

Advocating for Automation: Adapting Current Tools in Environmental Science through R

Hannah Podzorski

Why Automate?

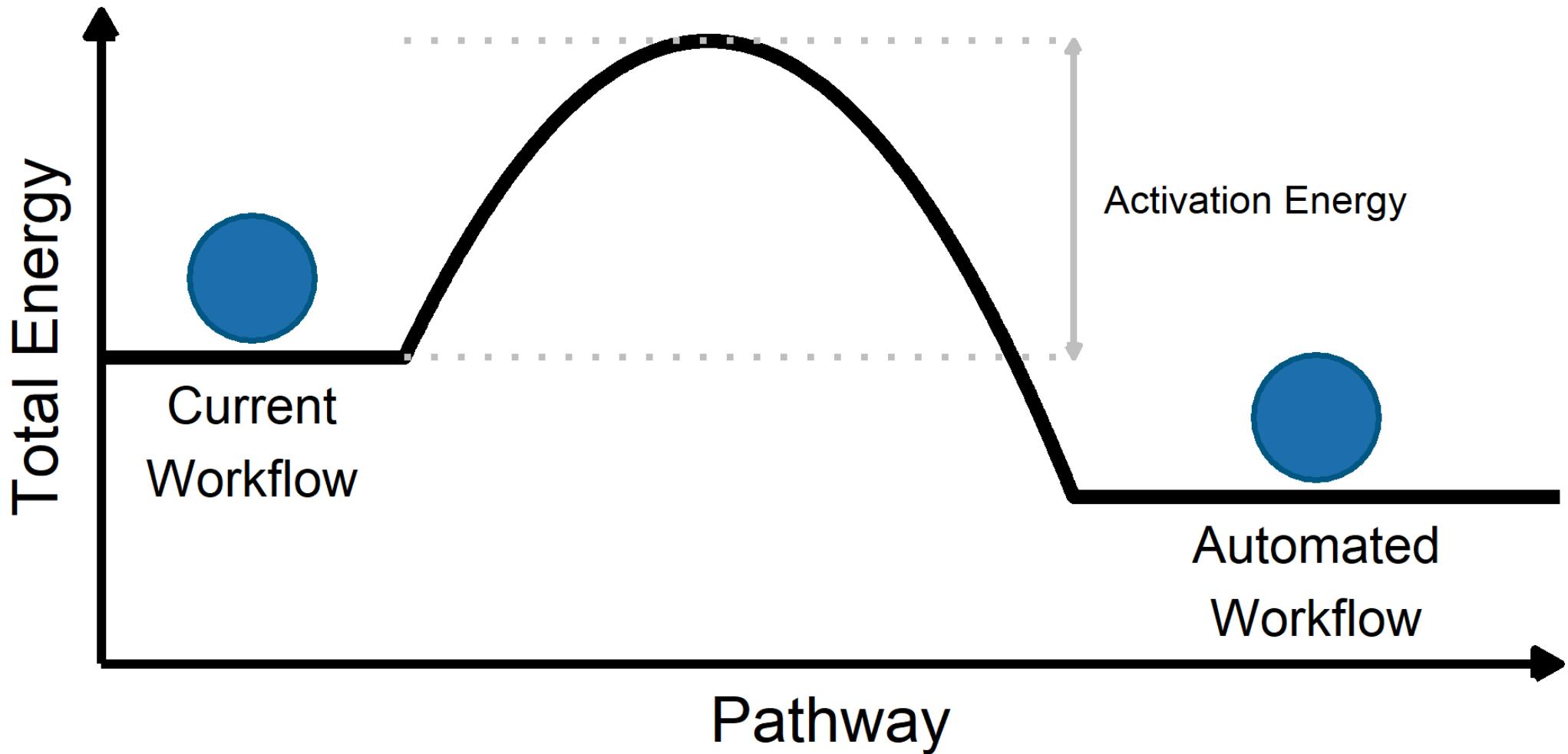


Programmer

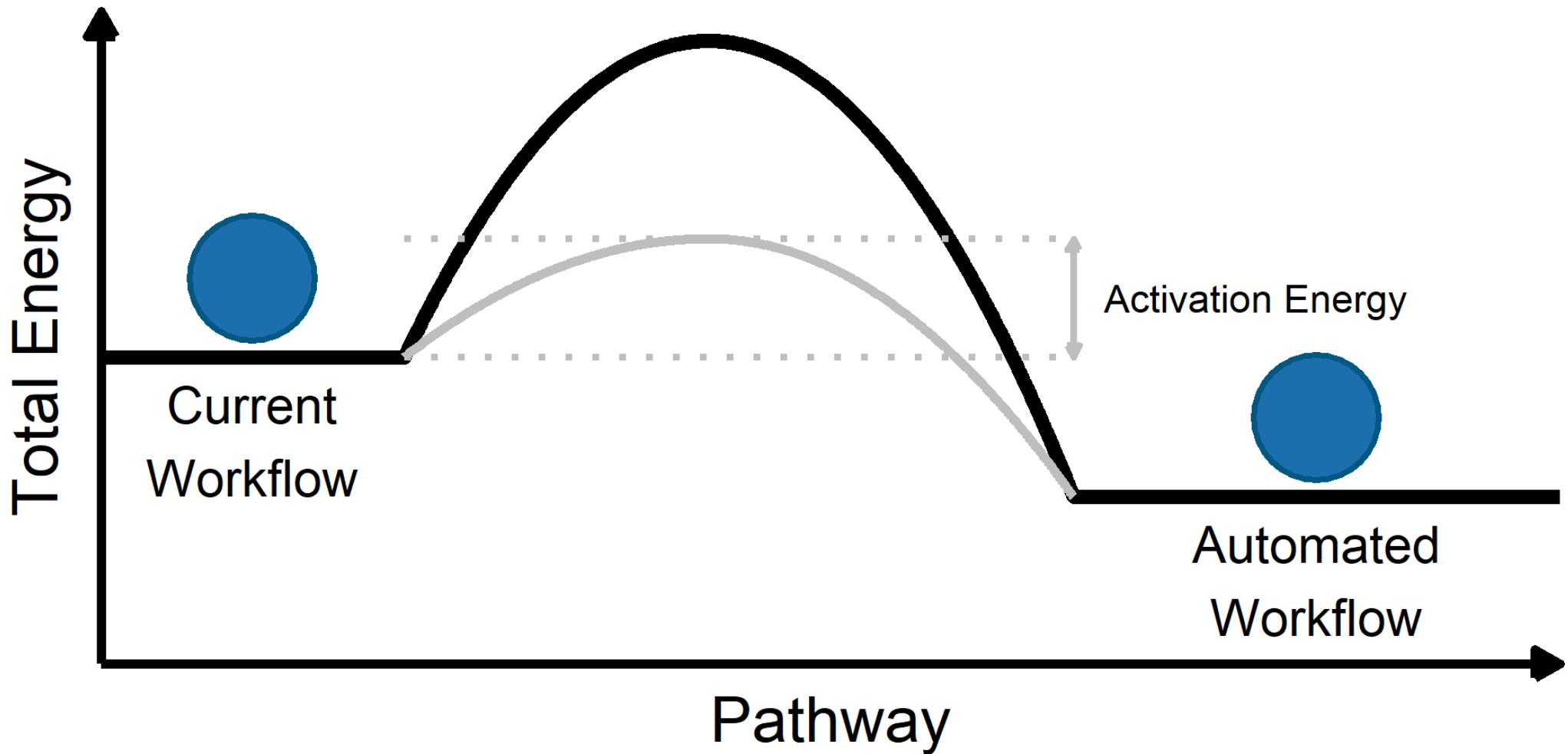
**Task that takes
5 minutes**

Can it be automated?

