



LARISSA WELTON

Orcid ID: 0000-0002-6374-8273

Natural History Museum, London

SPNHC 2022: Back to Basics: Museum
Techniques Skillshare
Wed, 8th June 2022

Presentation DOI:
[10.6084/m9.figshare.19919555](https://doi.org/10.6084/m9.figshare.19919555)

USING EXCEL MACROS FOR DATA QUALITY ASSURANCE: LESSONS LEARNED FROM A CODING BEGINNER

AGENDA

Why do we check data?

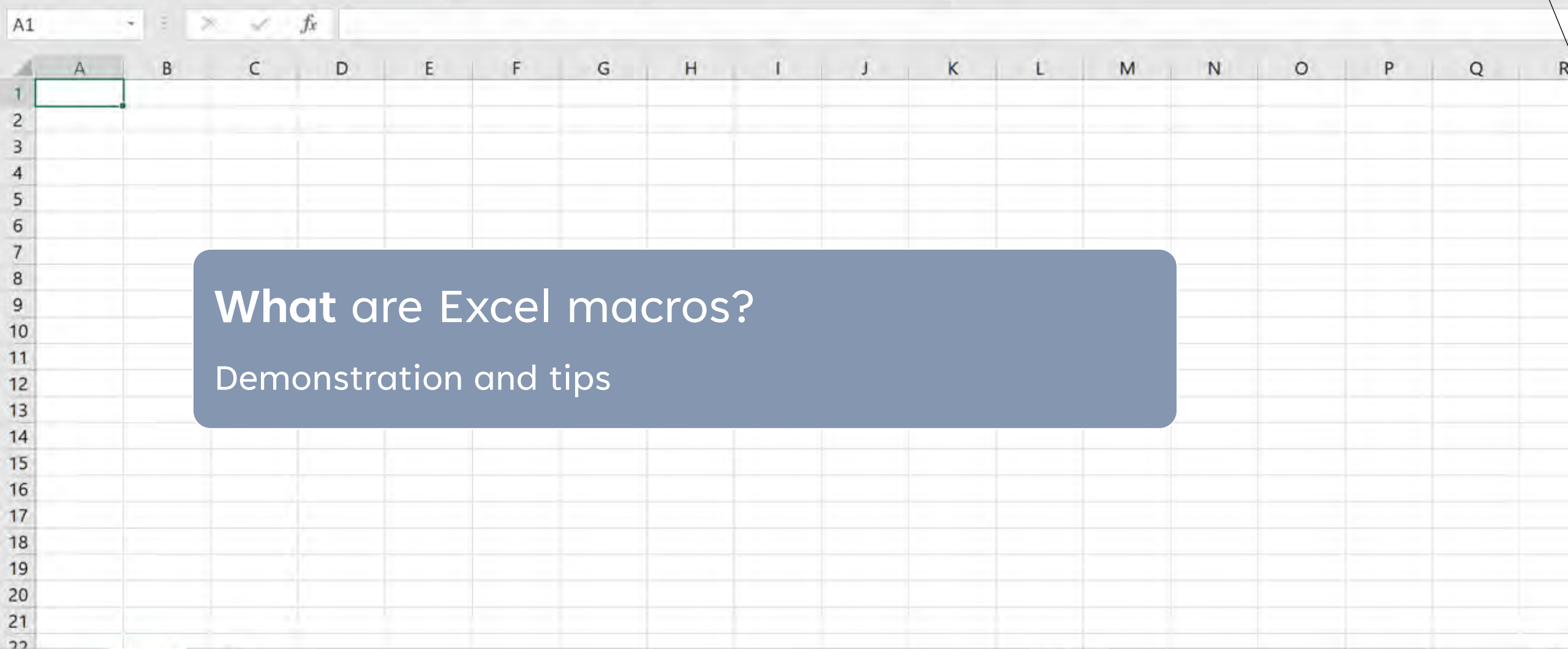
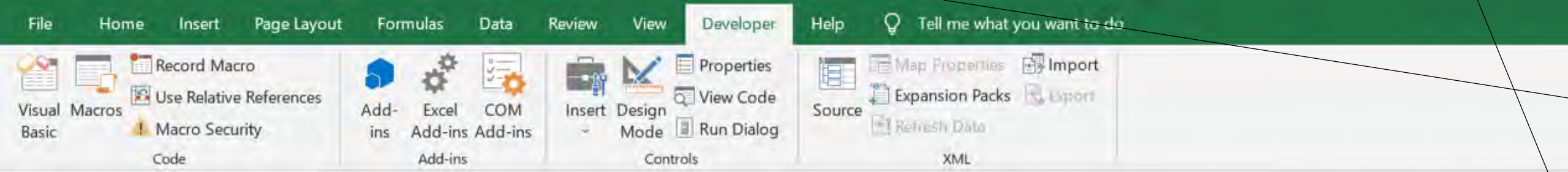
Instances you check data quality

