



School of Business

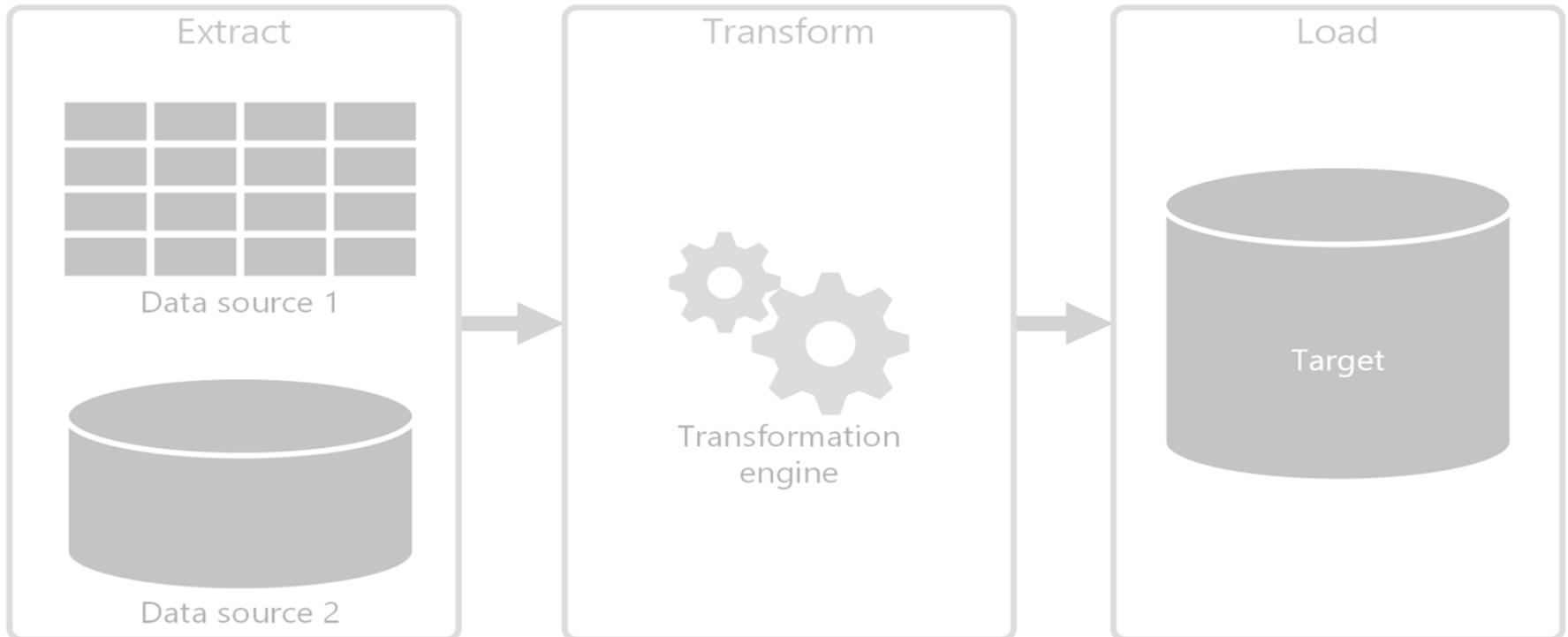
weARE Webinar: January 7, 2022

Data Automation

Karen Kitching
703-993-9038
kkitchin@gmu.edu

- A process used to collect data from various sources, transform the data depending on business rules, and load the data into a destination database.

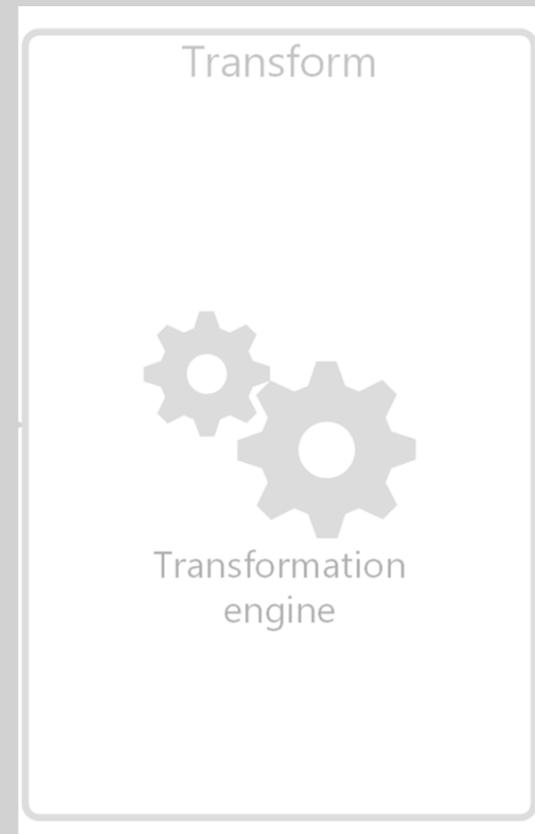
Source: <http://www.sql-tutorial.net/ETL.asp>



Transform

The data transformation may include:

- **filtering**
- **sorting**
- **aggregating**
- **joining data**
- **cleaning data**
- **changing data characteristics**
- **generating calculated data based on existing values**
- **validating data**
- **etc.**



ETL: Extract, transform, and load

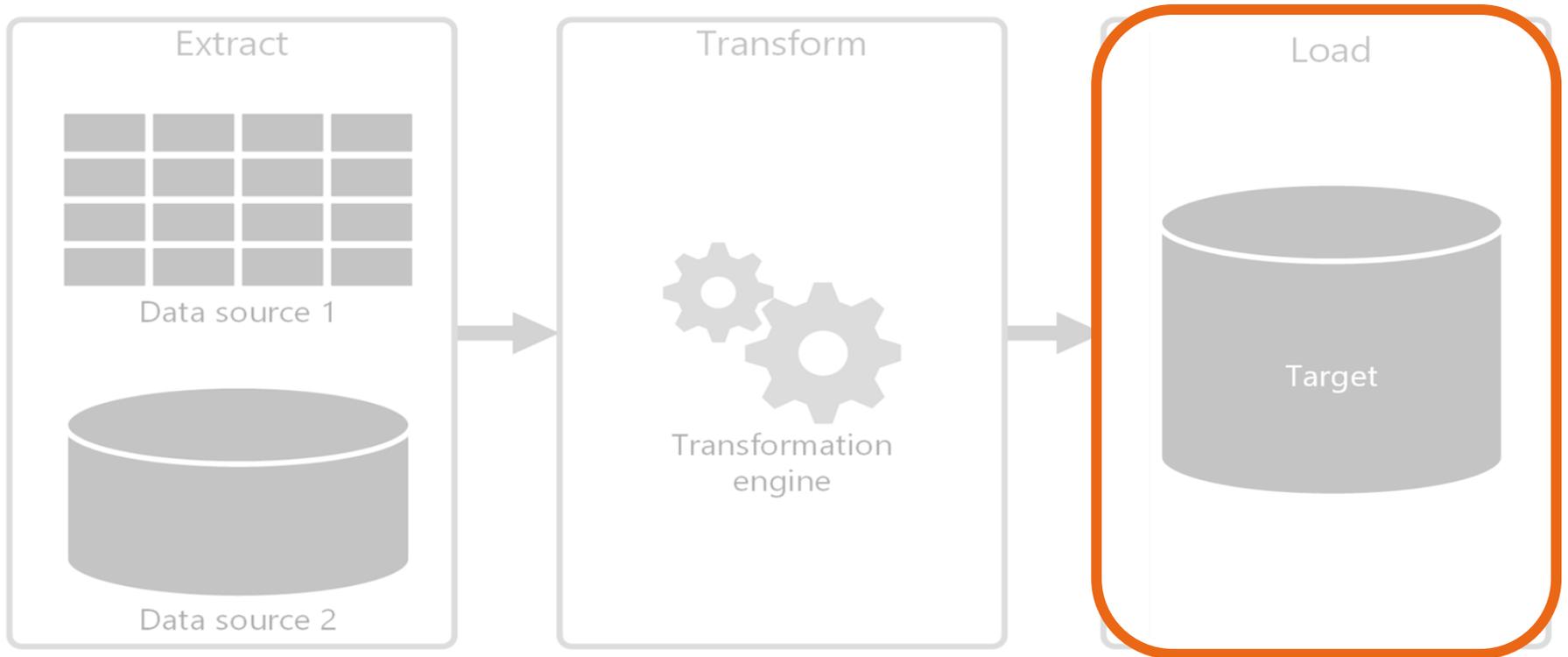


Image source: <https://docs.microsoft.com/en-us/azure/architecture/data-guide/relational-data/etl>

Automates data transformation, i.e., the ETL process



Why Alteryx?