Reducing the Activation Energy



Reducing the Activation Energy



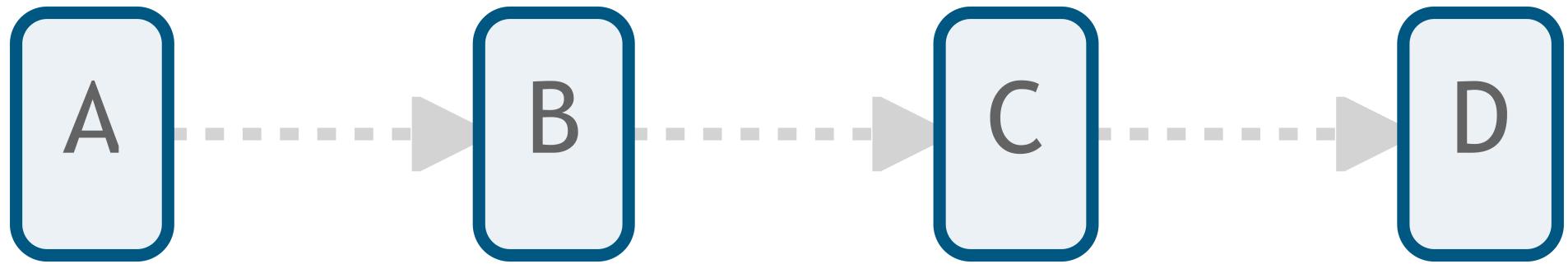
Disclaimer

Differences in Workflow

Reactionary Workflow

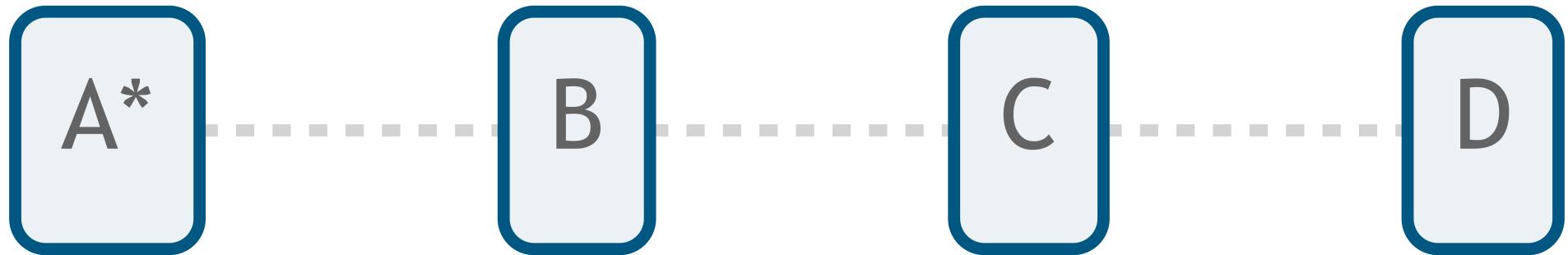


Automated Workflow

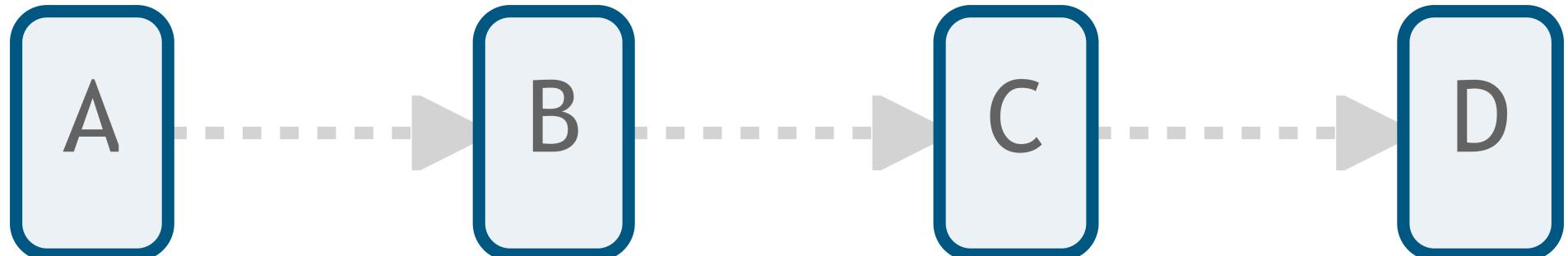


Differences in Workflow

Reactionary Workflow

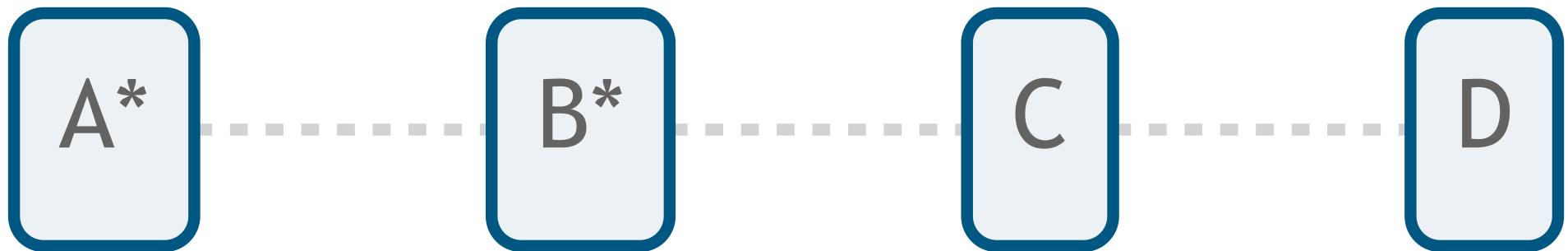


Automated Workflow

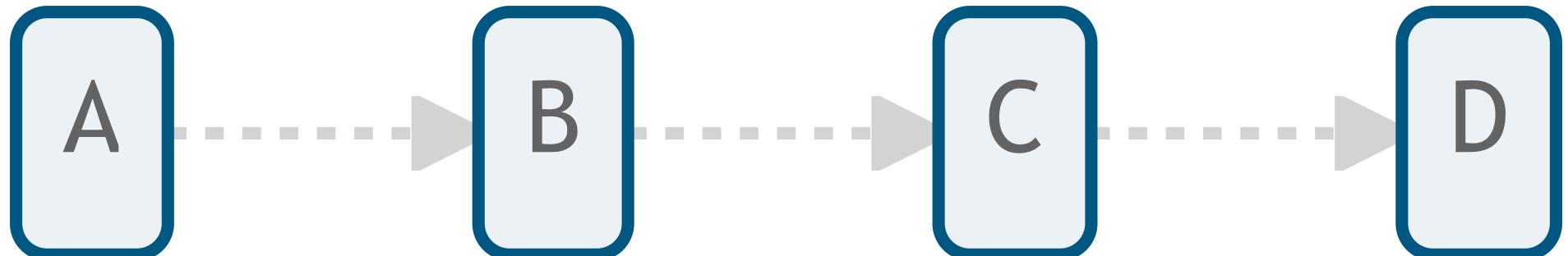


