INVOICE EXTRACTION USING AUTOMATION ANYWHERE COMMUNITY EDITION vA2019

AIM:

To create an efficient bot to extract data from invoices downloaded to a folder.

REQUIREMENTS:

Automation Anywhere Community Edition v2019.

OBJECTIVES:

- Create text files from the PDFs
- Store data from the text files
- Write data to the Excel worksheet

CHALLENGES:

Ever since the update on the AAE Community Edition from previous v.A11 to v.A2019, the challenges faced are:

- Bots can no longer be exported to the system. They are present only within the online cloud/repository of A2019 control room.
- The 'Extract Form Fields' command is not present in the v.A2019. Thus, data can be extracted from PDF to Excel worksheet indirectly till date. This eventually leads to erroneous data and I honestly hope the 'Extract Form Fields' command is added soon.

I'll show the steps and variables required to get started. A video will be logged in GitHub repository and the input files as well. So if you want to recreate it with your PDF integration, you're welcome to do so and send me feedbacks on how I can improve it. Let's get down to business.

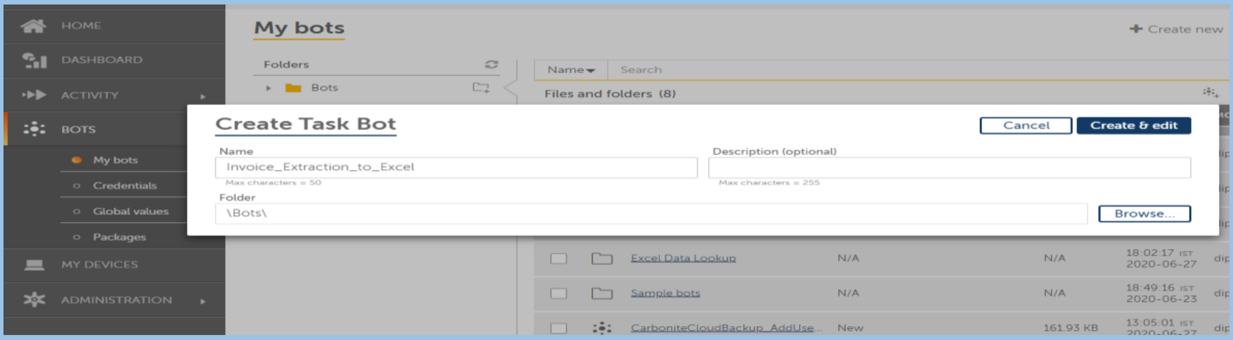
VARIABLES REQUIRED:

Sl. No.	**Variables**	**Type**	**Default Value**	**About**
1	Table_Value	Table		Holds the values of the text file in a table. Can be accessed with index values.

2	Space	String	“ ”	Will act as a space delimiter. Just enter a space in the default value.
3	ListforDiscount	List		Will hold string value which will be split later. Can be accessed with index.
4	DiscountValue	String		Will hold string value of discount %.
5	Split	List		Will hold string value which will be split later. Can be accessed with index.
6	TextFilename	Dictionary		Holds the file names in a folder. Can be accessed with keys. For a file name, use ‘name’ as key and for extension use ‘extension’ as key value.

STEPS:

Open the control room for community edition of Automation Anywhere and create new bot. Also keep the input files entirely to a different folder.



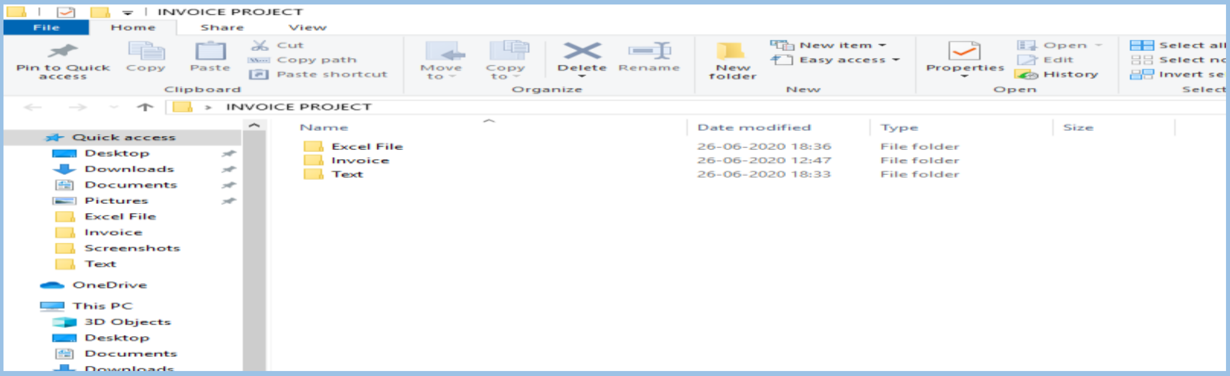
Create Task Bot

Name: Invoice_Extraction_to_Excel (Max characters = 50)

