

## Using Tableau Prep Builder

Tableau Prep is used by several Big Four accounting firms to automate the ETL process. In this exercise, you will **Extract** the General Journal file, **Transform** the data into the required format (a General Ledger and an Unadjusted Trial Balance), and subsequently **Load** the transformed data into Excel.

The journal entries for *A Byte of Accounting* are stored in an Excel sheet titled General Journal. Notice that the column headings are in the fourth row, and some rows do not have account numbers.

Begin with the General Journal

A Byte of Accounting, Inc. General Journal							
Note: You can only enter data into the yellow filled cells.							
Transaction	Date	Account	Name	Description	Debit	Credit	
01	Jun 01	1110	Cash	Investment from Mark Friedman	28,000.00		
01	Jun 01	3100	Capital Stock	Investment from Mark Friedman		28,000.00	
02	Jun 01	1211	Office Equip.	Hudson equipment Invoice BC3887	4,000.00		
02	Jun 01	2101	Accounts Payable	Hudson equipment Invoice BC3887		4,000.00	
08	Jun 14	1110	Cash	Services performed by Lucus Pictures	11,000.00		
08	Jun 14	4100	Computer & Consulting Revenue	Services performed by Lucus Pictures		11,000.00	

End with the General Ledger and the Unadjusted Trial Balance

Account Name	Transaction	Date	Description	Debit	Credit	Running Balance
1110 Cash	1	6/1/2018	Investment from Mark Friedman	\$28,000.00	\$ -	\$ 28,000.00
1110 Cash	3	6/1/2018	Hailey Computers 87245, ck6001	\$ -	\$ 7,000.00	\$ 21,000.00
1110 Cash	5	6/4/2018	Scanner, Jake Supplies, 54-541 ck6002	\$ -	\$ 125.00	\$ 20,875.00
1110 Cash	7	6/8/2018	Pd A/p ck6003	\$ -	\$ 640.00	\$ 20,235.00
1110 Cash	8	6/14/2018	Services performed by Lucus Pictures	\$11,000.00	\$ -	\$ 31,235.00
1211 Office Equip.	2	6/1/2018	Hudson equipment Invoice BC3887	\$ 4,000.00	\$ -	\$ 4,000.00
1211 Office Equip.	5	6/4/2018	Scanner, Jake Supplies, 54-541 ck6002	\$ 125.00	\$ -	\$ 4,125.00
1311 Computer Equip.	3	6/1/2018	Hailey Computers 87245, ck6001	\$ 7,000.00	\$ -	\$ 7,000.00
2101 Accounts Payable	2	6/1/2018	Hudson equipment Invoice BC3887	\$ -	\$ 4,000.00	\$ (4,000.00)
2101 Accounts Payable	4	6/2/2018	Avery Repairs, Invoice 25478	\$ -	\$ 725.00	\$ (4,725.00)
2101 Accounts Payable	6	6/8/2018	Zac Advertising,23547	\$ -	\$ 3,380.00	\$ (8,105.00)
2101 Accounts Payable	7	6/8/2018	Pd A/p ck6003	\$ 640.00	\$ -	\$ (7,465.00)
3100 Capital Stock	1	6/1/2018	Investment from Mark Friedman	\$ -	\$28,000.00	\$ (28,000.00)
4100 Computer & Consulting Revenue	8	6/14/2018	Services performed by Lucus Pictures	\$ -	\$11,000.00	\$ (11,000.00)
5030 Advertising Expense	6	6/8/2018	Zac Advertising,23547	\$ 3,380.00	\$ -	\$ 3,380.00
5040 Repairs & Maint. Expense	4	6/2/2018	Avery Repairs, Invoice 25478	\$ 725.00	\$ -	\$ 725.00

Account And Name	Balance
1110 Cash	31235
1211 Office Equip.	4125
1311 Computer Equip.	7000
2101 Accounts Payable	-7465
3100 Capital Stock	-28000
4100 Computer & Consulting Revenue	-11000
5030 Advertising Expense	3380
5040 Repairs & Maint. Expense	725

## Start Tableau Prep

The Tableau Prep Builder workspace consists of the Connections pane and three coordinated areas that help you interact with and explore your data:

- Flow pane: A visual representation of your operation steps as you prepare your data.
- Profile pane: A summary of each field in your data sample. See the shape of your data and quickly find outliers and nulls.
- Data grid pane: The row level detail for your data.

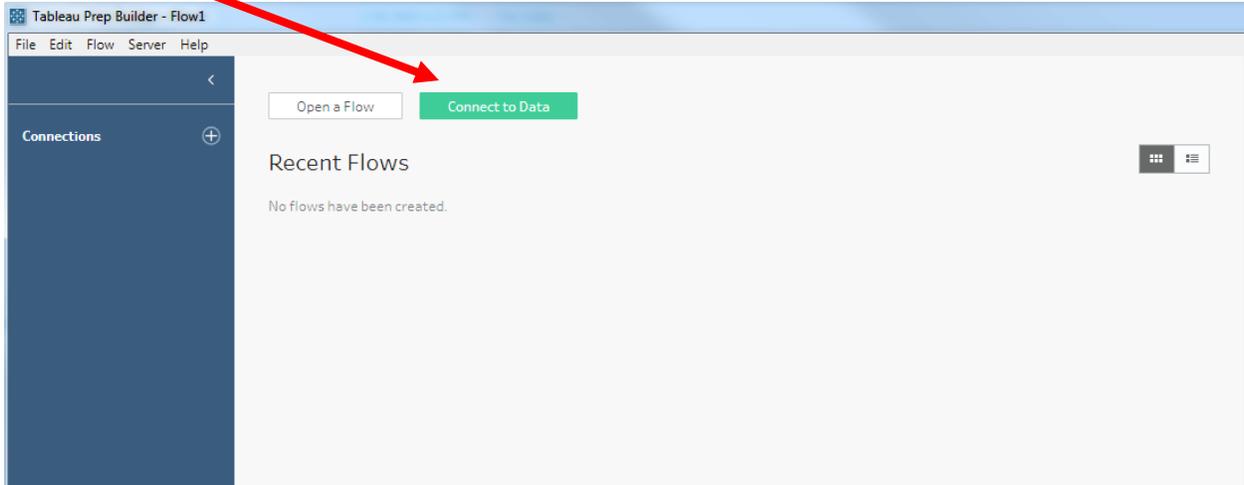
The screenshot displays the Tableau Prep Builder workspace. On the left is the 'Connections' pane showing a file named 'MF8916.xlsx'. Below it is the 'Tables' pane with a list of tables including 'Balance Sheet', 'Changes in Retained E...', 'General Journal Part 2', 'General Journal', 'Income Statement', and several 'Balance Sheet' and 'General Journal' tables. The main workspace is divided into three panes:

- Flow pane:** Shows a visual flow with 'General Journal' as the source and 'Clean 1' as the operation.
- Profile pane:** Displays a summary of the data fields: Transaction (40), Date (6), Account (9), Name (9), Description (9), Debit (9), and Credit (9). It includes a bar chart for the 'Transaction' field and a list of values for other fields.
- Data pane:** Shows a table of data with columns: Transaction, Date, Account, Name, Description, Debit, Credit, and Total Balance. The table contains several rows of data, including transactions from 06/01/2018 to 06/14/2018.

