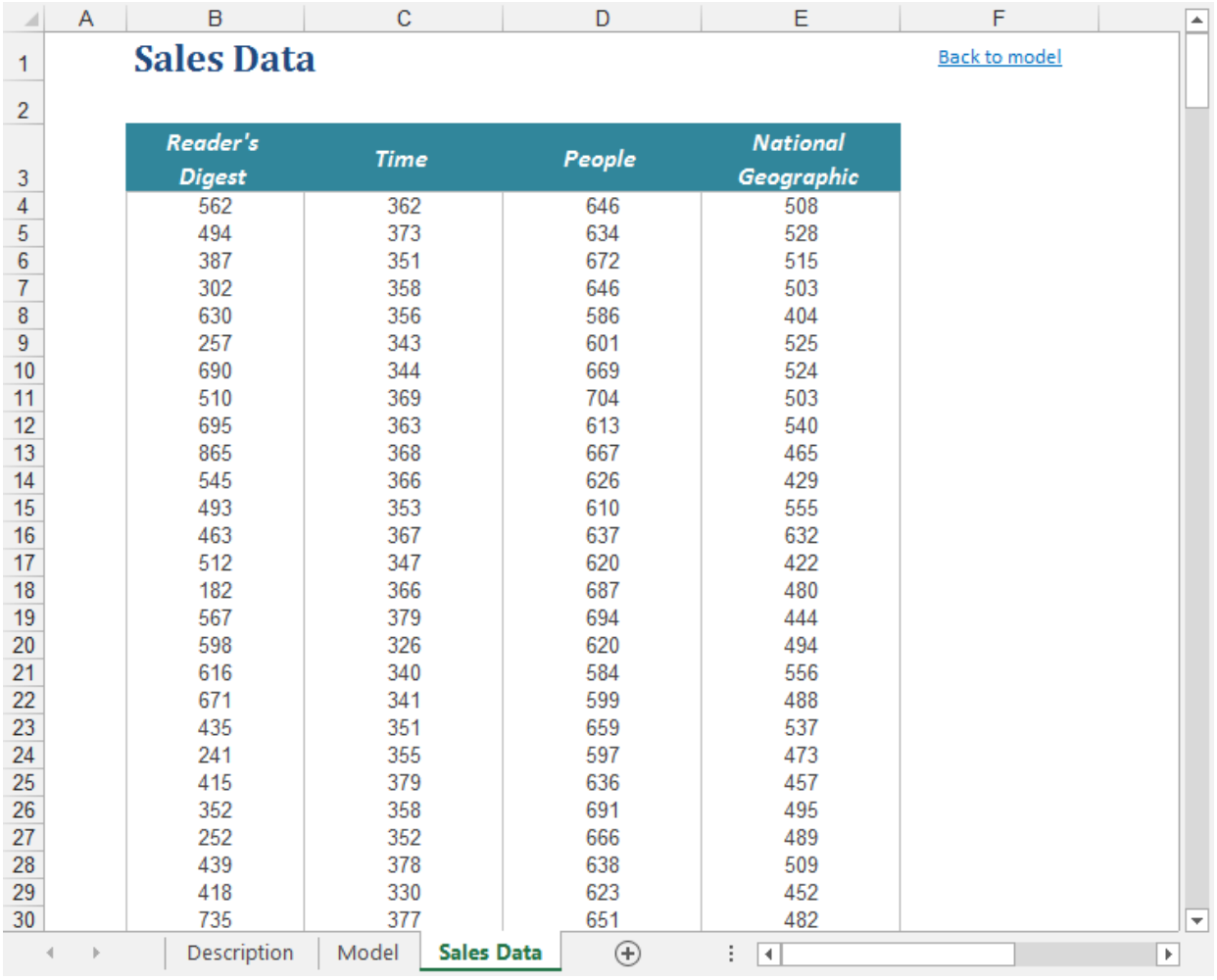
# One-Minute Spotlight

## BATCH FIT TOOL

The Batch Fit Tool automatically fits continuous probability distributions to multiple data series. Batch Fit is intended to save you time when you need to create several assumptions from historical data.

In the model Magazine Sales.xls (available in the Crystal Ball examples), you want to forecast the gross profit resulting from newsstand sales of four popular magazines. On the Sales Data worksheet, you have a series of 360 data points (daily sales data) for each magazine.



