Using Tableau Prep Builder

Tableau Prep is used by several Big Four accounting firms to automate the ETL process. In this exercise, you will **E**xtract the General Journal file, **T**ransform the data into the required format (a General Ledger and an Unadjusted Trial Balance), and subsequently **L**oad 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

End with the General Ledger and   the   Unadjusted Trial Balance

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.

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.
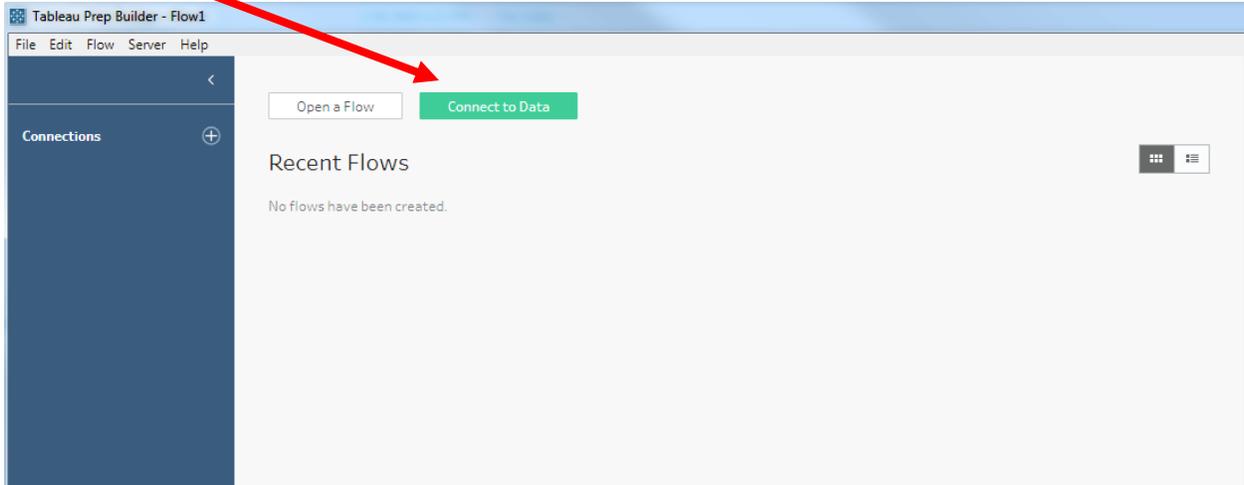
On the connections pane select Microsoft Excel.
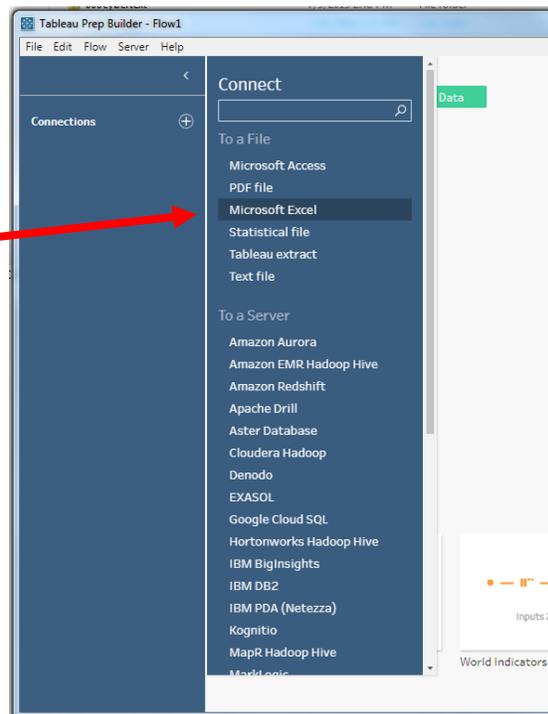
Select the file.

