Rapid Insight® | Simplifying data analysis

**Rapid Insight® Analytics Training Session:**

*The objective of this training document is to quickly get a user up to speed on the basic functionality of Rapid Insight® Analytics.*

*To aid in the learning the software, there are green buttons that prompt the user on what to do next. If a button is green, Rapid Insight is telling you that you should probably push that button to move forward with the analysis. To start out, follow the instructions below which explain how to create an analysis. Below that, each of the tabs and major features are defined.*

**CREATE AN ANALYSIS:**

An analysis can be opened in three simple steps:

1) *Click on the 'open dataset' button.* This will bring up the 'select a dataset' window. In here you need to click on the tab on the left for the type of file you would like to open. Next, select the file or database table that you would like to open.

2) *Next, click on the green button that says 'load variables'.* This will actually read in the data. Now you're ready to begin by creating an analysis.

3) *Click on the green button that says 'create analysis'.* Now you're ready to begin data mining and analyzing your data. Below is an explanation of each tab and each type of analysis you can do.

---

**VARIABLE STATISTICS TAB:**

Click on the variables statistics tab. In here we can see the statistics for each of our numeric variables in the dataset. Rapid Insight automatically calculates the mean, standard deviation, minimum value, maximum value, coefficient of variation, number of missings, number of observations, and the range of each variable.

---

**ANALYZE DATA TAB:**

In this tab you can get a view of the entire dataset. If you don't want to view all of the variables at once, select the 'clear all' button, and then click on each of the variables that you'd like to view.

> **Sort:** To sort the data by one or more columns, click on those columns (within the body of the column) and click the up or down arrow to sort in either direction. To sort by more than one variable, put the columns next to each other by clicking on the header for a column, holding the left mouse button down, moving the column next to the other column you'd like to sort by, and letting go of your mouse button. Highlight two or more columns and click one of the sort buttons. The data will be sorted first by the left-most columns first.
> **Aggregate:** Click on a column that you'd like to aggregate by (click within the body of the column). Next click on either the 'take means by' or the 'take sums

by' buttons. If the selected variable is categorical, it will take the means for each of the other variables, for each of the selected variable's categories. If the chosen variable is continuous, it will decile by the chosen variable and take the means for each decile for each of the other variables.

---

**UNIVARIATE TAB:**
In this tab, you can view the distribution of any variable in your dataset. Click on a variable and it will show you the distribution of that variable's values. If the variable is categorical, it will show you the counts/percentages for each of the categories of that variable. If the variable is continuous numeric, the software will create bins by taking the variables highest and lowest value and splitting the data into ranges based on the high/low values.

---

**REPORTING:**
Any graphs, charts, or analyses can be captured and put into a PowerPoint presentation or into a Rapid Insight® Report. To do this, click on 'reports' in the menu at the top of the screen. Then click on the first choice to 'launch report bar'. Click on any chart, graph, or analysis, and drag it into this report bar. Once you've collected all of the graphs of interest, export them into report by selecting 'report' at the bottom of the report bar (this is selected by default). This will put all of the collected graphs into the report contained in the 'report' tab. The other option is to put them into a PowerPoint presentation. To do this, click the 'ppoint' option at the bottom of the report bar. Click export and when prompted, name the PowerPoint presentation. Note that you must have Microsoft PowerPoint installed on your computer. The report bar should always be closed after using it. Do this by clicking on the red 'X' in the upper right hand corner of the report bar.

---

**MULTIVARIATE TAB:**
In this tab, you can view the relationships between any two or three variables. To do this, select a y variable and an x variable in the respective dropdown lists. The y variable must be continuous or binary (those are the only variables that will be shown in the y list). The x variable can be continuous, binary, or categorical. View relationships between any variables. By default RI will show the mean of the y variable for each category of the x variable (for categorical x variables). If the x variable is continuous, RI will sort and decile by the x variable, and show the mean of the x and the y for each decile. Use the report bar (explained above) to capture and report on any relationships. Be sure to close the report bar after using.

---

**CONTROL PANEL TAB:**
In the control panel tab, we can do a number of things:

1) **Choose a Y variable**. To use the automated mining tab and to build models, you'll need to choose a y variable. To do this, highlight a variable and then click on the down arrow toward the bottom left of the screen. You can also choose a y variable by simply double clicking on it.
2) **Set filters.** Click on any variable and then click the 'set' button in the last column.
3) **Change Variable Types.** When Rapid Insight reads in the data, it guesses on the variable type for each field. For example, if a variable is text and has more than 20 different unique values, Rapid Insight assumes that the variable is not a categorical variable. Pure text fields that are not marked as 'categorical', are not available for analysis. To change a 'text' field to 'categorical', simply click on that variable in the 'type' column and change its type using the drop down menu.
4) **Change model availability.** By default, each variable will be available in a model if it is found to be related to the y variable. You can change a variable's model availability by clicking on the 'model availability' column for that variable, and changing the availability to either 'available' or 'exclude'.

