Introduction to Data Science with Snowflake

Tatiana Petrache, Solution Engineer

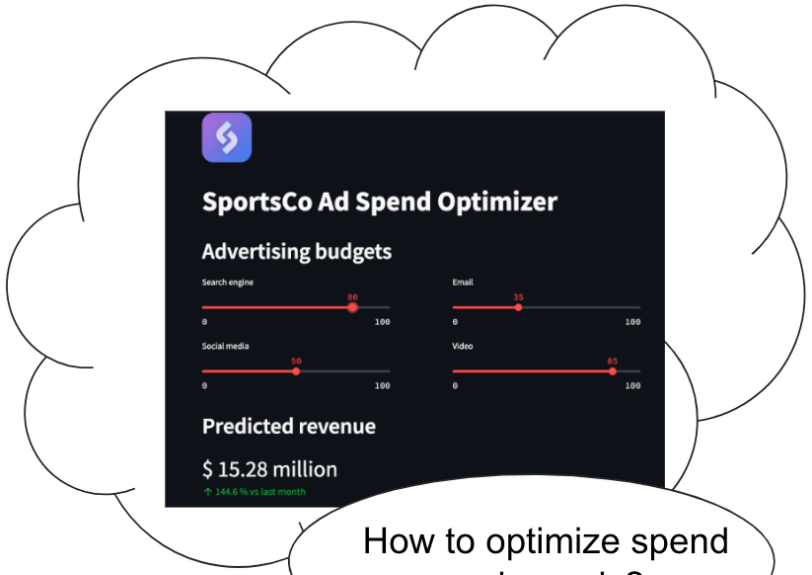
About me

- My data journey: DWH, BI, Big Data, Data Science & ML
- electronic & classical music, jazz
- books and podcasts

Agenda

- Step1: Fix your Data Foundation
- Step 2: Simplify your Data Pipeline
- Step 3: Remove bottlenecks with elastic processing
- Step 4: Build and deploy from your tool of choice
- Step 5: Make your model actionable in minutes

Every data science task starts with a business problem….

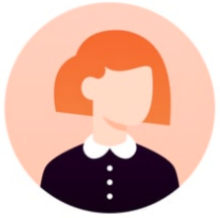


How to optimize spend across channels?

Stefan, CMO SportsCo

John, Data Engineer

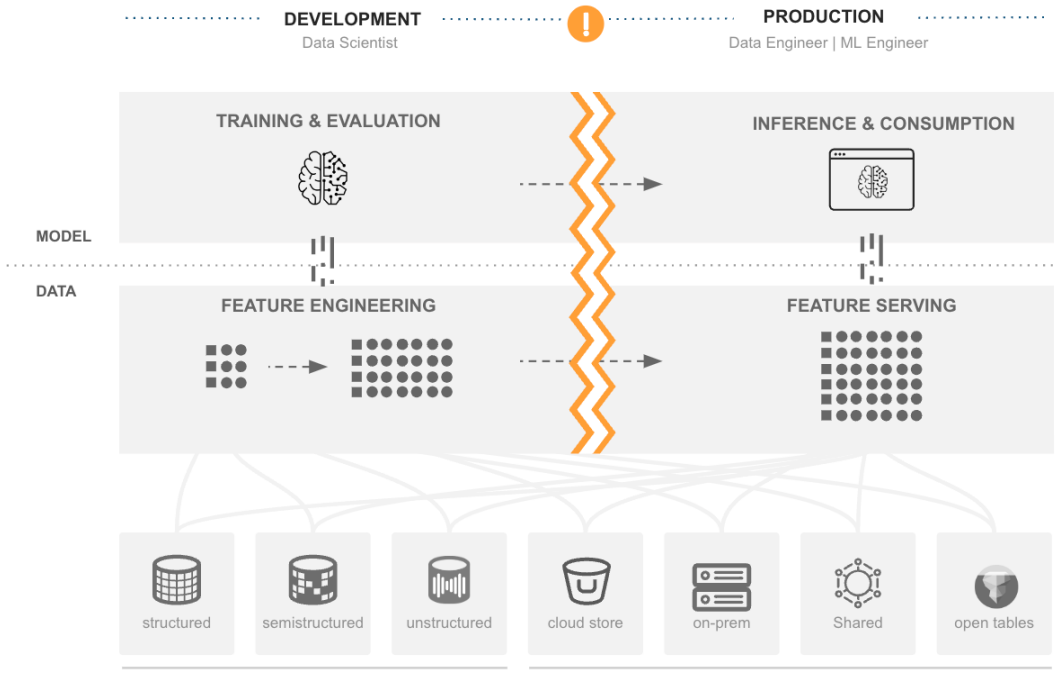
I am stuck with managing pipelines. It takes a lot of manual configuration…



