

Organizing Your Work with RStudio Projects



Charlotte Wickham

DATA SCIENTIST AND EDUCATOR

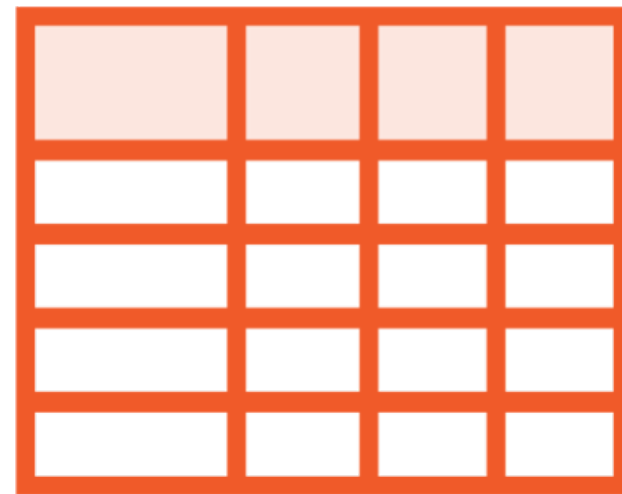
@cvwickham www.cwick.co.nz

Goal: Use Real Data for Report



summarize_sales.R

Uses small, fake
data



property_sales.csv

Real data



Multiple files

How to keep work
organized?

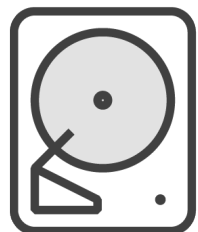
Disk



summarize_sales.R



property_sales.csv



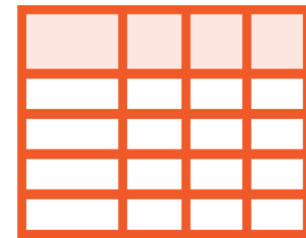
R Workspace

Unassigned objects

"Peaceful Glen", "Peaceful Glen"...

"Dwelling", "Dwelling", ...

229900, 349900, 235000, ...

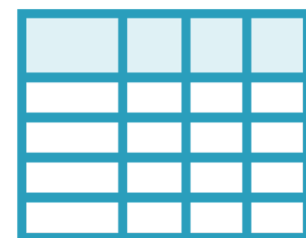


sales

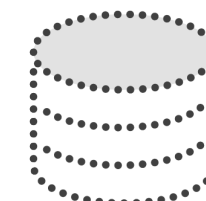
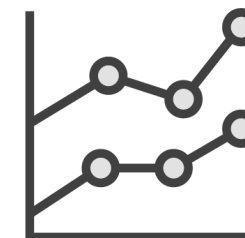
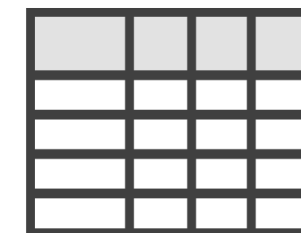
10

