

Contents

Introduction	2
Project Dependencies	2
Updating Variable	3
Deep Dive into Workflow	3
Build Data Table Activity	4
Assign Activity 1 st	4
Assign Activity 2 nd	5
For Each Activity 1 st	5
Read PDF Text Activity	5
Matches 1 st Activity	6
Matches 2 nd Activity	6
Matches 3 rd Activity	8
Matches 4 th Activity	9
Matches 5 th Activity	10
Assign 1 st Activity	11
Write Line 1 st Activity	11
Write Line 2 nd Activity	11
Parallel Activity	12
For Each 1 st Activity	12
Assign Activity	12
Add Data Row Activity	12
For Each 2 nd Activity	13
Assign Activity	13
Add Data Row Activity	13
For Each 3 rd Activity	14
Assign Activity	14
Add Data Row Activity	14
Assign 2 nd Activity	14
Write Range Activity	15

Introduction

This Pdf Data Extraction Model for Etisalat Ebills helps companies to extract Calls, SMS(Premium Parking and other kinds), purchases made from the Play Store, App Store, and other services activated. Using this data, we can easily analyze the data and find out exactly where to reduce the cost.

This Workflow will output a single Excel with 3 sheets each one for International, Local, and SMS/Purchase. Below shows the preview of the output once our workflow is completed.

Local Calls

	B	C	D	E	F	G	H
1	Calling Number	Date	Time	Dialled Number	Duration	Amount	PlanCost
2	43xxxxxx	5/1/2021	09:08:59	052xxxxxx	00:01:00	0.30	880.00
3	56xxxxxx	5/1/2021	10:45:45	055xxxxxx	00:01:00	0.30	100.00

International Calls

	A	B	C	D	E	F	G
1	Calling Number	Date	Time	Dialled Number	Country	Duration	Amount
2	4xxxxxx	5/4/2021	15:54:11	0039xxxxxxxxxx	Italy	00:01:00	2.12
3	4xxxxxx	5/5/2021	09:49:11	00973xxxxxx	Bahrain	00:07:00	14.84
4	4xxxxxx	5/5/2021	16:48:53	00917xxxxxxxxxx	India	00:01:00	2.4
5	50xxxxxx	5/31/2021	12:26:17	0041xxxxxxxxxx	Switzerland	00:23:18	0
6	50xxxxxx	5/31/2021	15:01:19	0032xxxxxxxxxx	Belgium	00:00:06	0

SMS

	A	B	C	D	E	F
1	Calling Number	Date	Time	Dialled Number	Duration	Amount
2	501xxxxxx	5/1/2021	05:51:45	RTAmParking	1100	11
3	50xxxxxx	5/17/2021	16:00:12	SharjahmParking	200	2
4	50xxxxxx	5/2/2021	18:34:32	Huwaei_App_Gallery	330	3.3
5	50xxxxxx	5/31/2021	13:10:49	AppleMLJG1H08M8	1099	10.99
6	54xxxxxx	5/1/2021	12:31:10	EtisalatVideoStore	300	3
7	56xxxxxx	5/1/2021	04:07:59	Applicateka	200	2
8	564xxxxxx	5/6/2021	01:45:08	MobiPlanet	1000	10
9	565xxxxxx	5/3/2021	01:07:06	Playzone	1100	11

Project Dependencies

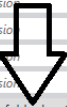
- Uiopath.PDF.Activities

Updating Variable

Before Starting the Process, you need to update 2 variables for the process to run.

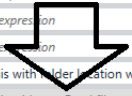
- **SourcePDFLocation String** variable
 - We need to provide the folder location where all the Etisalat Ebills(PDF) saved
 - For Example – “c:\user\username\Desktop\New Folder”

Name	Variable type	Scope	Default
dtLocalMobileSPSM	DataTable	Sequence	Enter a VB expression
dtInternationalCalls	DataTable	Sequence	Enter a VB expression
dtParkingSMS	DataTable	Sequence	Enter a VB expression
count	Int32	Sequence	Enter a VB expression
ListofPDFfiles	String[]	Sequence	Enter a VB expression
strdate	String	Sequence	Enter a VB expression
SourcePDFLocation	String	Sequence	*Replace this with folder location where all the PDF bills Save



