Job Market
Data Pipeline

YOUR NEXT JOB IS ON
WUZZUF

Explore Now

Eng. Dina Mohsen Gamal El-din

Data Source Overview

WUZZUF is one of the leading job portals in Egypt and the Middle East, connecting job seekers with employers across various industries.
It allows users to browse job opportunities, apply online, and explore detailed job descriptions in a user-friendly way.

Project Overview

This project builds an end-to-end data pipeline to collect, clean, and structure job market data from **WUZZUF**, a job portal without a public API.

The pipeline demonstrates how to:

Scrape job data from a real-world website
Handle both **static and dynamic web content**
Clean and normalize raw scraped data
Design **fact and dimension tables** suitable for analysis

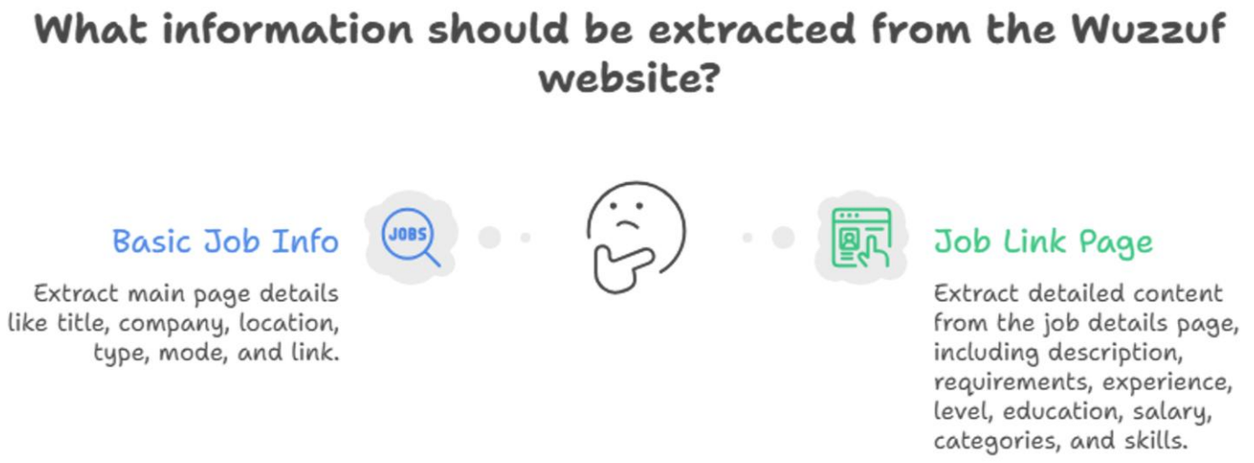
Goal

Transform raw job postings into **structured, analysis-ready datasets** that enable meaningful **job market insights**.

Project Components

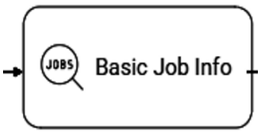
The pipeline consists of the following stages:

Extracted Job Market Data



Main Job Information

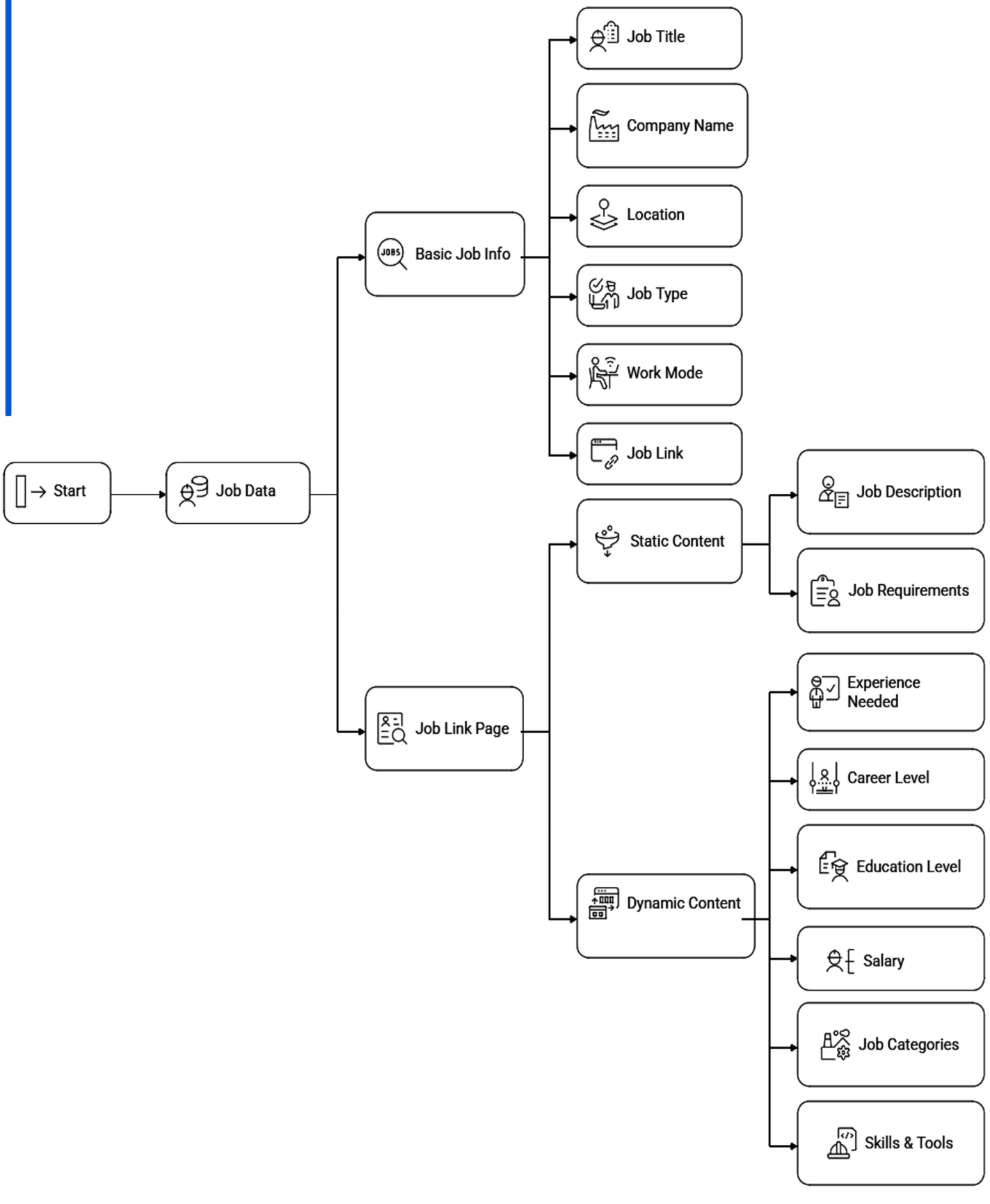
- Job Title
- Company Name
- Location
- Job Type (Full-time / Part-time)
- Job Link



