

# Justin Cooke / Data Engineer

## Contact

e: jstncooke@gmail.com   t: (613) 869-4251   b: jstnck.github.io   gh: github.com/jstnck

## Work Experience

### KAR Global / Data Engineer

JANUARY 2022 – NOVEMBER 2022 | TORONTO, ON

- **Datawarehouse Migration:** Successfully completed a full migration from Redshift to Snowflake, including all database objects, ETL processes and backups
- **Data Pipelines:** Lead data engineer on implementing dbt for in-database transformations, ETLs developed using Airflow, Kubernetes, Python.
- **Unit/integration Testing and Deployment:** Automated the testing process for new ETLs with pytest and tox, Sonarcloud code coverage, CI/CD using Azure DevOps.
- **Analytics Engineering:** Fully onboarded Analytics team to dbt, including developing data models, data quality testing with dbt, automated new model deployment by integrating dbt and Airflow.
- **AWS Cloud Solutions:** Managed IAM roles for Data Engineering processes, secrets management, and Kubernetes (EKS) configuration for data pipelines.

### Beam Data / Data Engineering Intern

MARCH 2021 – SEPTEMBER 2022 | TORONTO, ON

- **Internal Client:** Built an internal tool to automate the businesses surveys. Used EC2 and Python / Flask to connect the SurveyMonkey API and the Slack API. Results were stored in Postgres and automated dashboard creation using ReDash.
- **Start-up Client:** Productionized a machine learning recommender engine using Docker and Airflow, moving from nightly batch processing to on-demand recommendations in under 5 minutes. Automated external data collection via news and social media APIs and web scraping. Introduced version control and container registry via GitLab, infrastructure as Code using CloudFormation.

### Ontario Government, Min. of Children, Community & Social Services / Sr. Research Associate

AUGUST 2016 – JANUARY 2021 | TORONTO, ON

- **Project Management:** Lead Researcher and project manager on multi-year, large budget dataset integration project using multiple federal and provincial datasets, with complex privacy and governance requirements.
- **Analytics:** Build Web Portal and Dashboards for Ontario immigration data. Presented results of data analytics conference at Metropolis Immigration conferences. Data cleaning and analysis for policy, program development.
- **Privacy Impact Assessments:** Managed the development of a Privacy Impact Assessments governing a PII data pipeline, assessed by the Ontario Privacy Commission.

### The Conference Board of Canada / Economist

JUNE 2011 – MARCH 2015 | OTTAWA, ON

- Lead economist covering Canadian Territorial economies, conducted interviews with both local and national media.
- Designed, estimated, and maintained econometric and fiscal forecasting models.
- Created customized economic impact and policy studies.

## Key Skills

### PROGRAMMING

Python  
SQL  
Scala  
R  
JavaScript  
Bash

### DATA STORAGE

Postgres  
Elasticsearch  
Snowflake  
BigQuery  
Redis

### DATA PROCESSING

Spark  
Dask  
Apache Beam  
Kafka  
dbt

### CLOUD / DEVOPS

AWS / GCP  
Airflow  
Git  
GitLab  
Docker  
Terraform  
Kubernetes  
Argo

### DATA VISUALIZATION

Plotly / Dash  
Streamlit  
R Shiny  
PowerBI  
Superset

## Education

### Oregon State University

BSc, COMPUTER SCIENCE  
2024 | CORVALLIS, OR (Online)

### WeCloudData

DIPLOMA, DATA ENGINEERING  
2021 | TORONTO, ON

### Queen's University

MA, ECONOMICS  
2016 | KINGSTON, ON

### Guelph University

BComm, ECONOMICS  
2009 | GUELPH, ON