Select "Open".
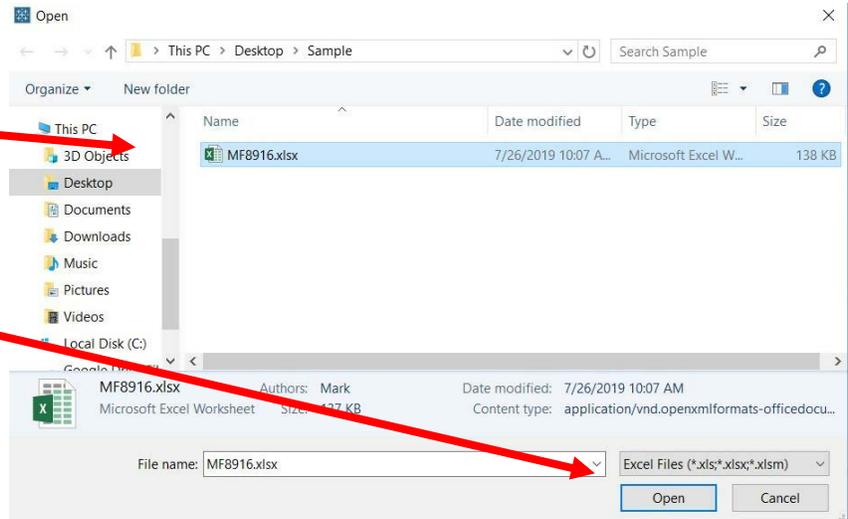
After you open the file, select the "General Journal" sheet on the Connections pane. Either double click on the sheet name or drag the sheet name to the canvas.
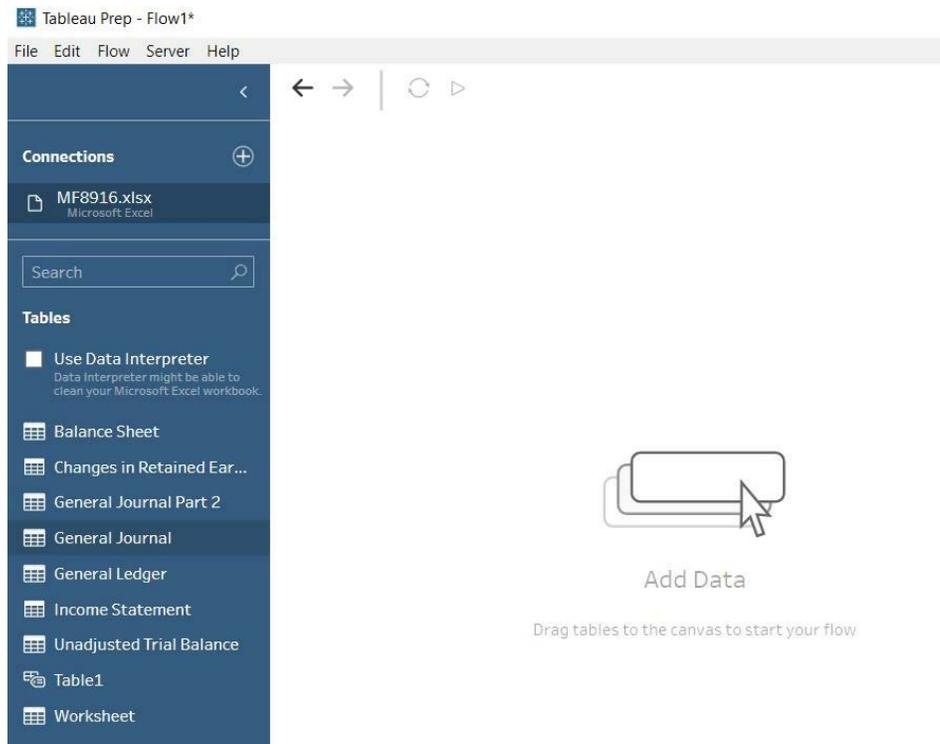
After selecting the sheet, "General Journal" will appear on the Flow pane.
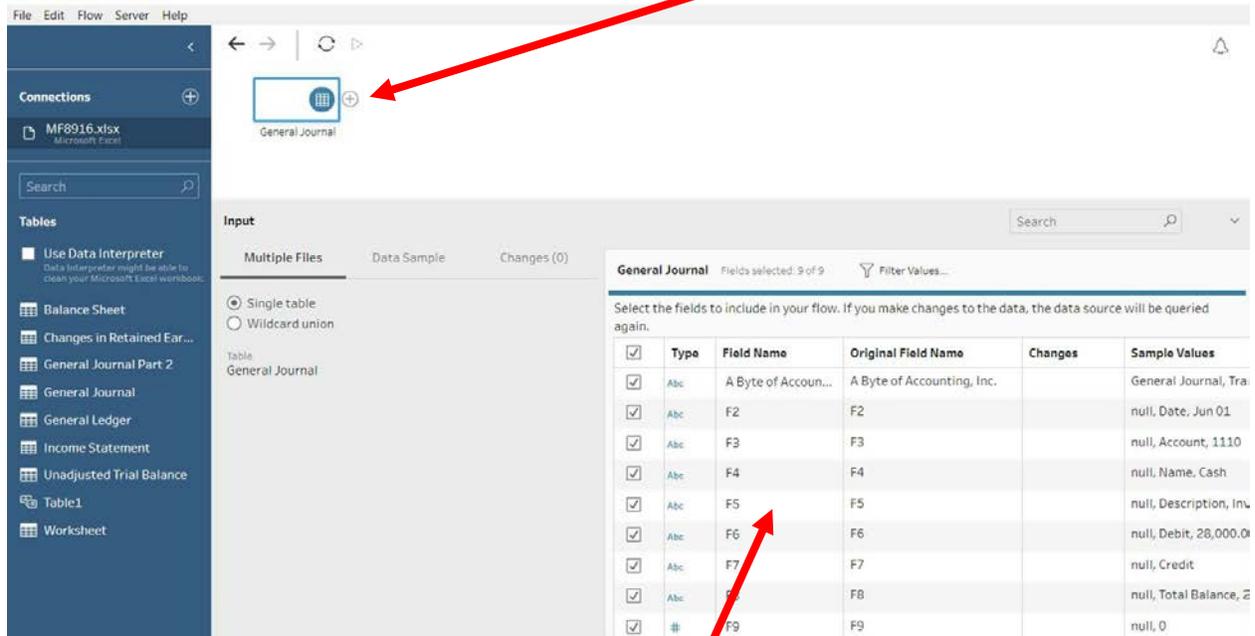
Sample data will be visible on the Profile pane.

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".

