
# JMP005: Patient Feedback Improving Patient Satisfaction

## Health Care Case

Endocrinology Group Patient Feedback

Produced by

Jane Oppenlander
Mary Ann Shifflet
Shirley Shmerling
Eric Stephens
Mia Stephens

# Improving Patient Satisfaction
Endocrinology Group Patient Feedback

**Key Ideas**: Exploratory Data Analysis and Process Improvement

## Background

A regional endocrinology specialty office is experiencing a decrease in the number of patients seen on a weekly basis. The office consists of four doctors, a nurse practitioner, two registered nurses (RNs), four licensed practical nurses (LPNs), and five administrative staff members. Among the clinic staff, there is some concern that this decrease in volume is due to patient dissatisfaction.

## The Task

A team is assembled from among the staff members to investigate potential causes of decreased patient volume. You are the internal consultant, brought in to assist the team with the analysis.

**The Data** ComplaintData.xls and PatientFeedback.jmp

ComplaintData.xls contains information on 262 customer complaints received over a three-month period. Complaint data is captured through the specialty office website.

| | |
|---|---|
| *Date* | Date complaint filed |
| *Complaint Type* | Reason for complaint |

PatientFeedback.jmp contains summary information from patient surveys administered over a three-month period.

| | |
|---|---|
| *Week #* | Week 1, 2, … |
| *Percent Recommend* | Percentage of responders who would recommend this office |
| *Wait* | Percent indicating wait time was acceptable |
| *Respect* | Percent indicating they were treated with respect |
| *Enough Information* | Percent indicating they received enough information from provider |

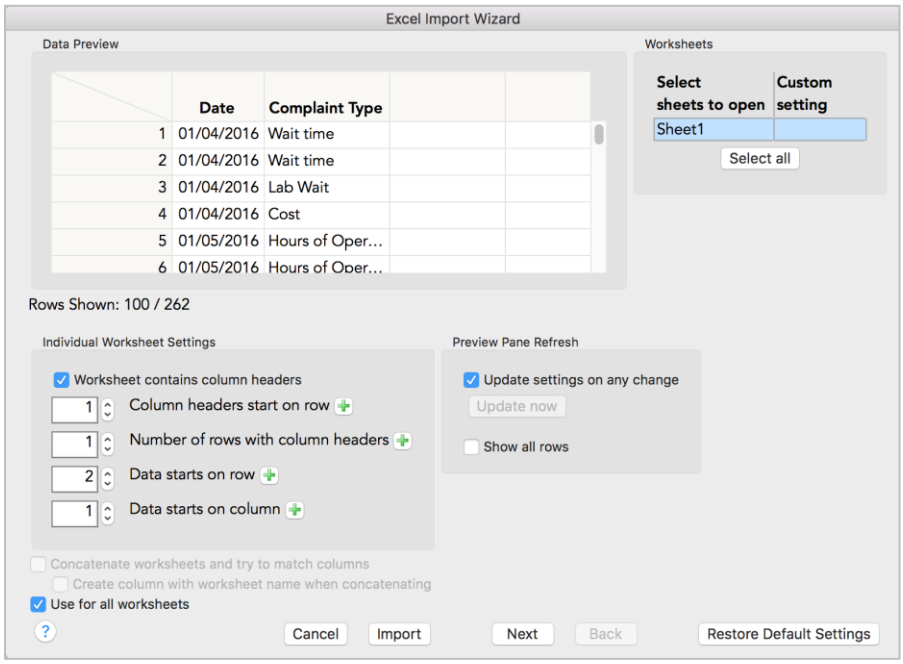
## Data Management

The team begins by looking at the patient complaint data. First, they import the data into JMP using the Excel Import Wizard, as shown in Exhibit 1. The first 15 rows are displayed in Exhibit 2.

The data values in both columns contain text or symbols. So, JMP assigns the Nominal modeling type to both variables, as denoted by the red bars ( ) next to the column names in the data table. Note that *Date* is stored in JMP with a date format. To see this, right click on the column head for *Date* in the data table and select **Column Info**.

