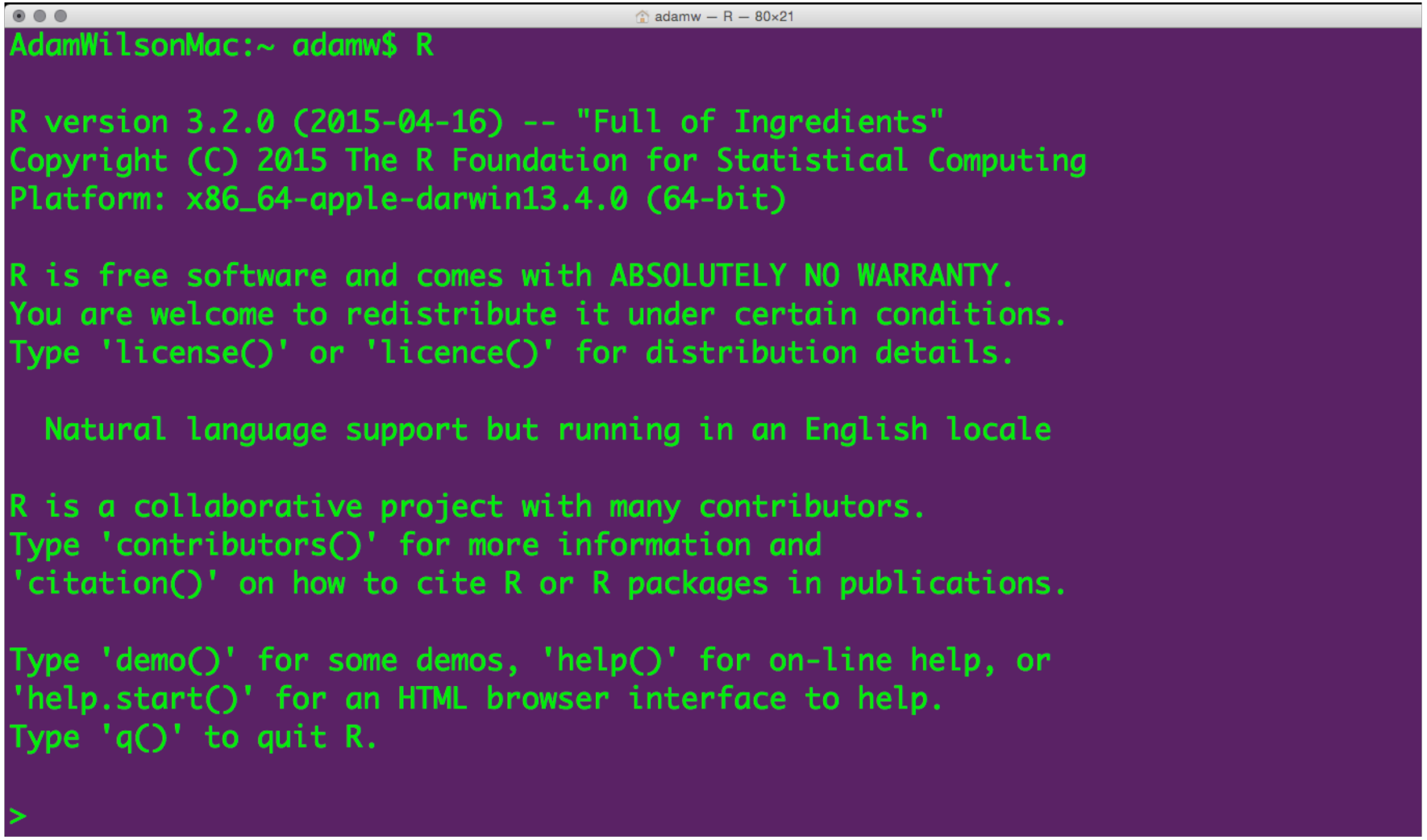
GEO 503: Spatial Data Science (with R)