What are Excel macros?

Demonstration and tips

How can we use them?

Case study from NHM: Slide digitisation



What are Excel macros?
Demonstration and tips

Microsoft Visual Basic for Applications

File Edit View Insert Format Debug Run Tools Add-Ins Window Help

Ln 2, Col 1

Project - VBAProject

VBAProject (FUNCRES.XLAM)
VBAProject (test workbook.xlsm)

- Microsoft Excel Objects
 - Sheet1 (Sheet1)
 - Sheet3 (Sheet4)
 - ThisWorkbook
- Modules
 - Module1

Properties - Sheet3

Sheet3 Worksheet

Alphabetic Categorized

(Name)	Sheet3
DisplayPageBreaks	False
DisplayRightToLeft	False
EnableAutoFilter	False
EnableCalculation	True
EnableFormatConditions	True
EnableOutlining	False
EnablePivotTable	False
EnableSelection	0 - xlNoRestrictions
Name	Sheet4
ScrollArea	
StandardWidth	8.09
Visible	-1 - xlSheetVisible

test workbook.xlsm - Module1 (Code)

(General) Filenames

```
Sub Filenames()  
'  
' Filenames Macro  
'  
'  
Range("A1:A535").Select  
Selection.TextToColumns Destination:=Range("A1"), DataType:=xlDelimited, _  
TextQualifier:=xlDoubleQuote, ConsecutiveDelimiter:=False, Tab:=True, _  
Semicolon:=False, Comma:=False, Space:=False, Other:=True, OtherChar _  
:="\", FieldInfo:=Array(Array(1, 1), Array(2, 1), Array(3, 1), Array(4, 1), Array(5, _  
1), Array(6, 1), Array(7, 1), Array(8, 1), Array(9, 1), Array(10, 1), Array(11, 1), Array(12 _  
, 1), Array(13, 1)), TrailingMinusNumbers:=True  
Columns("A:L").Select  
Range("L1").Activate  
Selection.Delete Shift:=xlToLeft  
Range("A1").Select  
End Sub
```

TIPS AND TRICKS

- If you can do it once, you can make a macro for it
- Best for actions you do repeatedly

- If you select a range while recording, it will select the same range every time
- Creating a new sheet – rename “ActiveSheet”
- Running a macro can’t be undone
- Use ‘ to write yourself notes
- Save as a macro-enabled workbook (.xlsm)
- Online resources abound!

CASE STUDY: SLIDE DIGITISATION AT THE NHM





Five barcodes encode information about the specimen



014231204_833514_1610664_157853_NonType.JPG



014231204_additional.JPG

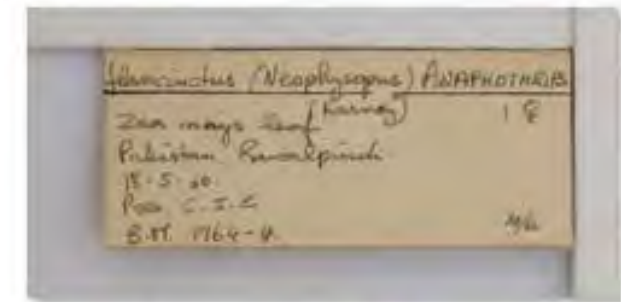
Software reads this and renames the file

