

PERSONUPPGIFTER

sofia.norberg@example.se

+46 70 213653

Örebro

FÄRDIGHETER

Plattformer och programmering

Databricks

Delta Lake

PySpark

DLT

Unity Catalog

Kvalitet och drift

Datamodellering

Testning

CI/CD

Övervakning

Dokumentation

Samarbete

Kravinsamling

Kodgranskning

Incidenthantering

Tvärfunktionellt arbete

UTBILDNING

Örebro universitet

Masterexamen i tillämpad datavetenskap

09/2014 - 06/2018

SPRÅK

Svenska

Engelska

CERTIFIERINGAR

Databricks Certified Data Engineer Professional

Certifieringsorganisation

03/2025

Workshop i datakvalitet: Lakehouse för underhållsplanering

Svenska Data Engineering-nätverket

10/2023

Sofia Norberg

Data Engineer med Databricks

SAMMANFATTNING

Data Engineer med Databricks inom industrin. Förbereder sensor- och arbetsorderdata för förebyggande underhåll och uppföljning av produktionsstopp.

ERFARENHET

Data Engineer med Databricks

Brukspuls Analys

Örebro

01/2022 - Nuvarande

Byggde Delta Live Tables för 340 000 maskinmätningar i 3 fabriker; underhållsledarna fick status 12 minuter tidigare vid skiftstart.

Flyttade 6,5 TB historik till Delta Lake med schemakontroll; omkörning efter källändring sjönk från 7 timmar till 1 timme 16 minuter.

Satte Unity Catalog-behörigheter för 15 analysteam och OT-data; säkerhetsgranskning bekräftade att råa signaler bara nådde 4 godkända roller.

Optimerade PySpark-joiner för 4 data scientists; träningen av reservdelsmodellen gick från 49 till 19 minuter.

Data Engineer med Databricks

Mälardalen Processdata

Örebro

04/2018 - 12/2021

Kopplade CDC från 4 produktionsdatabaser till lakehouse i tio minuters intervall; ihop med ERP-teamet minskade orderfördröjningen från 48 till 11 minuter.

Validerade serienummer och måttenheter på 2,8 miljoner veckomätningar; felaktig koppling mellan serviceorder och utrustning sjönk 39 procent.

Automatiserade avstängning av beräkningskluster efter körning i 8 flöden; oanvänd beräkningstid minskade 28 procent på två månader.

Följde 30 OEE-mått med 7 processtekniker; avvikande rapportvärden kunde spåras till källa före veckomötet.

PROJEKT

Lakehouse för underhållsplanering

Dataprodukt som stödjer Databricks, lakehouse och industri, med förvaltrade källor, versionshanterade transformationer och kontrollerbara kvalitetsregler.

Databricks

Delta Lake

PySpark

DLT