total_sales

40

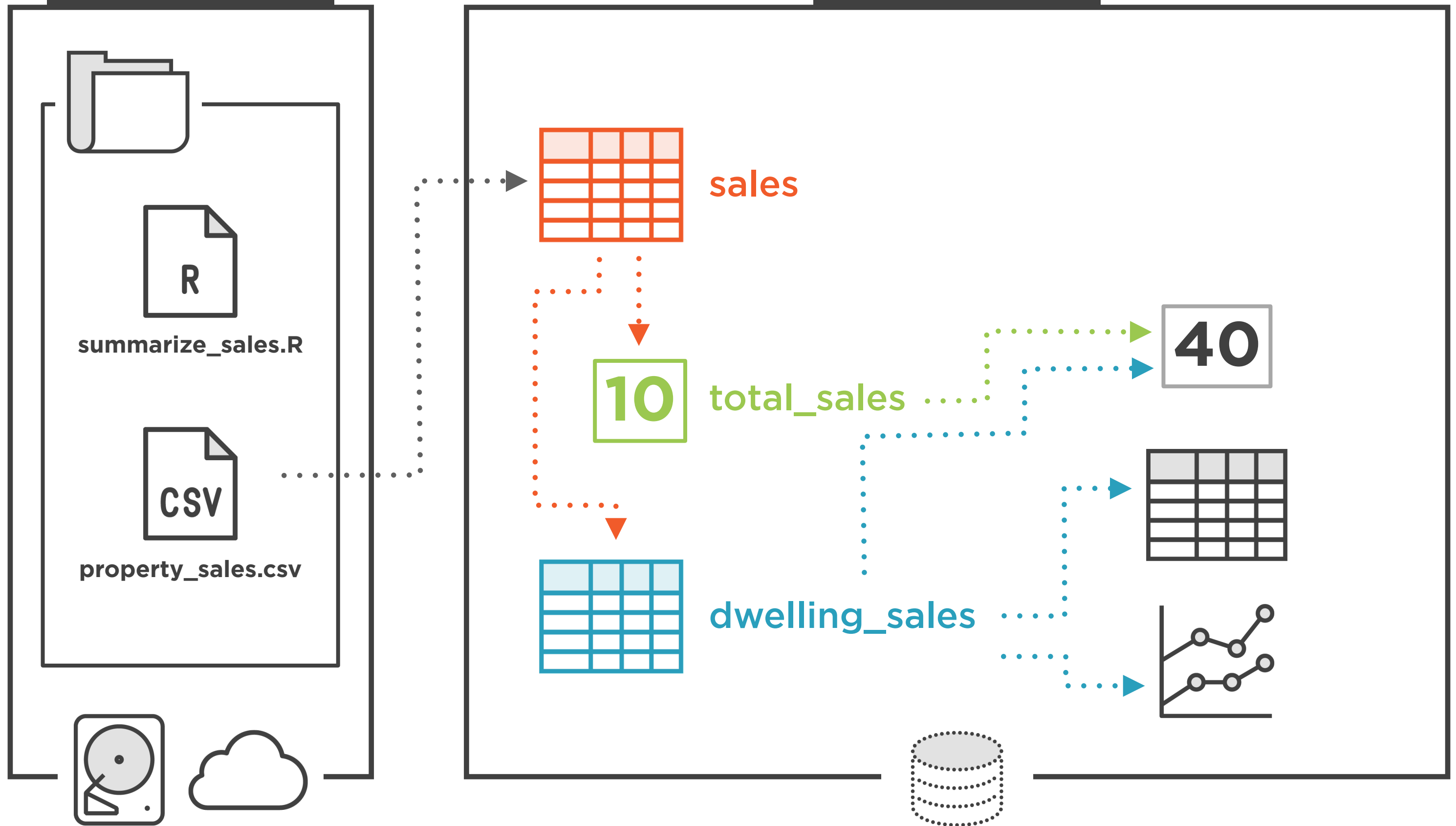


dwelling_sales



Disk

R Workspace



Overview

What is an RStudio project?

Creating an RStudio project

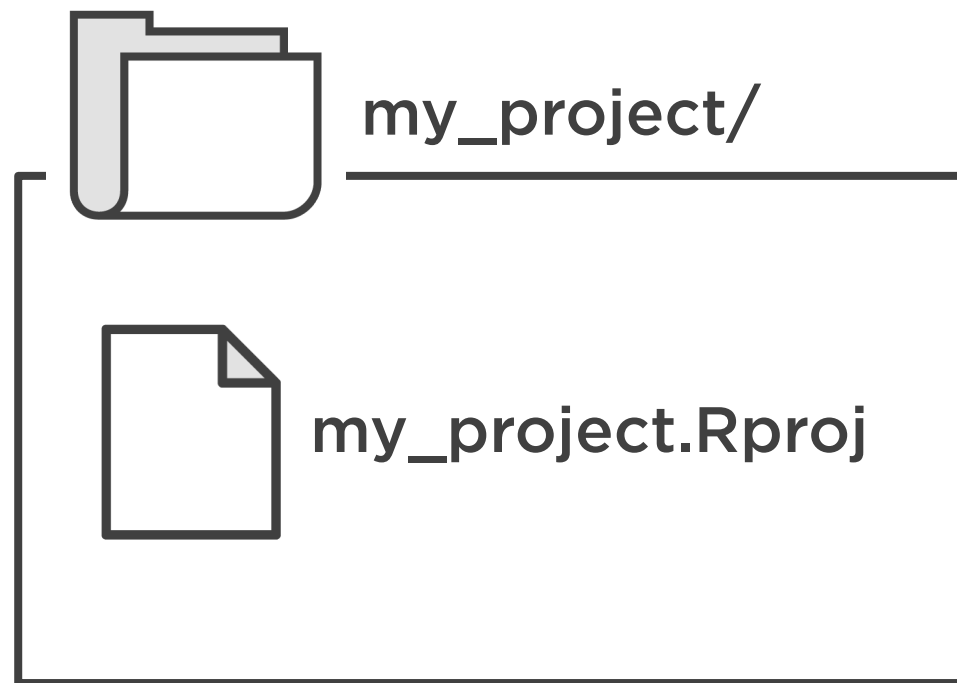
Adding files to an RStudio project

Referring to file in an R script

Reproducible projects

Creating a New RStudio Project

What is an RStudio Project?