Sales Data

Back to model

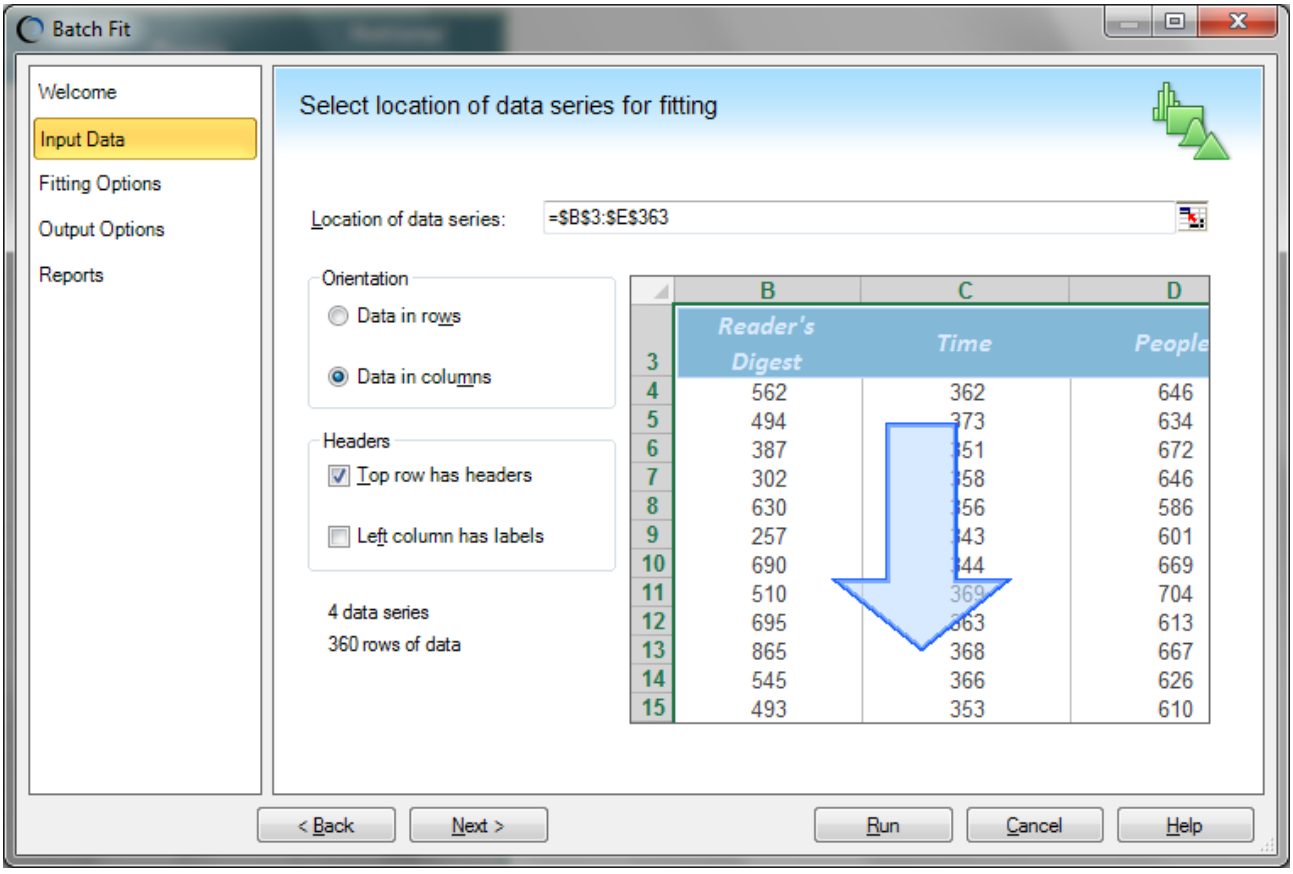
| Reader's Digest | Time | People | National Geographic |
|---|---|---|---|
| 562 | 362 | 646 | 508 |
| 494 | 373 | 634 | 528 |
| 387 | 351 | 672 | 515 |
| 302 | 358 | 646 | 503 |
| 630 | 356 | 586 | 404 |
| 257 | 343 | 601 | 525 |
| 690 | 344 | 669 | 524 |
| 510 | 369 | 704 | 503 |
| 695 | 363 | 613 | 540 |
| 865 | 368 | 667 | 465 |
| 545 | 366 | 626 | 429 |
| 493 | 353 | 610 | 555 |
| 463 | 367 | 637 | 632 |
| 512 | 347 | 620 | 422 |
| 182 | 366 | 687 | 480 |
| 567 | 379 | 694 | 444 |
| 598 | 326 | 620 | 494 |
| 616 | 340 | 584 | 556 |
| 671 | 341 | 599 | 488 |
| 435 | 351 | 659 | 537 |
| 241 | 355 | 597 | 473 |
| 415 | 379 | 636 | 457 |
| 352 | 358 | 691 | 495 |
| 252 | 352 | 666 | 489 |
| 439 | 378 | 638 | 509 |
| 418 | 330 | 623 | 452 |
| 735 | 377 | 651 | 482 |

