

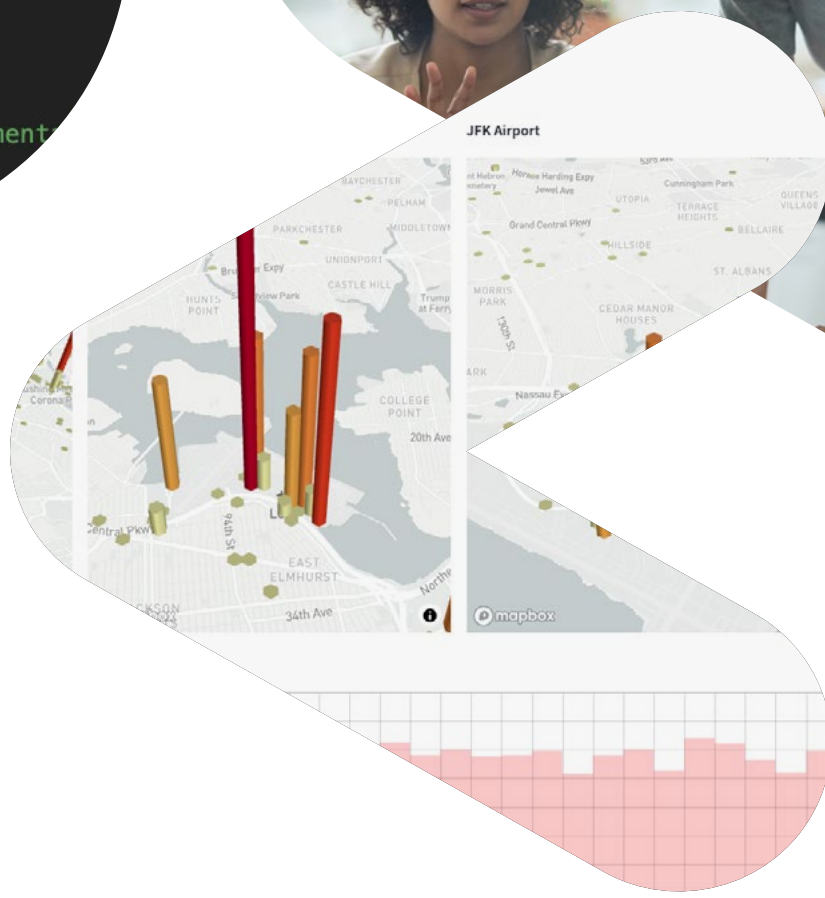
```
# Import required libraries
# Snowpark
from snowflake.snowpark.session import Session
from snowflake.snowpark.types import IntegerType
from snowflake.snowpark.functions import avg, sum, col, call_udf, lit, call_udf

import pandas as pd
import streamlit
import streamlit as st

# Create Snowpark DataFrames that loads data from Knoema: Environment
def load_data(session):
    # CO2 Emissions by Country
    snow_df_co2 = session.table("ENVIRONMENT.EDGARED2019")
    snow_df_co2 = snow_df_co2.group_by('Location Name')

    # Forest Occupied Land Area by Country
    snow_df_land = session.table("ENVIRONMENT")
    snow_df_land = snow_df_land.group_by('Location Name')

if "snowpark_session" not in st.session_state:
    session = Session.builder.configs({"warehouse": "test", "database": "test", "schema": "test", "role": "ACCOUNTADMIN"}).create()
else:
    session = st.session_state["snowpark_session"]
```



DATA APPLICATIONS IN PYTHON

How to bridge the gap between machine learning and business action



CHAMPION
GUIDES

EBOOK

TABLE OF CONTENTS

- 3** Executive summary
- 4** Why is it so difficult to get from data to action?
- 6** The new way: Python-based app dev in the Data Cloud
- 8** Build ML models in your preferred language at scale with Snowflake
- 10** Transform ML models into apps in minutes with Streamlit
- 12** Create better data apps faster with Snowflake and Streamlit

EXECUTIVE SUMMARY

Exponential data growth is our new normal. The question now is, How quickly can you turn data into useful, relevant information?

Data scientists take the first step, using machine learning (ML) and AI to make sense of the data deluge, building models that feed on large amounts of data to deliver valuable insights. But if business users don't understand the output of machine learning models or see how the results fit into their workflows, those valuable insights (and the work that went into producing them) may be left unused.