Differences in Workflow

Reactionary Workflow

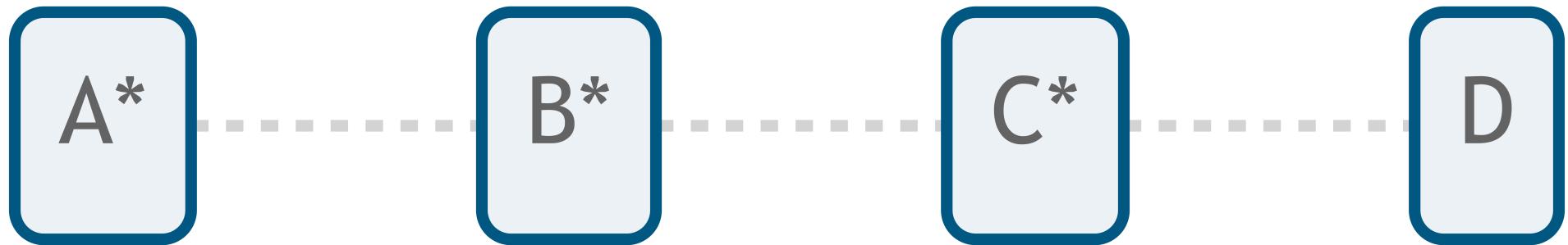


Automated Workflow

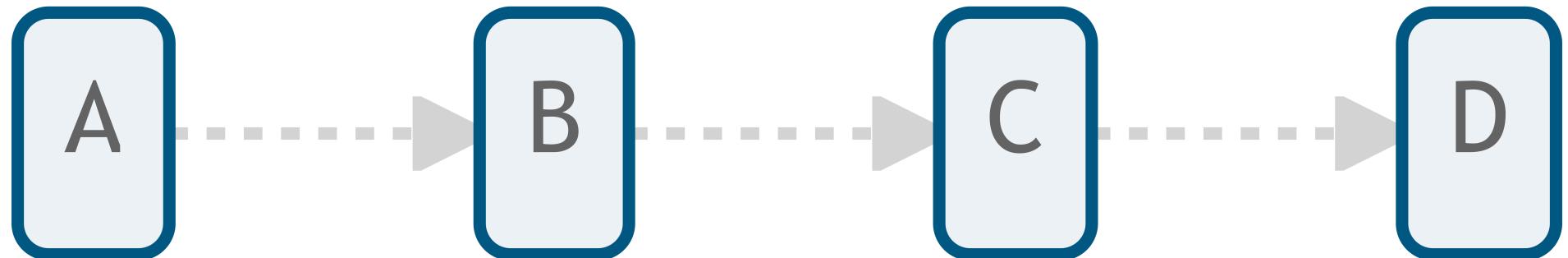


Differences in Workflow

Reactionary Workflow



Automated Workflow

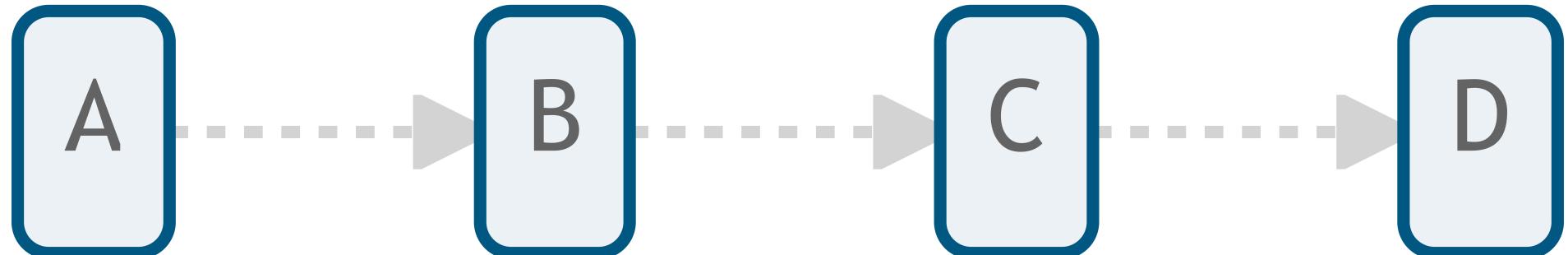


Differences in Workflow

Reactionary Workflow



Automated Workflow

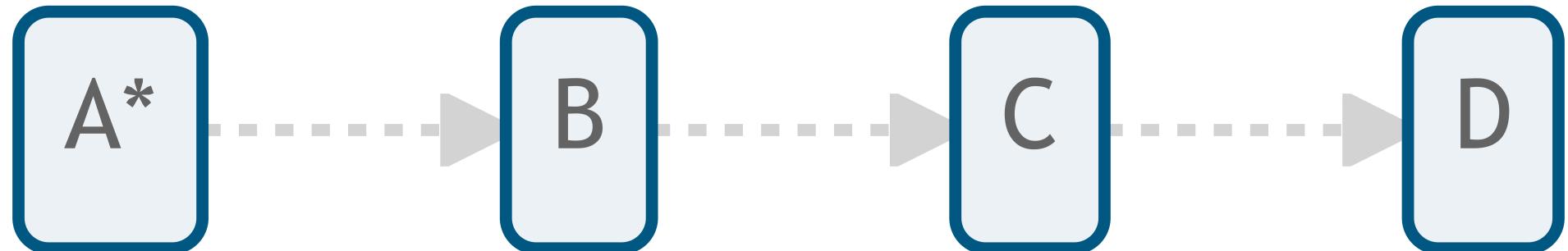


Differences in Workflow

Reactionary Workflow



Automated Workflow