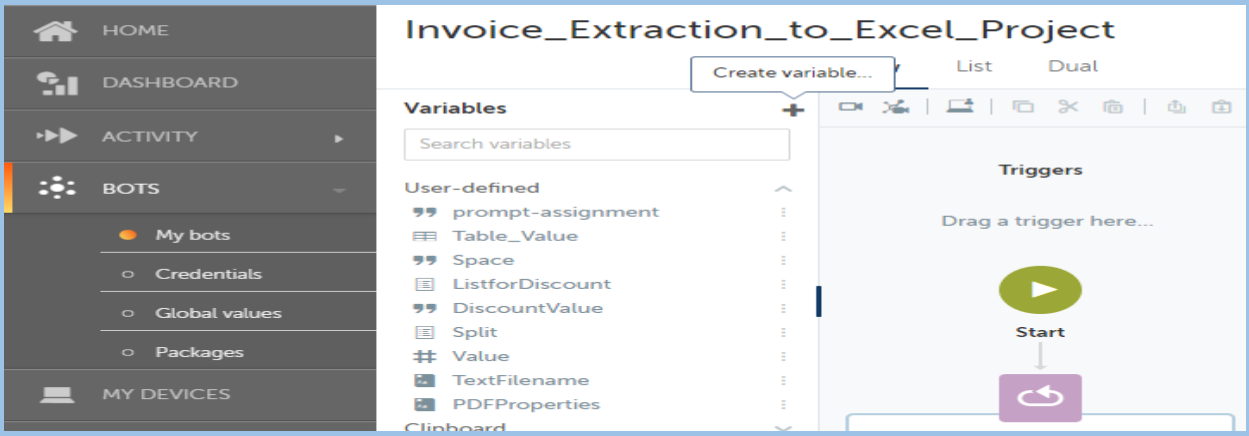
Description (optional): (Max characters = 255)

Folder: \Bots\ Browse...

Excel Data Lookup	N/A	N/A	18:02:17 IST 2020-06-27	dir
Sample bots	N/A	N/A	18:49:16 IST 2020-06-23	dir
CarboniteCloudBackup_AddUse...	New	161.93 KB	13:05:01 IST 2020-06-27	dir



Name	Date modified	Type	Size
Excel File	26-06-2020 18:36	File folder	
Invoice	26-06-2020 12:47	File folder	
Text	26-06-2020 18:33	File folder	



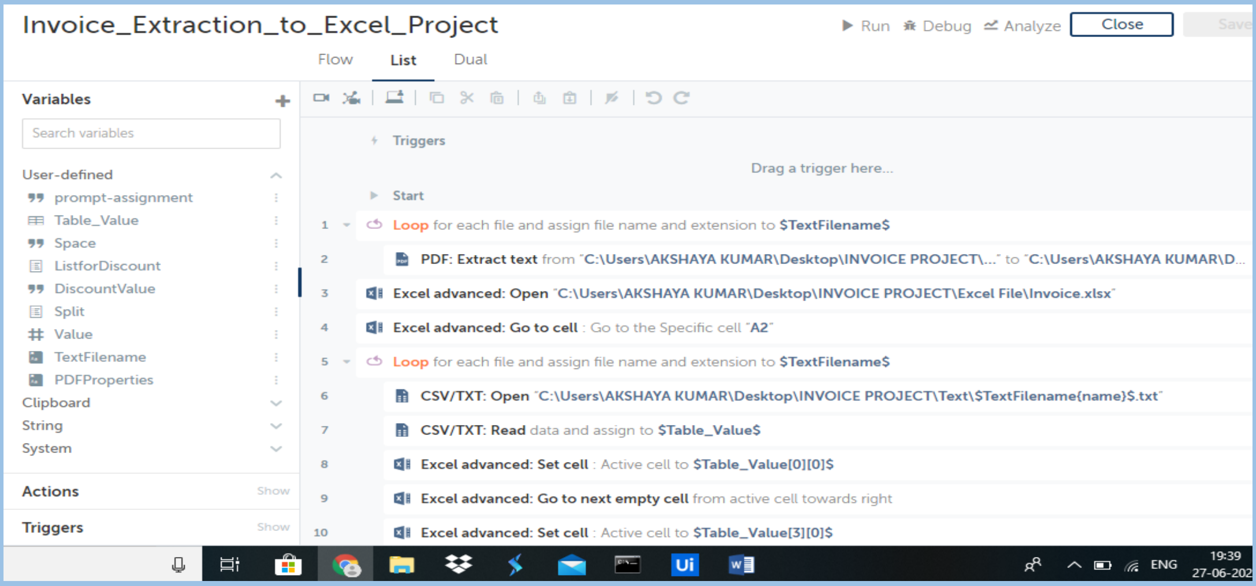
Invoice_Extraction_to_Excel_Project

Variables

User-defined
- prompt-assignment
- Table_Value
- Space
- ListforDiscount
- DiscountValue
- Split
- Value
- TextFilename
- PDFProperties

Clipboard

After creating the variable, let's get on with the following steps:



Invoice_Extraction_to_Excel_Project

Variables

User-defined
- prompt-assignment
- Table_Value
- Space
- ListforDiscount
- DiscountValue
- Split
- Value
- TextFilename
- PDFProperties

