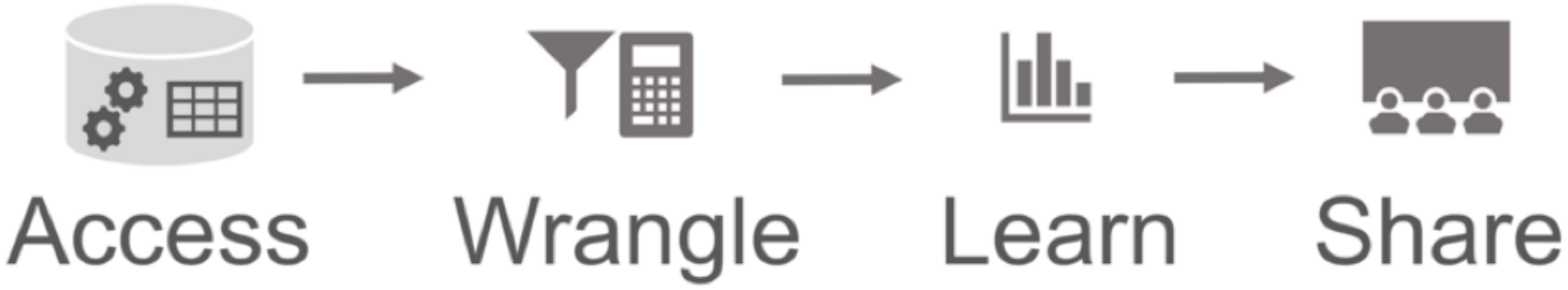
Databases with RStudio

Creating Shiny dashboards performant at scale

Using Databases with R

- You can use R with almost any data source



- But historically it's not always easy

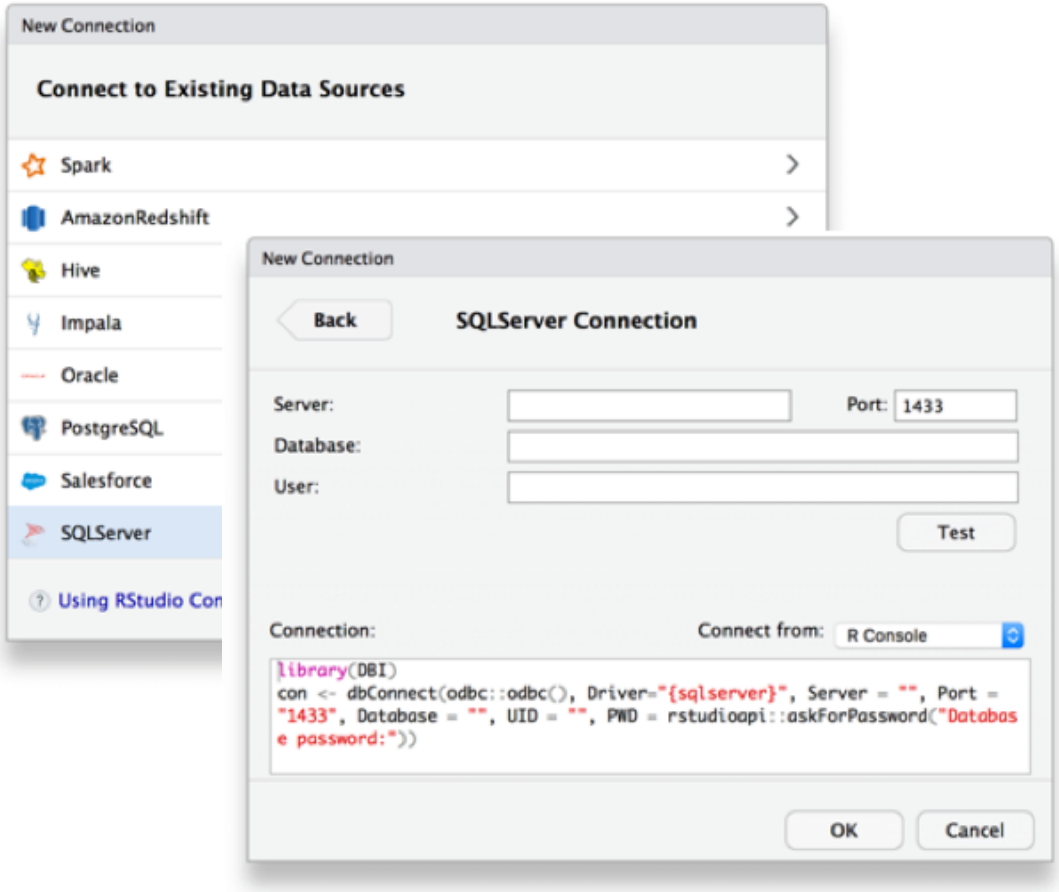
Obstacles when using databases with R

Obstacles	Examples
Hard to establish connections and access data	R users often use one tool for SQL and another tool for R
No consistent toolset or language	R users interchange SQL and R to talk to the exact same data
Setup and configuration is confusing	R users unable to access their databases
No centralized place to get information	R users get their information from package manuals

