**SUBHADEEP SEN**

DATA ENTHUSIAST | SENIOR CONSULTANT

**EVERYBODY LOVES A NICE STORY.
I TELL MINE WITH DATA!!**

## CONTACT INFO

📞 (+91)-908-353-1573
✉ subhadeep.sen@outlook.in
in www.linkedin.com/me
📍 Kolkata, West Bengal, India

## EDUCATION

**2012-2016**
**Bachelor of Technology**
Heritage Institute Of
Technology, Kolkata

**2005-2012**
**Secondary and Higher
Secondary**
Ramakrishna Mission Vidyapith,
Purulia

## TECHNICAL SKILLS

- PySpark
- Python
- Amazon Redshift
- Databricks
- AWS Athena, AWS Glue
- AWS Lambda, EC2
- Apache Airflow
- Docker
- SQL

## DOMAIN

- Insurance
- TeleCommunication

## PERSONAL DETAILS

Date Of Birth: Nov 28, 1994
Sex: Male
Nationality: Indian

## ABOUT ME

I am a Data enthusiast looking forward to work in a stimulating environment, where I can implement and enhance my skills and knowledge to meet both organizational and individual goals. Team player with strong desires to face new challenges and expand my skills set.

## EXPERIENCE

**OCT 2022 - PRESENT**

**CURRENT ORGANIZATION**
Ernst & Young Global Delivery Service India

**ERNST & YOUNG
GLOBAL DELIVERY
SERVICE INDIA
OCT 2022 - PRESENT**

**CLOUD DATA ENGINEER**
Worked as a Cloud Data Engineer, leading the Transformation Team in migrating the Claims from the legacy claim system into the strategic EIS system. Worked in setting up the AWS Infrastructure for end-to-end data pipeline using AWS s3, Glue, Athena and AWS Managed Services (Apache Airflow). Developed the customized transformation framework in Python, deployed as an AWS Glue Job. Developed Airflow DAGs for overall orchestration of the pipeline. Created Technical documentation in Confluence and developer stories in JIRA.

**TATA CONSULTANCY
SERVICES LIMITED
JAN 2021 - OCT 2022**

**DATA ENGINEER - PYSPARK**
Worked as Data Engineer for migrating existing Microsoft SQL Server database and its dependencies to a centralized repository in AWS. Debugged the existing SSIS packages and re-engineered the code in Pyspark. Developed Spark applications using Pyspark, Spark-SQL for data extraction, for analysizing and transforming data to uncover insights into product analytics and customer usage patterns. Developed Glue Crawler and Jobs for loading the data to corresponding tables in Athena for downstream systems. Experience in creating docker file, building docker images. Responsible for estimating the cluster size, monitoring and troubleshooting of the spark databricks cluster.

**TATA CONSULTANCY
SERVICES LIMITED
DEC 2019 - DEC 2020**

**DATA ENGINEER - PYTHON**
Worked as Python and AWS expert for developing customized python framework to automate back-end data preparation for BI Dashboard. Developed User-Defined functions(UDF) and scripts to interact with AWS s3,uploading data to s3 bucket, unloading data from AWS s3 to Amazon Redshift tables. Developed validation and syncing scripts to auto-sync updated data to BI dashboard and refresh.

**TATA CONSULTANCY
SERVICES LIMITED
DEC 2018 - NOV 2019**

**BACK-END DEVELOPER & ANALYST**
Worked as Amazon Redshift expert to design advanced Redshift Tables, developed optimized queries to manage large volume of data and generate back-end data for Analytical Reports of BI dashboard hosted in ec2 instance.

**TATA CONSULTANCY
SERVICES LIMITED
APR 2017 - NOV 2018**

**SOFTWARE DEVELOPER**
Member of the development Team for developing a Data Ware House for Client using SQL

## PROFESSIONAL SYNOPSIS

Overall 6+ years of IT experience with 4 years of experience as Data Engineer.Proficient in Analyzing User requirement and functional specifications. Well versed in Exploratory Data Analysis, developing Data pipeline and providing data solutions. Proficient in working with multiple AWS Cloud technologies. Team Player with excellent communication skills.

### PY-SPARK
Understanding of Spark framework, big data architectures and developing Data Pipeline.
Strong Experience in data engineering and building ETL pipelines on batch and streaming data using Pyspark, SparkSQL.
Working Experience with large volume of datasets
Good working exposure on Cloud technolgies of AWS - S3, EC2, Lambda, Glue, Athena.

### PYTHON
Experienced with core Python programming, developing customized scripts for automation, developing UDF(user-defined functions), ORMs(Object-Relational Mappers)
Used Python to develop customized framework to interact with AWS s3, Amazon Redshift for preparation of data on cloud.Experience in Python libraries - Numpy, Pandas, boto3.

### AMAZON REDSHIFT
Understanding of Amazon Redshift Architecture and hands on experience in developing advanced Tables and optimizing complex Queries in Amazon Redshift
Used materialized view and optimizing techniques of Redshift for dealing with large volume of data and faster report loading