Karina Baselay

Datos Personales

Dirección : Vicente Lopez, provincia de Buenos Aires.

Nacionalidad : Argentina.

Mail : baselay.karina@gmail.com

Estudios Realizados

Título Universitario: Ingeniera en Sistemas de Información, egresada de la Universidad Tecnológica Nacional (U.T.N).Graduada

Postgrado : Posgrado de Business Intelligence con Introducción al data mining- (U.T.N).Finalizado

Diplomatura: Diplomatura de Machine Learning con Python- (Centro de E-learning -U.T.N).Finalizado

Habilidades

- **Plataformas y Cloud:** Apache Spark (PySpark), Databricks Ecosystem, AWS (EMR, EC2, Redshift, Athena, Glue, S3, Iam, Rastreador, Sagemarker, Cloudwatch),Hadoop Ecosistema,DataLake ,DeltaLake y Datawarehouse on premise y cloud.
- **Metodologías:** Arquitectura Medallion (Bronze/Silver/Gold), CDC, SCD Tipo 1 y 2, Modelado Dimensional (Star Schema y Snowflake schema) basado en metodología kimball.
- **Bi** :microstrategy , Power Bi,ETL informatica Powercenter.
- **Bases de datos Relacionales** :Oracle /teradata, SQL.
- **Modernos Stack Tecnológicos** : Conocimientos de DBT, Snowflake,Databrick.
- **Hands-on experience with Databricks:** Medallion Aquitectura(Bronze, Silver, Gold), Delta Lake & Delta Tables, Ingestion (Full & Incremental), Auto Loader, Structured Streaming, Slowly Changing Dimensions (SCD Type 1 & 2), Change Data Capture (CDC) & MERGE operaciones, Lakeflow Pipelines (Spark Declarative), Orquestacion, Unity Catalog.

Certificaciones

- **Machine Learning con Python Centro Elearning UTN**

Link https://validator.centrodeelearning.com/validator/jlm6yR54YD

- **Spark and Python for BigData with Pyspark (Udemy)**

Link : https://www.udemy.com/certificate/UC-1K16REL0/

- **Python 3 (Udemy)**

Link : https://www.udemy.com/certificate/UC-25U435HE/

- **Aws Analytics –Athena Kinesis Redshift Glue (Udemy)**

Link : https://www.udemy.com/certificate/UC-ZJ6MZL09/

- **Databricks Certified Data Engineer Professional -Preparation**

https://www.udemy.com/certificate/UC-93829f0c-9271-4050-aa70-2b0bd567b61b/?utm_campaign=email&utm_medium=email&utm_source=sendgrid.com

Snowflake Master Class For Data Engineer-AWS -Zero to Expert

https://www.udemy.com/certificate/UC-d70bfee3-0c03-4e78-a297-9732e3289fae/

Idiomas

- Inglés : Upper Intermediate finalizado en International House

Experiencia

Desde Junio 2020 hasta la actualidad.

BBVA.

- **Sr Data Engineer**

Tech Stack:

Versionado :GIT / Bitbucket . Validaciones: Sonar ,Samuel,Artifactory
Herramientas de Integración Continua:Jenkins.
PySpark y Python .**Aws Athenea, Glue, Sagemarker,Cloudwatch.**
Hadoop y **Controlm**
IDE: Intellij / Pycharm
Base datos Relacional:Teradata ,Oracle ,**SQL** .
Jupyter notebooks.
Gobiernos de Datos :Datum .
Metodologías Ágiles/Jira
Certificación **Snowflake y Databricks**
Responsabilidades de mi rol