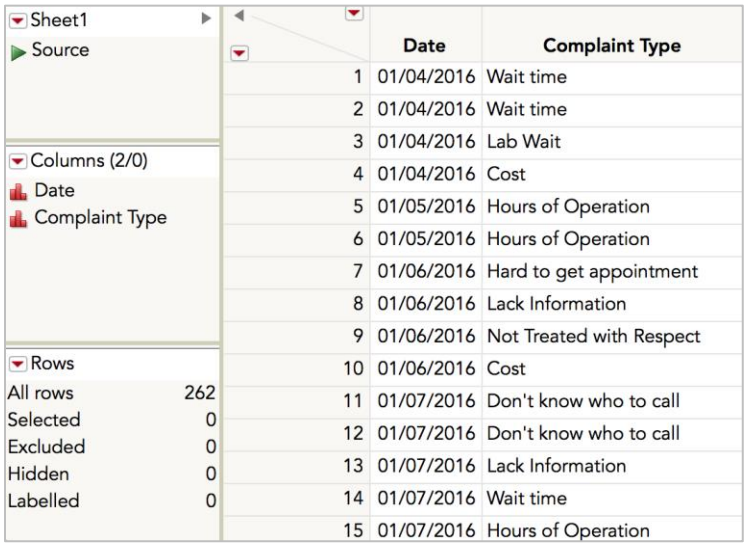
This file is saved as ComplaintData.jmp (using **File > Save As**).

**Exhibit 1** Importing the Complaint Data using the **JMP Excel Import Wizard**




*(Navigate to the file, and use **File > Open**. This launches the JMP Excel Import Wizard. Click Import to import the data into JMP.)*

**Exhibit 2** The Complaint Data in JMP




| | Date | Complaint Type |
|---|---|---|
| 1 | 01/04/2016 | Wait time |
| 2 | 01/04/2016 | Wait time |
| 3 | 01/04/2016 | Lab Wait |
| 4 | 01/04/2016 | Cost |
| 5 | 01/05/2016 | Hours of Operation |
| 6 | 01/05/2016 | Hours of Operation |
| 7 | 01/06/2016 | Hard to get appointment |
| 8 | 01/06/2016 | Lack Information |
| 9 | 01/06/2016 | Not Treated with Respect |
| 10 | 01/06/2016 | Cost |
| 11 | 01/07/2016 | Don't know who to call |
| 12 | 01/07/2016 | Don't know who to call |
| 13 | 01/07/2016 | Lack Information |
| 14 | 01/07/2016 | Wait time |
| 15 | 01/07/2016 | Hours of Operation |

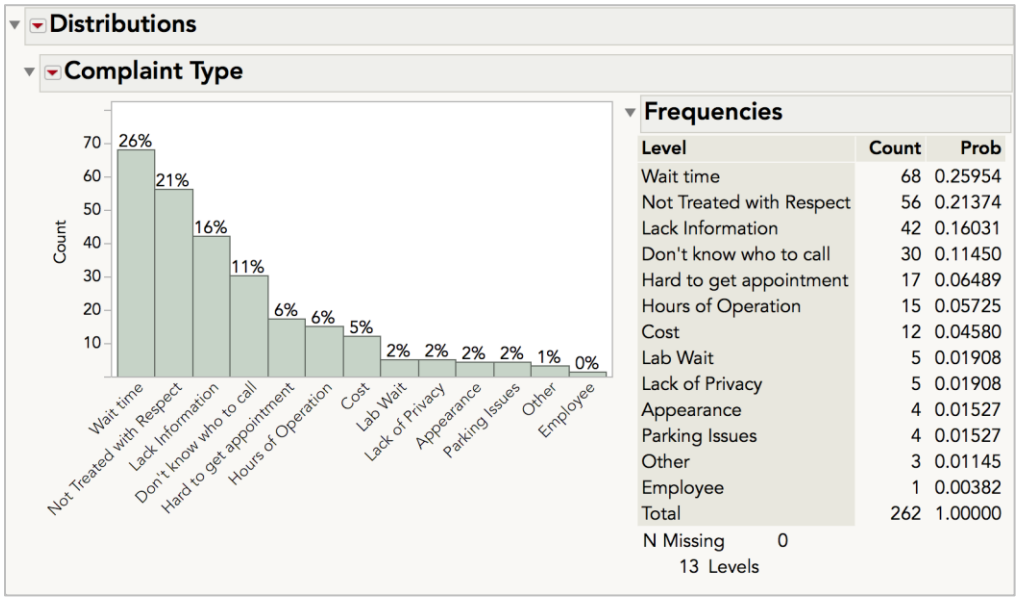
## Analyzing the Complaint Data

Since the office has not been in the habit of routinely looking at this data, the group is surprised by the volume of complaints over the last quarter. A Pareto Chart (see Exhibit 3) is used to summarize the data. This chart helps the team to prioritize the problems by the most frequently occurring complaint.