Users can manipulate data using repeatable drag and drop workflows without having advanced programming skills (such as SQL, Python, R)

How automation works



Inputs

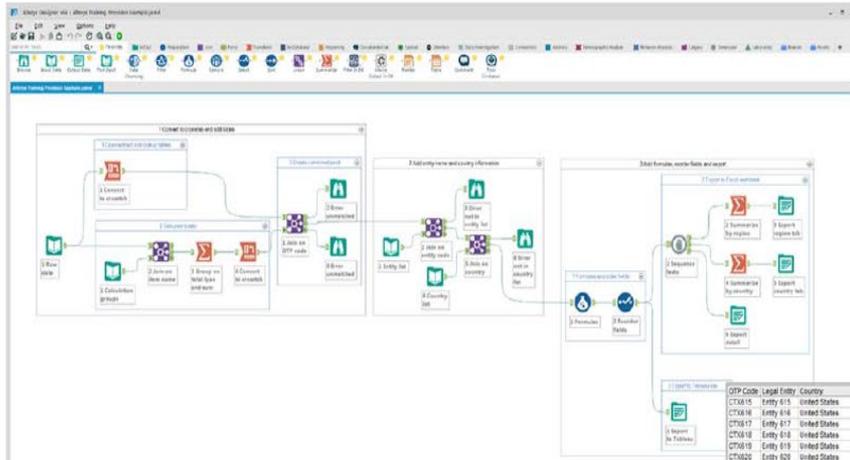
OTF Code	Legal Entity	Country	Total Pre-Tax Book Income
CT000	Entity 001	Australia	15,291,414.89
CT003	Entity 004	Australia	5,036,894.60
CT003	Entity 006	Australia	4,906,365.60
CT003	Entity 006	Australia	36,528.00
CT003	Entity 007	Australia	22,827,554.60
CT003	Entity 008	Australia	124,995.00
CT004	Entity 209	Australia	(52,187.62)
CT004	Entity 019	Australia	20,346.00
CT004	Entity 011	Australia	532,170.60
CT004	Entity 014	Australia	97,808.00
CT004	Entity 014	Australia	(4,290,862.00)
CT004	Entity 014	Australia	(74,246.00)
CT005	Entity 015	Australia	86,926,156.19
CT005	Entity 016	Australia	(76,124,945.00)
CT005	Entity 019	Bangladesh	67,915,266.00
CT005	Entity 017	Australia	3,245,226.00
CT005	Entity 019	Bangladesh	4,137,862.00
CT005	Entity 020	Bangladesh	102,844.00
CT007	Entity 001	Australia	8,203,246.84

OTF Code	Legal Entity	Country
CT000	Entity 001	Australia
CT003	Entity 004	Australia
CT003	Entity 006	Australia
CT003	Entity 006	Australia
CT003	Entity 007	Australia
CT003	Entity 008	Australia
CT004	Entity 209	Australia
CT004	Entity 019	Australia
CT004	Entity 011	Australia
CT004	Entity 014	Australia
CT004	Entity 014	Australia
CT004	Entity 014	Australia
CT005	Entity 015	Australia
CT005	Entity 016	Australia
CT005	Entity 019	Bangladesh
CT005	Entity 017	Australia
CT005	Entity 019	Bangladesh
CT005	Entity 020	Bangladesh

Country	Region	Statutory Tax Rate
Adriatic	Europe	0.1500
Algeria	AFRICA	0.2000
Angola	AFRICA	0.3000
Argentina	LATIN AMERICA	0.3500
Australia	Asia Pacific	0.3000
Austria	Europe	0.2500
Azerbaijan	Russia/Caucasus	0.2000
Bahrain	Middle East	0.2500
Bangladesh	Asia Pacific	0.2500
Barbados	LATIN AMERICA	0.2500
Belgium	Europe	0.2398
Bolivia	LATIN AMERICA	0.2500
Bosnia and Herzegovina	Europe	0.2000
Brazil	LATIN AMERICA	0.2400

Item Name	Calculation Group
Accounted Intercompany	Total Permanent Differences
Accounted RPI/RE	Total Permanent Differences
ADD 01	Total Permanent Tax Adjustments
Branch Profit Tax	Total Cash Tax Adjustment
Currency Gains/Loss	Total Permanent Differences
Contribution to State and Local Income Taxes	Total Permanent Differences
Deemed PRT AG	Total Permanent Differences
Deemed PRT Tax	Total Cash Tax Adjustment
Fiscal Unity Income Adj	Total Permanent Differences
Foreign Dividends	Total Permanent Differences
GST/VAT Disabled	Total Permanent Differences
IC Home Office Expenses	Total Permanent Differences
IC Interest	Total Permanent Differences
IC Management Fees	Total Permanent Differences
IC Royalty	Total Permanent Differences

Workflow



Outputs

OTF Code	Legal Entity	Country	Region	Total Pre-Tax Book Income	Pre-Tax Book Income Plus Perms	Statutory Tax Rate	Tax Liability	Total State and Local Income Taxes
CT0015	Entity 015	United States	North America	9,250,335	9,400,837	0	0	3,293,023
CT0016	Entity 016	United States	North America	(206,115,366)	(204,644,724)	0	(172,322,642)	(172,322,642)
CT0017	Entity 017	United States	North America	(2,365,628)	(2,365,628)	0	(793,978)	(793,978)
CT0018	Entity 018	United States	North America	17,855,719	17,855,719	0	6,238,813	6,238,813
CT0019	Entity 019	United States	North America	(6,973)	(6,973)	0	(2,440)	(2,440)
CT0020	Entity 020	United States	North America	1,103	1,103	0	407	407
CT0021	Entity 021	United States	North America	(8,999)	(8,999)	0	(2,449)	(2,449)
CT0022	Entity 022	United States	North America	301,817	301,817	0	105,566	105,566
CT0023	Entity 023	United States	North America	(4,238,400)	(5,078,048)	0	(1,762,617)	(1,762,617)
CT0024	Entity 024	United States	North America	(12,261)	(12,261)	0	(4,268)	(4,268)
CT0025	Entity 025	United States	North America	(7,422)	(7,422)	0	(2,346)	(2,346)
CT0026	Entity 026	United States	North America	32,420,716	37,623,543	0	13,171,739	13,171,739
CT0027	Entity 027	United States	North America	-	-	0	-	-
CT0028	Entity 028	United States	North America	(44,875,209)	(44,748,456)	0	(15,862,316)	(15,862,316)
CT0029	Entity 029	United States	North America	66,214,657	66,214,641	0	16,564,374	16,564,374
CT0030	Entity 030	United States	North America	14,628,289	14,627,544	0	3,224,640	3,224,640
CT0031	Entity 031	United States	North America	136,000	136,000	0	40,603	40,603
CT0032	Entity 032	United States	North America	3,800,803	3,800,803	0	3,322,822	3,322,822
CT0033	Entity 033	United States	North America	90,343,488	91,488,493	0	32,014,473	32,014,473
CT0034	Entity 034	United States	North America	602,202	602,206	0	333,282	333,282
CT0035	Entity 035	United States	North America	598,249	598,249	0	136,527	136,527
CT0036	Entity 036	United States	North America	8,195,697	8,210,604	0	3,174,516	3,174,516
CT0037	Entity 037	United States	North America	-	-	0	-	-
CT0038	Entity 038	United States	North America	(31,403,633)	(31,403,633)	0	(16,991,330)	(16,991,330)
CT0039	Entity 039	United States	North America	(1,256,453)	(1,256,453)	0	(481,983)	(481,983)
CT0040	Entity 040	United States	North America	(214)	(214)	0	(75)	(75)
CT0041	Entity 041	United States	North America	-	-	0	-	-
CT0042	Entity 042	United States	North America	-	-	0	-	-
CT0043	Entity 043	United States	North America	-	-	0	-	-
CT0044	Entity 044	United States	North America	29,938,100	29,938,100	0	7,307,666	7,307,666
CT0045	Entity 045	United States	North America	-	-	0	-	-
CT0046	Entity 046	United States	North America	-	-	0	-	-
CT0047	Entity 047	United States	North America	-	-	0	-	-
CT0048	Entity 048	United States	North America	607	607	0	(24)	(24)
CT0049	Entity 049	United States	North America	(32,814)	(32,814)	0	(11,485)	(11,485)
CT0050	Entity 050	United States	North America	(815,664)	(815,664)	0	(264,366)	(264,366)
CT0051	Entity 051	United States	North America	(605,924)	(607,478)	0	(126,478)	(126,478)

