Scatter Graphs

Scatter graphs enable you to interpret the relationship between multiple variables, if any of the variables are a good predictor of another, or if the variables change on their own. You can make scatter graphs go even further by adding clusters or trend lines.

Visualization Type	More Information
Category	Shows you a set of vertices (or nodes) connected by links called edges (or arcs), which can also have associated directions.
Scatter	Uses dots to represent values for two numeric variables where the position of each dot on the horizontal and vertical axis indicates values for an individual data point. Scatter plots are good to use if you want to see the relationships between variables.
Stacked Category	Based on the category graph where values are stacked by category.

Enhance Data in Visualizations with Advanced Analytics Functions

Advanced analytics are statistical functions that you apply to enhance the data displayed in visualizations.

The Analytics area in the Data panel contains standard analytics functions (for example, Clusters and Trend Line). You can use analytics functions as they are, or use them to create your own calculated columns that reference statistical scripts.

Topics:

About Advanced Analytics Functions

You can use advanced analytics functions, like forecast, cluster, and outliers, in charts to achieve better insights into your data.

46

Forecast

The forecast function uses linear regression to predict future values based on existing values along a linear trend.

You can set a number of time periods in the future for which you want to predict the value, based on your existing time series data. See Add Advanced Analytics Functions to Visualizations.

Oracle supports these forecast model types:

- **Auto-Regressive Integrated Moving Average (ARIMA)** - Use if your past time series data is nonseasonal but provides enough observations (at least 50, but preferably more than 100 observations) to explain and project the future.

- **Seasonal ARIMA** - Use if your data has a regular pattern of changes that repeat over time periods. For example, seasonality in monthly data might be when high values occur during summer months and low values occur during winter months.
- **Exponential Triple Smoothing (ETS)** - Use to analyze repetitive time series data that doesn't have a clear pattern. This model type produces an exponential moving average that takes into account the tendency of data to repeat itself in intervals over time.

Create a custom calculation using the `FORECAST` function to have more control over settings, or if you want to use the forecast in other visualizations. See Time Series Functions.

Clusters

46

The cluster function groups a set of objects in such a way that objects in the same group show more coherence and proximity to each other than to objects in other groups. For example, you can use colors in a scatter chart to show clusters of different groups. See Add Advanced Analytics Functions to Visualizations.

- **K-means clustering** - Use to partition "n" observations into "k" clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster.
- **Hierarchical clustering** - Use to create a hierarchy of clusters built using either an agglomerative (bottom-up) approach, or a divisive (top-down) approach.

Create a custom calculation using the `CLUSTER` function to have more control over settings, or if you want to use the cluster in other visualizations. See Analytics Functions.

Outliers

46

The outliers function displays data records that are located the furthest away from the average expectation of individual values. For example, extreme values that deviate the most from other observations fall into this category. Outliers can indicate variability in measurement, experimental errors, or a novelty. If you add outliers to a chart that already has clusters, then the outliers are depicted as different shapes.

Outliers can use K-means clustering or hierarchical clustering. See Add Advanced Analytics Functions to Visualizations.

Create a custom calculation using the `OUTLIER` function to have more control over settings, or if you want to use the outlier in other visualizations. See Analytics Functions.

Reference Lines

46

The reference lines function defines horizontal or vertical lines in a chart that correspond to the X-axis or Y-axis values. See Add Reference Lines to Visualizations.

- **Line** - You can choose to compute the line between average, minimum, or maximum. For example, in the airline industry, if passenger turnout is plotted against time, the reference line can show whether passenger turnout for a particular month is above or below average.
- **Band** - A band represents upper and lower range of data points. You can choose a custom option or a standard deviation function, and between average, maximum, and minimum. For example, if you're analyzing sales by month and you use a

custom reference band from average to maximum, you can identify months where sales are above average, but below the maximum.

Trend Lines

46

The trend line function indicates the general course of the metric in question. A trend line is a straight line connecting a number of points on a graph. A trend line helps you analyze the specific direction of a group of value sets in a visualization. See Add Advanced Analytics Functions to Visualizations.

- **Linear** - Use with linear data. Your data is linear if the pattern in its data points resembles a line. A linear trend line shows that your metric is increasing or decreasing at a steady rate.
- **Polynomial** - Use this curved line when data fluctuates. It's useful, for example, for analyzing gains and losses over a large dataset.
- **Exponential** - Use this curved line when data values rise or fall at increasingly higher rates. You can't create an exponential trend line if your data contains zero or negative values.

Use the `TRENDLINE` function to create custom calculations, or if you want to use the trend line in other visualizations. See Analytics Functions.

