# Eric Jalbert

Data Engineer
Markham, Ontario, Canada

Standing at the intersection of data engineering and analytics, I use my skills to resolve problems with a solution that is tailored for each situation. I am adept at getting a minimal product shipped or optimizing to perfection.

- ✉ ericmjalbert@gmail.com
- 🌐 ericmjalbert.com
- ⓞ ericmjalbert
- in ericmjalbert

## ⚙ Skills

### Data Analytics

- Used Bokeh/seaborn with Jupyter notebooks to create analytical stories for providing insights.
- Avid follower of the principles of reproducible research.
- Educated others on theory and application of hypothesis tests for different A/B testing scenarios.
- Conducted multiple-correlation analysis and regressions using R and Python.
- Simulated A/A tests results to compare Permutation tests, T-test, Mann-Whitney-Wilcoxon Test for performance with non-normal and normally distributed data.
- Presented deep analysis using scikit-learn to find hidden relations.

### Engineering

- Experience using a comprehensive techstack to perform ETL and data warehousing tasks like Kafka Connect, EC2, S3, Lambda, Kinesis, Redshift.
- Wrote scripts to automate manual tasks; used Shell

## 💼 Experience

### Data Engineer • Ada Support
June 2019 - April 2020

- Optimized ETL and Analytics Dashboarding layer to maintain a 99.9% data processing SLO.
- Created streaming ETL to process over 1k/records per second.
- Separated resource consumptions for specific services to reduce situational CPU peaks.
- Setup and managed Looker to enable data driven decisions.
- Grew into leading a team of 3 Data Engineers.

### Data Engineer • RubyLife
May 2018 - June 2019

- Architectured and implemented a scalable ETL that processed over 10GB of user data per day.
- Developed interactive web app to allow for efficient A/B test analysis.
- Fostered a culture around live documentation to improve communications between teams.
- Created payload validation system to reduce QA effort.
- Setup and optimized Spark server for Data Scientist usage.

### Sr. Data Analyst • theScore, Inc.
Dec 2017 - May 2018

- Increased analysis confidence and reproducibility by developing a custom postgreSQL integration testing framework inheriting functionality from Python unittest
- Simplified user-based analytics by creating a framework to map IDs across multiple sources
- Reduced Analyst workload by automating the generation of executive-level reports and A/B test documentations
- Used machine learning to remove seasonality from KPIs
- Enabled deep understanding of user reviews by writing an ETL to get reviews from app store and augment them with sentiment analysis and TF-IDF data.

### Data Engineer • theScore, Inc.
Jan 2017 - Dec 2017

- Simplified ETL scheduling by replacing complex CRON tasks with Apache Airflow.
- Reduced inefficiencies in daily ETL runtime by building a system to maintain data consistency between different databases.
- Improved user retention for theScore app by writing a team recommendation model based on GIS data.
- Optimized AWS Redshift performance by automating table VACUUM operations and by managing disk usage on slices.
- Created article categorization model using random forest.
- Coordinated analytics tracking framework transition.
- Practiced DRY principals by teaching colleagues how to use git and create modular work.
- Modernized tech stack by porting Python 2.7 to 3.6.