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

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

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)"

#	Account
#	Number (decimal)
✓ #	Number (whole) - default
🕒	Date & Time
📅	Date
Abc	String
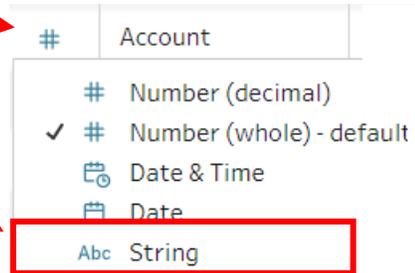
General Journal   Fields selected: 9 of 9   🔽 Filter Values...

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

☑	Type	Field Name	Original Field Name	Changes
☑	#	Transaction	Transaction	
☑	📅	Date	Date	
☑	Abc	Account	Account	ᴀ
☑	Abc	Name	Name	
☑	Abc	Description	Description	
☑	#	Debit	Debit	ᴀ
☑	#	Credit	Credit	ᴀ
☑	#	Total Balance	Total Balance	ᴀ
☑	#	F9	F9	

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

☑	#	Debit	F6		📝 ᴬ
☑	#	Credit	F7		📝 ᴬ
☑	#	Total Balance	F8		📝 ᴬ
☐	#	F9	F9		▣

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

General Journal

Add:
+ Clean Step
Σ Aggregate
▢ Pivot
⊘ Join
⊟ Union
▤ Script
⬡ Output

⬡ Insert Flow

Select "+ Clean Step"

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".
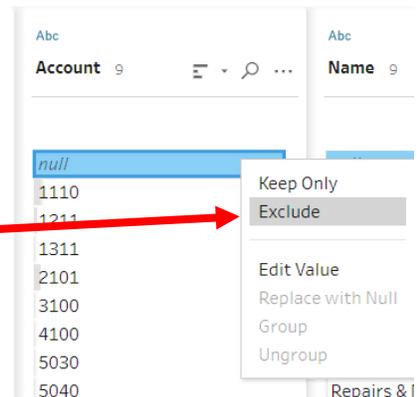
Step 4: Change all occurrences of "null" to "0".

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 8	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.

# Debit 9	# Credit 9
0	*null*
125	125
640	640
725	725
3,380	3,380
4,000	4,000
7,000	7,000
11,000	11,000
28,000	28,000

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	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

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

Account	Transaction Name	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"
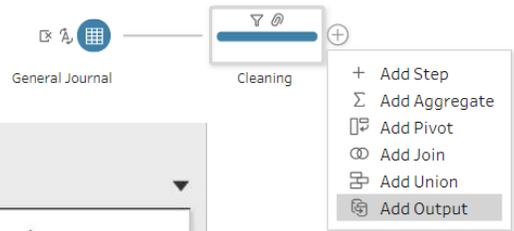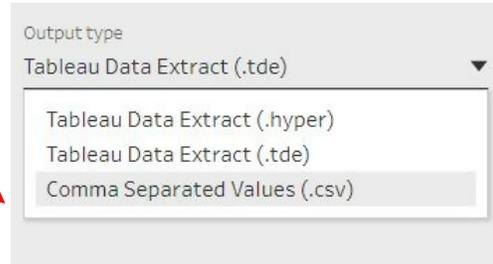
Select the .csv file type.

Select "Save to file" option button.
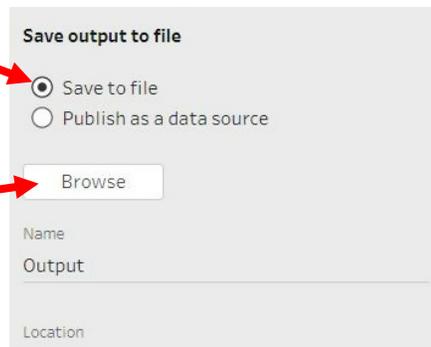
Select "Browse" button
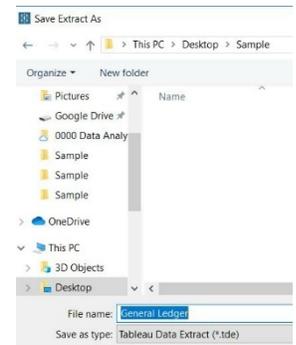
Select the location for the file and specify the file name, General
Ledger".

Select "Accept"
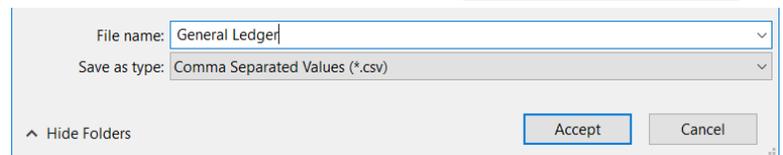
Select "Run Flow"
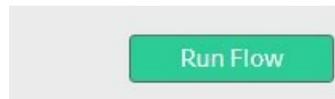
Select "Done".

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.

Select "Delimited."

Select "My data has headers".

Select "Next".

Select ",".