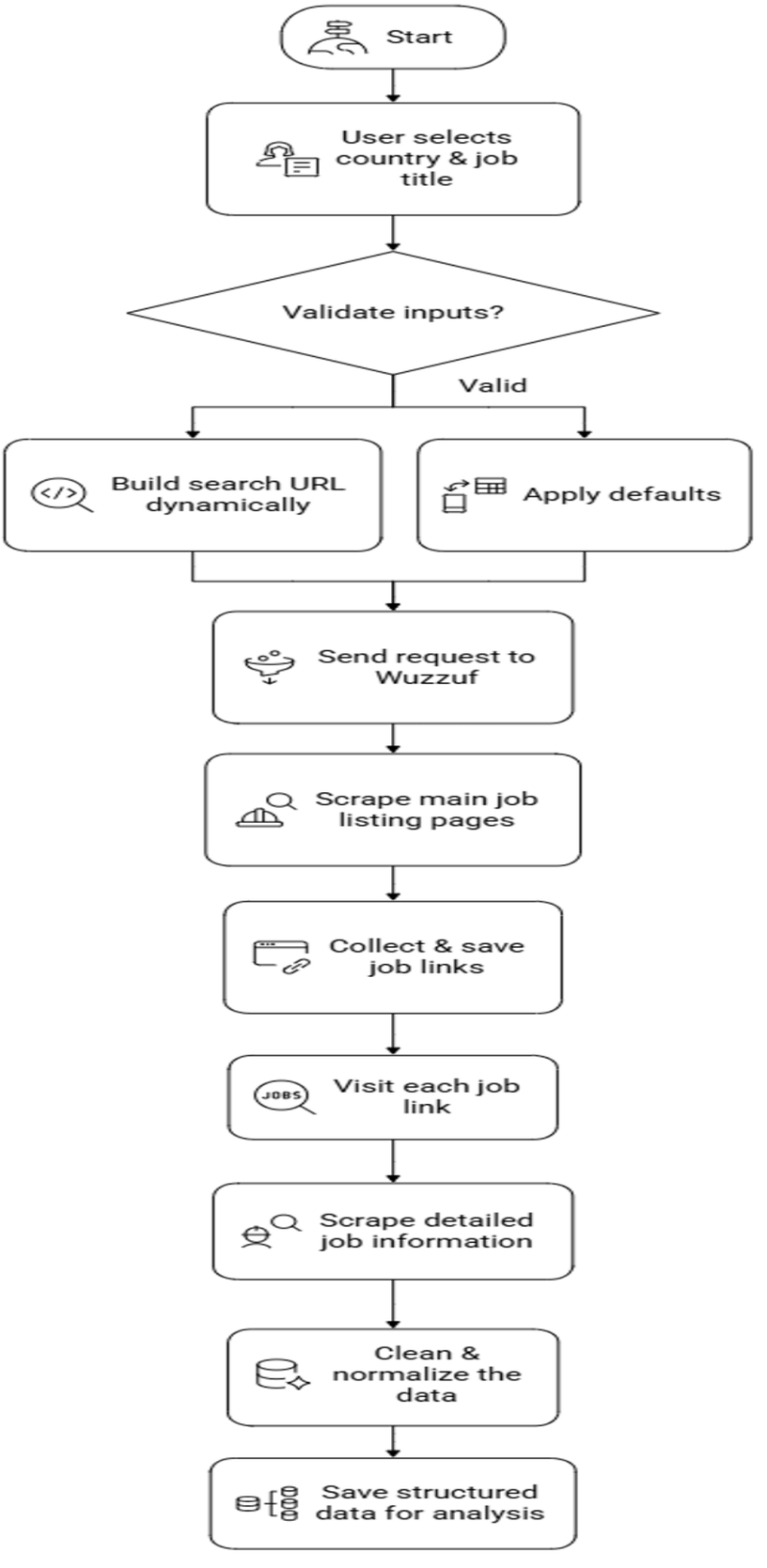
Detailed Job Information

- Job Description
- Job Requirements
- Required Experience
- Career Level
- Skills & Tools
- Job Categories