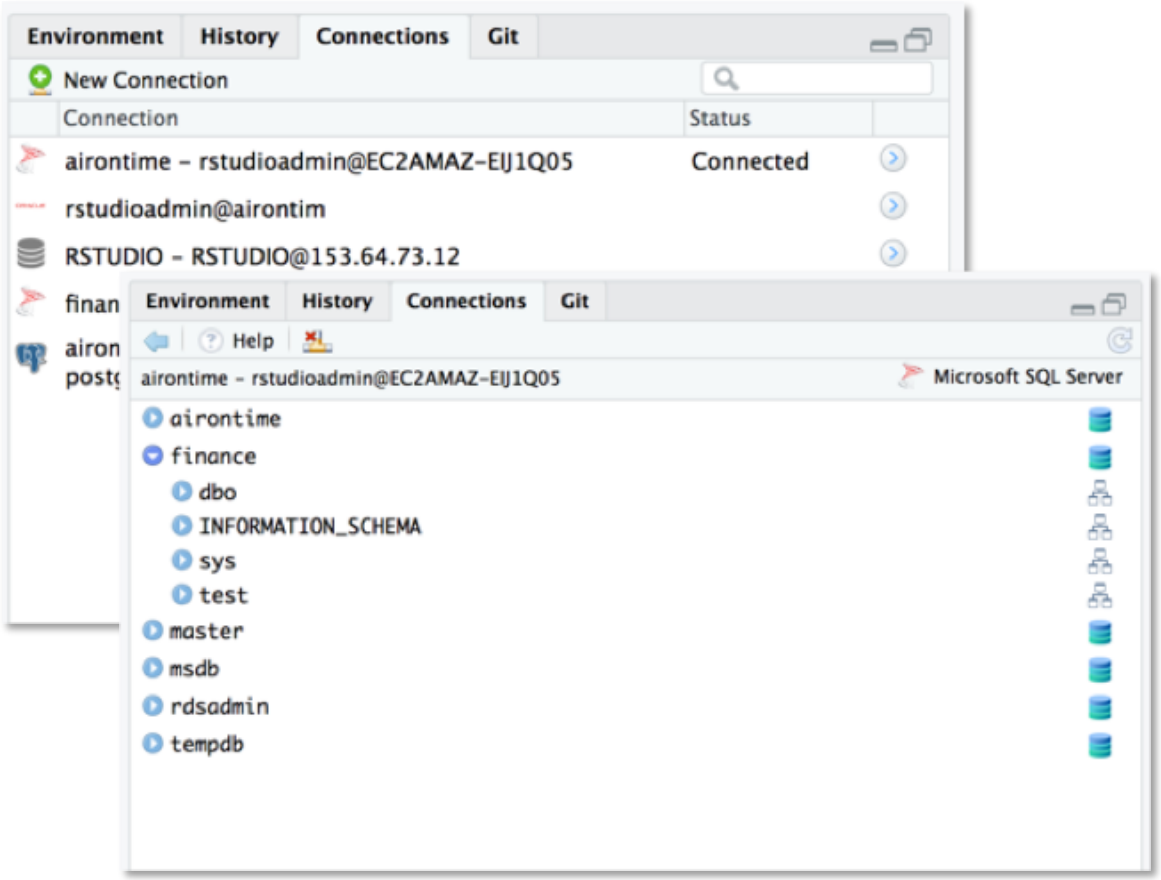
What’s new from RStudio

RStudio version 1.1 new features

Connection wizard

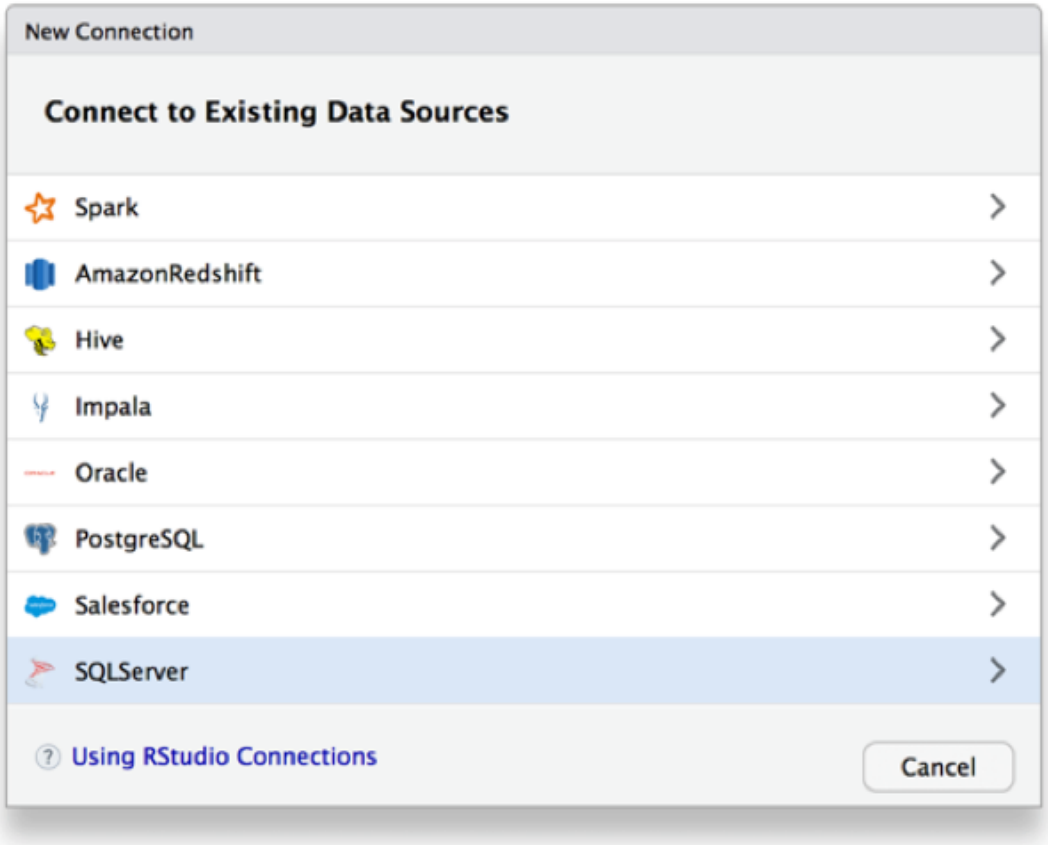


Connections tab



Connection methods

1. ODBC drivers
2. Data source names (DSN)
3. R Snippets
4. R Packages



Improved R Packages

- DBI
  - Standard database interface for R
  - Query your database with SQL
  - Do database operations
- dplyr
  - Query your database with R code
  - Generalized SQL backend
  - dbplyr [New!] translates dplyr syntax to SQL for specific databases