Select "Finish".

|	A	B	C	D	E	F	G	H	I
1	Account	Name	Transactio	Date	Descriptio	Debit	Credit	Total Balance	
2	1110	Cash	1	01-Jun-18	Investmen	28000	0	28000	
3	3100	Capital Sto	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 &	4	02-Jun-18	Avery Repa	725	0	725	
9	2101	Accounts F	4	02-Jun-18	Avery Repa	0	725	-725	
10	1211	Office Equ	5	04-Jun-18	Scanner, Ja	125	0	125	
11	1110	Cash	5	04-Jun-18	Scanner, Ja	0	125	-125	
12	5030	Advertising	6	08-Jun-18	Zac Advert	3380	0	3380	
13	2101	Accounts F	6	08-Jun-18	Zac Advert	0	3380	-3380	
14	2101	Accounts F	7	08-Jun-18	Pd A/p ck6	640	0	640	
15	1110	Cash	7	08-Jun-18	Pd A/p ck6	0	640	-640	
16	1110	Cash	8	14-Jun-18	Services pe	11000	0	11000	
17	4100	Computer	8	14-Jun-18	Services pe	0	11000	-11000	
18									

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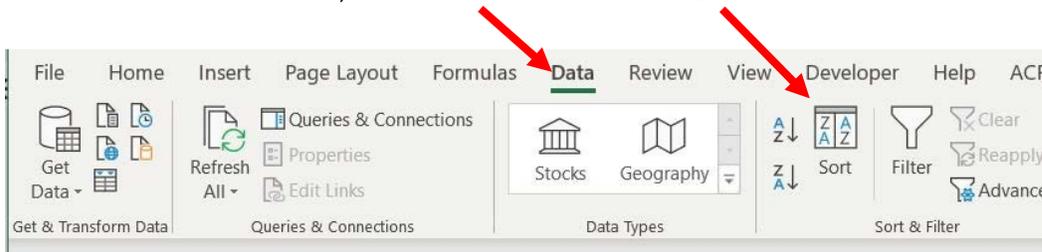
In the sorting dialog, first select the "Account" field and select the order "Smallest to Largest". Then click "Add Level" button on top and do the same for "Transaction".

Press OK.
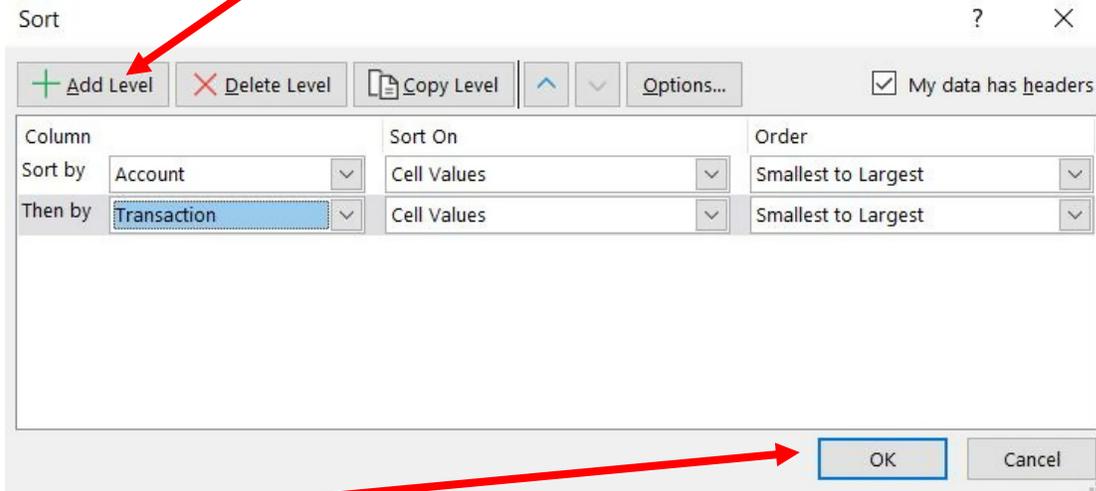
Press OK.

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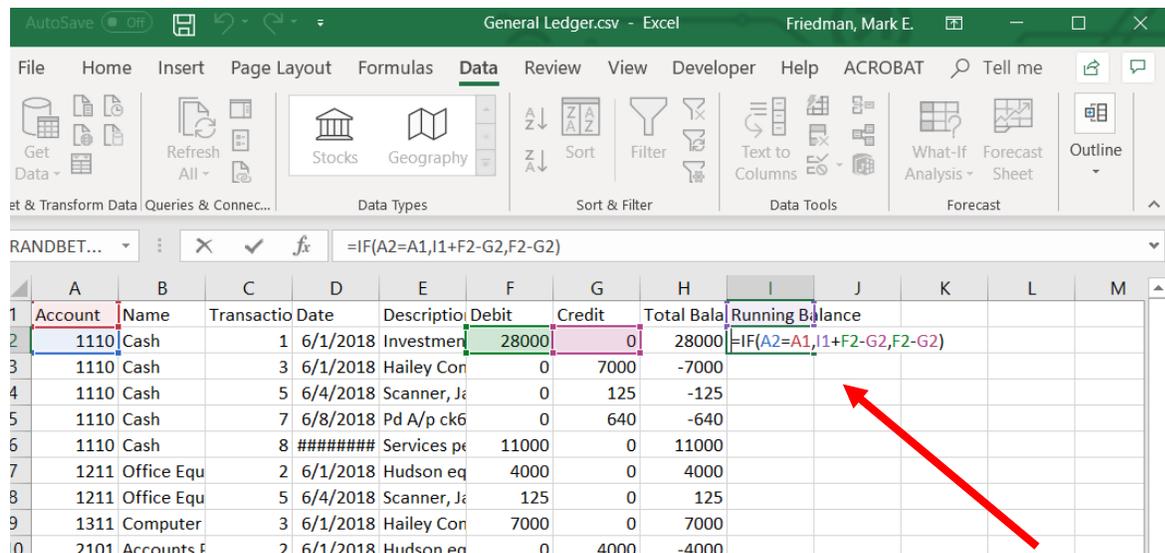
Copy the formula from cell I2 to the bottom of the data.