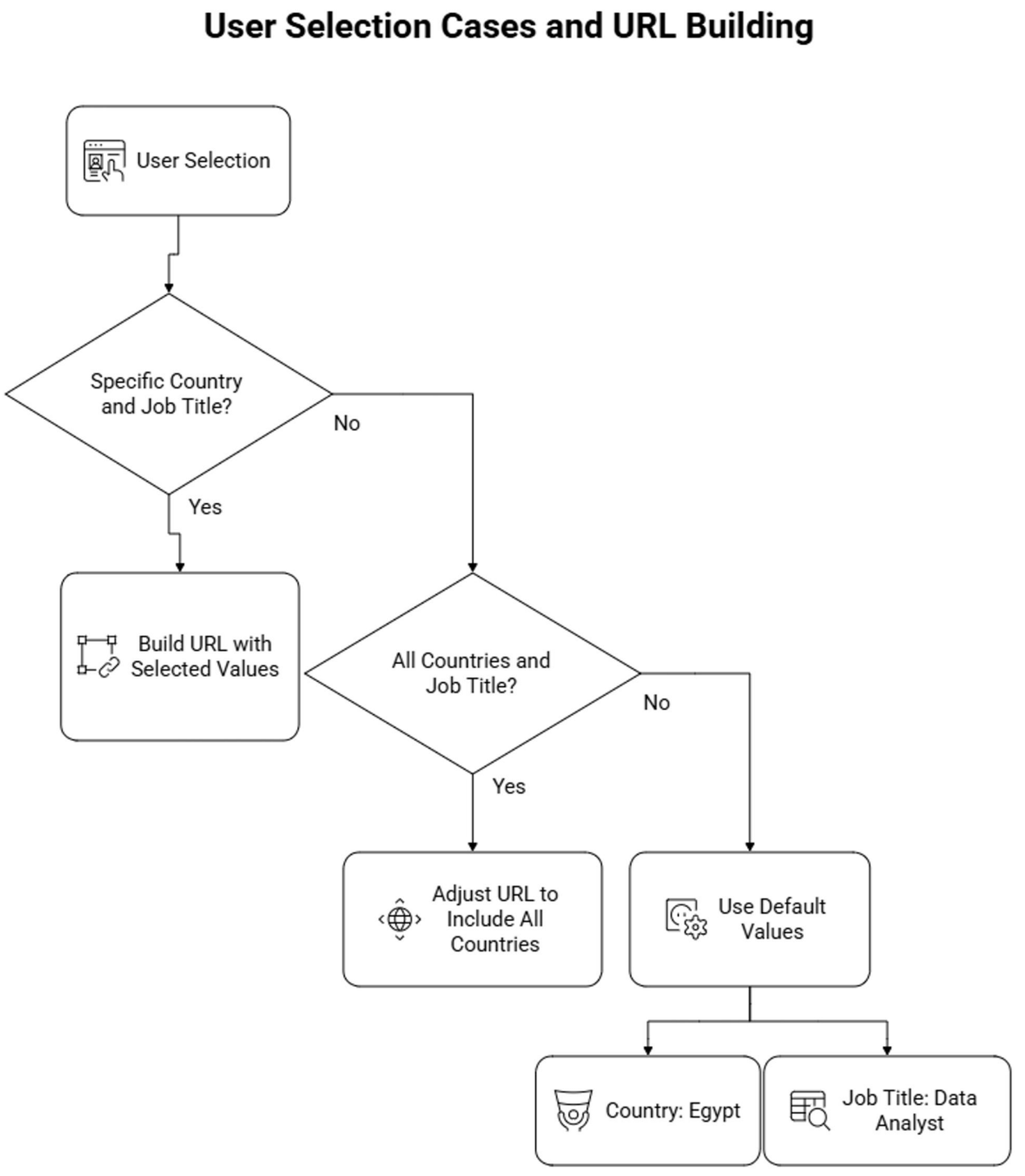
Information Extraction Process from Wuzzuf