- **ExcelexportAndOutputLocation String** variable
 - We need to update this variable with the excel file location with the file name and extension
 - For Example – “c:\user\username\Desktop\data.xlsx”

Name	Variable type	Scope	Default
dtLocalMobileSPSM	DataTable	Sequence	Enter a VB expression
dtInternationalCalls	DataTable	Sequence	Enter a VB expression
dtParkingSMS	DataTable	Sequence	Enter a VB expression
count	Int32	Sequence	Enter a VB expression
ListofPDFfiles	String[]	Sequence	Enter a VB expression
strdate	String	Sequence	Enter a VB expression
SourcePDFLocation	String	Sequence	*Replace this with folder location where all the PDF bills Save
ExcelexportAndOutputLocation	String	Sequence	*Replace this with you Excel file name location with extension



Deep Dive into Workflow

We will now look into all the activities and how they work in our workflow to extract the exact data we require for our analytics.

Activities that are used to achieve this workflow are as follows and we will looking into how these activities are used in detail in the further section and sections will flow sequentially mirroring the workflow

- Build Data Table
- Add Data Row
- Assign
- For Each
- Read PDF Text
- Matches
- Write Line
- Parallel Flow
- Write Range

Build Data Table Activity

We are creating 3 “Build Data Table” activity each one for International, Local, and SMS/Purchase.

- Local “Build Data Table” contains the following fields and created “**dtLocalMobileSPSMS**” as output variable

	Calling Number (Int32)	Date (DateTime)	Time (String)	Dialed Number (String)	Duration (String)	Amount (String)	PlanCost (String)
+							
x							

- International “Build Data Table” contains the following fields and created “**dtInternationalCalls**” as output variable

	Calling Number (Int32)	Date (DateTime)	Time (String)	Dialed Number (String)	Country (String)	Duration (String)	Amount (String)	PlanCost (String)
+								
x								

- SMS/Purchase “Build Data Table” contains the following fields and created “**dtParkingSMS**” as output variable

	Calling Number (Int32)	Date (DateTime)	Time (String)	Dialed Number (String)	Duration (String)	Amount (String)	PlanCost (String)
+							
x							

Assign Activity 1st

- Creating 1st Assign activity for the incremental purpose to keep track of how many PDFs are being processed
- Creating **Count** as **Int32** variable

Assign Activity 2nd

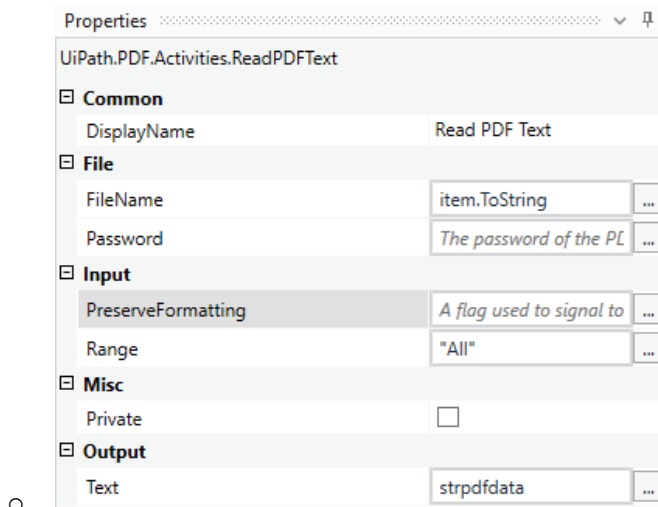
- Creating 2nd Assign activity for getting all the PDFs from the folder with 2 new variables created
 - 1st Variable - **ListofPDFfiles** as **String Array**
 - 2nd Variable – **SourcePDFLocation** as **String** and initialize it with the folder location which contains all the PDFs.

