

JMP Instructional data set 002

Cost of Goods

Relevant JMP platforms and statistical methods

Graph Builder: Time series graph; bar chart
Tabulate: Table of summary statistics
Formula Editor: Creating new variables

Problem statement

The U.S. Bureau of Labor Statistics retains historic data on the prices consumers pay for common goods such as food and utilities. Economists analyze these data to produce studies that help the government and financial institutions get visibility into the costs that families are incurring and how those costs are changing. This information is extremely valuable to guide various legislative and monetary policies.



One such summarized dataset consists of the average monthly prices from April 2003 through Feb 2026 on a set of goods averaged over a sample of locations across the country [1].

Data

IDS_002.jmp

Month-Year	Month and Year (04/2003 – 03/2026)
Bananas	Average cost of a lb of bananas
Oranges	Average cost of a lb of navel oranges
Tomatoes	Average cost of a lb of field grown tomatoes
Bread	Average cost of a lb of white pan bread
Milk	Average cost of a gallon of whole fortified milk
Eggs	Average cost of a dozen of large eggs
Chicken	Average cost of a lb of whole chicken
Beef	Average cost of a lb of ground chuck beef
Electricity	Average cost of a KWH of electricity delivered as a home utility
Gas_Auto	Average cost of a gallon of regular unleaded automotive gasoline
Gas_Home	Average cost of a therm (100 cubic feet) of natural gas as home utility

Note: The U.S. BLS gathers a tremendous amount of cost data beyond just these items, and it's a substantial effort to aggregate the vast sources of data across the country into a single cost number for each item in a given month. As a result, it's not uncommon to have missing data for a specific item in some months.

Exercises

1. Create a time series graph of the monthly cost for each of the items. Provide a few short sentences describing some of the key features these graphs show for each item (e.g., Describe the general trend. Describe the variation around that general trend. Is the variation large/small? Is the variation consistent in any way, or is it sporadic? Are there time periods when the prices significant decrease/increase much more than other time periods? Any seasonal patterns? Any time periods where prices stand out as being very different from the rest?

Instructions:

Graph > Graph Builder. Place the Month-Year variable on the X axis, and 'Bananas' on the Y axis. Choose to display both the points and a connect line by selecting those options in the Graph Palette.  

Under the Red Triangle Menu, choose Redo > Column Switcher. Select 'Bananas' from the Initial Column list, and select the eleven variables: 'Bananas', ..., 'Gas_Home' for the Replacement Columns. Use the Column Switcher added to the graph window to view the monthly cost for any of the items.

2. Add a 'Linear Regression Line' to the time series graphs. Does this equation do a reasonable job at modeling/describing the change in prices over time for any of these items? If so, describe what features are being captured by this type of model. Describe what features are not being captured by this type of model? Do you think these models could be useful in any way at predicting the prices of any of these items in the future?

Instructions:

Select 'Line of Fit' icon  from the graph palette.

3. Create a new set of variables quantifying the % change in the monthly price of each item from 12-months prior.

Instructions:

Create a new column (Cols > New Columns). In the New Column dialog box, choose Formula in the Column Properties drop down menu. Select Edit Formula. In the Formula Editor, create the formula:

$$\left(\frac{\text{Bananas} - \text{Lag}(\text{Bananas}, 12)}{\text{Lag}(\text{Bananas}, 12)} \right) \cdot 100$$

Note: the 'Lag' function is in the Row category in the function list on the left.

Name the column 'Bananas_12 month % Chg'.

Repeat this for the other ten items.

4. Create times series bar charts of the 12-month % change for each of the items. Describe some features these graphs reveal (e.g., Are there items where the 12-month % change is consistent over many years? Which items experience the most variation in % price change? Which items experience periods with very large increase or decrease? What are some items and time periods where the % price change is sustained for many years? Do different items tend to experience large increases/decreases during similar time periods?).

Instructions:

Graph > Graph Builder. Place the Month-Year variable on the X axis, and 'Bananas_12 month % Chg' on the Y axis. Choose Bar graph from the Graph Palette. 

Under the Red Triangle Menu, choose Redo > Column Switcher. Select 'Bananas_12 month % Chg' from the Initial Column list, and select the eleven new variables: 'Bananas_12 month % Chg', ..., 'Gas_Home_12 month % chg' for the Replacement Columns. Use the Column Switcher added to the graph window to view the monthly cost for any of the items.

Note: As you change which variable to display, the Y axis rescales so the data values fill the graph region. You can manually rescale the Y axis by floating the cursor at the very bottom and/or top of the Y axis until it turns into an upward/downward facing hand, then clicking and dragging. To assist in comparing data across the different items, you can lock the scale at a particular min/max you set. Select Lock Scales under the red triangle next to Graph Builder title. Then choose Retain Axis Settings under the red triangle next to the Column Switcher.

5. Create a table of the average 12-month % change for each of the items. Identify the three highest and three lowest average values. What are some factors important to include when using these summary statistics to describe how cost of these goods have been changing over time?

Instructions:

Analyze > Tabulate. Select all eleven 12-month % Chg variables and place in the Drop zone for rows. Replace the Sum with the Mean by dragging 'Mean' from the list of statistics and dropping on top of one of the 'Sum' labels in the table.

6. Considering all the goods/services a typical family may purchase and how the cost of these varies across the country, what are some other data that would be informative to examine to help develop an even deeper understanding of the cost-of-living families experience?

Supporting materials

1. Average price data from U.S. Bureau of Labor Statistics
<https://www.bls.gov/charts/consumer-price-index/consumer-price-index-average-price-data.htm>

** Please adhere to any citation requirements for distribution and use of these data.*