Name	Status	Date	Type
014231204_833514_1610664_157853_NonType.JPG	☁	19/10/2021 14:56	JPG File
014231205_833514_1610664_157853_NonType.JPG	☁	19/10/2021 14:56	JPG File
014231206_833514_1610664_157853_NonType.JPG	☁	19/10/2021 14:56	JPG File
014231207_833514_1610664_157853_NonType.JPG	☁	19/10/2021 14:56	JPG File
014231208_833514_1610664_157853_NonType.JPG	☁	19/10/2021 14:56	JPG File
014231209_833514_1610664_157853_NonType.JPG	☁	19/10/2021 14:56	JPG File
014231210_833514_1610664_157855_NonType.JPG	☁	19/10/2021 14:56	JPG File
014231211_833514_1610664_157855_NonType.JPG	☁	19/10/2021 14:56	JPG File
014231212_833514_1610664_157855_NonType.JPG	☁	19/10/2021 14:56	JPG File
014231213_833514_1610664_157855_NonType.JPG	☁	19/10/2021 14:56	JPG File
014231214_833514_1610664_157855_NonType.JPG	☁	19/10/2021 14:56	JPG File
014231215_833514_1610664_157855_NonType.JPG	☁	19/10/2021 14:56	JPG File
014231216_833514_1610664_157894_NonType.JPG	☁	19/10/2021 14:56	JPG File
014231217_833514_1610664_7900975_NonType.JPG	☁	19/10/2021 14:56	JPG File
014231218_833514_1610664_157882_NonType.JPG	☁	19/10/2021 14:56	JPG File
014231219_833514_1610664_157662_NonType.JPG	☁	19/10/2021 14:56	JPG File
014231220_833514_1610664_157662_NonType.JPG	☁	19/10/2021 14:56	JPG File
014231221_833514_1610664_157662_NonType.JPG	☁	19/10/2021 14:56	JPG File
014231222_833514_1610664_157662_NonType.JPG	☁	19/10/2021 14:56	JPG File
014231224_833514_1610664_157662_NonType.JPG	☁	19/10/2021 14:56	JPG File

532 items



✔ 014231204_833514_1610664_157853_NonType.JPG



✔ 014231204_additional.JPG

Spotting errors in filenames means better data in CMS

Software reads this and renames the file

Clipboard Font Alignment Number Styles Cells Editing Analysis Sensitivity

Calibri 11 Bold Italic Underline Font Color Background Color

General Number

Conditional Formatting Format as Table Cell Styles

Insert Delete Format

Sort & Filter Find & Select

Analyze Data Sensitivity

B4 012816892_833498_1611107_157865_Type.JPG

	A	B	C	D	E	F	G	H	I
1	Please use this Instruction Sheet to help you								
2									
3	Please insert your filenames here:	Final image filenames	Reverse image filenames						
4		012816892_833498_1611107_157865_Type.JPG	012816892_additional.JPG						
5		014231204_833514_1610664_157853_NonType.JPG	014231204_additional.JPG						
6		014231205_833514_1610664_157853_NonType.JPG	014231205_additional.JPG						
7		014231206_833514_1610664_157853_NonType.JPG	014231206_additional.JPG						
8		014231207_833514_1610664_157853_NonType.JPG	014231207_additional.JPG						
9		014231208_833514_1610664_157853_NonType.JPG	014231208_additional.JPG						
10		014231209_833514_1610664_157853_NonType.JPG	014231216_additional.JPG						
11		014231210_833514_1610664_157855_NonType.JPG	014231217_additional.JPG						
12		014231211_833514_1610664_157855_NonType.JPG	014231263_additional.JPG						
13		014231212_833514_1610664_157855_NonType.JPG	014231273_additional.JPG						
14		014231213_833514_1610664_157855_NonType.JPG	014231276_additional.JPG						
15		014231214_833514_1610664_157855_NonType.JPG	014231277_additional.JPG						
16		014231215_833514_1610664_157855_NonType.JPG	014231278_additional.JPG						
17		014231216_833514_1610664_157894_NonType.JPG	014231279_additional.JPG						
18		014231217_833514_1610664_7900975_NonType.JPG	014231280_additional.JPG						
19		014231218_833514_1610664_157882_NonType.JPG	014231281_additional.JPG						
20		014231219_833514_1610664_157662_NonType.JPG	014231282_additional.JPG						
21		014231221_833514_1610664_157662_NonType.JPG	014231283_additional.JPG						

