



Rockwell Automation Document

Readme Document



Table of Contents

Section 0: Pre-requisites	3
Section 1: Installation.....	5
Section 2: Configuration	5
Section 3: Execution.....	5
Section 4: Post Execution Compilation	6
Section 5: Cell Formulae	10

Section 0: Pre-requisites

To work with this framework, following setup has to be complete

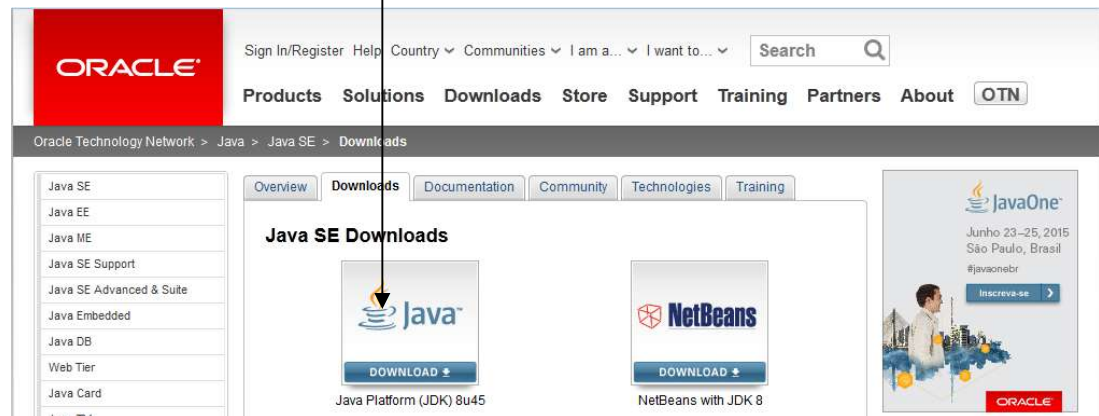
Chrome browser has to be installed in the machine

JDK has to be installed and the path to jdk/bin has to be set in environment variables

- a. Download JDK from the following url -

<http://www.oracle.com/technetwork/java/javase/downloads/index.html>

Click on Java Platform (JDK)



Click on the Accept License Agreement Radio button and click on JDK link for windows depending on the OS version [32/64 bit]

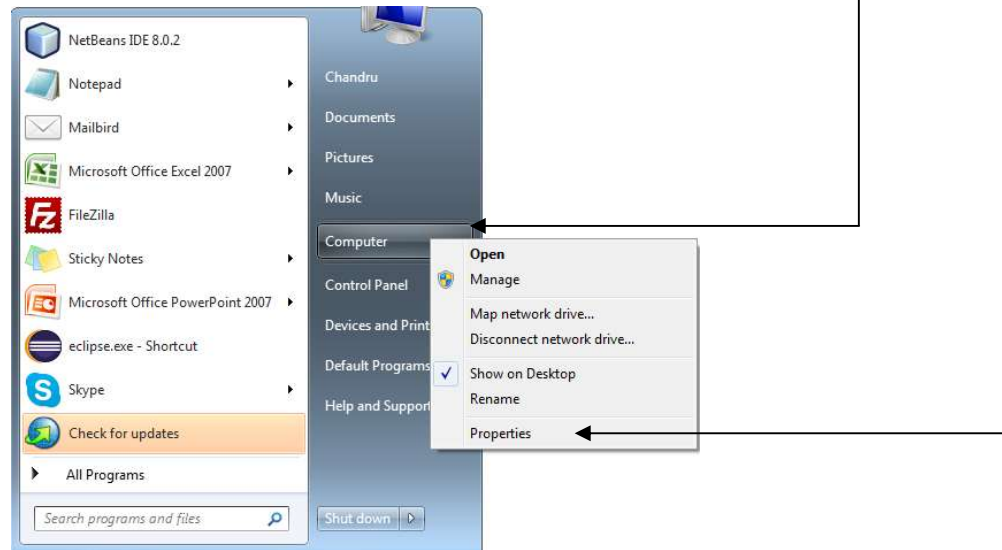
Looking for JDK 8 on ARM?

JDK 8 for ARM downloads have moved to the [JDK 8 for ARM download page](#).

