# Preparing Data for Use

## Pooja Gandhi

LEAD ANALYTICS ENGINEER | PLURALSIGHT
2X TABLEAU ZEN MASTER

@DrexelPooja  thedataduo.com

# Overview

Data interpreter

Transform your data

Union your data

# Data Interpreter

# What does data interpreter do?

**Gives you a head start when cleaning your data**

**Detects titles, notes, footnotes, empty cells**

**Bypasses flaws to identify actual fields and values in your dataset**

**Detect tables and sub-tables within your worksheets**

**Available for**

- Microsoft Excel
- Text files
- PDF files
- Google Sheets

# Example

| Annual Report | | | |
|---|---|---|---|
| Year | Order ID | Order Date | Ship Date |
| 2017 | CA-2017-152156 | 11/8/17 | 11/11/17 |
| | CA-2017-152156 | 11/8/17 | 11/11/17 |
| | CA-2017-138688 | 6/12/17 | 6/16/17 |
| 2016 | US-2016-108966 | 10/11/16 | 10/18/16 |
| | US-2016-108966 | 10/11/16 | 10/18/16 |
| 2015 | CA-2015-115812 | 6/9/15 | 6/14/15 |
| | CA-2015-115812 | 6/9/15 | 6/14/15 |
| | CA-2015-115812 | 6/9/15 | 6/14/15 |
| | CA-2015-115812 | 6/9/15 | 6/14/15 |
| | CA-2015-115812 | 6/9/15 | 6/14/15 |
| | CA-2015-115812 | 6/9/15 | 6/14/15 |
| | CA-2015-115812 | 6/9/15 | 6/14/15 |

# Demo

In this demo, you will learn –

How to use data interpreter

# Keep in Mind

**Data interpreter is not available**

- If the data source is already in the format Tableau can interpret
- If the data contains many rows and many columns
  - 2000 columns
  - 3000 rows and 150 columns
- If the data source is not supported

# Transform Your Data

# Transform Your Data

| | | | |
|---|---|---|---|
| Rename | Create Calculations | Pivot | Change Data Type |
| Hide | Split | Sort | Alias Fields |

# Demo

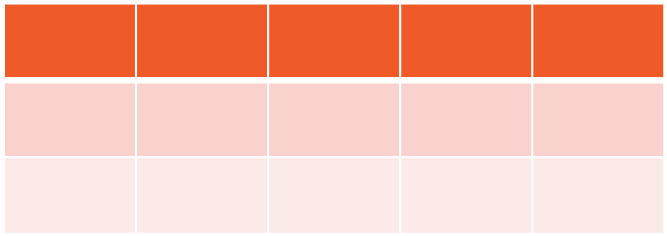
In this demo, you will learn –

How to transform your data

# Union Your Data

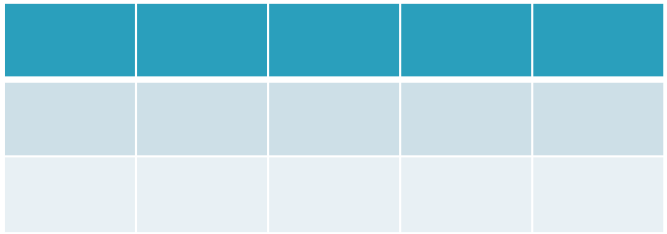
# Union Your Data

Table 1

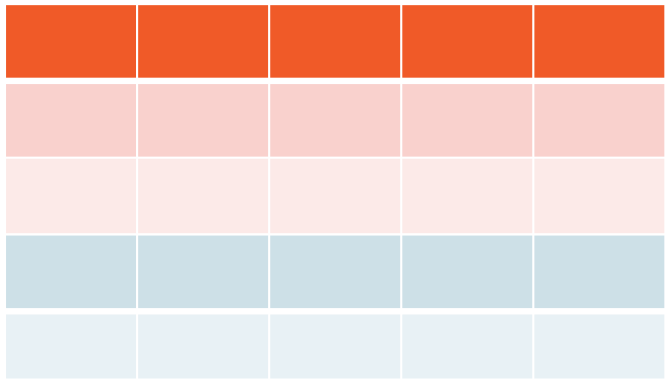


+

Table 2



Single Table



# Demo

In this demo, you will learn –

Union your data

# Summary

In this module, we learnt –

Data interpreter

Transform your data

Union your data

# Things to Remember

Use data prep if your data needs more cleaning

To union your data, the tables must come from the same connection

Fields generated as a result of a union can be used in a pivot

Stored procedures can not be unioned