Emma, Data Scientist

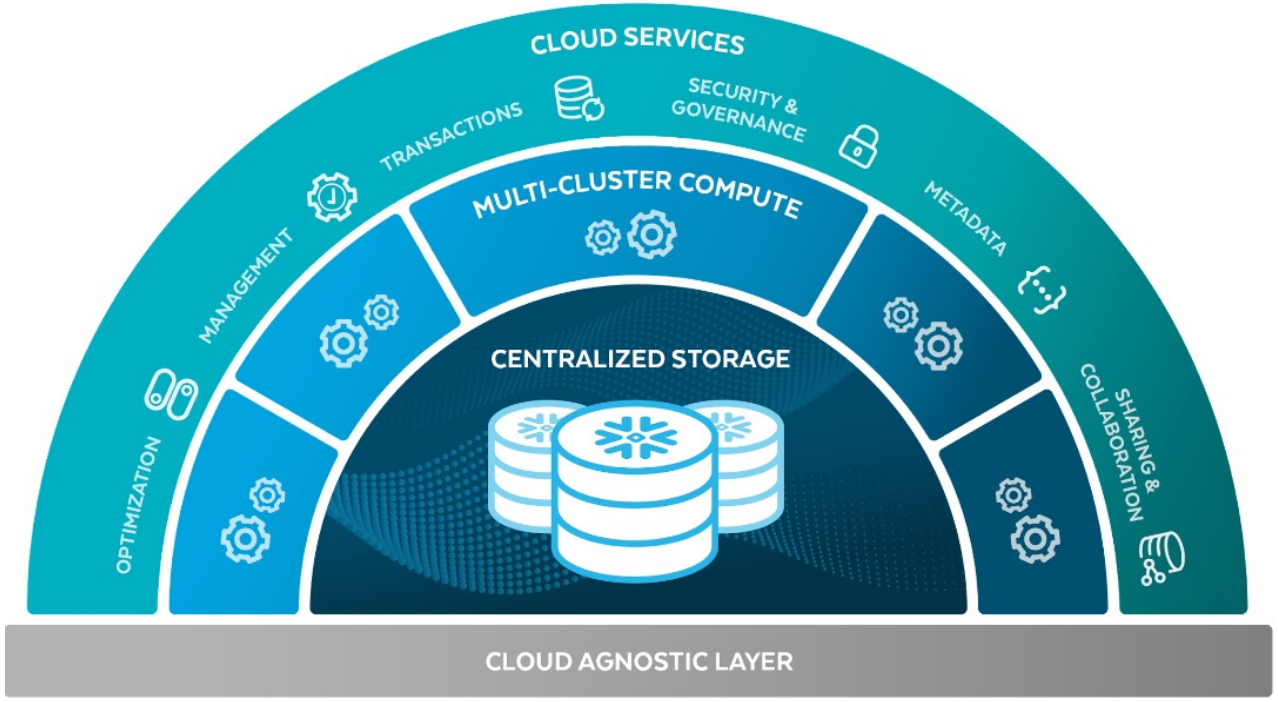
I am waiting for John to finish his work… and now I have to learn front-end programming :(

