

# Alon Honig

New York, New York

Senior Data Engineer at JUUL Labs

---

## Summary

- ▶ Data specialist with experience in application development, systems migration, database maintenance, and code refactoring
  - ▶ Capabilities in all stages of the Data Life Cycle (DLC):  
Generation | Warehousing | Distribution | Archiving | Purging
  - ▶ Proven facilitator between Business and Engineering. Managed full-stack development of software.
  - ▶ Functional in multiple environments including cloud, on-prem and data center. Written, committed and deployed code in Python, SQL, C++, Shell, Spark, Groovy, and Javascript. Experienced in relational databases, No-SQL, and Graph: Postgres, MySQL, MongoDB, Hive, SQLite, Kafka, Gemfire, Solr, Neo4j, and Redis. DevOps experience in git, Jira, Jenkins and Make.
- 

## Experience

Senior Data Engineer at JUUL Labs

August 2018 - Present (1 year 2 months)

Data Engineer at VOLANT TRADING

March 2017 - August 2018 (1 year 6 months)

- ◆ Designed, developed, and supported python file downloading application for position reconciliation, fee accounting, and product listing. Application manages 30+ FTP, SFTP, HTTP, SCP data feeds and initiates, reconciles, and validates 250+ file downloads on a daily basis for 3 business units. Created Front-end credential management using PostgreSQL, bootstrap, jquery, and flask REST API. Log oriented - multiprocessing architecture enabled download tracking, analytics, and self-stabilization.
- ◆ Engineered next generation System Of Record (SOR) that ingests data from CME, OCC, CFE, ICE, ISO, and Bloomberg in python. Architected data model for instruments, exchanges, companies, terms, equities, options, futures and currencies in SQLAlchemy Object Relational Mapping (ORM). Automated scripting constructs 69 source tables and 11 derived tables in 5GB PostgreSQL database. Object taxonomy of source, listing, type, and asset class provided by 10 Enum categories. Continuous Extract, Load, Transform (ELT) Lambda architecture ensured state recovery in case of failure. Sqlite snapshot functionality enabled rapid testing for feeder

applications. Supported updates include corporate actions, contract adjustments, delistings, and expiry.

- ◆ Refactored C++ System of Record (SOR) interface to read from Sqlite. Created mapping between Sqlite data model and Trading RPC API.
- ◆ Managed legacy MySQL System Of Record (SOR) for 1+ million financial instruments. Added 15 new financial products including intercommodity spreads and interest rate futures using Groovy.
- ◆ Developed service that monitors disk utilization across 25+ colocated servers in USA and Asia in python. Search indexing built for files, directories, users, permissions, datetime, and size in PostgreSQL. Analytics dashboard built in Javascript, Flask, Bootstrap, JQuery, and SBAdmin2.

### Senior Consultant, Data Science at Capco

April 2015 - March 2017 (2 years)

- ◆ Developed system management software for customer search application in python. Built functionality for create, read, update, delete (CRUD) using Flask REST API. Supports rebuild, backup, batch-load and search operations on Solr database. Executed Testing framework using Insomnia REST client.
- ◆ Researched and integrated Parts-Of-Speech (POS) booster for customer banking FAQ search. Created framework for meta-parameter optimization using Rank-Biased Overlap (RBO) in Solr with python NLTK on 2GB Portuguese Corpus.
- ◆ Architected, coded, and tested customer rewards optimization framework using Customer Lifetime Value (CLV), Customer Acquisition Cost (CAC), and Google Survey in Spark. Built API for KPI reporting engine in Flask. Metrics included Conversion Rate (CVR), Cost Per Conversion (CPC), and Return on Investment (ROI).
- ◆ Developed Extract-Transform-Load (ETL) process for 8+ million SEC filing repository in python and MySQL. Established hierarchical classification scheme of 12 groups for 417 filing types. Extracted 27 corporation and individual entity identifiers from filings and stored them in 5+ million MongoDB database.
- ◆ Created data repository of 18+ million prices for 9,000+ stocks listed in 15+ exchanges using MySQL and R.
- ◆ Architected banking data model for customers, employees, corporates and counterparties containing over 500 fields, 7 payment systems, and 18 tables. Generated 60+ million records in test data.
- ◆ Created full-cycle project plan for resourcing, developing, integrating, and hand-off of Big Data customer segmentation application. Addressed multiple project factors including staffing, timelines, deliverables, dependencies, architecture, risk, and stakeholder management. Designed, modeled, and implemented 500 TB+ customer segmentation model with python, java, Shell, Hive, and Hadoop.

### Analyst, Data Scientist at AIG

October 2014 - April 2015 (7 months)

- ♦ Engaged in risk management consulting for property and casualty clients with annual premiums of \$1-\$90 million by synthesizing narratives, identifying cost drivers, and validating results. Regions studied include US, Asia, Latin America, Europe in industries such as Defence, Retail and Logistics. Helped manage output of 10 person global research team.
- ♦ Synthesized data dictionary for workers compensation database containing 24 tables, 265 columns and over 400 million records using R and SQL. Researched data lineage, integrity, and taxonomy in 1,000+ table data warehouse.
- ♦ Composed strategy note for senior management on developing human resource analytics. Compiled personal experience, industry reports, and subject matter expert opinion to identify risks and highlight opportunities associated with current strategy.

### Senior Analyst at Aon Hewitt

May 2011 - September 2014 (3 years 5 months)

- ♦ Executed analytics consulting engagements for corporate clients with revenues of \$50 million – \$20 billion by reducing costs and managing risks associated with providing employee benefits such as healthcare, long-term disability, and term life insurance.
- ♦ Completed benchmarking analyses on client expense levels and identified cost saving opportunities. Created, managed, and maintained repository for fees paid by clients in Excel and Access. Utilized regression model to find cost drivers in R and SQL. Drafted, tested, and deployed benchmarking algorithm in Excel that reduced time spent by over 80%. Ran training sessions for 40+ colleagues on use of tool.
- ♦ Developed Monte Carlo simulation in R and C++ that maps benefit risk to \$50+ million insurance contracts. Lead clients to use predictive modeling in risk management framework. GUI enabled optimization of insurance contracts for client data.
- ♦ Led redesign of Excel based employment benefit underwriting system by focusing on user experience and complexity reduction. Engineered backward compatibility with two legacy systems. Modularized workflow by eliminating duplication, automating data mapping, and regularizing data attributes. Enhanced functionality includes date customization, automated reconciliation and improved dashboarding.

---

## Education

University of Toronto

MA, Economics, 2010 - 2011

Western University

BA, Financial Economics, 2006 - 2010

---

# Alon Honig

New York, New York

Senior Data Engineer at JUUL Labs

---

**Linked**  **Recruiter**