FINDING DUPLICATES AND MISSING BARCODE NUMBERS

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P		
1	UID	Duplicates	Missing	LOC	Thysan	Drawer	TAX	Species	Species	SITE	Country	Country	TYPE	Type iss	FILE	All Issu		
17	14231218	CORRECT	CORRECT										CORRECT	NonType	CORRECT	JPG	CORRECT	
18	14231219	CORRECT	ISSUE	=IF(A19-A18<>1,"ISSUE","CORRECT")										CORRECT	NonType	CORRECT	JPG	ISSUE
19	14231221	CORRECT	CORRECT										CORRECT	NonType	CORRECT	JPG	CORRECT	
20	14231222	CORRECT	CORRECT										CORRECT	NonType	CORRECT	JPG	CORRECT	
21	14231223	CORRECT	CORRECT	833514	23	CORRECT	1610664	Anaphoth	CORRECT	157662	Australia	CORRECT	NonType	CORRECT	JPG	CORRECT		
22	14231224	CORRECT	CORRECT	833514	23	CORRECT	1610664	Anaphoth	CORRECT	157662	Australia	CORRECT	NonType	CORRECT	JPG	CORRECT		
23	14231225	CORRECT	CORRECT	833514	23	CORRECT	1610664	Anaphoth	CORRECT	157662	Australia	CORRECT	NonType	CORRECT	JPG	CORRECT		
24	14231226	CORRECT	CORRECT	833514	23	CORRECT	1610664	Anaphoth	CORRECT	157662	Australia	CORRECT	NonType	CORRECT	JPG	CORRECT		
25	14231227	CORRECT	CORRECT	833514	23	CORRECT	1610664	Anaphoth	CORRECT	157662	Australia	CORRECT	NonType	CORRECT	JPG	CORRECT		
26	14231228	CORRECT	CORRECT	833514	23	CORRECT	1610664	Anaphoth	CORRECT	157662	Australia	CORRECT	NonType	CORRECT	JPG	CORRECT		
27	14231229	CORRECT	CORRECT	833514	23	CORRECT	1610664	Anaphoth	CORRECT	157662	Australia	CORRECT	NonType	CORRECT	JPG	CORRECT		
28	14231230	CORRECT	CORRECT	833514	23	CORRECT	1610664	Anaphoth	CORRECT	157662	Australia	CORRECT	NonType	CORRECT	JPG	CORRECT		
29	14231231	CORRECT	CORRECT	833514	23	CORRECT	1610664	Anaphoth	CORRECT	157662	Australia	CORRECT	NonType	CORRECT	JPG	CORRECT		
30	14231232	CORRECT	CORRECT	833514	23	CORRECT	1610664	Anaphoth	CORRECT	157662	Australia	CORRECT	NonType	CORRECT	JPG	CORRECT		
31	14231233	CORRECT	CORRECT	833514	23	CORRECT	1610664	Anaphoth	CORRECT	157662	Australia	CORRECT	NonType	CORRECT	JPG	CORRECT		
32	14231234	CORRECT	CORRECT	833514	23	CORRECT	1610664	Anaphoth	CORRECT	157662	Australia	CORRECT	NonType	CORRECT	JPG	CORRECT		
33	14231235	CORRECT											Australia	CORRECT	NonType	CORRECT	JPG	CORRECT
34	14231236	ISSUE											Australia	CORRECT	NonType	CORRECT	JPG	ISSUE
35	14231236	CORRECT											Australia	CORRECT	Type	CORRECT	JPG	CORRECT
36	14231237	CORRECT	CORRECT	833514	23	CORRECT	1610664	Anaphoth	CORRECT	157662	Australia	CORRECT	NonType	CORRECT	JPG	CORRECT		