Interactive applications that visualize ML results can bridge the gap between the data and code used by data science teams and the insights business users need to take action. Yet, for the majority of business teams, it is simply out of reach to have a dedicated team of developers building internal business apps. And data scientists have little time to learn the additional programming skills required to create interactive apps.

What if data scientists and other Python developers could build applications using nothing but Python? And what if they could rapidly share and iterate on these interactive applications, allowing business teams to immediately mobilize data and put ML-powered insights to work?

With Snowflake and Streamlit, there is now a solution to go all the way from raw data to data applications. Read on to learn why the ML-to-action gap exists and how you can bridge it by:

- Using Snowflake to build and deploy ML models at scale
- Using Streamlit to build interactive apps in Python
- Leveraging Streamlit-built apps to quickly and securely iterate with business teams to ensure your ML-derived insights deliver maximum value

WHY IS IT SO DIFFICULT TO GET FROM DATA TO ACTION?

ML models can help you get the most from your data by transforming huge amounts of data and highlighting trends and insights. Building a high-performing ML model is not a data scientist's end goal; the end goal is to have the model predictions drive business value. That requires business users to turn ML-powered insights into action.

Unfortunately, there's a gap between data scientists and business users when it comes to leveraging ML-powered insights. Data scientists spend much time and effort creating models that can offer great insight, but business users find it difficult to understand and trust the findings because they don't use familiar business terms and visualizations.

Without that understanding and trust in model results, business stakeholders are less likely to incorporate the ML insights into their workflows. That means the extensive data science work involved in identifying and preparing relevant data sets, building accurate models, testing and refining models, and analyzing outputs has gone to waste.

Both sides know the business is missing out on the chance to fully realize powerful ML insights. So what's kept organizations from successfully bridging the ML-to-action gap?



Data scientists work with Python; business teams work with campaigns, channels, stages, forecasts, trends, and more. To find common ground and drive business success, they need a solution that quickly turns data science findings into formats familiar to business users.

DATA SCIENCE ANALYSIS IS NOT EASILY ACCESSIBLE

Data scientists and engineers love to work in Python because it's easy to learn and use, it has an active community and many machine learning libraries, and they can develop and share Python code with other data scientists using notebooks and other interactive development environments (IDEs).

Business users are unlikely to speak Python or know how to set up and run a notebook, so to share results, data scientists often take analysis from

their notebook and put screenshots in slides or docs. These static screenshots are missing all of the context originally available as code in the notebook, which leaves business users—who range from ML novices to somewhat savvy—wishing the information was represented in a way that helped them understand and trust how results were generated. It's difficult for business users to trust the data science results when they can't easily grasp the process and model logic, as well as what data was involved and how it was used.

The static results docs don't offer business users a way to adjust the data to confidently show business cases to customers and C-suite stakeholders.

This method of sharing analysis and results is also out of alignment with today's fast-moving organizations. Data scientists can't easily collect feedback from business stakeholders to update their ML models, nor can business users engage with data in near-real time to see how trends are affected by changing business conditions.

BUILDING AND MAINTAINING APPS TAKES TOO MUCH EXTRA EFFORT

Applications that visualize data and display results with intuitive, interactive UIs are an excellent way to help transform predictions into action. But application development—and application maintenance—traditionally requires web dev and front-end programming knowledge as well as dedicated development time.

Data scientists could work on building out front-end UIs that help business users more readily consume the data and results, but that is a time-consuming effort which requires learning additional programming languages and unfamiliar dev tools. In addition, data scientists would have to think about devOps and dedicate time to managing and updating apps for as long as they're in use. On the business side, teams also have limited time to learn complex data science models and rarely have dedicated developer resources to build apps from models provided by data scientists.

Clearly, what's needed is a solution that allows data scientists to rapidly explore data, build interactive applications using their preferred language—Python—and share them with business teams so they can put ML-powered insights to work immediately.



THE NEW WAY: PYTHON-BASED APP DEV IN THE DATA CLOUD

With Snowflake and Streamlit, you can do more than just bridge the gap between data scientists and business users—you can close it completely.

This new approach leverages the Snowflake platform's fast data access, elastically scalable data processing, and flexible language support to accelerate ML workflows with minimal operational overhead. It also adds the previously missing piece: Streamlit's open-source, pure-Python application framework, which allows you to turn data and models into shareable web apps without front-end dev expertise.