Clipboard
String
System

Triggers

Start

1. Loop for each file and assign file name and extension to $TextFilename$
2. PDF: Extract text from "C:\Users\AKSHAYA KUMAR\Desktop\INVOICE PROJECT\..." to "C:\Users\AKSHAYA KUMAR\D...
3. Excel advanced: Open "C:\Users\AKSHAYA KUMAR\Desktop\INVOICE PROJECT\Excel File\Invoice.xlsx"
4. Excel advanced: Go to cell : Go to the Specific cell "A2"
5. Loop for each file and assign file name and extension to $TextFilename$
6. CSV/TXT: Open "C:\Users\AKSHAYA KUMAR\Desktop\INVOICE PROJECT\Text\$TextFilename{name}$.txt"
7. CSV/TXT: Read data and assign to $Table_Value$
8. Excel advanced: Set cell : Active cell to $Table_Value[0][0]$
9. Excel advanced: Go to next empty cell from active cell towards right
10. Excel advanced: Set cell : Active cell to $Table_Value[3][0]$

Image 1: Steps a

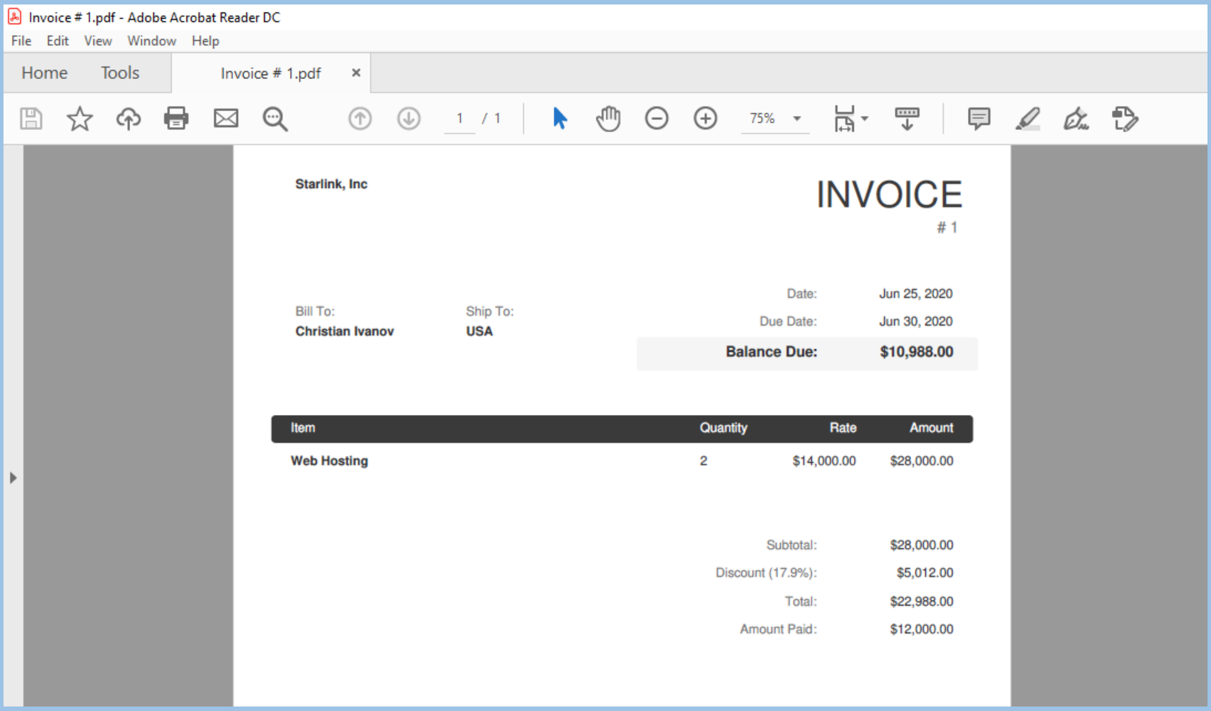
Step1: **Loop** for each file and assign file name and extension to $TextFilename$

→Loop Type: Iterator: 'For each file in a folder'. Provide folder path & add variable 'TextFilename' to the **Assign file name and extension to this variable**. Keep rest as it is.

Inside the loop:

a) '**PDF: Extract text**' command. Provide path for PDF file with 'TextFilename{name}' so it looks like this, [C:\Users\AKSHAYA KUMAR\Desktop\INVOICEPROJECT\Invoice\**$TextFilename{name}$**.pdf]. Text Type: Plain. Page range: All pages. In the 'Export data to text file Desktop file' provide the path with 'TextFilename', such that it looks like: [C:\Users\AKSHAYA KUMAR\Desktop\INVOICE

PROJECT\Text\**$TextFilename{name}$**.txt]. Check the 'Overwrite files with the same name' check box.


