

# **Excel**

## **Data Preparation**

Alish Bista - 08/20/2025

# Basic Steps

# Basic Steps

## Doing The Data Laundry

- Collect/Import the raw data
- Explore the data. Sort and filter as you see fit.
- Format the data
- Eliminate duplicates and errors
- Enrich or Summarize if necessary

# SubTotal

# SubTotal

## Checking All Levels

- Automatically groups and calculates data
- Works with sorted data - groups similar items together first
- Calculation options: Sum, Average, Count, Max, Min
- **Examples:**
  - Sales by region
  - Expenses by department

# Lookups

# Lookups

## Tell Me More

- Use relationships between tables to enrich data
- Using a common key, one can go from one table and look for associated records in another
- ***Functions:***
  - VLOOKUP
  - HLOOKUP
  - XLOOKUP

# VLOOKUP

## The Most Looked Up Lookup Function

- Syntax: *VLOOKUP(lookup\_value, table\_array, col\_index\_num, [range\_lookup])*
- **Arguments:**
  - *lookup\_value* - value to search for
  - *table\_array* - table or dataset to be searched
  - *col\_index\_num* - column number with return value
  - *range\_lookup* - FALSE for exact match or TRUE for approximate match

# HLOOKUP

## The Unsung One

- Syntax: *HLOOKUP(lookup\_value, table\_array, **row\_index\_num**, [range\_lookup])*
- **Arguments:**
  - **row\_index\_num** - row number with return value
  - Searches data horizontally

# XLOOKUP

## The Rising Star!

- Syntax: *XLOOKUP(lookup\_value, lookup\_array, return\_array, [if\_not\_found], [match\_mode], [search\_mode])*
- **Arguments:**
  - *lookup\_value* - value to search for
  - *lookup\_array* - array or range to search
  - *return\_array* - array or range to return
  - *[if\_not\_found]* - Value to return when valid match is not defined

# INDEX/MATCH

# INDEX/MATCH

The Matchmaking Power Couple

- **INDEX**(table\_array, [row\_num], [column\_num]) - Returns value at specific position
- **MATCH**(lookup\_value, lookup\_array, [match\_type]) - Finds position of value
- **Together:** **INDEX**(table\_array, **MATCH**(lookup\_value, lookup\_array, FALSE), [column\_num])

# Pivot Tables

# Pivot Tables

Pivot. Pivot. Pivot!!!

- Helps in summarizing data
- Aggregates and populates data in dynamic summary tables
- Easily produces basic aggregations like Average, Sum, and Count

# Points of Pivot

## Getting Ready To Pivot

- Generally, fields in a table have two types of information:
  - **Data:** Contains numerical data to be summarized
  - **Category:** Describe the data
- In a **Pivot Table,**
  - One or more data fields get summarized
  - Category fields appear in Rows, Columns, or Filters

# Parts of Pivot

## Parts that Pivot Pack

- **Column Labels:** Field in the column orientation (appearing horizontally)
- **Row Labels:** Field in the row orientation (appearing vertically)
- **Values:** Cells that contain summary data
- **Filter:** Field used to limit the data being summarized Category fields
- **Subtotals:** Row or Column that displays subtotals for detail cells
- **Grand Total:** Row or Column that displays total for all detail cells
- **Group:** A collection of items treated as a single entity
- **Item:** Elements in a field that appears as row or column header

# Q&A