- Arquitectura de Datos y (Spark**):** Diseñe y desarrollé pipelines ETL/ELT altamente escalables utilizando **Python (PySpark)** para la ingesta y procesamiento distribuido de datos masivos en el Data Lake (AWS S3 / HDFS).
- Migración de Arquitectura Core(on premise a Cloud):Desarrollo y ajustes de pipeline y procesos spark para migrar de **Hadoop** a **Aws**.
- Catálogo**:** Implementé gobernanza de datos mediante **AWS Glue Data Catalog** y consultas con **Amazon Athena.**
- Diseño de Capas e Ingesta**:** Implementé estrategias de ingesta mediante cargas **Full e Incrementales** , aplicando patrones de arquitectura de capas (**Bronze/Silver/Gold**) y optimizando el almacenamiento en formatos eficientes (Avro y **Parquet**).
- Modernización y Validación de Datos (SQL Avanzado): Realicé validación, testing y consistencia de datos aplicando **SQL avanzado** sobre sistemas transaccionales tradicionales (**Teradata, Oracle**) y notebooks **Jupyter .**
- Orquestación y Operaciones de Producción**:** Automatice y schedule jobs via Controlm . Realice monitoreo de procesos usando **Cloudwatch** y resolución de incidentes y mantenimiento correctivo/preventivo de flujos analíticos en entornos productivos críticos.
- Implemente la tokenización de datos para seguridad del modelo.
- Participe en proyecto de Data Regresión (Hdfs oracle y mongo) para analizar transacciones de clientes(tarjetas y transferencias de movimientos) y categorizarlos para luego ser consumidos por app móvil.
- Prácticas DevOps (CI/CD**):** Deploy de pipeline automatizados de **CI/CD** utilizando **Jenkins con Docker, Git y Artifactory**, garantizando despliegues estables y control de versiones.
- Liderazgo Técnico y Negocio**:** Colaboré activamente con Managers y Stakeholders para el refinamiento de requerimientos técnicos, traduciendo lógicas de negocio complejas en arquitectura de datos eficientes.

Desde Noviembre 2018 Junio 2020
Cencosud
Sr Data Engineer BI ETL Data Warehouse/Cloud AWS
Tech Stack :
BI Microstrategy
Big Data
Pyspark , Python
Teradata Bteq and fastload, **SQL**
Unix Bash script
Gitlab Repository
Aws EC2 instances, EMR (Elastic Map Reduce), S3, Rastreador, IAM , Cli , Glue Catalog, Athenea , Redshift.
Descripción de mi rol :

- Migración de Arquitecturas Core (On-Premise a Cloud)**:** Desarrollo y ejecución de la migración del Data Warehouse corporativo desde infraestructuras **Teradata** hacia la nube de **AWS**, garantizando la continuidad operativa del negocio,aplicando patrones de arquitectura de capas (**Bronze/Silver/Gold**)
- Procesamiento Distribuido a Gran Escala (Big Data)**:** Diseñe e implementé pipelines ETL/ELT de alto rendimiento utilizando **Python y Apache Spark** sobre clústers de **Amazon EMR** y procesamiento en instancias EC2, gestionando volúmenes masivos de datos corporativos.
- Catálogo**:** Implementé gobernanza de datos mediante **AWS Glue Data Catalog** y consultas con **Amazon Athena**.
- Ingeniería de Data Warehouse Teradata**:** Responsable del modelado lógico y físico del repositorio central, creación de objetos de base de datos distribuidos y optimización de cargas críticas de datos mediante utilitarios de alta velocidad (**FastLoad, BTEQ**).
- Automatización y Orquestación Enterprise**:** Desarrollé scripts avanzados en **Bash Unix** y configuré flujos de trabajo automatizados a nivel empresarial utilizando **CAWA (CA Workload Automation)** para la sincronización de procesos batch.
- Análisis diseño y desarrollo de pipeline de ingestas de datos con jobs con Python Apache Spark y conocimientos avanzados de Databricks.
- Soporte a Inteligencia de Negocios**:** Mantuve la capa semántica y de reportes corporativos sobre la plataforma **MicroStrategy**.