**A directory,
with a special an RStudio project file**

Creating a new project, adds the .Rproj file

Manipulating Files in an RStudio Project



Operating System

File management tools



RStudio IDE

Files pane

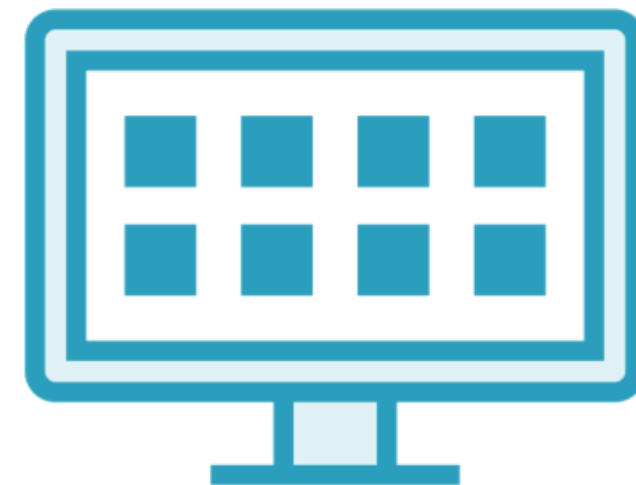
Closing and Opening RStudio Projects

Opening an RStudio Project



RStudio IDE

Project menu



Operating System

Double-click .Rproj file

Referring to Files in R

```
read_csv(path)
```

Import a CSV File

Location of file on disk described by `path`

```
read_csv("/Users/pluralsight/Documents/  
property_sales/data/property_sales.csv")
```

Absolute Paths

Location relative to top-most directory on disk

```
read_csv("/Users/pluralsight/Documents/  
property_sales/data/property_sales.csv")
```

Absolute Paths

Location relative to top-most directory on disk

```
read_csv("~/Documents/  
property_sales/  
data/property_sales.csv")
```

Absolute Paths

Location relative to top-most directory on disk

Shorthand for users home directory ~/

```
read_csv("~/Documents/  
property_sales/  
data/property_sales.csv")
```

Absolute Paths

Location relative to top-most directory on disk

Shorthand for users home directory ~/


```
read_csv("~/Documents/  
property_sales/  
data/property_sales.csv")
```

