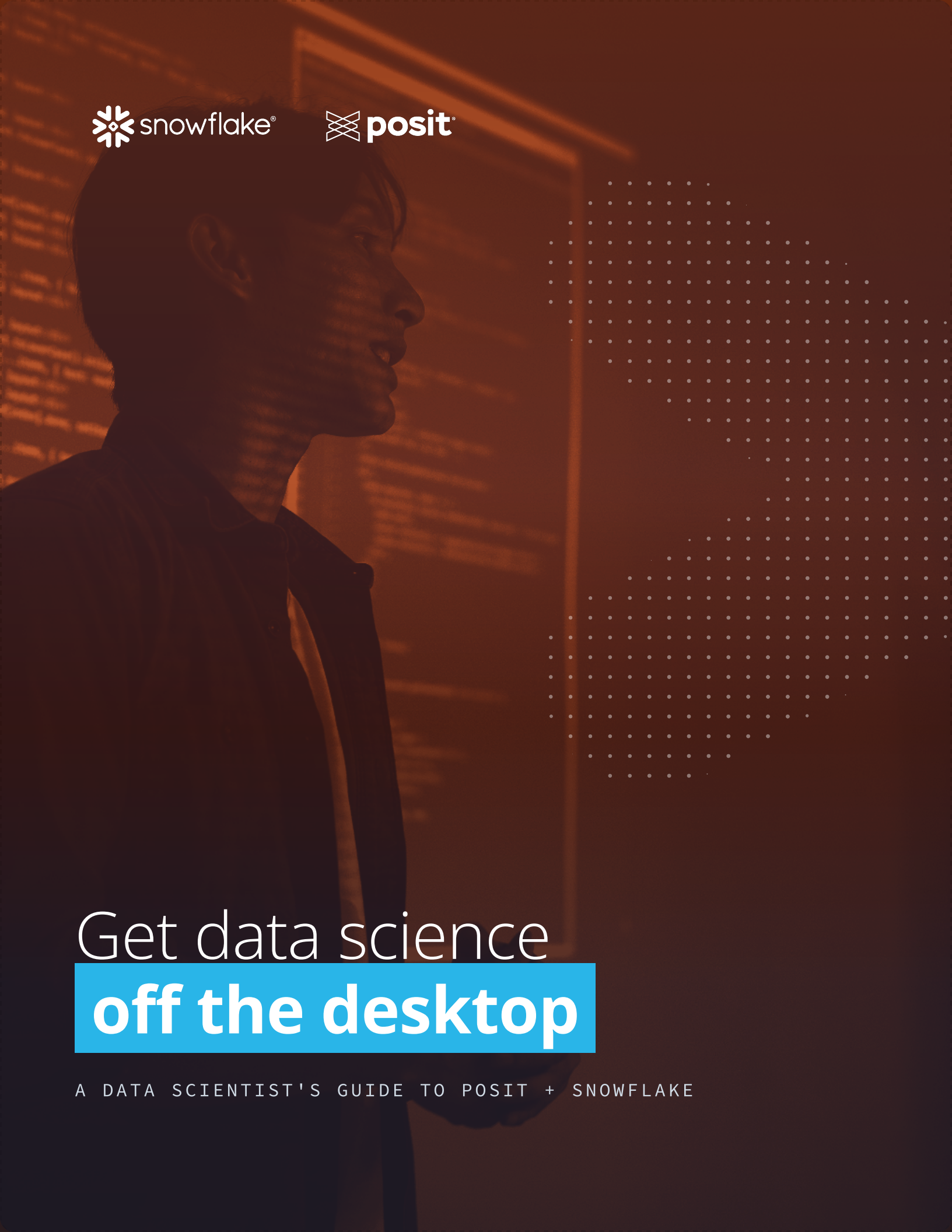


snowflake® posit®

Get data science **off the desktop**

A DATA SCIENTIST'S GUIDE TO POSIT + SNOWFLAKE

TABLE OF CONTENTS

Get data science off the desktop

EXECUTIVE SUMMARY 03

01
The problem you deal with 04

02
Tools that bring Snowflake into your R and Python session 05

03
Where you write the code 06

04
Proof from teams like yours 08

05
What changes for you, day to day 11

06
Go further 12

ABOUT
About Posit and Snowflake 13

Executive summary

You work around the same tradeoff every week: the open-source tools you want on one side, locked-down enterprise data on the other. **Posit removes that tradeoff:** you run R and Python against your Snowflake data, on Snowflake compute, with AI support from Snowflake Cortex and the credentials you have, inside the governance boundary IT already approved for Snowflake itself, with no new security review to win first.

