Quick introductory project.
(1) Log on to VCL following instructions.

(2) Double click Enterprise Miner icon (wait a minute or so for EM to open).

(3) Look at the properties panel at the left.
   - Right click on data sources and click on “Create Data Source” pop-up
   - Click “next”
   - Double click AAEM on left to open your folder then click on Airline table
   - Click OK.
   - Click “next” until you get to step 5
   - Hold down CTRL and click on Lair, Air and Date then click Explore
   - Select Actions->plot then select the line plot and next
   - Give AIR the role of Y and DATE the role of X (click in role white space) then select finish.
   - Discuss the plot and whether a regression approach would work well
   - Make a plot of Lair versus date. Which plot looks better – why?
   - Close the explore window.

(4) Mark AIR as rejected, Lair as target. Mark the “dummy variable” mon12 as rejected (why?).

(5) Click next until you get to the last window then click finish.

(6) Click Diagrams in the properties panel then right click and create diagram. Call it Airline.

(7) Drag the Airline DATA SOURCE into the diagram window.

(8) Above the diagram window click the Explore subtab then drag its “graph explore” icon into the diagram. This is now a “node.” Connect the Airline data node to the graph explore node.

(9) Right click the graph explore node and select run. Select “results” when it finishes and explore that window.

(10) Click the top right (plot) icon and plot Lair by t as you did before.

(11) Close the graphs and Graph Explore results window. Click the Model subtab.

(12) Drag a regression (third from right) node into your diagram and connect the Airline data source to it with an arrow.

(13) Run the Regression node. Click Results when it completes and investigate its contents.
(14) Close the results window.

(15) Drag another Graph Explore window into the diagram. Connect the regression node to the new Graph Explore node. As before, graph the predicted values and actual values by date, overlaid on the same graph. You will need to check the “allow multiple roles” checkbox to declare two variables to have the Y role for overlaying.

Detail:
To get the graph of means in 5% chunks, SAS does this
(1) sort by descending predicted values
(2) pick out the best (for example) 5% of predicted values and average them.
   Note: look at the output window and you will see it is an average of 8 observations then for a while then back to 8. \{8 7 7 7 8 7 7 7 8 7 7 7 etc.\} Now 144(0.05) = 7.2 and 5(7.2 )=36 = 8+7+7+7+7 so they are using integer numbers of points that sum correctly at certain intervals.
(3) average the target values associated with these predicted values. For example at depth 5, these would not necessarily be the highest 8 Lair values.