Sanne de Vries | Data Engineer met focus op schaalbare dataplatformen

Keizersgracht 123, Amsterdam, Noord-Holland, Nederland, 1015 CJ
+31 6 12345678 • sanne.devries@example.com • https://sanne-devries.nl
@sannedevries • @sannedevries

Basis Informatie

Data Engineer met 6+ jaar ervaring in het ontwerpen en bouwen van betrouwbare, schaalbare dataplatformen. Sterk in het vertalen van bedrijfsbehoeften naar robuuste ETL/ELT-pipelines en datamodellen.

- Ervaren met cloud-native oplossingen (AWS, Azure) en moderne data stacks zoals Spark, dbt en Airflow
- Focus op datakwaliteit, governance en het leveren van betrouwbare analyses voor BI- en ML-teams
- Teamspeler die graag kennis deelt en junior collega's begeleidt

Opleidingen

Technische Universiteit Delft **https://www.tudelft.nl**

Master, Data Science en Technologie, Score: 8.2 *sep 2015–jul 2017*

- Specialisatie in grootschalige dataverwerking en datamodellering
- Afstudeeronderzoek naar schaalbare streaming-pipelines voor sensordata
- Ervaring opgedaan met Python, Spark, Kafka en cloudplatformen

Cursussen: Gedistribueerde systemen, Databasetechnologie, Machine Learning, Big Data Architectuur, Cloud Computing, Data-ethiek en Privacy

Werkervaring

Rabobank **https://www.rabobank.nl**

Senior Data Engineer *jan 2021–Heden*

- Ontwierp en implementeerde een cloud-native dataplatform op AWS voor de afdeling Risk & Compliance
- Bouwde schaalbare ETL-pipelines met Python, Spark en Airflow, goed voor circa 2 TB dagelijkse data
- Introduceerde dbt voor datamodellering en verbeterde de datakwaliteit met geautomatiseerde tests
- Werkt samen met data scientists en BI-analisten om dataproducten van idee tot productie te brengen
- Bewaakt governance en privacy (AVG) en begeleidt junior data engineers

Trefwoorden: AWS, Spark, Airflow, dbt, Data Governance

KPN **https://www.kpn.com**

Data Engineer *mrt 2018–dec 2020*

- Ontwikkelde batch- en streaming-pipelines met Kafka, Spark en Python voor netwerk- en klantdata
- Migreerde een on-premise datawarehouse naar Azure Data Lake en Synapse
- Automatiseerde datakwaliteitscontroles en monitoring, waardoor incidenten met 30% daalden
- Werkte in een agile team en droeg bij aan architectuurkeuzes en code reviews

Trefwoorden: Kafka, Azure, Python, Streaming, Datakwaliteit

Achmea **https://www.achmea.nl**

Junior Data Analist *sep 2017–feb 2018*

- Bouwde rapportages en dashboards in Power BI voor schade- en klantdata
- Schreef SQL-queries en automatiseerde rapportages met Python
- Ondersteunde data-engineering bij het opschonen en modelleren van brondata

Trefwoorden: SQL, Power BI, Python, Rapportage

Talen

Nederlands: Moedertaal of tweetalige vaardigheid

Engels: Volledige professionele vaardigheid — **Trefwoorden**: Cambridge C1, Werktaal

Duits: Beperkte werkvaardigheid — **Trefwoorden**: Goethe B1, Leesvaardig

Vaardigheden

Data Engineering: Expert — **Trefwoorden**: Python, SQL, Spark, Airflow, dbt

Cloudplatformen: Geavanceerd — **Trefwoorden**: AWS, Azure, Databricks, Snowflake

Datawarehousing: Geavanceerd — **Trefwoorden**: Redshift, Synapse, BigQuery, Kimball

Streaming en messaging: Geavanceerd — **Trefwoorden**: Kafka, Kinesis, Event Hubs

Datagovernance: Intermediair — **Trefwoorden**: Datakwaliteit, Lineage, AVG, Metadata

Programmeren: Geavanceerd — **Trefwoorden**: Python, Scala, Bash, Git

Onderscheidingen

Data Engineer of the Year

Rabobank — *dec 2023*

Erkend voor het ontwerpen van een schaalbaar en kostenefficiënt dataplatform dat meerdere teams binnen Risk & Compliance ondersteunt.

Certificaten

AWS Certified Data Analytics - Specialty — **https://aws.amazon.com/certification/**

Amazon Web Services — *jun 2022*

Microsoft Certified: Azure Data Engineer Associate **https://learn.microsoft.com/en-us/certifications/azure-data-engineer/**

Microsoft — *nov 2020*

Publicaties

Data Magazine NL — **https://www.datamagazine.nl/artikelen/schaalbare-datapipelines**

Schaalbare datapipelines in de cloud — *mrt 2022*

- Praktijkartikel over het ontwerpen van betrouwbare en schaalbare datapipelines
- Gaat in op orchestratie, datakwaliteit en kostenbeheersing in cloudomgevingen
- Gebaseerd op ervaringen met AWS, Spark, Airflow en dbt

Referenties

Dr. Marleen Jansen — **+31 15 278 1234**

Hoogleraar Data Science — *m.jansen@tudelft.nl*

Sanne combineert sterke technische vaardigheden met een scherp oog voor datakwaliteit en governance. Zij is een betrouwbare teamspeler die complexe dataproblemen gestructureerd aanpakt.

Projecten

EnergieDashboard — **https://github.com/sannedevries/energieda**

Open-source dashboard voor het inzichtelijk maken van energieverbruik in Nederlandse wijken — *jan 2023–*

- Verzamelt en verwerkt openbare energie- en weerdata via API's
- Pipeline gebouwd met Python, dbt en DuckDB, georkestreerd met Airflow
- Dashboard in Streamlit voor bewoners en beleidsmakers

Trefwoorden: Open Data, Duurzaamheid, Python, Streamlit

Interesses

Duurzaamheid: Energie, Circulaire economie, Open data

Sport: Hardlopen, Wielrennen, Klimmen

Muziek: Gitaar, Festivals

Vrijwilligerswerk

Stichting Code voor Nederland **https://www.codevoornederland.nl**

Vrijwilliger Data Mentor *sep 2021–Heden*

- Begeleidt jongeren en zij-instromers bij hun eerste stappen in data-analyse en Python
- Organiseert maandelijkse workshops over datakwaliteit en open data
- Droeg bij aan een open-source datapipeline voor het verwerken van gemeentelijke open data