---

**AUTOMATED MINING TAB:**
In the automated mining tab, the software will automatically find relationships between variables. In order to do this, a Y variable must be chosen in the control panel (the first tab). The Y variable is the variable that you will be analyzing. When entering the automated mining tab, you will be shown a list of all variables available for analysis. Click on the button that says 'automine it'. Rapid Insight will then analyze all of the variables to see which ones are statistically related to the y variable. To view these relationships, click on the button that says 'view relationships'. This will bring up a window similar to the multivariate analysis tab. At the top of the screen, you'll see a drop down box labeled 'related variables'. Open this drop down to choose the variable that you'd like to view the relationship for.

---

**CREATE VARIABLES TAB:**
In this tab, you can create variables. To create a variable, first select the type of variable you'd like to create. Next define the variable's definition. After this, name the variable in the box labeled 'new variable name'. Finally click the 'create this variable' tab. This variable will then be available to you everywhere in the analysis. See the help menu for documentation on how to create each variable type.

---

**FREQUENCY ANALYSIS TAB:**
In this tab, you can view the frequency distributions for any categorical variables.

1) *Single Variable Category Frequencies:* Just highlight the variables for which you'd like to view the frequency distributions for. Click the green right hand arrow, and then click the 'submit' button. The frequency distributions for the

first selected variable will be shown in the workspace below. To view any of the other variables that you have selected, click the '+' button to expand that variable.

2) *Two Way Cross Frequencies:* To view cross tabulations between variables, select the two-way cross frequencies option. Select any variables that you'd like to analyze, and then choose one variable which you'd like to view these variables cross tabulated with. For each variable combination, RI will show four sets of numbers. First it will show the raw numbers, or the actual counts of the # of occurrences for each of the combinations of categories for the two variables. Next it will show the row percentages. Here it will show for each category of the row variable, the percentage breakdown for each category of the column variable (each row category adds to 100%). Next it will show the column percentages. Here it will show for each combination of the column variable, the percentage breakdown for each category of the row variable (each column category adds to 100%). Finally it will show the total percentages. Here it will show the percentage breakdown for each combination of the row and column variables. The frequency data can be put into a report or exported to Excel.

---

**MEANS ANALYSIS TAB:**
In this tab you can view the means of any chosen continuous or binary variables by any categorical or binary variable. To do this, select the variables that you'd like to examine the means of. Next select a categorical or binary variable for which you'd like to view the means of those variables by. Click submit, and you'll see the means analysis in the workspace below. The means analysis for the first selected variable will be shown. To view any of the other variables that you have selected, click the '+' button to expand that variable. The means data can be put into a report or exported to Excel.

---

**CORRELATION ANALYSIS TAB:**
In this tab, you can view the correlations between any variables. Click on as many variables as you'd like, and Rapid Insight will build a correlation matrix for the chosen variables. The positive correlations will be shown in green and the negative correlations will be shown in red. The correlation matrix can be exported to Excel.

---

**PROFILING ANALYSIS TAB:**
In this group, you can compare any two groups to see how the two groups different. For example, you might want to compare high value customers to low value customers. The only variables that are available to us in this tab are binary (zero – one) variables. So if you want to do a profile on high value vs. low value customers, we'll need to first go into the 'create variables' tab, and create a binary variable for high value/low value customers. In other words, we'll want to define high value customers as a one, and low value customers as a zero. When you enter the profile analysis tab, click on the variable

that you'd like to use to profile. You can then name the two groups (for example, we would name the high value and low value customer groups). Next click on the 'auto profile' button. Rapid Insight will then go and find all of the variables in the dataset for which these two groups are statistically different. It will then show each of the two groups, along with the averages for each variable for which these two groups are different. The profiles can be viewed graphically by clicking on the 'view profiles graphically' button. The correlation data can be put into a report or exported to Excel.

---

**MODEL TAB:**
In this tab, you can build statistical models. In order to start modeling, you need to have first chosen a Y variable. If you haven't done this yet, go back to the 'control panel' tab (first tab) and choose a Y variable. Once you have a Y variable chosen, you can click on the 'begin regression analysis' button. Once in the regression window, you can build a model automatically by clicking on the 'build model automatically'. This button will only work if you first went into the 'automated mining' tab, and ran automine on the y variable. Click this button and Rapid Insight will build a model for you automatically. You can also build a model by selecting the variables that you'd like in the model and then clicking the 'run regression' button.

---

**REPORTS TAB:**
In the reports tab, you can export any graphs and charts that you have collected in the report bar. Remember, the report bar can be launched from the main menu of the application, and any charts and graphs can be dragged and dropped into the report bar. To export them into the report tab, simply click export in the report bar (note that in the report bar the 'report' needs to be selected, and not 'ppoint'. This report can be changed and modified and then printed out or saved. By default, the report will be saved in rich text format. If you'd like to save as a word document, click on the 'Word Format' option in the upper right hand corner.

*For a more detailed explanation on any of these features, click on the help option in the main menu of the Rapid Insight® Analytics application.*