A Pareto Chart is a special case of a bar chart, where the events are displayed in descending order of frequency of occurrence. So, the most frequently occurring event is displayed on the far left and the least frequently occurring event is displayed on the far right.

The group learns that there were 262 complaints in the last quarter, and that the top three issues, accounting for 63% of the complaints, are Wait Time, Not Treated with Respect, and Lack information.

**Exhibit 3** Pareto Chart of Complaint Data




*(Analyze > Distribution; Select Complaint Type as Y, Columns, and click OK. For a horizontal layout select Stack under the top red triangle. To customize the graph, as shown above, select the following options from the red triangle next to Complaint Type: Histogram Options > Count Axis and Show Percents, Display Options > Axes on Left, and Order By > Count Descending. Note that Pareto Charts can also be produced from Analyze > Quality and Process and from Graph > Graph Builder.)*

Rather than speculate on causes of the negative feedback, the team decides to listen to the "voice of the customer" by using a short survey.

> "Voice of the customer", or VOC, is a term used in quality improvement to refer to the process of obtaining feedback on customer likes, dislikes and concerns.

Over a three-month period, patients seen in the office are given a short survey as they finish their appointments. The patients are asked the following questions:

- Was the wait time to be seen today acceptable?
- Were you treated with respect by the provider and office staff?
- Did you receive enough information from the provider?

In addition, all patients are asked:

- Would you recommend this office to others?

A 7-point Likert scale, commonly used in survey research to indicate participant's agreement with a statement, is used for all responses. A seven-point scale was used where, 1 = Totally Disagree to 7 =

Totally Agree. So, lower ratings indicate the patient is less satisfied, and higher ratings mean the patient is more satisfied.

The team decides to focus on the percentage of satisfied patients. They decide that patients responding with a 6 or a 7 on the Likert scale will be classified as satisfied, while patients responding with a 1 – 5 will be classified as not satisfied. At the end of the three-month period, for each question, the percentage of satisfied patients is calculated for each week. These data are found in the file PatientFeedback.jmp**.**

### Analyzing the Patient Satisfaction Data

The primary measure the team will use to assess patient satisfaction is *Percent Would Recommend*. The team starts by producing summary statistics for the variable *Percent Would Recommend* (see Exhibit 4).

**Exhibit 4** Summary Statistics – Percent Would Recommend

| | Percent Would Recommend |
|---|---|
| Mean | 57.8 |
| Std Dev | 8.9 |
| Median | 56.0 |
| Interquartile Range | 16.0 |
| Min | 42.0 |
| Max | 71.0 |

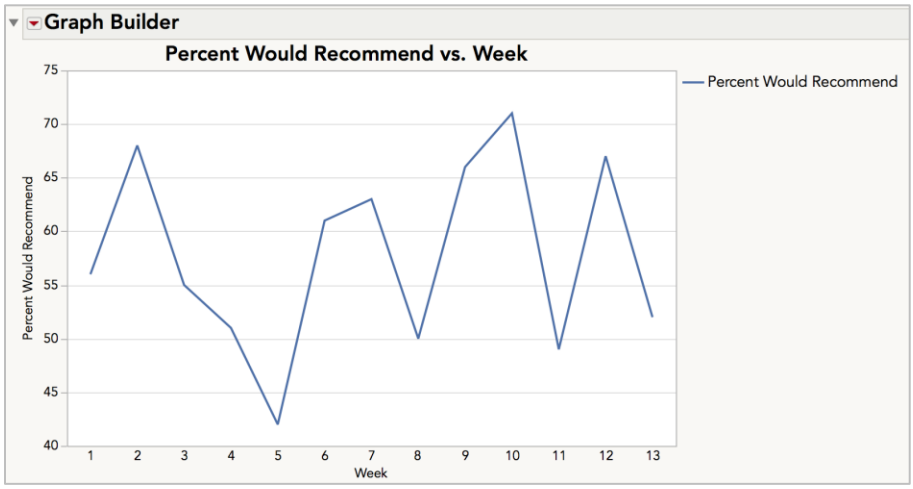
*(Analyze > Tabulate; Drag Percent Would Recommend to Drop Zone for Columns. Drag Summary Statistics to Drop Zone for Rows. To change the decimal format, click on Change Format box in lower left corner. Choose Use the Same Decimal Format, choose Fixed Dec, enter 1 under Number of Decimals, and click OK. Click Done to close the Tabulate control panel)*

