


3

Chapter 3 : Excel Basics = Cleaning & Combining Data

- Find and replace → CTRL + H
- Leading and trailing space → TRIM() function
- Check duplicates → Conditional Formatting
- Splitting Columns → text to column converter wizard
- Combining Data from two tables (Very Important in Interviews)
 - **VLOOKUP()** - this function is used to look up a value in a table by searching for a corresponding value in the leftmost column of another table.
 - Syntax : = VLOOKUP(lookup_value, table_Array, col_index_num, [range_lookup])
 - =VLOOKUP([movie_id], Financials[#All], 2, FALSE)
 -=VLOOKUP(Movies[@[movie_id]:[movie_id]], Financials[#All], 4, FALSE)
 - Movies[@[movie_id]:[movie_id]] → this is to lock  the lookup_value
 - Limitation of VLOOKUP:
 - it searches only the first columns of a reference table for a matching value
 - Adding a new column to the reference table can cause errors by shifting column indexes and changing the referenced column.
 - food[#All], food[#Data], food[#Header] → to get the table information
 - **INDEX()** -
 - Syntax : = INDEX(array, row_num, [column_num])
 - = index(food_table_to_look, row_num, col_num)

- **MATCH()** -
 - Syntax : =MATCH(lookup_value, lookup_array, [match_type])
 - = match("item_to_find", food[item], 0)
 - = match("price", food[#Headers])
- **INDEX()_MATCH()** - are two powerful Excel functions that are often used together to search for & retrieve data from a table or range.
 - Attempts: for the budget column
 - =INDEX(Financials, 1, 2)
 - =INDEX(Financials, MATCH([@[movie_id]], Financials[movie_id], 0), 2)
 - =INDEX(Financials, MATCH([@[movie_id]], Financials[movie_id], 0), MATCH("budget", Financials[#Headers],0))
 - =INDEX(Financials, MATCH(Movies[@[movie_id]:[movie_id]], Financials[[movie_id]:[movie_id]], 0), MATCH(Movies[#Headers], [budget]], Financials[#Headers],0))
 - Financials[movie_id]—the effect is the same as
 - Financials[movie_id]:[movie_id]—> but here we are doing lock when we drag the column to other
 - Attempts: For revenue column — dynamically from revenue col
 - =INDEX(Financials, MATCH([@[movie_id]:[movie_id]], Financials[[movie_id]:[movie_id]], 0), MATCH("revenue", Financials[#Headers],0))
 - =INDEX(Financials, MATCH(Movies[@[movie_id]:[movie_id]], Financials[[movie_id]:[movie_id]], 0), MATCH(Movies[#Headers], [revenue]], Financials[#Headers],0))
 - Same as for Unit & Currency columns as :
 - =INDEX(Financials, MATCH(Movies[@[movie_id]:[movie_id]], Financials[[movie_id]:[movie_id]], 0), MATCH(Movies[#Headers], [unit]], Financials[#Headers],0))
 - =INDEX(Financials, MATCH(Movies[@[movie_id]:[movie_id]], Financials[[movie_id]:[movie_id]], 0), MATCH(Movies[#Headers], [currency]], Financials[#Headers],0))

- **XLOOKUP()** - Unlike VLOOKUP, which only searches in the leftmost column of a table, XLOOKUP is capable of searching any column.
- Syntax : =XLOOKUP(lookup_value, lookup_array, return_array, [if_not_found], [match_mode])
- Attempts for budget column:
 - =XLOOKUP([@[movie_id]],Financials[#All],[movie_id],Financials[#All],[budget],"Not Available",0)
 - =XLOOKUP(Movies[@[movie_id]:[movie_id]],Financials[#All],[movie_id]:[movie_id],Financials[#All],[budget],"Not Available",0)
- Same dynamic change for revenue, Unit, currency columns respectively as follows:
 - =XLOOKUP(Movies[@[movie_id]:[movie_id]],Financials[#All],[movie_id]:[movie_id],Financials[#All],[revenue],"Not Available",0)
 - =XLOOKUP(Movies[@[movie_id]:[movie_id]],Financials[#All],[movie_id]:[movie_id],Financials[#All],[unit],"Not Available",0)
 - =XLOOKUP(Movies[@[movie_id]:[movie_id]],Financials[#All],[movie_id]:[movie_id],Financials[#All],[currency],"Not Available",0)
 - Look at @the reference:

The screenshot shows an Excel spreadsheet with a table of movies. The table has columns for movie_id, title, industry, release_year, imdb_rating, studio, language_id, budget, revenue, unit, and currency. A dialog box for the XLOOKUP function is open, showing the following arguments:

- Lookup_value: 101
- Lookup_array: Financials[All] (movie_id)
- Return_array: Financials[All] (budget)
- [if_not_found]: "Not Available"
- [match_mode]: 0

The dialog box also includes a search range and a match mode section. The formula result is 1.

