Wuzzuf Dashboard: Power Query Transformation Steps Documentation

This document details all transformation steps applied to the Wuzzuf job-postings dataset (`Wuzzuf-Jobs-Posting.csv`) in Power BI using Power Query Editor for the JobPulse Dashboard project. Each step includes a description, the action performed, and a placeholder for a screenshot. The transformations prepare the data for analysis in Power BI, creating a main `Wuzzuf-Jobs-Posting` table and a `Skills` table.

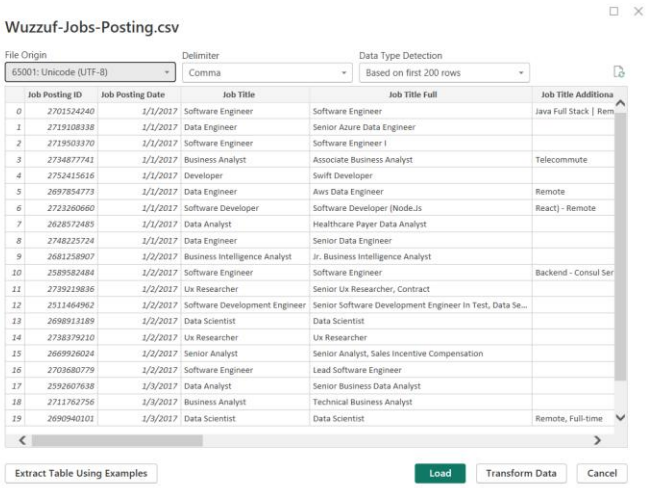
Load the Dataset

Import the raw `Wuzzuf-Jobs-Posting.csv` file into Power BI Power Query to begin transformations.

Steps:

- Opened Power BI Desktop.
- Clicked **Home** > **Get Data** > **Text/CSV**.
- Selected `Wuzzuf-Jobs-Posting.csv` from the local directory.
- Clicked **Transform Data** to open the Power Query Editor.
- Verified that all columns (e.g., `Job Posting ID`, `Job Title`, `Job Skills`, `Years of Experience`, `Number of Applicants`) loaded correctly.

Screenshot Placeholder:



Remove Blank Rows

Remove any completely blank rows to ensure data quality and avoid errors in downstream calculations.

Steps:

- In Power Query Editor, right-clicked the table.
- Selected **Remove Rows** > **Remove Blank Rows**.
- Checked the row count to confirm rows were removed.

Handle Missing Values in Numerical Columns

Replace null values in numerical columns (`Years of Experience`, `Number of Applicants`) with 0 to ensure accurate calculations.

Steps:

- Selected `Years of Experience` column.
- Right-clicked > **Replace Values** > replace `null` with `0`.
- Repeated for `Number of Applicants`.
- Verified no nulls remained using filters.

Filter Out Nulls in Critical Columns

Remove rows where critical columns (`Job Posting Date`, `Job Title`, `Job Position Level`) are null to maintain data integrity.

Steps:

- Filtered `Job Posting Date`, `Job Title`, and `Job Position Level` columns.
- Unchecked `(null)` values from the filter options.

Clean Job Skills Column (Remove Brackets and Quotes)

Description: Clean the `Job Skills` column by removing square brackets (`[]`) and quotes (`"`).

Steps:

- Selected `Job Skills` column.
- Replaced `[` and `]` with empty string.
- Replaced `"` with empty string.
- Trimmed whitespace using **Transform > Trim**.