For Each Activity 1st

- Loop through all the PDF files and performing the below list of activities on every PDF file
 - Assigning **item** with **ListofPDFfiles** variable type **Object**

Read PDF Text Activity

- Converts the PDF into Text format and outputs the same to a **String** variable called **strpdfdata**



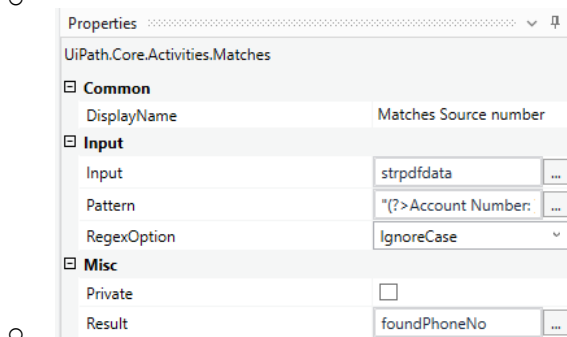
Matches 1st Activity

- This Activity will match the Mobile/landline number the bill associated
- It takes Input **strpdfdata** from the previous activity “Read PDF Text” and outputs the result to **foundPhoneNo** to **IEnumerable<Match>** type variable



- The above Regex will match all the patterns in the string out from the “Read PDF Text” Activity, below screenshot shows the exact output.

Account Number: 056XXXXXX



Matches 2nd Activity

- This Activity will match all the local calls made within the country and extract all the line items
- It takes Input **strpdfdata** from the previous activity “Read PDF Text” and outputs the result to **foundMatchesLMcalls** to **IEnumerable<Match>** type variable.

Matches 3rd Activity

- This Activity will match all the International Calls and extract all the line items
- It takes Input **strpdfdata** from the previous activity “Read PDF Text” and outputs the result to **foundMatchesInternationalCalls** to **IEnumerable<Match>** type variable.

RegEx Builder

Test Text

RegEx	Value	Quantifiers		
Digit	\d	At least one (1 or more)	+	+ Δ ▾ ✕
Whitespace	\s	Exactly	1	+ Δ ▾ ✕
Any word character	\w	At least one (1 or more)	+	+ Δ ▾ ✕
Whitespace	\s	Exactly	1	+ Δ ▾ ✕
Digit	\d	At least one (1 or more)	+	+ Δ ▾ ✕
Whitespace	\s	Exactly	1	+ Δ ▾ ✕
Advanced	\d\d\d\d\d	Exactly	1	+ Δ ▾ ✕
Whitespace	\s	Exactly	1	+ Δ ▾ ✕
Advanced		Exactly	1	+ Δ ▾ ✕
Whitespace	\s	Exactly	1	+ Δ ▾ ✕
Digit	\d	At least one (1 or more)	+	+ Δ ▾ ✕
Whitespace	\s	Exactly	1	+ Δ ▾ ✕
Advanced		Exactly	1	+ Δ ▾ ✕
Whitespace	\s	Exactly	1	+ Δ ▾ ✕
Any word character	\w	At least one (1 or more)	+	+ Δ ▾ ✕
Whitespace	\s	Exactly	1	+ Δ ▾ ✕
Advanced	\d\d\d\d\d	Exactly	1	+ Δ ▾ ✕
Whitespace	\s	Exactly	1	+ Δ ▾ ✕
Advanced	(\d+)	Exactly	1	+ Δ ▾ ✕
Advanced	\.	Exactly	1	+ Δ ▾ ✕
Advanced	(\d+)	Exactly	1	+ Δ ▾ ✕

Full Expression

(\d+)\s(w+)\s(\d+)\s\d\d\d\d\d\d\d\d\d\s(\d+)\s(\s(w+)\s\d\d\d\d\d\d\d\d\d\s(\d+)\s(\d+))

Regex Option

☒ IgnoreCase

Save

Cancel

- Using the Regex on the below output to match all the lines

```
01 May 2021 10:51:38 Ĩ 0091XXXXXXXXX Í India 00:00:11 0.00
```


Matches 4th Activity

- This Activity will match all the SMS and extract all the line items
- It takes Input **strpdfdata** from the previous activity “Read PDF Text” and outputs the result to **foundMatchesforParkingSMS** to **IEnumerable<Match>** type variable