| movie_id | title | industry | release_year | imdb_rating | studio | language_id | budget | revenue | unit | currency |
|----------|---|-----------|--------------|-------------|---------------------------|-------------|--------|---------|------|----------|
| 101 | K.G.F: Chapter 2 | Hollywood | 2019 | 8.2 | Paramount Pictures | 1 | 5 | 5 | 5 | 5 |
| 102 | Doctor Strange in the Multiverse of Madness | Hollywood | 2022 | 7.0 | Universal Pictures | 5 | 5 | 5 | 5 | 5 |
| 103 | Thor: The Dark World | Hollywood | 2013 | 7.0 | Universal Pictures | 5 | 5 | 5 | 5 | 5 |
| 104 | Thor: Ragnarok | Hollywood | 2017 | 8.0 | Universal Pictures | 5 | 5 | 5 | 5 | 5 |
| 105 | Thor: Love and Thunder | Hollywood | 2022 | 6.0 | Universal Pictures | 5 | 5 | 5 | 5 | 5 |
| 106 | Sholay | Hollywood | 2019 | 8.4 | Marvel Studios | 5 | 5 | 5 | 5 | 5 |
| 107 | Dhruva Dhanraaj Le Jayenge | Hollywood | 2018 | 8.4 | Marvel Studios | 5 | 5 | 5 | 5 | 5 |
| 108 | 3 Idiots | Hollywood | 2019 | 8.5 | Not Available | 5 | 5 | 5 | 5 | 5 |
| 109 | Kabhi Khushi Kabhie Gham | Hollywood | 2019 | 8.5 | Not Available | 5 | 5 | 5 | 5 | 5 |
| 110 | Bajirao Mastani | Hollywood | 2019 | 8.5 | Not Available | 5 | 5 | 5 | 5 | 5 |
| 111 | The Shawshank Redemption | Hollywood | 1994 | 9.3 | Government of West Bengal | 7 | 5 | 5 | 5 | 5 |
| 112 | Inception | Hollywood | 2010 | 8.8 | Not Available | 5 | 5 | 5 | 5 | 5 |
| 113 | Interstellar | Hollywood | 2014 | 8.6 | Not Available | 5 | 5 | 5 | 5 | 5 |
| 115 | The Pursuit of Happyness | Hollywood | 2019 | 8.4 | Marvel Studios | 5 | 5 | 5 | 5 | 5 |
| 116 | Gladiator | Hollywood | 2000 | 8.5 | Not Available | 5 | 5 | 5 | 5 | 5 |
| 117 | Titanic | Hollywood | 1997 | 7.6 | Mythri Movie Makers | 2 | 5 | 5 | 5 | 5 |
| 118 | It's a Wonderful Life | Hollywood | 2019 | 8.5 | Not Available | 5 | 5 | 5 | 5 | 5 |
| 119 | Avatar | Hollywood | 2009 | 7.9 | Universal Pictures | 5 | 5 | 5 | 5 | 5 |
| 120 | The Godfather | Hollywood | 1972 | 9.2 | Paramount Pictures | 5 | 5 | 5 | 5 | 5 |
| 121 | The Dark Knight | Hollywood | 2008 | 9.0 | Syncopy | 5 | 5 | 5 | 5 | 5 |
| 122 | Schindler's List | Hollywood | 1993 | 9.0 | Universal Pictures | 5 | 5 | 5 | 5 | 5 |
| 123 | Jurassic Park | Hollywood | 1993 | 8.2 | Universal Pictures | 5 | 5 | 5 | 5 | 5 |
| 124 | Parasite | Hollywood | 2019 | 8.5 | Not Available | 5 | 5 | 5 | 5 | 5 |
| 125 | Avengers: Endgame | Hollywood | 2019 | 8.4 | Marvel Studios | 5 | 5 | 5 | 5 | 5 |
| 126 | Avengers: Infinity War | Hollywood | 2018 | 8.4 | Marvel Studios | 5 | 5 | 5 | 5 | 5 |
| 127 | Pathan Parichal | Hollywood | 1955 | 8.3 | Government of West Bengal | 7 | 5 | 5 | 5 | 5 |
| 128 | Taare Zameen Par | Hollywood | 2007 | 8.3 | Not Available | 5 | 5 | 5 | 5 | 5 |
| 129 | Munna Bhai M.B.B.S. | Hollywood | 2003 | 8.1 | Vinod Chopra Productions | 1 | 5 | 5 | 5 | 5 |
| 130 | PK | Hollywood | 2014 | 8.1 | Vinod Chopra Films | 1 | 5 | 5 | 5 | 5 |
| 131 | Sanju | Hollywood | 2018 | NULL | Vinod Chopra Films | 1 | 5 | 5 | 5 | 5 |
| 132 | Pushpa: The Rise - Part 1 | Hollywood | 2021 | 7.6 | Mythri Movie Makers | 2 | 5 | 5 | 5 | 5 |
| 133 | RRR | Hollywood | 2022 | 8 | DV Entertainment | 2 | 5 | 5 | 5 | 5 |
| 134 | Baahubali: The Beginning | Hollywood | 2015 | 8 | Arka Media Works | 2 | 5 | 5 | 5 | 5 |
| 135 | The Kashmir Files | Hollywood | 2022 | 8.3 | Fee Studios | 1 | 5 | 5 | 5 | 5 |
| 136 | Bajrangi Bhaijaan | Hollywood | 2015 | 8.1 | Salman Khan Films | 1 | 5 | 5 | 5 | 5 |