A STEP FURTHER

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	UID	Duplicates	Missing	LOC	Thysan	Drawer	TAX	Species	Species	SITE	Countr	Countr	TYPE	Type iss	FILE	All Issu
68	14231269	CORRECT	CORRECT		Not found	ISSUE	633875	Chaetana	CORRECT	157662	Australia	CORRECT	NonType	CORRECT	JPG	ISSUE
219	14231420	CORRECT	CORRECT	833498	29	CORRECT	1611046	Chaetana	CORRECT		Not found	ISSUE	NonType	CORRECT	JPG	ISSUE
330	14231531	CORRECT	CORRECT	8385998	Not found	ISSUE	1611085	Charassot	CORRECT	157663	Costa Rica	CORRECT	Type	CORRECT	JPG	ISSUE
331	14231532	CORRECT	CORRECT	833498	29	CORRECT	179852085	Not found	ISSUE	157663	Costa Rica	CORRECT	Type	CORRECT	JPG	ISSUE
332	14231533	CORRECT	CORRECT	833498	29	CORRECT	1611085	Charassot	CORRECT	20145663	Not found	ISSUE	Type	CORRECT	JPG	ISSUE
352	14231553	CORRECT	CORRECT	833498	29	CORRECT	632584	Chirothrip	CORRECT	0	Not found	ISSUE	NonType	CORRECT	JPG	ISSUE
387	14231588	CORRECT	CORRECT	833498	29	CORRECT	632584	Chirothrip	CORRECT	157798	Kenya	CORRECT	ParaType	ISSUE	JPG	ISSUE

1	IRN	Taxonomy
2	7804933	Charassothrips -- Thysanoptera; Insecta
3	1866869	Phlaeothripinae -- Thysanoptera; Insecta
4	2015442	Merothripidae -- Thysanoptera; Insecta
5	2015513	Thripidae -- Thysanoptera; Insecta
6	2015519	?merothripidae -- Thysanoptera; Insecta
7	2015601	?aelothripidae -- Thysanoptera; Insecta
8	2015470	?aelothripidae -- Thysanoptera; Insecta
9	1607439	Hoplothrips ulmi -- Phlaeothripinae; Phlaeothripidae; Thysanoptera; Insecta
10	1611181	Chirothrips aculeatus -- Thripinae; Thripidae; Thysanoptera; Insecta
11	1611180	Thrips aculeatus -- Thripinae; Thripidae; Thysanoptera; Insecta
12	1614322	Thrips orientalis -- Thripinae; Thripidae; Thysanoptera; Insecta
13	1606663	Haplothrips rubromarginalis -- Phlaeothripinae; Phlaeothripidae; Thysanoptera; Insecta
14	1613331	Rhamphothrips karnyi -- Thripinae; Thripidae; Thysanoptera; Insecta
15	1607438	Thrips ulmi -- Phlaeothripinae; Phlaeothripidae; Thysanoptera; Insecta
16	1603420	Mesandrothrips inquilinus -- Phlaeothripinae; Phlaeothripidae; Thysanoptera; Insecta
17	1603421	Neoheegeria fuscicornis -- Phlaeothripinae; Phlaeothripidae; Thysanoptera; Insecta
18	1605311	Arrhenothrips lewisi -- Phlaeothripinae; Phlaeothripidae; Thysanoptera; Insecta
19	1606296	Thrips aculeatus -- Phlaeothripinae; Phlaeothripidae; Thysanoptera; Insecta
20	1613553	Scolothrips pallidus -- Thripinae; Thripidae; Thysanoptera; Insecta
21	1614068	Thrips bianchi -- Thripinae; Thripidae; Thysanoptera; Insecta

```
=XLOOKUP(G:G,Tax!A:A,Tax!B:B,"Not found")
```