Invoice # 1.pdf - Adobe Acrobat Reader DC
File Edit View Window Help
Home Tools Invoice # 1.pdf
1 / 1
75%

Starlink, Inc

INVOICE

1

Bill To:
Christian Ivanov

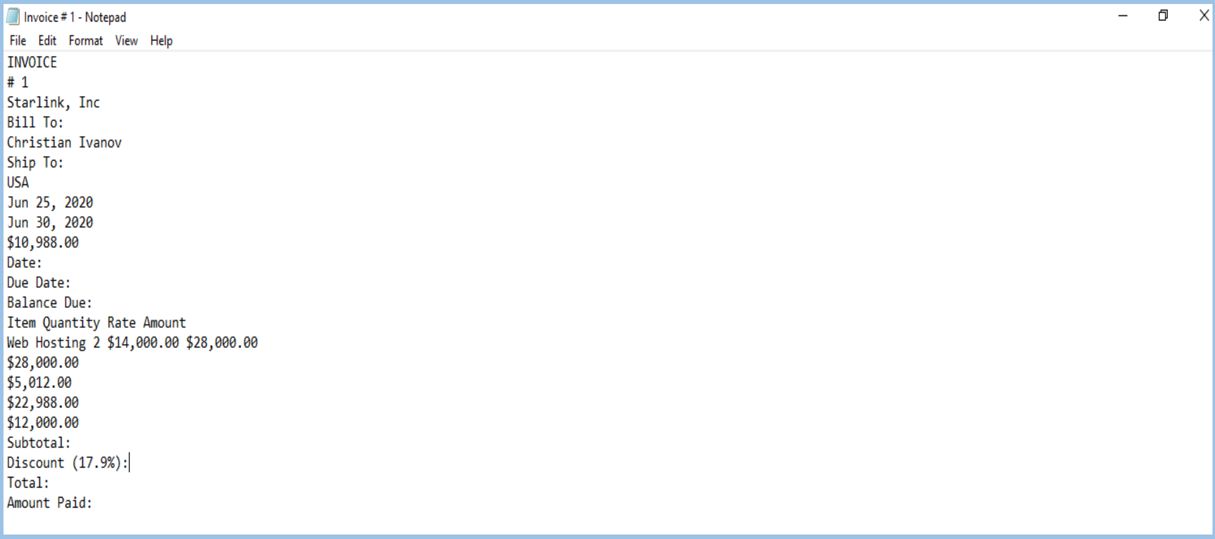
Ship To:
USA

Date:	Jun 25, 2020
Due Date:	Jun 30, 2020
Balance Due:	$10,988.00

Item	Quantity	Rate	Amount
Web Hosting	2	$14,000.00	$28,000.00

Subtotal:	$28,000.00
Discount (17.9%):	$5,012.00
Total:	$22,988.00
Amount Paid:	$12,000.00

Image 2: Input PDF



```
INVOICE
# 1
Starlink, Inc
Bill To:
Christian Ivanov
Ship To:
USA
Jun 25, 2020
Jun 30, 2020
$10,988.00
Date:
Due Date:
Balance Due:
Item Quantity Rate Amount
Web Hosting 2 $14,000.00 $28,000.00
$28,000.00
$5,012.00
$22,988.00
$12,000.00
Subtotal:
Discount (17.9%):
Total:
Amount Paid:
```

Image 3: Output Text file from the loop in Step 1 in 'plain' format.

Step 2: '**Excel advanced: Open'** command. Session : Default. Provide Desktop file path of the excel file where you wanna store the data. Specify sheet name as "Sheet1". Open in: 'Read-write mode'. Check the 'Sheet contains a header' checkbox.

Step 3: ‘**Excel advanced: Go to cell’** command. Session name: Default. Specify cell as : “A2”.

Step 4: ‘**Loop’** command. Loop Type: Iterator: ‘For each file in a folder’. Provide folder path that contains the pdf invoice, something like this: [C:\Users\AKSHAYA KUMAR\Desktop\INVOICEPROJECT\Text]. Assign ‘TextFilename’ variable to it. Leave the rest and apply.

➔ Inside the loop:

a) ‘**CSV/TXT: Open’** command. Set session name as “Default”. Provide the file path where the text files are saved [C:\Users\AKSHAYA KUMAR\Desktop\INVOICEPROJECT\Invoice\**$TextFilename{name}$**.txt]. Check the ‘Contains header’ check box. Delimiter as ‘Newline’. Then, apply the changes.

b) ‘**CSV/TXT: Read’** command. Set session name as “Default”. Assign ‘Table_Value’ variable to store the values of the text file in table format.

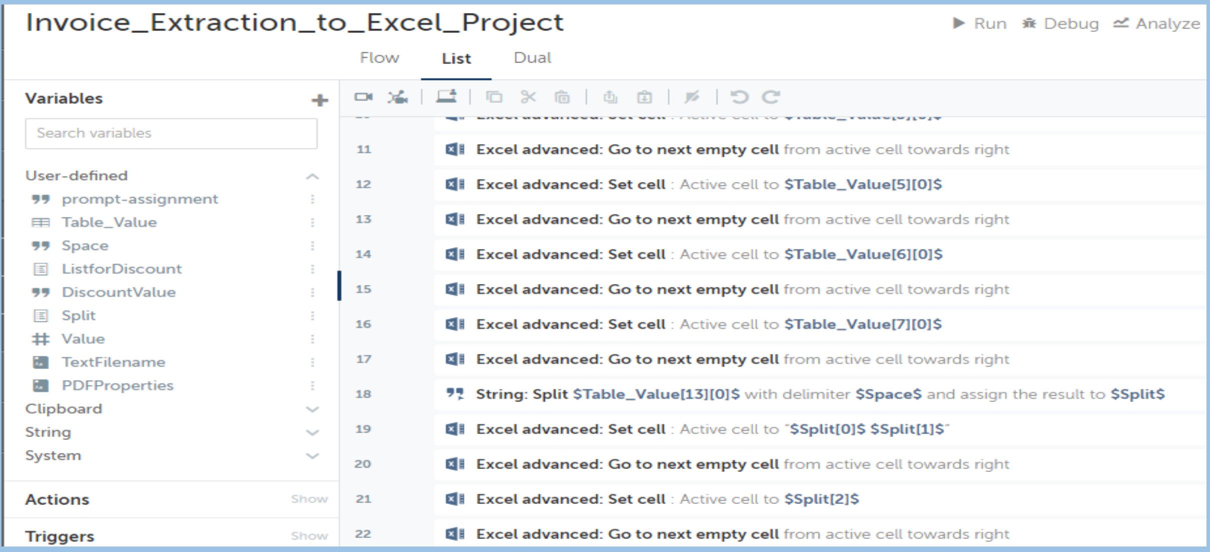
c) ‘**Excel advanced: Set cell’** command. Set ‘Active cell’ and cell value as “**$Table_Value[0][0]$**”.