And a lot of challenges…



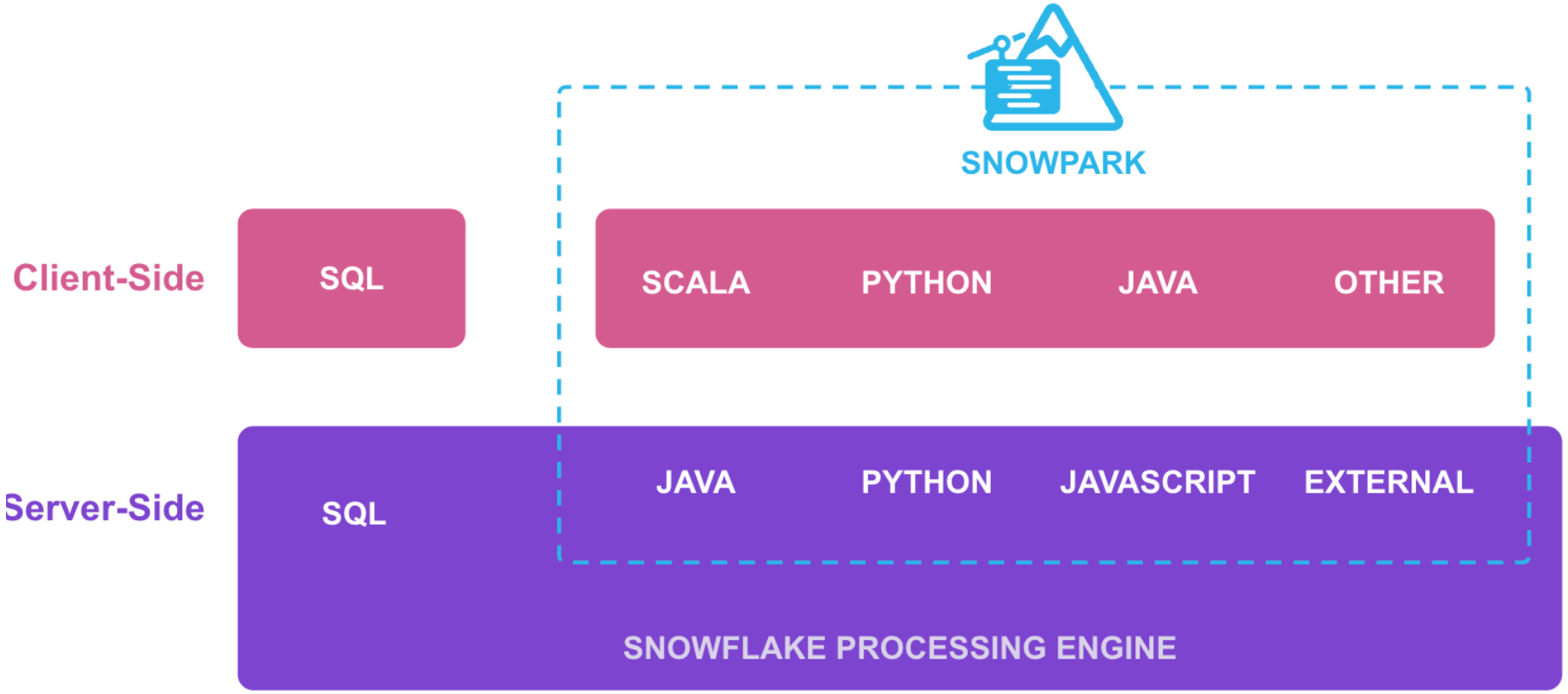
- Siloed teams & tools
- Processing complexity
- Broken data foundation

Why Snowflake?



- ✓ Fully managed
- ✓ Multi-language, elastic engine
- ✓ Unified Data Access

Meet Snowpark



Why Snowpark for Python?

