



IWIN-DAP: An Excel Macro to Analyze CIMMYT International Wheat Trial Data

Zakaria KEHEL, Thomas PAYNE, and CIMMYT Global Wheat Program colleagues

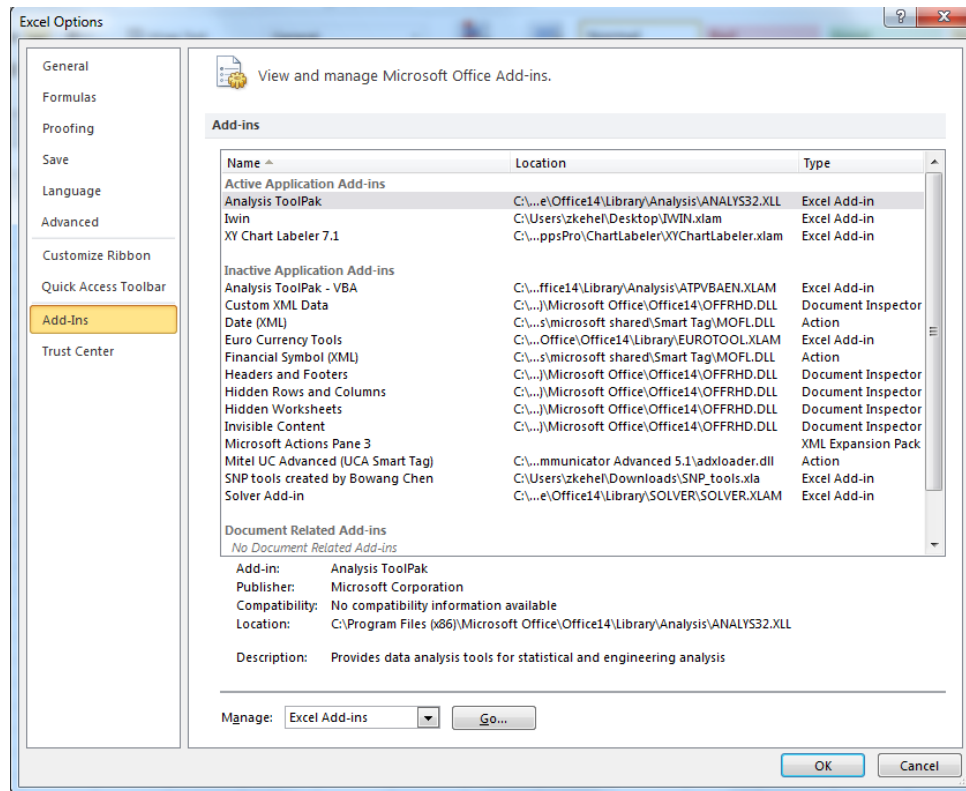
CIMMYT Mexico: z.kehel@cgiar.org, t.payne@cgiar.org

Background

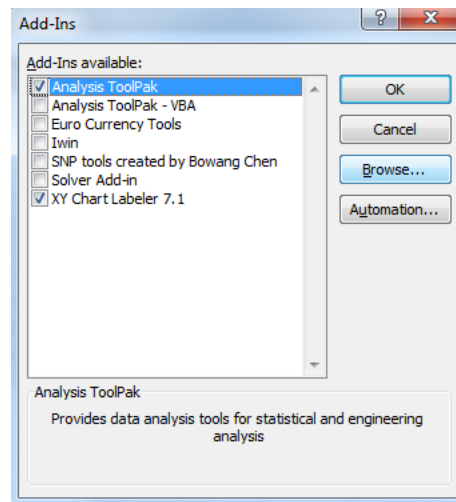
This tutorial explains how to use the IWIN-DAP, which is an Add-In to Microsoft Excel 2010, to analyze data from CIMMYT's International Wheat Information System. IWIN-DAP is an illustration of an alternative, quick and easy approach allowing the use of Excel to analyze multi-location trials from CIMMYT international nurseries. This tutorial explains, step by step, how to install the Add-in, download the data from CIMMYT website, and run a variety of analyses. The macro operates at the single-trait level.

