

# Khalid Abdelrazk

Junior Data Engineer | Data Analyst

Phone: +20 101 556 5612

Email: khalidabdelrazk4@gmail.com

LinkedIn: linkedin.com/in/khalidabdelrazk

GitHub: github.com/khalidabdelrazek

Location: Cairo, Egypt

## Summary

**Junior Data Engineer & Business Intelligence Analyst** with 1+ year of experience managing and extracting insights from large-scale healthcare and business datasets (**100K–1M+ records**). Delivered 7+ end-to-end analytics projects that optimized operations and reduced delays by up to 40%. Skilled in building **ETL pipelines**, designing executive **Power BI/Tableau dashboards**, and performing root-cause analysis. Currently specializing in **Microsoft Azure cloud infrastructure** and scalable data management to drive data-driven decision-making.

## Skills

- **Data Engineering & Architecture:** Microsoft Azure, ETL Pipelines, Data Modeling, REST APIs, SQL Server.
- **Programming Languages:** SQL (Advanced - Joins, CTEs, Window Functions), Python (Pandas, NumPy), C++, Dart.
- **Data Analysis & BI:** Exploratory Data Analysis (EDA), Statistical Analysis, KPI Tracking, Root Cause Analysis, Power BI (DAX), Tableau, Microsoft Excel (Advanced Formulas, Pivot Tables).
- **Machine Learning:** Customer Segmentation, Predictive Modeling, Feature Engineering, Imbalanced Data Handling (Scikit-learn).
- **Software Development:** Flutter, Mobile App Development, Git/GitHub, Data Structures & Algorithms.
- **Soft Skills:** Stakeholder Communication, Data Storytelling, Cross-Functional Collaboration.

## Experience

- **DEPI — Data Engineering & Data Science Trainee** *Nov 2025 – Present*
  - Selected for the intensive **Microsoft Data Engineer** program, specializing in **cloud architecture** and **SQL-based data warehousing**.
  - Developing end-to-end **Data Pipelines** and **Big Data** processing workflows using **Advanced Python** and **Microsoft Azure Data Engineer Associate** curriculum.
  - Implementing **AI-driven solutions** for data engineering tasks, focusing on automated deployment and scalable data management.
- **Integrated Diagnostics Holdings — Data Scientist Intern** *Aug 2025 – Sep 2025*
  - Processed and analyzed **500K+ healthcare and laboratory records** using **Python** and **SQL**, identifying operational bottlenecks through advanced **EDA** and data cleaning.
  - Engineered **6+ Power BI dashboards** to track real-time **KPIs**, reducing reporting overhead and increasing visibility into laboratory test lifecycle stages.
  - Discovered recurring delay patterns that led to a **15% projected improvement** in turnaround times (TAT) through enhanced monitoring and escalation protocols.
  - Utilized **Scikit-learn** for **Customer Segmentation** and **Predictive Modeling** to forecast demand, improving resource allocation efficiency by **18%**.
- **2B Company — CRM Intern (Siebel)** *Sep 2024 – Oct 2024*
  - managed large-scale customer datasets to **maintain high data quality standards** and system efficiency.
  - Collaborated on **troubleshooting and system testing**, producing technical documentation that improved CRM adoption and reduced user errors.

## Projects

- **IDH Daily Visits Analysis | Excel, Power BI, Python**
  - **Standardized and cleaned 100K+ branch visit records** using Python to **validate and standardize input records** for downstream reporting.
  - Developed **automated BI visualizations** to monitor daily visits, unique visit rates, and branch-level KPIs.
  - **Identified operational performance gaps**, leading to a **~20% increase** in branch efficiency.
- **IDH Test Delay Analysis | Power BI, SQL Server**
  - Investigated 500K+ laboratory records **using SQL Server to identify bottleneck patterns in the testing lifecycle**.
  - Designed dashboards visualizing delays across five key departments, **reducing monthly reporting time by 30 hours and enhancing cross-departmental communication**.
  - Enabled **real-time monitoring protocols, resulting in a ~15% reduction in turnaround time (TAT)**.
- **IDH House Call Prediction | Python, Pandas, Scikit-learn**
  - **Developed predictive models on 1M+ records** to forecast house call demand using advanced machine learning techniques.
  - Executed feature engineering and resampling (SMOTE) to handle highly imbalanced datasets and improve model precision.
  - **Optimized resource allocation**, enhancing forecasting effectiveness by **~18%**.
- **Customer Churn Analysis | Python, Pandas, Power BI (GitHub)**
  - Evaluated **extensive consumer data** to detect behavioral churn patterns and identify high-risk groups.
  - **Conducted customer segmentation** to deliver actionable insights for targeted marketing and retention strategies.
- **HR Employee Attrition Analysis | Python, SQL Server, Power BI (GitHub)**
  - Built a robust ETL and analytics pipeline for 10K+ employee records, automating data ingestion and cleaning.
  - Identified key attrition drivers through statistical analysis, supporting HR in data-driven workforce planning.
- **Sales Analytics Dashboard | Excel (GitHub)**
  - Analyzed 60K+ transactions to extract business insights, discovering that **71.8% of profit is generated during weekdays**.
  - Found that the **50+ age demographic contributes ~40% of total profit**, informing targeted sales strategies.

# Education & Certifications

---

Bachelor of Engineering – Telecommunications Engineering (In Progress)

GPA 3.1

Beni-Suef University, Beni-Suef

Expected Graduation: *June 2026*

## Courses :

1. **Microsoft Data Engineer Associate** *Nov 2025 – Present*  
*DEPI* — Specializing in Azure Cloud Infrastructure, Data Pipelines, and **relational database administration**.
2. **Data Analysis** *Jun 2025 – Dec 2025*  
*Vocat Academy* — Advanced data manipulation and analysis using **Python (Pandas, NumPy, Matplotlib), Power Bi**, Excel and SQL.
3. **Data Analysis Training** *Jul 2025 – Aug 2025*  
*NTI* — Python, SQL, Power BI, Tableau, data cleaning, EDA, visualization, KPI dashboards, and actionable insights.
4. **Machine Learning** *Aug 2025 – Sep 2025*  
*Self-paced* — Covering supervised and unsupervised learning, model evaluation, and real-world applications of machine learning in data-driven decision-making.
5. **Advanced Microsoft Excel for Engineers** *Jun 2024 – Sep 2024*  
*Self-paced* — Focused on formulas, pivot tables, data visualization, conditional formatting, and basic automation for engineering data tasks.
6. **C Programming, Data Structures & Algorithms** *Jan 2024 – May 2024*  
*Online Course* — Strengthened **computational efficiency** and logic skills with arrays, linked lists, trees, and sorting/searching algorithms.
7. **Flutter Mobile App Development** *Dec 2024 – Jun 2025*  
*Route Academy* — Training in building responsive cross-platform apps using Flutter, with hands-on projects and clean architecture patterns.

## Languages

---

- Arabic : Native
- English : Professional Working Proficiency