RegEx Builder

Test Text

RegEx	Value	Quantifiers
Digit	\d	At least one (1 or more)
Whitespace	\s	Exactly 1
Any word character	\w	At least one (1 or more)
Whitespace	\s	Exactly 1
Digit	\d	At least one (1 or more)
Whitespace	\s	Exactly 1
Advanced	\d\d\d\d\d\d	Exactly 1
Whitespace	\s	Exactly 1
Advanced		Exactly 1
Whitespace	\s	Exactly 1
Advanced	(?<=)(.+?)(?=)	At least one (1 or more)
Whitespace	\s	Exactly 1
Advanced	f	Exactly 1
Digit	\d	At least one (1 or more)
Whitespace	\s	Exactly 1
Advanced	(\d+)	Exactly 1
Advanced	\s	Exactly 1
Advanced	(\d+)	Exactly 1

Full Expression

(\d+)\s(\w+)\s(\d+)\s\d\d\d\d\d\d\s(?<=|)(.+?)(?=|)\sf(\d+)\s(\d+)\s(\d+)

Regex Option

☒ IgnoreCase

Save Cancel

- Using the Regex on the below output to match all the lines

30 May 2021 05:32:22 Ì RTA mParking Í 1100 11.00

Properties

UiPath.Core.Activities.Matches

Common

DisplayName Matches for Parking SMS

Input

Input strpdfdata

Pattern "(\\d+)\\s(\\w+)\\s(\\d+)\\s"

RegexOption IgnoreCase

Misc

Private

Result foundMatchesforParki

Matches 5th Activity

- This Activity will match the Monthly Cost of the Plan and extract it
- It takes Input **strpdfdata** from the previous activity “Read PDF Text” and outputs the result to **foundMatchesforParkingSMS** to **IEnumerable<Match>** type variable

RegEx Builder

Test Text

RegEx	Value	Quantifiers
Advanced ⓘ	(?<=Service Rentals)	Exactly 1
Advanced ⓘ	(\d+)	Exactly 1
Advanced ⓘ	\.	Exactly 1
Literal ⓘ	(\d+)	Exactly 1

Full Expression

(?<=Service Rentals)(\d+)\.(\d+)

Regex Option

☒ IgnoreCase

Save Cancel

- Using the Regex on the below output to match exact details

Service Rental 100.00 100.00 5.00

Properties

UiPath.Core.Activities.Matches

Common

DisplayName Matches for Monthly Plan c

Input

Input strpdfdata

Pattern "(?<=Service Rentals)"