Course Objectives

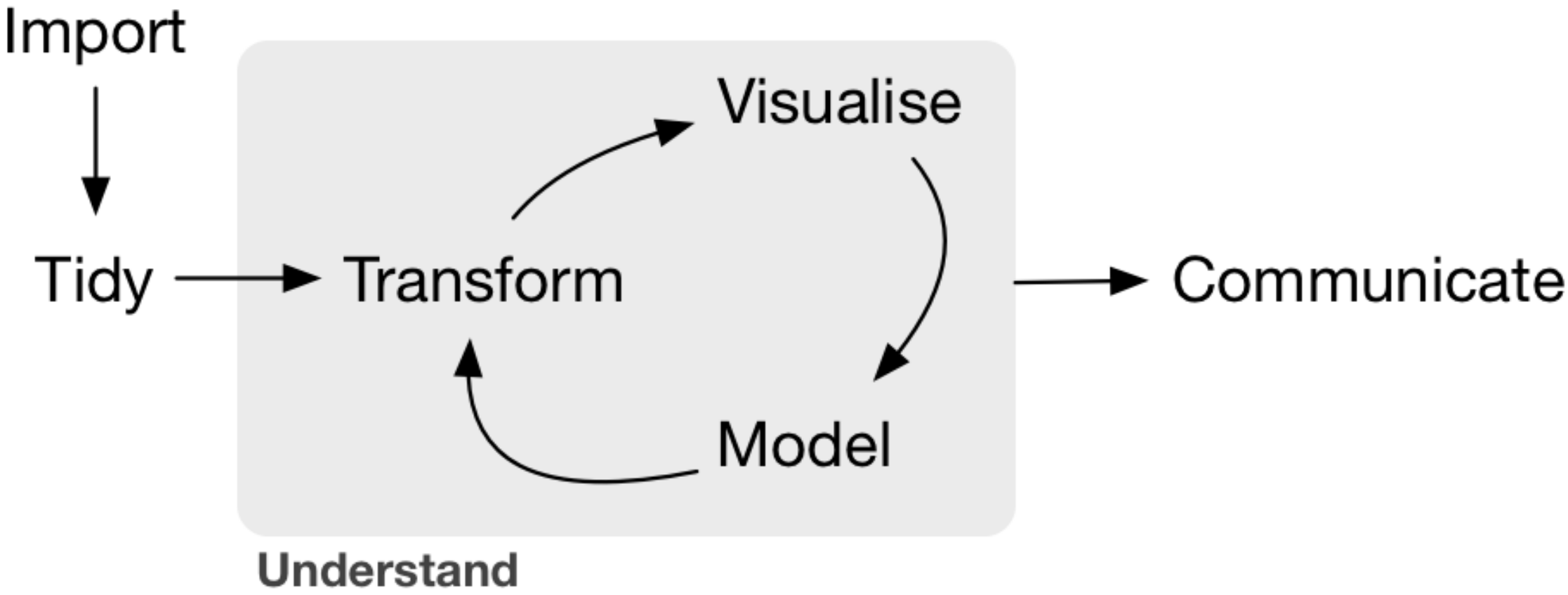
4 Learning Objectives

- Become familiar with R programming language
- Learn to code geospatial analyses
- Learn to develop custom data visualization (especially spatial)
- Learn to develop reproducible research workflows

This course is NOT

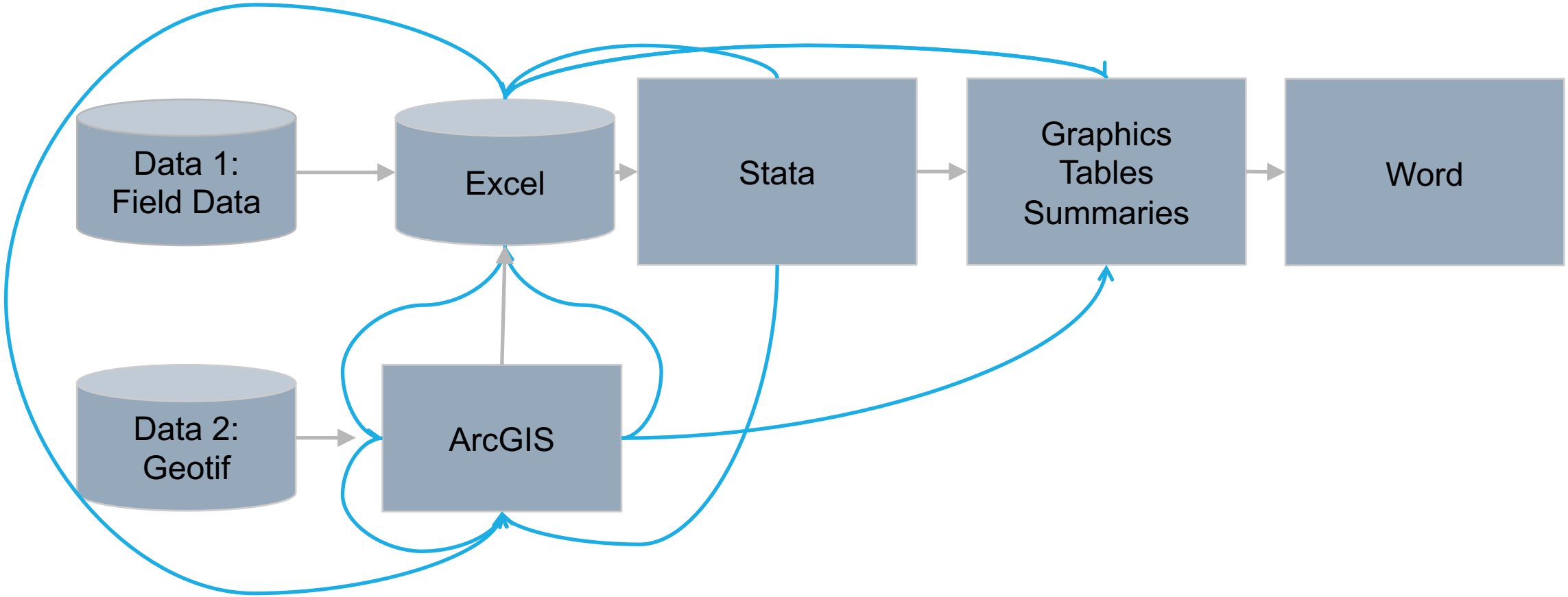
- A statistics course (see GEO 505, etc.).
- We will focus on workflow and methods (‘how’ not ‘why’)

