

Reshape data table

Using the Sample LA dataset, here is the data before Reshape. Note: you can select a subset of columns.

Current selection: Sample [LA]

Optional: Select columns. If no selection, all columns are used in order. Dependent variable must be first, drag to reorder.

Optional: Select/deselect all columns. To delete multiple items in selection box, use Control or Shift key to select them, then press DELETE key

Sort column names in dropdown

Optional: Select integer/character variables to make discrete:

AllCause75 tmin tmax Month Day Year

Show 10 entries Search:

	AllCause75	PM2.5	tmin	tmax	MAXRH	Month	Day	Year
1	151	38.4	36	72	68.8	1	1	2007
2	158	17.4	36	75	48.9	1	2	2007
3	139	19.9	44	75	61.3	1	3	2007
4	164	64.6	37	68	87.9	1	4	2007

Now, we want to reshape the data table to see the day trend from year to year. Using the following settings:

Reshape data table

Columns to gather, and spread using values in 'Columns to unite'

AllCause75 x PM2.5 x tmin x tmax x MAXRH x

Put all columns into gather box above, then unselect the columns to unite.

Columns to unite, and to be removed

Year x

Reshape Results will be displayed in same table above.

When click Reshape button, the data table is replace by:

Optional: Select integer/character variables to make discrete:

Month Day AllCause75_Year2007 AllCause75_Year2008 AllCause75_Year2009 AllCause75_Year2010 tmax_Year2007 tmax_Year2008 tmax_Year2009 tmax_Year2010 tmin_Year2007 tmin_Year2008 tmin_Year2009 tmin_Year2010

Show 10 entries Search:

	Month	Day	AllCause75_Year2007	AllCause75_Year2008	AllCause75_Year2009	AllCause75_Year2010	MAXRH_Year2007
1	1	1	151	146	153	146	68.8
2	1	2	158	166	155	158	48.9
3	1	3	139	150	163	150	61.3
4	1	4	164	156	135	126	87.9

Due to different number of days in February, there is a missing row in the result. We simply use **Clean Rows** to delete the rows with missing data:

Number of rows with missing data = 1

Impute Data	Impute the columns displayed in above table.
Clean Rows	Clean all rows in original table even if only a subset are selected.

The reshaped data columns have **_Year*** appended at the end: this column names contains time sensitive info that can be used for Bayesian analysis.

Describe

CAT

- Analyze
- Bayesian
- Causal
- Correlations
- Granger

Input

Constraints and model

Column names contain time-series defined by the following pattern. Note: * is a special character indicating integer value in column names.

For example, check Reshape data table or Add lagged columns

The default timing filter **_Year* _Month* _Day*** works fine for this dataset, although **_Year*** also works in this case. Above checkbox is check to indicate that black list constraints should be generated when performing **bnlearn**. After Run, we see that no arcs are going from a later Year to previous Year:

Run

Output

Bayesian Network diagram.

Network discovered by bnlearn using model hc

Number of blacklist items defined by Column names containing time-series pattern: 190

An arrow between two variables shows that they are informative about each other.

The diagram is a directed acyclic graph (DAG) representing a Bayesian network. The nodes are labeled with variable names, including: MAXRH_Year2007, PM2.5_Year2007, AllCause75_Year2007, tmin_Year2007, tmax_Year2007, MAXRH_Year2008, PM2.5_Year2008, AllCause75_Year2008, tmin_Year2008, tmax_Year2008, Day, MAXRH_Year2009, PM2.5_Year2009, AllCause75_Year2009, tmin_Year2009, tmax_Year2009, MAXRH_Year2010, PM2.5_Year2010, AllCause75_Year2010, tmin_Year2010, and tmax_Year2010. There is also a 'Month' node at the top. Arrows indicate dependencies, showing that variables in a given year are influenced by variables in the same year and the previous year. For example, MAXRH_Year2007 influences PM2.5_Year2007, AllCause75_Year2007, and tmin_Year2007. The 'Month' node influences tmin_Year2007, tmax_Year2007, and tmin_Year2008. The 'Day' node influences tmin_Year2008 and tmax_Year2008. There are no arrows pointing from later years to earlier years, indicating that the network respects the temporal order.