✉️ katie.sipos@gmail.com
🧑‍💻 github.com/ohitsmekatie

👋 Katie Sipos

- A data/analytics engineer and analyst with 10+ years of experience in technology roles at PlanetScale, GitHub, and Microsoft
- Advanced SQL, Python and Git skills
- Experience building and maintaining data pipelines for large datasets to the scale of billions of rows with tools like Airflow, Fivetran, dbt, BigQuery, Hightouch
- Curious learner who loves a challenging data problem to solve
- Excellent cross-team collaboration, experienced with cross-functional stakeholders like finance, marketing, sales, operations, and product
- Thrives in a fast-paced autonomous workplace, outstanding self-prioritization skills

---

Skills and technologies

- **Technologies:** SQL (advanced), Git (advanced), Python + data science libraries (intermediate)
- **Data tools:** Fivetran, Airflow, dbt, Hightouch, Segment, BigQuery, Snowflake
- **Marketing & Sales tools:** Hubspot, Salesforce, Madkudu, Google Analytics
- **Data workflow & visualization**: Hex, Looker, Mode, Jupyter
- **DataOps**: GitHub, CI/CD, testing best practices
- **Data modeling:** advanced understanding of data modeling techniques (star schema, OBT)
- **Executive presence:** engaged with leadership team directly, presenting and putting together board & investor slides

Experience

PlanetScale, Analytics Engineer & Analyst   *August 2021 - April 2024*

- **Built, maintained, and scaled the company's data infrastructure as the first data person**
  - Established BigQuery data warehouse for unified business data, limiting company access by utilizing least privilege principles via Terraform and ensuring security compliance by storing only business-critical PII
  - Created and maintained a robust dbt project, following best practices for modeling, macros, and testing which led to an increase in data quality with 95% project test coverage. Utilized incremental syncs and nightly snapshots to reduce project build time by 30%
  - Maintained Airflow setup that sent a 10TB+ application database to BigQuery which allowed sales and marketing data to be enriched with product data, giving the business a more holistic view of the customer
- **Partnered with teams across marketing, sales, finance, product, and engineering to improve insights and inform business decisions**
  - Defined, modeled, and operationalized a feature utilization score which stored customer product usage over time. Utilized by marketing team to improve mid-funnel nurture emails which led to a 12% increase in the number of users who connected to a database within the first 14 days on platform
  - Analyzed customer storage, usage, and pricing data which informed the rollout of a new plan that made up 40% of paid plan revenue stream within first year of launching
  - Designed and implemented Segment track events for key product conversions which led to a more informed lead qualification process, increasing marketing qualified lead to initial sales discovery call conversion by 15%

- - Analyzed user retention to inform onboarding email campaigns and new technical content, leading to a 5% increase in customers retained between week 0 and 1
    - Created re-runnable parameterized reports in Hex and Mode which led to a decrease in one-off data requests and an increase in stakeholder self-sufficiency
  - **Defined and executed the companies data processes while working to build an engaged culture around data**
    - Influenced company strategy and roadmap direction by leading monthly product data reviews with executive leadership team
    - Improved the company's understanding of the business by developing, socializing and visualizing a set of core product and business metrics. "Company Health" dashboard was viewed by 60% of company at least once per week

GitHub, Senior Data Analyst   *October 2018 - August 2021*

- **Facilitated insights on growth and open source software (OSS) teams as embedded data partner**
  - Collaborated with teammates to analyze data for GitHub's "State of the Octoverse", a yearly report on the state of the open source community which had millions of viewers each day
  - Defined and operationalized a OSS community health metric, a measurement of an OSS community's health used by open source maintainers of projects like Homebrew to understand contributor actions and improve project workflows
  - Conducted free-to-paid conversion analysis to help growth team understand upgrade patterns and identify new experimentation opportunities
- **Streamlined existing processes and promoted a culture of curiosity around data**
  - Ran GitHub's Analytics Primer, a course aimed at increasing data literacy and self-service of data via SQL and tool education. Trained 40+ Product Managers with a 95% completion rate
  - Refined and managed the data team's request and intake process, reducing the average time to ticket response and closure, leading to an increase in tickets opened by external stakeholders

GitHub, Senior Product Manager   *November 2015 - October 2018*
Data analysis, developing metrics, building Looker dashboards, and user research. Worked across many different functions of the business like Enterprise, GitHub Education, the user profile, user authentication, OSS communities

MagmaLabs, Technical Product Lead   *November 2014 - November 2015*
Client management & team mentorship. Worked mainly with GoPro on the product management side of GTM function for new product lines

ModCloth, Technical Project Manager   *November 2009 - November 2014*
Large scale project implementation & remote team management. Key projects include migration from Quickbooks to Oracle for the finance team and the rollout of a new warehouse management system for 250+ warehouse employees

Education

- [Analytics Engineers Club](): a 10 week program aimed at learning analytics engineering practices
- [Uplimit](): immersive 4 week programs in advanced SQL, analytics engineering with dbt, data modeling for the modern data warehouse, and intermediate Python
- Community College of Allegheny County: Associate degree, Data Analytics