Pipeline Workflow



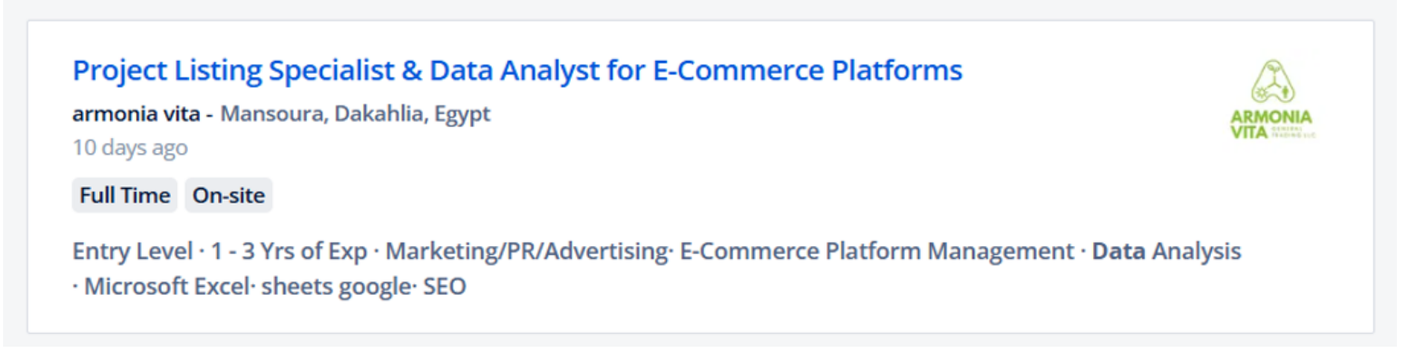
User Input & Selection



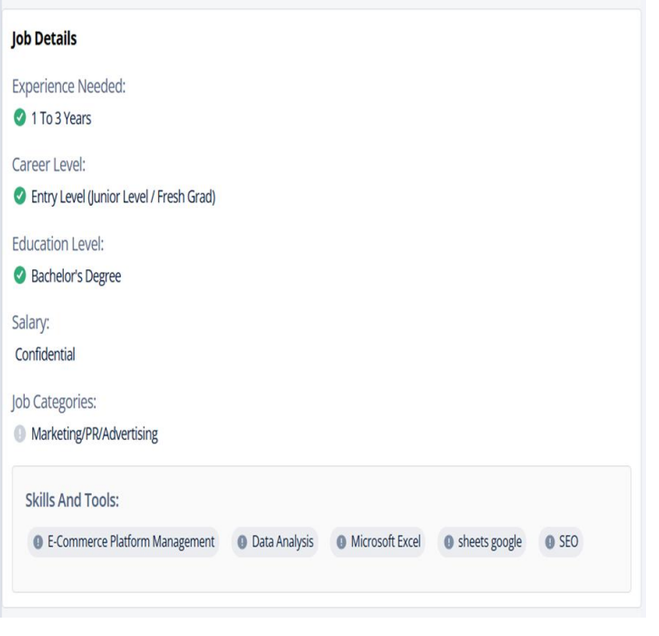
Web Scraping Process

Example from the Website

Main Page

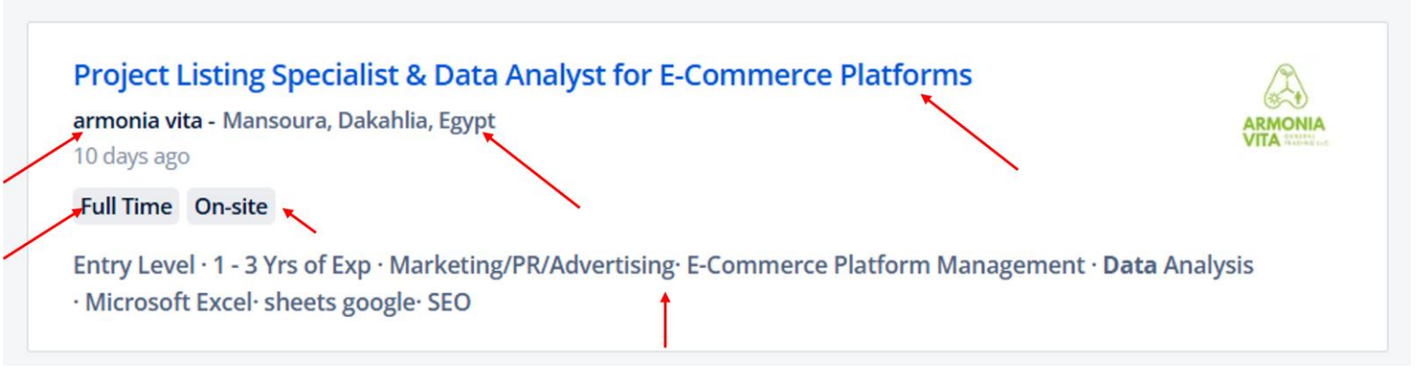


Job Details

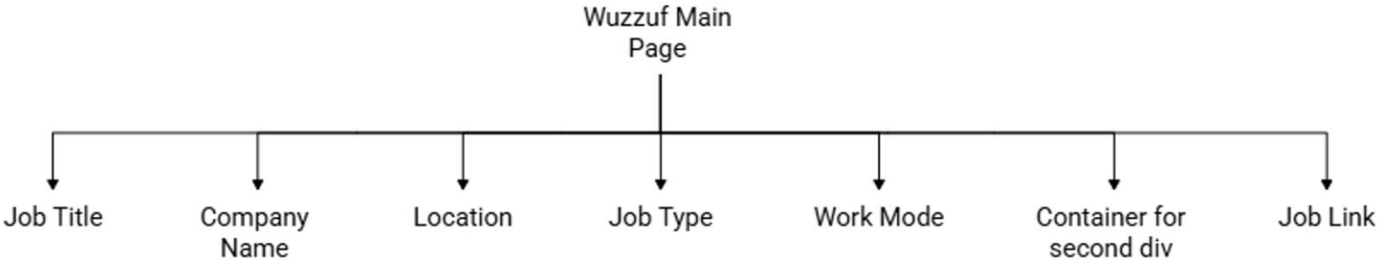


Web Scraping Process

Main Page Scraping

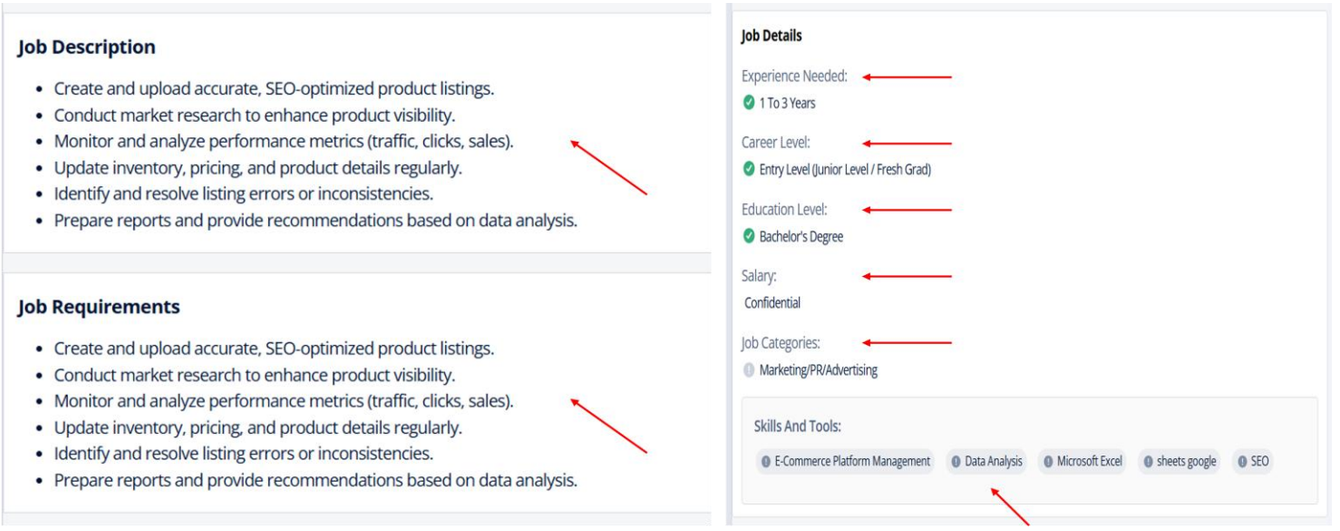


Wuzzuf Main Page Job Information Extraction

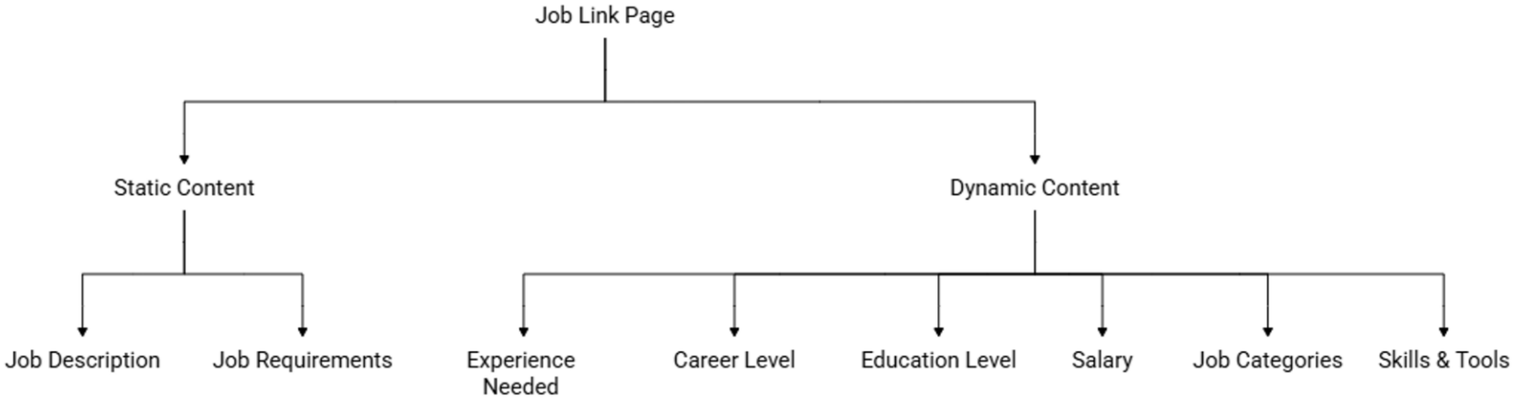


Web Scraping Process

Job Details Scraping



Job Link Page Content Extraction



Data Cleaning & Modeling

Raw DataFrames

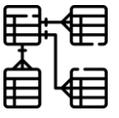
After scraping, the project produces multiple raw datasets:

- Main page data
- Job details data
- Description & requirements data

Data Cleaning

- Handle missing values
- Remove unnecessary characters
- Drop irrelevant columns
- Standardize text formats