A.1.

Visibly determine how many unique customers are included in both tables.

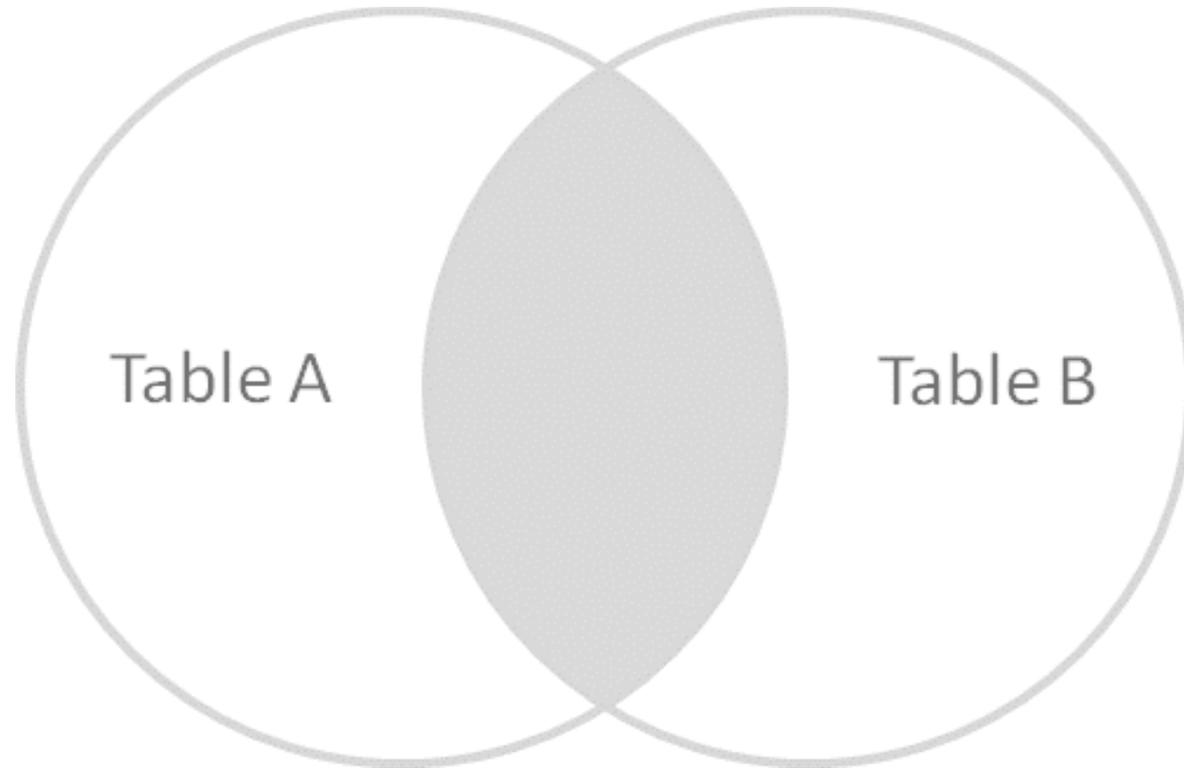
tblAge

Name	Age
Karen	60
Juan	42
Jay	39
Xiao	27
Connie	71

tblOccupation

Name	Occupation
Rhiana	Accountant
Juan	Nurse
Jay	Carpenter
Jing	Teacher
Connie	Technologist

Inner join



INNER JOIN

How many records would be in the output?

tblAge

Name	Age
Karen	60
Juan	42
Jay	39
Xiao	27
Connie	71

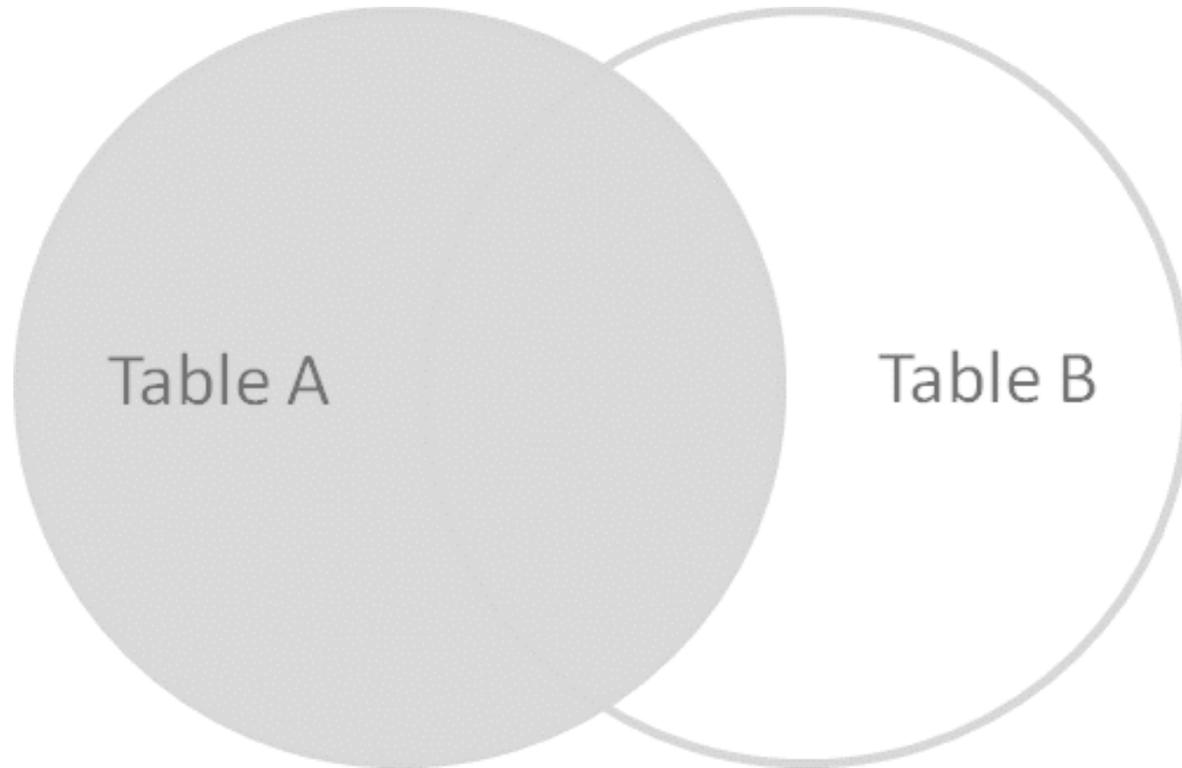
tblOccupation

Name	Occupation
Rhiana	Accountant
Juan	Nurse
Jay	Carpenter
Jing	Teacher
Connie	Technologist