Desde Junio 2015 hasta Octubre 2018.
IBM.
Senior Data Engineer Bi ETL Powercenter Data warehouse / Big Data.
Tech Stack : **Teradata** ,Oracle, **SQL** , **ETL Powercenter**,**Microstrategy**, Unix, Big Data.
Descripción de mi Rol:

- Modelado y Diseño Dimensional: Diseñe e implemente modelos de datos lógicos y físicos bajo enfoques dimensionales (**Star Schema / Copo de Nieve**), garantizando estructuras óptimas para el análisis de grandes volúmenes de datos históricos.
- Ingeniería de ETL con Informatica PowerCenter**:** desarrolle, optimize y realicé el mantenimiento evolutivo de flujos complejos de extracción, transformación y carga (ETL) utilizando **Informática PowerCenter** con base de datos **Teradata**.
- Adopción de Tecnologías Big Data (Hadoop Eco): Participé en la transición hacia entornos distribuidos, ejecutando la creación, particionamiento y carga eficiente de estructuras analíticas masivas utilizando **Apache Hive y Hadoop**.
- Tuning y Optimización de Performance**:** Realice *performance tuning* en base de datos teradata del Data Warehouse, optimizando scripts SQL, índices y estructuras de almacenamiento para reducir tiempos de procesamiento en cargas batch críticas.Y desarrolle nuevos script en sql para crear nuevos objetos en la base de datos de Teradata.

Desde Septiembre 2014 hasta Mayo 2015.
Directv .
Sr Analista Técnico ETL SAS de Campañas
Tech Stack: UNIX,Shell scripting , **Oracle** 11g ,SAS Enterprise Guide ,SAS Data integration,**SQL** .
Tareas de mi Rol:

- Análisis y Desarrollo de nuevas campañas usando ETL SAS Data Integration, SAS Guide bajo motor de base de Datos Oracle 11g.

Desde Mayo 2013 hasta Septiembre 2014.
Tenaris.
Analista Técnico Data Integración -ODS Industrial (Operational Data Stores).
Tech Stack: UNIX,Shell scripting , Control-M , **Oracle** Pl-sql, Oracle 11g ,Etl DataStage
Tareas:

- Análisis y modelado de datos de Integración entre PSImetals (sistema Industrial) y ODS(Operational Data Stores) Industrial.
- Modelado físico y lógico de datos.
- Desarrollo de ETL.
- Desarrollo de Oracle Pl-sql , complejos queries,procedure, packages,vistas.

Desde Junio 2011 hasta abril 2013.

Oracle.

Sr BI ,Data Warehouse ETL Powercenter, Obi Apps, SQL

Tech Stack: **Oracle** 10g, **Plsql, SQL,** Control-m, DAC,OBIEE, **ETL Informatica Powercenter** 8.6.1, Unix, Shell scrip Unix.

Responsabilidades de mi rol :

- Desarrollo ETL customizado**:** usando oracle bi apps diseño y construcción de flujos de datos ETL customizados para el Siebel Data Warehouse utilizando **Informatica PowerCenter 8.6.1**, integrando múltiples fuentes desde sistemas transaccionales OLTP de siebel y datos provenientes del Datawarehouse de Telcel.
- Optimización de Negocio**:** Adaptación del modelo de datos de Siebel Marketing y su motor de segmentación para la incorporación y análisis de nueva información estratégica.
- Modelado de Datos**:** Modelado físico y lógico de data marts bajo esquemas **Estrella (Star)** y **Copo de Nieve (Snowflake)**.
- Estrategias de Integración**:** Implementación de técnicas avanzadas de tratamiento de datos, incluyendo *Slowly Changing Dimensions* (**SCD Tipo 1 y 2**) y *Change Data Capture* (**CDC**).
- Orquestación y Administración**:** Gestión y ejecución de cargas de trabajo mediante **Oracle BI Datawarehouse Administrator Console (DAC)** v.10.1.3.4.
- Programación Base de Datos**:** Desarrollo y optimización de procedimientos almacenados y lógica de negocio en **Oracle PL/SQL**.

Desde abril del 2010 hasta junio 2011.

Telecom Personal

Analista Programador Sr/Ssr Pl-sql y SQL.

Tech Stack: **Oracle** 10g,Pl-sql,Unix,Shell scripting Unix, Erwin

Tareas:

- Modelización de Base de Datos ORACLE con ERWIN
- Performance y Tuning de consultas Sql en ORACLE 10g.
- Analista programador PL/SQL realizando tareas de análisis, desarrollos pruebas unitarias ,trabajando con packages,funciones y procedimientos de base de datos.
- Unix SHELL SCRIPTING.