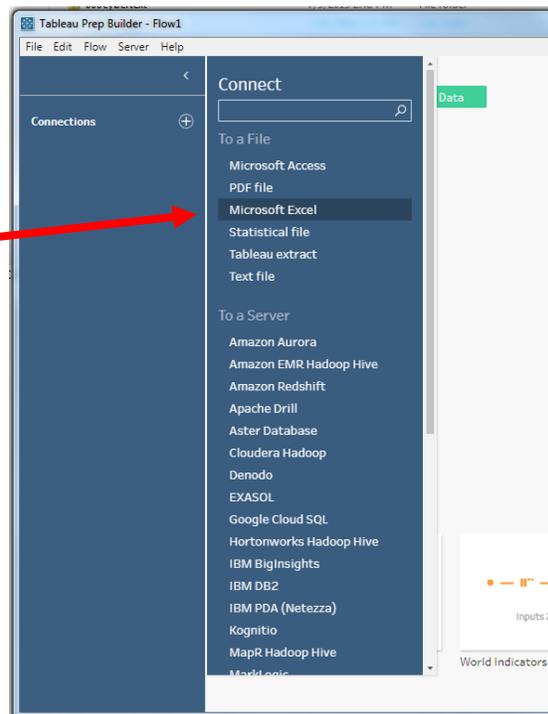
After you connect to your data and begin building your flow, you add steps in the Flow pane. These steps function as a lens into the structure of your data, as well as a summary of operations that is applied to your data. Each step represents a different category of operations that you define.

### Step 1: Input Excel file

To open Excel file via Tableau Prep, press green button “Connect to Data” on the Flow pane.

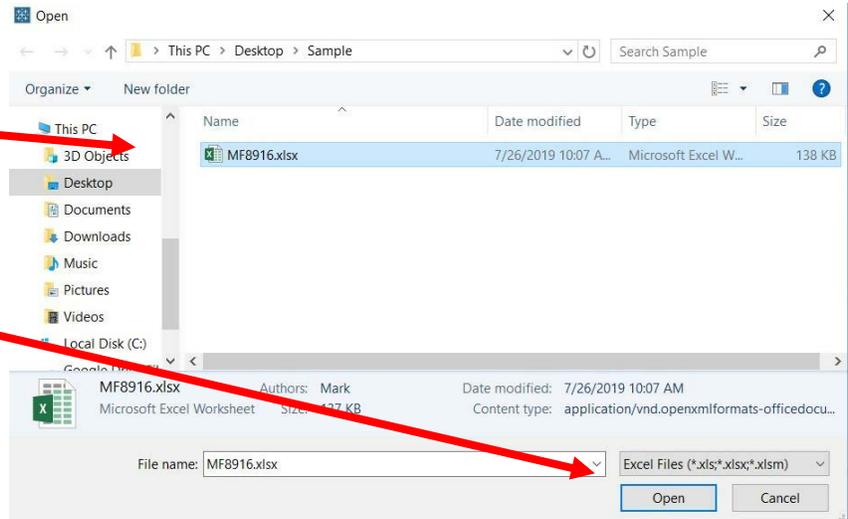


On the connections pane select Microsoft Excel.

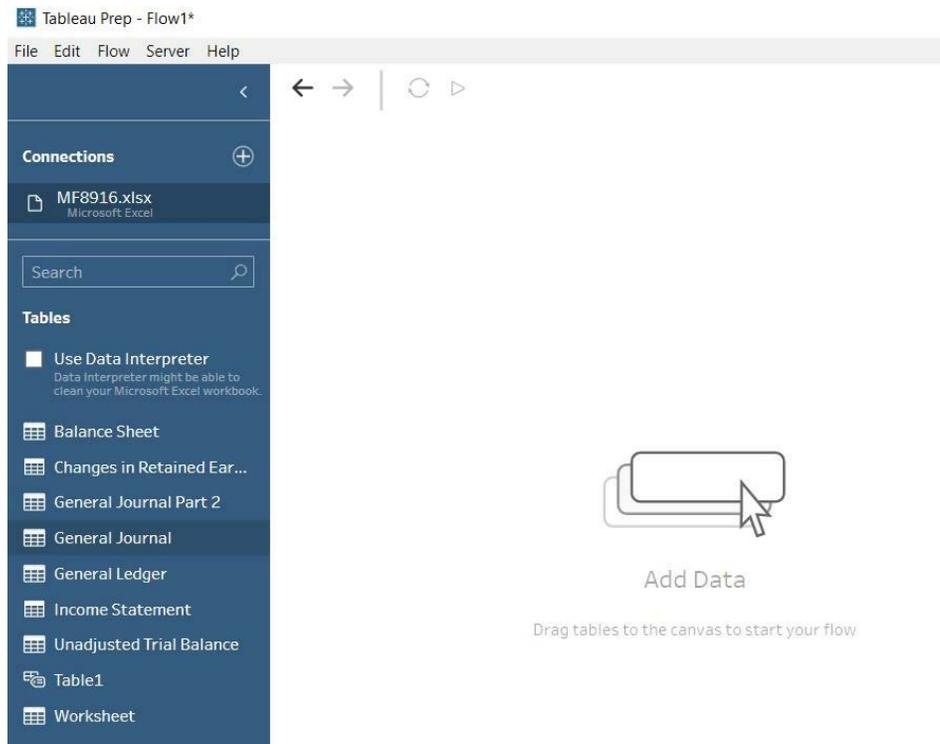


Select the file.

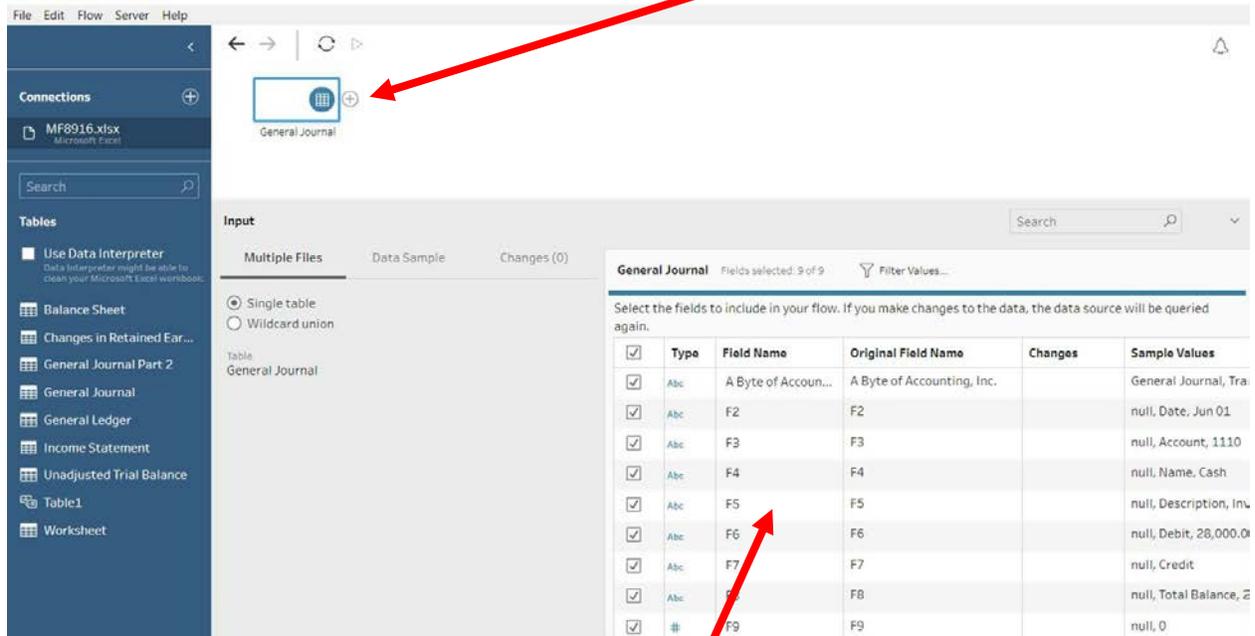
Select "Open".



After you open the file, select the "General Journal" sheet on the Connections pane. Either double click on the sheet name or drag the sheet name to the canvas.



After selecting the sheet, “General Journal” will appear on the Flow pane.