RegexOption IgnoreCase

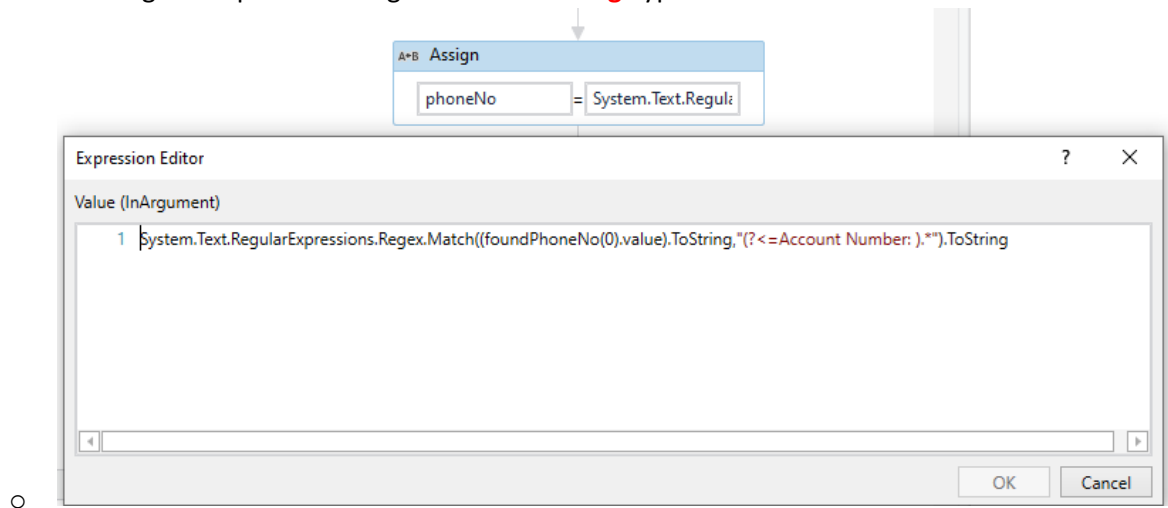
Misc

Private

Result foundMonthlyplanCos

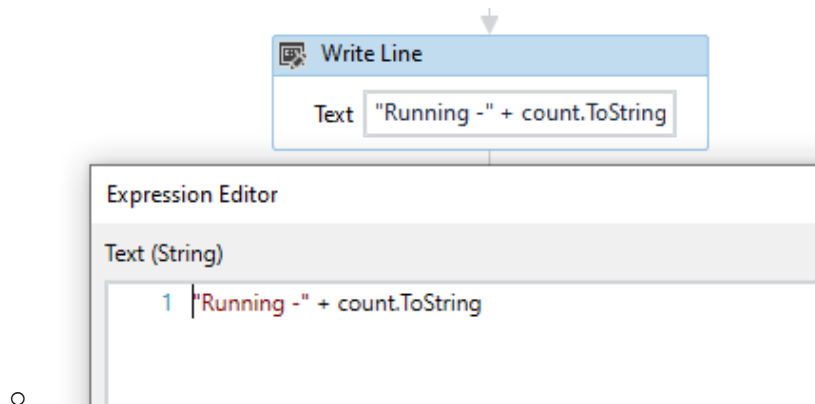
Assign 1st Activity

- We are further extracting the phone number to get the exact value using `System.Text.RegularExpressions.Regex.Match` to **String** Type **variable PhoneNo**



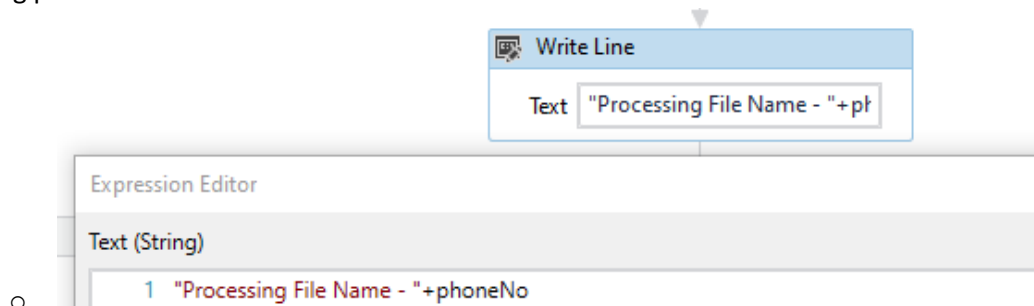
Write Line 1st Activity

- For our logging purpose, this activity will write line count of PDF which is currently being processed



Write Line 2nd Activity

- For our logging purpose, this activity will display the phone number associated with the PDF being processed.



Parallel Activity

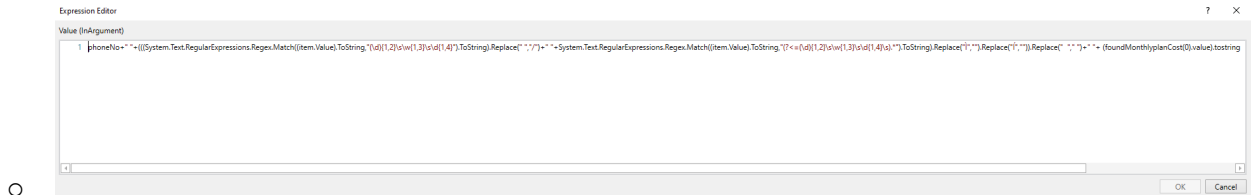
- Using this Activity we will run further 3 more For EACH loop to match multiple lines and Add them to the Data Table using ADD Data Row Activity all these activities will run parallelly.