	A	B	C	D
1	Reverse UID	Additional	Is Reverse UID in Final UID list?	Issues
62	14231546	additional.JPG		1 CORRECT
63	14231547	additional.JPG		1 CORRECT
64	14231636	additional.JPG		1 CORRECT
65	14231637	additional.JPG		0 ISSUE
66	14231638	additional.JPG		1 CORRECT
67	14231639	additional.JPG		1 CORRECT
68	14231640	additional.JPG		1 CORRECT
72	14231644	additional.JPG		1 CORRECT
73	14231645	additional.JPG		1 CORRECT
74	14231646	additional.JPG		1 CORRECT
75	14231647	additional.JPG		1 CORRECT
76	14231648	additional.JPG		1 CORRECT
77	14231649	additional.JPG		1 CORRECT
78	14231650	additional.JPG		1 CORRECT
79	14231651	additional.JPG		1 CORRECT
80	14231652	additional.JPG		1 CORRECT
81	14231653	additional.JPG		1 CORRECT

```
=COUNTIF(FinalCalcTable[UID],A65)
```


A STEP FURTHER

	C	D	E	F	G	H	I	J	K	L	M	N
1												
2		Type Status ▾	Number of Specimens		Country ▾	Number of Specimens		Thysanoptera Drawer Number ▾	Number of Specimens		Species name ▾	Number of Specimens
3		NonType	482		Australia	79		23	60		0	1
4		ParaType	1		Belgium	2		29	472		Anaphothrips sudanensis Trybor	60
5		Type	51		Brazil	2		Not found	2		Chaetanaphothrips clarus (Mou	73
6		(blank)			Brunei	1		Grand Total	534		Chaetanaphothrips machili Hoo	2
7		Grand Total	534		China	3					Chaetanaphothrips orchidii (Mo	75
8					Colombia	15					Chaetanaphothrips Priesner, 19:	10
9					Costa Rica	58					Chaetanaphothrips signipennis (23
10					Cyprus	4					Chaetanaphothrips theiperdus (12
11					Dominica	22					Chaetisothrips chavesi Mound &	1
12					England	15					Chaetisothrips gardeniae (Craw	16
13					Ethiopia	6					Chaetisothrips reticulatus (Craw	2
14					Fiji	7					Chaetisothrips striatus (Hood, 1	1
15					France	20					Charassothis in comparabilis (L	2
16					Germany	7					Charassothis macroseta Gold:	15
17					Greece	1					Charassothis piperaffinis (Joh:	22
18					Grenada	7					Charassothis urospathae Hoo	18
19					Guadeloupe	3					Chilothrips pini Hood, 1916 -- Th	4
20					Honduras	3					Chirothrips aculeatus Bagnall, 1'	62
21					Hungary	2					Chirothrips africanus Priesner, 1	14
22					India	14					Chirothrips ah Girault, 1929 -- Th	25
23					Indonesia	3					Chirothrips Haliday, 1836 -- Thri	95
24					Italy	1					Not found	1
25					Kenya	34					Grand Total	534
26					Kiribati	1						
27					Malaysia	21						

Input

Report

Finalcalc

ReverseCalc

Loc

Tax

Site





RESULTS

Errors found	342 mistakes found using the macro 192 in one event; 150 during daily checks Average of 0.67 errors caught per digitiser per day
Time saved	45.5 minutes saved per digitiser per day 5.5 minutes to do duplicate & missing checks. 40 minutes to do additional checking
Extra specimens imaged	168 hours saved in project 9432 extra specimens could be imaged in the time saved
Cost-benefit	£2,985.81 net saving, my time spent on the macro factored in

RESULTS

Description	Value	Category
SC1 digitiser salary - Developer (FEC)	£35,735.00	Fixed
Working days per year	220	Fixed
Working hours per day	7.2	Fixed
Cost per digitiser per minute	£0.38	Calculated
Minutes saved per day by using macro	45.5	Fixed
Cost saved per day per digitiser using macro	£17.11	Calculated
Person days worked on Thysanoptera project using macro	222	Variable
Cost saved by using macro over whole project	£3,797.97	Calculated
Estimated hours taken to develop and document macro	36	Fixed
Estimated cost to develop and document macro	£812.16	Calculated
Net cost saved over project	£2,985.81	Calculated

