

# Tony Jacobs

Phoenix, Arizona, United States



tjacobs.datascience@gmail.com



[linkedin.com/in/tony-jacobs-696559184](https://www.linkedin.com/in/tony-jacobs-696559184)

## Summary

Certified and accomplished Data Engineer with an extensive background in provisioning and setting up data platform technologies that are on-premises and in the cloud. I'm familiar with managing and securing the flow of structured and unstructured data from multiple sources. Whether it be relational databases, non-relational databases, data streams, or file stores I'm successfully able to ensure that data services securely and seamlessly integrate with other data platform technologies or application services.

I primarily focus on data-related tasks within the top 3 Cloud offerings (AWS, Azure, and GCP). Primary responsibilities include using services and tools to ingest, egress, and transform data from multiple sources. All the while managing, monitoring, and ensuring the security and privacy of data meet industry standards and expectations.

My scope of work goes well beyond looking after a database and the hosted server. Majority of my time is devoted to data wrangling which is the process to get, ingest, transform, validate, and clean up data to meet business requirements. My skill set adds tremendous value to both business intelligence and data science projects. Data wrangling can consume a lot of time. When I wrangle data, projects move more quickly because data scientists and artificial intelligence folks can focus on their own areas of work.

### ACTIVE CERTIFICATIONS:

AWS Cloud Certified - Solutions Architect Associate

Microsoft Certified - Azure Data Engineer Associate

Google Cloud Certified - Associate Cloud Engineer

## Experience



### AWS Data Engineer

Carestarter

Jan 2021 - Jul 2021 (7 months)

- Assisted in support case resolution to increase SES email limit from 200/per hour to 25000.
- Generated weekly and monthly Web-App Adoption metrics for management and outside parties.
- Multiple POCs/Dev projects consisting of a mixture of the following AWS Services:
  - 1) Cloudformation - Infrastructure as Code
  - 2) Analytics (Athena, QuickSight)
  - 3) Application Integration (API Gateway, App Sync, SNS, SQS, AppFlow, Step Functions)
  - 4) Compute (EC2, ECR, Fargate, Lambda)
  - 5) Database (Aurora, RDS SQL Server/Postgres, DynamoDB)
  - 6) DevTools (Cloud9, CloudShell, CodeCommit, CLI)
  - 7) Data Pipeline - Sequence, schedule, run, and manage recurring data processing workloads



### Azure & GCP Data Engineer

Y&L Consulting, Inc.

Dec 2019 - Jan 2021 (1 year 2 months)

- Azure technical architect and hands-on data engineer for multiple POC cloud engagements to migrate legacy systems to the cloud and data lakes for data science teams, reporting, and analytics.
- Utilized Azure Data Factory to integrate external data to a data lake as well as to import and export data from Azure blob storage, AWS S3 buckets, and parquet files.
- Ingestion, Data Profiling, Data Cleansing, Transformations, Job Scheduling/Execution, and Dimension & Fact table creation using Azure Databricks, Python, and SQL based on in-memory architecture methodology.
- Implemented micro services, data lake ingestion of data and BI functions using the power U-SQL scripting language provided by Azure Data Lake Analytics and Python.
- Created real-time streaming Operational Data Store environment with Azure event hubs and stream analytics that directly streamed into MS Power BI for corporate reporting.
- Spun up HDInsight clusters and used Hadoop ecosystem tools like Kafka, Spark and Databricks for real-time analytics streaming, sqoop, pig, hive and CosmosDB for batch jobs.
- Various AWS, Azure, and GCP Cloud proof-of-concept work.

## **Big Data Oil & Gas Consultant**

Enverus

Jun 2019 - Dec 2019 (7 months)

Provide big data advice, recommendations, proto-types, and database migration assistance for high-value customer data feeds on the cloud. I'm familiar with various big data architectures, etl methodologies, and various devops practices in support of activities within AWS, Azure, and GCP cloud platforms.

