King Saud University
College of Computer and Information Sciences
Department of Information Technology
IT 362 – Data Science
2st Semester 1447 H

Tech Job Skills Analysis
in Saudi Arabia

Logbook

	Students Name	ID
Section #66849 Group #6	Sara Aljuraybah	445201345
	Haya Alomar	445203016
	Fatimah AlHebshi	445200431
	Dalal alghumlas	445200220

Supervised by: Dr. Reem Alqifari

Phase	Task	Date	Decisions & Rationale	Challenges & Solutions	Implementation
1: Data Collection Research and Assessment	API	4 Feb	We collected job data using the JSearch API instead of web scraping to ensure structured, reliable, and consistent data.	We faced API quota limitations and addressed them by optimizing query parameters and adjusting the number of retrieved pages.	Python (requests + pandas) was used to retrieve and normalize the data. The final dataset includes 144 job records with 31 features.
	Dataset Expansion	12 Feb	We increased the number of queries and pages to improve dataset size and diversity.	API quota limitations required optimizing query parameters.	Additional job roles were added to the search list, resulting in 1,024 records.
	Data Quality Assessment	13 Feb	We reviewed the dataset to ensure it was correctly loaded and properly formatted before proceeding with analysis.	During the review, we identified encoding issues affecting the readability of some text fields. The issue was resolved by adjusting the file encoding to UTF-8.	The dataset was reloaded and verified after updating the encoding format to ensure correct text representation.
	Literature Review	8 Feb	We searched for academic studies related to skill extraction from job advertisements.	Some studies were not relevant, so we refined the search keywords.	Selected relevant papers and summarized their problem, dataset, methods, and findings.