Relative Paths

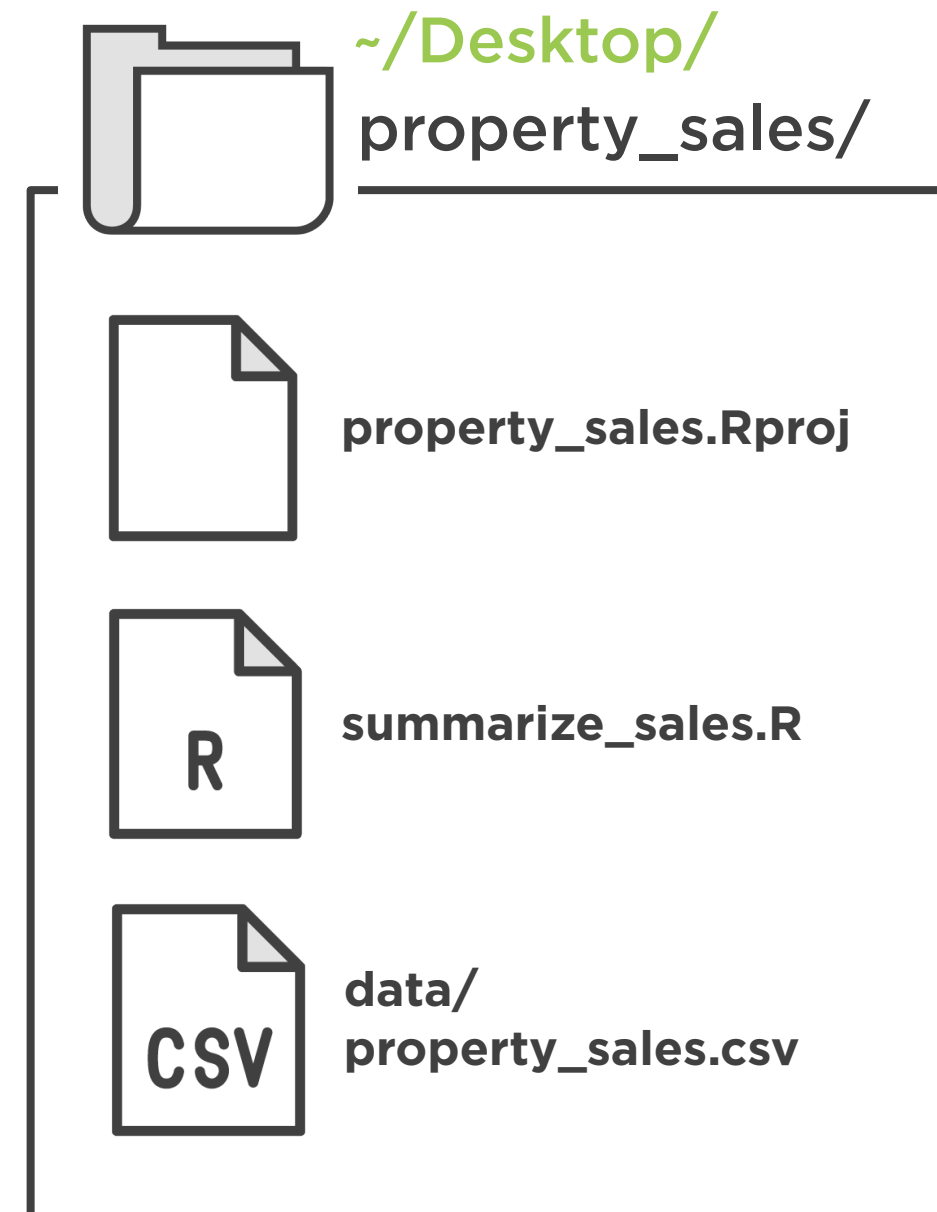
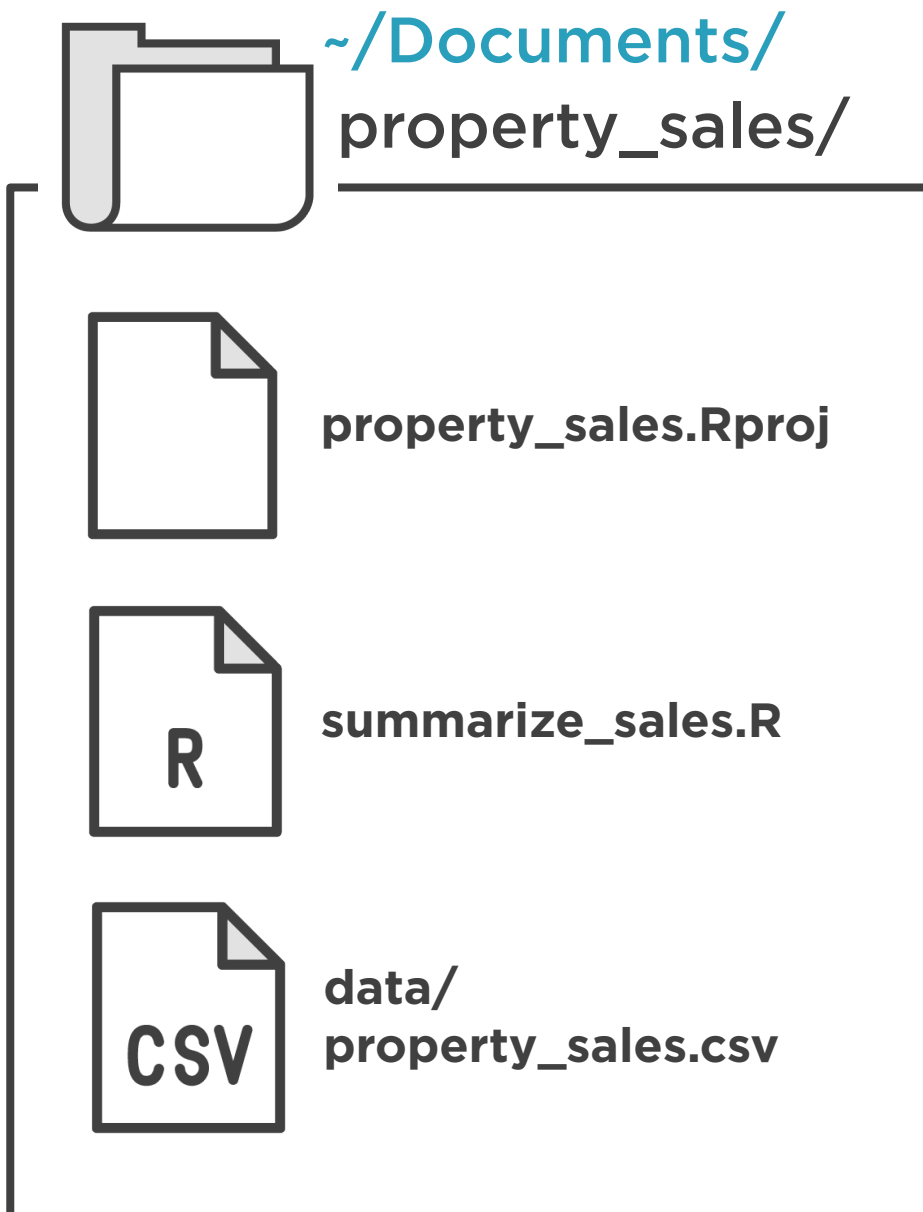
Location relative to **current working directory**

```
read_csv("data/property_sales.csv")
```