Installing the Macro

Open Excel. Go to File > Options, then Add-Ins, then click Go.



In the resulting window, click Browse.



Navigate to the folder that contains the IWIN-V1 file, and select the file. In the Add-Ins window, select IWIN-V1 and then click OK.

Downloading the Data

Data can be downloaded from the CIMMYT International Wheat Information System website (<http://apps.cimmyt.org/wpgd/Cycles.aspx>), where they are arranged by nursery and year.

Follow the steps in the web page: 1- Choose the nursery. 2- Get the cycle for the nursery and select a year by clicking on it in the list. 3- Search the nursery and select one from the list. 4- Search traits and download the Excel file.

Data are downloaded in Excel format with each trait covered by one worksheet containing the following information:

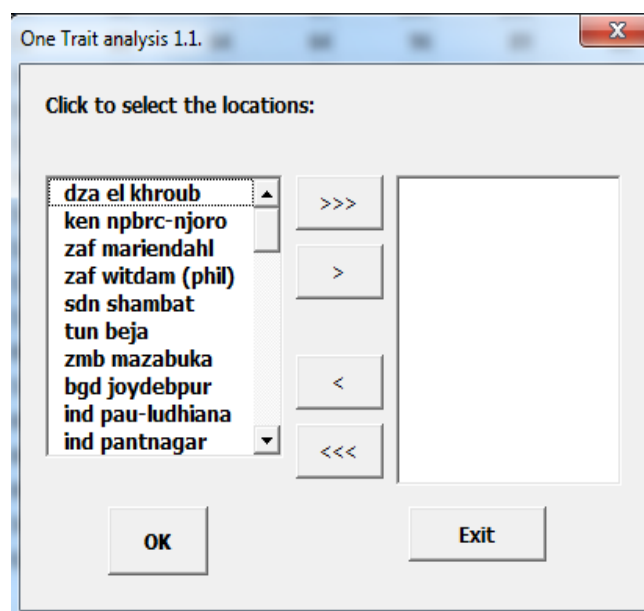
- date of last update (row 1)
- name of nursery (row 3)
- complete name of trait (row 4)
- locations (row 6 , starting at column 4)
- number of original observations / occurrence number for the location (row 7 , starting at column 4)
- unique identifier for the location (row 9 , starting at column 4)
- names of genotypes (column 2, starting at row 10)
- selection history of genotypes (column 3, starting at row 10)
- phenotypic data (column 4, starting at row 10)

See the example below for more details.