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".
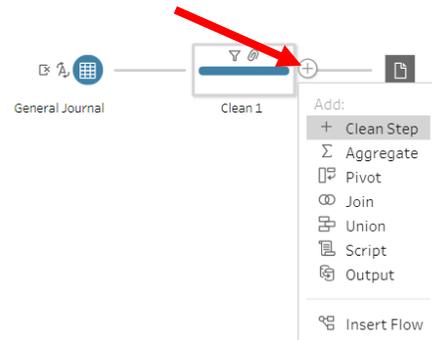
Select "Create Calculated Field".
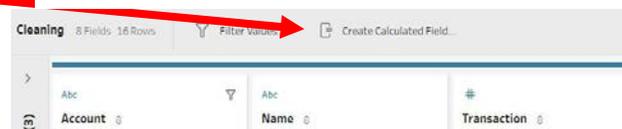
Enter "Account and Name" as the "Field Name".

Field Name

Account and Name

[Account]+" "+[Name]

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

Select "Save".

Account And Name   8

1110 Cash
1211 Office Equip.
1311 Computer Equip.
2101 Accounts Payable
3100 Capital Stock
4100 Computer & Consultin...
5030 Advertising Expense
5040 Repairs & Maint. Expe...
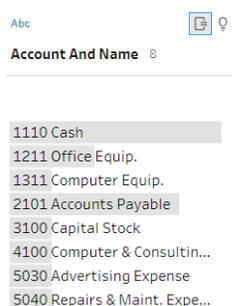
If you receive an error message the field "Account" is still coded as a numeric.

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"

Select "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".

Clean 2

Clean 1

General Journal   Output

Rename Step
Add Description
Preview in Tableau Desktop
Edit Step Color   ▶

Remove

General Journal   Clean 1
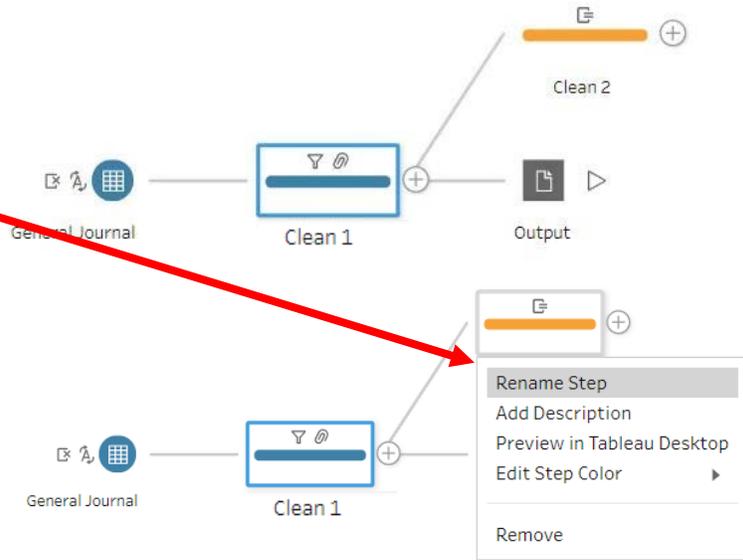
Type "Account + Name" and press "Enter".

Account +Name

Add a description
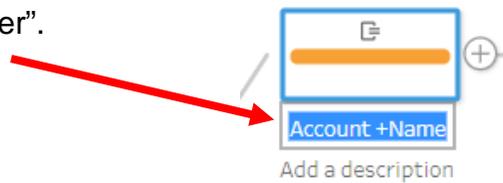
Click on the "+" to the right of "Account + Name".

Select "Add Aggregate".

Account +Name

+   Add Step
Σ   Add Aggregate
▯⋤   Add Pivot
◯◯   Add Join
⊞   Add Union
⊡   Add Output