Sample data will be visible on the Profile pane.

## Step 2: Cleaning the data

Remember that the field names are in the fourth row.

Sometimes Tableau can find the headers by checking the box to the left of “Cleaned with Data Interpreter”.

The screenshot shows the Tableau Prep interface. On the left, the 'Connections' pane shows 'MF8916.xlsx' with the 'Cleaned with Data Interpreter' checkbox checked. The 'Tables' list includes 'General Journal'. The 'Input' pane shows 'General Journal' selected. The main pane displays a table with the following columns: Type, Field Name, Original Field Name, Changes, and Sample Values. The 'Field Name' column is highlighted with a red box, and the 'Original Field Name' column is also highlighted with a red box. A red arrow points from the 'Cleaned with Data Interpreter' checkbox to the 'Field Name' column header.

Type	Field Name	Original Field Name	Changes	Sample Values
#	Transaction	Transaction		1, null
#	Date	Date		06/01/2018, null
#	Account	Account		1,110, 3,100, null
Abc	Name	Name		Cash, Capital Stock
Abc	Description	Description		Investment from M
#	Debit	Debit		28,000, null
#	Credit	Credit		null, 28,000
#	Total Balance	Total Balance		28,000, -28,000, 0
#	F9	F9		0

If it works, the field names will be displayed correctly.

If not, double click on the “Field Name” and change the names as shown below.

The first table shows the 'Field Name' column with the value 'Accounting, Inc.' highlighted in blue. A red arrow points from the 'Field Name' header to this cell. The second table shows a list of field names and their original names. A red arrow points from the 'Field Name' header to the first row of this table.

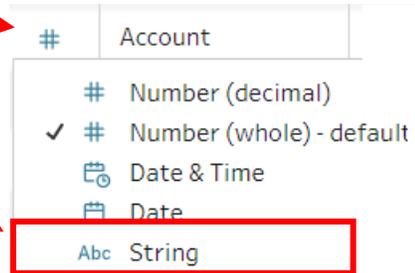
Field Name	Original Field Name
Transaction	A Byte of Accounting, Inc.
Date	F2
Account	F3
Name	F4
Description	F5
Debit	F6
Credit	F7
Total Balance	F8
F9	F9

In a report the field “Account” will be combined with the field “Name”. Strings can be easily combined with other strings and we will not be doing any math calculations on the field “Account”.

All number fields that will be added, subtracted, multiplied, divided must be coded as “Number (decimal)”.

Change the field “Type” by selecting the “Type” to be changed and selecting the desired “Type”.

- “Account”: “String”
- “Debit”: “Number (decimal)”
- “Credit”: “Number (decimal)”
- “Total Balance”: “Number (decimal)”



**General Journal** Fields selected: 9 of 9 Filter Values...

Select the fields to include in your flow, apply a filter, or change data types. To see and clear in the flow pane.

<input checked="" type="checkbox"/>	Type	Field Name	Original Field Name	Changes
<input checked="" type="checkbox"/>	#	Transaction	Transaction	
<input checked="" type="checkbox"/>	📅	Date	Date	
<input checked="" type="checkbox"/>	Abc	Account	Account	🔗
<input checked="" type="checkbox"/>	Abc	Name	Name	
<input checked="" type="checkbox"/>	Abc	Description	Description	
<input checked="" type="checkbox"/>	#	Debit	Debit	🔗
<input checked="" type="checkbox"/>	#	Credit	Credit	🔗
<input checked="" type="checkbox"/>	#	Total Balance	Total Balance	🔗
<input checked="" type="checkbox"/>	#	F9	F9	

Deleting the check from the box on the left of the field “F9” deletes the column.

<input checked="" type="checkbox"/>	#	Debit	F6	
<input checked="" type="checkbox"/>	#	Credit	F7	
<input checked="" type="checkbox"/>	#	Total Balance	F8	
<input type="checkbox"/>	#	F9	F9	

In the “Flow Pane” select the “+” on the right of the “General Journal” input.

Select “+ Clean Step”

General Journal

- Add:
- + Clean Step
- Σ Aggregate
- Pivot
- Join
- Union
- Script
- Output
- Insert Flow

**Flow Pane**

**Profile Pane**

**Data Pane**

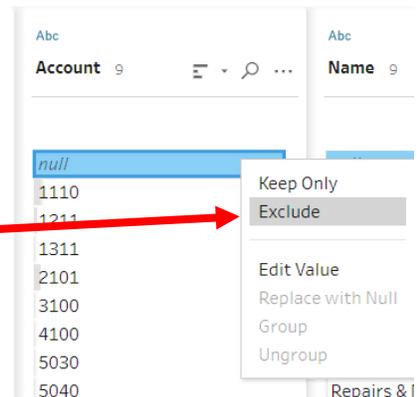
Transaction	Date	Account	Name	Description	Debit	Credit
1	06/01/2018	1,110	Cash	Investment from M...		
1	06/01/2018	3,100	Capital Stock	Investment from l...		
1	06/02/2018	1,211	Accounts Payable	Avery Repairs, Invoice 25478	125	125
1	06/04/2018	1,311	Advertising Expense	Hailey Computers 87245, ck...	640	640
1	06/08/2018	2,101	Capital Stock	Hudson equipment Invoice B...	725	725
1	06/14/2018	3,100		from Mark Fried...	3,380	3,380
1		4,100		3	4,000	4,000
1		5,030		e Supplies, 54-5...	7,000	7,000
1		5,040		formed by Lucu...	11,000	11,000
1				ng.23547	28,000	28,000

**Step 3:**

**Exclude all rows that do not contain data**

There are several rows that do not contain an account number. We can modify the working dataset so it includes only those rows that contain account numbers that are not "null".

Right click on the "null" value in the "Account" field and select "Exclude".

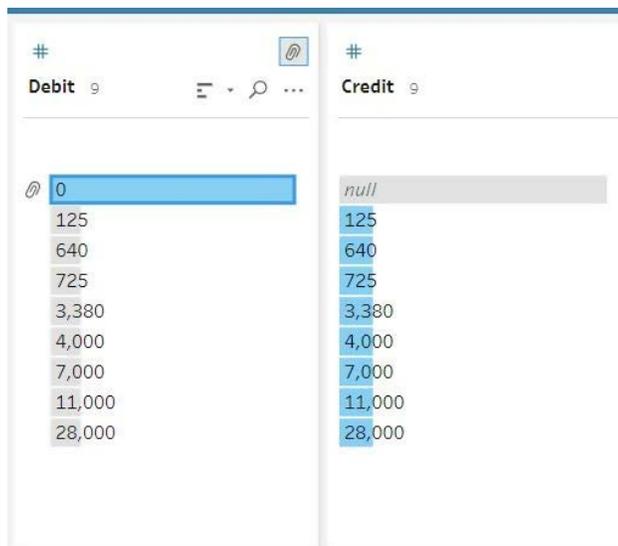


**Step 4: Change all occurrences of “null” to “0”.**

There are cells in the “Debit” and “Credit” columns that are coded as “null” in the “Data Pane”.

Transaction	Date	Account	Name	Description	Debit	Credit	Total Balance
1	06/01/2018	1110	Cash	Investment from M	28,000	null	28,000
1	06/01/2018	3100	Capital S	Investment from M	null	28,000	-28,000
2	06/01/2018	1211	Office E	Hudson equipment	4,000	null	4,000
2	06/01/2018	2101	Account	Hudson equipment	null	4,000	-4,000
3	06/01/2018	1311	Comput	Hailey Computers &	7,000	null	7,000

Double click the null value in “Debit” and change the “null” to a “0” and press “Enter”. Do the same to the “Credit” field.