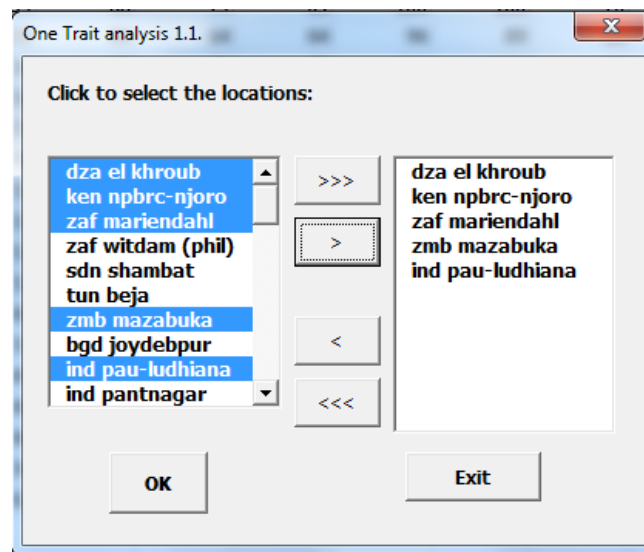
CIMMYT INT. WDMS PGD VER. 1.0		Date : 4/30/2013 11:50:20 AM					
Table of Phenotypic Data							
Trial : 8TH ITYN							
Trait : DAYS_TO_ANTHESIS (days)							
Location	Name	Selection History	el khroub	npbrc-njoro	mariendahl	dam (phil)	sdn shambat
Obs/Occ ->			74/1	75/2	25/3	25/4	75/6
Genotype							tun beja
			[19005]	[12302]	[10004]	[10005]	[12704]
							[19402]
1	YOREME 75	X2802-38N-3M-7N-5M-0Y	143	65	99	95	52
2	NAVOJOA	X2802-38N-3M-6N-6M-0Y	142	64	97	95	52
3	BACUM	X2832-24N-3M-7N-4M-0Y	140	60	95	95	51
4	ARABIAN	X2802-38N-5M-6N-6M-0Y	142	65	96	100	53
5	JUPATECO F 73	II30842-31R-2M-2Y-0M-0MEX	102	67	96	96	53
6	MAPACHE	X2802-F-12M-1N-1M-0Y	142	64	103	97	51
7	MAPACHE	X2802-F-12M-1N-1M-0Y	139	63	103	96	53
8	SETTER	X2802-70N-3M-1N-2M-0Y	139	74	109	110	58
9	KOALA	X2091-100Y-101B-2N-2M-3Y-1M-0Y	153	67	106	105	54
10	MEXICALI C75	CM470-1M-3Y-0M	109	60	99	89	53
11	YOCO RESEL	X1648-RESEL	139	67	104	100	58
12	LINCE	X1648-8N-3M-0Y	141	65	96	95	50
13	RAHUM	X2148-1N-1M-0Y	137	66	96	91	55
14	M1A	X2148-4N-2M-2Y-1M-0Y	142	63	94	89	49
15	SIETE CERROS T66	II8156-1M-2R-4M-0Y	105	65	106	100	53
16	OCTO BULK/BUSH	X7233-32M-3Y-2M-0Y	136	66	103	98	52
17	IGA42/IGA42	X8319-A-3Y-1M-0Y	136	60	89	84	52
18	FS1897	Selection History doesn't exist	137	64	104	95	52
19	BEAGLE 2	X1530-RESEL	134	67	104	109	60
20	CINNAMON	X2802-68N-5M-0N	136	64	97	98	53
21	DRIRA	X7110	147	70	112	115	60
22	UC8825	Selection History doesn't exist	137	69	113	111	66
23	6TA204/BCO90	D71042-51W-0W	141	73	116	117	66
24	BVR-T/ARM	74CT301-6	138	67	103	99	54
25	LOCAL CHECK	Selection History doesn't exist	89	64	97	93	62
MEAN			133.84	65.56	101.48	98.88	54.88
MIN			89	60	89	84	49
MAX			153	74	116	117	66
STD			15.03	3.4	6.41	8.03	4.57

Preparing the Data and Data Analysis

- Open the Excel file and select the worksheet containing the data for the trait to be analyzed.
- Press **Ctrl** + **o**. A window will appear listing all locations available for the selected trait.



- Click on locations in the left-hand column to select them for analysis (to deselect, click again).
- Click **>** to move the selections to the right-hand list. To move all locations to the right-hand list, click **>>>**.
- Locations can be removed from the right-hand list by selecting them and clicking **<** or clicking **<<<** to remove them all.
- Click **OK** to start analyzing.



- Choose a check by clicking on the name of a genotype or its selection history and then clicking **OK**.