General Journal   Clean 1   Output
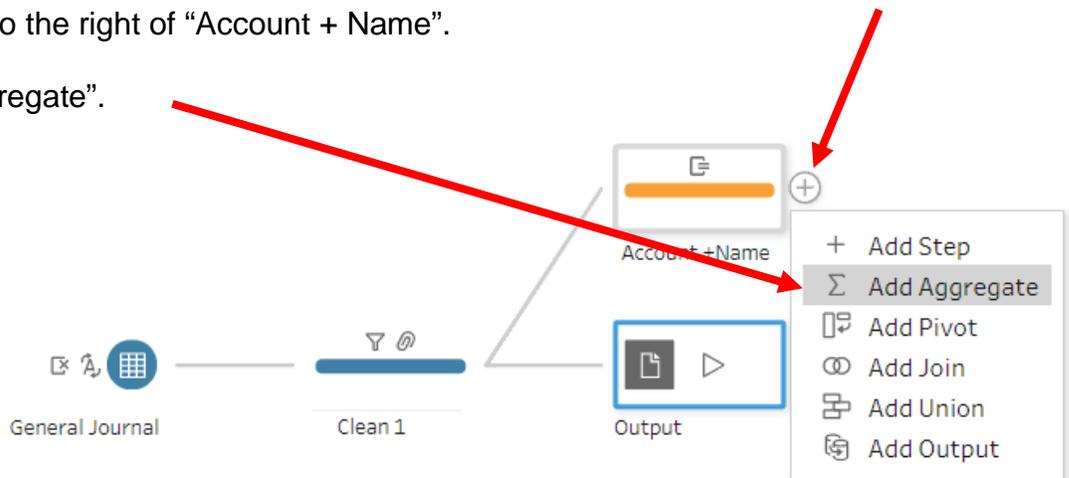
Drag "Account and Name" field
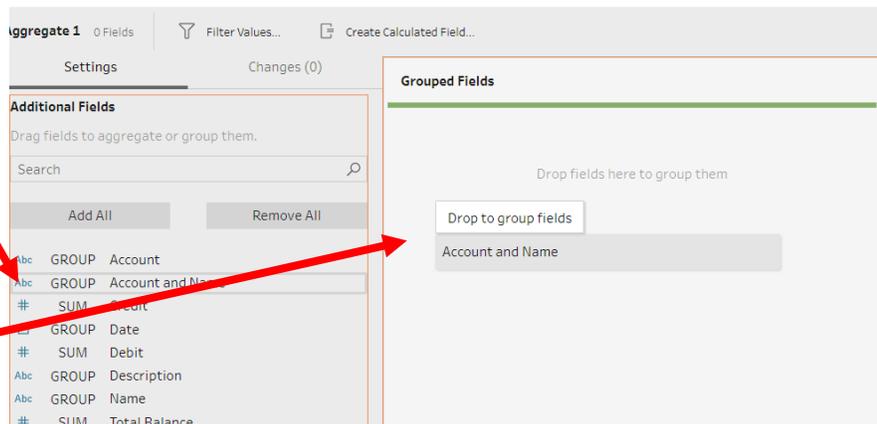
to "Grouped Fields"

Drag "Total Balance" field
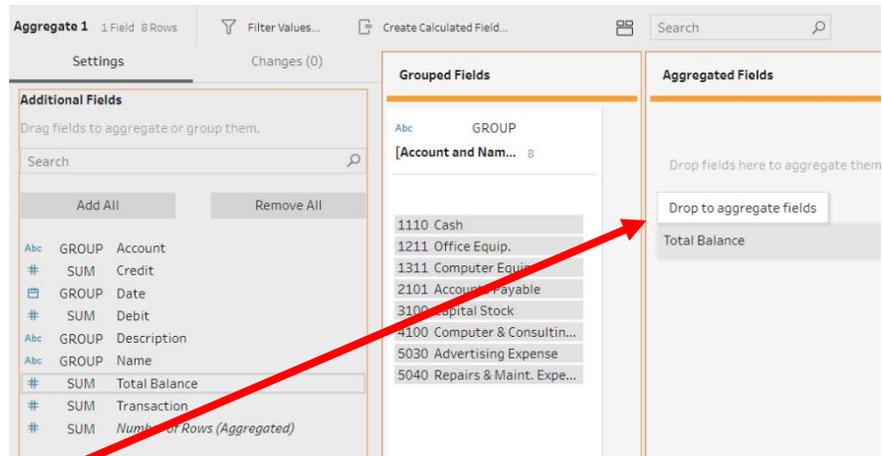
to "Aggregated Fields".

Double click on "Total Balance" and rename the filed "Balance"
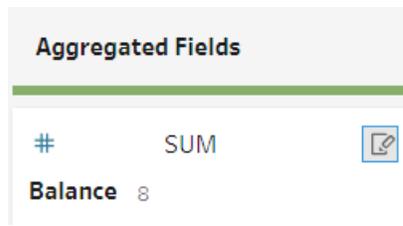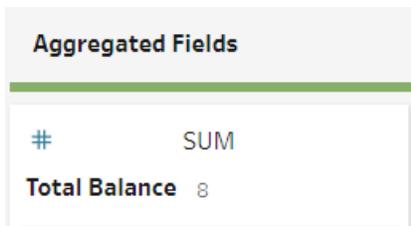
**Step 10: Output the aggregated balances to the Unadjusted Trial Balance**

Click on the plus sign of our Aggregate tile and select "Add Output".
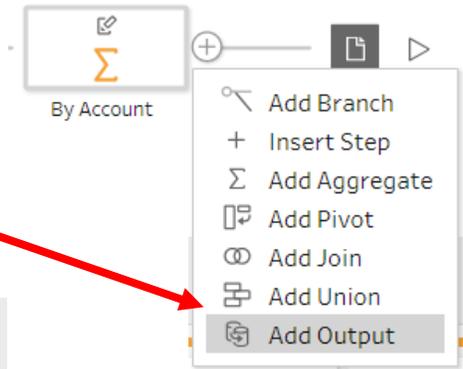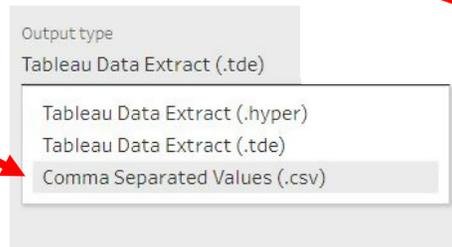
Select the .csv file type.

