Manual – QAvg v 1.3 Add-in

Jakub Rybacki, jakub.rybacki@gmail.com

Warsaw, 2014

Table of Contents

[Introduction 3](#_Toc406682453)

[Installation 3](#_Toc406682454)

[Keyboard Shortcuts 3](#_Toc406682455)

[Data Aggregation 4](#_Toc406682456)

[Important – Horizontal/Vertical Input/Output 4](#_Toc406682457)

[Important – Limitation 6](#_Toc406682458)

[Listing to Higher Frequency 7](#_Toc406682459)

[Important 7](#_Toc406682460)

[Interpolation 8](#_Toc406682461)

[Important – Removing Interpolation 9](#_Toc406682462)

[Median Stem and Leaf 10](#_Toc406682463)

# Introduction

The aim of this manual is to present QAvg – Excel Aggregation & Interpolation Add-in. As its name indicate, this application calculated quarterly averages, as it was created, however with lapse of time it has grown bigger and bigger.

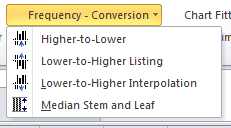
Presently final consumer of this application would receive a flexible machine allowing for:

* Data aggregation to lower frequency by a wide range of functions (sum, average, product)
* Interpolation of low frequency data to higher frequency by a various algorithms including: linear interpolation, polynomial interpolation, spline interpolation

Finally as an additional option Steam and Leaf median were introduced.

# Installation

Installation of QAvg add-in would create following menu on add-ins tab inside the ribbon:



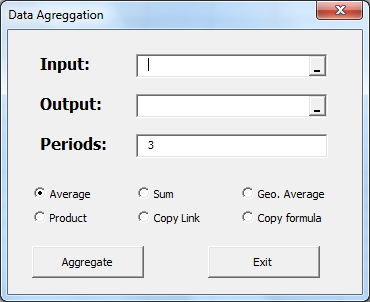
Also following keyboard combinations would be overridden:

## Keyboard Shortcuts

|  |  |
| --- | --- |
| Shortcut | Action |
| CTRL + SHIFT + F6 | Data aggregation to lower frequency |
| ALT + SHIFT + F6 | Data Interpolation |
| CTRL + ALT + F6 | Transcription of low frequency data into higher frequency (one step before interpolation) |

# Data Aggregation

Selection of ***Higher-to-Lower*** option, or pressing **CTRL + SHIFT + F6** combination wouldtrigger following menu:



Selection of aggregating function is quite intuitive, program would select every n-th element based on data inputted in *Periods* field. In case of:

* Average, sum, Product – data would be addressed directly (e.g. AVERAGE($A$2:$A$5)), allowing for copying/moving data to the other worksheet without a fear of data loss.
* Copy Link/Copy Formula – formula/link of every n-th observation would be copied. Similarly like in case of aggregating functions, also link is copied based on direct addressing, on the other hand copy formula creates exact copy of formula in the source cells (which are sensible for copying).

In the *Input* field source data is inputted – please note that add-in allows for selection of multiple series, which could be both row-oriented or column-oriented (macro assumes that data is column-oriented when number of rows is greater than number of columns, oppositely macro assumes that data is row-oriented in case of greater number of columns than rows). It is also important to notice that application would not aggregate incomplete date (e.g. macro would deny calculating result 1Q, if we had only outcomes from January and February).

*Output* field shall contain 2x1 range (or 1x2 if we want to locate data horizontally), where output would be printed. Note that it does not matter how long is output range, every cell from input range, which allows for creation of complete aggregate would be included.