d) ‘**Excel advanced: Go to next empty cell”** command and set ‘Active cell’ towards ‘right’.

e) ‘**Excel advanced: Set cell’** command. Set ‘Active cell’ and cell value as “**$Table_Value[3][0]$**”.

f) ‘**Excel advanced: Go to next empty cell”** command and set ‘Active cell’ towards ‘right’.

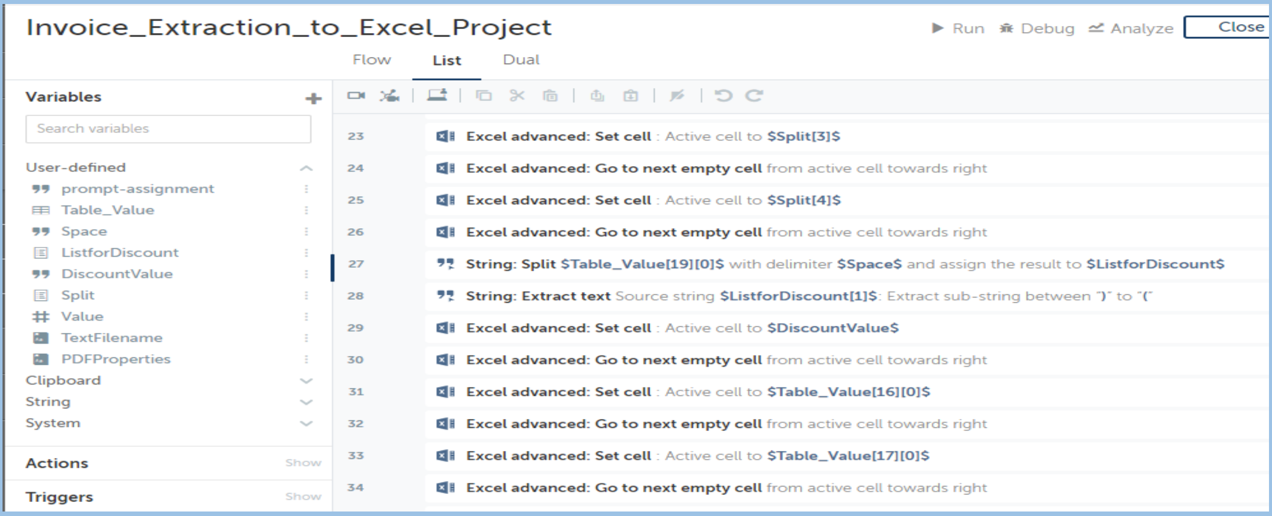
g) ‘**Excel advanced: Set cell’** command. Set ‘Active cell’ and cell value as “**$Table_Value[5][0]$**”.



```
Excel advanced: Go to next empty cell from active cell towards right
Excel advanced: Set cell : Active cell to $Table_Value[5][0]$
Excel advanced: Go to next empty cell from active cell towards right
Excel advanced: Set cell : Active cell to $Table_Value[6][0]$
Excel advanced: Go to next empty cell from active cell towards right
Excel advanced: Set cell : Active cell to $Table_Value[7][0]$
Excel advanced: Go to next empty cell from active cell towards right
String: Split $Table_Value[13][0]$ with delimiter $Space$ and assign the result to $Split$
Excel advanced: Set cell : Active cell to "$Split[0]$ $Split[1]$"
Excel advanced: Go to next empty cell from active cell towards right
Excel advanced: Set cell : Active cell to $Split[2]$
Excel advanced: Go to next empty cell from active cell towards right
```

Image 4: Steps b

h) ‘**Excel advanced: Go to next empty cell”** command and set ‘Active cell’ towards ‘right’.
i) ‘**Excel advanced: Set cell’** command. Set ‘Active cell’ and cell value as “**$Table_Value[6][0]$**”.
j) ‘**Excel advanced: Go to next empty cell”** command and set ‘Active cell’ towards ‘right’.
k) ‘**Excel advanced: Set cell’** command. Set ‘Active cell’ and cell value as “**$Table_Value[7][0]$**”.
l) ‘**Excel advanced: Go to next empty cell”** command and set ‘Active cell’ towards ‘right’.
From the next step we have to split texts and set them.
m) ‘**String: Split’** command. Source string: “**$Table_Value[13][0]$**” . Delimiter: “**$Space$”** variable. ‘Assign the output to list variable’ as “$Split$”.
n) ‘**Excel advanced: Set cell’** command with session name as “Default”. Cell option as ‘Active cell’ and cell value as “**$Split[0]$ $Split[1]$”.**
o) ‘**Excel advanced: Go to next empty cell”** command and set ‘Active cell’ towards ‘right’.
p) ‘**Excel advanced: Set cell’** command with session name as “Default”. Cell option as ‘Active cell’ and cell value as “**$Split[2]$”.**

