# Weather Underground Import

Import Weather Underground data into IBM SPSS Modeler


Product: IBM® SPSS® Modeler

Extension type: Utility

**Table of Contents**

## Description:

This extension imports weather data for a vector of valid 3 or 4 digit Airport codes or a valid Weather Station ID (examples: "BUF", "ORD", "VABB" for Mumbai). The extension will return a record of weather information for a given date at each location.

## Requirements:

- SPSS Modeler v16.0 or later
- R: http://www.r-project.org/
- 'R Essentials for SPSS Modeler' plugin: https://developer.ibm.com/predictiveanalytics/downloads/

## Installation:

Close SPSS Modeler. Save the .cfe file in the CDB folder of the IBM SPSS Modeler installation directory for Windows and Linux. The copy should reside in that same folder and not in a sub-folder.

For example, for Windows 7 the default location is "C:\ProgramData\IBM\SPSS\Modeler\16\CDB". If the ProgramData folder is hidden type the path manually.

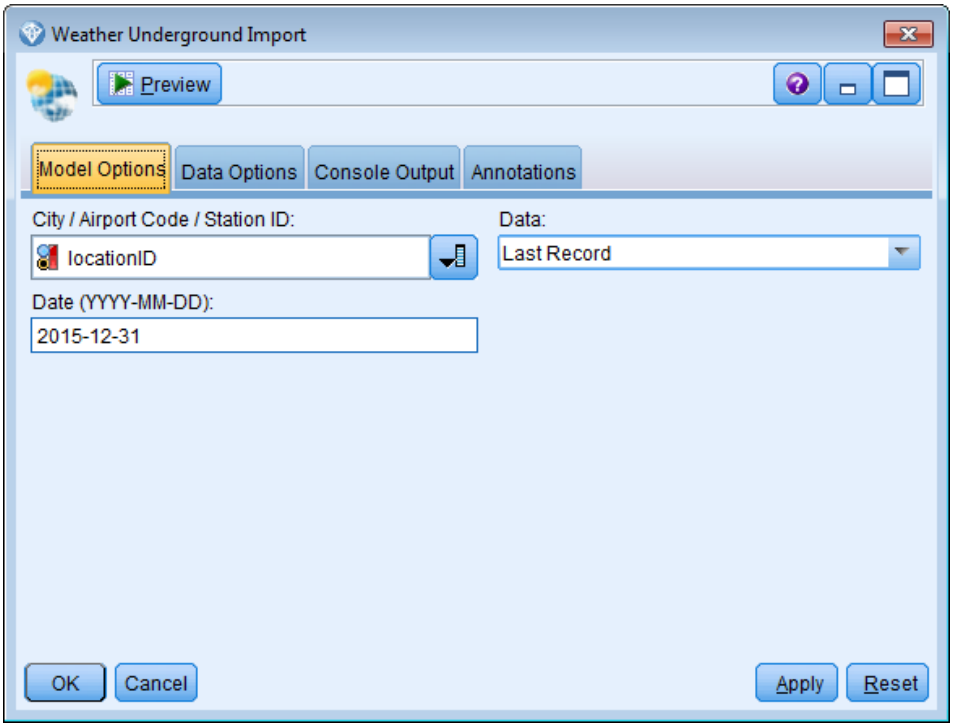
Restart SPSS Modeler: the node will now appear in the Field Ops palette.

## R Packages used:

- 'plyr' package created by Hadley Wickham Lang https://cran.r-project.org/web/packages/plyr/
- 'weatherData' package created by Ram Narasimhan https://cran.r-project.org/web/packages/weatherData/

## User Interface

- Double click on the node to get to the options. There are the following fields:
  - City/Airport Code/Station ID: click the drop down to select the column containing the location
  - Data: Choose the specific field of data to be imported for each location or the last record available for all data points (default)
  - Date (YYYY-MM-DD): Enter the date for the requested weather data.