Familiar Programming Constructs

Use familiar syntax with DataFrame abstraction

Rich Ecosystem

Easy access to thousands of packages with automated dependency management

Secure Processing

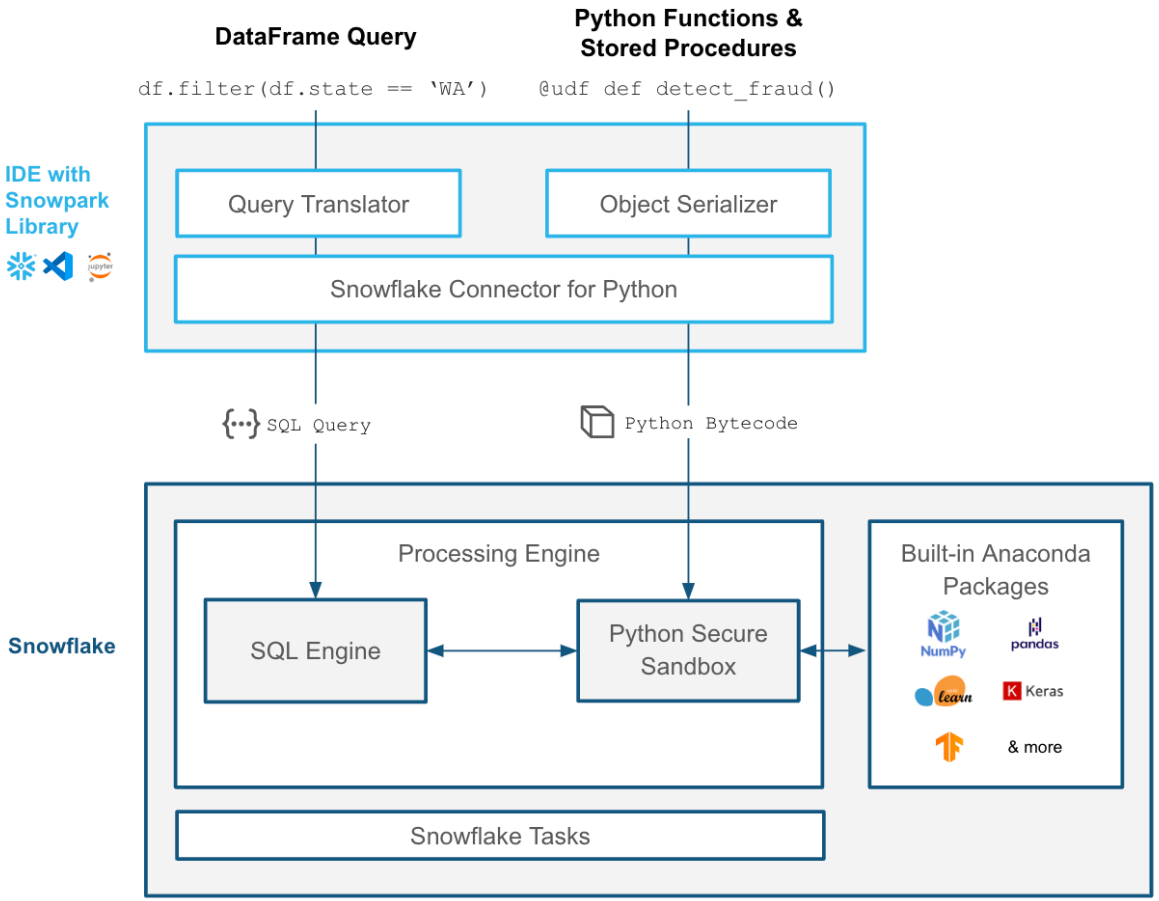
Build with confidence in a highly secure, sandboxed environment

How does it work?