dplyr syntax

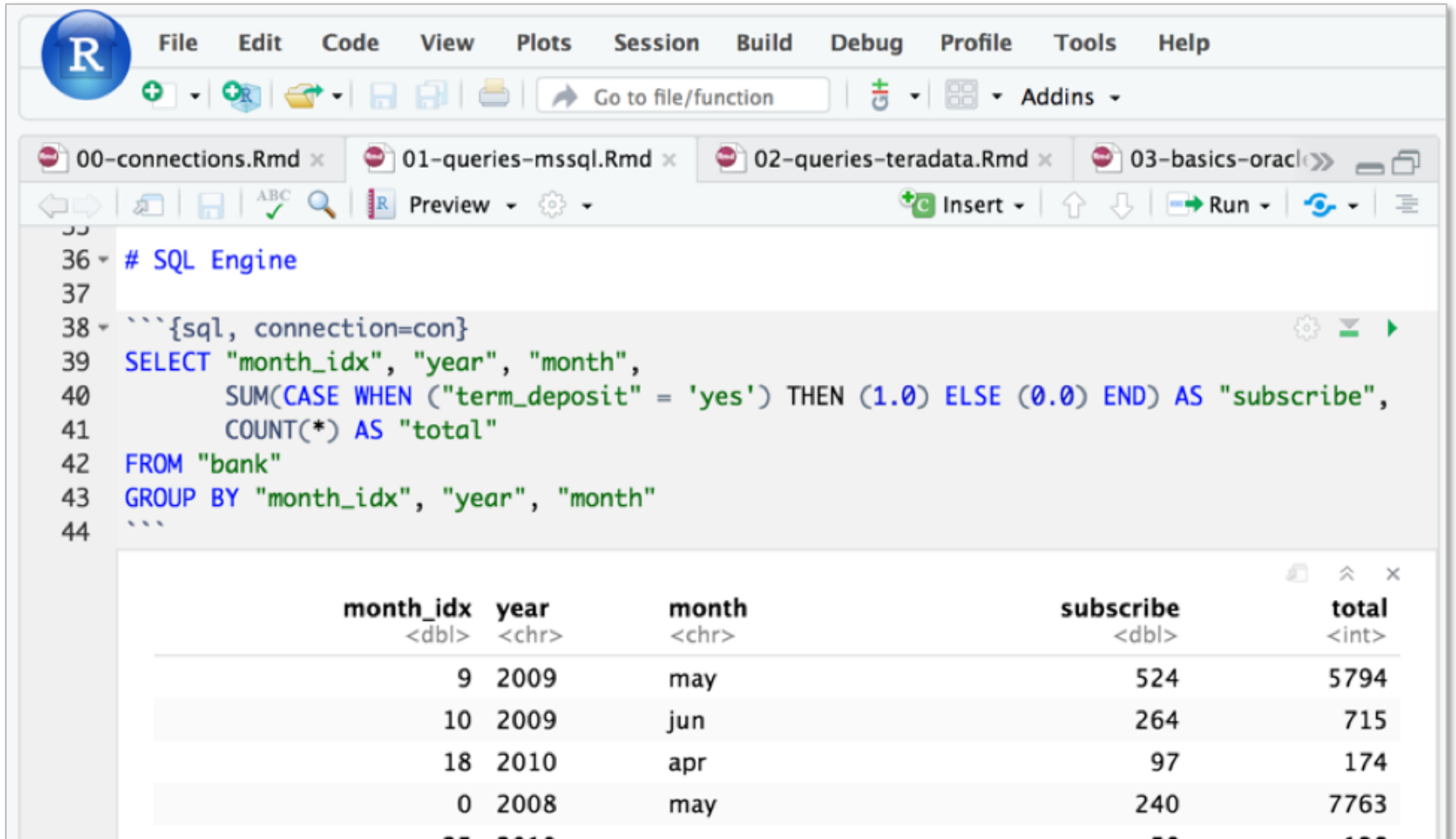
Use dplyr verbs

```
flights %>%
  group_by(name) %>%
  tally()
```

Write SQL statements

```
SELECT "name",
COUNT(*) AS "n"
FROM "vwFlights"
GROUP BY "name"
```

SQL Language Engine for:
R Markdown and R Notebooks



New R package for connecting with ODBC

odbc [New!]

- Connects R to any data source via ODBC
- Bring your own driver or use RStudio Professional Drivers
- DBI compliant
- Actively developed
- Designed for performance

Database connection methods

odbc [New!]