Transaction	Date	Account	Name	Description	Debit	Credit	Total Balance
1	06/01/2018	1110	Cash	Investment from Mark Friedman	28,000	0	28,000
1	06/01/2018	3100	Capital Stock	Investment from Mark Friedman	0	28,000	-28,000
2	06/01/2018	1211	Office Equip.	Hudson equipment Invoice BC3887	4,000	0	4,000
2	06/01/2018	2101	Accounts Payable	Hudson equipment Invoice BC3887	0	4,000	-4,000
3	06/01/2018	1311	Computer Equip.	Hailey Computers 87245, ck6001	7,000	0	7,000

### Step 5 – Reorder the fields

In the “Data Pane” select the field “Account” and drag it to the first field position.

Transaction	Account	Date	Name	Description	Debit	Credit	Total Balance
1	1110	06/01/2018	Cash	Investment from Mark Friedman	28,000	0	28,000
1	3100	06/01/2018	Capital Stock	Investment from Mark Friedman	0	28,000	-28,000
2	1211	06/01/2018	Office Equip.	Hudson equipment Invoice BC3887	4,000	0	4,000
2	2101	06/01/2018	Accounts Payable	Hudson equipment Invoice BC3887	0	4,000	-4,000
3	1311	06/01/2018	Computer Equip.	Hailey Computers 87245, ck6001	7,000	0	7,000

In the Data Pane” select the field “Name” and drag it between the field “Account” and the field “Transaction”

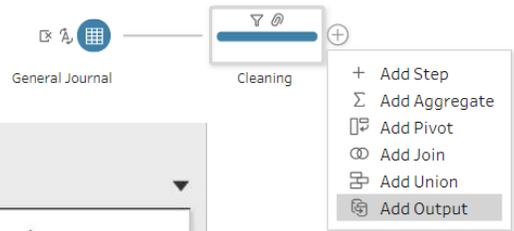
Account	Transaction	Date	Name	Description	Debit	Credit	Total Balance
1110	1	06/01/2018	Cash	Investment from Mark Friedman	28,000	0	28,000
3100	1	06/01/2018	Capital Stock	Investment from Mark Friedman	0	28,000	-28,000
1211	2	06/01/2018	Office Equip.	Hudson equipment Invoice BC3887	4,000	0	4,000
2101	2	06/01/2018	Accounts Payable	Hudson equipment Invoice BC3887	0	4,000	-4,000
1311	3	06/01/2018	Computer Equip.	Hailey Computers 87245, ck6001	7,000	0	7,000

Now the fields are in the order that we need for the General Ledger.

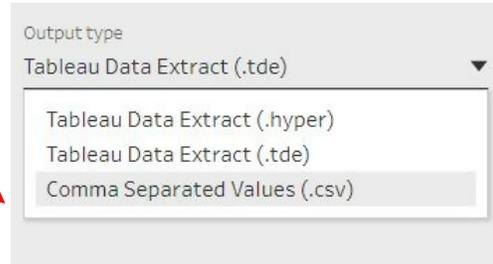
Account	Name	Transaction	Date	Description	Debit	Credit	Total Balance
1110	Cash	1	06/01/2018	Investment from Mark Friedman	28,000	0	28,000
3100	Capital Stock	1	06/01/2018	Investment from Mark Friedman	0	28,000	-28,000
1211	Office Equip.	2	06/01/2018	Hudson equipment Invoice BC3887	4,000	0	4,000
2101	Accounts Payable	2	06/01/2018	Hudson equipment Invoice BC3887	0	4,000	-4,000
1311	Computer Equip.	3	06/01/2018	Hailey Computers 87245, ck6001	7,000	0	7,000

### Step 6: Output the cleaned data that will be used to make the General ledger

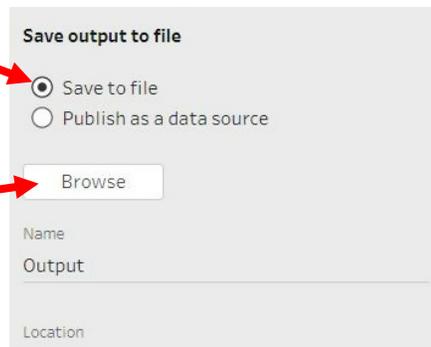
Click on the plus sign of our cleaning operation tile and select "Add Output"



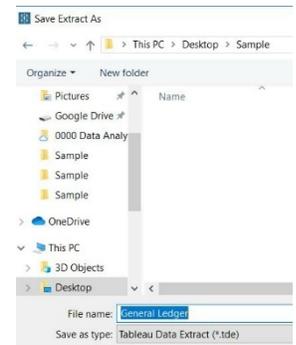
Select the .csv file type.



Select "Save to file" option button.

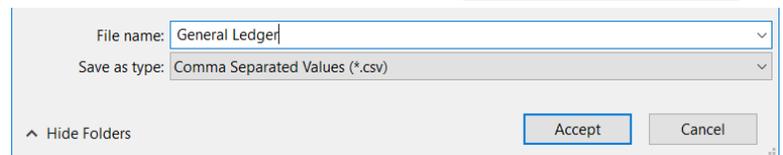


Select "Browse" button

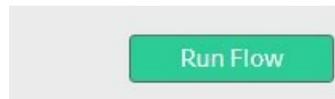


Select the location for the file and specify the file name, General Ledger".

Select "Accept"



Select "Run Flow"



Select "Done".

#### Finished Running Flow



General Ledger.csv

Total time 00:00

Done

### Step 7: Create General Ledger from the exported file and add a Running Balance field

Open the “General Ledger.csv” file that was exported on the previous step.

Text Import Wizard - Step 1 of 3

The Text Wizard has determined that your data is Delimited.  
If this is correct, choose Next, or choose the data type that best describes your data.

Original data type

Choose the file type that best describes your data:

- Delimited - Characters such as commas or tabs separate each field.
- Fixed width - Fields are aligned in columns with spaces between each field.

Start import at row: 1 File origin: 65001 : Unicode (UTF-8)

My data has headers.

Preview of file C:\Users\mfriedma\Desktop\Sample\General Ledger.csv.

```

1 Account,Name,Transaction,Date,Description,Debit,Credit,Total Balance
2 1110,Cash,1,6/1/2018,Investment from Mark Friedman,28000,0,28000
3 3100,Capital Stock,1,6/1/2018,Investment from Mark Friedman,0,28000,-28000
4 1211,Office Equip.,2,6/1/2018,Hudson equipment Invoice BC3887,4000,0,4000
5 2101,Accounts Payable,2,6/1/2018,Hudson equipment Invoice BC3887,0,4000,-4000
6 1311,Computer Equip.,3,6/1/2018,Hailey Computers 87245, ck6001",7000,0,7000
    
```

Buttons: Cancel, < Back, Next >, Finish

Annotations:

- Select “Delimited.”
- Select “My data has headers”.
- Select “Next”.

Text Import Wizard - Step 2 of 3

This screen lets you set the delimiters your data contains. You can see how your text is affected in the preview below.