Data scientists can now take their Snowflake data, models, and generated predictions and use Streamlit to build apps that present the findings in a business-friendly interface. This streamlined path bridges the gap between machine learning and business action, delivering big benefits for both data practitioners and business users. Here's a look at five of the top benefits:

ACCELERATE MODEL DEVELOPMENT

Snowflake provides a single point of access to a global network of trusted data and has native support for structured, semi-structured, and unstructured data so data scientists spend less time finding and preparing data sets. This data can be easily transformed into automated data pipelines and ML models with [Snowpark](#), Snowflake's developer framework.

BUILD (AND UPDATE) APPS IN PURE PYTHON

Data scientists can now build both ML models and interactive apps. There's no front-end experience required and no HTML, CSS, JavaScript, or HTTP requests to work on, and the apps can automatically update with the latest insights. Streamlit also provides a jump-start to quickly create prototypes, iterate with stakeholders, and make changes on the fly—in Python—to improve apps' effectiveness.

SHARE INSIGHTS IN A WAY EVERYONE UNDERSTANDS

Apps help business stakeholders visualize data and ML insights using familiar formats and terms. But they can also help data scientists educate stakeholders and gain their trust by showing the data and logic involved and explaining why a specific algorithm was used.

WHAT IS STREAMLIT?

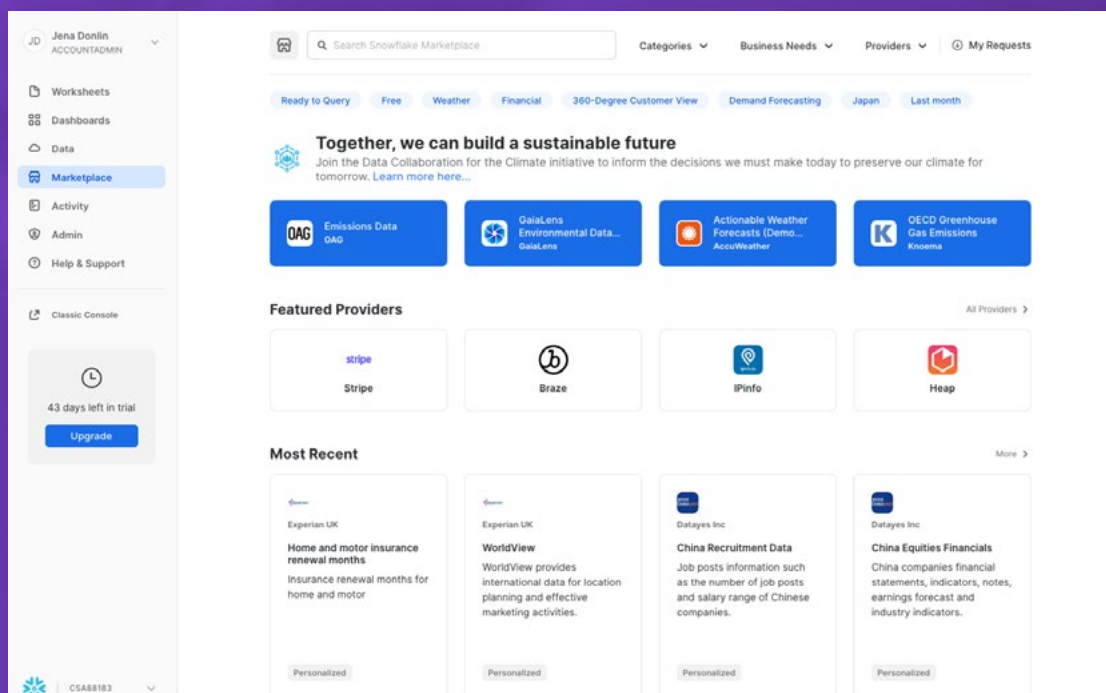
Streamlit is an [open-source](#) Python framework that enables data scientists to create data applications in minutes without front-end development expertise. Data scientists and ML teams can build an app in a few lines of code with Streamlit's API and see it automatically update as they iteratively save the source file. They can then instantly deploy interactive apps and effortlessly share and manage them directly from Streamlit.