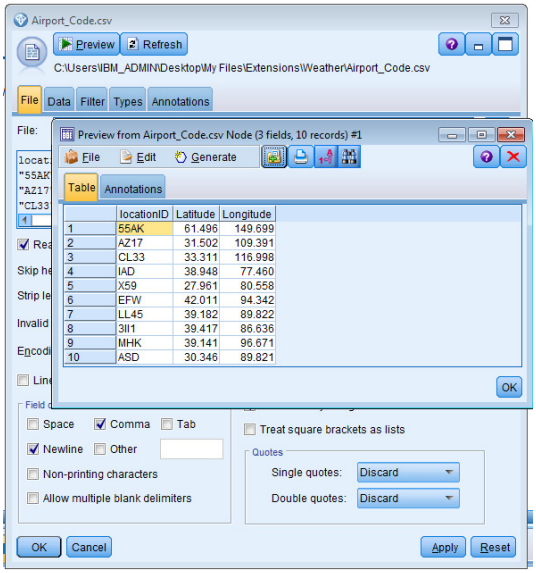


## Example

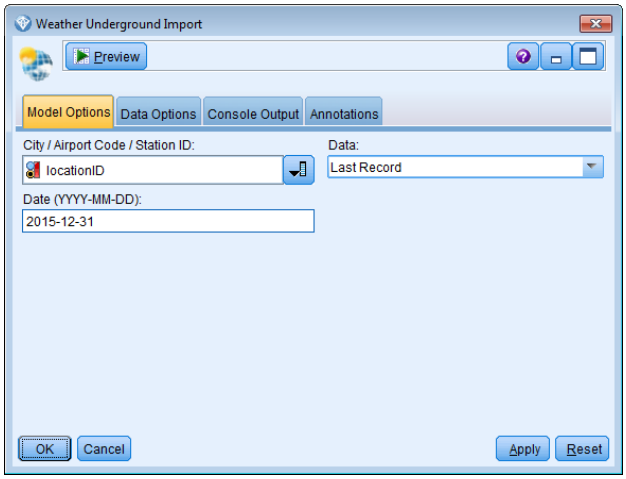
This example will demonstrate how to import weather data for a column of valid locations in a CSV file. The CSV used in this example can be found in the example folder of this extension's GitHub Repository.

**User Input**

1. In the sources palette, add a Var File node to the stream and read Airport_Code.csv. The locationID column of this dataset has the valid location identifier required for this extension.




2. Add the Weather Underground Import extension from the Field Ops palette and connect it to the Var File node. Select locationID for the City/Airport Code/Station ID field, and enter a date in the format YYYY-MM-DD. Finally, choose the type of data to import in the Data field or leave as the default Last Record.



3. From the Output palette, connect a Table to the exension
4. Click on the Table and run the stream


**Results**

If you used the sample text above, your results should match the table below:

Table (18 fields, 26 records) #3

File Edit Generate

Table Annotations

| | locationID | Latitude | Longitude | Time | Temperature | Dew.Point | Humidity | Sea Level Pressure |
|---|---|---|---|---|---|---|---|---|
| 1 | 55AK | 61.496 | 149.699 | $null$ | $null$ | $null$ | $null$ | $null$ |
| 2 | AZ17 | 31.502 | 109.391 | 11:56 PM | 37.000 | 30.900 | 79.000 | 30.000 |
| 3 | CL33 | 33.311 | 116.998 | $null$ | $null$ | $null$ | $null$ | $null$ |
| 4 | IAD | 38.948 | 77.460 | 11:52 PM | 43.000 | 30.900 | 62.000 | 30.150 |
| 5 | X59 | 27.961 | 80.558 | 11:53 PM | 73.000 | 71.100 | 93.000 | 30.110 |
| 6 | EFW | 42.011 | 94.342 | 11:55 PM | 23.400 | 19.200 | 84.000 | 30.310 |
| 7 | LL45 | 39.182 | 89.822 | 11:55 PM | 27.100 | 23.400 | 86.000 | 30.360 |
| 8 | 3II1 | 39.417 | 86.636 | 11:53 PM | 30.000 | 23.000 | 75.000 | 30.370 |
| 9 | MHK | 39.141 | 96.671 | 11:52 PM | 23.000 | 17.100 | 78.000 | 30.510 |
| 10 | ASD | 30.346 | 89.821 | 11:00 PM | 77.000 | 71.600 | 83.000 | 30.150 |
| 11 | 3GM | 43.034 | 86.198 | 11:55 PM | 28.000 | 17.100 | 63.000 | 30.160 |
| 12 | 23MN | 43.641 | 92.022 | 11:55 PM | 18.300 | 14.400 | 85.000 | 30.160 |
| 13 | 09M | 33.991 | 90.078 | $null$ | $null$ | $null$ | $null$ | $null$ |
| 14 | NC49 | 35.528 | 77.197 | 11:55 PM | 57.200 | 55.400 | 94.000 | 30.110 |
| 15 | 09NE | 41.333 | 96.194 | 11:55 PM | 19.200 | 15.300 | 85.000 | 30.360 |
| 16 | 2Q9 | 37.610 | 118.006 | 11:56 PM | 16.000 | 6.100 | 65.000 | 30.240 |
| 17 | 93OH | 41.044 | 81.186 | 11:55 PM | 29.800 | 23.000 | 76.000 | 30.180 |
| 18 | OK62 | 34.249 | 97.209 | 11:55 PM | 35.600 | 32.000 | 87.000 | 30.490 |
| 19 | 60OR | 43.426 | 123.280 | 11:53 PM | 28.000 | 27.000 | 96.000 | 30.320 |
| 20 | PS61 | 41.691 | 79.366 | 11:55 PM | 26.600 | 21.200 | 80.000 | 30.170 |

OK

This weather data can be used to build models or do analysis involving data from different sources.

## Important Links

### Learn

- Learn more about SPSS software.
- Visit developerWorks Business analytics for more technical analytics resources for developers.
- Create your Bluemix account here: https://console.ng.bluemix.net/solutions/watson
- The Comprehensive R Archive Network is the main site for the R project and each R package. The help pages and manuals that are associated with optimx, nlmrt, and Rcgmin are detailed. Numerous references are provided.
- Read "Do I need to learn R?" (Catherine Dalzell, developerWorks, September 2013) to learn why R is a valuable tool for data analytics that was expressly designed to reflect the way that statisticians think and work.
- "Calling R from SPSS" describes how to use R code inside IBM SPSS Modeler 16.
- Read "Create new nodes for IBM SPSS Modeler 16 using R" to learn how to create new extensions easily.

## Discuss

- Visit the IBM SPSS Community to share tips and experiences with other IBM SPSS developers.
- Follow developerWorks on Twitter to be among the first to hear about new resources.