Delimiters

- Tab
- Semicolon
- Comma
- Space
- Other:

Treat consecutive delimiters as one

Text qualifier: "

Data preview

Account	Name	Transaction	Date	Description	Debit	Credit
1110	Cash	1	6/1/2018	Investment from Mark Friedman	28000	0
3100	Capital Stock	1	6/1/2018	Investment from Mark Friedman	0	28000
1211	Office Equip.	2	6/1/2018	Hudson equipment Invoice BC3887	4000	0
2101	Accounts Payable	2	6/1/2018	Hudson equipment Invoice BC3887	0	4000
1311	Computer Equip.	3	6/1/2018	Hailey Computers 87245, ck6001"	7000	0

Buttons: Cancel, < Back, Next >, Finish

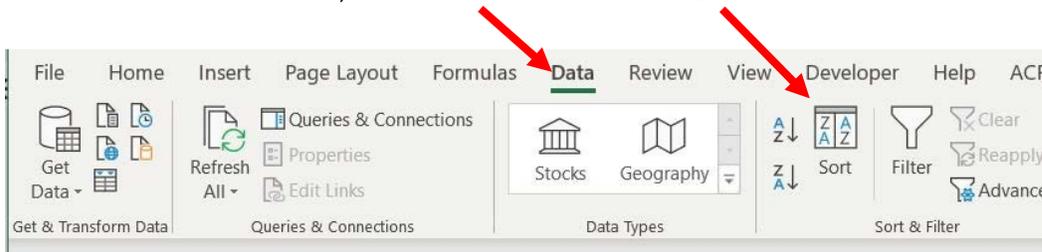
Annotations:

- Select “,”.
- Select “Finish”.

	A	B	C	D	E	F	G	H	I
1	Account	Name	Transactio	Date	Description	Debit	Credit	Total Balance	
2	1110	Cash	1	01-Jun-18	Investmen	28000	0	28000	
3	3100	Capital Stc	1	01-Jun-18	Investmen	0	28000	-28000	
4	1211	Office Equ	2	01-Jun-18	Hudson eq	4000	0	4000	
5	2101	Accounts F	2	01-Jun-18	Hudson eq	0	4000	-4000	
6	1311	Computer	3	01-Jun-18	Hailey Con	7000	0	7000	
7	1110	Cash	3	01-Jun-18	Hailey Con	0	7000	-7000	
8	5040	Repairs & I	4	02-Jun-18	Avery Repa	725	0	725	
9	2101	Accounts F	4	02-Jun-18	Avery Repa	0	725	-725	
0	1211	Office Equ	5	04-Jun-18	Scanner, Ja	125	0	125	
1	1110	Cash	5	04-Jun-18	Scanner, Ja	0	125	-125	
2	5030	Advertising	6	08-Jun-18	Zac Advert	3380	0	3380	
3	2101	Accounts F	6	08-Jun-18	Zac Advert	0	3380	-3380	
4	2101	Accounts F	7	08-Jun-18	Pd A/p ck6	640	0	640	
5	1110	Cash	7	08-Jun-18	Pd A/p ck6	0	640	-640	
6	1110	Cash	8	14-Jun-18	Services pe	11000	0	11000	
7	4100	Computer	8	14-Jun-18	Services pe	0	11000	-11000	
8									

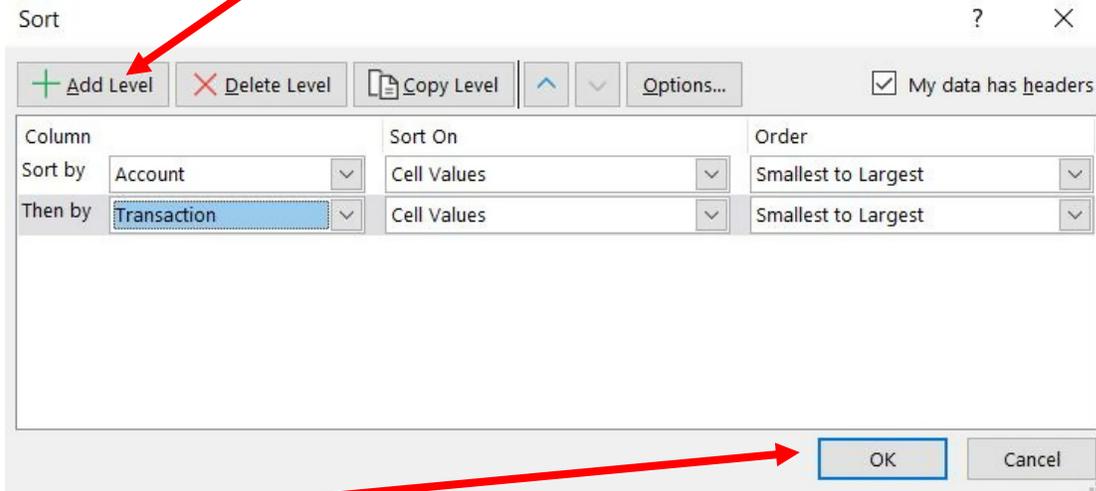
Data looks good now, but we need to sort by **“Account”** and **“Transaction”** in **“Ascending”** order.

With the cursor in cell A1, select **“Data”** and then **“Sort”**



In the sorting dialog, first select the “Account” field and select the order “Smallest to Largest”. Then click “Add Level” button on top and do the same for “Transaction”.

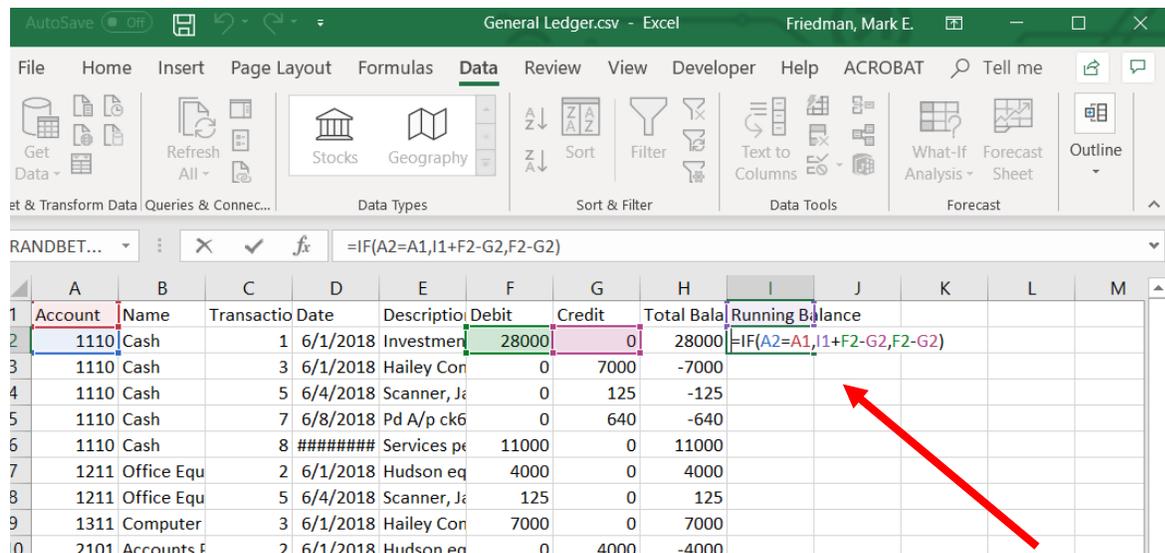
Press OK.



Press OK.

A General Ledger has a running balance within each account. To make Running Balance calculation, type column header “Running Balance” into the first row of the first empty column (cell I1).

In the cell under header (I2) put in this formula: =IF(A2=A1,I1+F2-G2,F2-G2)



Copy the formula from cell I2 to the bottom of the data.

Copy the data on the sheet and “Paste Values” into a new sheet in the initial Excel file. Rename the sheet “General Ledger”.

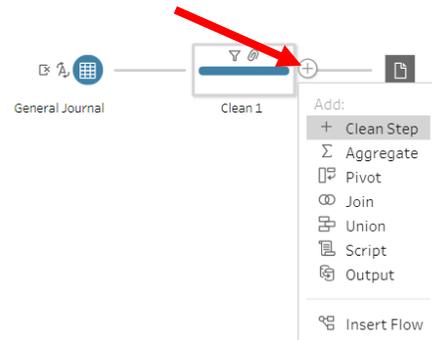
Delete the “Total Balance” column – it is not needed anymore.