Description | Model | Sales Data

## OPENING THE TOOL AND SELECTING THE INPUTS

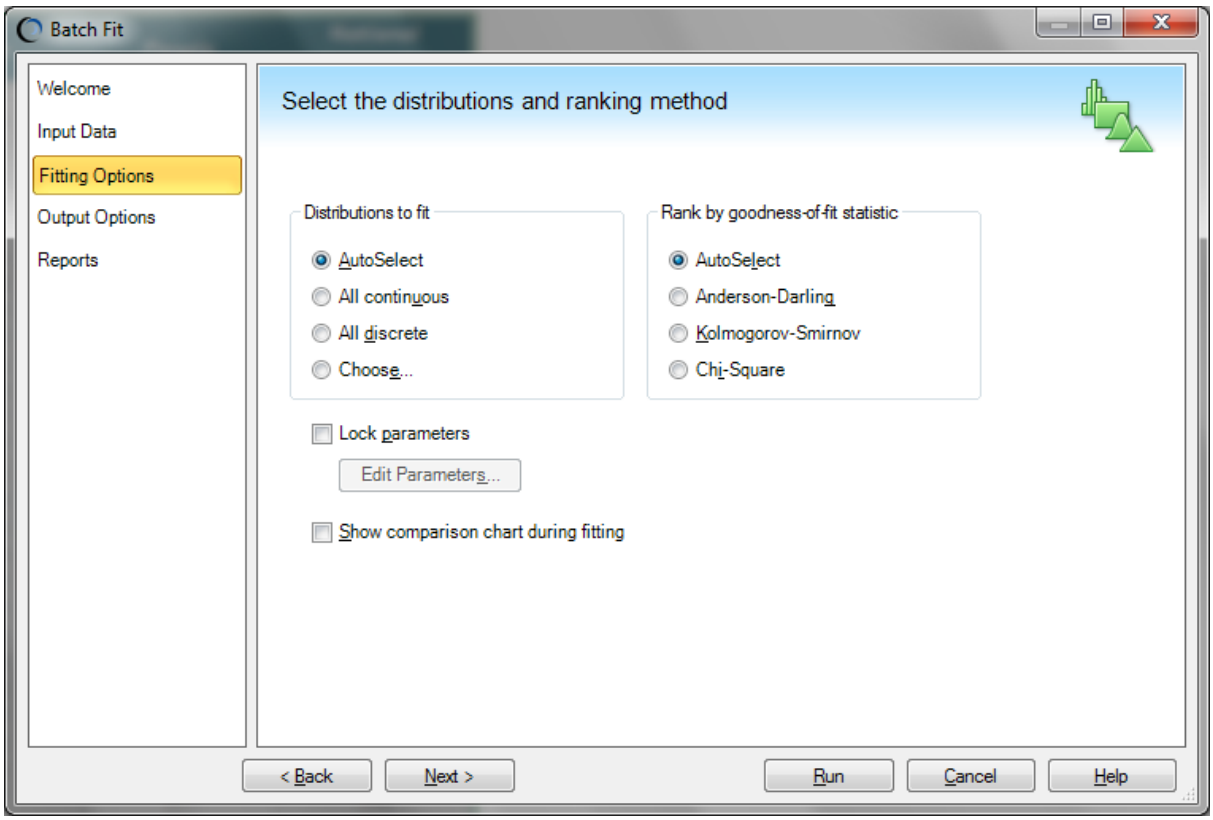
You open the Batch Fit tool through the Run -> Tools menu.

When the tool opens, you will see the Welcome screen. Click on Next to view the Input Tool step. Here you define the location of your data table and select the Fitness Criteria. Your data series must be contiguous (in adjacent rows or columns) in either rows or columns. Click on Next to move to Fitting Options.