Java SE Development Kit 8u45		
You must accept the Oracle Binary Code License Agreement for Java SE to download this software.		
<input type="radio"/> Accept License Agreement <input checked="" type="radio"/> Decline License Agreement		
Product / File Description	File Size	Download
Linux x86	146.89 MB	jdk-8u45-linux-i586.rpm
Linux x86	166.88 MB	jdk-8u45-linux-i586.tar.gz
Linux x64	145.19 MB	jdk-8u45-linux-x64.rpm
Linux x64	165.24 MB	jdk-8u45-linux-x64.tar.gz
Mac OS X x64	221.98 MB	jdk-8u45-macosx-x64.dmg
Solaris SPARC 64-bit (SVR4 package)	131.73 MB	jdk-8u45-solaris-sparcv9.tar.Z
Solaris SPARC 64-bit	92.9 MB	jdk-8u45-solaris-sparcv9.tar.gz
Solaris x64 (SVR4 package)	139.51 MB	jdk-8u45-solaris-x64.tar.Z
Solaris x64	95.88 MB	jdk-8u45-solaris-x64.tar.gz
Windows x86	175.98 MB	jdk-8u45-windows-i586.exe
Windows x64	180.44 MB	jdk-8u45-windows-x64.exe

- b. After download, you can double click on the exe to install JDK

- c. To Set environment variables, click on windows start -> Right click on computer -> Go to properties

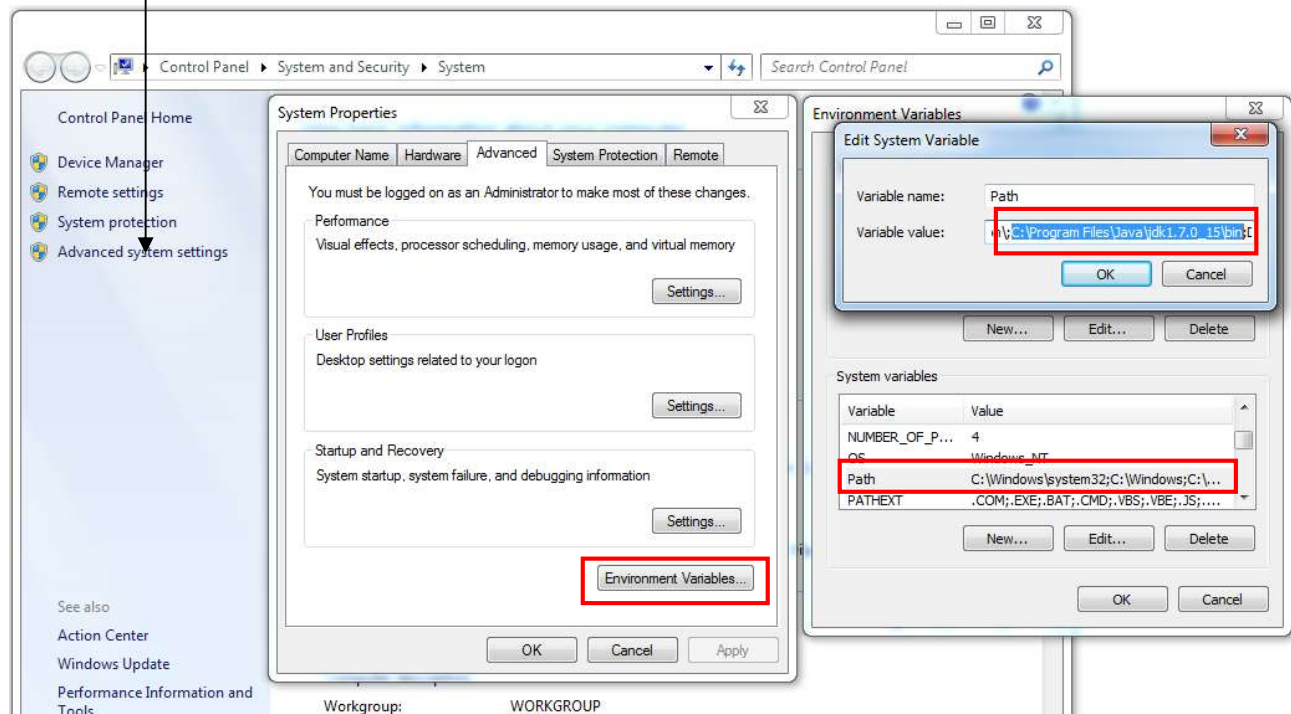


Click on Advanced system settings to open up System Properties.