What is Data Science?

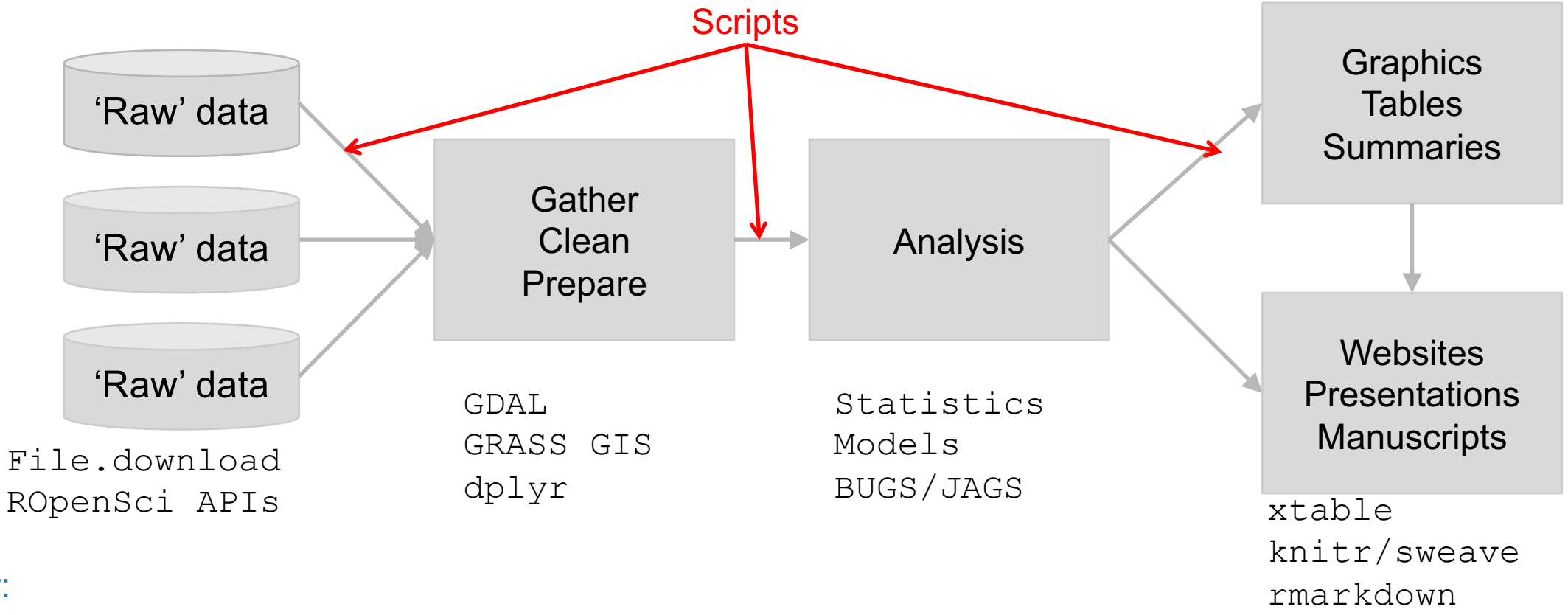


--Grolemund & Wickham, R for Data Science, O'Reilly 2016

Typical GUI Workflow



Organized and repeatable workflow (and some example commands)



Advisor:

- *I've updated the field data with a few more locations, please re-run that analysis…*
- *New satellite data are available, can you update that figure?*

Sure, I can do that this afternoon…

Adapted from Gandrud (2014) *Reproducible Research with R and RStudio*.