The mean (the average) and the median (the middle-most value) are measures of the central tendency. It is clear that the overall patient satisfaction is much lower than the team expected, with roughly 58% of patients on average reporting that they would recommend the office to others.

The std dev (standard deviation) and the interquartile range (the range of the middle 50% of the values) are measures of spread or variability in patient satisfaction from one week to the next. The min (minimum) and max (maximum) values, which provide another indication of variability, show that the weekly percentage of satisfied patients ranges from 42% to 71%.

Run charts (see Exhibit 5) provide a graphical indication of the central tendency and the variability over time. Run charts also allow time-dependent patterns, such as the decrease in week 5, to be easily detected. (We revisit whether there are patterns over time in the exercises.)

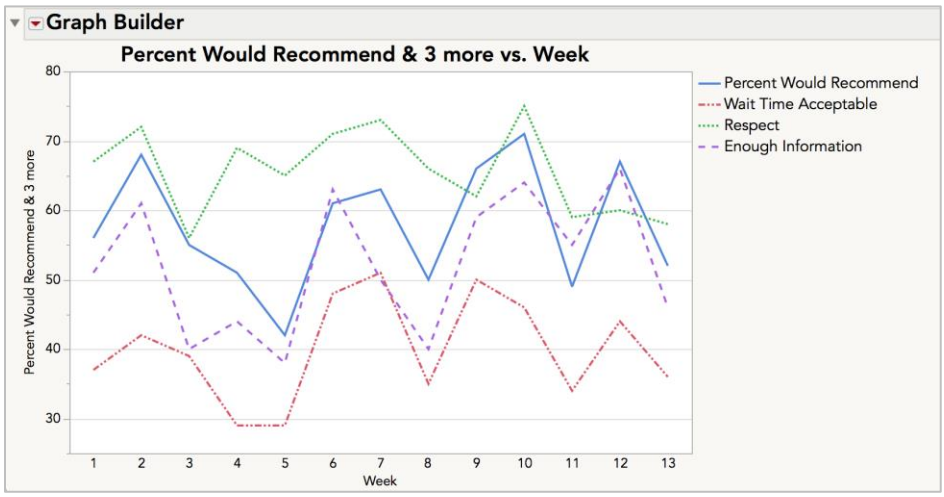
**Exhibit 5** Run Chart – Percent Would Recommend




*(Graph > Graph Builder; Drag Percent Would Recommend to the Y zone and Drag Week to the X zone. Select the Line icon at the top of the graph. Right-click on the y-axis, choose Axis Settings from menu, and check the box for Major Gridlines. Note that to apply a 0 to 100% scale (not shown here) Make Minimum = 0, Maximum = 100, Increment = 10, Tick Marks = 1. Click OK to accept axis setting changes, and click Done to close the Graph Builder control panel.)*

Run charts for the other three measures, along with *Percent Would Recommend*, are given in Exhibit 6. For ease of comparison, the four measures are plotted on the same graph. The values for these variables tend to move together, with the exception of *Respect*. For example, as *Percent Would Recommend* goes up *Enough Information* generally goes up, and as *Percent Would Recommend* goes down *Enough Information* generally goes down.