Snowflake acquired Streamlit in March 2022. Learn more about Streamlit at <https://streamlit.io>.



SNOWFLAKE MARKETPLACE: YOUR COLLABORATION SUPERSTORE

As businesses open up their data in the Data Cloud, your apps can bring in data from the broader ecosystem to enhance ML initiatives and drive next-level collaboration. [Snowflake Marketplace](#) provides data sets from hundreds of providers, while [Snowflake Secure Data Sharing](#) capabilities allow organizations to collaborate with data while maintaining full control over their own data resources.



Snowflake Marketplace

GIVE BUSINESS USERS THE CHANCE TO PLAY WITH DATA

Interactive Streamlit apps encourage users to engage with the data and models. Input boxes, sliders, and more let business users interact with model results, answer their own questions, and incorporate relevant, up-to-date findings into their forecasts, predictions, and strategic plans.

COLLABORATE OFTEN AND EFFICIENTLY

When you are able to create an app in a few minutes and share it with a click, collaboration gets much easier. You can collect quick feedback from business stakeholders, update the model and Streamlit app, and ensure the business team has what it needs to make confident decisions. More experiments will progress from prototype to production when business teams thoroughly understand the available insights.

Let's take a closer look at how you can use Snowflake and Streamlit to improve the process of building ML models at scale and transforming them into shareable data apps.

BUILD ML MODELS IN YOUR PREFERRED LANGUAGE AT SCALE WITH SNOWFLAKE

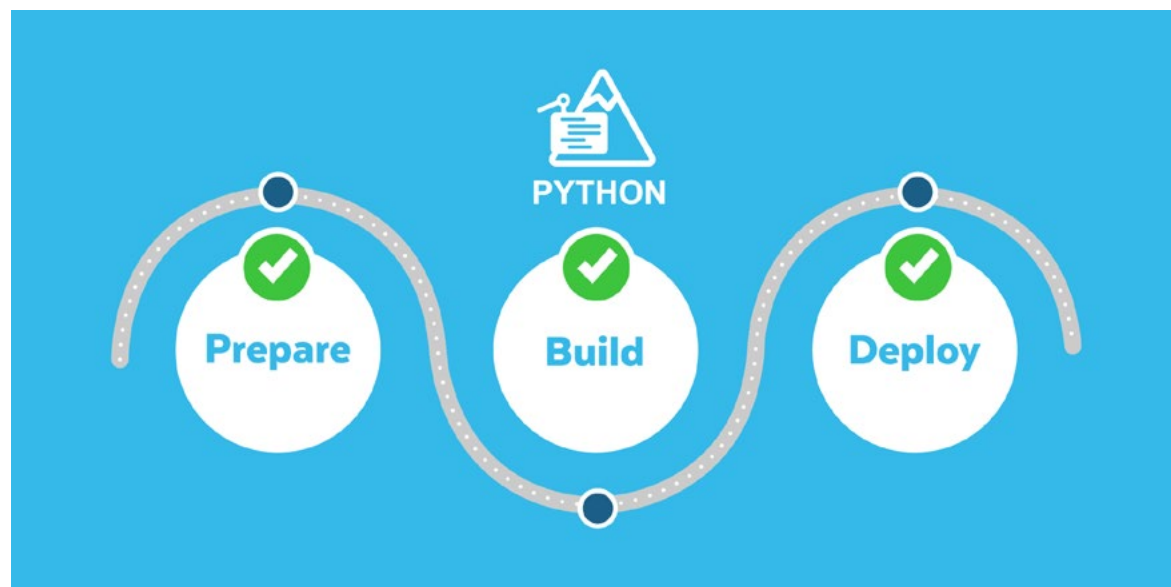
PLANNING AND ACCESSING DATA

First and foremost, align with the business team on the overall objective. Not only will this give you critical details for the data sets to use in the model and the questions they want to answer, but you'll also get information that can help you build an app to match the team's specific needs and stakeholder preferences.

Snowflake simplifies the task of getting access to data sources for use in ML models. The Data Cloud provides a single, consolidated source for all data so you spend less time hunting for relevant data and requesting access. You can also incorporate third-party data from Snowflake Marketplace with the knowledge that you'll get clean, trusted data to enhance the predictive power of your model.