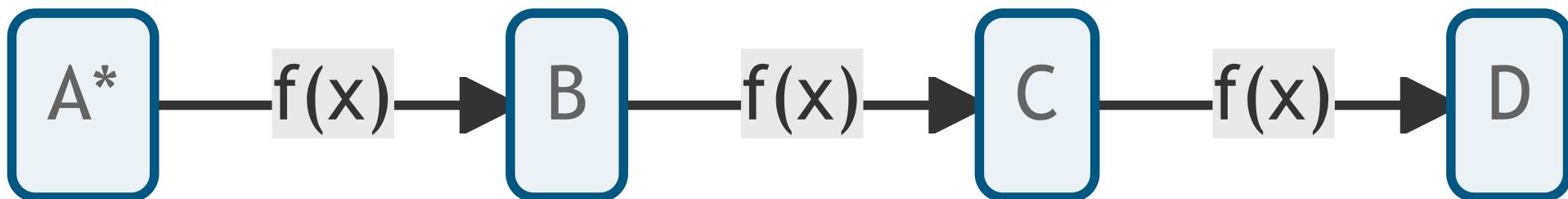


Differences in Workflow

Reactionary Workflow



Automated Workflow



Where to Start?

{openxlsx}

```
1 write.csv(data, "data.csv")
```

{openxlsx}

```
1 write.csv(data, "data.csv")
2
3 openxlsx::write.xlsx(data, "data.xlsx")
```

{openxlsx}

```
1 write.csv(data, "data.csv")
2
3 openxlsx::write.xlsx(data, "data.xlsx")
```

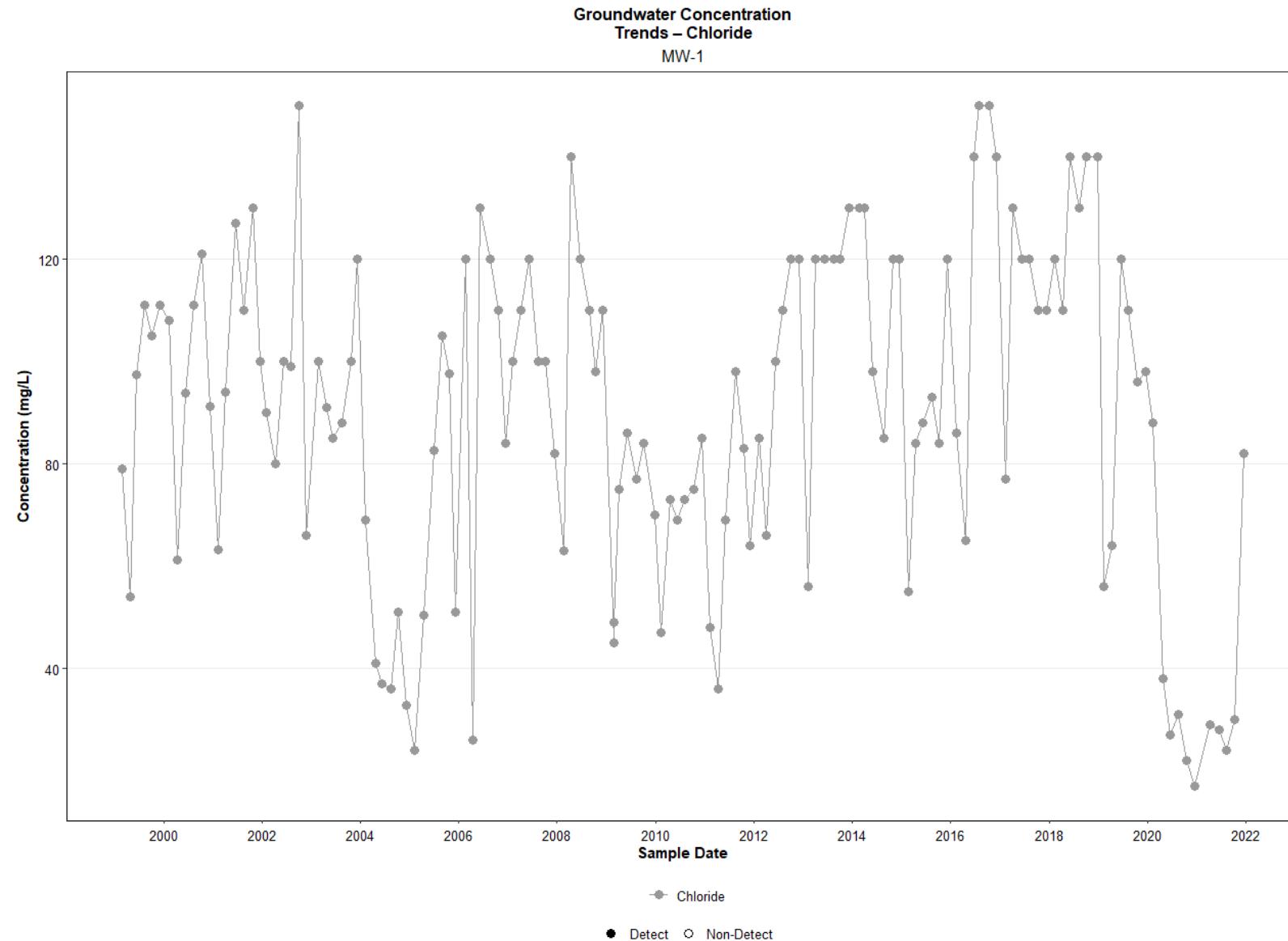
Well ID	Chemical Name	Units	Samples	First Date Sampled	Last Date Samples	Minium Sample	Mean Sample	Maximum Sample
MW-1	Calcium	mg/L	139	03/02/1999	12/15/2021	10	50.64	83
MW-1	Chloride	mg/L	138	03/02/1999	12/15/2021	17	89.68	150
MW-1	Fluoride	mg/L	138	03/02/1999	12/15/2021	0.12	0.33	0.63
MW-1	Magnesium	mg/L	138	03/02/1999	12/15/2021	3	16.04	26
MW-1	Potassium	mg/L	138	03/02/1999	12/15/2021	3.7	7.68	13
MW-1	Sodium	mg/L	138	03/02/1999	12/15/2021	27	75.83	120

