# Zac Yang
## Data Engineer Software Engineer

📞 6264935215
✉ zacyang0701@gmail.com
🌐 https://www.linkedin.com/in/zac-yang/

## Personal Profile

Data/Software Engineer with a demonstrated experience on Big data solutions including databases (SQL/No-SQL) and ETL/Data pipelines using cloud services. Solid understanding of fundamental Computer Science concepts. Familiar with requirement gathering and SDLC. Great programming experience in multi-language such as Python, Java, SQL, PowerShell. Excellent problem-solving skills and software engineering habits, including peer code reviews, unit testing, etc.

## Core Skills

- OLTP/OLAP/ETL/ELT
- Azure ADF/ADL/Scope script
- AWS S3/Glue/Lambda/Athena

- Data Modeling/Data Analyzing
- Git/Docker
- Java/Python

- TSQL/PLSQL/MySQL
- PowerBI/Tableau
- Agile/Scrum/SDLC

## Career Summary

**2021 − present**  **Data Engineer**
                    **Capital Group**

*Role outline*

Serve in the Investment Group. Implement ETL logic to meet business requirements from stakeholders. Technologies include but are not limited to AWS S3, Lambda, Glue, Athena, Postgres SQL and CloudWatch.

*Achievements*

- Generated Glue jobs by analyzing data from different sources. Such as Parquet,XML,Json, CSV, CloudWatch logs and third-party API.

- Utilized Glue jobs and Lambda functions to handle data processing logics with different kinds of triggers such as SQS and Glue catalog events.

- Automated ETL process with Glue jobs by using AWS CDK which extract CloudTrail log into parquet files and dump them into Athena DB.

- Initialed and managed an AWS glue pipeline for processing sensitive data. Converted IRR, MoM,FundReturnRatio and other Excel functions in Pyspark. More than 20 Lambda / Glue jobs were included for data staging, ELT processing, and result validation. Trigger the pipeline via SQS whenever the dependency datasets are ready.

- Persistent the ETL results into Postgre DB via AWS Glue jobs. Partitioned and indexed monthly coming dataset with pre-defined store procedures to improve performance.

- Rendered Postgre data into Tableau visuals into different workspaces, and implemented Role-level-Security to limit different data access requests.

**2019 − 2021**  **Software Developer**
                 **Microsoft**

*Role outline*

Work as a software engineer under Office IP Antispam team. Response for maintaining and upgrading the ETL/ELT data pipeline, Lens/ADF jobs, and stream dataset in Cosmos and ADLS with Scope script and Kusto queries. As well as generate Ad-hoc visuals with varieous tools based on different requirements.

*Achievements*

- Migrated Cosmos uploader/downloader with ScopeSDK, T-SQL, and PowerShell from old JScript logic. Equipped with alert and pipeline monitor to enhance durability.

- Troubleshoot issues, errors in ETL services. Generated probes, monitors, alerts for different components and services with C#,Scope language, and PowerShell.

- Built and maintained a big data platform (ADF/ADLS) in Cosmos Scope script as per industry standards and best practices.

- Performed DB admin activities on the multiple DB Servers. Optimized data pipeline by rewriting new logic and SQL schema, gained both performance and service durability. Reduced up to 20% storage cost with archiving and re-indexing stale data as well as increase the query performance.

- Rescued customers' sensitive data from old DB backup files online. Praised by the managers.

- Optimized slow running scope queries by using scope hints. Generate 10+TB level data daily bases.

Implemented Ad-hoc data status monitors and dashboards with PowerBI and Lens platform. Delieverd reports to end users weekly and monthly.

| 2018 − 2019 | **Data Engineer** |
|---|---|
| | **Conagra Brands Inc.** |

*Role outline*

Serve under Cost to Serve team. Generated annual reports from the sketch. Implemented data modeling, user validation and dynamic filters based on stakeholders' requirements.

*Achievements*

- Designed and implemented the whole ETL (Extract, Transform, Load) data process with various transactions within SSIS packages to implement business rules and logic.

- Generated PowerBI reports, dashboards with Star schema. Implemented complex requirements in PowerBI by using DAX and M language.

- Reduced refresh waiting time by using PowerBI API in PowerShell scripts, deployed, automated the scripts on CtrlM Servers.

- Created customized visuals with R libraries. Embedded the PowerBI visuals within Sales-force homepage with PowerBI-JavaScriptlibraries.

- Accumulated agile Product Management experience - a full gamut of creating stories, generated PowerBI reports and dashboards, working through user acceptance testing (UAT).

- Hold scrum meetings as a scrum master. Wrote various documents such as data mapping document, work detail descriptions for KT sessions, and future development.

| 2013 − 2015 | **SQL BI developer** |
|---|---|
| | **KSDHC Inc.** |

*Role outline*

As a SQL BI engineer, utilize the MS BI stack (SSIS/SSRS/SSAS) to perform ETL processes with a big scale of data. Includes data cleansing, data labeling and aggregation, etc. Generate different visuals with Microsoft SSRS.

*Achievements*

- Created logical and physical data models based on the requirements utilizing Erwin and NaviCat.

- Improved the performance of the official website by optimizing slowly running queries, utilized indexes, partitions, and deadlock monitor via SQL Profiler and DTA.

- Normalized database tables to avoid redundancy and DML anomalies.

- Developed complex SQL scripts and procedures for data profiling and auditing purposes. Combined C# logic in SSIS packages.

- Utilized Microsoft SSIS/SSAS/SSRS to handle ETL logic, create ad-hoc reports with both tabular model and cube model.

- Developed various Dashboards and Stories using advanced Tableau features including calculated fields, parameters, table calculations, row-level security, R integration, and dashboard actions while dealing with Cloud data in SSRS.

## Education

| | |
|---|---|
| **2015 - 2018** | **Oregon State University** (Master degree in computer science) |
| **2010 - 2014** | **Liaoning Normal University** (Bachelor degree in computer science and technology) |

## Certificates

| | |
|---|---|
| **Microsoft:** | **Azure Database Administrator Associate** (Since 2019) |
| **Microsoft:** | **Data Analyst Associate** (Since 2019) |