

9 of 54

Ways of Being Effective

- Seek first to understand.
 - Be humble about how much you know (ontological humility).
 - Use simple techniques (raw counts).
 - Plot early and often.
- Simplify.
 - Prefer models you understand rather than those that are simply sophisticated.
 - Interpretability is credibility.

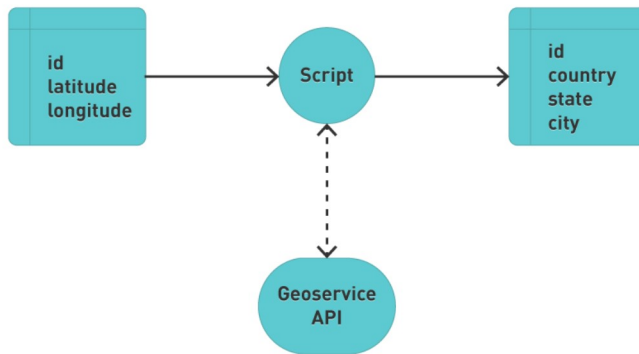
20 of 54

Labor Philosophy

- Fail fast
 - Soup to nuts: Build entire pipeline to make sure it will work.
 - Minimum viable algorithm: Build a simple predictive model to determine viability.
 - Think in RAM (downsampling)
 - Effort vs. progress: Be critical of your own work.
- Hofstadter's law: The work will take 20 percent longer than expected, even if you take this into account.
- Come correct.
 - Eliminate competing explanations.
 - Know the angles and shortcomings.
 - Anticipate questions.

24 of 54

Software Carpentry: Example



32 of 54

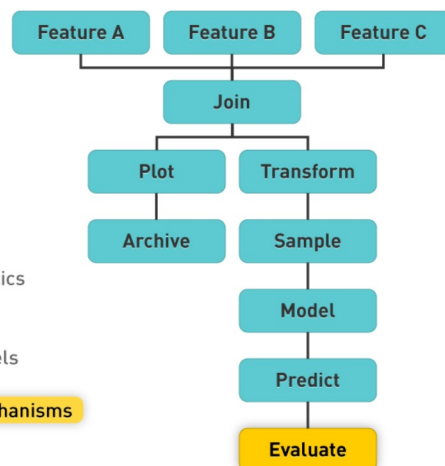
Software Carpentry

- Incremental processes
 - Data format contract
 - One task
 - Testable code

43 of 54

Architecture of Happiness

Prototypical Workflow



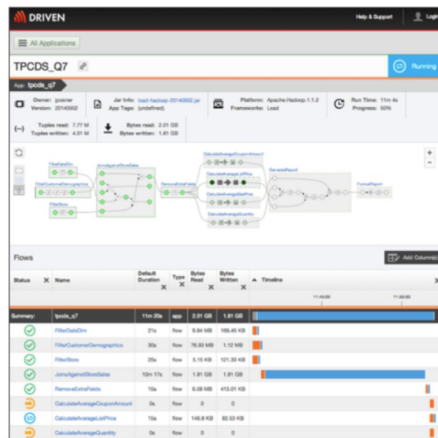
Structural Elements

Atomic feature generation
Robust, automated joins
Reusable plotting battery
Archived analyses and diagnostics
Reproducible transformations
Stratified sampling
Interchangeable statistical models
Prediction at scale

Live and offline evaluation mechanisms

46 of 54

Job Schedulers



48 of 54

Job Schedulers



Automated Execution
Rich arguments
Tracked configuration
Workflow runners