INNER JOIN OUTPUT (FLAT FILE)

Name	Age	Occupation
Juan	42	Nurse
Jay	39	Carpenter
Connie	71	Technologist

Left outer join



LEFT OUTER JOIN

How many records would be in the output?

tblAge

Name	Age
Karen	60
Juan	42
Jay	39
Xiao	27
Connie	71

tblOccupation

Name	Occupation
Rhiana	Accountant
Juan	Nurse
Jay	Carpenter
Jing	Teacher
Connie	Technologist

LEFT OUTER JOIN OUTPUT (FLAT FILE)

Name	Age	Occupation
Karen	60	
Juan	42	Nurse
Jay	39	Carpenter
Xiao	27	
Connie	71	Technologist

RIGHT OUTER JOIN

How many records would be in the output?

tblAge

Name	Age
Karen	60
Juan	42
Jay	39
Xiao	27
Connie	71

tblOccupation

Name	Occupation
Rhiana	Accountant
Juan	Nurse
Jay	Carpenter
Jing	Teacher
Connie	Technologist

RIGHT OUTER JOIN OUTPUT (FLAT FILE)

Name	Age	Occupation
<u>Rhiana</u>		Accountant
Juan	42	Nurse
Jay	39	Carpenter
Jing		Teacher
Connie	71	Technologist

Exercise: Understanding Joins and Unions



There are seven unique customers. We want seven records in our output file.

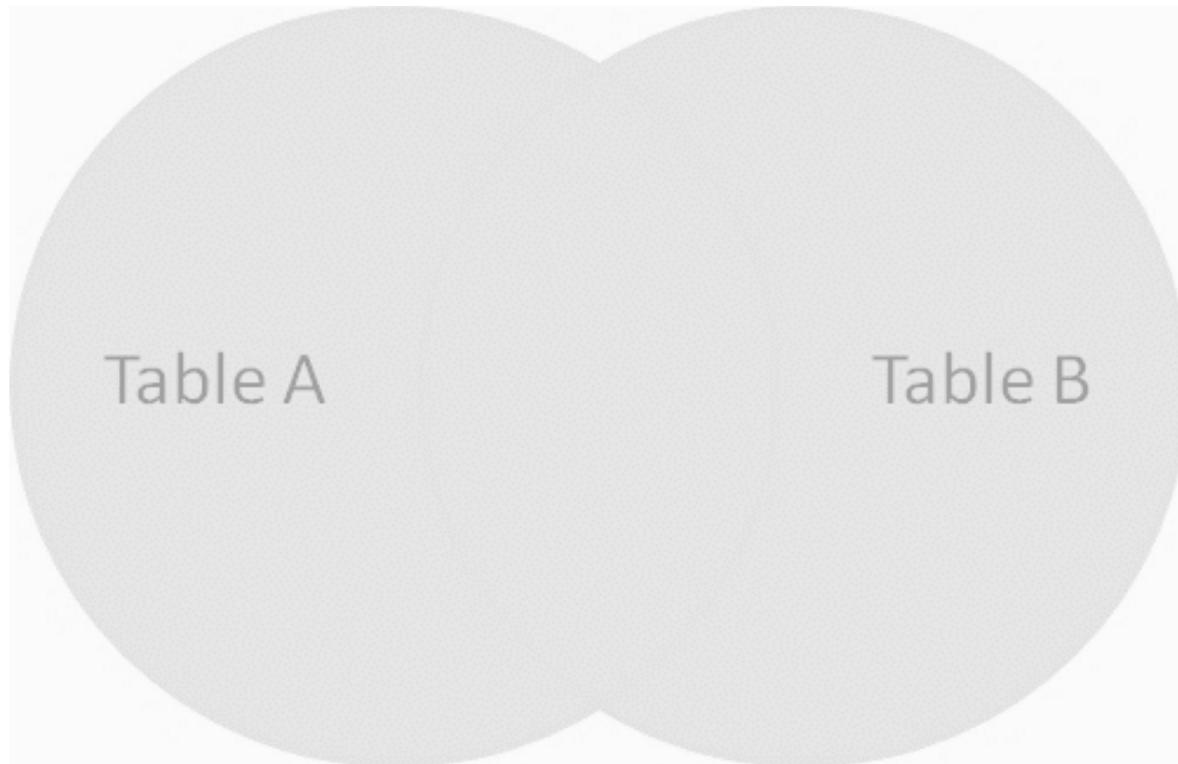
tblAge

Name	Age
Karen	60
Juan	42
Jay	39
Xiao	27
Connie	71

tblOccupation

Name	Occupation
Rhiana	Accountant
Juan	Nurse
Jay	Carpenter
Jing	Teacher
Connie	Technologist

Full outer join



FULL OUTER JOIN OUTPUT (FLAT FILE)

Name	Age	Occupation
Karen	60	
Juan	42	Nurse
Jay	39	Carpenter
Xiao	27	
Connie	71	Technologist
<u>Rhiana</u>		Accountant
Jing		Teacher

Exercise: Understanding Joins and Unions



A.2.

Visibly determine how many unique salespeople had sales in either January or February.

tblSales-January

Name	Sales
Masauda	149
Heather	172
Terry	110

tblSales-February

Name	Sales
Masauda	126
Lynn	138
Heather	141

A.3.

You want to generate a list of sales by salesclerk by month to compute commissions due.

tblSales-January

Name	Sales
Masauda	149
Heather	172
Terry	110

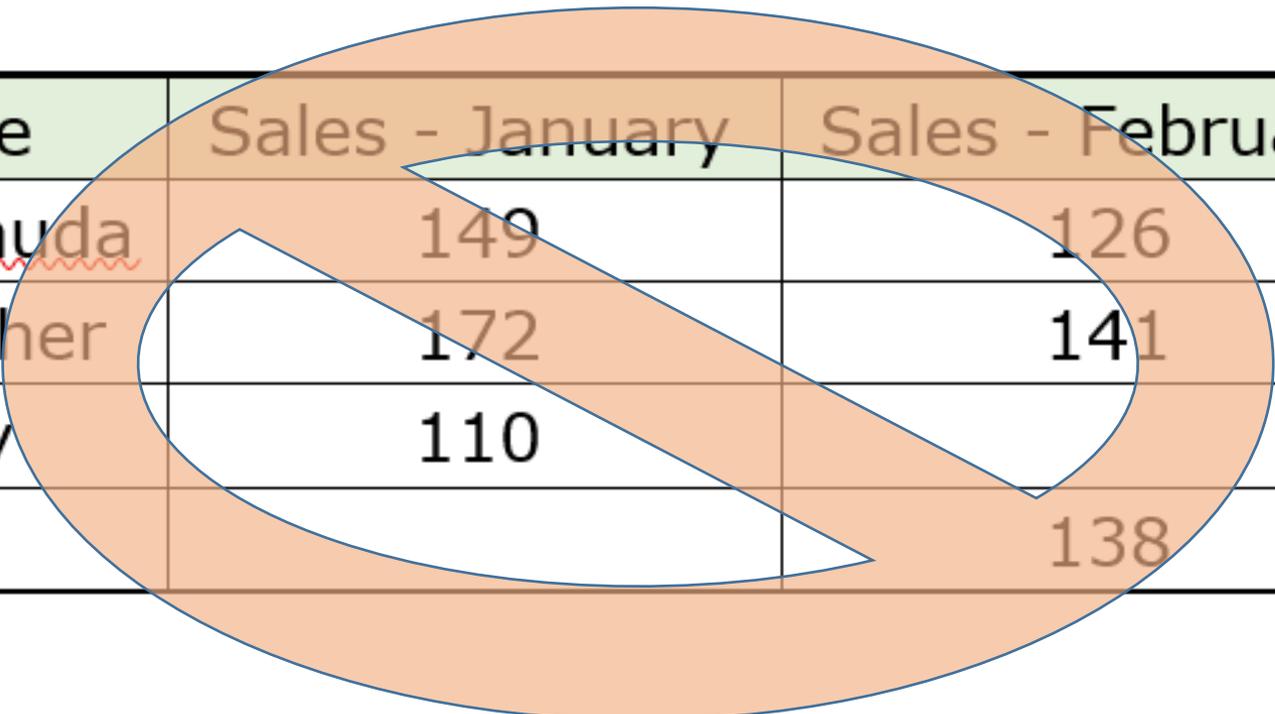
tblSales-February

Name	Sales
Masauda	126
Lynn	138
Heather	141

Is an outer join appropriate?