For Each 1st Activity

- Loop through the **foundMatchesLMcalls** output from Local Call Matches

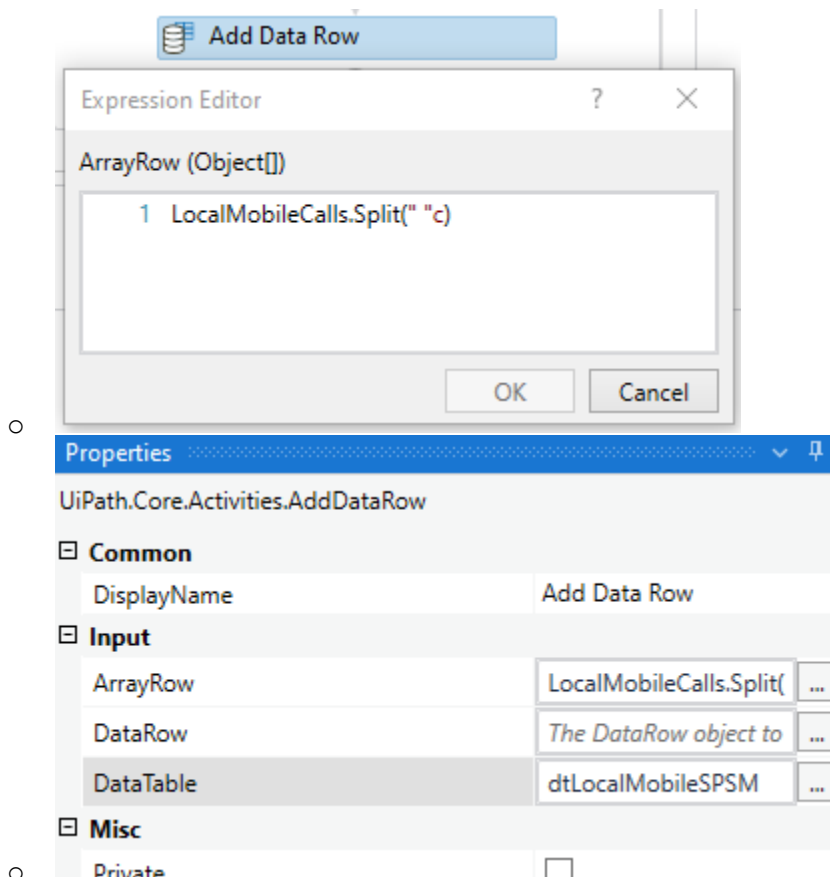
Assign Activity

- Using the Assign Activity we will manipulate and extract the data in the correct format and assign it to **LocalMobileCalls String** Variable



Add Data Row Activity

- Using this Activity we will add the formatted from the **LocalMobileCalls** assign activity by splitting them by "Space" and adding them to "Array Row" and then add **dtLocalMobileSPSM** to the DataTable