Click on Environment Variables

Double click on Path environment variable under System variables

Provide the path of java jdk installation [normally c:\Program Files\Java\jdk] and click on OK. Click on OK to close the rest of the windows.

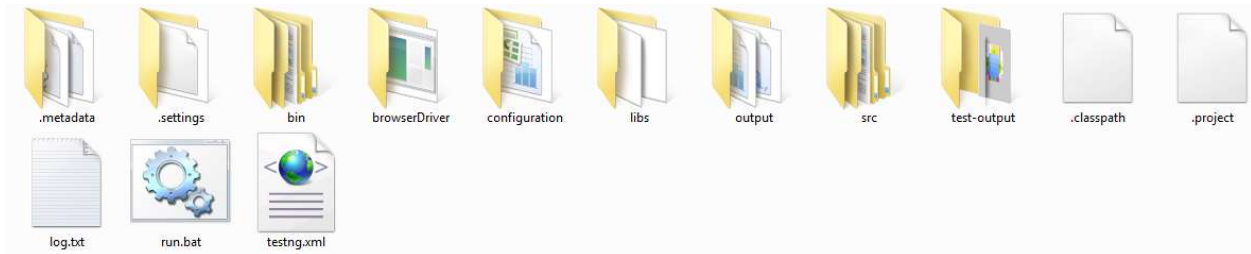


To check if the environment variable is working fine, open cmd prompt and type
Java -version

This command should show the version of Java

Section 1: Installation

Double click on “setup.exe” to install the framework. Default installation path is “C:\RockwellWebCrawler”. You can customize the path of installation. Once the setup is done you will get the following folders and files



Section 2: Configuration

- The distributor excel file has to be placed under the output folder. The name of the distributor file has to be **distributorInputFile**
- Initial configurations have to be done before the script is executed. Open the config.xlsx under configuration folder and enter the Sheet Name, starting and ending row numbers of the distributorInputFile.xlsx

Section 3: Execution

Double click on “run.bat” to launch chrome browser and crawl through the website to collect data.

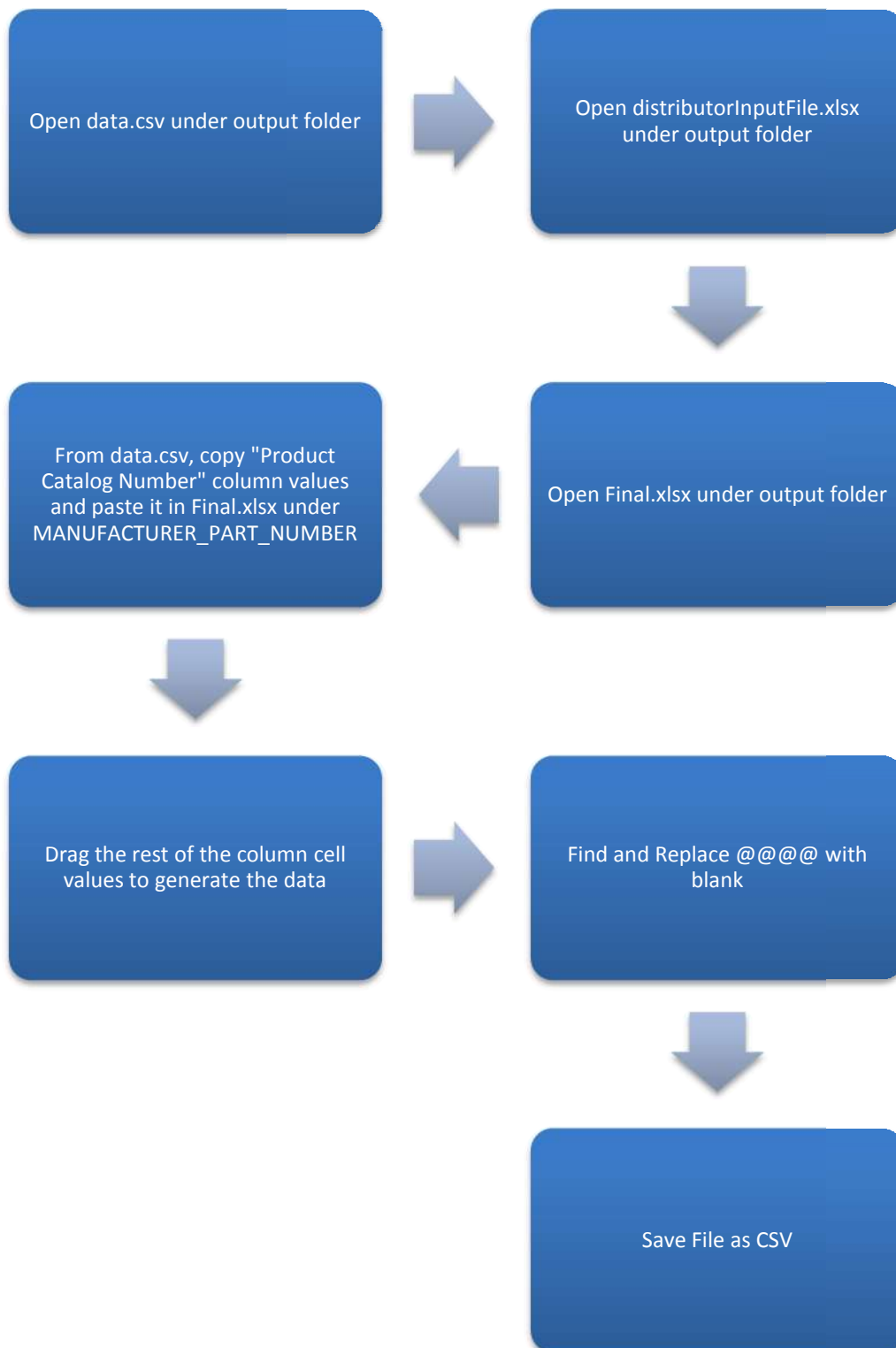
The execution might take about 8 – 12 hours depending on the system configuration and internet bandwidth.