BUILDING DATASETS FOR MACHINE LEARNING MODELS

Before an ML model can be presented as an interactive app, data must be prepared and feature engineering by selecting the right variables from raw data to provide the most relevant inputs to the ML models. Using Snowpark and the Snowflake elastic performance engine, all necessary development and data processing can be done with near-zero operations at scale.



Data scientists can use Snowflake for model training and inference.

BUILDING ML MODELS USING PYTHON INSIDE SNOWFLAKE

Once you've constructed a historical data set and completed feature engineering, you can use [Snowpark for Python](#) (currently in public preview) to train your ML model on the prepared data set and deploy it as a function that runs where the data lives. Using the Snowpark API, you can quickly create secure connections between Snowflake and your favorite IDE or notebook. Then, you can:

- **Query and transform Snowflake data using familiar Python DataFrames**
- **Execute model training and inference directly inside the Snowflake engine leveraging integrated open-source libraries from Anaconda**
- **Automate the end-to-end ML pipeline using Snowflake Streams and Tasks**

Under the hood, Snowpark allows data scientists to work in their preferred language while getting the benefits of Snowflake's elastic performance engine to handle the heavy processing requirements.

MAKING THE MODEL AVAILABLE IN SNOWFLAKE

To deploy your trained model, just register the model file as a Python user defined function (UDF). Once registered, users with access to the function can get new predictions by simply calling the function along with the new records that need to be scored.

Because the model is available as a Python UDF in Snowflake, anyone with the correct access level can also get a prediction from inside the worksheet. This means analysts and other users can use SQL to generate model predictions without the need to establish connections to an external IDE, install open-source packages, or manage a separate environment.

WHAT IS SNOWPARK AND WHY IS IT USEFUL?

Snowpark is a developer framework for Snowflake. It enables data scientists and developers to use their preferred language—including Java, Scala, and Python—to execute data pipelines, ML workflows, and data apps quickly and more securely from a single platform using deeply integrated, DataFrame-style programming and custom functions.

[Snowpark for Python](#) lets you tap into Python's rich ecosystem of open-source packages and libraries with automated dependency management so you can code fast and deploy faster. You can also push custom Python code to Snowflake, running it seamlessly in a highly secure sandboxed environment built into the Snowflake processing engine.

Learn more about Snowpark for Python [here](#), or dive right in with the [Snowpark QuickStart](#).



TRANSFORM ML MODELS INTO APPS IN MINUTES WITH STREAMLIT

Once you have deployed the ML model, it's time to convert it into an interactive app that's easy for business teams to consume. Using Streamlit, you can build that app and share it with internal users—without leaving the comforts of your familiar Python environment.

Today, using Snowflake and Streamlit open-source, you can create an interactive app by:

- **Installing the Streamlit open-source Python library.**
- **Connecting to data in Snowflake using the Snowflake Connector for Python.**
- **Building your app and editing it locally.** The Streamlit API will render your existing code as a web app, but you can also insert snippets to add components that make your app more engaging, like interactive sliders, input boxes, bar charts, and tabbed displays. No need to wrestle with HTML or CSS.
- **Deploying and sharing applications on your own.** Sharing your app is as easy as clicking “Share” in Streamlit and sending the URL to the business team. If the business users request changes—maybe they prefer sliders to input boxes—the Streamlit app can be edited live.

The Snowflake–Streamlit integration currently in development will allow you to:

- **Build a Streamlit app in Snowsight:** Build interactive apps in Python within a Snowsight worksheet.

- **Deploy in Snowflake:** Deploy and run your application entirely on Snowflake's governed platform. There's no need to move or copy data out of Snowflake.
- **Share with Snowflake users:** Securely share your app with other Snowflake users, collect feedback, and make updates quickly to iterate faster and more efficiently. Because all this is happening in Snowflake, role-based access controls are consistently enforced.
- **Monetize using the Native App Framework:** The Snowflake Native Application Framework, currently in private preview, provides a streamlined path to monetize apps through Snowflake Marketplace, where you can make your app available to other Snowflake customers and open new revenue opportunities.