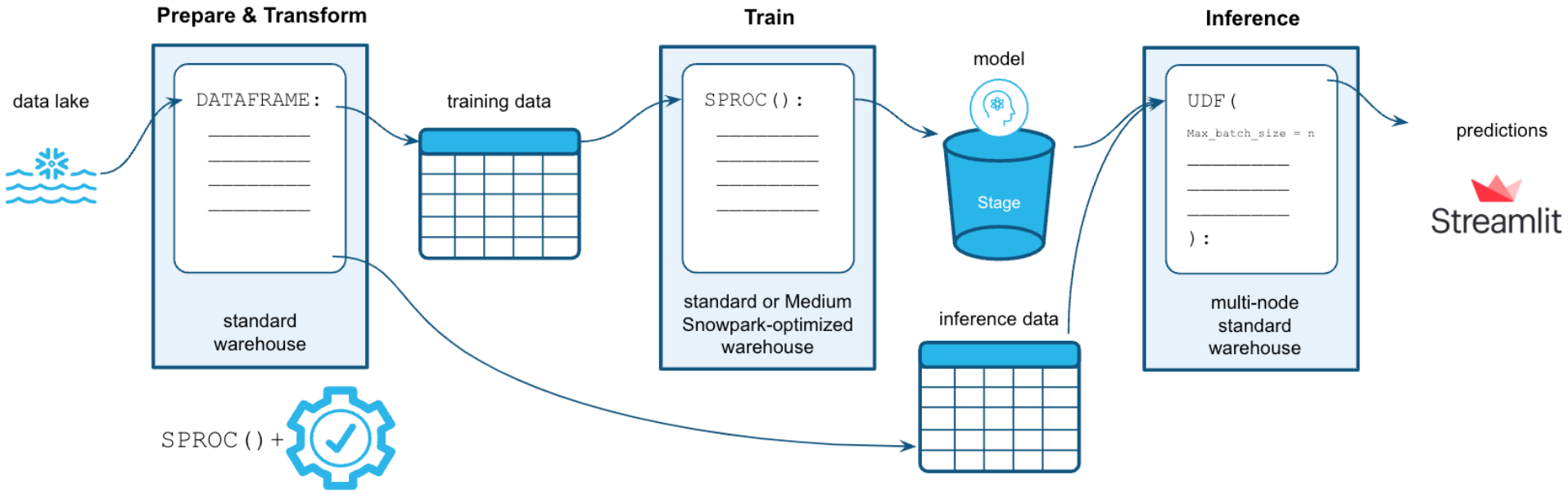
> Develop in IDE of Choice
Use VSCode, Jupyter or any client with a python kernel

> Process with security & scalability
Bring processing to governed data for DataFrame operations and custom Python code

> Automate code execution
Orchestrate your code with using SProcs with Snowflake Tasks