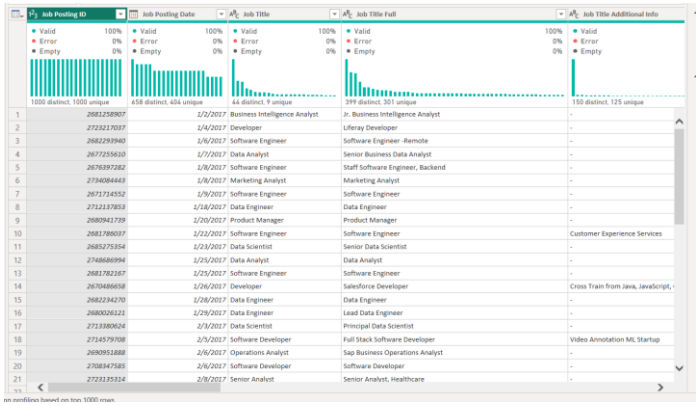
Split Job Skills Column into Rows

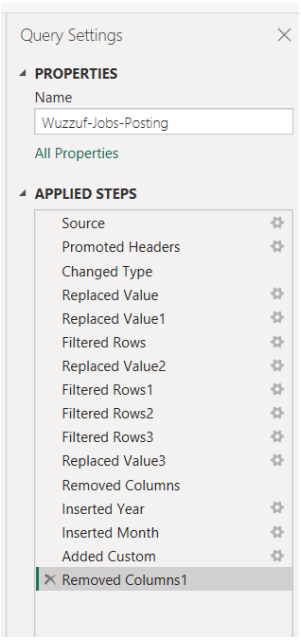
Split the `Job Skills` column by comma into separate rows for skill-level analysis.

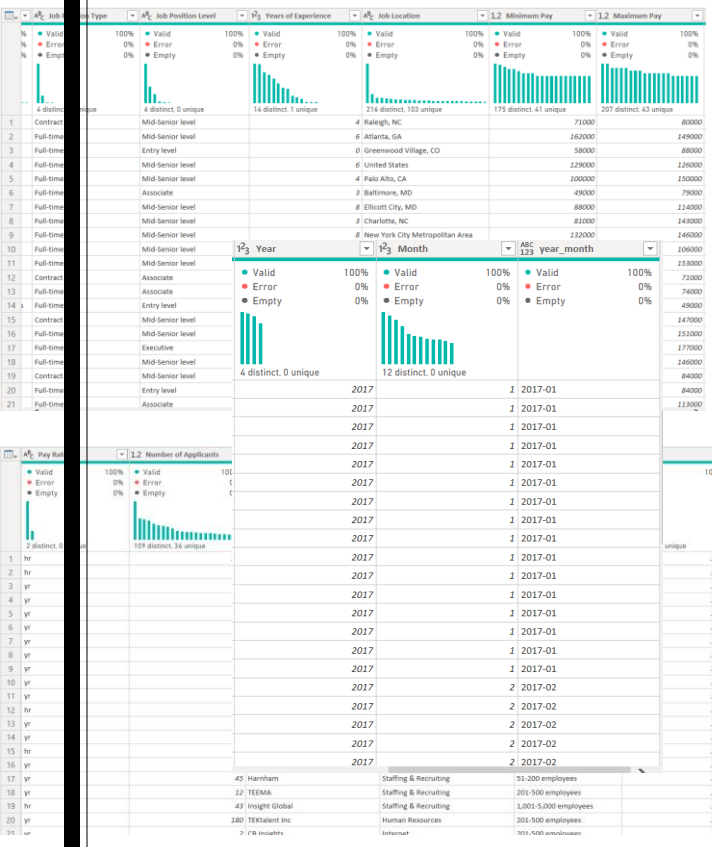
Steps:

- Selected `Job Skills` column.
- Clicked **Split Column** > **By Delimiter** > comma.
- Selected **Split into Rows** under Advanced Options.
- Renamed column to `Skill`.

Screenshots:







Trim and Clean Skills

Ensure `Skill` column has no extra spaces or residual characters after splitting.

Steps:

- - Selected `Skill` column.
- - Used **Transform > Trim** and optionally **Clean**.
- - Verified skill values are clean.

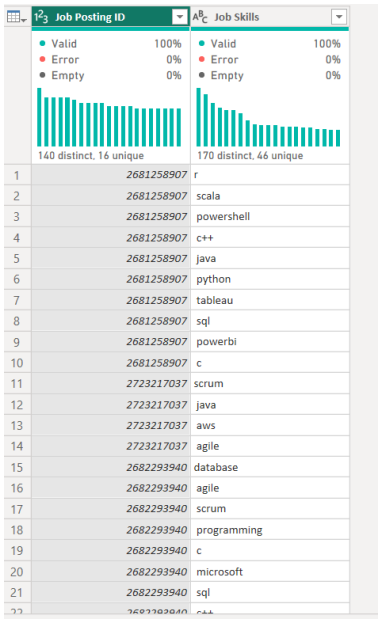
Create Skills Table

Create a dedicated `Skills` table with only `Job Posting ID` and `Skill`.