This guide covers the tools, the IDEs, and the AI assistant built into that workflow, and what happens when teams use them. One team cut deployment from weeks to hours and turned a full-day weekly report into one that runs on its own. Another analyzed 30,000 employee survey comments that SQL alone couldn't handle.



01 | The problem you deal with

Your data is in Snowflake. Your tools aren't. So you pull a sample to your laptop, run your analysis, and hope the sample looks like production by the time you're done. Or you wait on IT to provision a VM, install a driver, and grant you a role, which can take days that you don't have.

Each workaround costs you time: extracting data means securing and syncing a copy, and each new authentication step delays your next model. It doesn't make your analysis better, and adds time you don't have.

The workaround hits hardest on the step that takes the most time. Importing and tidying data (joining tables, recoding fields, and reconciling formats) is up to 70% of a data science project, before any real analysis starts. Doing that work against a local sample instead of the governed source risks drift, and it means reconciling your results against production again once you're done.

Then, even when you finish the analysis, the handoff can block it. Without an engineering team to deploy it, a finished model or dashboard ships as a screenshot in a slide deck instead of something stakeholders can open and interact with. Weeks can pass between finishing the work and a stakeholder opening it. Some of it goes unused, or the business has already made the call without you by the time it's ready.

02 | Tools that bring Snowflake into your R and Python session

Data science work breaks down into three steps: import and tidy the data, understand it through transformation, visualization, and modeling, and communicate what you've learned as a finished data product. **These packages cover all three, running on Snowflake compute, without changing how you write R or Python.**



dbplyr translates dplyr syntax into Snowflake SQL, so the joins and recodes that make up most of your import-and-tidy work run as a query against the warehouse instead of pulling data down first.



orbital converts tidymodels (R) and scikit-learn (Python) models into native SQL for in-database scoring, so you ship a model as a running data product instead of waiting on an engineering team to deploy it as a separate API. Predictions run inside Snowflake's query engine, where the data already lives, so scoring doesn't move rows out to a separate model server.



skiLift is a DBI-compliant R connector that authenticates and communicates with Snowflake through its SQL API. Reading and writing data takes nothing beyond installing the package. (Read more)



skiPatrol is an R interface to Snowflake's ML platform: model registry, feature store, experiments, and monitoring. Models you build in R run in Snowflake alongside everyone else's, instead of sitting in a separate environment. (Read more)