14246-4017
枯树枝
云南勐腊县
1988.11.17
卓运发



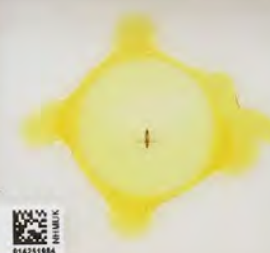
Dinothrips spinosus
(Schmutz)
det. Han Yun-fa

Selwyn & Davis
Danzon
K. DENCAL.



TEREBRANTIA.
Physothrips 1916-5
leproyi Bagn.
Co-TYPES 288
Flowers of tea
Feb. 6. 1909 (leproyi)

NO. 1582 ♀
Isoneurothrips orientalis
Bagn.
BM 1972-401
FLOYD ANDRE



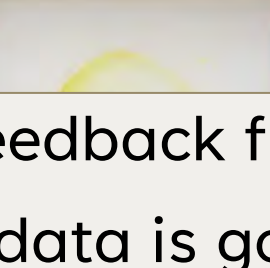
INDIA
Coimbatore
HOST Jasmine
Flowers
DATE 21-23
COLL. R. Ayyar

Haplothrips tritici Kurdj.
1966



Wheat
IRAQ: Arbil

MEXICO



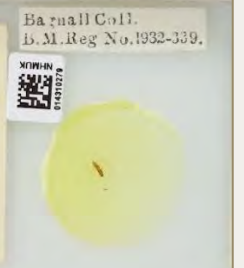
THYSANOPTERA

THYSANOPTERA



Palmyra, Ill.
Macoupin Co
July 31, 1963
Sm. Br. STARR
JEWELWEED
ILL. NAT. HIST. SUR.
Balsam

ITALY
Utrio del Cavallo
(Vesuvio)
Robinia 20.v. 1917
C. Minozzi leg.



- Good feedback from team
- Assured data is good quality
- Especially important as use for newer team members
- Saves time = saves money
- Applied to future projects – Ceratopogonidae slides, Micro moths, Voříšek Chrysomelidae



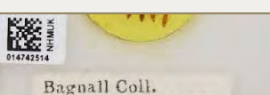
Thrips dilatatus
Bagn.

Eupathithrips ♀
affinis Bagn.



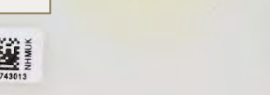
19. vi. 1983
LAM. 1728

Balsam
J. DOUGLAS HOOD
Bur. Biol. Surv., U.S. Dept. Agr.



Leaf galls of *Ficus nitida*.

de Paris



Haplothrips juncorum, Bagn.
BM. 1933. 602.
British Museum (Nat. Hist.).

Bregmatothrips iridis
Watson



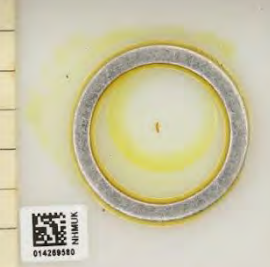
Washington, D.C.
10-1-25
Iris leaves from Turnbridge Wells, Eng.
R. G. Cogswell

Nesothrips propinquus ♀
(Bagn.)



Bird's nest (Raven)
AUSTRALIA: N.S.W.
Lake George
7. ii. 1967.
L.A. Mound 710

Thrips major Uzel.
1 ♀
G. D. Morrison det.
Sept. 1938.



Krios River
of Kibani
Cyprus.
G. A. Mavroustakis
Sept., 1937.



TAKEAWAYS

- Introduction to what Excel macros are and how to use them
- Case study at NHM showed increase in efficiency and data quality
- Even non-experts can create a custom tool for data quality assurance
- Give it a go!



THANK YOU

Questions?

l.welton@nhm.ac.uk