**Exhibit 6** Run Chart – All Four Measures



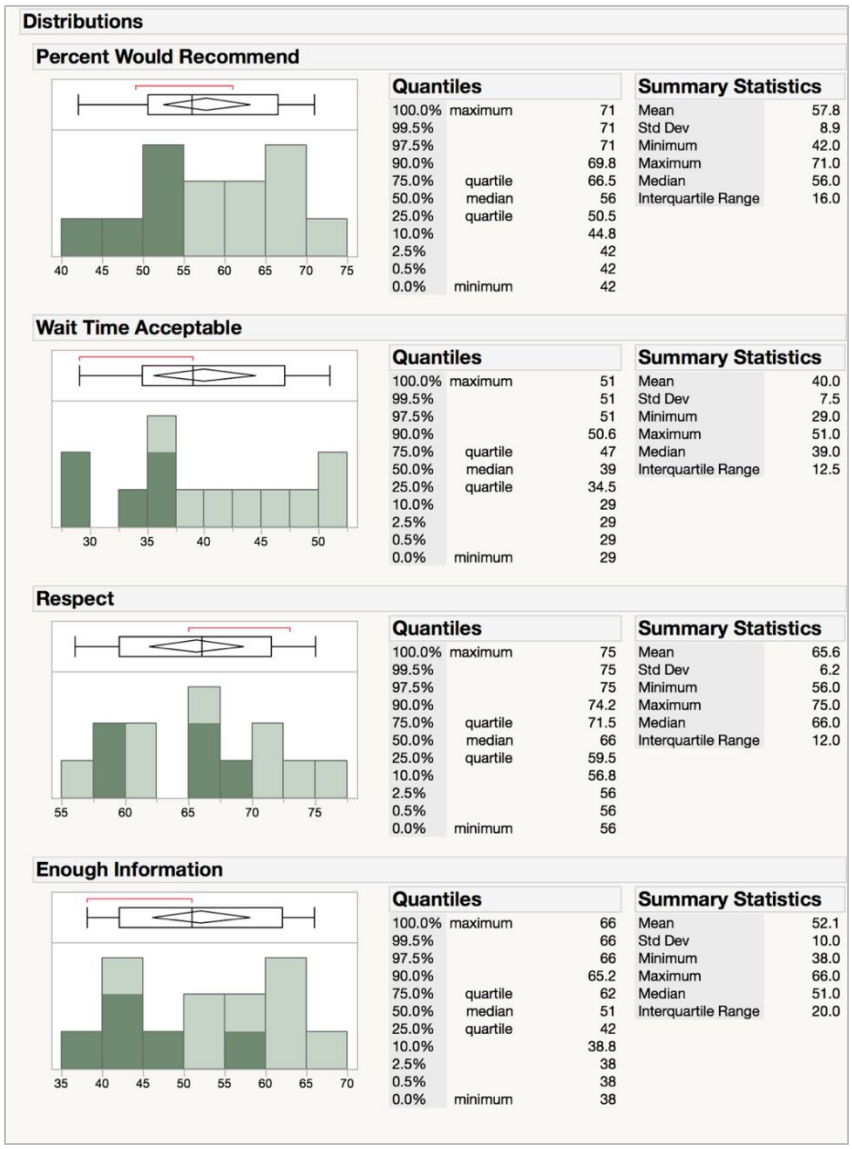

*(From the Graph Builder report produced in Exhibit 4, select Show Control Panel from the red triangle to reopen the control panel. Then, one at a time, drag each response to the drop zone just inside y-axis. To change the line styles, right-click on the response names in the legend on the right and select Line Style.)*

Histograms show the shape, center, and spread of the distributions. Histograms for the four responses, generated from the **Distribution** platform, are shown in Exhibit 7. The low values for *Percent Would*

*Recommend* are shaded, indicating that those rows are selected in the data table. Since the histograms are dynamically linked, the values for the selected rows are highlighted in the other histograms as well.

This confirms that low values for *Percent Would Recommend* generally correspond to low values for the other variables (except *Respect*).

**Exhibit 7** Distributions with Dynamic Linking




*(Analyze > Distribution; Drag all 4 variables to Y, Columns, and click OK. Click on red triangle next to Distributions and select Stack. Hold shift key and click on each of the lowest bars for Percent Would Recommend.)*

Along with the histograms, the **Distribution** platform provides summary statistics for each of the measures (see Exhibit 7). However, a more effective way to compare the summary statistics for the four variables is to use **Tabulate**, as shown in Exhibit 8.