Programming gives you access to more computer power.

The computer is incredibly fast, accurate, and stupid. Man is unbelievably slow, inaccurate, and brilliant. The marriage of the two is a force beyond calculation.

-- Leo Cherne

The World of Work: Careers and the Future

Course Topics (just a reminder…)

Introduction to R

Introduction to Git

Data Wrangling (dplyr, tidyr, etc)

Handling Spatial Vector Data

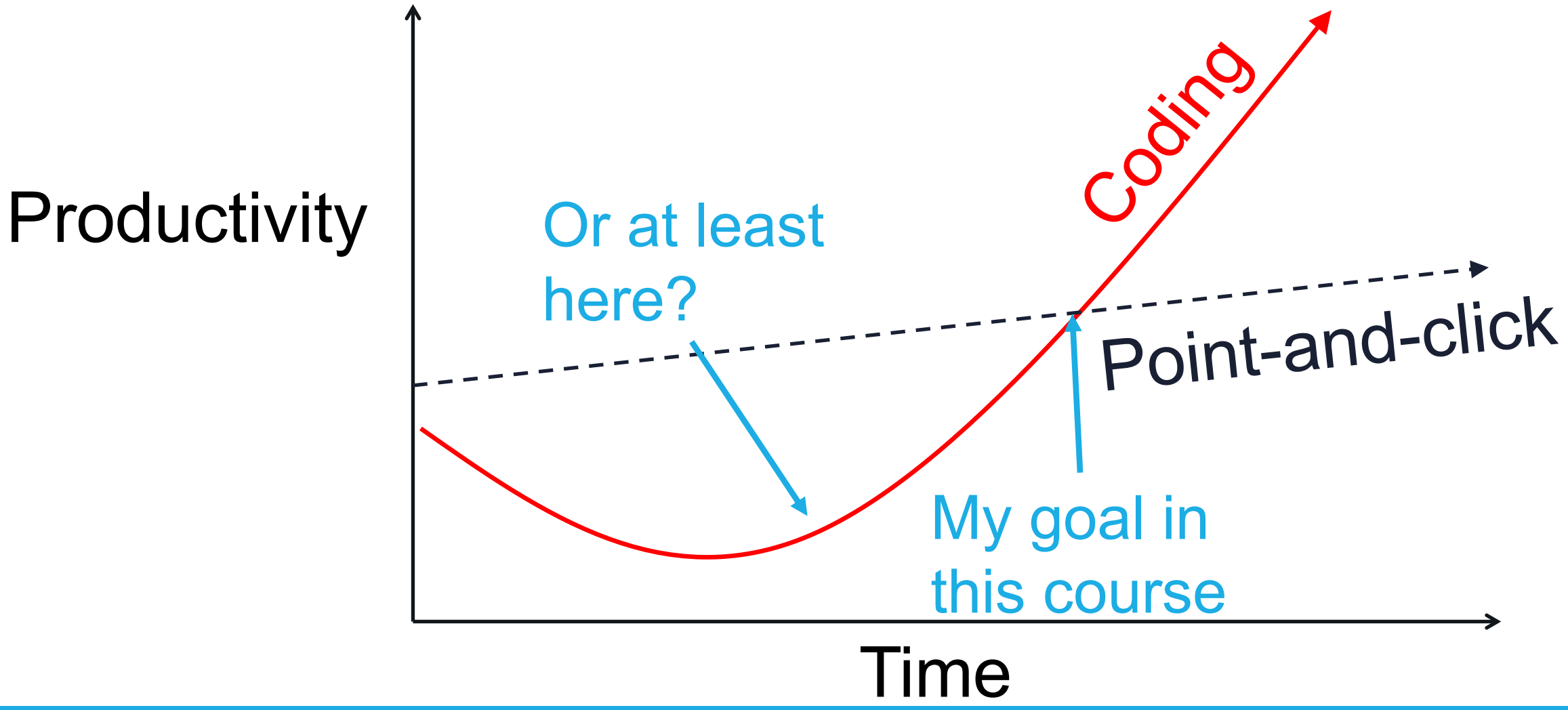
Handling Spatial Raster Data

Generating dynamic research outputs (including web publishing and interactive figures)

Interacting with APIs

Building and summarizing statistical models

From Graphical User Interface (GUI) to scripting/programming



Important products from this course

1. Public repository of worked examples illustrating solutions to problems in spatial data science (case studies)
2. Public website (and underlying repository) demonstrating your coding abilities in an independent project

Put these on your CV and share with future employers/advisors!

Project Presentations next week