Data Set: Denim.jmp

Summary

This data set is the result of an experiment conducted by a company that manufactures blue jeans. Denim fabric naturally contains starch, creating stiffness in the fabric. Most customers find this stiffness uncomfortable, so denim manufacturers subject the fabric to a variety of washing treatments to remove some of the starch and make it feel "worn." Although the feel of the fabric is important, the company is also concerned about the strength of the treated fabric, which is measured by a count of the destroyed threads.

In this lab you will study which, if any, of the categorical factors is related to the severity of the thread wear of the denim. You'll copy output and summarize your findings in a report (discussion questions and required output are in italics below).

The Denim Data

Open the file **Denim.jmp** from the JMP **Sample Data** directory. The data set contains information on 98 samples of denim, including:

- The lot number of the fabric.
- The method used to treat the fabric.
- The size of the load washed (in lbs.).
- Whether or not the fabric was sandblasted.
- The thread wear (an actual count of broken threads between 1 and 10).
- The thread wear, categorized as low, moderate or severe.
- The resulting starch content (%).

Data Types

The size of the wash load, the thread wear measured and the starch content are numeric variables (they have the continuous modeling type in JMP). The remaining variables are all categorical (these have nominal or ordinal modeling types).

Check in the **Columns** panel to the left of the data table to make sure that **Method** is shown as a **nominal** variable 🔴 (values belong to categories but the order is not important) and **Thread Wear** is shown as an **ordinal** variable 📊 (numeric or character data - values belong to ordered categories).

Right-click on the icon and change the modeling type if necessary.

Is the method of aging the denim independent of thread wear severity? What do you think? Write a brief response to this question in your report. Include your reasons for believing that an association might or might not be present.

Exploring Categorical Data in JMP®

Go to **Analyze > Fit Y by X**. Select **Method** for the **X, Factor** and **Thread Wear** (not Thread Wear Measured) for the **Y, Response**.

The key in the bottom corner of the dialog window shows what type of analysis JMP will conduct.

For a categorical X, Factor and continuous Y, Response, JMP conducts a Oneway analysis.

In this case, we have two categorical variables. So, JMP conducts a **Contingency** analysis.
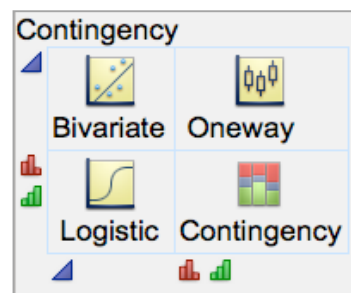
When you click **OK**, you'll see a mosaic plot and a two-way table called a **Contingency Table**. You'll also see test results at the bottom of the window. We'll omit discussion of these tests for now.

If your table shows values other than the counts in each cell, make your table easier to read by using the red triangle next to the heading **Contingency Table** to hide all values except the counts.

By looking at either the graph or the table, does this analysis confirm or refute your original belief about an association between the method of aging and the severity of the thread wear?

Copy the graphs and contingency table into your report and write a brief response to this question. Include your reasons for believing that an association might or might not be present based on the graphs and/or table.

Comparing Percentages

One way to assess whether there is an association between two variables is to calculate percentages, either in individual rows or individual columns, for the entries in the table.

Use the red triangle next to **Contingency Table** to show column percentages (**Col %**). *Copy this contingency table into your report.*

Now hide the column percentages and show row percentages (**Row %**). *Copy this new contingency table into your report.*

Are row percentages or column percentages reflected in the mosaic plot?

*Do either row or column percentages suggest dependence or independence of **method** and **thread wear**? Summarize your observations based on these percentages in your report.*

Exploring Other Factors

How about other factors among the categorical data? Repeat the steps above using the variable **Sand Blasted?** as the **X, Factor**.

*Were your ideas about independence of **sand blasted** and **thread wear** confirmed or refuted?*

Include graphs and some numerical information to support your answer. Remember to report your answers in the context of the problem.