Add Advanced Analytics Functions to Visualizations

You can apply advanced analytics functions to your workbook's visualizations.

1. On the Home page, select a workbook, click the **Actions menu**, and then select **Open**.
2. Click the **Analytics** icon in the Data panel.
3. Drag and drop an advanced analytic function from the **Analytics** pane to a visualization.

Add Reference Lines to Visualizations

You can use advanced analytics reference lines to identify the range of data element values in a visualization.

1. On the Home page, select a workbook, click the **Actions menu**, and then select **Open**.
2. In the Data Panel, click the **Analytics** icon .
3. Drag and drop **Reference Line** into a visualization. Alternatively, you can double-click **Reference Line** to add it to the selected visualization.
4. In the properties pane select the **Analytics** tab.
5. In the **Method** row, click the value to select **Line** or **Band**.
6. Click the current reference **Function** and select the value that you want to use.
   - When you choose the **Line** method, select a reference function, and enter appropriate values if required.
   - When you choose the **Band** method, select one or more reference functions, and enter appropriate values if required.
7. Click **Save**.

The Residuals dataset is outputted when you select these model and algorithm combinations.

Model	Algorithms
Numerics	Linear Regression Elastic Net Linear Regression CART for Numeric Prediction
Binary Classification	CART (Decision Tree)
Multi Classificatin	CART (Decision Tree)

Statistics

46

This dataset's metrics depend upon the algorithm used to generate it. Note this list of metrics based on algorithm:

- Linear Regression, CART for Numeric Prediction, Elastic Net Linear Regression - These algorithms contain R-Square, R-Square Adjusted, Mean Absolute Error(MAE), Mean Squared Error(MSE), Relative Absolute Error(RAE), Related Squared Error(RSE), Root Mean Squared Error(RMSE).
- CART(Classification And Regression Trees), Naive Bayes Classification, Neural Network, Support Vector Machine(SVM), Random Forest, Logistic Regression - These algorithms contain Accuracy, Total F1.

This dataset is outputted when you select these model and algorithm combinations.

Model	Algorithm
Numeric	Linear Regression Elastic Net Linear Regression CART for Numeric Prediction
Binary Classification	Logistics Regression CART (Decision Tree) Naive Bayes Neural Network Random Forest Support Vector Machine
Multi Classification	Naive Bayes Neural Network Random Forest Support Vector Machine

Summary

This dataset contains information such as Target name and Model name.

The Summary dataset is outputted when you select these model and algorithm combinations.

How Can I Use Oracle Machine Learning Models in Oracle Analytics?

Oracle Analytics allows you to register and use Oracle machine leaning models from Oracle Database or Oracle Autonomous Data Warehouse.

Using Oracle machine learning models with Oracle Analytics greatly increases the level of predictive analytics that you can perform on datasets because the data and the model reside in the database, the data scoring is performed in the database, and the resulting dataset is stored in the database. This allows you to use the Oracle machine learning execution engine to score large datasets.

You can register and use Oracle machine learning models from these database data sources:

- Oracle Autonomous Data Warehouse
- Oracle Database

46

In Oracle Analytics you can register any of the database's Oracle machine learning models in the mining classes Classification, Regression, Clustering, Anomaly, or Feature Extraction that were created using the Oracle Machine Learning for SQL API (OML 4SQL). Your Oracle Analytics user role and database permissions determine the Oracle machine learning models that are available for you to register and use.

You can also create predictive models in Oracle Analytics. See What Are Oracle Analytics Predictive Models?

Register Oracle Machine Learning Models in Oracle Analytics

The Oracle machine learning models must be registered in Oracle Analytics before you can use them to predict data. You can register and use models that reside in your Oracle Database or Oracle Autonomous Data Warehouse data sources.

1. On the Home page, click **Page Menu**, and then click **Register ML Model**.

   This option is available for users with the BI Service Administrator or DV Content Author role.

2. In the Register an ML Model dialog, select a connection.

   In the Select a Model to Register dialog, you'll see the database's Oracle machine learning models in the mining classes Classification, Regression, Clustering, Anomaly, or Feature Extraction that were created using the Oracle Machine Learning for SQL API (OML 4SQL).
   If needed, click **Create Connection** to create a connection to the Oracle Database or Oracle Autonomous Data Warehouse data source containing the Oracle machine learning model that you want to use.

3. In the Select a Model to Register dialog, click the model that you want to register and review the model's information. For example, the model class and algorithm used to build the model, the target the model predicts, the columns the model is trained on, model predictions, and parameters.

4. Click **Register**.

5. From the Home page, click **Navigator**, and then click **Machine Learning** to confirm that the model was imported successfully.