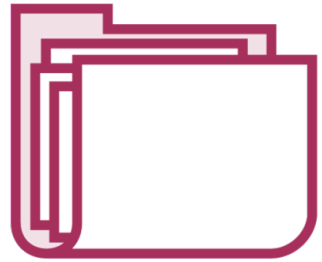
Relative Paths

Location relative to current working directory

Referring to Files Reproducibly in R



Referring to Files Reproducibly



Use projects to set current working directory



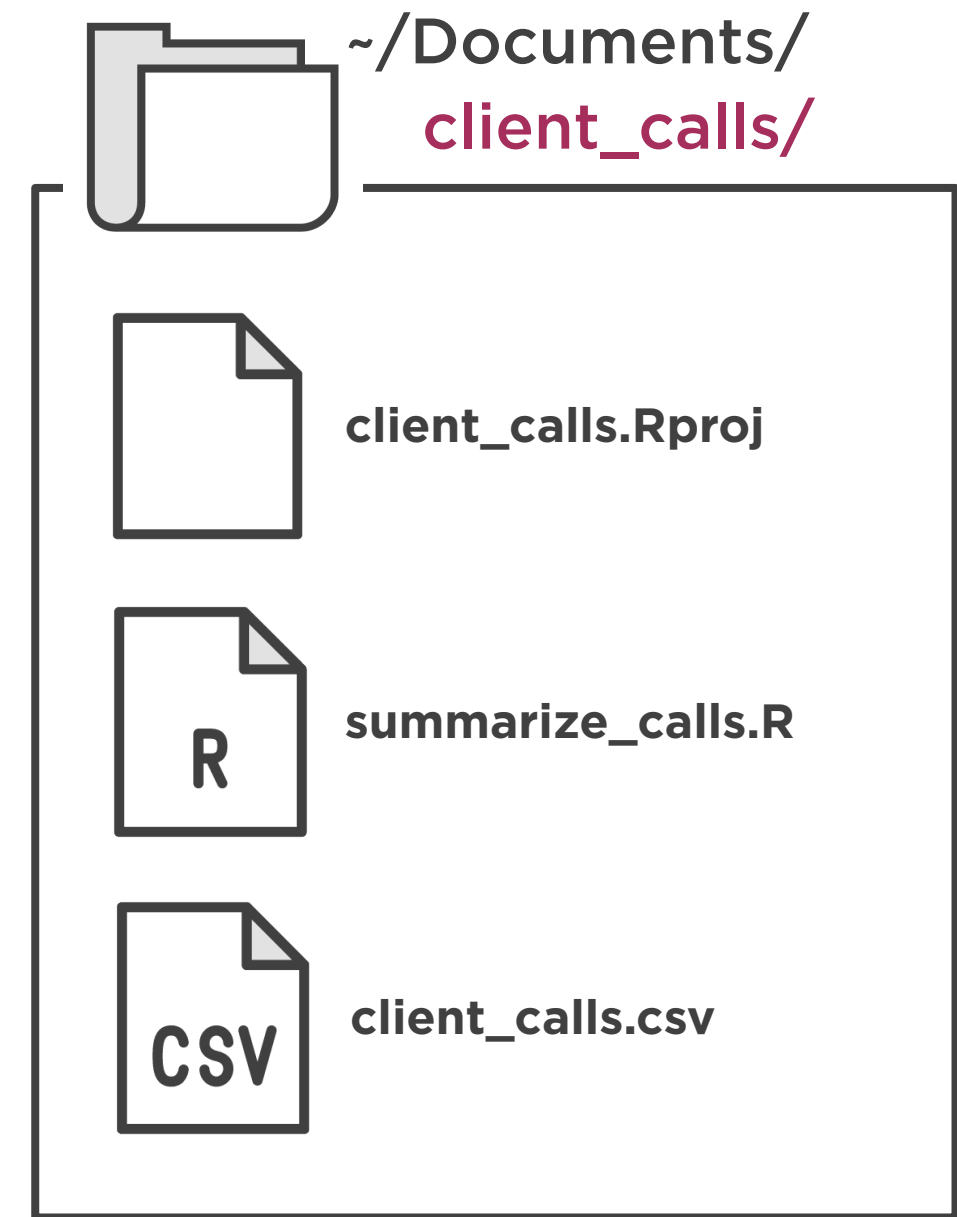
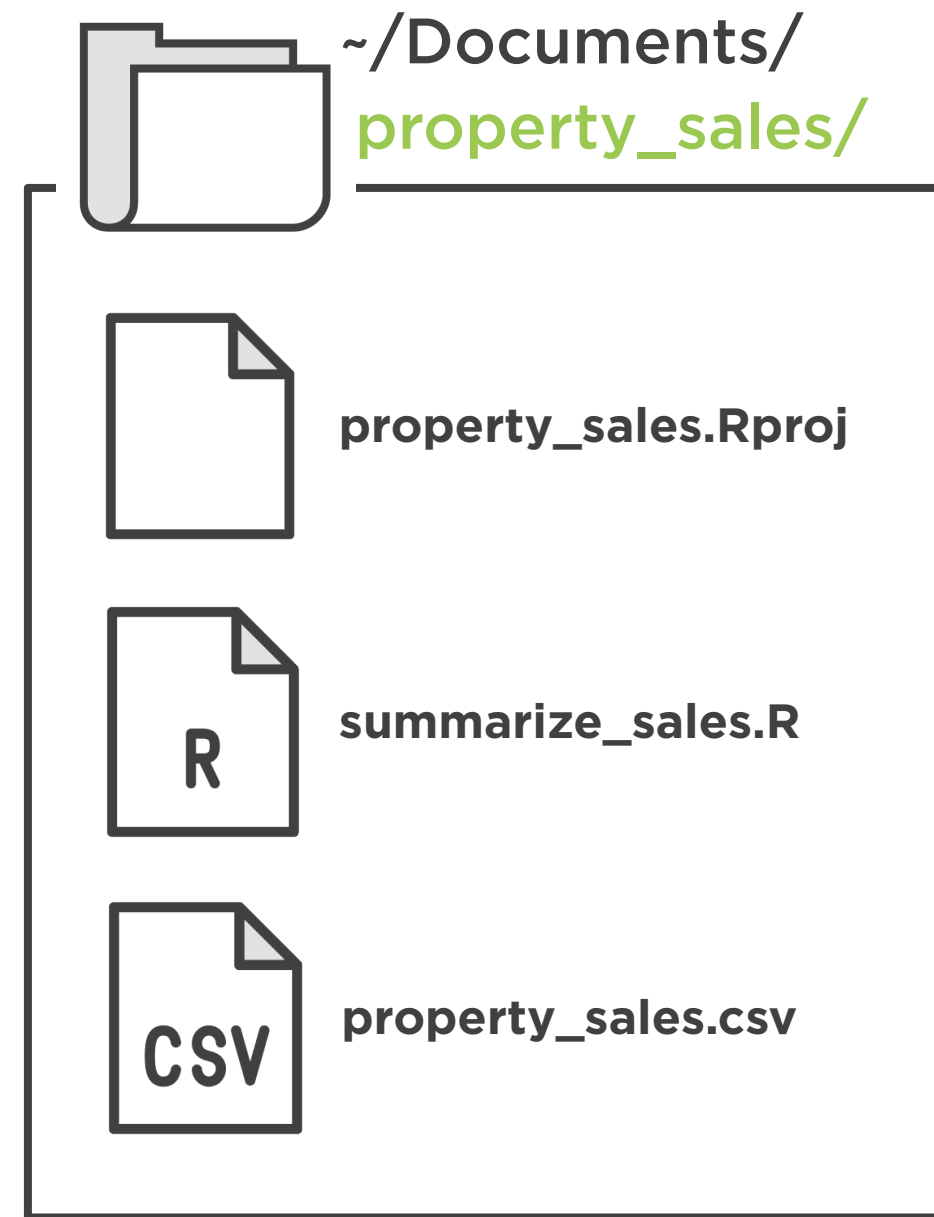
Use relative paths to refer to files



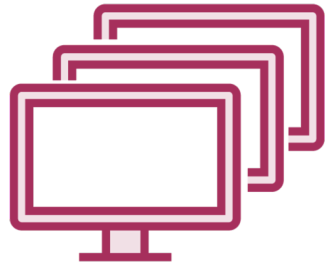
Together, project is reproducible regardless of location