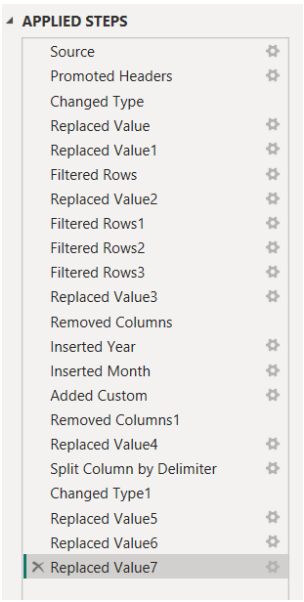
Steps:

- - Right-clicked query > **Duplicate**.
- - Renamed query to `Skills`.
- - Removed all columns except `Job Posting ID` and `Skill`.
- - Removed duplicates from `Skills` table.

Screenshot:



	Job Posting ID	Job Skills
	Valid 100%	Valid 100%
	Error 0%	Error 0%
	Empty 0%	Empty 0%
	140 distinct, 16 unique	170 distinct, 46 unique
1	2681258907	r
2	2681258907	scala
3	2681258907	powershell
4	2681258907	c++
5	2681258907	java
6	2681258907	python
7	2681258907	tableau
8	2681258907	sql
9	2681258907	powerbi
10	2681258907	c
11	2723217037	scrum
12	2723217037	java
13	2723217037	aws
14	2723217037	agile
15	2682293940	database
16	2682293940	agile
17	2682293940	scrum
18	2682293940	programming
19	2682293940	c
20	2682293940	microsoft
21	2682293940	sql



Extract Year and Month from Job Posting Date

Add `Year`, `Month`, and `Year-Month` columns for time-based filtering.

Steps:

- - Selected `Job Posting Date`.
- - Added `Year` and `Month` columns from **Add Column > Date**.
- - Created `Year-Month` using a custom column: `Text.From([Year]) & "-" & Text.PadStart(Text.From([Month]), 2, "0")`.

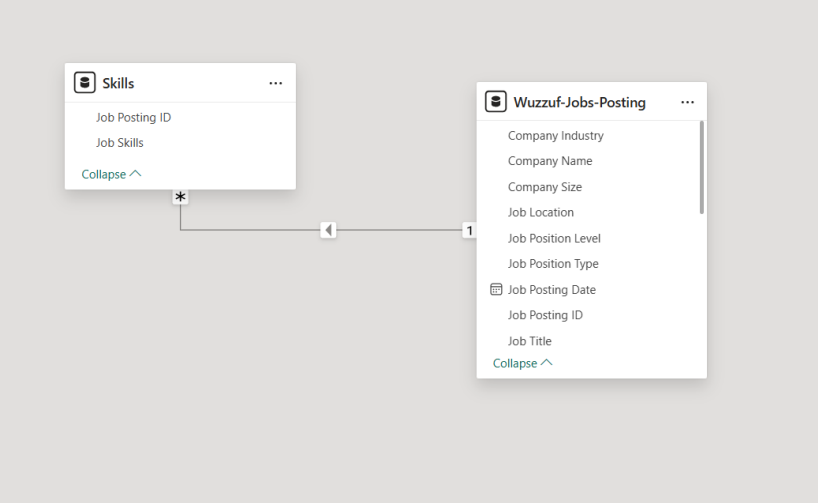
Load Queries to Data Model

Load both `Wuzzuf-Jobs-Posting` and `Skills` queries to Power BI data model.

Steps:

- - Clicked **Close & Apply** in Power Query Editor.
- - Confirmed both queries are added to the model.

Screenshot:



Notes

- **Dataset:** The `Wuzzuf-Jobs-Posting.csv` file contains 25,116 rows with columns like `Job Posting ID`, `Job Skills`, and `Years of Experience`. Transformations ensure clean, usable data.

- **Performance:** Splitting `Job Skills` into rows increases the `Skills` table size. Monitor query performance and consider filtering by `Job Posting Date` (e.g., 2020–2021) if slow.

- **Screenshots:** Capture each screenshot in Power BI Power Query Editor, ensuring the **Applied Steps** pane is visible to show the context. Save screenshots as PNG or JPG and insert them here.

- **Deliverable:** Include this documentation and screenshots in the `JobPulse_Deliverables` folder, alongside `Wuzzuf-Job-Posting-Cleaned.xlsx` and `JobPulse-Dashboard.pbix`.