FULL OUTER JOIN OUTPUT (FLAT FILE)

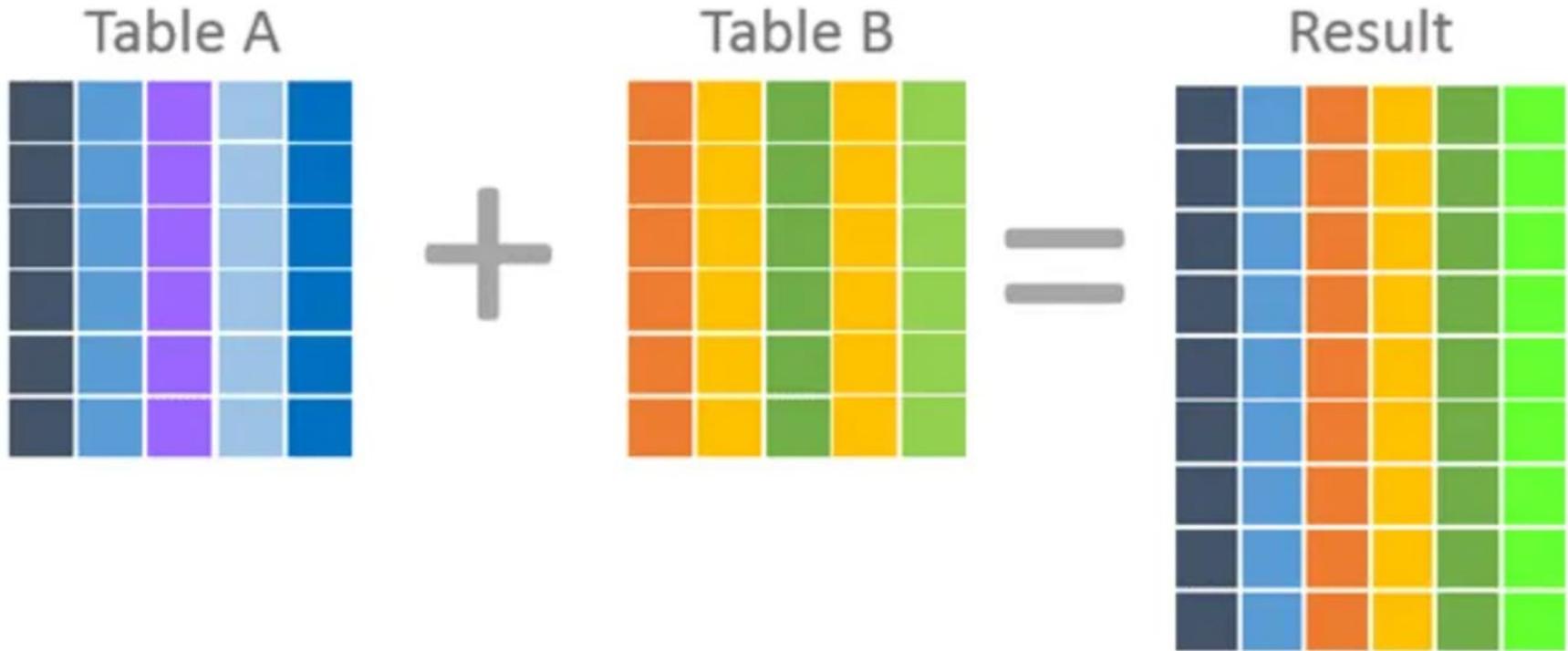
Name	Sales - January	Sales - February
Masauda	149	126
Heather	172	141
Terry	110	
Lynn		138



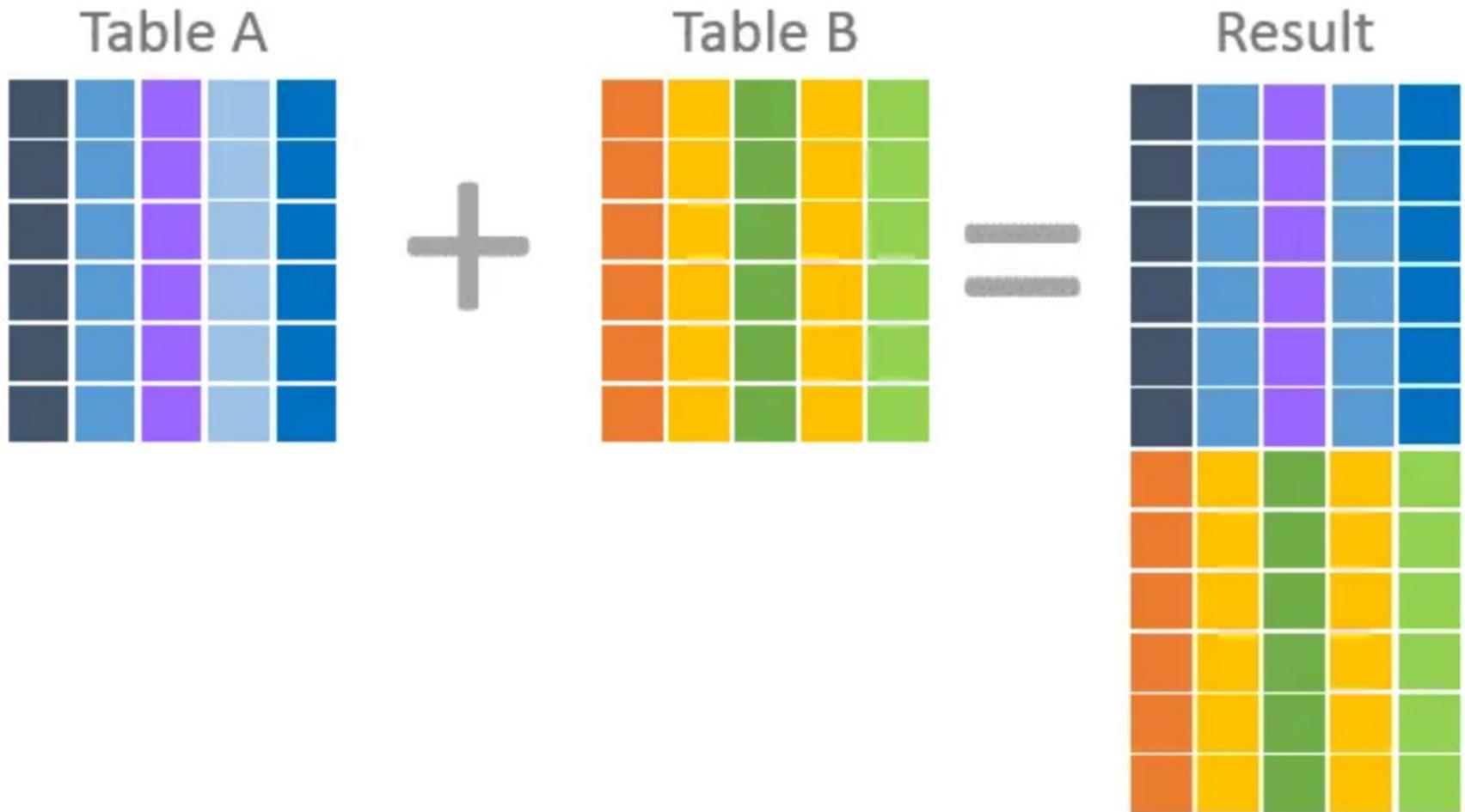
UNION (FLAT FILE)

