Jenna Jordan

Data Engineer

City of Boston Analytics Team (Department of Innovation & Technology)

- Initiated & led the engineering team's migration of over 200 legacy ETL pipelines to dbt, with redesigned orchestration workflows & data warehouse schemas
  - » This resulted in the team's first data catalog, jumpstarted data governance efforts, provided oversight over and analysis of both pipelines and data, and improved the use of both computational resources and engineers' development time
  - » Presented and advocated for this work to the team, the CIO, the department, and at a conference
- Built out custom ETL pipelines to integrate data trapped across many different sources, enabling key metrics to be tracked (e.g. time for affordable housing projects to be approved) and complicated manual processes to be automated (e.g. joining together permitting, housing, and development data)
  - » Advised on structural & process changes to involved source systems across different departments

Data Engineer, Analyst, & Computational Methods Instructor Sep 2020 - Nov 2021

Independent Contractor for: Network Contagion Research Institute; Bright Wolf

- Built custom ETL pipelines to ingest data from social media platforms (e.g. Telegram), enabling this OSINT data to be integrated into PushShift and analyzed for disinformation/extremism
- Taught workshops on computational skills (git, Bash, SQL, python) to NCRI's undergraduate interns
- Built a streamlit web app to visually analyze Bright Wolf's marketing data from Salesforce, email campaigns, and company data enrichment, enabling the sales team to see overarching trends
  - » Built the ETL pipelines to ingest data from these sources via APIs into a local database

Skills

Languages: python, SQL, bash, YAML/YAQL, jinja, regex, XPath, R (tidyverse), SPARQL

Featured libraries: dbt, pandas, streamlit, plotly, SQLAlchemy, Great Expectations

Data Stores: PostgreSQL (+ PostGIS), SQLite, duckdb, JSON, XML, RDF (knowledge graphs)

Other tools: git, GitHub, Civis Platform, Jira, ArcGIS Pro, OpenRefine

Concepts: data modeling (3NF, star schema), relational and analytical database design, data orchestration (ETL pipelines), data management and governance, FAIR data principles

Publications & Presentations

From coast to coast: implementing dbt in the public sector
talk presented at Coalesce (by dbt Labs); San Diego, CA — Oct 2023

Interactive Data Visualizations in Python
workshop published in The Carpentries Lesson Incubator — Aug 2021

Python can be tidy too: pandas recipes for normalizing data
poster presented online for PyCon 2020 — May 2020

Put Relational Databases in Your Data Curation Toolbox
poster presented at ASIS&T Annual Meeting; Melbourne, AUS — Oct 2019

Education & Certificates

Master of Science, Library and Information Science
University of Illinois at Urbana-Champaign — May 2020

Bachelor of Arts, Journalism, Political Science (with honors)
University of North Carolina at Chapel Hill — May 2015

dbt Developer, *certificate from dbt Labs* — Oct 2023

Analytics Engineering with dbt, *certificate from CoRise* — Mar 2022

Certified Instructor, *certificate from Software Carpentries* — Feb 2020