Keeping Your Work on Projects Separate

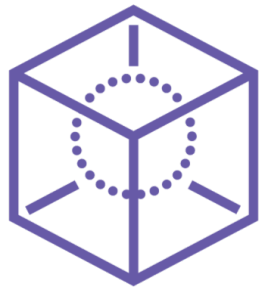
Use Separate Directories for Separate Projects



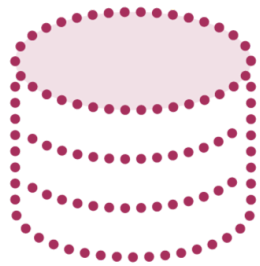
Keeping Your Work on Projects Separate



You can have more than one RStudio project open



R sessions are independent

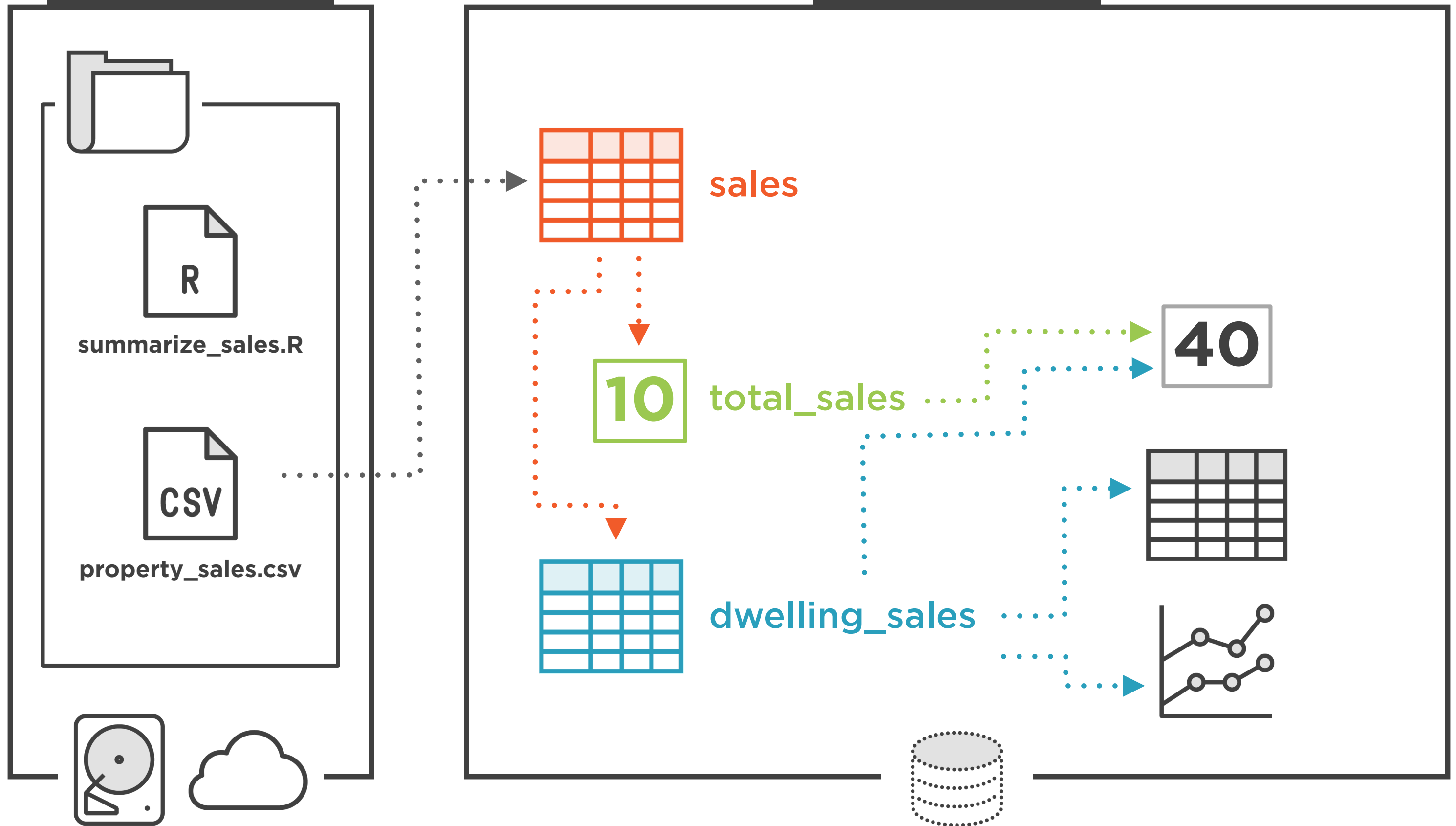


Workspace in one session doesn't effect any others

Saving Files Back to Disk

Disk

R Workspace



```
write_csv(x, path)
```

Save a CSV File