Name	Sales	Source
<u>Masauda</u>	149	<u>tblJanuary</u>
Heather	172	<u>tblJanuary</u>
Terry	110	<u>tblJanuary</u>
<u>Masauda</u>	126	<u>tblFebruary</u>
Lynn	138	<u>tblFebruary</u>
Heather	141	<u>tblFebruary</u>

Joins add columns



Unions add rows





alteryx | The Thrill
of Solving

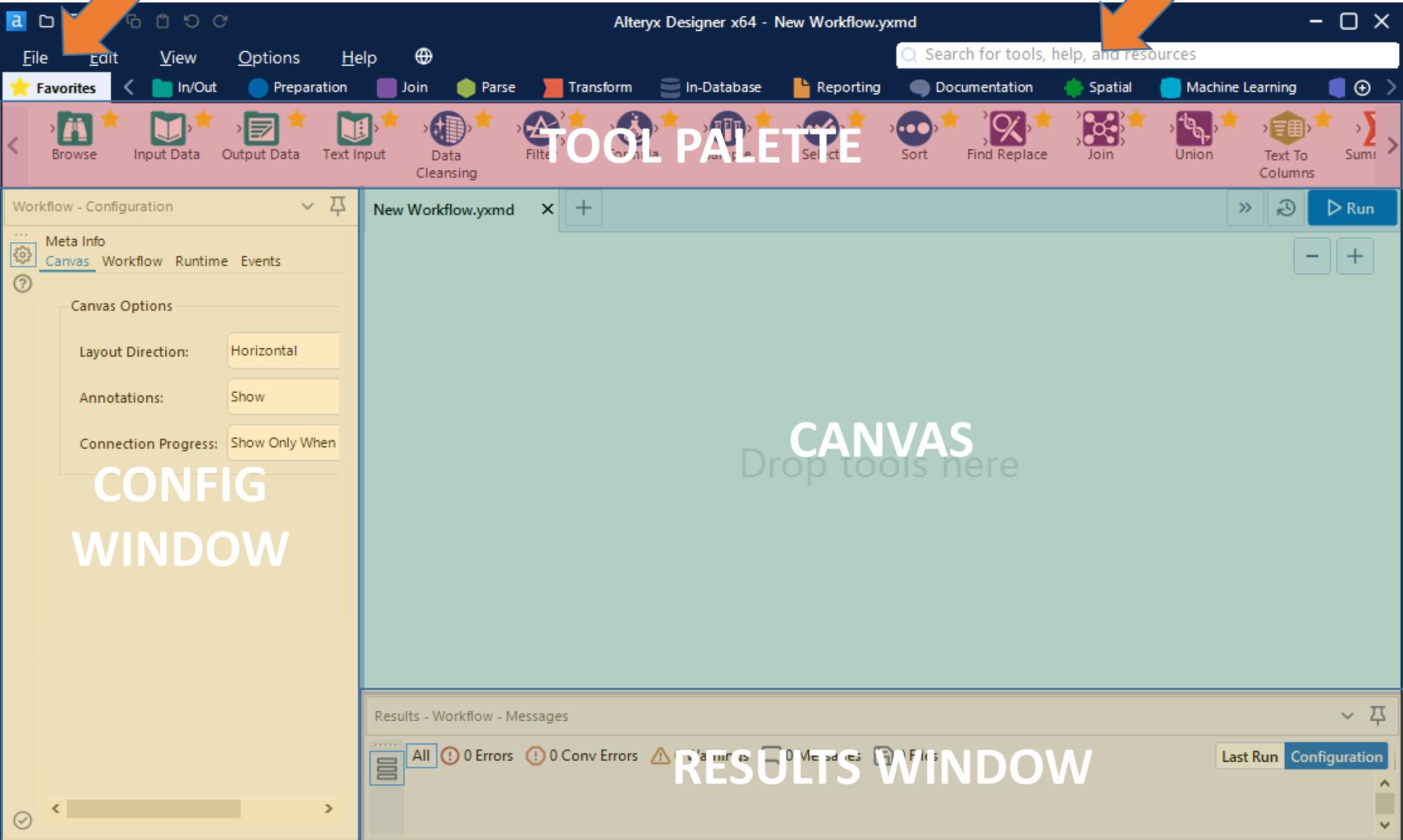
THE AGE OF THE BADASS ANALYST

Seize the day. Explore modern approaches to data analytics that unleash your brilliance.

**BECOME A
BADASS ANALYST.**

Prep and blend data
in minutes, not days.

Alteryx primary components

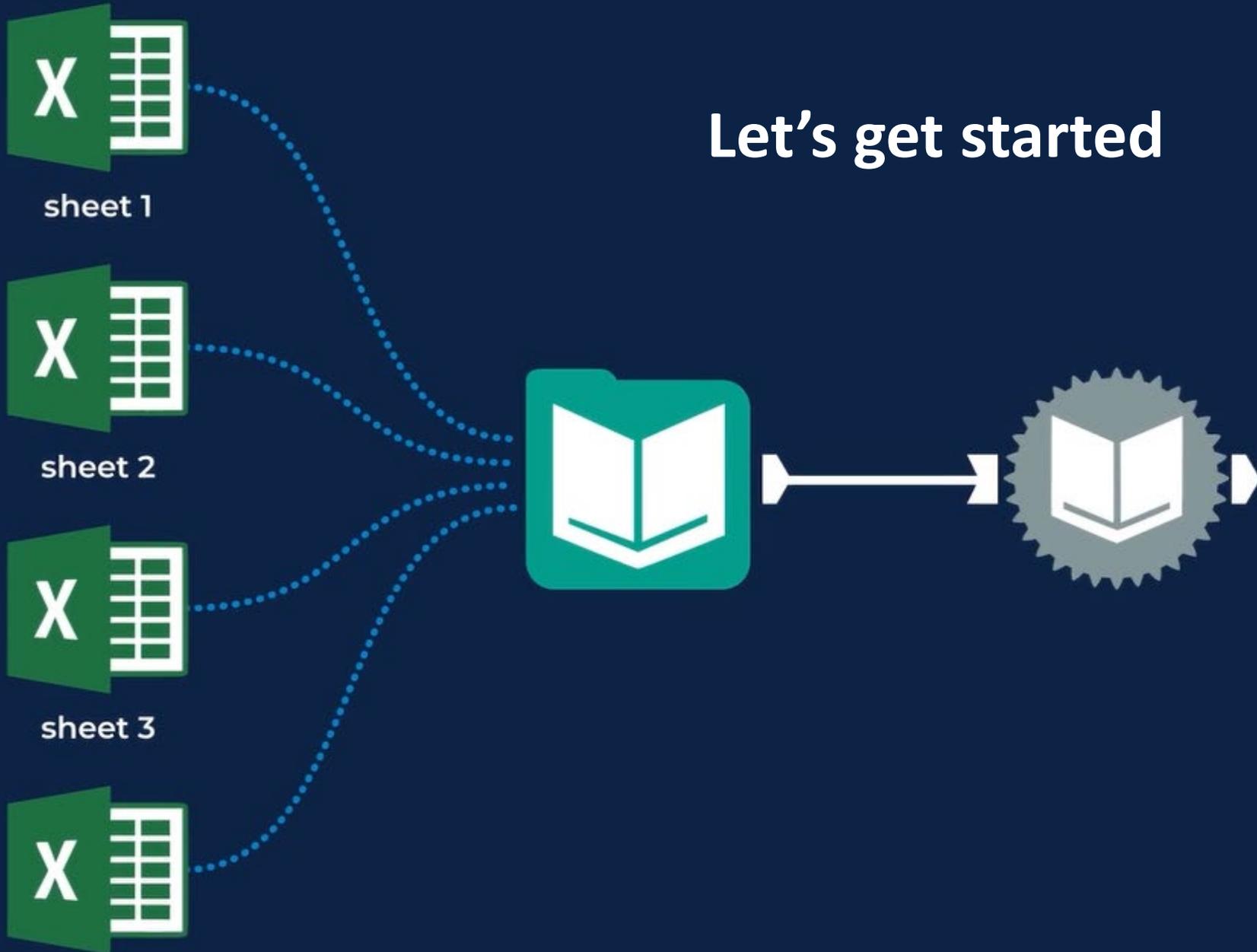


TOOL PALETTE

CONFIG WINDOW

RESULTS WINDOW

Let's get started

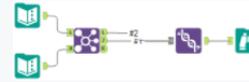
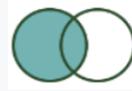


Hands-on Exercise

Joins and Unions.yxmd

Opening this file will initiate the Alteryx application with the tblAge, tblOccupation, tblSales-January, and tblSales-February data.

