

JACQUES GROBLER

DATA ENGINEER

Cape Town, Western Cape, 7925

082 550 4745

jacquesgrobler579@gmail.com

SUMMARY STATEMENT

I've been a data engineer since 15th January 2019 and confident with PostgreSQL/SQL, Python, Apache Airflow, Apache Spark and AWS.

PROJECTS

There are three main personal projects that I worked on involving apache-spark, apache airflow and creating a star schema in a Redshift data Warehouse (these projects can be seen in my GitHub repo <https://github.com/JacquesGrobler/udacity-data-engineering>):

AIRFLOW PROJECT

- Used Apache Airflow to Pull data from an AWS s3 bucket and populate fact and dimension tables in AWS Redshift.
- Technologies used include apache airflow, python, PostgreSQL and AWS S3.

SPARK PROJECT

- Used PySpark to pull data from an AWS s3 bucket, create fact and dimension tables and then push it back to an s3 bucket.
- Technologies used include PySpark, Python, AWS S3

DATA WAREHOUSE PROJECT

- Created a python script in this project which connects to AWS Redshift, creates two staging tables, a fact table and dimension tables and then loads the tables with data from an s3 bucket.
- Technologies used include PySpark, Python, AWS S3

WORK EXPERIENCE

JUNIOR DATA ENGINEER AT 2U CAPE TOWN

- Data modelling in our PostgreSQL data warehouse: Created tables and views to be used by other departments such as marketing for analyses.
- Data virtualization in Denodo: Serves as a data centre that connects to many different sources. Data is then replicated from there to our data warehouse.
- Data visualization in Tableau: Created reports that have helped analyse our business decisions such as a cost comparison between two call systems.
- Master data management in Salesforce including ETL jobs involving Microsoft visual studio and SSIS.

FEB 2020 - APRIL 2020

APRIL 2020

MARCH 2020

FEB 2020

15TH JAN 2019 - PRESENT

EDUCATION

UDACITY, ONLINE: DATA ENGINEERING NANODEGREE.

JAN 2020 - JUNE 2020

STELLENBOSCH UNIVERSITY: BACHELOR'S DEGREE IN INDUSTRIAL ENGINEERING.

FEB 2015 - DEC 2018

TECHNICAL SKILLS ACQUIRED

TOP 3 SKILLS:

- PostgreSQL
- Python
- Apache Airflow

DATA STORAGE AND WAREHOUSING:

- AWS Redshift and S3 buckets

DATA WRANGLING AND PARALLEL PROCESSING:

- PySpark and AWS EMR

ETL:

- Microsoft Visual Studio
- SSIS

MASTER DATA MANAGERMENTS:

- Salesforce

DATA VISUALISATION:

- Tableau

DATA VIRTUALIZATION:

- Denodo

ABOUT ME

I have always been a determined and passionate person and have always been willing to push myself. I was a hostel prefect at my high school as well as athletics captain, I was a mentor at my university hostel (Helshoogte, Stellenbosch University) and also a leader at my church, where I would facilitate a small group get together of about 10 to 15 students.

I started my data engineering journey on the 15th of Jan 2019 and am loving it. I have learned a lot at my current company and seek to continue this growth. I am very dedicated and would truly want to be of value to your company.

My LinkedIn profile: <https://www.linkedin.com/in/jacques-grobler/>

THANK YOU