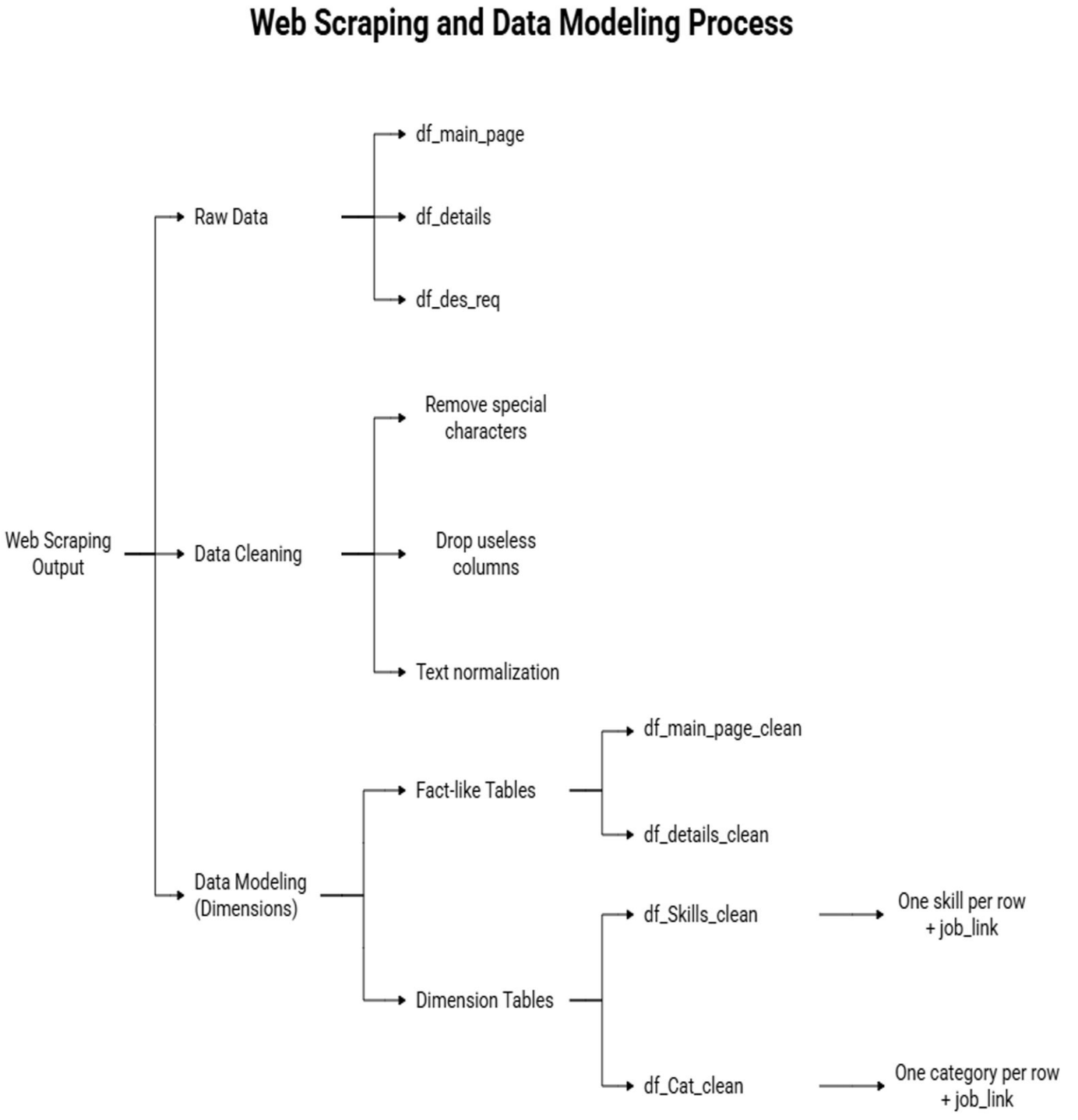
Data Modeling



Multi-valued columns (**Skills** and **Job Categories**) were normalized by splitting them into **separate dimension tables** to ensure clean and analyzable data.

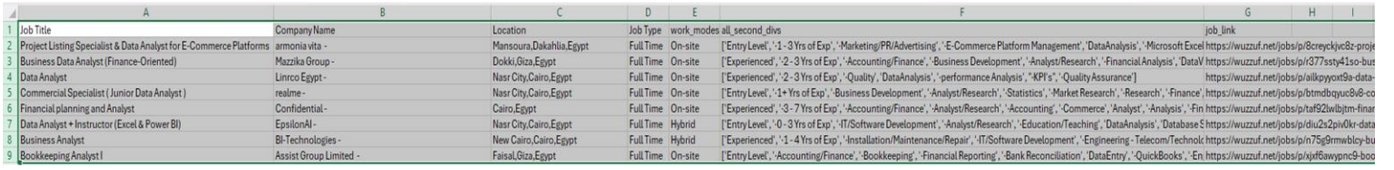
- One skill per row
- One category per row
- Linked using the job URL

Data Cleaning & Modeling



Data Transformation Results

Main Page – Raw Data (Before Cleaning)