Note:

Every time, you execute the framework, data.csv and Final.xlsx File will be overwritten. If you need a copy of the earlier execution, you need to backup the output folder.

Also, whenever there are changes/additions in the distributorInputFile, you need to replace the same under output folder

Section 4: Post Execution Compilation



Step 1: Open data.csv

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	Status	Pri/ Sec	Level 1	L1 Type	Level 2	L2 Type	Level 3	L3 Type	Level 4	L4 Type	Level 5	L5 Type	Friendly U Navigation	Page Head	Subcat Lis	Short Des	Overview	Product D	Product Catalog	N
2	PROD	P	Circuit & LC	Circuit Bre SC	Miniature SSC	1489 Theri	T						/Circuit-ar 1489 Theri	1489 Theri	/resource /resource	/resource Bulletin 1: 1489-M1C005				
3	PROD	P	Circuit & LC	Circuit Bre SC	Miniature SSC	1489 Theri	T						/Circuit-ar 1489 Theri	1489 Theri	/resource /resource	/resource Bulletin 1: 1489-M1C010				
4	PROD	P	Circuit & LC	Circuit Bre SC	Miniature SSC	1489 Theri	T						/Circuit-ar 1489 Theri	1489 Theri	/resource /resource	/resource Bulletin 1: 1489-M1C016				
5	PROD	P	Circuit & LC	Circuit Bre SC	Miniature SSC	1489 Theri	T						/Circuit-ar 1489 Theri	1489 Theri	/resource /resource	/resource Bulletin 1: 1489-M1C020				
6	PROD	P	Circuit & LC	Circuit Bre SC	Miniature SSC	1489 Theri	T						/Circuit-ar 1489 Theri	1489 Theri	/resource /resource	/resource Bulletin 1: 1489-M1C030				
7	PROD	P	Circuit & LC	Circuit Bre SC	Miniature SSC	1489 Theri	T						/Circuit-ar 1489 Theri	1489 Theri	/resource /resource	/resource Bulletin 1: 1489-M1C040				
8	PROD	P	Circuit & LC	Circuit Bre SC	Miniature SSC	1489 Theri	T						/Circuit-ar 1489 Theri	1489 Theri	/resource /resource	/resource Bulletin 1: 1489-M1C050				
9	PROD	P	Circuit & LC	Circuit Bre SC	Miniature SSC	1489 Theri	T						/Circuit-ar 1489 Theri	1489 Theri	/resource /resource	/resource Bulletin 1: 1489-M1C060				
10	PROD	P	Circuit & LC	Circuit Bre SC	Miniature SSC	1489 Theri	T						/Circuit-ar 1489 Theri	1489 Theri	/resource /resource	/resource Bulletin 1: 1489-M1C070				
11	PROD	P	Circuit & LC	Circuit Bre SC	Miniature SSC	1489 Theri	T						/Circuit-ar 1489 Theri	1489 Theri	/resource /resource	/resource Bulletin 1: 1489-M1C080				