**Step 8: Create a new field that combine the fields “Account” and “Name”**

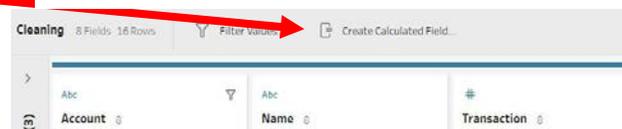
The “Unadjusted Trial Balance” will summarize the “Total Balance” by account number with the account name attached for clarity.

We now want to combine the fields “Account” and “Name” and we would like to have two spaces between the fields.

Click on the plus sign on the right of the “Clean 1” step and select “Clean Step”.



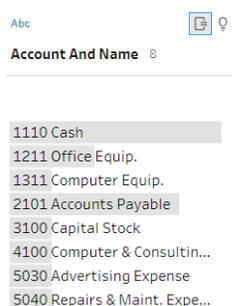
Select “Create Calculated Field”.



Enter “Account and Name” as the “Field Name”.

Enter the formula: [Account]+' "[Name]

Select “Save”.

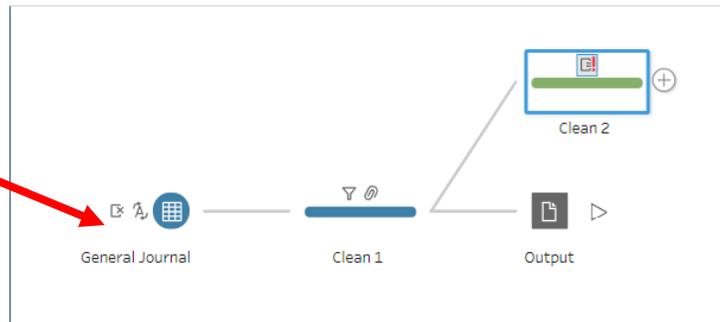


If you receive an error message the field “Account” is still coded as a numeric.

Function '+' can't accept parameters: (number, string).

Follow these steps if you must change the field “Account” from a numeric to a string.

Select the first step, “General Journal”.



Select the “Type” field for “Account”

**General Journal** Fields selected: 8 of 9 Filter Values...

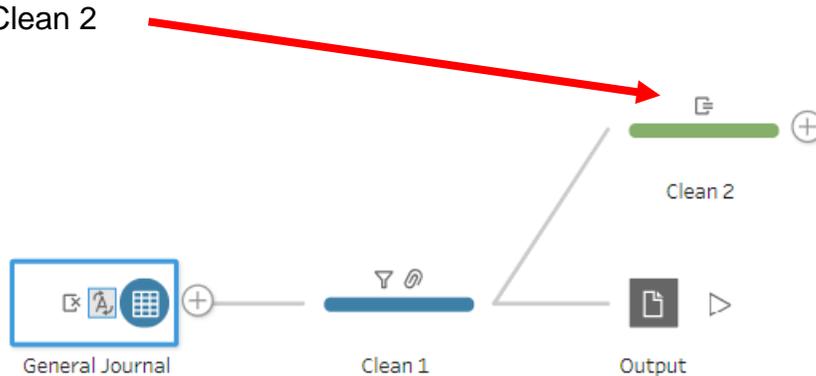
Select the fields to include in your flow, apply a filter, or change data type

<input type="checkbox"/>	Type	Field Name	Original Field Name
<input checked="" type="checkbox"/>	#	Transaction	Transaction
<input checked="" type="checkbox"/>	📅	Date	Date
<input checked="" type="checkbox"/>	#	Account	Account

Select “String”

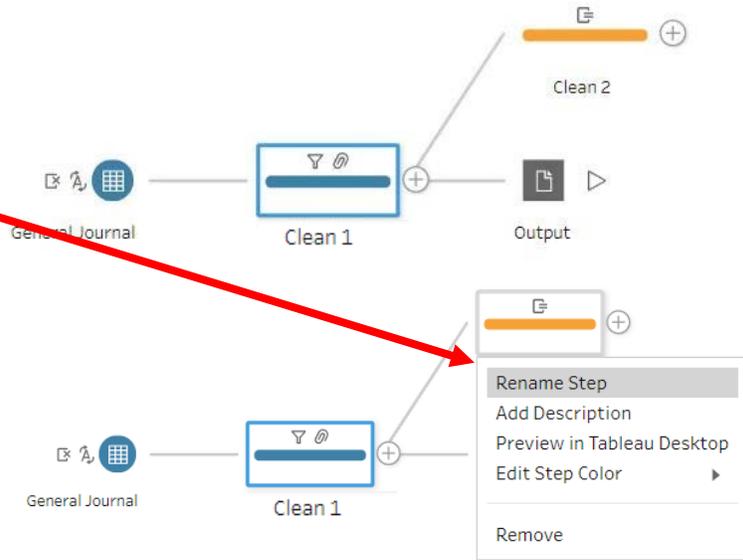
- # Number (decimal)
- # Number (whole) - default
- 📅 Date & Time
- 📅 Date
- Abc String

Return to the step Clean 2

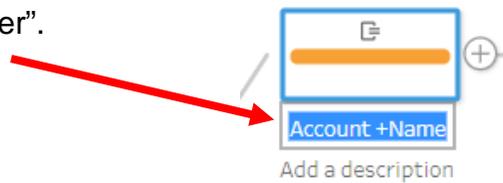


**Step 9: Aggregate balances and export “Unadjusted Trial Balance”**

Right Click on the “Clean 2” in the “Flow Pane” and select “Rename Step”.

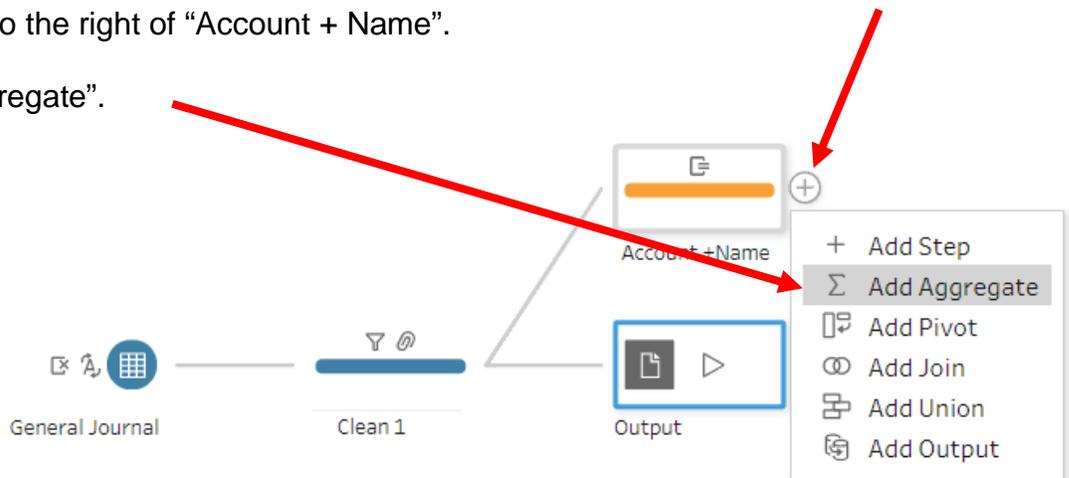


Type “Account + Name” and press “Enter”.