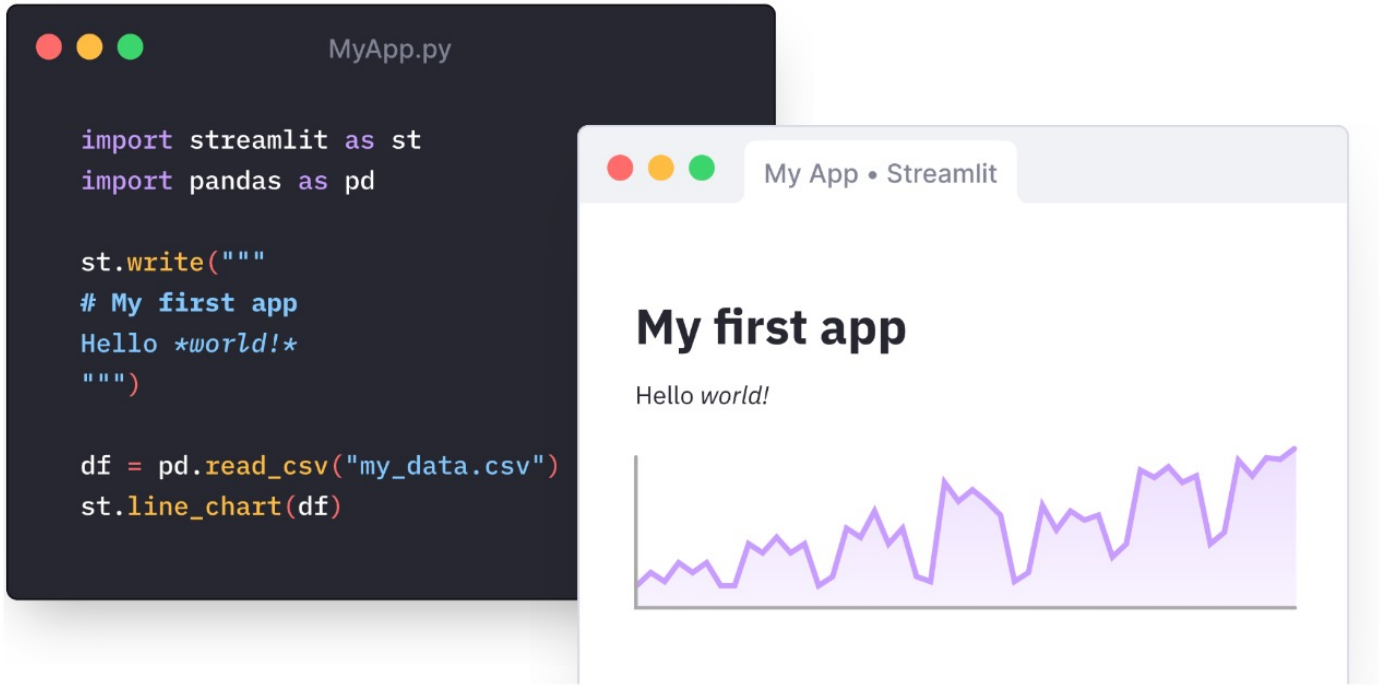


Workshop flow



Streamlit

The fastest tool for data scientists and ML engineers to create a visual, interactive data apps.
It gets you from a simple Python script to a data app in minutes!



Conclusion

Bring the code to data!
Familiar syntax
Rich ecosystem
Scalable
Fully managed
Secured, governed

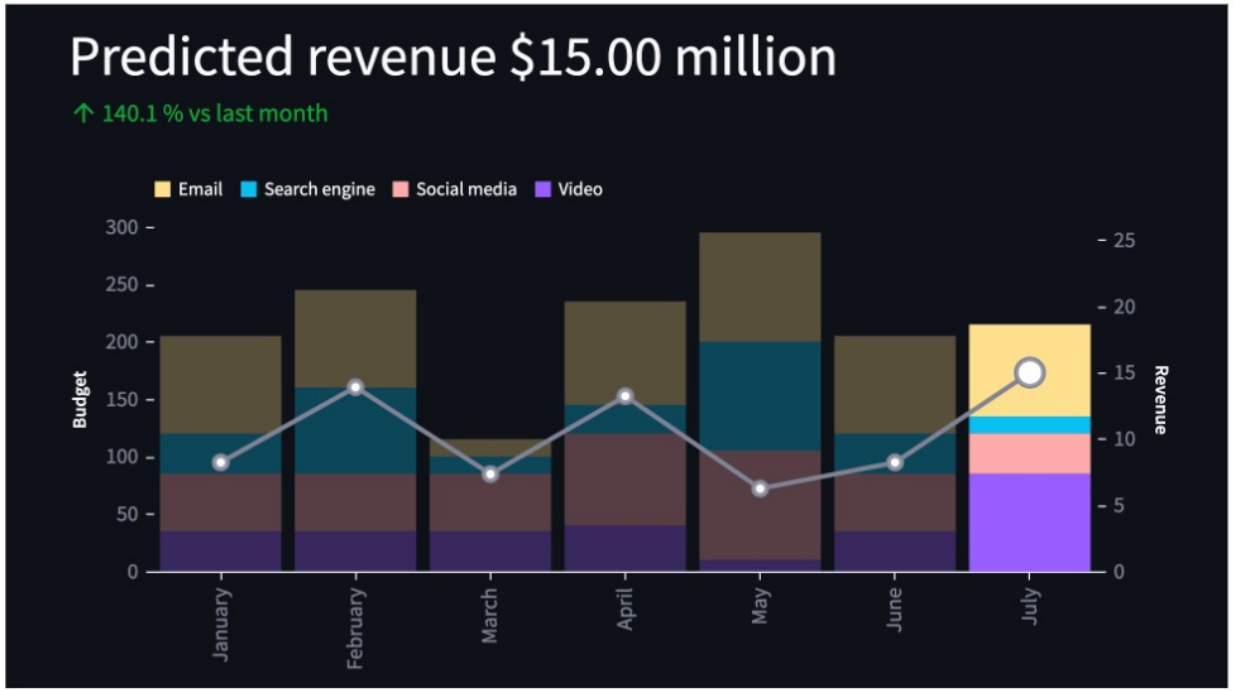
Snowflake and Snowpark

Data engineering at scale
Automated
data pipelines
Easy monitoring and
debugging

John, Data Engineer

Fast learning curve
Easy deployment
Automated package
management
Fast web app development

Emma, Data Scientist



Stefan, CMO SportsCo

```
if questions:
    try:
        answer()
    except RuntimeError:
        pass
else:
    print("Thank You.")
```