Introduction

* Scenario: Experiment data based on weather temperature for March 2016.
* Data Variables: Temperature values (in decimal) and dates (in string)
* Total number of observation and variables: 1 date column and 10 temperature columns.
* Some research on the background: weather forecasting uses machine learning to predict future weather temperatures and conditions which would be broadcasted to the media.
* Questions related to the data / hypothesis: what is the temperature distribution between the dates and does it corelate to weather conditions (winter vs summer temperatures).

Data Cleaning

* Any cleaning needed: Yes. To remove unnecessary columns to keep the data small and manageable.
* How to clean data: Use row and column filters. Use of comprehensive search and sort for null values. <https://forum.knime.com/t/count-null-value-in-the-column/17429/3>
* How to check whether data is clean: Use missing value column filter.
* Any reason for missing value: Wrong data type or user did not add a value for that column value resulting in a NULL value.