Desde Julio del 2004 hasta abril del 2010.

Telecom

ETL Powercenter Datawarehouse developer.

Tech Stack: **Oracle** 10g,Pl-sql,Unix,Shell scripting Unix, Erwin, **Informatica Powercenter**.

Tareas de mi Rol:

- Modelado Dimensional**:** Diseño y construcción de modelos dimensionales y datamarts para Data Warehouse, implementando objetos de base de datos optimizados en **Oracle**.
- Desarrollo ETL Avanzado Y modernización del Datawarehouse: Diseño, desarrollo y mantenimiento evolutivo de flujos de datos (*mappings* y *workflows*) robustos utilizando **Informatica PowerCenter 8.6.1**.
- Ingesta y ETL con Oracle pl-sql**:** Implementación de estrategias de extracción y carga masiva de datos mediante **SQL*Loader** y **Tablas Externas** y desarrollo de ETL con plsql (luego este datawarehouse se modernizó con Powercenter)
- Optimización de Rendimiento**:** *Performance & Tuning* de consultas SQL y código PL/SQL complejo, garantizando eficiencia en entornos **Oracle 10g** con grandes volúmenes de transacciones.
- Automatización y Scripting**:** Creación y mantenimiento de scripts en **Unix Shell Scripting** para la automatización de procesos operativos y de carga.
- Soporte de Producción**:** Cobertura y atención en esquemas de **guardias pasivas** para asegurar la alta disponibilidad y estabilidad de los desarrollos en producción.

Desde agosto del 2003 hasta julio del 2004.

BIS. Consultora de Sistemas Oracle Network Certified Partner.

Oracle Forms Reports 6.i developer ,Pl-sql.

Tech Stack : **Oracle**, Forms,Reports 6.0, PL-sql, Unix.

Tareas:

- Desarrollo de módulo a medida de Sistema de Comercio Exterior basado en la aplicación Oracle Financial .
-

Desde Junio del 2002 hasta Agosto del 2003 .

SMS .Consultora de Sistemas.

Analista Programador developer 6i ,Pl-sql

Tech Stack : Oracle, Forms,Reports 6.0, PL-sql, Unix.

Tareas :

- Desarrollo y Programación de un sistema para telecomunicaciones en herramientas ORACLE DEVELOPER :FORMS 6.0 y Reports 6.0

Desde Octubre del 2000 – Hasta Junio del 2002.

LUA Seguros

Analista Programador Developer 6 ,Pl-sql

Tech Stack : **Oracle**, Forms,Reports 6.0, PL-sql, Unix.

Tareas:

- Desarrollo y Mantenimientos de Sistemas de Seguros.

Desde Octubre de 1999- Hasta Octubre 2000.

Telefónica

Programador Pl-Sql

Tech Stack : **Oracle**, Pl-sql Unix..

Tareas:

- Desarrollo y Programación de procesos de Sincronización de Bases de Datos, realizando TRIGGERS / PACKAGES / STORED PROCEDURE en Base de Datos ORACLE 8i con PL/SQL.

Desde diciembre de 1997- Hasta octubre de 1999.

Consultex Informatica.

Programer ORACLE Pl/sql

Tech Stack : **Oracle**, Pl-sql

Tareas:

- Desarrollo y Mantenimiento de Sistemas varios basados en Planeamiento y control de gestión.
- Desarrollos de funciones,procedimientos y paquetes en ORACLE pl/sql

Desde Octubre de 1996 Hasta Diciembre de 1997

Fravega.

Programador Informix-4gl

Tech Stack : Informix On line 5.0, Informix 4Gl, UNIX.

Tareas :

- Análisis y Programación en INFORMIX – 4gl bajo entorno UNIX para sistema de Recursos humanos módulo sueldos.

Febrero de 1995- Octubre de 1996

ORGANIZACIÓN y GESTIÓN

Programador Informix-4gl Jr.

Tech Stack : Informix On line 5.0, Informix 4Gl, UNIX.

Tareas :

- Programación en INFORMIX – 4gl On Line 5.0 bajo entorno UNIX .