

# Jacob Yablonski

(949) 309 - 1955 | [jyablonski9@gmail.com](mailto:jyablonski9@gmail.com)

[GitHub](#) | [LinkedIn](#)

---

Data Analyst with 1+ years experience looking to step into a more challenging role operating in a modern data stack. Experience includes Python, dbt, Airflow, Terraform, and AWS Cloud with a background using containers & serverless functions to build scalable and robust data solutions.

## PROFESSIONAL EXPERIENCE

---

### Mintel

Data Operations Analyst

Chicago, IL

June 2021 - Present

- Served as a first point of contact for all reporting needs, data quality & validation checks, and ad-hoc analyses for some of our digital marketing data which fed into client-facing data products.
- Improved operational efficiency by building a series of ECS Tasks to run scheduled Python & R scripts which queried Elasticsearch, Snowflake, and GraphQL and sent out automated reports.
- Managed the integration process for mapping transformed data from Snowflake into Looker to help enable a self-service system for the Operations team & other internal users to access their data needs.
- Worked with cross-functional teams and product managers to turn business rules and requirements into technical specifications for our pipelines & products.
- Supported other projects involving ELT workflows using tools such as Fivetran, Snowflake, and dbt to break down nested JSON data into relational models used for downstream analytics.

## PERSONAL PROJECTS

---

**NBA ELT Pipeline** - featured on my [GitHub Profile](#) & [Slide Deck Overview](#)

- Built an end-to-end data pipeline to power an R Shiny Web Dashboard on the NBA Season, which sources data daily from multiple different APIs to display current metrics, trends, and other stats.
- Used ECS Fargate to run the containerized Python Script which scrapes the data, validates the schema, and stores it to a PostgreSQL Database as well as S3 for historical re-ingestion if needed.
- Utilized dbt Cloud to execute SQL transformations and shape the source data into dashboard-ready tables, with over 250 tests in place on 25 end models to monitor data quality & integrity.
- Implemented an ML Pipeline using scikit-learn to predict team win probability %s for upcoming games.
- 100% of AWS Infrastructure is built via Terraform, with EventBridge-triggered Lambda functions set up to send out email alerts if any script or cloud hardware failures occur.
- Integrated separate CI / CD steps across the project components to test new changes and release deployments upon PR merges via GitHub Actions.
- Built an Airflow Proof of Concept to illustrate the DAG workflow if the project was to be implemented in an orchestration tool, but it isn't used in production because of cost.

## EDUCATION

---

### California State University, Fullerton

B.A. in Business Administration, concentration in Business Analytics

Major GPA: 3.73

Aug. 2015 - May 2019

## SKILLS & SOFTWARE

---

- Python, R, SQL (Postgres, Snowflake), Git, Airflow, dbt, Docker, Terraform, AWS Cloud (EC2, RDS, S3, ECS, SQS, SNS, EventBridge, IAM, Lambda, API Gateway), Elasticsearch, Spark (PySpark), CI / CD (GitHub Actions), Grafana, Sonarcloud, Looker, Metabase, Jira, Linux & Bash.