Select "Save to file" option button.
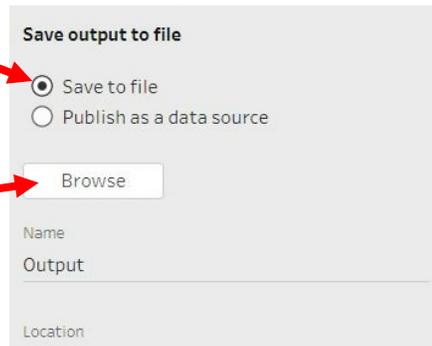
Select "Browse" button

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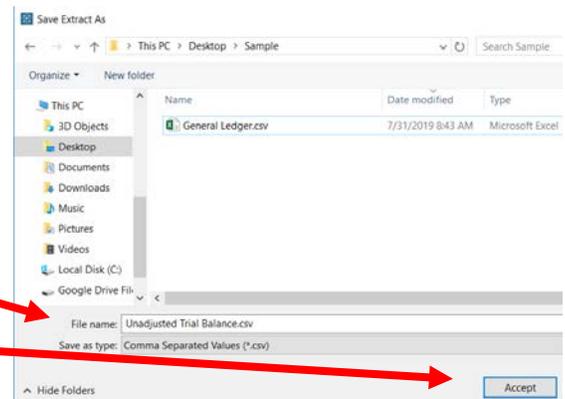
Select "Accept".

Select "Run Flow"
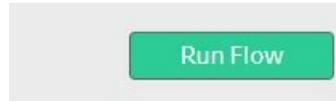
Select "Done".

Finished Running Flow

General Ledger.csv

Total time 00:00

Done

Final flow in Tableau Prep should look like this.
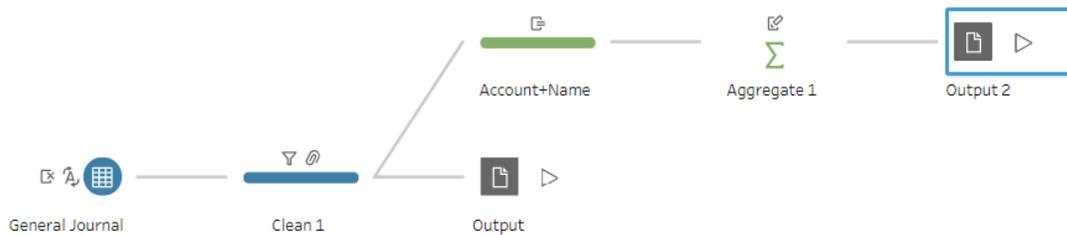
Step 11: Open the file in Excel

Double click on the file "Unadjusted Trial Balance.csv" file in "File Explorer".
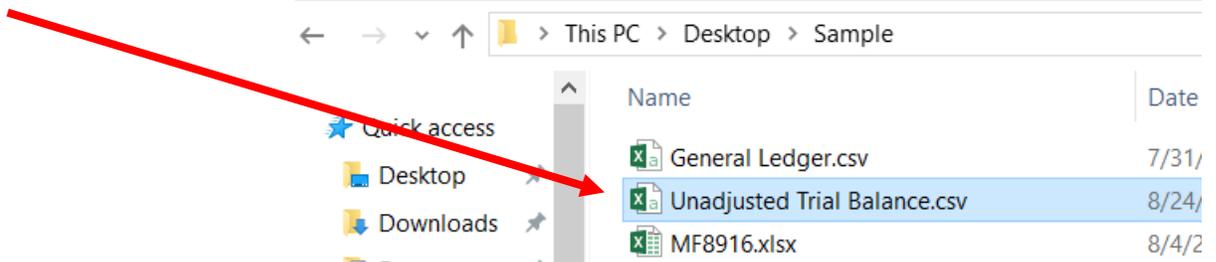
And the file opens in Excel.