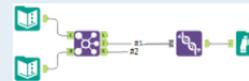
Left Outer Join: All records from the L input, including the records that joined with the R input.



To do a Left Outer Join, connect the J and L outputs of the Join tool to the Union tool.

Connect the J output first to establish the combined table schema.

Right Outer Join: All records from the R input including the records that joined with the L input.



To do a Right Outer Join, connect the J and R outputs of the Join tool to the Union tool.

Connect the J output first to establish the combined table schema.

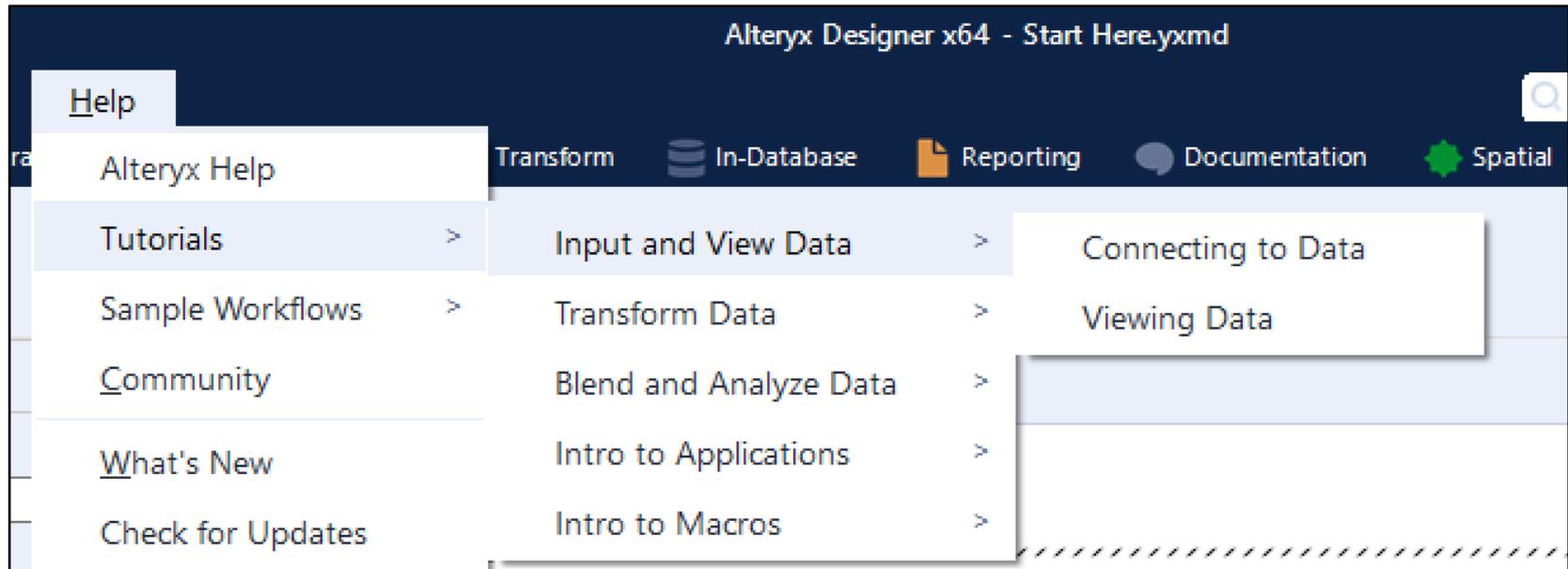
- **Workflows (.yxmd)**

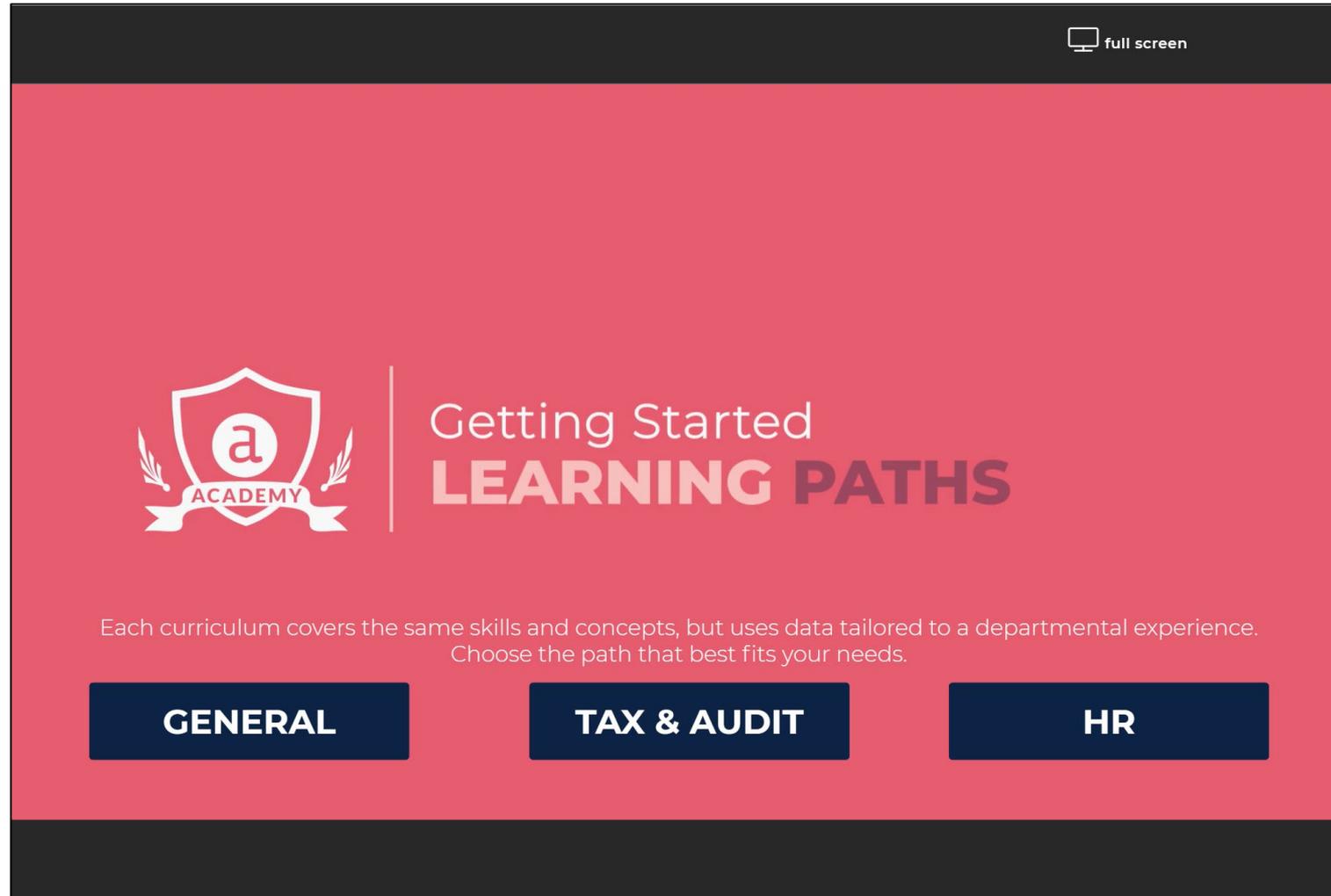
- Workflows without a copy of the data
- > **File > Save**

- **Packaged Workflows (.yxzp)**

- Same as .yxmd plus a copy of the data
- First save your workflow (.yxmd)
- > **Option > Export Workflow**

- In Alteryx: Help tab > Tutorials





full screen



Getting Started
LEARNING PATHS

Each curriculum covers the same skills and concepts, but uses data tailored to a departmental experience.
Choose the path that best fits your needs.