## **Senior Data Engineer**

Drillinginfo

Aug 2014 - Jun 2019 (4 years 11 months)

- Requirements gathering and analysis for new project development.
- Architecture: Conceptual, Logical, and Physical Data modeling.
- Data Preparation - Data Profiling & Data Cleansing.
- Created complicated T-SQL to clean, convert, transfer data between systems.
- Heavy datawarehouse development to support ODS, DW Facts & Dimensions.
- Design, develop, deploy, maintain, monitor, and tune ETL solutions for both incremental/full loads.
- Schedule jobs and monitoring to execute and track daily, weekly, and monthly ETL/ELT loads.
- Identify and troubleshoot business processes, data feeds, and data quality issues.
- Amazon Web Services proof-of-concept work.

## **BI Stack (SSIS,SSAS,SSRS) Datawarehouse Engineer**

EZCORP

Jun 2013 - Aug 2014 (1 year 3 months)

- Gather and analyze business requirements and translate business needs into long-term business intelligence solutions leveraging the Microsoft technology BI Stack (SSIS, SSAS, and SSRS).
- Master Data Services support (MDS) and Data Quality Services (DQS).
- Multi-Dimensional Cube Development.
- Tabular Model Cube Development.
- Report design and generation.

- Developed SSAS Cubes (Multi-Dimensional/Tabular) as well as associated reports with appropriate dimensions and measures, configured attributes of dimension table with hierarchy relationships, deployed KPIs on SSRS and different metrics, added new measure groups and dimensions when necessary.
- Create PowerPivot models and PowerPivot reports, publish reports onto SharePoint, and deploy models into SQL Server SSAS instances.
- Create parameterized SSRS reports and PowerPivot workbooks using data from Tabular model and Multi-Dimensional models to prove the functionality of databases.
- Power BI Development to support a self-service BI solution based on Excel.
- Provide technical assistance to the team on the various Microsoft BI Stack technologies.
- Tailor solutions for Management dashboards for financial reporting and operational business intelligence.
- Create reports in Power BI preview portal utilizing the SSAS Tabular via Analysis connector.
- Configure and optimize BI environment, which includes relational, OLAP, and data migration components.
- Troubleshoot BI Environment problems and tune for performance.
- Diagnose and resolve application bottlenecks and performance issues.
- Define and enforce coding standards, define styles and best practices.



## **Sr. Data Architect/Analyst**

Amherst Holdings

Jan 2013 - Jun 2013 (6 months)

- Project Plan creation.
- Data and Requirements gathering, analysis, & review (Requirements document).
- Data Dictionary design and creation.
- Data Modeling (Conceptual, Logical, and Physical Models based on requirements).
- ETL Design (data profiling, extraction, transformation, & loading).
- Fact and Dimension table creation for consumption by Tableau, SSRS, & LogiAnalytics.
- Matching Algorithm creation for binding multiple datasets for data mining.
- Predictive Analytics for positional bond trading (training set, test set).



## **Information Management Specialist (Consultant, Initiate MDM- Master Data)**

IBM

Jan 2012 - Jan 2013 (1 year 1 month)

- MDM Implementations | Clients: Adventist Healthcare | Advocate Healthcare | Choice Hotels.
- Customer requirements gathering, analysis, & review.
- Configure Initiate MDM member data model.
- Configure Algorithms (Standardize data, bucketing, & comparison functions).
- Clean data extracts (data profiling, extraction, transformation, & loading).
- Data derivation, weight generation, & bulk cross matching.
- Instance/Hub/Client Tools creation and deployment.



## **BI Architect (SSIS, SSAS, and SSRS)**

AMD

Aug 2011 - Jan 2012 (6 months)

- ASIA Sales/Inventory/ASP SSAS Cube Development.
- Requirements document analysis resulting in re-defining scope and deliverables.

- Identification of facts and dims with associated source-to-mapping document deliverable.
- Conceptual, logical, and physical data model deliverables.
- Cube specifications document deliverable.
- ETL (SSIS) Package deliverable.
- SSAS Cube creation.