Data from California's Groundwater Ambient Monitoring and Assessment Program

{officer}

```
1 library(officer)
2
3 plot <- rvg::dml(ggobj = plot)
4
5 ppxt <-read_pptx() %>%
6   add_slide() %>%
7   ph_with(plot, ph_location(left = 1.3, top = 0.4, width = 8.75, height = 6
8
9 print(ppxt, "./R/Fig-Example.pptx")
```

{officer}



Delayed Gratification

ProUCL

- Statistical Software for Censored Environmental Data
 - Calculates Upper Confidence Limits (UCLs)
- Developed by the U.S Environmental Protection Agency (EPA)



ProUCL Automation

ProUCL Output

UCL Statistics for Data Sets with Non-Detects				
User Selected Options				
Date/Time of Computation	ProUCL 5.11/21/2021 7:42:37 PM			
From File	All_Metals_Input.xls			
Full Precision	OFF			
Confidence Coefficient	95%			
Number of Bootstrap Operations	2000			
Lead.mg/kg (du-2.0.5)				
General Statistics				
Total Number of Observations	18	Number of Distinct Observations	17	
		Number of Missing Observations	0	
Minimum	20	Mean	66.56	
Maximum	211	Median	52.5	
SD	44.71	Std. Error of Mean	10.54	
Coefficient of Variation	0.672	Skewness	2.125	
Normal GOF Test				
Shapiro Wilk Test Statistic	0.803	Shapiro Wilk GOF Test		
5% Shapiro Wilk Critical Value	0.897	Data Not Normal at 5% Significance Level		
Lilliefors Test Statistic	0.168	Lilliefors GOF Test		
5% Lilliefors Critical Value	0.202	Data appear Normal at 5% Significance Level		
Data appear Approximate Normal at 5% Significance Level				
Assuming Normal Distribution				
95% Normal UCL		95% UCLs (Adjusted for Skewness)		
95% Student's-t UCL	84.89	95% Adjusted-CLT UCL (Chen-1995) 89.53		
		95% Modified-t UCL (Johnson-1978) 85.77		

Conclusions

- Communication is key.
- It's ok to start small.
- All skill sets welcome!