## Important – Horizontal/Vertical Input/Output

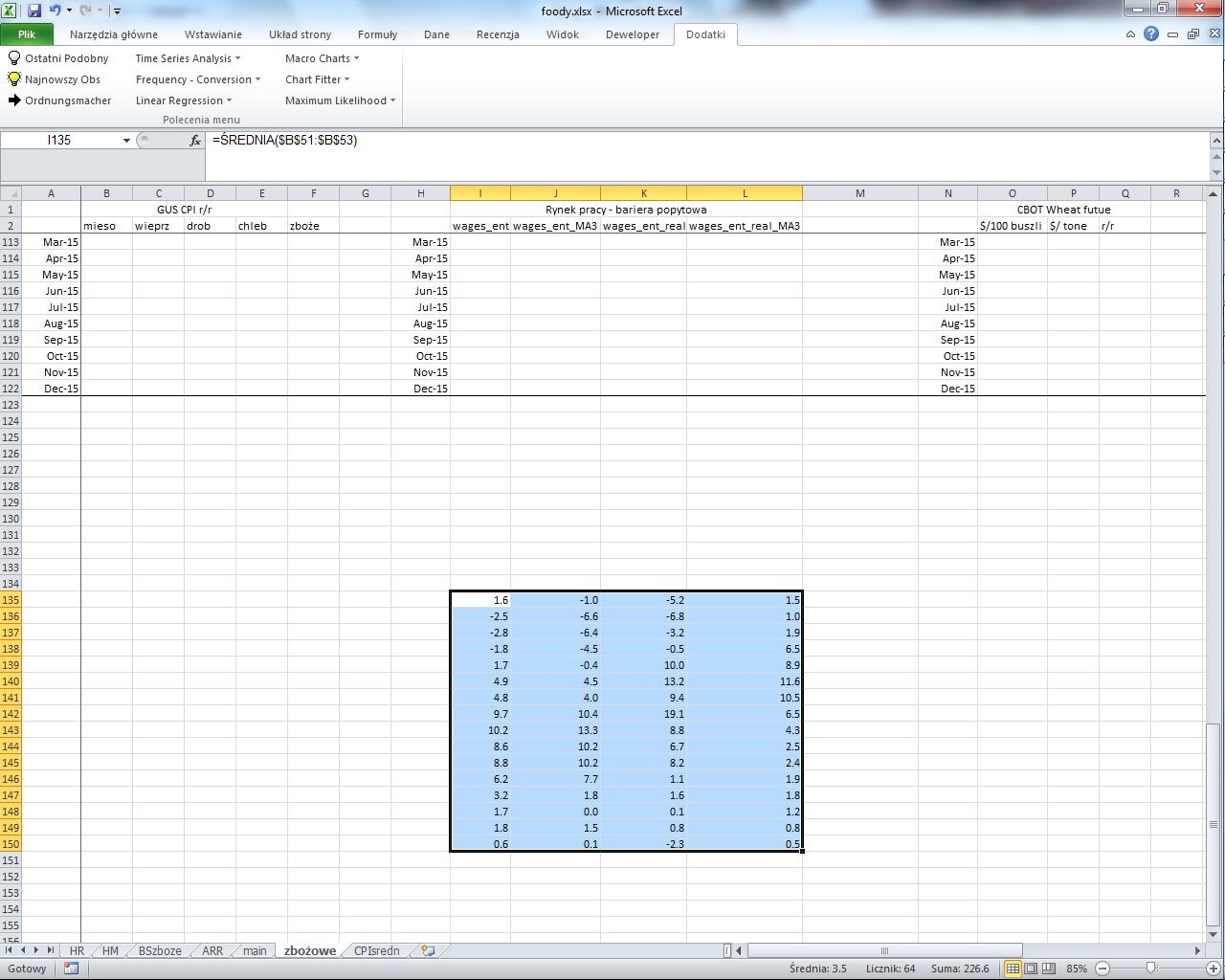
Note that this add-in works with input and output stored vertically (data put in one column) or horizontally (data put in one row). The mechanism of selection between rows/columns is quite simple: software would recognize that input is stored in column(s), when bigger amount of rows (than columns) would be selected. Similarly when bigger amount of columns was selected software would recognize that series are stored in rows.

Similar mechanism works with output field: In case of greater number of rows in output field software would print output in column(s). Otherwise software would print output in row(s).