## SELECTING YOUR FITTING OPTIONS

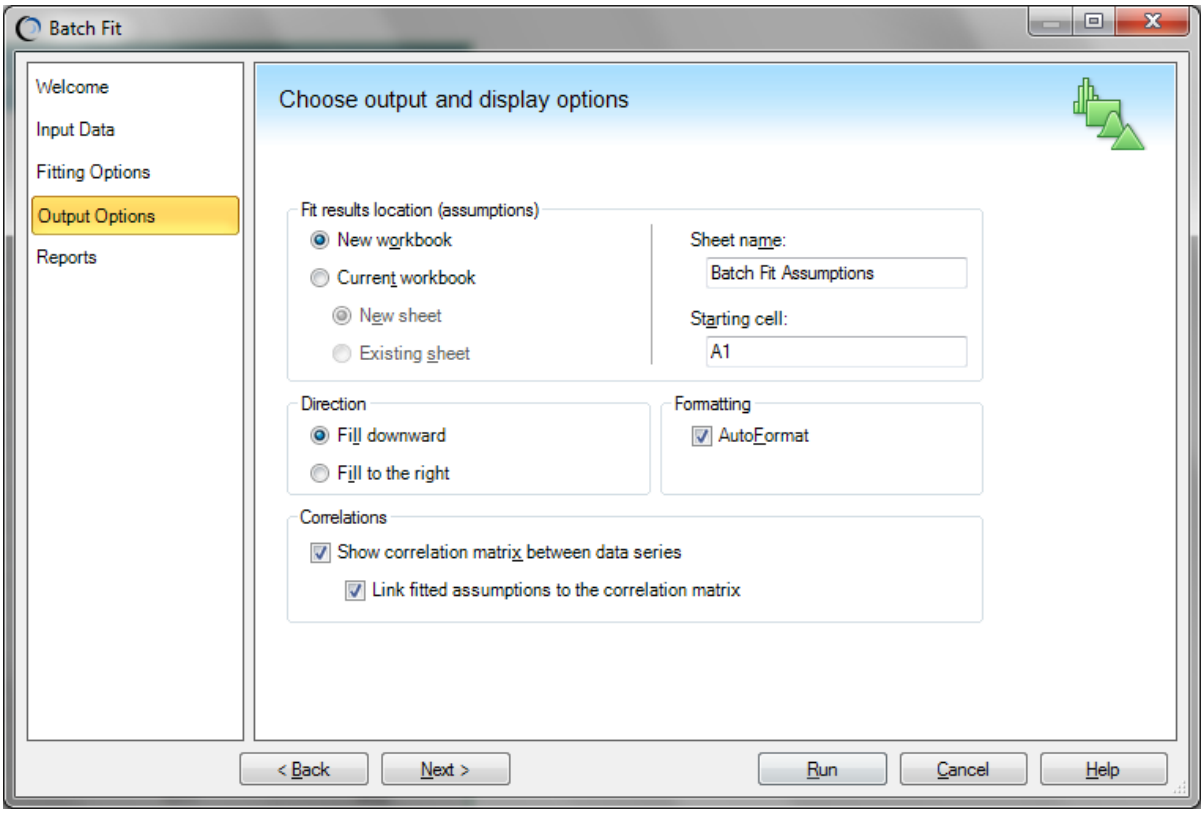
The tool lists all of the Crystal Ball continuous and discrete that may be fitted to your data series. The AutoSelect setting will calculate the fits for the appropriate distributions using the appropriate ranking method. For example, the tool will only use the Chi-Square method for fitting to Discrete distributions.




After you have selected the distributions to fit, click on Next to select your Output Options.

