

Advanced Data Preparation Using IBM SPSS Modeler

Course Outline

1: Using functions to cleanse and enrich data

- Use date functions
- Use conversion functions
- Use string functions
- Use statistical functions
- Use missing value functions

2: Using additional field transformations

- Replace values with the Filler node
- Recode continuous fields with the Binning node
- Change a field's distribution with the Transform node

3: Working with Sequence Data

- Use cross-record functions
- Count an event across records
- Expand a continuous field into a series of continuous fields with the Restructure node
- Use geospatial and time data with the Space-Time-Boxes node

4: Sampling, partitioning and balancing data

- Draw simple and complex samples with the Sample node
- Create a training set and testing set with the Partition node
- Reduce or boost the number of records with the Balance node

5: Improving Efficiency

- Use database scalability by SQL pushback
- Process outliers and missing values with the Data Audit node
- Use the Set Globals node

- Use parameters
- Use looping and conditional execution