	A	B	C	D	E	F	G	H	I
1	Job Title	Company Name	Location	Job Type	work_modes	all_second_divs	job_link		
2	Project Listing Specialist & Data Analyst for E-Commerce Platforms	armonia vita -	Mansoura,Dakahlia,Egypt	Full Time	On-site	['Entry Level', '-1 - 3 Yrs of Exp', '-Marketing/PR/Advertising', '-E-Commerce Platform Management', 'DataAnalysis', '-Microsoft Excel	https://wuzzuf.net/jobs/p/8creyckjvc8z-proje		
3	Business Data Analyst (Finance-Oriented)	Mazzika Group -	Dokki,Giza,Egypt	Full Time	On-site	['Experienced', '-2 - 3 Yrs of Exp', '-Accounting/Finance', '-Business Development', '-Analyst/Research', '-Financial Analysis', 'DataV	https://wuzzuf.net/jobs/p/r377ssty41so-bus		
4	Data Analyst	Linrco Egypt -	Nasr City,Cairo,Egypt	Full Time	On-site	['Experienced', '-2 - 3 Yrs of Exp', '-Quality', 'DataAnalysis', '-performance Analysis', '-KPI's", '-Quality Assurance']	https://wuzzuf.net/jobs/p/ailkpyyoxt9a-data-		
5	Commercial Specialist (Junior Data Analyst)	realme -	Nasr City,Cairo,Egypt	Full Time	On-site	['Entry Level', '-1+ Yrs of Exp', '-Business Development', '-Analyst/Research', '-Statistics', '-Market Research', '-Research', '-Finance'	https://wuzzuf.net/jobs/p/btmdbqyuc8v8-co		
6	Financial planning and Analyst	Confidential -	Cairo,Egypt	Full Time	On-site	['Experienced', '-3 - 7 Yrs of Exp', '-Accounting/Finance', '-Analyst/Research', '-Accounting', '-Commerce', 'Analyst', '-Analysis', '-Fin	https://wuzzuf.net/jobs/p/taf92lwlbjtm-finar		
7	Data Analyst + Instructor (Excel & Power BI)	EpsilonAI -	Nasr City,Cairo,Egypt	Full Time	Hybrid	['Entry Level', '-0 - 3 Yrs of Exp', '-IT/Software Development', '-Analyst/Research', '-Education/Teaching', 'DataAnalysis', 'Database S	https://wuzzuf.net/jobs/p/diu2s2pin0kr-data		
8	Business Analyst	BI-Technologies -	New Cairo,Cairo,Egypt	Full Time	Hybrid	['Experienced', '-1 - 4 Yrs of Exp', '-Installation/Maintenance/Repair', '-IT/Software Development', '-Engineering - Telecom/Technolc	https://wuzzuf.net/jobs/p/n75jg9rmwblcy-bu		
9	Bookkeeping Analyst I	Assist Group Limited -	Faisal,Giza,Egypt	Full Time	On-site	['Entry Level', '-Accounting/Finance', '-Bookkeeping', '-Financial Reporting', '-Bank Reconciliation', 'DataEntry', '-QuickBooks', '-En	https://wuzzuf.net/jobs/p/xjxf6awypnc9-boo		

Main Page – Cleaned Data (After Cleaning)

	A	B	C	D	E	F	G	H
1	Job Title	Company Name	Location	Job Type	work_modes	job_link	Job Level	Experience Needed
2	Project Listing Specialist & Data Analyst for E-Commerce Platforms	armonia vita	Mansoura,Dakahlia,Egypt	Full Time	On-site	https://wuzzuf.net/jobs	Entry Level	1 - 3 Yrs of Exp
3	Business Data Analyst (Finance-Oriented)	Mazzika Group	Dokki,Giza,Egypt	Full Time	On-site	https://wuzzuf.net/jobs	Experienced	2 - 3 Yrs of Exp
4	Data Analyst	Linrco Egypt	Nasr City,Cairo,Egypt	Full Time	On-site	https://wuzzuf.net/jobs	Experienced	2 - 3 Yrs of Exp
5	Commercial Specialist (Junior Data Analyst)	realme	Nasr City,Cairo,Egypt	Full Time	On-site	https://wuzzuf.net/jobs	Entry Level	1+ Yrs of Exp
6	Financial planning and Analyst	Confidential	Cairo,Egypt	Full Time	On-site	https://wuzzuf.net/jobs	Experienced	3 - 7 Yrs of Exp
7	Data Analyst + Instructor (Excel & Power BI)	EpsilonAI	Nasr City,Cairo,Egypt	Full Time	Hybrid	https://wuzzuf.net/jobs	Entry Level	0 - 3 Yrs of Exp
8	Business Analyst	BI-Technologies	New Cairo,Cairo,Egypt	Full Time	Hybrid	https://wuzzuf.net/jobs	Experienced	1 - 4 Yrs of Exp
9	Bookkeeping Analyst I	Assist Group Limited	Faisal,Giza,Egypt	Full Time	On-site	https://wuzzuf.net/jobs	Entry Level	

Cleaning

Remove unnecessary characters

Drop irrelevant columns

Note

The missing Experience Required value in record #9 is expected. This field was not available in the raw data, so it appears empty after string processing.

Data Transformation Results

Job Details– Raw Data (Before Cleaning)

	A	B	C	D	E	F	G
1	Experience_Needed	Career_Level	Education_Level	Salary	Job_Categories	Skills_And_Tools	job_link
2	1 To 3 Years	Entry Level (Junior Level / Fresh Grad)	Bachelor'S Degree	Confidential	['Marketing/PR/Advertising']	['E-Commerce Platform Managemer	https://wuzzuf.net/jobs/p/8creyckjvc8z-project-listing-specialist-data-analyst-for-e
3	2 To 3 Years	Experienced (Non-Manager)	Bachelor'S Degree	Confidential	['Accounting/Finance', 'Business	['Financial Analysis', 'Data Visualizat	https://wuzzuf.net/jobs/p/r377ssty41so-business-data-analyst-finance-oriented-m
4	2 To 3 Years	Experienced (Non-Manager)	Not Specified	Confidential	['Quality']	['Data Analysis', 'performance Analy	https://wuzzuf.net/jobs/p/ailkpyyoxt9a-data-analyst-linrco-egypt-cairo-egypt
5	More Than 1 Year	Entry Level (Junior Level / Fresh Grad)	Bachelor'S Degree	Confidential	['Business Development', 'Analyst	['Statistics', 'Market Research', 'Rese	https://wuzzuf.net/jobs/p/btmdbqyuc8v8-commercial-specialist-junior-data-analys
6	3 To 7 Years	Experienced (Non-Manager)	Bachelor'S Degree	Confidential	['Accounting/Finance', 'Analyst/R	['Accounting', 'Commerce', 'Analyst',	https://wuzzuf.net/jobs/p/taf92lwlbjtm-financial-planning-and-analyst-cairo-egypt
7	0 To 3 Years	Entry Level (Junior Level / Fresh Grad)	Bachelor'S Degree	Confidential, C	['IT/Software Development', 'Anal	['Data Analysis', 'Database SQL Serv	https://wuzzuf.net/jobs/p/diu2s2piv0kr-data-analyst-instructor-excel-power-bi-eps
8	1 To 4 Years	Experienced (Non-Manager)	Bachelor'S Degree	Confidential, Pi	['Installation/Maintenance/Repai	['Business Analysis', 'business', 'Busi	https://wuzzuf.net/jobs/p/n75g9rmwblcy-business-analyst-bi-technologies-cairo-e
9	Not Specified	Entry Level (Junior Level / Fresh Grad)	Not Specified	8000 To 20000	['Accounting/Finance']	['Bookkeeping', 'Financial Reporting',	https://wuzzuf.net/jobs/p/xjxf6awypnc9-bookkeeping-analyst-i-assist-group-limited

