : "2016-1-29",
      location : "OAC",
      goal : [
        "Learn",
        "Connect",
        "Have Fun"
      ]
    };
  }
});
```

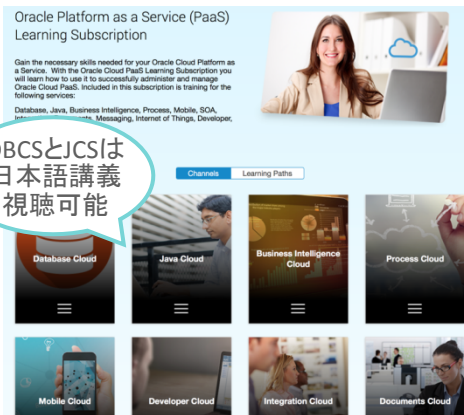


Oracle Cloud ラーニング・サブスクリプション

Oracle Cloud (PaaS) の導入から運用まで、クラウドを活用するために必要なスキルを身につけた "Cloud Ready" なエンジニアを育成するための学習プラットフォーム

- Oracle Cloud Service の活用に必要なスキルを学習できるビデオ・トレーニング
- 製品アップデートに応じて最新のコンテンツに更新
- 1年間のサブスクリプション形式

Oracle Platform as a Service ラーニング・サブスクリプション



多彩な Oracle PaaS の活用方法をトータルにカバー。すべての Oracle PaaS 技術者にオススメです！

学習内容:

- Oracle Cloud Platform as a Service の使用方法
- Oracle PaaS の多様なサービスを活用する利点
- Oracle PaaS の運用管理
- Oracle PaaS を使ったアプリケーションの開発とデプロイ
- 既存のアプリケーションの Oracle PaaS への移行 など

【対応サービス】

Database Cloud, Java Cloud, Business Intelligence Cloud, Process Cloud, Mobile Cloud, Integration Cloud, Documents Cloud, Messaging Cloud, Internet Of Things Cloud, SOA Cloud, Database Backup Cloud, Developer Cloud

定価: 116,856 円 (税込)

50% Off

特別価格: **58,428** 円 (税込)

1ユーザー/1年間利用可能

【ご注意】 Oracle Platform as a Service (PaaS) ラーニング・サブスクリプションの最小購入ユーザー数は 5 です。本特別価格は、2015 年 12 月 31 日までにご購入される方に対して適用されます。また、他の割引契約、またはキャンペーンと併用することはできません。

ただいま 無償体験版公開中！

Oracle Cloud インスタンスの作成やクラウド上の Oracle Database, WebLogic Server の起動方法など、技術者が円滑に Oracle Cloud (PaaS) をはじめるためのポイントを学習できる『Getting Started』を視聴可能

アクセスはこちらから

education.oracle.co.jp/cls_paas

オラクルユニバーシティ
お問い合わせ窓口



TEL 0120-155-092

URL <http://www.oracle.com/jp/education/>



Integrated Cloud

Applications & Platform Services

ORACLE®