- Connects R to any data source via ODBC
- Bring your own driver or use RStudio Professional Drivers
- DBI compliant
- Actively developed
- Designed for performance

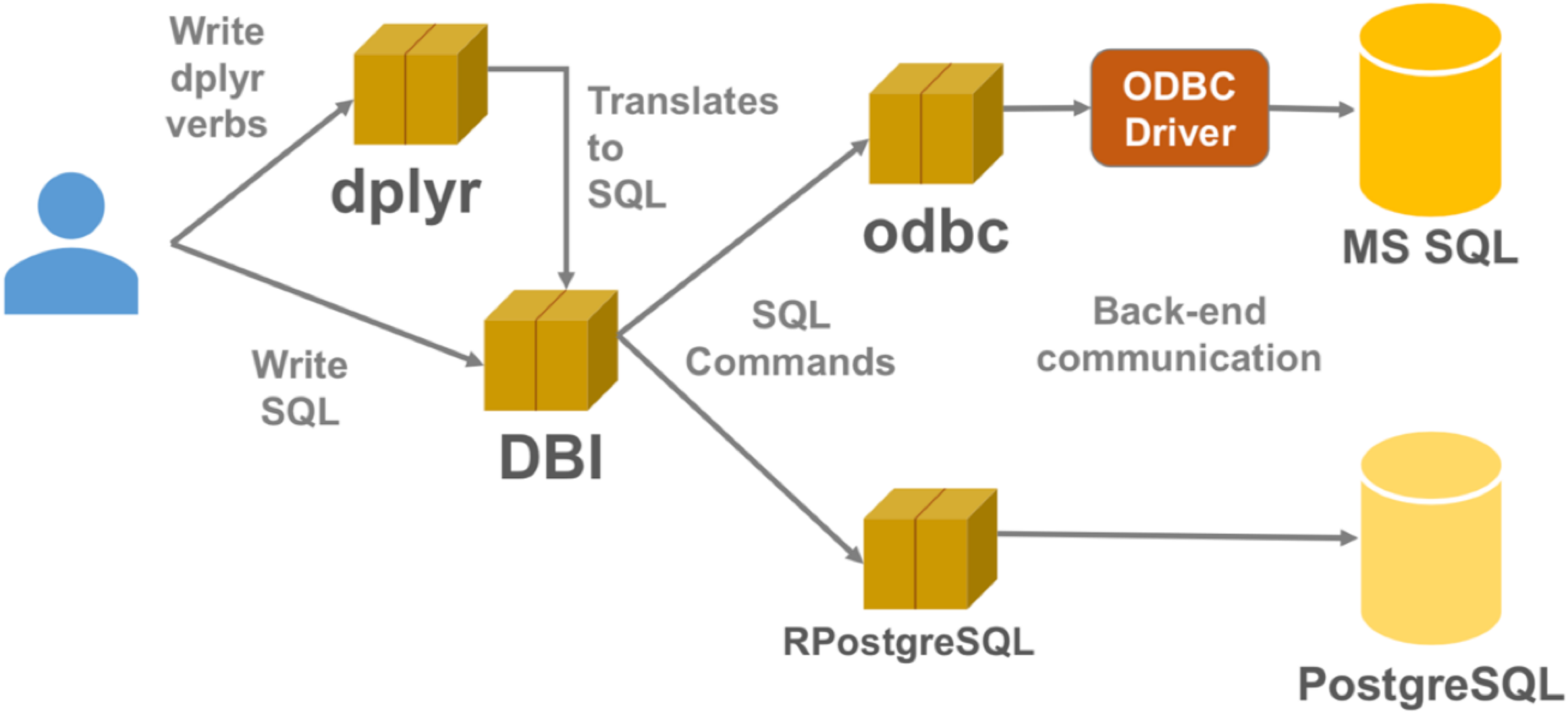
R Packages

- RSQLite
- rpostgres
- bigrquery
- […]