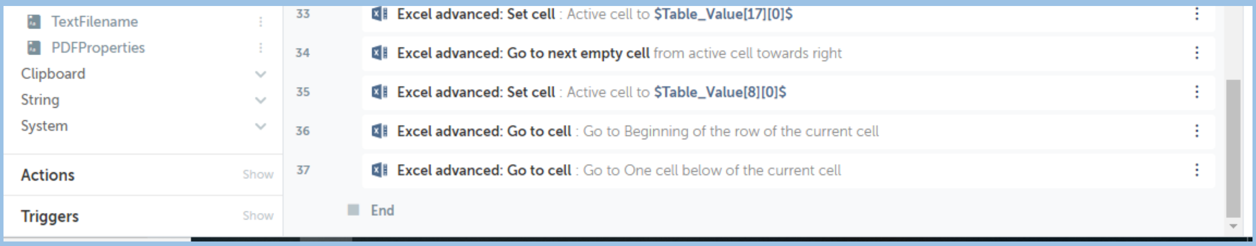


```
Excel advanced: Set cell : Active cell to $Split[3]$
Excel advanced: Go to next empty cell from active cell towards right
Excel advanced: Set cell : Active cell to $Split[4]$
Excel advanced: Go to next empty cell from active cell towards right
String: Split $Table_Value[19][0]$ with delimiter $Space$ and assign the result to $ListforDiscount$
String: Extract text Source string $ListforDiscount[1]$: Extract sub-string between ")" to "("
Excel advanced: Set cell : Active cell to $DiscountValue$
Excel advanced: Go to next empty cell from active cell towards right
Excel advanced: Set cell : Active cell to $Table_Value[16][0]$
Excel advanced: Go to next empty cell from active cell towards right
Excel advanced: Set cell : Active cell to $Table_Value[17][0]$
Excel advanced: Go to next empty cell from active cell towards right
```

Image 5: Steps c

q) ‘**Excel advanced: Go to next empty cell”** command and set ‘Active cell’ towards ‘right’.
r) ‘**Excel advanced: Set cell’** command with session name as “Default”. Cell option as ‘Active cell’ and cell value as “**$Split[3]$”.**

s) ‘**Excel advanced: Go to next empty cell”** command and set ‘Active cell’ towards ‘right’.
t) ‘**Excel advanced: Set cell’** command with session name as “Default”. Cell option as ‘Active cell’ and cell value as “**$Split[4]$”.**
u) ‘**Excel advanced: Go to next empty cell”** command and set ‘Active cell’ towards ‘right’.
v) ‘**String: Split’** command. Source string: “**$Table_Value[19][0]$**” . Delimiter: “**$Space$”** variable. ‘Assign the output to list variable’ as “$ListforDiscount$”.
w) ‘**String: Extract text’** command. Source string as variable “”. Get characters from ‘Before and/or after’: Start after text: (and end before text:). Trim the extracted text and also remove Enter from the extracted text. Assign the output to variable ‘DiscountValue’.
x) ‘**Excel advanced: Set cell’** command with session name as “Default”. Cell option as ‘Active cell’ and cell value as **“$DiscountValue$”.**
y) ‘**Excel advanced: Go to next empty cell”** command and set ‘Active cell’ towards ‘right’. Do the same for the rest of the cells.
z) Finally to go the beginning of the row. ‘**Excel advanced: Go to cell’** with cell option ‘Active cell’ to one cell below. Use the same step to go to the ‘beginning of the row’. End the command list.


TextFilename
PDFProperties
Clipboard
String
System
Actions Show
Triggers Show

Excel advanced: Set cell : Active cell to $Table_Value[17][0]$
Excel advanced: Go to next empty cell from active cell towards right
Excel advanced: Set cell : Active cell to $Table_Value[8][0]$
Excel advanced: Go to cell : Go to Beginning of the row of the current cell
Excel advanced: Go to cell : Go to One cell below of the current cell
End