Let’s look at this mechanism with a help of few examples. First Let’s look at the column-oriented worksheet example. By selecting one cell or one column output would be printed in columns.

|  |  |
| --- | --- |
| **Selected one cell** | **Selected one column** |

Note that every row/column selected in the Input series were aggregate in line with the direction of output field, what is presented at the following picture (4 input series were selected):



Dealing with the same series group, in case where only one row was selected, output is printed in four rows. This was presented at the picture below (right one):

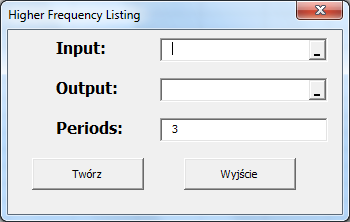
|  |  |
| --- | --- |
| **Selected one row** |  |

## Important – Limitation

Please remind that macro works only in one worksheet (input and output have to be in the same area), fortunately thanks to direct addressing, output can be easily transformed to the other worksheets by cut/paste combination (**Ctrl+X** i **Ctrl+V).**

# Listing to Higher Frequency

Selection of ***Lower-to-Higher Listing*** option or, pressing **CTRL + ALT + F6** would trigger following interface:



Interface works similarly like in case of data aggregation, directly linking lower frequency series into multiplication of number in ***periods*** field. An example is provided below:

|  |  |
| --- | --- |
|  |  |

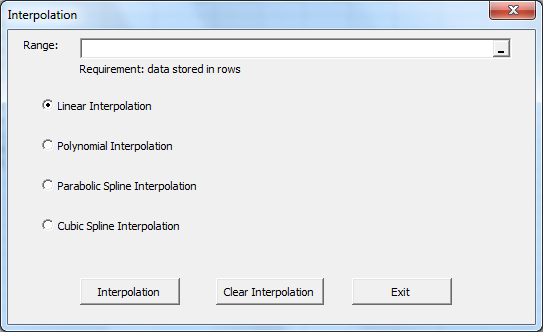
## Important

Interface works similarly like data aggregation thus it also has similar limitations:

* Programs requires that input and output are in the same worksheet
* Direct link allows for cutting/copying and pasting in other fields
* Mechanism of input/output orientation selection (between rows/columns) is create in the same manner like in Data aggregations

# Interpolation

Selection of ***Interpolation*** option, orpressing  **ALT + SHIFT + F6** would trigger following interface:



As it is written in the form program works only with data stored in rows. Data point may be stored as values or formula.

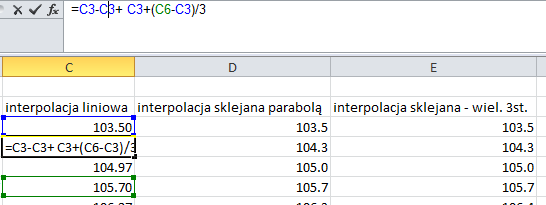
Programs have following algorithms implemented:

* Linear Interpolation – connecting every pair of point with a straight line.
* Polynomial interpolation – fitting n-th root polynomial to n starting points (reliable only in very small samples)
* Parabolic/Cubic Spline Interpolation – fitting field between two points based on their values as well as values of their predecessor/ predecessors.

Example of three interpolation techniques (Linear interpolation – blue, Parabolic spline interpolation – red, Cubic spline interpolation – green)

## Important – Removing Interpolation

Program allows for quick creation/removal of interpolation without concern about safety of written data. This is due to a fact, that module creates special identification sign for interpolated cells: in formula panel it adds and subtracts beginning cell (we may expect no one would write such formula).

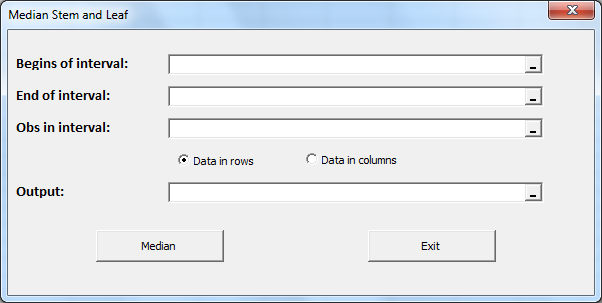


Program adds and removes basic element in the formula to identify that this was a part of interpolation

Manual changes of cell content may result in misleading identification of interpolated cells (it would be recognized as data points).

# Median Stem and Leaf

Selection of ***Median Stem and Leaf*** would trigger following interface:



Variables in the field **Begin/end of interval** should be sorted ascendingly, field ***Obs in interval*** shall contain frequency of observation in the interval.

*Output* field shall contain a single cell.