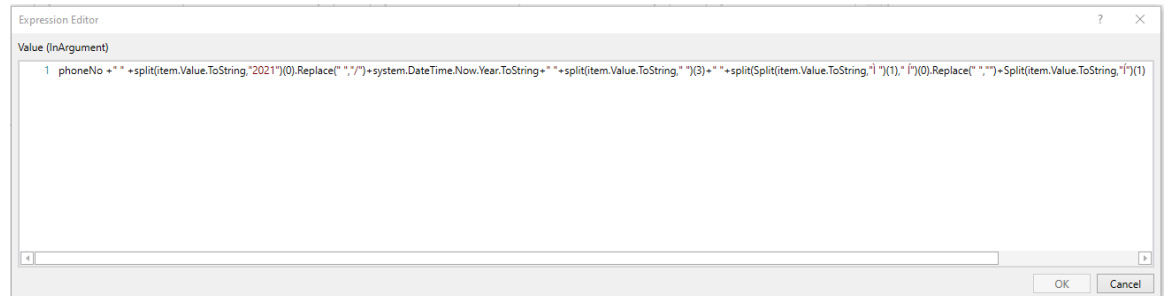


For Each 2nd Activity

- Loop through the **foundMatchesforParkingSMS** output from Local Call Matches

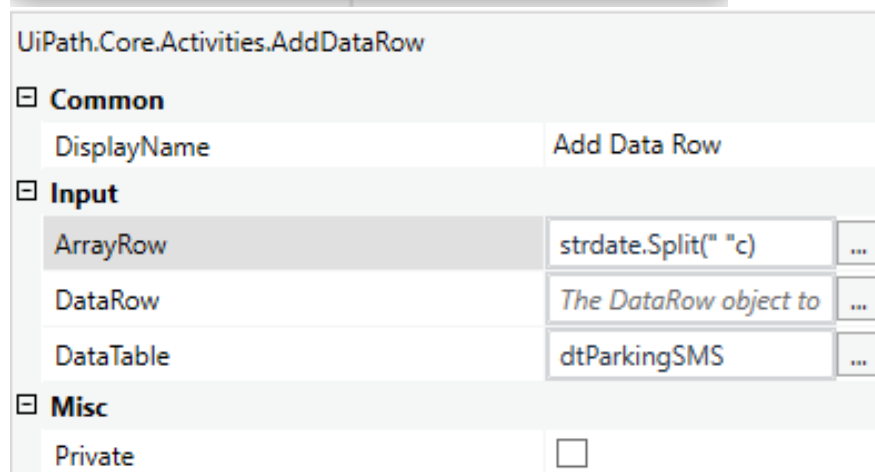
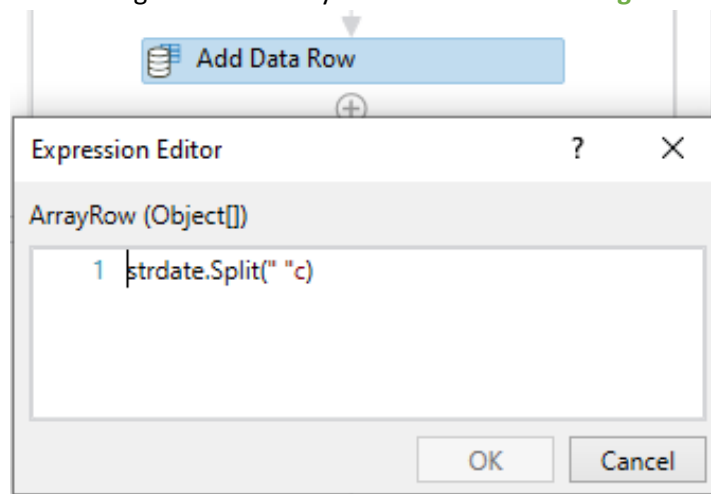
Assign Activity

- Using the Assign Activity we will manipulate and extract the data in the correct format and assign it to **strdate String** Variable



Add Data Row Activity

- Using this Activity we will add the formatted from the **strdate** assign activity by splitting them by "Space" and adding them to "Array Row" and add **dtParkingSMS** to DataTable

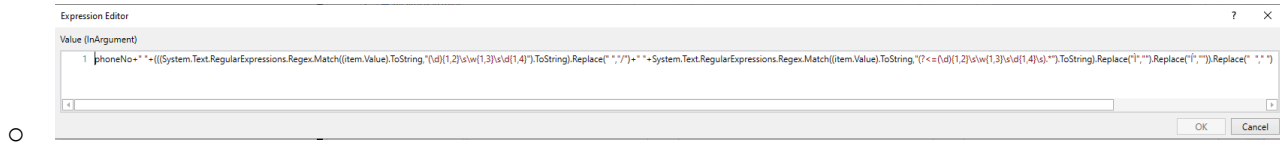


For Each 3rd Activity

- Loop through the **foundMatchesInternationalCalls** output from Local Call Matches