12	PROD	P	Circuit & LC	Circuit Bre SC	Miniature SSC	1489 Theri	T						/Circuit-ar 1489 Theri	1489 Theri	/resource /resource	/resource Bulletin 1: 1489-M1C100				
13	PROD	P	Circuit & LC	Circuit Bre SC	Miniature SSC	1489 Theri	T						/Circuit-ar 1489 Theri	1489 Theri	/resource /resource	/resource Bulletin 1: 1489-M1C130				
14	PROD	P	Circuit & LC	Circuit Bre SC	Miniature SSC	1489 Theri	T						/Circuit-ar 1489 Theri	1489 Theri	/resource /resource	/resource Bulletin 1: 1489-M1C150				
15	PROD	P	Circuit & LC	Circuit Bre SC	Miniature SSC	1489 Theri	T						/Circuit-ar 1489 Theri	1489 Theri	/resource /resource	/resource Bulletin 1: 1489-M1C160				
16	PROD	P	Circuit & LC	Circuit Bre SC	Miniature SSC	1489 Theri	T						/Circuit-ar 1489 Theri	1489 Theri	/resource /resource	/resource Bulletin 1: 1489-M1C200				
17	PROD	P	Circuit & LC	Circuit Bre SC	Miniature SSC	1489 Theri	T						/Circuit-ar 1489 Theri	1489 Theri	/resource /resource	/resource Bulletin 1: 1489-M1C250				
18	PROD	P	Circuit & LC	Circuit Bre SC	Miniature SSC	1489 Theri	T						/Circuit-ar 1489 Theri	1489 Theri	/resource /resource	/resource Bulletin 1: 1489-M1C300				
19	PROD	P	Circuit & LC	Circuit Bre SC	Miniature SSC	1489 Theri	T						/Circuit-ar 1489 Theri	1489 Theri	/resource /resource	/resource Bulletin 1: 1489-M1C320				
20	PROD	P	Circuit & LC	Circuit Bre SC	Miniature SSC	1489 Theri	T						/Circuit-ar 1489 Theri	1489 Theri	/resource /resource	/resource Bulletin 1: 1489-M1C350				
21	PROD	P	Circuit & LC	Circuit Bre SC	Miniature SSC	1489 Theri	T						/Circuit-ar 1489 Theri	1489 Theri	/resource /resource	/resource Bulletin 1: 1489-M1C400				
22	PROD	P	Circuit & LC	Circuit Bre SC	Miniature SSC	1489 Theri	T						/Circuit-ar 1489 Theri	1489 Theri	/resource /resource	/resource Bulletin 1: 1489-M1C500				
23	PROD	P	Circuit & LC	Circuit Bre SC	Miniature SSC	1489 Theri	T						/Circuit-ar 1489 Theri	1489 Theri	/resource /resource	/resource Bulletin 1: 1489-M1C600				
24	PROD	P	Circuit & LC	Circuit Bre SC	Miniature SSC	1489 Theri	T						/Circuit-ar 1489 Theri	1489 Theri	/resource /resource	/resource Bulletin 1: 1489-M1C630				
25	PROD	P	Circuit & LC	Circuit Bre SC	Miniature SSC	1489 Theri	T						/Circuit-ar 1489 Theri	1489 Theri	/resource /resource	/resource Bulletin 1: 1489-M1D005				

