

# AADITYA SURESH DESHMUKH

## DATA ANALYST

[GitHub](#) | [LinkedIn](#)

Phone: 0493303028

Email: [aadityadeshmukh55@gmail.com](mailto:aadityadeshmukh55@gmail.com)

### SUMMARY

---

As a Data Analyst, I specialize in uncovering insights and solving problems through smart data-driven solutions. At SafeWork NSW, I helped create a powerful business lookup tool with 90% accuracy, showcasing my ability to deliver practical results. I bring expertise in Python, SQL, Spark, and cloud technologies like AWS and GCP to build scalable data pipelines and tackle massive datasets. With a passion for innovation and a knack for turning complex data into clear, actionable insights, I'm ready to make an impact wherever data meets opportunity.

### SKILLS

---

PROGRAMMING: PostgreSQL, Python, R, Snowflake, Spark, Databricks, SAP,

BI TOOLS: Power BI, Tableau, MS-Excel

KNOWLEDGE SETS: Probability and statistics, data mining, data analytics, DBMS, machine learning optimization, Neural Networks, big data Technologies, Text mining, natural language processing, recommendation system, Regression and tree based algorithms. Web scraping.

### WORK EXPERIENCE

---

SAFEWORK - NSW

August 2023 - November 2023 (3 Months)

DATA ANALYST (CONTRACT)

- Our team have developed a business lookup tool for SafeWork monitoring.
- The tool can extract most businesses in any industrial sector, such as manufacturing and real estate in NSW, with more than 90% accurate results.
- It has additional features such as filtering the business by Postal code to filter business of specific area.

### EDUCATION

---

MASTER OF DATA SCIENCE AND INNOVATION

2022-2024

UNIVERSITY OF TECHNOLOGY, SYDNEY

CGPA : 4.49/7

PG DIPLOMA IN DATA SCIENCE

2019 - 2021

IMS PROSCHOOL, MUMBAI

CGPA: 8.0/10

BACHELOR OF MECHANICAL ENGINEERING

2014 – 2018

MIT, AURANGABAD

CGPA 7.64/10

### PROJECTS

---

ELT DATA PIPELINE WITH AIRFLOW

- Tools - GCP Cloud Composer, PostgreSQL, DBeaver, dbt
- Data was loaded into the GCP Postgres raw table using Airflow DAGs. Employing dbt, created a data warehouse with raw, staging, warehouse, and DataMart layers for ad-hoc analysis.

NEW YORK TAXI SERVICE DATA PROCESSING USING SPARK

- Tools - Microsoft Azure, Databricks
- Loaded **730 million records data** from Azure blob storage to DBFS. Used Spark for data transformation, cleaning, analysis, and applied ML models to predict the fare, facing significant computational challenges

### AWARDS AND CERTIFICATIONS

---

- **Best Project Award** (Special Commendation) In MDSI Innovation Labs 2023
- Business Analytics Certificate, Advanced Machine Learning & Artificial intelligence certificate