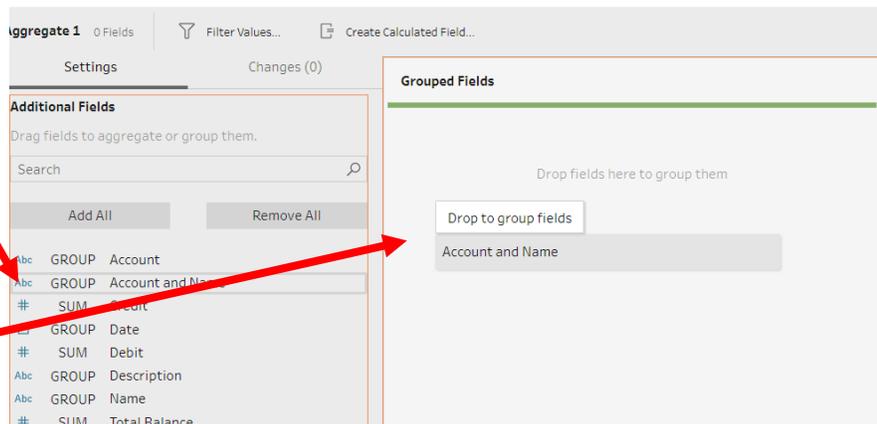


Click on the “+” to the right of “Account + Name”.

Select “Add Aggregate”.

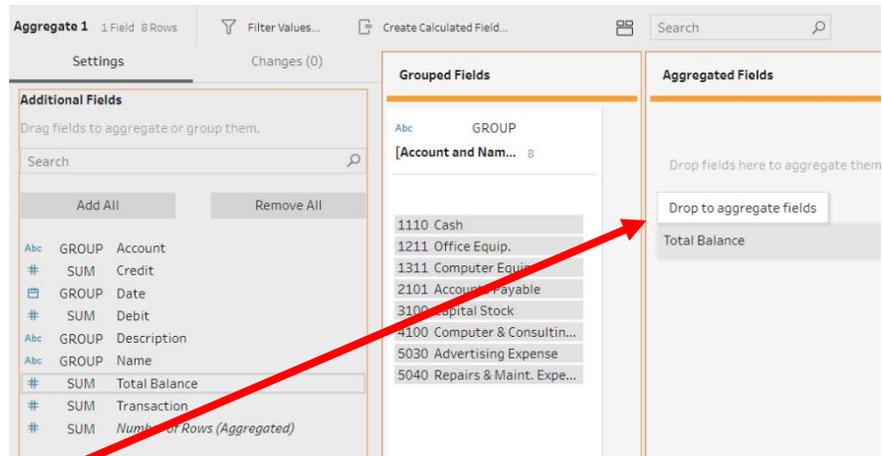


Drag "Account and Name" field



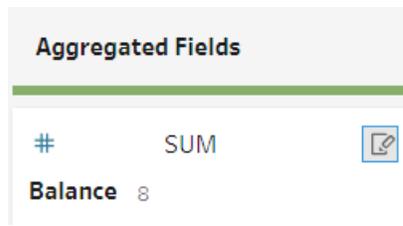
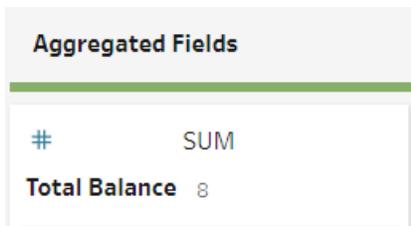
to "Grouped Fields"

Drag "Total Balance" field



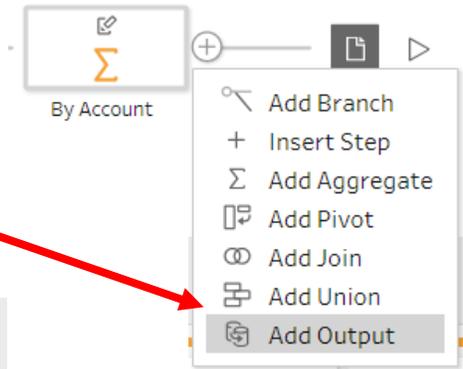
to "Aggregated Fields".

Double click on "Total Balance" and rename the field "Balance"

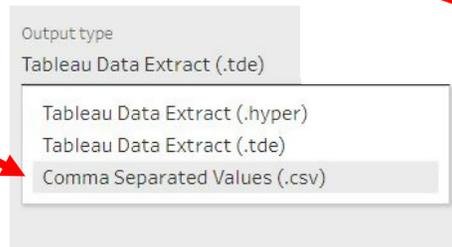


### Step 10: Output the aggregated balances to the Unadjusted Trial Balance

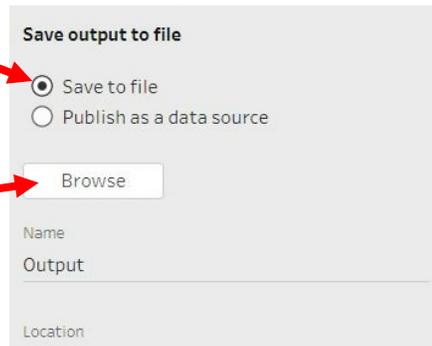
Click on the plus sign of our Aggregate tile and select "Add Output".



Select the .csv file type.

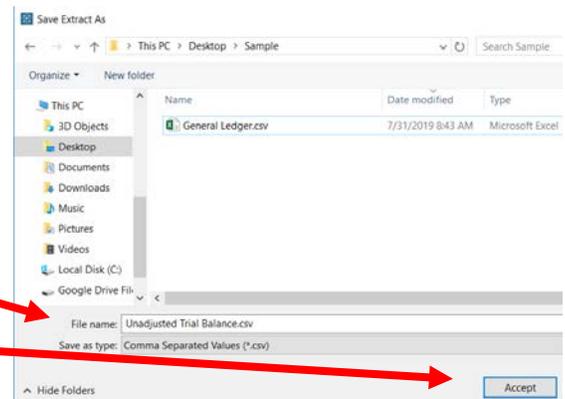


Select "Save to file" option button.



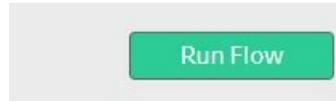
Select "Browse" button

Select the location for the file and specify the "File name".



Select "Accept".

Select "Run Flow"



Select "Done".

Finished Running Flow

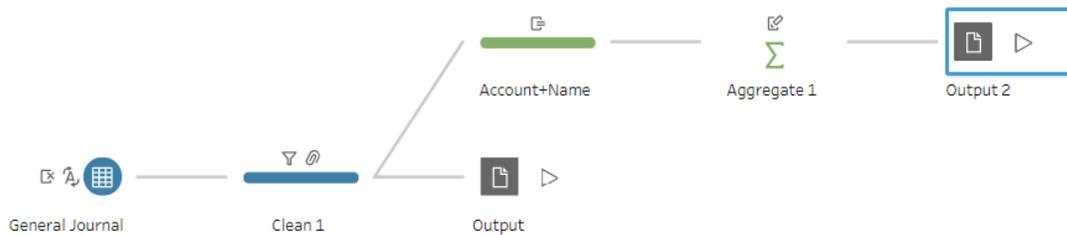


General Ledger.csv

Total time 00:00

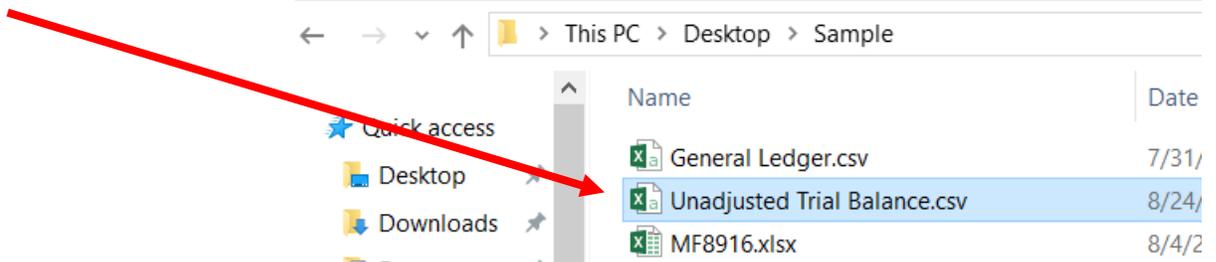
Done

Final flow in Tableau Prep should look like this.