Step 2: Open distributorInputFile

A	B	C	D	E	F	G	H	I	J	K
1	TRC_PART_NUMBER	MANUFACTURER_NAME	BRAND_NAME	UPC	MANUFACTURER_PART_NUMBER	UNSPSC	ACTIVE(Active='Y' InActive='N'	Planned Obsolescence='P'	Obsolete='O' Item Alert='T'	Unlisted='U' Non-price
2	ALBR1000855T07	Rockwell Automation	Allen-Bradley	66207476701	1000-855T-07	32151503	Y			
3	ALBR1000855T09	Rockwell Automation	Allen-Bradley	66207476703	1000-855T-09	32151503	Y			
4	ALBR1000855T12	Rockwell Automation	Allen-Bradley	66207475030	1000-855T-12	32151503	Y			
5	ALBR1000855T13	Rockwell Automation	Allen-Bradley	66207475031	1000-855T-13	32151503	Y			
6	ALBR1000855T19	Rockwell Automation	Allen-Bradley	61259866977	1000-855T-19	32151503	Y			
7	ALBR1000855T25	Rockwell Automation	Allen-Bradley	88563013177	1000-855T-25	32151503	Y			
8	ALBR100A12ND3	Rockwell Automation	Allen-Bradley	66207229603	100-A12ND3	39121529	Y			
9	ALBR100C09A10	Rockwell Automation	Allen-Bradley	66207260821	100-C09A10	39121529	Y			
10	ALBR100C09B10	Rockwell Automation	Allen-Bradley	66207260822	100-C09B10	39121529	Y			
11	ALBR100C09D01	Rockwell Automation	Allen-Bradley	66207260951	100-C09D01	39121529	Y			
12	ALBR100C09D10	Rockwell Automation	Allen-Bradley	66207260820	100-C09D10	39121529	Y			
13	ALBR100C09D200	Rockwell Automation	Allen-Bradley	66246871835	100-C09D200	39121529	Y			
14	ALBR100C09D300	Rockwell Automation	Allen-Bradley	66246871836	100-C09D300	39121529	Y			
15	ALBR100C09D400	Rockwell Automation	Allen-Bradley	66246871837	100-C09D400	39121529	Y			
16	ALBR100C09D01	Rockwell Automation	Allen-Bradley	66246886443	100-C09D01	39121529	Y			
17	ALBR100C09D10	Rockwell Automation	Allen-Bradley	66246886374	100-C09D10	39121529	Y			
18	ALBR100C09D300	Rockwell Automation	Allen-Bradley	66246886714	100-C09D300	39121529	Y			
19	ALBR100C09D400	Rockwell Automation	Allen-Bradley	66246886646	100-C09D400	39121529	Y			

Step 3: Open Final.xlsx

The screenshot shows the Microsoft Excel interface with the following details:

- Title Bar:** Final.xlsx - Microsoft Excel
- Ribbon:** Home, Insert, Page Layout, Formulas, Data, Review, View, Developer, Team.
- Formulas Tab:**
 - Function Library:** AutoSum, Fill, Sort & Filter, Find & Select.
 - Formula Auditing:** Trace Precedents, Trace Dependents, Show Formulas, Show Input Cells.
 - Calculation:** Calculate Now, Calculate External Links.
 - Formula Options:** General, Number, Text, Logical, Date & Time, Financial, Accounting, Conditional Formatting, Format as Table, Cell Styles, Merge & Center, Wrap Text.
- Formula Bar:** E2, =data.csv!\$T\$2
- Worksheet:**
 - Columns:** A, B, C, D, E, F, G, H, I, J, K, L, M, N.
 - Rows:** 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20.
 - Cell E2:** Contains the formula =data.csv!\$T\$2.
 - Cell G2:** Contains the text: ACTIVE[Active='Y', InActive='N', Planned, Obsolescence='P', Obsolete='O', Item Alert='T', Unlisted='U' Non-].
 - Cell H2:** Contains the text: DISPLAY_O, NUNE.
 - Cell I2:** Contains the text: ORIG_SORT, 129.
 - Cell J2:** Contains the text: Level 1, Product Id.
 - Cell K2:** Contains the text: L1 Type, Product Id.
 - Cell L2:** Contains the text: Level 2, Product Id.
 - Cell M2:** Contains the text: L2 Type, Product Id.
 - Cell N2:** Contains the text: Level 3, Product Id.

Step 4: Copy product catalog numbers from data.csv

[illegible]

Step 5: Paste the values under MANUFACTURER_PART_NUMBER in Final.xlsx

The screenshot displays the Microsoft Excel interface with the following details:

- Title Bar:** Final.xlsx - Microsoft Excel
- Ribbon Tabs:** Home, Insert, Page Layout, Formulas, Data, Review, View, Developer, Team.
- Home Tab Groups:**
 - Clipboard:** Cut, Copy, Paste, Format Painter.
 - Font:** Calibri, size 11, Bold (B), Italic (I), Underline (U), Color, Background Color.
 - Paragraph:** Bullets, Numbering, Indentation, Wrap Text, Merge & Center.
 - Alignment:** Left, Center, Right, Justify, Top, Bottom, Middle.
 - Number:** General, Currency (\$), Percentage (%), Fraction (/), Date, Time, Text, Special.
 - Conditional Formatting:** Highlight Cells Rules, Cell Style Rules, New Cell Rule.
 - Format as Table:** Table Styles.
 - Cell Styles:** Cell Styles.
 - Insert:** Rows, Columns, Tables, Charts, Links, Pictures, Comments.
 - Delete:** Delete Cells, Delete Rows, Delete Columns.
 - Format:** Font, Fill, Borders, Conditional Formatting, Styles.
 - Cells:** AutoSum, Fill, Clear, Sort & Filter, Find & Select.
- Formula Bar:** Shows the active cell E2 containing the value 1489-M1C005.
- Spreadsheet Content:**

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	TRC_PART_NUMBER	MANUFACTURER_NAME	BRAND_NAME	UPC	ART_NUMBER	UNSPSC	ACTIVE(Active='Y' 'Inactive='N' Planned Obsolescence='P' 'Obsolete='O' Item Alert='T' Unlisted='U' Non-	DISPLAY_O		Level 1 Product Id	L1 Type	Level 2 Product Id	L2 Type	Level 3 Product Id
2	ALBR1489M1C005	Rockwell Automatio	Allen-Bradley	88563061962	1489-M1C005	39121603	Y	Y	129	Circuit & Load F C		Circuit Breaker SC		Miniature Ci
3					1489-M1C010									
4					1489-M1C016									
5					1489-M1C020									
6					1489-M1C030									
7					1489-M1C040									
8					1489-M1C050									
9					1489-M1C060									
10					1489-M1C070									
11					1489-M1C080									
12					1489-M1C100									
13					1489-M1C130									
14					1489-M1C150									
15					1489-M1C160									
16					1489-M1C200									
17					1489-M1C250									
18					1489-M1C300									
19					1489-M1C320									
20					1489-M1C360									
- Status Bar:** Sheet1 | Sheet2 | Sheet3 | Count: 172 | 100%

Step 6: Drag the other columns to fill the data

Final.xlsx - Microsoft Excel

Home Insert Page Layout Formulas Data Review View Developer Team

Clipboard Font Alignment Number

General Conditional Formatting Format as Table Cell Styles Insert Delete Format Cells

AutoSum Fill Clear Sort & Find & Filter Select Editing

A2 =INDEX([distributorInputFile.xlsx]Sheet1!\$A\$2:\$A\$20466,MATCH(E2,[distributorInputFile.xlsx]Sheet1!\$E\$2:\$E\$20466,0))

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
	TRC_PART_NUMBER	MANUFACTURER_NAME	BRAND_NAME	UPC	ART_NUMBER	UNSPSC	ACTIVE(Active='Y' 'Inactive='N' Planned Obsolescence='p' 'Obsolete='O' Item Alert='T' Unlisted='U' Non-	DISPLAY_O	ORIG_SORT	Level 1 Product Id	L1 Type	Level 2 Product Id	L2 Type	Level 3 Product Id
1	ALBR1489M1C005	Rockwell Automatio	Allen-Bradley	88563061962	1489-M1C005	39121603	Y	Y	129	Circuit & Load F C		Circuit Breaker SC		Miniature Ci
2	ALBR1489M1C010				1489-M1C010									
3	ALBR1489M1C016				1489-M1C016									
4	ALBR1489M1C020				1489-M1C020									
5	ALBR1489M1C030				1489-M1C030									
6	ALBR1489M1C040				1489-M1C040									
7	ALBR1489M1C050				1489-M1C050									
8	ALBR1489M1C060				1489-M1C060									
9	ALBR1489M1C070				1489-M1C070									
10	ALBR1489M1C080				1489-M1C080									
11					1489-M1C100									
12					1489-M1C130									
13					1489-M1C150									
14					1489-M1C160									
15					1489-M1C200									
16					1489-M1C250									
17					1489-M1C300									
18					1489-M1C320									
19					1489-M1C350									

Sheet1 Sheet2 Sheet3

Drag outside selection to extend series or fill; drag inside to clear

Count: 10 100%

Section 5: Cell Formulae

TRC_PART_NUMBER

INDEX('[distributorInputFile.xlsx]Sheet1'!\$A\$2:\$A\$20466,MATCH(E2,'[distributorInputFile.xlsx]Sheet1'!\$E\$2:\$E\$20466,0))