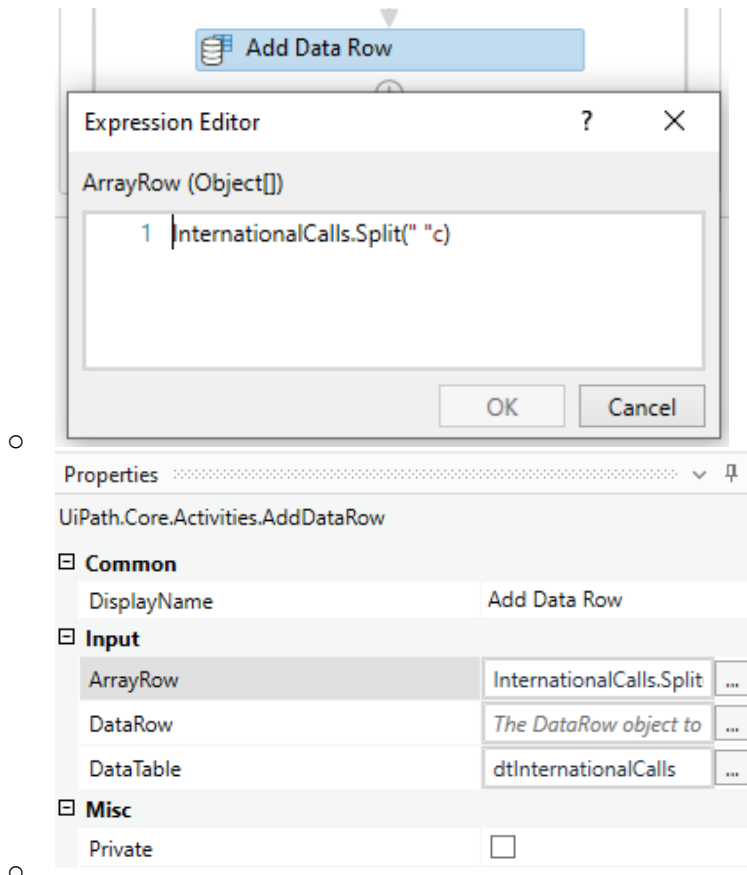
Assign Activity

- Using the Assign Activity we will manipulate and extract the data in the correct format and assign it to **InternationalCalls** String Variable



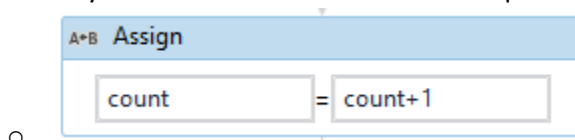
Add Data Row Activity

- Using this Activity we will add the formatted from the **InternationalCalls** assign activity by splitting them by "Space" by adding them to "Array Row" and add **dtInternationalCalls** to DataTable



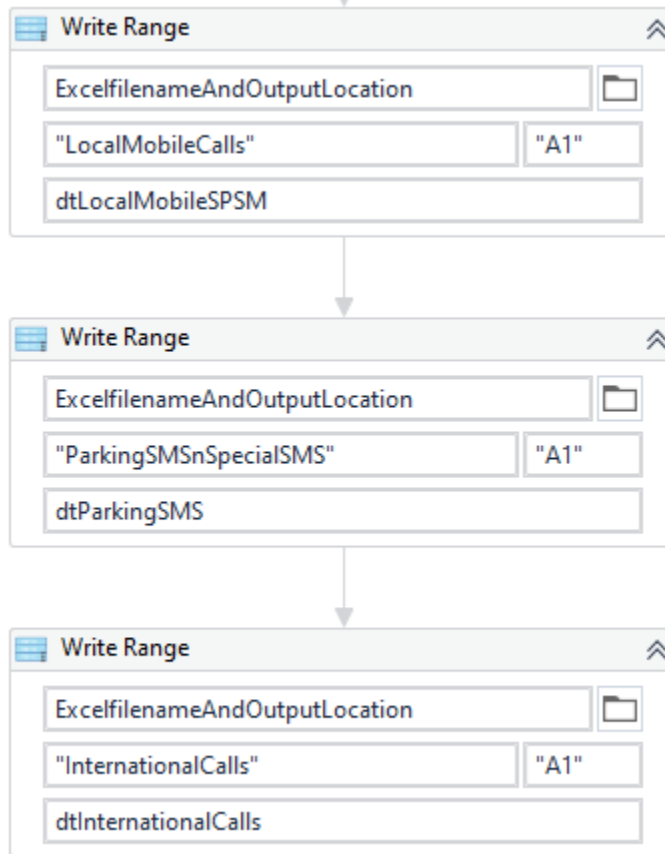
Assign 2nd Activity

- This activity will increment the counter to keep track of how many files has been processed



Write Range Activity

- We now use 3 Write Range Activity to add the DataTables to the respective sheet for output



○