### Step 11: Open the file in Excel

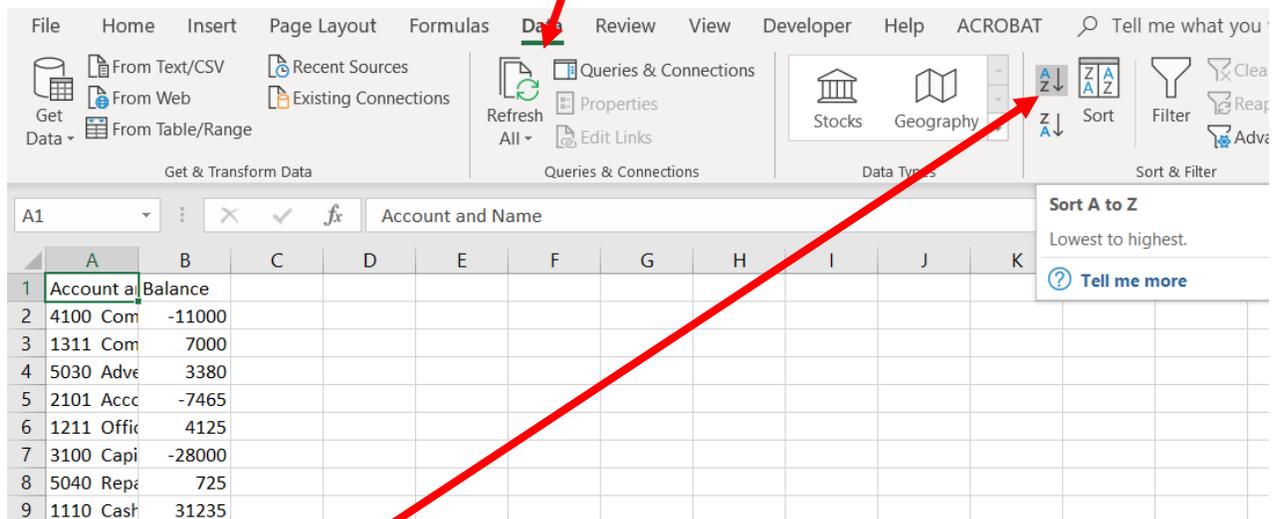
Double click on the file “Unadjusted Trial Balance.csv” file in “File Explorer”.



And the file opens in Excel.

	A	B	C
1	Account and	Balance	
2	4100 Com	-11000	
3	1311 Com	7000	
4	5030 Adve	3380	
5	2101 Acco	-7465	
6	1211 Offic	4125	
7	3100 Capi	-28000	
8	5040 Repa	725	
9	1110 Cash	31235	
10			

With the cursor in cell “A1” select “Data”.



Select “A to Z”.  
Change the column widths.

**Step 12: Transfer the data to the original MF8916.xlsx file**

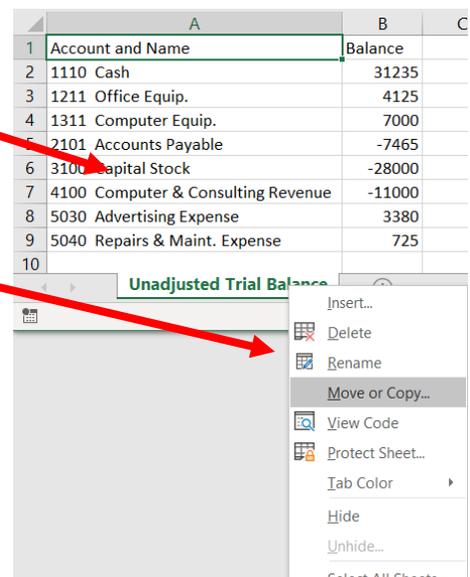
## 1) Copy and Paste

- a) Open the initial excel file “MF8916.xls”.
- b) If a sheet “Unadjusted Trial Balance does not exist, name a new sheet “Unadjusted Trial Balance”.
- c) Return to the “Unadjusted Trial Balance” file.
- d) Highlight all the data in the “Unadjusted Trial Balance”.
- e) Select “Copy”.
- f) Return to the file “MF8916.xls”.
- g) On the “Unadjusted Trial Balance Sheet” select “Paste Values”.
- h) **Save the files.**

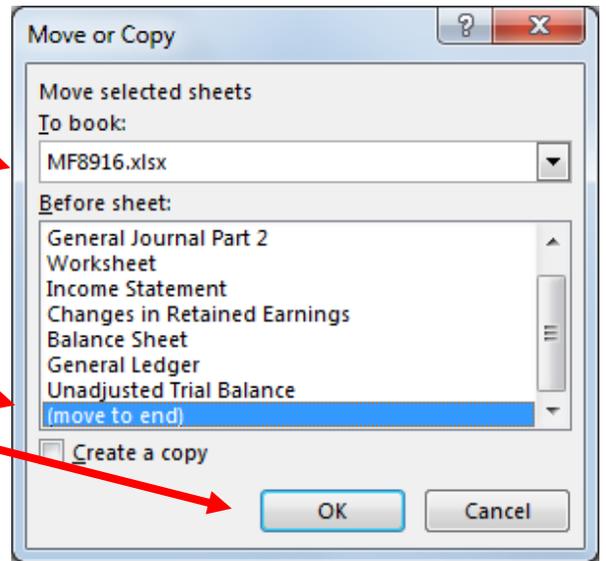
## 2) Move the sheet into the “MF8916.xls” file.

- a) Right click on the “Unadjusted Trial Balance” tab.

- b) Select “Move or Copy...”.



c) Select the name of the original file.



d) Scroll till the end and click on “(move to end)”.

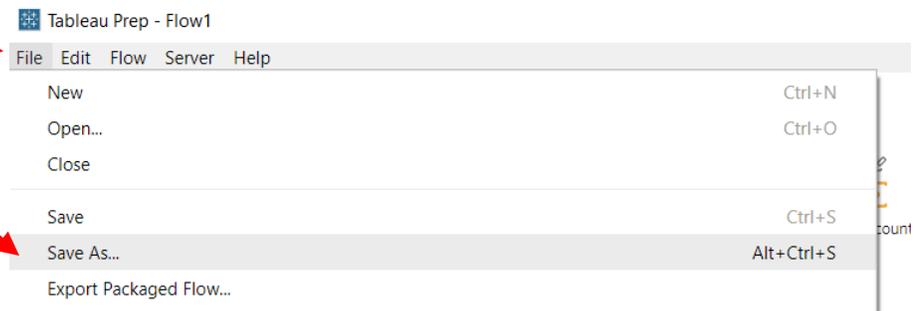
e) Then click “OK”.

f) The Unadjusted Trial Balance has been created successfully.

g) Save” and “Close” the Excel file.

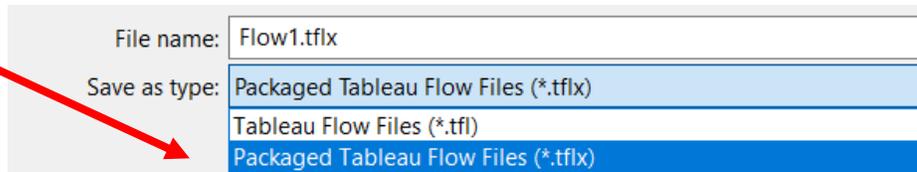
**Step 13: Share the workflow and the related input files.**

Select “File”.



Select “Save As”.

Change the type to “Packaged Tableau Flow Files (\*.tflx)”.



Select “Save”.



Save the file as a normal Tableau Prep file,” \*.tfl”.