**Exhibit 8** Summary Statistics – All Four Measures

| | Percent Would Recommend | Wait Time Acceptable | Respect | Enough Information |
|---|---|---|---|---|
| Mean | 57.8 | 40.0 | 65.6 | 52.1 |
| Std Dev | 8.9 | 7.5 | 6.2 | 10.0 |
| Median | 56.0 | 39.0 | 66.0 | 51.0 |
| Interquartile Range | 16.0 | 12.5 | 12.0 | 20.0 |
| Min | 42.0 | 29.0 | 56.0 | 38.0 |
| Max | 71.0 | 51.0 | 75.0 | 66.0 |

*(From the Tabulate report produced in Exhibit 5, select Show Control Panel from the red triangle to reopen the control panel. Then, drag Wait Time Acceptable to the Drop Zone for Columns area next to Percent Would Recommend. Repeat for each measure.)*

These measures confirm some of what the team learned from the patient complaint data. A low percentage of patients respond that the wait time is acceptable and that they are given enough information from the provider. These are clearly two important issues that the team must address.

To better understand whether long wait times and not receiving enough information might be related to low recommendations, the team looks at the correlations among the four variables. A table of correlations, along with the corresponding scatterplot matrix, is shown in Exhibit 9.

A correlation is a numeric measure of the strength of the linear relationship between two numeric variables. A correlation value close to one indicates a strong positive relationship, a value close to negative one indicates a strong negative relationship, and a value close to 0 indicates weak/no relationship. While not proving causation, a high correlation can, in some cases, be an indicator of a potential cause-and-effect relationship.

The scatterplot matrix (bottom, in Exhibit 9) provides a visual representation of the association between the pairs of variables. The graphs in the first row of the scatterplot matrix represent the relationship between *Percent Would Recommend* and each of the other measures. The tighter (narrower) the red ellipses, the stronger the correlation.

The high correlations between *Acceptable Wait Time* and *Percent Who Would Recommend* (0.85) and between *Enough Information* and *Percent Would Recommend (*0.82) suggests potential to improve *Percent Would Recommend* by solving these two problems.

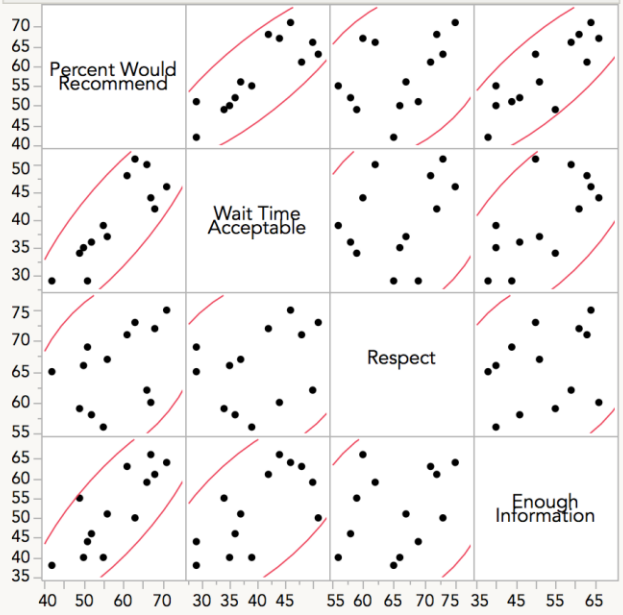
**Exhibit 9** Correlations

**Multivariate**

**Correlations**

| | Percent Would Recommend | Wait Time Acceptable | Respect | Enough Information |
|---|---|---|---|---|
| Percent Would Recommend | 1.0000 | 0.8452 | 0.4128 | 0.8242 |
| Wait Time Acceptable | 0.8452 | 1.0000 | 0.3189 | 0.6779 |
| Respect | 0.4128 | 0.3189 | 1.0000 | 0.3058 |
| Enough Information | 0.8242 | 0.6779 | 0.3058 | 1.0000 |

**Scatterplot Matrix**




*(Analyze > Multivariate Methods > Multivariate; Select all variables of interest as Y, Columns and click OK.)*

## Summary

In this case we:

- Use a Pareto chart to identify the most frequently occurring complaints.
- Use run charts to plot measures over time.
- Use tabular summaries to produce basic descriptive statistics.
- Use histograms and dynamic linking to explore relationships between variables.
- Use correlations and scatterplots to explore and quantify the relationships between variables.