	A	B	C	D	E	F	G	H
1	Name	Selection History	dza el kh	ken npbr	zaf marier	zaf witdar	sdn gezira	sdn sham
2	YOREME 75	X2802-38N-3M-7N-5M-0Y	2.71	6.47	2.21	2.15	0.45	0.68
3	NAVOJOA	X2802-38N-3M-6N-6M-0Y	2.9	3.91	3.38	1.74	0.37	0.36
4	BACUM	X2832-24N-3M-7N-4M-0Y	4.03	6.41	2.23	1.79	0.38	0.16
5	ARABIAN	X2802-38N-5M-6N-6M-0Y	2.84	6.04	2.87	2.78	0.48	0.52
6	JUPATECO F 73	II30842-31R-2M-2Y-0M-0MEX	3.57	1.83	0.29	1.78	0.76	0.48
7	MAPACHE	X2802-F-12M-1N-1M-0Y	3.35	7.09	2.81	1.56	0.46	1.96
8	MAPACHE	X2802-F-12M-1N-1M-0Y	3.94	6.59	2.57	1.44	0.36	0.72
9	SETTER	X2802-	3.5	1.83	2.06	0.52	0.54	
10	KOALA	X2091-	4.75	1.78	2.2	0.53	1.74	
11	MEXICALI C75	CM470	3.75	0.28	0.43	0.63	0.12	
12	YOCO RESEL	X1648-	5.38	2.26	1.57	0.53	0.27	
13	LINCE	X1648-	6.2	2.78	2.65	0.41	0.76	
14	RAHUM	X2148-	6.87	2.32	1.99	0.57	0.41	
15	M1A	X2148-4N-2M-2Y-1M-0Y	3.92	6.76	2.56	1.98	0.41	0.93
16	SIETE CERROS T66	II8156-1M-2R-4M-0Y	4.49	0.98	1.41	0.44	0.47	0.15
17	OCTO BULK/BUSH	X7233-32M-3Y-2M-0Y	3.01	6.06	2.63	2.6	0.31	0.26
18	IGA42/IGA42	X8319-A-3Y-1M-0Y	2.15	4.95	1.26	1.27	0.29	0.08
19	FS1897	Selection History doesn't exist	2.81	5.91	2.53	2.02	0.5	0.68
20	BEAGLE 2	X1530-RESEL	4.01	6.28	2.54	3.24	0.61	1.27
21	CINNAMON	X2802-68N-5M-0N	2.39	5.64	2.12	1.01	0.53	0.27
22	DRIRA	X7110	2.84	6.3	1.77	1.63	0.82	1.19
23	UC8825	Selection History doesn't exist	3.35	3.61	1.2	1.53	0.56	1.26
24	6TA204/BCO90	D71042-51W-0W	4.2	5.5	1.07	1.27	0.65	1.17
25	BVR-T/ARM	74CT301-6	4.26	7.51	1.92	2.82	0.49	0.73
26	LOCAL CHECK	Selection History doesn't exist	4.07	4.7	3.37	2.86	0.59	0.1
27	MEAN		3.31	5.32	2.08	1.87	0.51	0.67
28	MIN		2.15	0.98	0.28	0.43	0.29	0.08
29	MAX		4.49	7.51	3.38	3.24	0.82	1.96
30	SDev		0.67	1.62	0.82	0.71	0.13	0.51
31	SUM		82.80	132.99	51.99	46.81	12.68	16.81
32	COUNT		25.00	25.00	25.00	25.00	25.00	25.00
33	VARIANCE		0.45	2.64	0.66	0.51	0.07	0.26

Input

Choose the check.

\$A\$26

OK Cancel

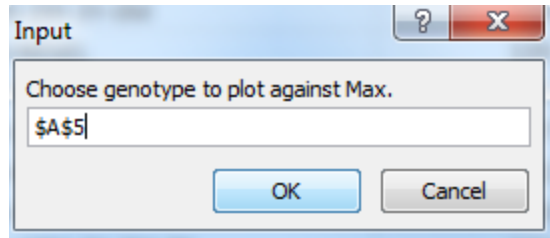
- Click **Yes** to choose another check or **No** to continue the analysis.