Image 6: Steps d

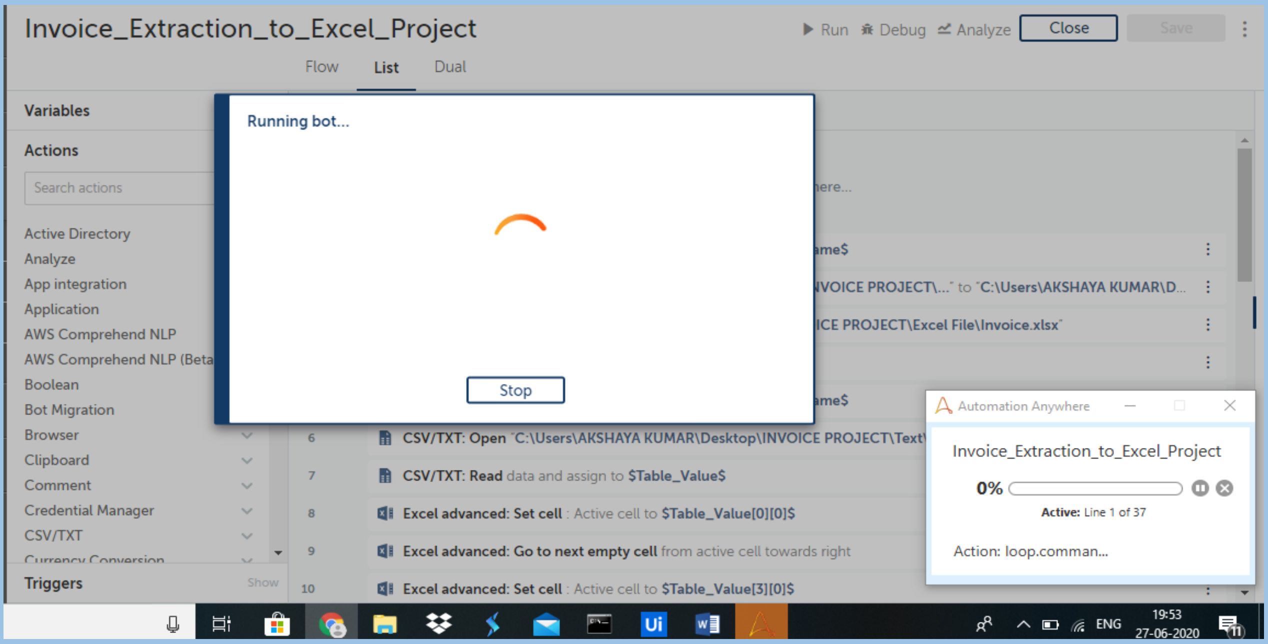
Now, finally save the task and run it!

For better debugging, it is suggested to break the entire bot into test bots and use the ‘**Message Box**’ command to check if the values you need to print are exactly the same.

The screen recorded video of the bot running is present.

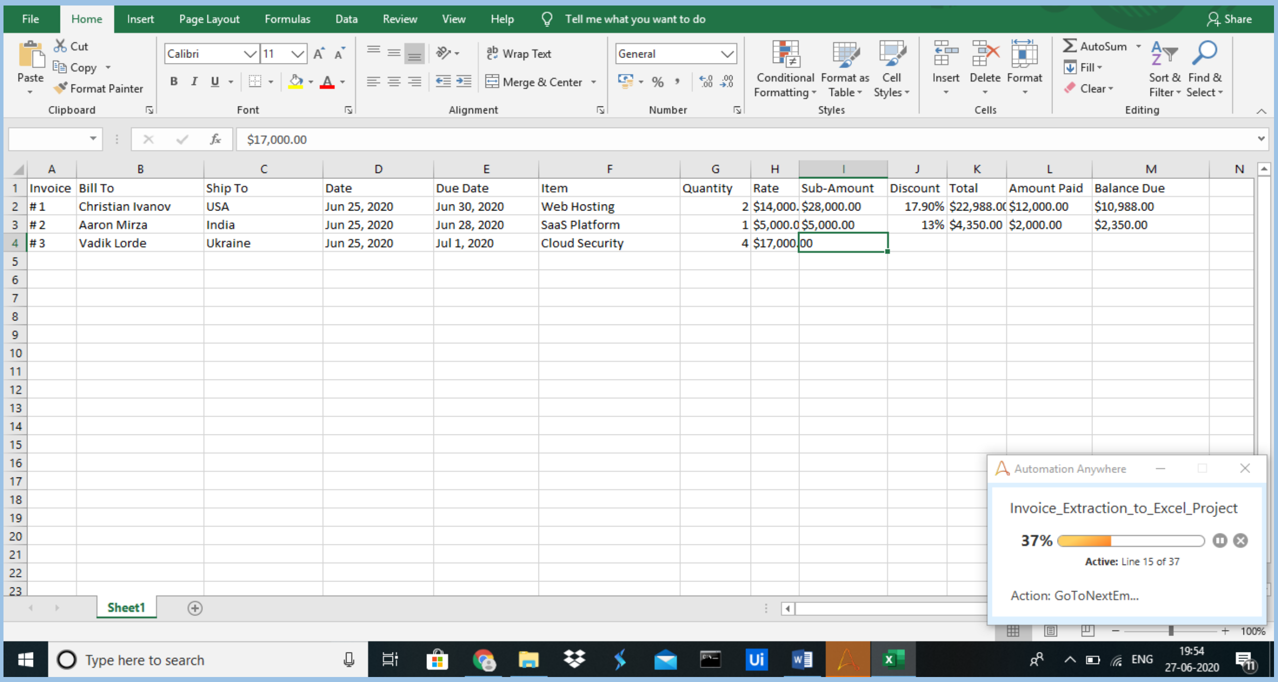
IMPROVEMENTS:

As soon as the ‘Extract Form Field’ or an alternate is added in the A2019 version a better more efficient bot can be created.