RJDBC

- Connect to any data source via JDBC
- Requires JDBC driver
- Requires Java

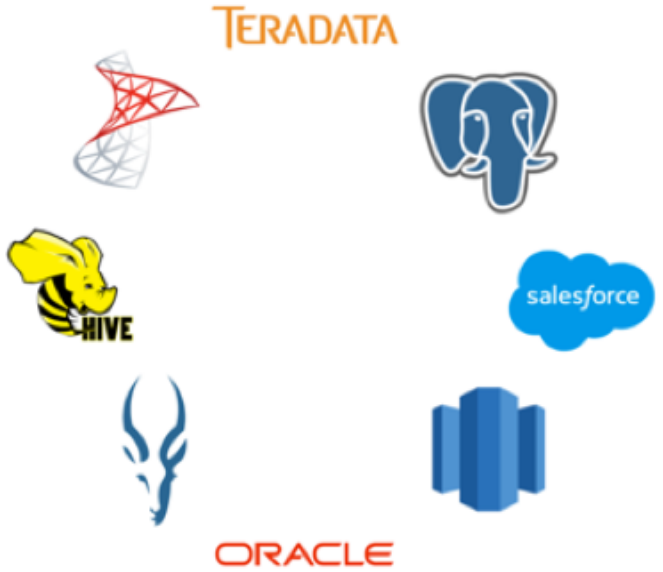
Establish a database connection with RStudio



RStudio Professional Drivers

RStudio Professional Drivers

- Coverage
  - Connect to many popular databases
- Convenience
  - Easy to install and configure
- Consistent
  - Use with all RStudio pro products
- Professional
  - Supported for production level work

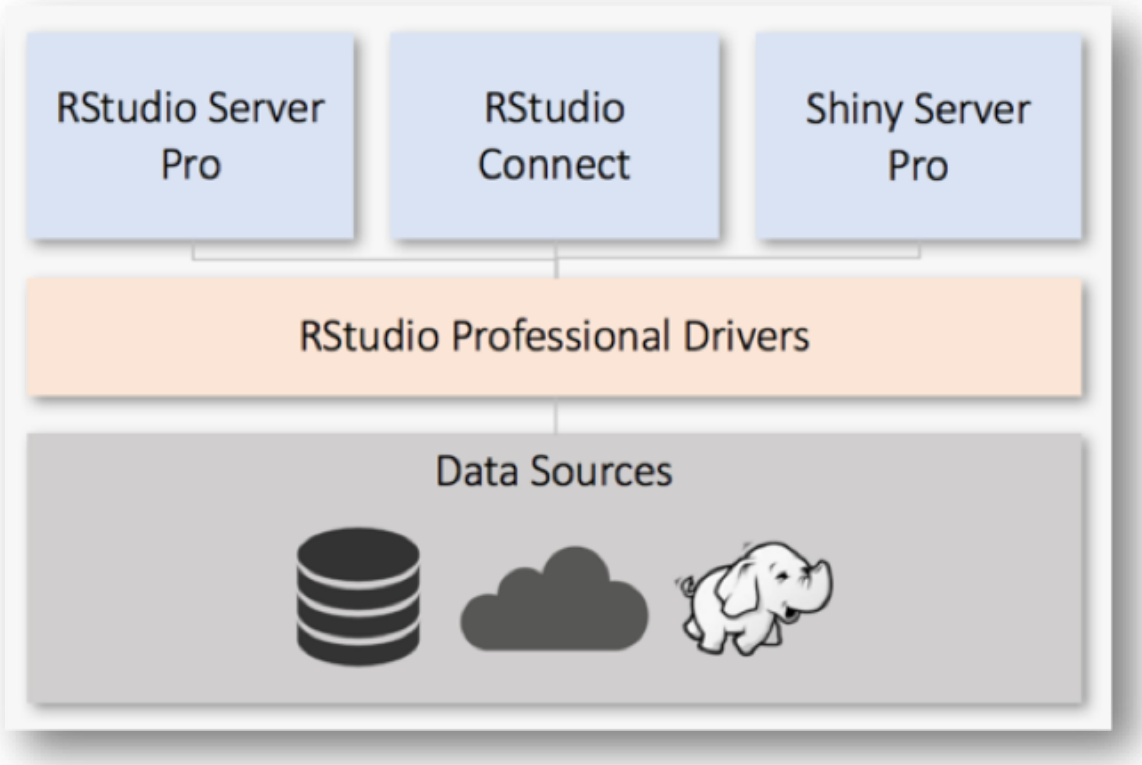


Teradata
PostgreSQL
Salesforce
Redshift
Oracle
Impala
Hive
SQL Server

RStudio Professional Drivers in Production

- Using R and databases in production
  - Shiny applications
  - R Markdown reports