## **QlikView Consultant/Developer**

### **NCSOFT**

Apr 2011 - Aug 2011 (5 months)

- Migration of Cognos Reports to QlikView Reports. (Contract)
- Analysis Phase – Identified and discussed QlikView system functional & technical requirements, proposed possible solutions with positive/negatives and provided a recommended solution, specified the final agreed upon detailed solution requirements.
- Planning Phase – Established the QlikView project management elements of the project, including communications, roles and responsibilities, initial timeline, processes and procedures
- Design Phase – Conceptual, logical and physical data model development. Developed a 3-Tier Mixed (qvds & QlikMarts) blueprint (design document) for the construction of the QlikView solution (QVD Architecture, QVD Reload/Load Methodology, Publisher, and Server implementation & performance), including critical elements such as data sources and integration, application workflow, infrastructure, based on requirements.
- Checkpoint Phase – Assessed the design requirements against resources, timeline, and estimate to completion, making modifications when appropriate.
- Build Phase – Developed and integrated QlikView solution components to meet design specifications via an iterative development process.
- Test Phase – Testing of the QlikView system across multiple levels (unit, system, functional, user acceptance, performance, security), refined the solution to ensure acceptance.
- Deploy Phase – Preparation of finished QlikView solution, including design automation, developed required documentation (user guide, admin guide, user training.
- Review Phase – Project team discussions to gather lessons learned and assess project performance in regard to data accuracy & application performance.



## **Sr. Database Engineer**

### **HomeAway**

Jan 2009 - Apr 2011 (2 years 4 months)

- Requirements Gathering and Data Modeling.
- MS BI Stack (SSIS, SSAS, SSRS).
- Worked with business users to define and analyze problems.
- Created conceptual, logical, and implemented physical data models.
- Develop and Manage Reporting Solutions.
- Implement SSRS schedules, subscriptions, and security using Report Manager.
- Report creation using Report Designer.
- Dataset creation (w/ filters, groups, sort, aggregates, parameters).
- Tableau Reporting Solutions.
- Tableau Server Product and Capacity Key activation.
- Tableau Licensing Task automation.
- WMS for Integration with Tableau deployment.

- Tableau Product Key management.



## **DBA (Admin, Development, BI) / Datawarehouse Architect / .Net Programmer**

Dell Technologies

Jul 2004 - Jan 2009 (4 years 7 months)

- Database Object Creation and Implementation.
- Conceptual, logical, and physical data model creation for data marts/data warehouse.
- Implementation of tables, views, triggers, functions, stored procedures, constraints, indexes, full-text searches, as well as the creation of user-defined types to support database activities.
- SQL Server Performance Monitoring and Troubleshooting.
- Gathering of performance and optimization data via SQL Server Profiler, Database Engine Tuning Advisor, and DMVs.
- Monitoring and resolution of blocks and deadlocks.
- Diagnosis and resolution of database server errors.
- Monitoring of SQL Server Agent job history to identify failures, outcome details, and job activities.

## **Education**



### **The University of Texas at San Antonio**

Bachelor's degree, Information Systems

2001 - 2005

## **Licenses & Certifications**



### **AWS Cloud Certified - Solutions Architect Associate - Amazon Web Services (AWS)**

Issued Jun 2017 - Expires Jul 2020

JKNMJC91L1BQ1HGX



### **Qlikview Designer I/II - Qlik**



### **Tableau Certified Consultant - Tableau Software**



### **IBM Initiate MDM - Certified Professional - IBM**



### **MCP - SQL Server Database Design - Microsoft**



### **1. Google Cloud Platform Big Data and Machine Learning Fundamentals -**

Coursera

J5A52XSVDV7W



### **2. Leveraging Unstructured Data with Cloud Dataproc on Google Cloud Platform**

- Coursera

USVU6Y3W5H77



**3. Serverless Data Analysis with Google BigQuery and Cloud Dataflow -**

Coursera

VEYMD767FDAC



**4. Serverless Machine Learning with Tensorflow on Google Cloud Platform -**

Coursera

ZNNM8NHX2GC



**5. Building Resilient Streaming Systems on Google Cloud Platform - Coursera**

3QKFYL27JK9Z



**Microsoft Certified: Azure Data Engineer Associate - Microsoft**

Issued May 2020 - Expires May 2022

H422-8280



**Google Cloud Certified - Associate Cloud Engineer - Google**

Issued Nov 2019 - Expires Nov 2021

CSQnzR



**Google Cloud Specialization: Data Engineering, Big Data, and Machine Learning on GCP Specialization - Google**

SEPVHDP5B3C4

## **Skills**

Data Engineering • SQL • Databases • Oil & Gas • Microsoft SQL Server • Data Architecture • Data Science • Business Intelligence (BI) • Data Analysis • Data Modeling