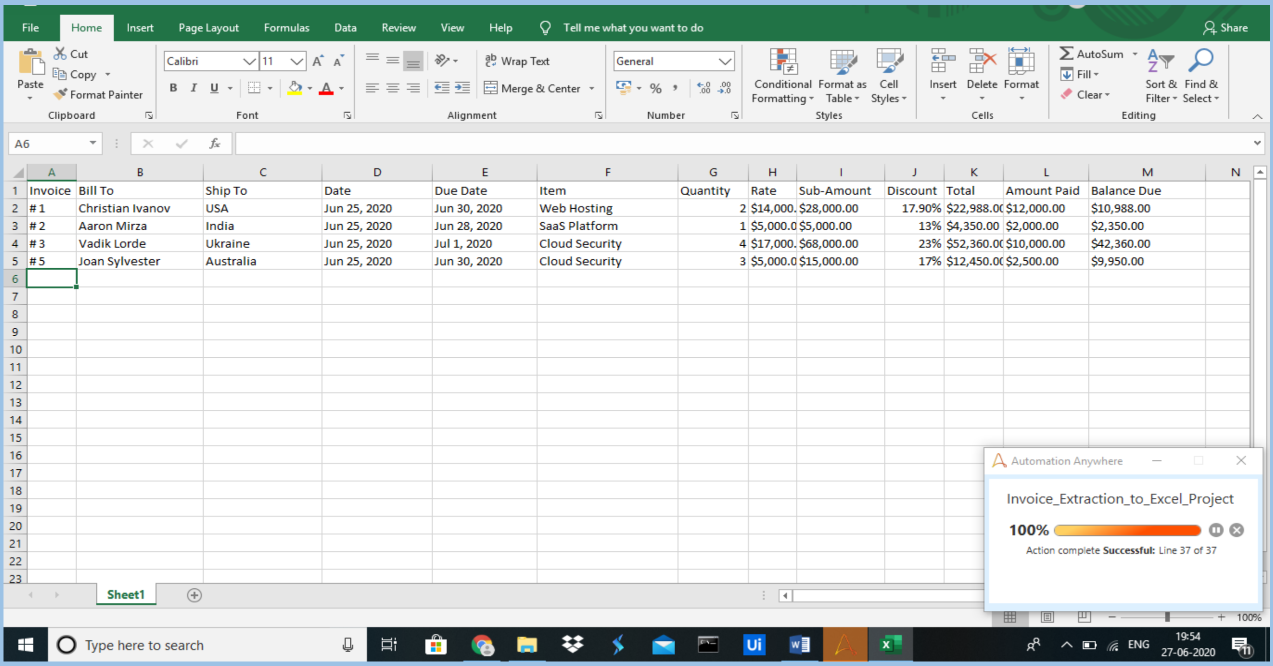
```
6   CSV/TXT: Open "C:\Users\AKSHAYA KUMAR\Desktop\INVOICE PROJECT\Text\
7   CSV/TXT: Read data and assign to $Table_Value$
8   Excel advanced: Set cell : Active cell to $Table_Value[0][0]$
9   Excel advanced: Go to next empty cell from active cell towards right
10  Excel advanced: Set cell : Active cell to $Table_Value[3][0]$
```

Image 7: Bot Running image 1



	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Invoice	Bill To	Ship To	Date	Due Date	Item	Quantity	Rate	Sub-Amount	Discount	Total	Amount Paid	Balance Due
2	# 1	Christian Ivanov	USA	Jun 25, 2020	Jun 30, 2020	Web Hosting	2	$14,000.	$28,000.00	17.90%	$22,988.0(	$12,000.00	$10,988.00
3	# 2	Aaron Mirza	India	Jun 25, 2020	Jun 28, 2020	SaaS Platform	1	$5,000.0	$5,000.00	13%	$4,350.00	$2,000.00	$2,350.00
4	# 3	Vadik Lorde	Ukraine	Jun 25, 2020	Jul 1, 2020	Cloud Security	4	$17,000.	00				

Image 8: Bot Running image 1



Invoice	Bill To	Ship To	Date	Due Date	Item	Quantity	Rate	Sub-Amount	Discount	Total	Amount Paid	Balance Due
# 1	Christian Ivanov	USA	Jun 25, 2020	Jun 30, 2020	Web Hosting	2	$14,000.	$28,000.00	17.90%	$22,988.0(	$12,000.00	$10,988.00
# 2	Aaron Mirza	India	Jun 25, 2020	Jun 28, 2020	SaaS Platform	1	$5,000.0	$5,000.00	13%	$4,350.00	$2,000.00	$2,350.00
# 3	Vadik Lorde	Ukraine	Jun 25, 2020	Jul 1, 2020	Cloud Security	4	$17,000.	$68,000.00	23%	$52,360.0(	$10,000.00	$42,360.00
# 5	Joan Sylvester	Australia	Jun 25, 2020	Jun 30, 2020	Cloud Security	3	$5,000.0	$15,000.00	17%	$12,450.0(	$2,500.00	$9,950.00

Automation Anywhere

Invoice_Extraction_to_Excel_Project

100%

Action complete **Successful:** Line 37 of 37

Image 8: Bot Running image 1

LINK FOR CREATING INVOICES:

https://invoice-generator.com/#/9