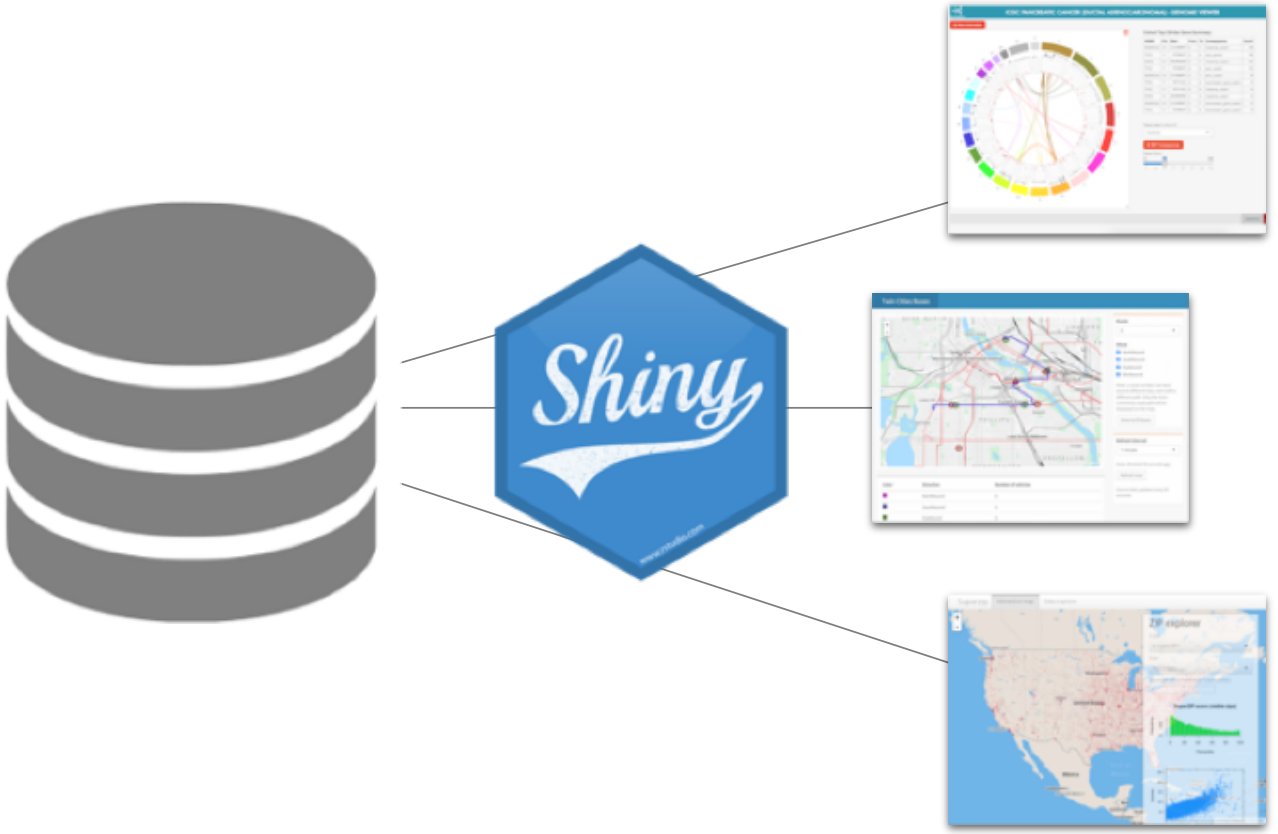
- Benefits of pro drivers
  - Support for the entire workflow
  - Professional license
  - Consistent dev/test/prod environment



Databases with Shiny

Shiny applications and databases

- Develop and deploy Shiny applications that depend on databases
- Use RStudio professional products when hosting Shiny in production systems



Push computation to the database

Database Demos

- Using R with Databases
  - A deep dive into connections and dplyr
- Using Shiny with Databases
  - Simple Shiny app
  - Extracting data into a dashboard
  - Pushing compute into a dashboard
- Scoring data
  - Using API's
  - Using R and SQL

- https://github.com/nwstephens/databases-and-shiny

Learn more at db.rstudio.com