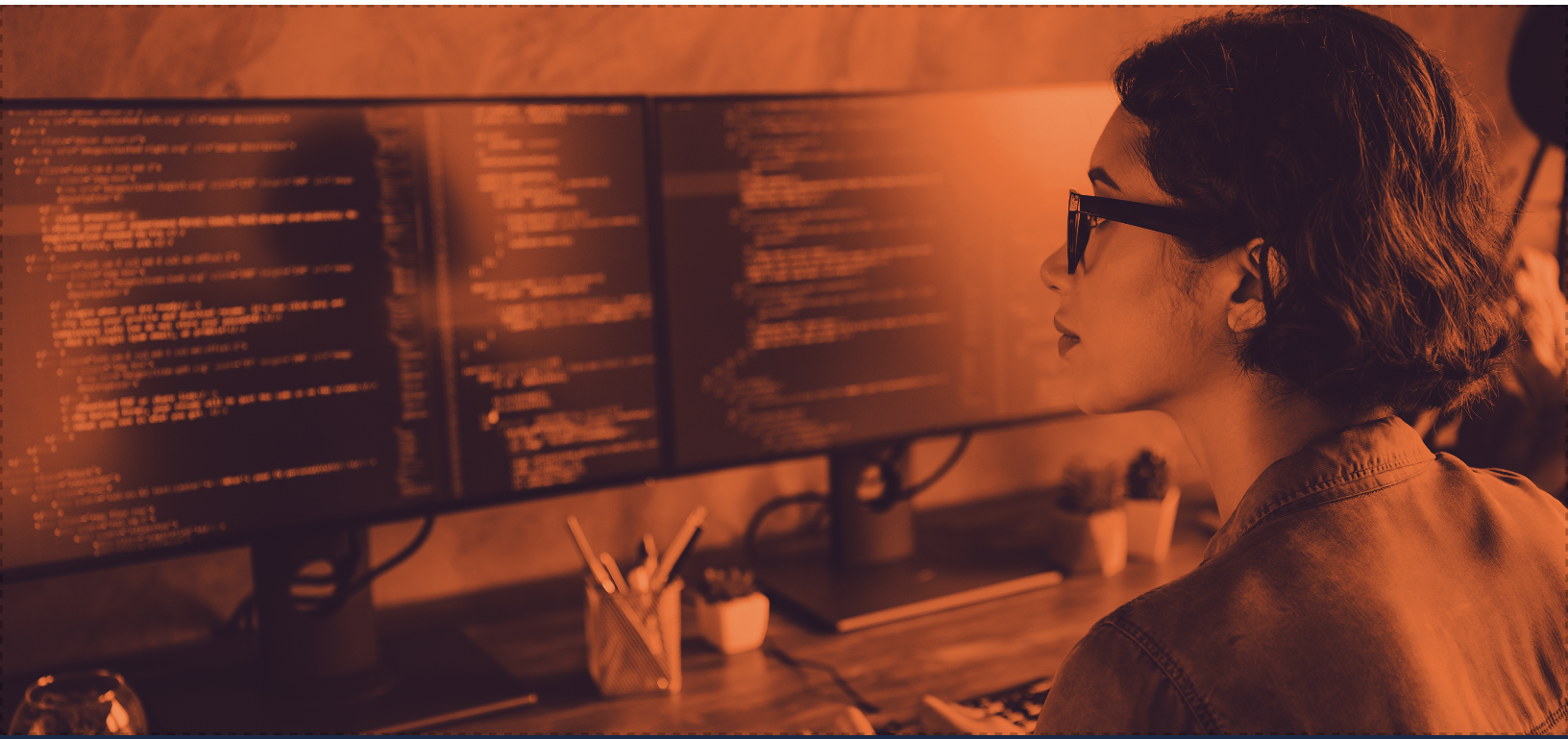


querychat, paired with Posit Connect and Snowflake Cortex, turns a viewer's plain-English question into Snowflake SQL using your team's own semantic model, and renders the answer as a table or chart. It ships once, as an interactive data product, instead of a dashboard you keep rebuilding on request.

03 | Where you write the code

Posit Workbench gives you RStudio, VS Code, Jupyter, and Positron in one place, so you pick the IDE that fits the task instead of switching platforms to switch languages. When you enable Posit Assistant inside Positron, it configures Snowflake Cortex as its LLM backend, using authorization you already have, as long as your Snowflake role has Cortex models enabled, with no API key for you to request, generate, or rotate.

Because Posit Assistant runs on Cortex, your interactions follow the governance rules Snowflake already has in place, whether your team runs the Posit Team Native App or Posit Connected Apps. Cortex gives Posit Assistant access to Claude and GPT models, so you pick the model for the task without opening a new vendor relationship. (More on Cortex integration.)



Continued from **03 | Where you write the code**



Posit Assistant grounds each suggestion in the actual code, data, and environment behind it, so a reviewer can rerun it and confirm the result. When your team has defined a Snowflake semantic model, Assistant reads it too, so suggested joins and field names match the definitions your team agreed on instead of a guess. (More on the Snowflake skill)



When the work is ready, you publish it from your IDE to Posit Connect in one click: a Shiny or Streamlit app, a Quarto report, an API, or a scheduled job. On the Native App, Connect runs it inside Snowflake, and each viewer sees the data their own Snowflake role allows, so you don't rebuild it for another stack or wait on an engineering handoff.



On the native app, idle Posit Workbench sessions auto-suspend after a timeout that your admin sets, so you're not paying for compute you forgot to close. Platform updates land automatically in scheduled maintenance windows, so you're never the one tracking down a patch.

Posit Assistant also checks in before steps that are expensive or risky. In this demo, it recommended a hyperparameter sweep, and the user stopped it to validate on a hold-out set before spending the compute.

04 | Proof from teams like yours

Brilliant Earth

From an engineering handoff to one click

Brilliant Earth's data science team had no governed way to take a model from a notebook into production. Marketing wanted campaign answers that spanned disconnected data sources, and shipping anything to production meant an engineering handoff and a wait.

The team rebuilt the workflow inside Snowflake: they develop in Positron and deploy to Posit Connect, in the same security perimeter as the data. Their marketing mix model runs as an interactive Streamlit app that stakeholders can open. The team's weekly reports, which used to take a full day each Monday, now run on their own, using the Anthropic and Cortex APIs. The team ships in hours what used to take weeks.