## SELECTING YOUR OUTPUT OPTIONS

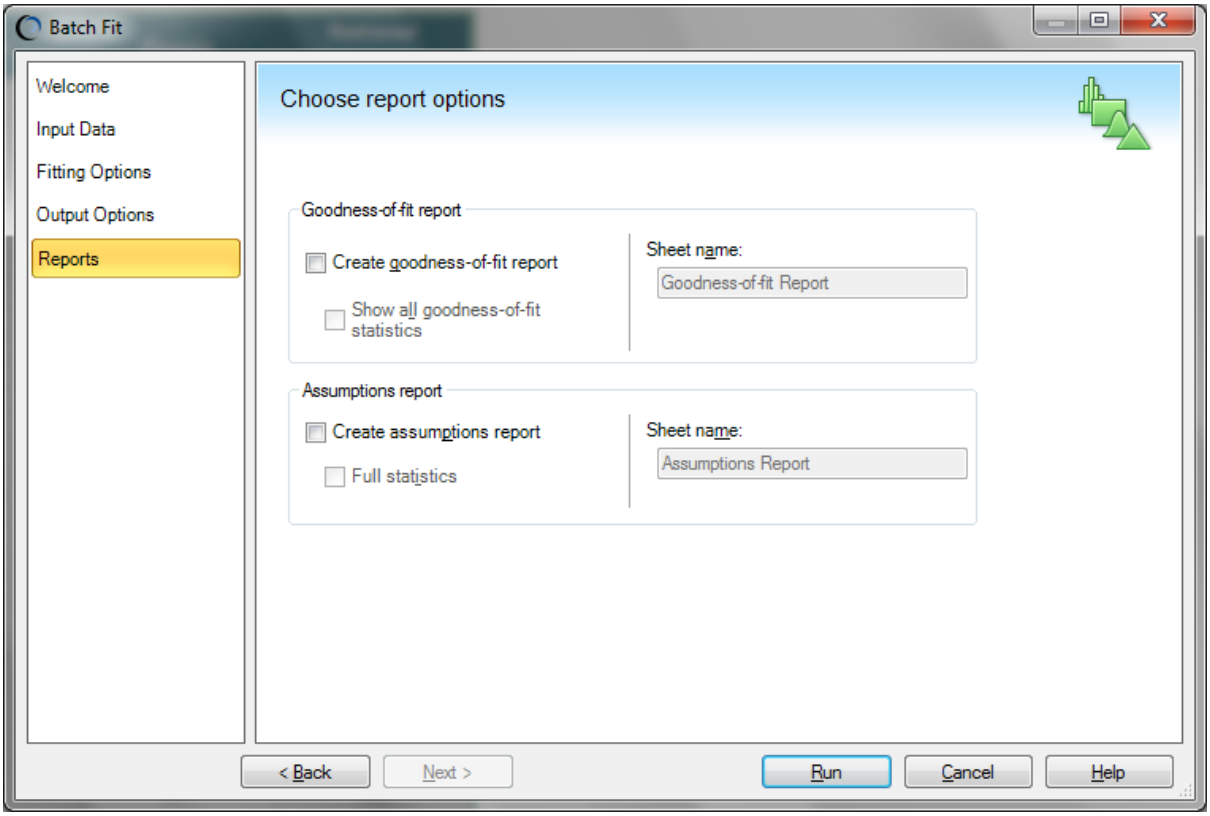
Choose where you want your assumptions to be created, in a new workbook or in the current workbook. You can optionally choose to output a correlation matrix, showing the relationships between each of the data series (i.e. magazine sales).




In this example, we'll select the two options to calculate correlations between the data series and link them to the assumptions. Click on Next to select your Report Options.

## SELECTING YOUR REPORT OPTIONS

In the final step (shown below), you can choose among several reports that let you see the details of the Batch Fit process, as well as details for the created assumptions.



## EXAMINING THE BATCH FIT RESULTS

When the Batch Fit tool runs, it fits each distribution to each column of data. The results shown below were posted into a new workbook, but you can have them output to your model worksheet.

| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| 1 | **Data Series:** | Reader's Digest | Time | People | National Geographic | | |
| 2 | **Distribution:** | 499.46 | 354.66 | 637.5 | 480.71 | | |
| 3 | **Best Fit:** | Neg Binomial | Binomial | Discrete Uniform | Gamma | | |
| 4 | **Rank Methods:** | | | | | | |
| 5 | Chi-Square | 23.9515 | 33.3925 | 33.0667 | | | |
| 6 | Anderson-Darling | | | | 0.0893 | | |
| 7 | **P-Value:** | 1.000 | 0.854 | 1.000 | 0.997 | | |
| 8 | | | | | | | |
| 9 | **Correlations:** | Reader's Digest | Time | People | National Geographic | | |
| 10 | Reader's Digest | 1 | | | | | |
| 11 | Time | 0.0165508 | 1 | | | | |
| 12 | People | -0.012844488 | -0.005857389 | 1 | | | |
| 13 | National Geographic | -0.012360848 | 0.060698976 | -0.039167797 | 1 | | |
| 14 | | | | | | | |
| 15 | | | | | | | |
| 16 | | | | | | | |
| 17 | | | | | | | |
| 18 | | | | | | | |

Batch Fit Assumptions

Row 2 contains the new Crystal Ball assumption for each series, with the distribution type described in Row 3. Rows 5 and 6 show the goodness-of-fit statistics for the best fitting distribution, and Row 7 shows the P-value of the fit (when the value can be calculated). The calculated correlation coefficients are listed in a matrix at the bottom of this sheet (Rows 10-13); these coefficients are also defined within the assumptions.

Using Crystal Ball's Copy and Paste commands, you can copy the assumptions in row 3 to the appropriate cells in your model.

**Note:** If your historical data have a time-series element, trend, or seasonality, you should use CB Predictor instead of the Fit Distribution functionality.

For more information or to contact us, browse to http://www.oracle.com/technetwork/middleware/crystalball/overview/index.html.