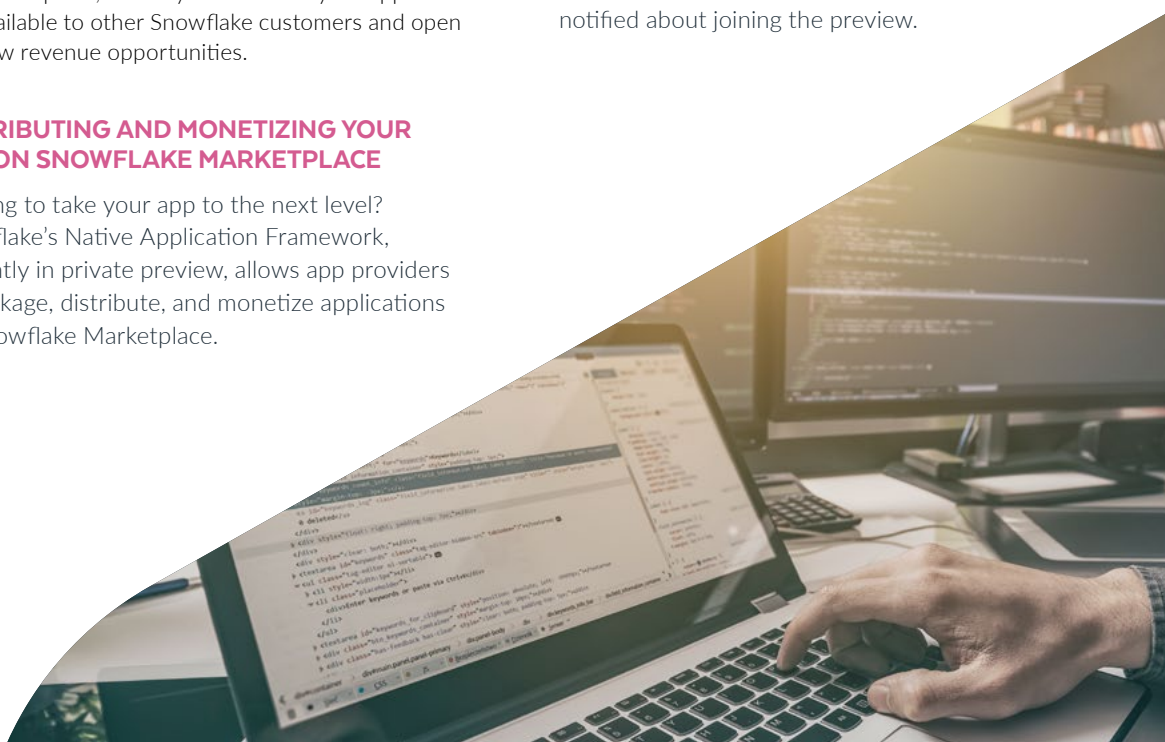
DISTRIBUTING AND MONETIZING YOUR APP ON SNOWFLAKE MARKETPLACE

Looking to take your app to the next level? Snowflake's Native Application Framework, currently in private preview, allows app providers to package, distribute, and monetize applications on Snowflake Marketplace.

The Native Application Framework will include the Streamlit object (currently in development), making it easy to add a user interface to the applications you build and publish on Snowflake Marketplace. And in the native app deployment model, compute costs and sensitive data remain on the customer's side, which can help you deploy faster and with better margins.

For customers, native apps give them the benefit of control: control over their data and what an application can do in their accounts. All of this helps customers speed up their time to value.

Learn more about native apps [in this blog](#) or visit snowflake.com/native-apps to be notified about joining the preview.