MANUFACTURER_NAME

INDEX('[distributorInputFile.xlsx]Sheet1'!\$B\$2:\$B\$20466,MATCH(E2,'[distributorInputFile.xlsx]Sheet1'!\$E\$2:\$E\$20466,0))

BRAND_NAME

INDEX('[distributorInputFile.xlsx]Sheet1'!\$C\$2:\$C\$20466,MATCH(E2,'[distributorInputFile.xlsx]Sheet1'!\$E\$2:\$E\$20466,0))

UPC

INDEX('[distributorInputFile.xlsx]Sheet1'!\$D\$2:\$D\$20466,MATCH(E2,'[distributorInputFile.xlsx]Sheet1'!\$E\$2:\$E\$20466,0))

MANUFACTURER_PART_NUMBER

'data.csv'!\$T\$2

UNSPSC

INDEX('[distributorInputFile.xlsx]Sheet1'!\$F\$2:\$F\$20466,MATCH(E2,'[distributorInputFile.xlsx]Sheet1'!\$E\$2:\$E\$20466,0))

ACTIVE(Active='Y' InActive='N' Planned Obsolescence='P' Obsolete='O' Item Alert='T' Unlisted='U' Non-price Maintained='K' Withdrawn='W' Pending To Delete='X')

INDEX('[distributorInputFile.xlsx]Sheet1'!\$G\$2:\$G\$20466,MATCH(E2,'[distributorInputFile.xlsx]Sheet1'!\$E\$2:\$E\$20466,0))

DISPLAY_ONLINE

INDEX('[distributorInputFile.xlsx]Sheet1'!\$H\$2:\$H\$20466,MATCH(E2,'[distributorInputFile.xlsx]Sheet1'!\$E\$2:\$E\$20466,0))

ORIG_SORT

INDEX('[distributorInputFile.xlsx]Sheet1'!\$I\$2:\$I\$20466,MATCH(E2,'[distributorInputFile.xlsx]Sheet1'!\$E\$2:\$E\$20466,0))

Level 1 Product Id

INDEX('[data.csv]data'!\$C\$2,MATCH(E2,'[data.csv]data'!\$T\$2,0))

L1 Type

INDEX('[data.csv]data'!\$D\$2,MATCH(E2,'[data.csv]data'!\$T\$2,0))

Level 2 Product Id

INDEX('[data.csv]data'!\$E\$2,MATCH(E2,'[data.csv]data'!\$T\$2,0))

L2 Type

INDEX('[data.csv]data'!\$F\$2,MATCH(E2,'[data.csv]data'!\$T\$2,0))

Level 3 Product Id

INDEX('[data.csv]data'!\$G\$2,MATCH(E2,'[data.csv]data'!\$T\$2,0))

L3 Type

INDEX('[data.csv]data'!\$H\$2,MATCH(E2,'[data.csv]data'!\$T\$2,0))

Level 4 Product Id

INDEX('[data.csv]data'!\$I\$2,MATCH(E2,'[data.csv]data'!\$T\$2,0))

L4 Type

INDEX('[data.csv]data'!\$J\$2,MATCH(E2,'[data.csv]data'!\$T\$2,0))

Level 5 Product Id

INDEX('[data.csv]data'!\$K\$2,MATCH(E2,'[data.csv]data'!\$T\$2,0))

L5 Type

INDEX('[data.csv]data'!\$L\$2,MATCH(E2,'[data.csv]data'!\$T\$2,0))

Friendly URL

INDEX('[data.csv]data'!\$M\$2,MATCH(E2,'[data.csv]data'!\$T\$2,0))

Navigation Text

INDEX('[data.csv]data'!\$N\$2,MATCH(E2,'[data.csv]data'!\$T\$2,0))

Page Heading

INDEX('[data.csv]data'!\$O\$2,MATCH(E2,'[data.csv]data'!\$T\$2,0))

Subcat Listing image

INDEX('[data.csv]data'!\$P\$2,MATCH(E2,'[data.csv]data'!\$T\$2,0))

Short Description Image

INDEX('[data.csv]data'!\$Q\$2,MATCH(E2,'[data.csv]data'!\$T\$2,0))

Overview Tab Image

INDEX('[data.csv]data'!\$R\$2,MATCH(E2,'[data.csv]data'!\$T\$2,0))

Product Description

INDEX('[data.csv]data'!\$S\$2,MATCH(E2,'[data.csv]data'!\$T\$2,0))