Pins: Three Ways

Ian Lyttle

2022-08-04

Table of contents

Preface

The purpose of this book is to make a brief demonstration of the pins package using R and Python, and to imagine how it might be used with JavaScript.

Pins helps you manage sharing data with yourself, others, or even CI processes. There are two levels of abstraction:

- **pin**: a "thing" to be shared as a file. It could be a data frame, a model, a nested list (dictionary, object). If it can be serialized to a file, it can be pinned. Some serializations, such as CSV, JSON, and arrow, are common to multiple languages (R, Python, JavaScript), so can be used for cross-language collaboration. Other serializations are specific to a language (pickle for Python, rds for R).
- **board**: a collection of pins hosted at a "place". A board could be hosted at Azure Blob Storage, an Amazon S3 Bucket, RStudio (soon to be Posit) Connect, a local filesystem, a remote URL, …

Rest of the book

In the rest of the book I (plan to):

- use R to:
  - create a board.
  - write a data frame as a pin, using the arrow format.
  - read the pin into a data frame.
- use Python to:
  - read the data-frame pin written using R.
  - write a pandas data-frame as a pin using the arrow format.
- use JavaScript to:
  - read the data-frame pins written using R and Python, using arquero, which supports the arrow format.

Perspectives

I have some ideas for the conculsions I might come to in the course of writing the rest of this material. That said, I'll want to make some *actual observations* before calling for any action. I'll update this section as I go.