## Concepts and Tools

| **Statistical Methods/Tools** | **Data Management Tasks/Concepts** | **JMP Platforms/Features** | **Quality Tools** |
|---|---|---|---|
| Bar Charts | Import Excel Spreadsheet | Distribution Platform | Pareto Chart |
| Run Charts | | Graph Builder | Voice of the Customer |
| Histograms | | Tabulate | |
| Scatterplots | | Multivariate Methods | |
| Summary Statistics | | | |
| Correlations | | | |

## Exercises

1. Use the PatientFeedback.jmp data set to answer the following questions:

    a. Is the distribution for *Percent Would Recommend* symmetric?
    b. Based on the shape of the distribution for *Percent Would Recommend*, which measure of centrality, mean or median, is the most representative?
    c. For the run chart in Exhibit 5 (*Percent Would Recommend*) the default *y*-axis scaling is roughly 40% - 70%. Why might it make more sense to change the axis to 0 – 100%?
    d. Reproduce the Pareto chart given in Exhibit 3 using the **Pareto Plot** platform instead of the **Distribution** platform (Analyze > Quality and Process > Pareto Plot). What are the top 5 patient complaints? What percentage of all complaints do the top 5 complaints account for?
    e. Recreate the run chart shown in Exhibit 6. Identify the weeks in which *Percent Would Recommend* was particularly high or low. Are the other variables also high or low? Does this tell you anything about cause and effect?

2. Open the Fitness.jmp dataset in the JMP Sample Data directory (under Help > Sample Data Library).

    a. Create a scatterplot matrix, and find the correlations among the continuous variables following the directions provided in this case.
    b. Which pair of variables has the strongest positive correlation (and what is the value)?
    c. Which pair of variables has the strongest negative correlation (and what is the value)?
    d. What does this negative correlation indicate?

3. Open the ComplaintData.jmp file, and regenerate the Pareto Plot displayed in Exhibit 3. There appears to be a data quality issue with "*Don't know who to call*". Use Cols > Recode to fix this issue, and re-run the analysis. Does this change the conclusions regarding the most serious issues?

4. The data set ComplaintDataEx.jmp contains data collected for a three-month period after steps were taken to address the top two complaints.

    a. Create a Pareto Plot for *Complaint Type*. Was there improvement in wait time? Are there any other key issues that need to be addressed?
    b. Create a new formula column for day of the week. Create both a graph and a tabular summary of complaints by day of the week. Describe what you learn. Are there more complaints on some days than others? Do some complaint types occur more frequently on some days than on others? How could these issues be addressed?
    c. Create a new formula column for week. Create a run chart for complaints by week. Describe what you observe.
    d. Open question: Explore the data using other tools available in JMP. Can we learn anything else from these data?

## Discussion Questions

1. The file ComplaintData.jmp contains the complaint type and the date of the complaint. This information was downloaded from the specialty office website.
    a. Discuss the quality of this information for understanding customer complaints.
    b. What additional information would be useful?
    c. What information should be captured by the complaint data collection system?
2. The data in the file PatientFeedback.jmp contains summarized satisfaction results based on customer surveys.
    a. Discuss how these data were collected and summarized.
    b. From a data collection perspective what would you do differently (if anything)?
    c. What would you recommend in terms of collection of patient satisfaction information going forward?
3. Discuss the difference between correlation and causation. Specifically, why does correlation not imply causation? Provide at least one example where there is a strong correlation but clearly no evidence of causation.
4. How is causation established? Discuss what actions might be taken to establish causation.
5. In this case, the team has identified some potential causes for poor patient satisfaction ratings. What are potential next steps to address these issues?
6. After changes are implemented to address the issues, how will you verify that changes you've made have improved patient satisfaction and weekly number of patients seen? How will you verify that these issues don't return (i.e. that patients continue to be satisfied and that the number of patients seen doesn't drop over time)?