Data table **x**

Location of file on disk described by **path**

```
ggsave(filename)
```

Save a ggplot2 Plot

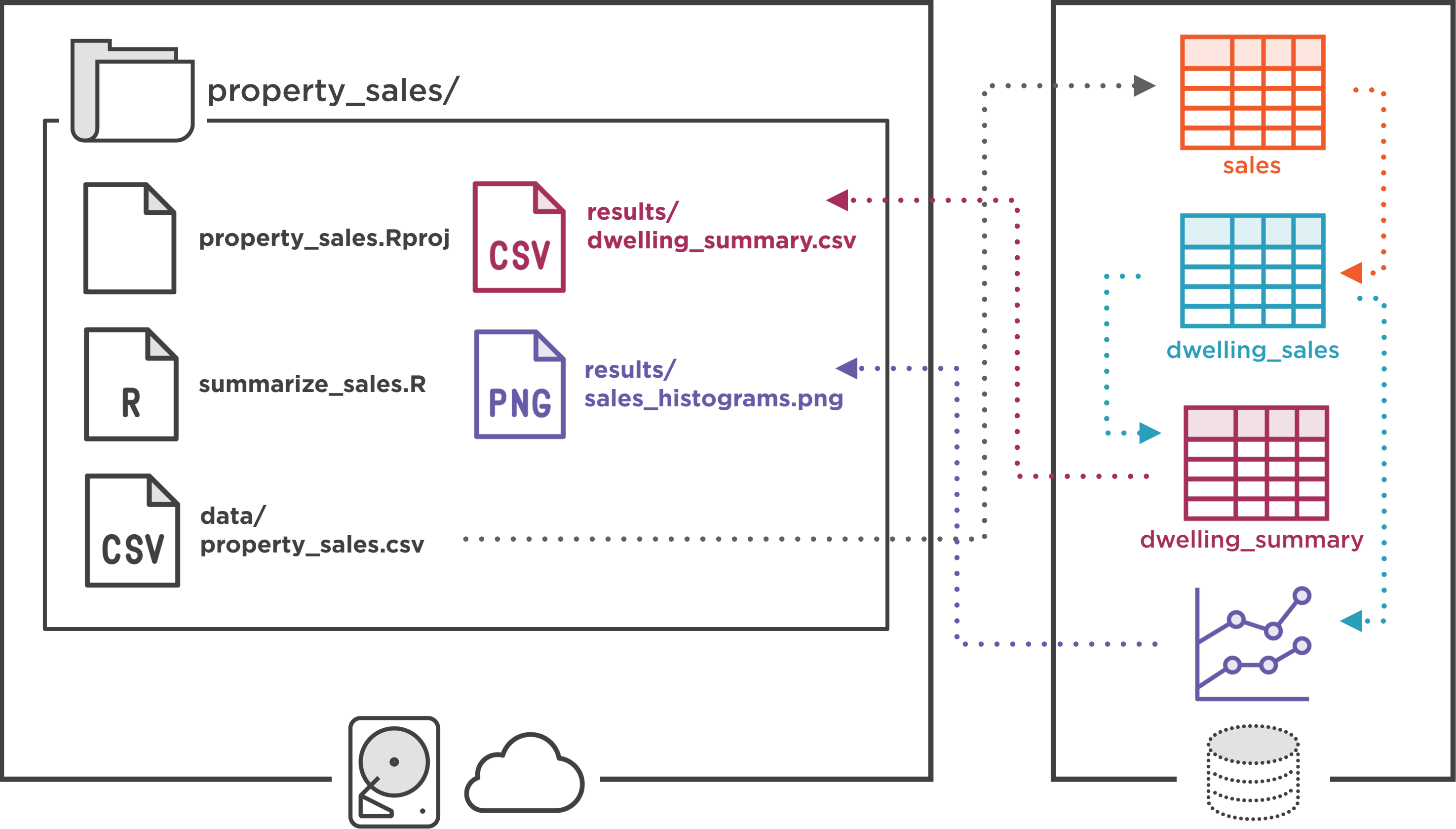
Location of file on disk described by **filename**

Saves last plot created

Wrapping Up

Disk

R Workspace



Summary

Making a directory an RStudio project:

- Easy to open
- Sets working directory
- Encourages relative file paths
- Shareable

Up Next:

**Communicating and Documenting Your
Analysis with R Markdown**