CREATE BETTER DATA APPS FASTER WITH SNOWFLAKE AND STREAMLIT

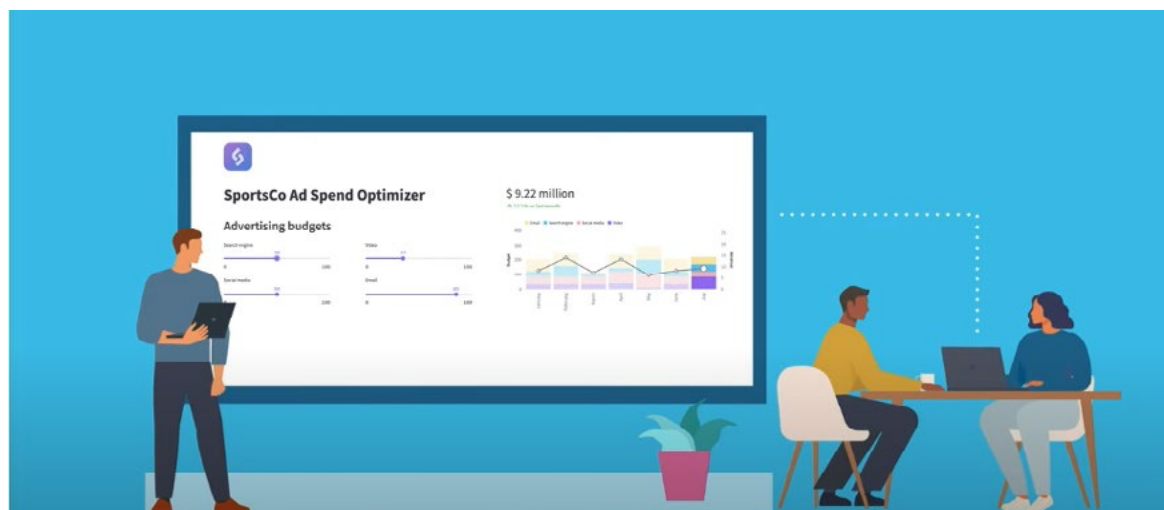
The Snowflake–Streamlit combination brings to life a whole new way of collaborating with business teams via interactive apps while taking advantage of Snowflake scale, performance, and governance.

Data scientists creating internal apps for business teams is just one use case; consider these examples as well:

- **External applications:** When prototyping an app, developers want short feedback loops with initial customers to quickly find out what works and adjust accordingly. Iterating on an app UI can be slow and involve front-end expertise to make and publish updates.

With Snowflake and Streamlit, app developers can create quick prototypes and share directly with potential customers or test groups. Once they have a preferred version, they can choose whether to continue working in Streamlit or move to another tool to fully customize the app's look and interaction style.

- **Retail:** A retailer can create a Streamlit app to help store management teams view predictions of high traffic times and use those insights to adjust staff levels and inventory accordingly.
- **Financial services:** A bank can build an interactive app for analysts that visualizes holdings insights and also updates in near-real time when new data comes in.



With Streamlit, data scientists can use familiar Python language to build interactive, highly visual web apps that resonate with teams.

- **Healthcare:** A hospital can design a Streamlit app to give patient care coordination teams insights into various outreach efforts so they can focus their efforts for maximum positive outcomes.

With Snowflake and Streamlit, there is no reason for ML models to sit idle or for companies to avoid putting ML-powered insights to work.

The Snowflake Data Cloud makes data more readily accessible and available for analysis and Snowflake's elastic performance engine makes the necessary processing a breeze. With Streamlit, data scientists can rapidly develop interactive apps using familiar Python constructs, giving them a powerful way to present ML insights in a format that business users can understand and embrace. Business teams can engage more deeply and directly with the data that influences critical business decisions. Most importantly, the interactive apps can facilitate collaboration on data—and inspire new actions.

TAKE THE NEXT STEP: EXPLORE THE POWER OF SNOWFLAKE AND STREAMLIT

- **Build an app:** Build your own application with Snowflake and Streamlit by following this [QuickStart demo](#)
- **Get excited:** Get a sneak peek at the integrated experience of Streamlit in Snowflake (currently in development) [with this demo](#) from Adrien Treuille, Head of Streamlit at Snowflake
- **Get inspired:** See what others have built on Streamlit at <https://streamlit.io/gallery>



ABOUT SNOWFLAKE

Snowflake enables every organization to mobilize their data with Snowflake's Data Cloud. Customers use the Data Cloud to unite siloed data, discover and securely share data, and execute diverse analytic workloads. Wherever data or users live, Snowflake delivers a single data experience that spans multiple clouds and geographies. Thousands of customers across many industries, including 506 of the 2021 Forbes Global 2000 (G2K) as of April 30, 2022, use Snowflake Data Cloud to power their businesses. Learn more at [snowflake.com](https://www.snowflake.com)



© 2022 Snowflake Inc. All rights reserved. Snowflake, the Snowflake logo, and all other Snowflake product, feature and service names mentioned herein are registered trademarks or trademarks of Snowflake Inc. in the United States and other countries. All other brand names or logos mentioned or used herein are for identification purposes only and may be the trademarks of their respective holder(s). Snowflake may not be associated with, or be sponsored or endorsed by, any such holder(s).