

Rohan Kumar

Computer Scientist

✉ rohanvaswani16@gmail.com ☎ +923362591060 📍 Gulf residency near teen talwar Clifton block 8 Karachi
in <https://www.linkedin.com/in/rohanvaswani16>


Professional Experience

Jr. Data Analyst, Data Check Limited I am working on an SQL Developer(Oracle) database, and also working on ETL/ELT tools. we are running daily cron jobs in Talend. I am also working on apache- airflow to create a scheduler for pipelines.	02/2023 – present Karachi, Pakistan
Jr. Business Intelligence Analyst, Sindh microfinance Bank Ltd I am working on an Oracle SQL developer database and using ETL/ELT tools, creating KPIs for datasets and mining through python.	12/2022 – 01/2023 Karachi, Pakistan
Database Administrator intern, State Bank of Pakistan Learn Validate new MicroStrategy® reports & documents, track/fix bugs via SQL queries on Oracle and MySQL RDBMS Collaborated in designing security solutions and analyzing the sensitive data concern in the company. Provide security solutions regarding data security and how to protect against cyber security	08/2021 – 09/2021 Karachi, Pakistan

Projects

Pipeline Data with Airflow, Consumer Financial Protection Bureau we already have an API available for consumption which has parameters. We have frequent requests from our internal stakeholders to be provided with an up-to-date dashboard for their monitoring purpose. of State, From Date To Date, you can sample fetch data from it. 1) we used this list for Recursion on API to get all states' Data. 2) API Hit URL 3) Scrapping of Data for Washington State for Toda 4) Dumping Data into Google Sheets	
ETL Data warehousing using Pyspark, SCD2 implementation The task involves creating MySQL procedures and scripts to generate the schema and migrate data from an OLTP (Online Transaction Processing) system to an OLAP (Online Analytical Processing) system using a star schema for all the dimensions, including date, and fact. Additionally, the task requires implementing a PySpark Slowly Changing Dimension (SCD) on the employees' dimension using either SparkSQL or Spark Native API. To validate the SCD2 implementation, the task requires demonstrating test cases on the employee dimension, including a scenario where a new record is inserted and another scenario where a record is updated. The SCD before and after the test cases should be shown to ensure that the implementation is correct.	
Display with Data with Streamlit, Financial Consumer Complaints across the state. We were made to take it to the next level and create a Streamlit dashboard to showcase financial consumer complaints across the state. The dashboard will enable users to understand state-wise complaints and actionable items.	
Easy Rentable My Final year project is to build React Native application. We save their costs 80% - 85% which they used for a temporary time, where they can rent out their products, and earn from their products through the application.	
Store Furniture Data Prediction, Data Science We predicted previous data of store using python Jupiter library on Anaconda software.	

Education

Bachelor of Science (B.S) - Computer Science  Shaheed Zulfikar Ali Bhutto Institute of Science and Technology, Karachi	09/2018 – 07/2022 Karachi, Pakistan
---	--

Skills

MySQL / Oracle SQL database / SQL Microsoft Server	Data Science Data visualization, Predictive analyzing, Deep learning, Python	Artificial intelligence Matlab Jupyter Notebook, Deep learning.
Apache-Airflow open-source platform used for programmatically authoring, scheduling, and monitoring workflows or data pipelines.	python Scrapping scrapping data from website and store it in database.	Apache Spark is a fast and general-purpose distributed computing engine used to process and analyze large datasets in parallel across a cluster of computers.
Data Mining RapidMiner(ETL/ELT) , Talend, R, Python.	NoSQL that provide a flexible and scalable way to store and manage unstructured or semi-structured data.	

Certificates

Certified Data Engineering By KarachiDotAI	Introduction to Databases  by Meta	Foundations: Data, Data, Everywhere  by Google
Python for Data Science, AI & Development  by IBM		