With the cursor in cell "A1" select "Data".
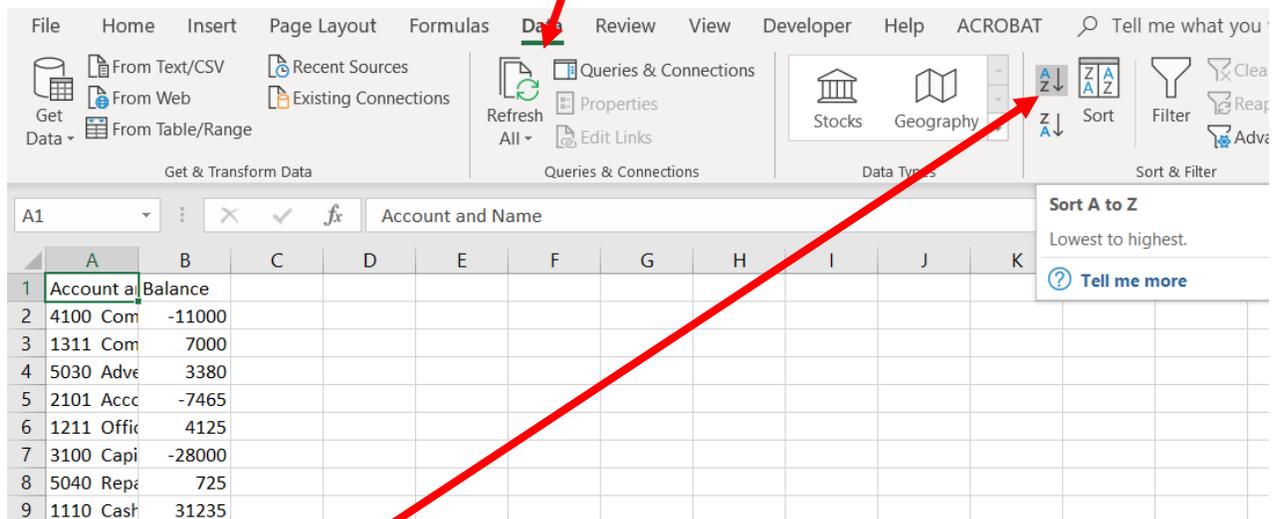
Select "A to Z".
Change the collumn widths.

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…".
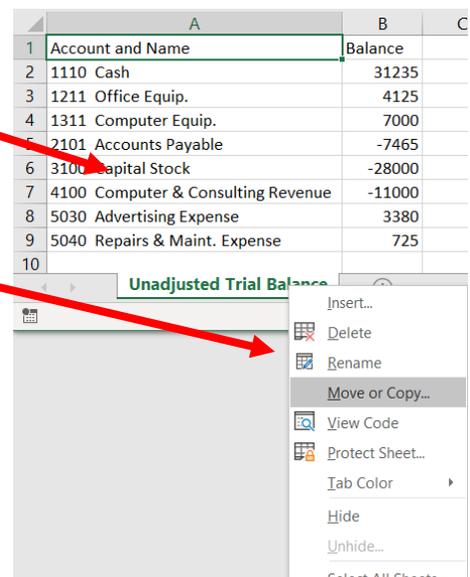
c) Select the name of the original file.
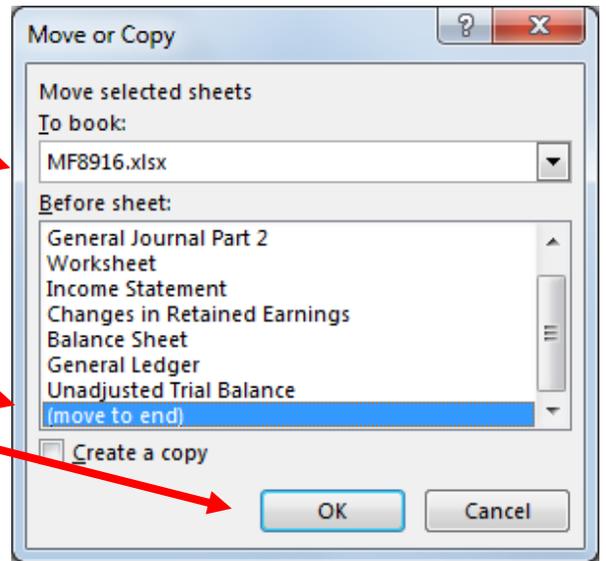
d) Scroll till the end and click on "(move to end)".

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".
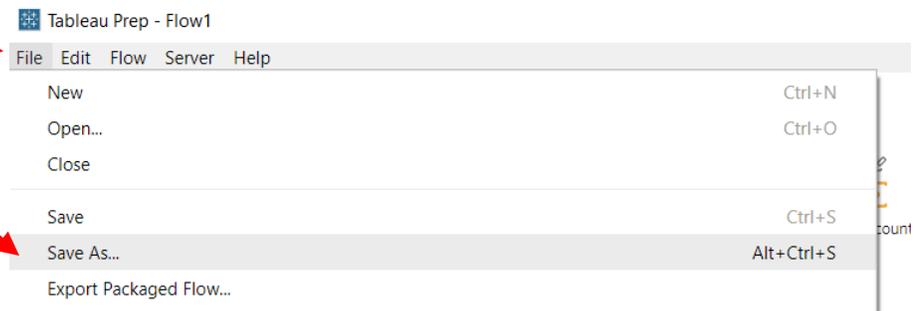
Select "Save As".

Change the type to "Packaged Tableau Flow Files (*.tflx)".
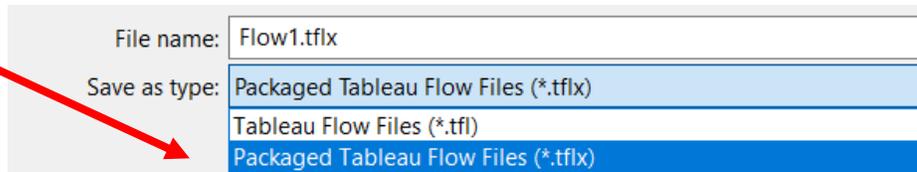
Select "Save".

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