M.S. Óscar Azeem Becerril Domínguez

Curriculum Vitae

Information

Born	October 22, 1992 (29 years old). Mexico. CDMX	Phone:	(Mex+52) 55-10-78-98-72
LinkedIn	https://www.linkedin.com/in/oscarazeem/	Email:	osdobecerril@esimez.mx

Professional Profile

Data Engineer and Business Intelligence Developer

Experience as Data Engineer and Business Intelligence developer. Building, maintaining and automating ETL processes and pipelines complying with business rules and cases. Employing several technologies such as Informatica PowerCenter (IPC), Exadata (Oracle), Teradata for single node computation, and Hadoop (HDFS), Spark (Pyspark), Scala for distributed workload. Proficient knowledge working with Unix, coding Bash, and KSH scripts, using built-in tools such as AWK and RegEx for data cleaning. Knowledgeable with SQL and PL/SQL.

Work Experience

Current / 07-2021
BBVA, *Data Engineer*,
Activities: Use of the Big Data Datio platform. Schemas creation and definition of Avro and Parquet files for the Hadoop HDFS Data Lake. Use of Spark (Pyspark) for data ingestion and validation. System control version utilizing Bitbucket and Jira for management. A DevOps culture through Jenkins, Jira (Atlassian), Jfrog Artifactory, and Bitbucket integration. Creating, maintaining, and implementing batch JOBs through the BMC Control-M scheduling software.

07-2021 / 06-2018
BBVA - Softtek, *Business Intelligence Developer - ETL*,
Activities: Use of the pmrep and pmcmd embedded commands from Informatica PowerCenter (IPC) inside BASH/KSH scripts to automate workflow execution, extraction, importation, and exportation. Implementing IPC transformations such as Expression, Filter, LookUp, UpdateStrategy, Joiner, Union, SQL, Stored procedure, Normalizer. PL/SQL query translation to IPC mapping/workflows complying with business rules. Retrieval information from multiple sources like Mainframe computers, Data Lake repositories, RDBMS databases (Oracle, Teradata). Creating, maintaining, and implementing batch JOBs through the BMC Control-M scheduling software.

06-2018 / 03-2018
AT&T - Ibérica, *Jr. Data Analyst*,
Activities: Develop machine learning models such as Random Forest, Polynomial Regression, and Support Vector Machines to obtain the 3G and 4G cellular cells' failure probability for each day, using Python Scikit-Learn libraries. Data cleaning applying Python Pandas and Regex for CSV files. Loading of information to SQL Server for reporting and further analysis.

Education

2016-2018 **M.S. Telecommunications Engineering**, *Instituto Politécnico Nacional*, Sección de Estudios de Posgrado e Investigación (SEPI) ESIME, Zacatenco.
Master's Degree. Computer networks and Machine Learning speciality.

2011-2016 **Communications and Electronics Engineering**, *Instituto Politécnico Nacional*, ESIME, Zacatenco.
Bachelor's Degree. Computation speciality with professional license

Certifications

2020-2021 **Scaled Agile Framework SAFe 5**, *Certified SAFe 5 Practitioner*, Scrum practitioner.

Computational Skills and Agile Software Development Methodologies

Languages Python, SQL, PL/SQL, BASH, KSH, Scala

DBMSs Teradata, Oracle, MySQL, Hadoop HDFS

BI Informatica PowerCenter (IPC), BMC Control-M Scheduling, Remedy, Jira.

Platforms Unix, Windows

Tools & Libraries Git, Github, Bitbucket, Jupyter Notebook, ScikitLearn, Numpy, Pandas, Matplotlib, Seaborn, Teradata Studio, Oracle SQL Developer, MySQL Workbench, TexStudio, Office Suite (Excel, Word, Power Point).

DevOps Jenkins, Jfrog Artifactory

Languages

English level C1 (Advanced) endorsed by the EF SET organization.