

Insights in visualizations for two continuous fields

Last Updated: 2024-02-29

Insights for two continuous fields are available when a visualization involves two continuous fields and an optional categorical group or points field.

Overview [↗](#)

Use visualizations such as scatter plot for the two continuous fields, possibly sliced by categories of the group field. The main goal is to detect any relationship between the continuous field and include the categorical group field as well. The results contain predictive strength of discovered relationship, the relationship description provided by fit lines, and any points with large discrepancy from the fit lines as meaningful differences.

Algorithms [↗](#)

12.8

IBM® Cognos Analytics computes multiple regression models that involve one of the continuous fields as the response, and the other continuous fields as an explanatory field. The optional categorical group field is used as a model factor. In addition to additive model contributions corresponding to the explanatory field, Cognos Analytics considers the square of the explanatory field and any interaction terms that include a factor. A regression model that provides an optimal fit for the data is selected from a number of possible models. The corresponding fit line is derived from a linear or quadratic model. In the case where an optional categorical group field has been supplied, it can produce a different line or quadratic curve for each category of the factor. A factor with up to three categories is currently considered in order not to overload the visualization.

Each point in a visualization represents a number of rows in the data and it is defined by the **Points** field. Corresponding row counts that are based on the response field define frequency weights that are used for building the regression models. Regression weights are used independently of frequency weight when Cognos Analytics computes the regression models.

Details [↗](#)

Two continuous fields

When Cognos Analytics applies multiple linear regression for two continuous fields, one is chosen as the response and the other as an explanatory field in the model. Cognos Analytics considers both linear and quadratic model terms. If the quadratic model is significant based on the F test and its relative predictive strength improvement is more than 10% over the linear model, Cognos Analytics reports its predictive strength and displays the quadratic curve based on the computed model. This curve displays predicted values of the response based on the corresponding values of

the explanatory field. Otherwise, linear predictor model is considered. If it is significant and its predictive strength is larger than 10%, Cognos Analytics reports its predictive strength and displays a line representing the predicted values of the response field based on the corresponding explanatory values. If the linear model does not qualify, the mean is reported as the fit line and no relationship is reported between the two continuous fields.

When linear or quadratic relationship is detected, Cognos Analytics also inspects the differences between predicted and observed values of the response field. These differences are called residuals and Cognos Analytics conducts studentized residuals test to detect outliers. Points with large departure from the discovered relationships are displayed under meaningful differences in the corresponding chart.

Restrictions

The following table describes the conditions that determine whether insights are suggested for this algorithm.

Response	Explanatory	Group	Weight	Points	Insight
Exactly 1	Exactly 1	N/A	Optional	Optional	Predictive strength
Summarization level = any	Continuous		Continuous	Any	Fit line
Continuous					Meaningful differences

Categorical group field

When a categorical group field is specified in addition to two continuous fields, it is used as a factor in the multiple linear regression where one of the two continuous fields is chosen as the response field and the other as an explanatory field. Cognos Analytics considers linear and quadratic model terms for continuous explanatory combined with contributions from the factor. If the quadratic model or linear model that include the factor is significant based on the F test and its relative predictive strength improvement is more than 10% over the linear model with continuous explanatory only, Cognos Analytics generates four extra models. These models include all possible interactions of continuous explanatory and factor. A model with the maximum adjusted R-squared that is also significant is selected as the final model. It is used to create a fit line for each category of the categorical predictor. Otherwise, the linear model with continuous explanatory is tested for significance and reported if its predictive strength is greater than 10%. If the linear model does not qualify, no reliable relationship among fields is established and the overall mean is reported as the fit line.

When a reliable relationship is detected, Cognos Analytics also checks for difference between the predicted and observed values of the response field. Cognos Analytics conducts studentized residuals test to detect outliers and display them under meaningful differences in the corresponding chart.

Restrictions

The following table describes the conditions that determine whether insights are suggested for this algorithm.

Response	Explanatory	Group	Weight	Points	Insight
Exactly 1	Exactly 1	Exactly 1	Optional	Optional	Predictive strength
Summarization level = any	Continuous	Categorical	Any	Any	Fit line
Continuous					Meaningful differences

Regression weights field

An optional continuous field can be used to specify regression weights for the model. Regression weight for an available value corresponds to influence of the observation on the computed model parameters.

Parent topic:

→ [Insights in visualizations](#)