One Trait Analysis

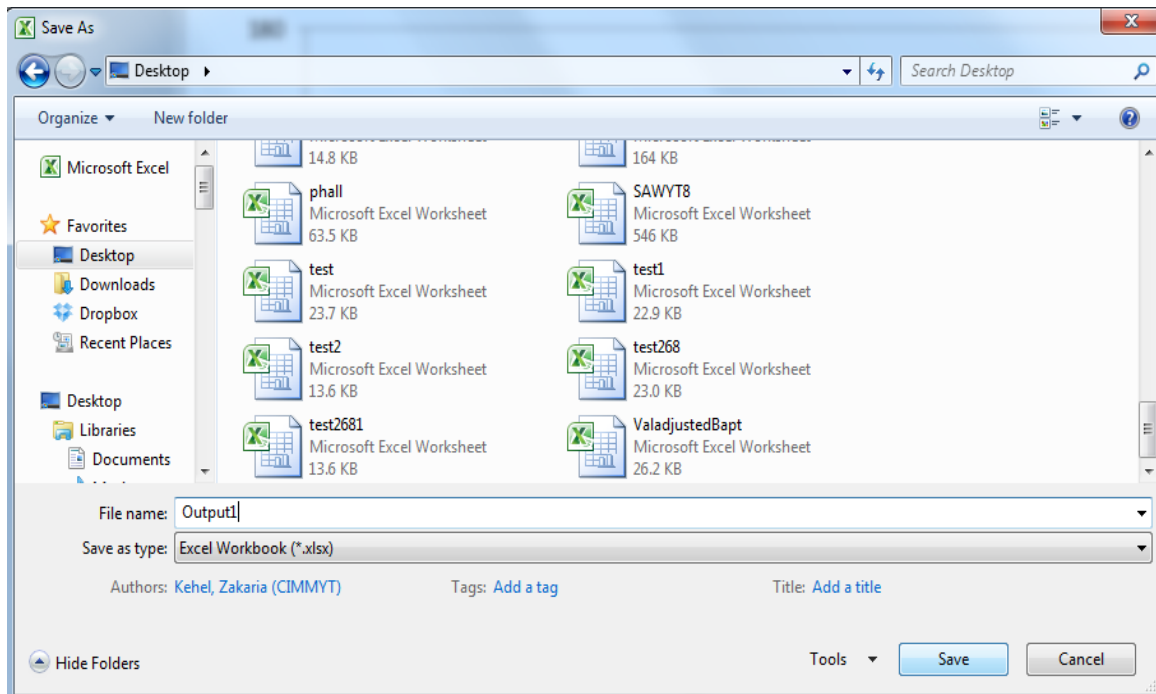
Do you want to choose another check? (Click No to Continue analysis.)

Yes No

- Choose a genotype to plot against the check and the maximum and average of the trait in each location by clicking on the genotype name, its selection history, or its rank, then clicking **OK**.



- A window will appear with options for saving the output file and exit Excel. Choose a folder, give a name to the output file, and click Save.



Type of analysis IWIN-DAP is carrying out

The IWIN-DAP processes one trait at a time and runs the following analyses on selected locations:

- Descriptive statistics: Mean maximum, minimum, standard deviation, variance, rank, and count for locations and genotypes.
- Percentage of checks: value of trait of a genotype at a given location as a percentage of the value of the chosen checks.
- Summary of checks: average percentage of all chosen checks and a plot of these percentages.
- Correlation: Pearson correlation between selected locations.
- ANOVA:

One-factor ANOVA tests the null hypothesis that location does not influence the trait. In the resulting ANOVA table, a P-value for “between locations” of less than 0.01 means that there is less than a 1% chance that the difference between locations is due to chance, and that therefore, we can reject the null hypothesis and say that there is a significant difference between locations. Two-factor ANOVA divides the total variation into variations due to genotype, location, and their interactions. Two tables for one and two factors ANOVA are produced.

- Trait stability:

The stability test runs the Eberhart and Russel 1966 stability indices (regression coefficient, B , and deviation mean squares, SD^2) and draws a scatter plot of average trait versus SD^2 . The trait values of a genotype in a given location are regressed on the mean of the location. The deviation mean squares measure describes the contribution of a genotype to the genotype by environment interaction (G.E). interactions. SD^2 is considered a stability parameter, as it is highly related to the remaining unpredictable part of variability of a genotype across locations. A genotype is stable when its deviation from regression on the environmental index is small.

- Plot maximum, average for each location, check, and chosen genotype.
- Eigenvalues and percentages of explained variance from principal components analysis (PCA) analysis on genotypes and locations.
- PCA scores for genotypes and locations and a bi-plot.

Contact Information

To report errors, concerns or offer other comments, please contact Zak Kehel (z.kehel@cgiar.org) or Tom Payne (t.payne@cgiar.org).