Job Details Main Table (After Cleaning)

	A	B	C	D	E
1	Experience_Needed	Career_Level	Education_Level	Salary	job_link
2	1 To 3 Years	Entry Level (Junior Level / Fresh Grad)	Bachelor'S Degree	Confidential	https://wuzzuf.net/jobs/p/8creyckjvc8z-project-listing-specialist-data-analyst-for-e-commerce-platforms-armonia-vita-dakahlia-egypt
3	2 To 3 Years	Experienced (Non-Manager)	Bachelor'S Degree	Confidential	https://wuzzuf.net/jobs/p/r377ssty41so-business-data-analyst-finance-oriented-mazzika-group-giza-egypt
4	2 To 3 Years	Experienced (Non-Manager)	Not Specified	Confidential	https://wuzzuf.net/jobs/p/ailkpyyoxt9a-data-analyst-linrco-egypt-cairo-egypt
5	More Than 1 Year	Entry Level (Junior Level / Fresh Grad)	Bachelor'S Degree	Confidential	https://wuzzuf.net/jobs/p/btmdbqyuc8v8-commercial-specialist-junior-data-analyst-realme-cairo-egypt
6	3 To 7 Years	Experienced (Non-Manager)	Bachelor'S Degree	Confidential	https://wuzzuf.net/jobs/p/taf92lwlbjtm-financial-planning-and-analyst-cairo-egypt
7	0 To 3 Years	Entry Level (Junior Level / Fresh Grad)	Bachelor'S Degree	Confidential, Commission	https://wuzzuf.net/jobs/p/diu2s2piv0kr-data-analyst-instructor-excel-power-bi-epsilonai-cairo-egypt
8	1 To 4 Years	Experienced (Non-Manager)	Bachelor'S Degree	Confidential, Profit Share,	https://wuzzuf.net/jobs/p/n75g9rmwblcy-business-analyst-bi-technologies-cairo-egypt
9	Not Specified	Entry Level (Junior Level / Fresh Grad)	Not Specified	8000 To 20000 EGP Per Mc	https://wuzzuf.net/jobs/p/xjxf6awypnc9-bookkeeping-analyst-i-assist-group-limited-giza-egypt

Job Category – Dimension Table

	A	
1	Job_Categories	job_link
2	Marketing/PR/Advertising	https://wuzzuf.net/jobs/p/8creyckjvc8z-proj
3	Accounting/Finance	https://wuzzuf.net/jobs/p/r377ssty41so-busi
4	Business Development	https://wuzzuf.net/jobs/p/r377ssty41so-busi
5	Analyst/Research	https://wuzzuf.net/jobs/p/r377ssty41so-busi
6	Quality	https://wuzzuf.net/jobs/p/ailkpyyoxt9a-data-
7	Business Development	https://wuzzuf.net/jobs/p/btmdbqyuc8v8-co
8	Analyst/Research	https://wuzzuf.net/jobs/p/btmdbqyuc8v8-co
9	Accounting/Finance	https://wuzzuf.net/jobs/p/taf92lwlbjtm-finan

Skills & Tools Dimension Table

	A	
1	Skills_And_Tools	job_link
2	E-Commerce Platform Management	https://wuzzuf.net/jobs/p/8creyckjvc8z-proje
3	Data Analysis	https://wuzzuf.net/jobs/p/8creyckjvc8z-proje
4	Microsoft Excel	https://wuzzuf.net/jobs/p/8creyckjvc8z-proje
5	sheets google	https://wuzzuf.net/jobs/p/8creyckjvc8z-proje
6	SEO	https://wuzzuf.net/jobs/p/8creyckjvc8z-proje
7	Financial Analysis	https://wuzzuf.net/jobs/p/r377ssty41so-busi
8	Data Visualization	https://wuzzuf.net/jobs/p/r377ssty41so-busi
9	Business Intelligence	https://wuzzuf.net/jobs/p/r377ssty41so-busi

Cleaning & Modeling

Remove unnecessary characters

Drop irrelevant columns

Separate dimension tables for skills and categories

Thanks!

Feel free to check out the project and connect with me

🔗 GitHub
💼 LinkedIn