Read the full story



> The infrastructure now lets us move at the speed the business requires, while keeping insights governed and reproducible.
>
> — MATTHIAS MUELLER, DIRECTOR, DATA SCIENCE, BRILLIANT EARTH



Trusted across regulated and high scale industries

Posit is an Elite Snowflake Partner, with joint Snowflake customers in insurance (Beneva), sports (NHL), retail (Brilliant Earth), healthcare and life sciences (NMDP), and more.

Continued from **04 | Proof from teams like yours**

Pinterest
Analysis SQL alone couldn't do

Pinterest's People Insights and Analytics team works with sensitive employee data they weren't allowed to download, and had no centralized environment for their R users. Preparing a single dataset took five separate steps before an analyst could start analyzing it.

Posit Workbench, running in the Posit Team Native App inside Snowflake, gave analysts a place to combine SQL and R against the same governed data. With SQL and R together, the team ran regression-based analysis on unstructured text, something SQL alone couldn't do, reading 30,000 employee survey comments instead of running keyword counts on them.

Read the full story

> The adoption of Posit Workbench in Snowflake has unlocked new levels of efficiency for the People Insights and Analytics Team while reinforcing our commitment to secure, scalable data practices.
>
> **– TREVOR FRY, LEAD DATA ANALYST, PINTEREST**



Continued from **04 | Proof from teams like yours**

NHL
From a CSV relay to one governed workspace

The NHL's three-person data science team pieced fan analytics together across disconnected tools: download CSVs locally, connect to a hosted R environment through a VM, run the code, export results back out as CSVs, then upload those into a BI tool for anyone else to see. By the time an insight reached a stakeholder, the data behind it was stale.

The team runs its everyday modeling inside Posit Workbench Advanced, deployed with the Posit Team Native App: one governed workspace where R and Python run against the league's fan data inside Snowflake. The team doesn't hand off files between systems or wait on a CSV export to see the result.

Read the full story

> We have a streamlined process for analyzing and modeling these spikes and their impacts as the data becomes available.
>
> — JAMIE SHIVE, DIRECTOR OF DATA SCIENCE, NHL



05 | What changes for you, day to day

You update models on a schedule that your team sets.

You work under your own Snowflake role in Workbench and Connect, so you see the data you've been granted, with no separate access layer to request.

You write R, Python, and SQL in the same IDE, against the same data, without moving any of it.

You get an AI assistant (Posit Assistant) tuned to your codebase without opening a new security review, because it runs inside the boundary around your data.

Your model ships as a finished data product (an API, a Shiny or Streamlit app, a scheduled report) without an engineering handoff.

Usage of what you built shows up under your name, not a shared login, so the people deciding what gets built next can see that it's yours.

Code, data connections, and environment ship together, so a teammate can continue what you built without relearning it, whether you're out sick or have moved to a new project.



06 | Go further

Start your free trial: download the Native App from the Snowflake Marketplace. It's a 30-day self-serve trial of Workbench and Connect together. Not ready to install? Learn more first.



Watch the demo: an end-to-end workflow running natively inside Snowflake.



Posit + Snowflake customer stories



Quickstart: analyze data with R using Posit Workbench + Snowflake



Quickstart: analyze data with Python using Posit Workbench + Snowflake



Quickstart: run tidymodels prediction workflows inside Snowflake with orbital



Quickstart: build and deploy an interactive dashboard with Posit Team and Cortex



Talk to a Posit expert about your data and use case.

About Posit and Snowflake

Snowflake provides the AI Data Cloud, a global network where organizations mobilize data at scale. Its Native App Framework and Snowpark Container Services let partners like Posit run their applications inside a customer's own Snowflake account, giving regulated industries a single governance boundary for both data and the tools that work on it.

Posit (creators of RStudio) is the data science platform for R and Python, giving organizations in financial services, insurance, life sciences, government, and academia the tools to develop, deploy, and govern analytic work. Snowflake and Posit support work like the National Marrow Donor Program, which analyzes a registry of 42 million potential donors to find lifesaving matches.

Snowflake named Posit its 2026 Global Product Innovation Partner of the Year and 2026 AI/ML Data Science Tooling Partner of the Year, and named Posit a "One to Watch" in its 2026 Modern Marketing Data Stack report.