GENERAL **TAX & AUDIT** **HR**

[Learning Paths - Alteryx Community](#)

full screen

- Getting Started
 - Using Designer
 - Benefits of Designer
- Input Data
 - Text Input
 - Datatypes
 - Select
 - Browse
 - Filter
 - Sort
 - Sample
 - Practice Exercise 1
- Text to Columns
 - Unique
 - Union
- Join**
- Practice Exercise 2
- DateTime
 - Rows vs Columns
- Functions & Expression Editor
 - Formula
 - Function Types
 - Summarize
- Practice Exercise 3
- Transpose & Cross Tab
 - Find Replace
 - Append Fields
 - Output Data
- Practice Exercise 4
- Capstone Project

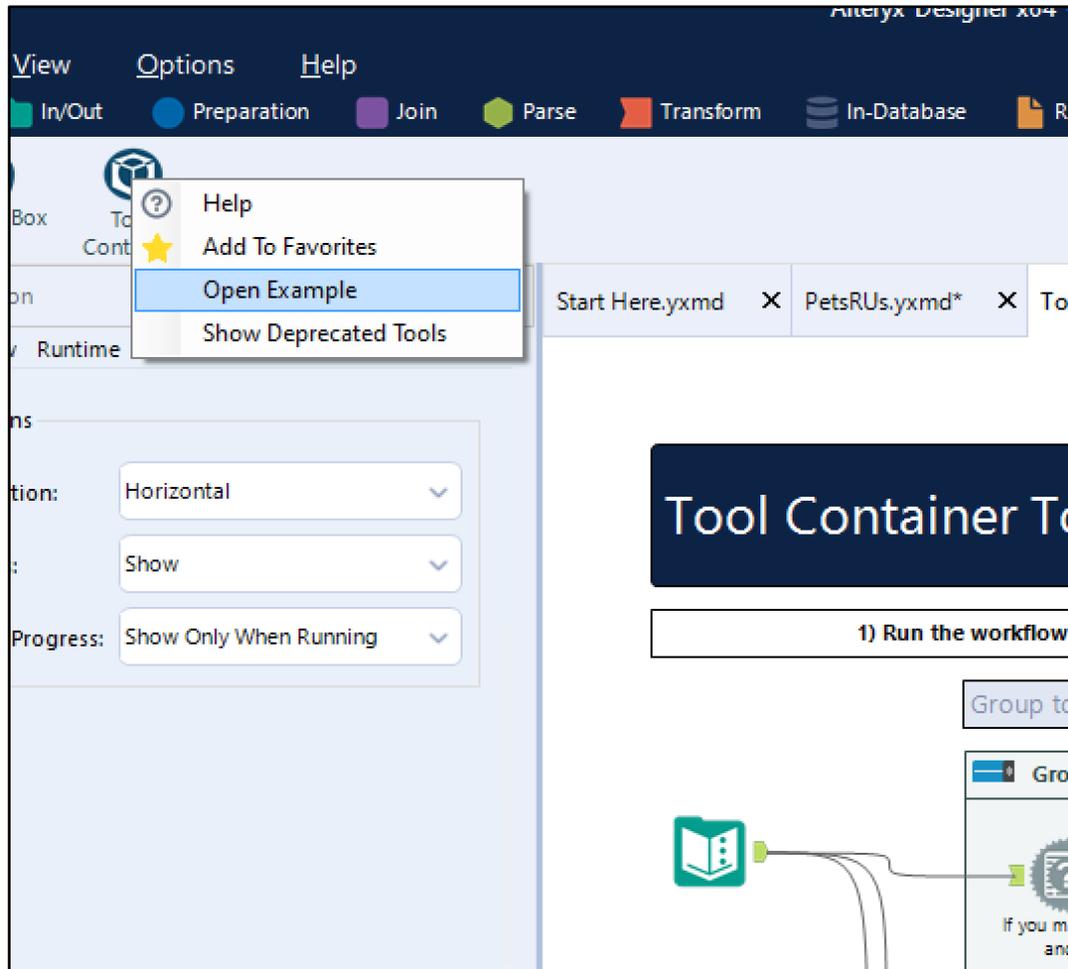
JOIN

Combine two data streams based on common fields (or record position). In the joined output, each row will contain the data from both inputs.

If you need to combine data horizontally, the **Join** tool can utilize a common field or combine by position. If both incoming datasets share a common column, joining on that column can be used to match rows of data. Alternatively, if you are confident that the row order of the datasets match, you can join by record position. This tool is very powerful and makes it easy to work with multiple data sources or combine disparate data streams.

- Interactive Learning
- Tool Mastery
- TRY IT!

- Right click on any tool; select “Open Example”



Thank You!

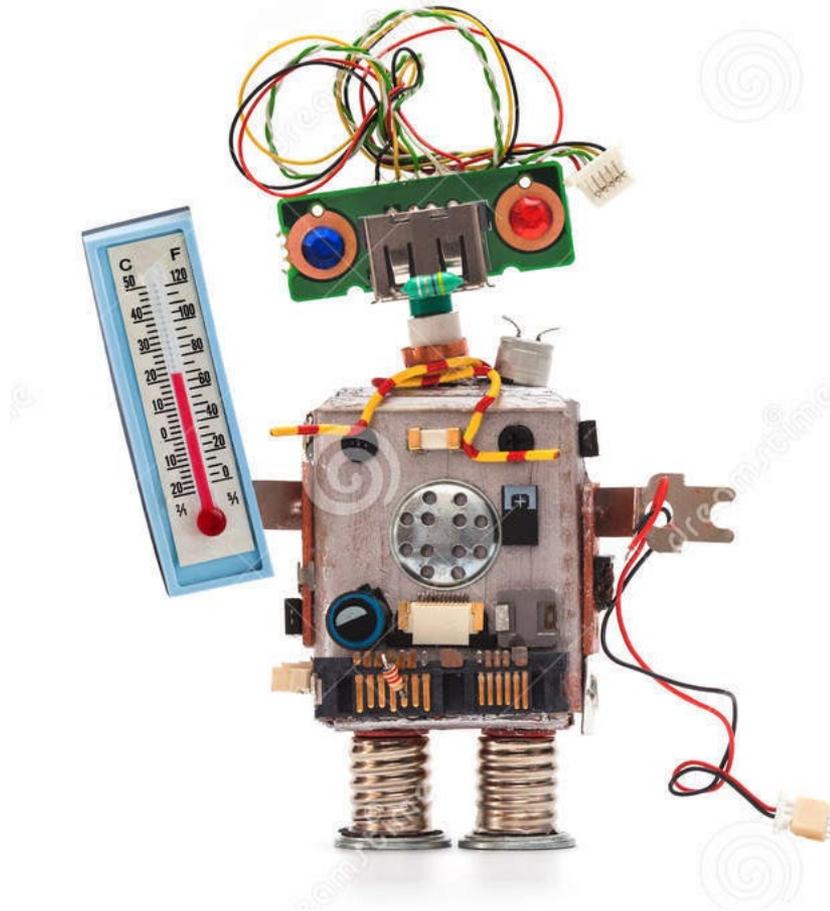


Image from cartoondealer.com