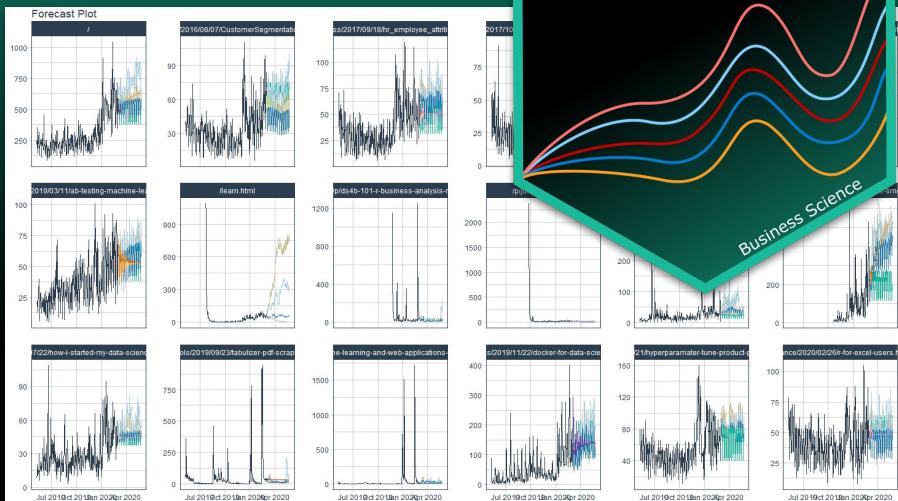


How to Forecast Time Series

3 Tips for Scalable
Forecasting
200 Lines of Full Code Demo

By Business Science

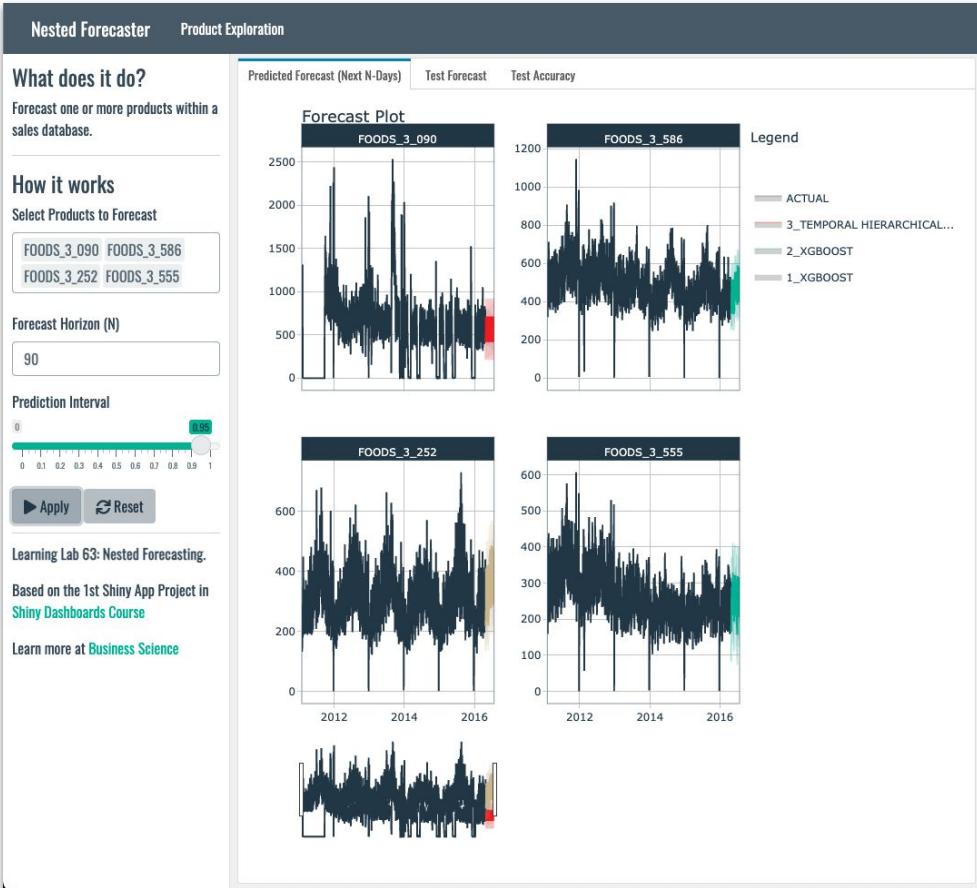
FEATURING: Modeltime



Agenda



1. Business Problem
2. 3 Pro Tips for Scaling Forecasts
3. Modeltime Innovations
4. Full Code Tutorial:
Learn my process.
5. Special Offers on
Full Course



Business Problem

Product Sales Forecasting for Walmart



Walmart: Many products

We're forecasting for Walmart.

Data comes from the M5 Competition.

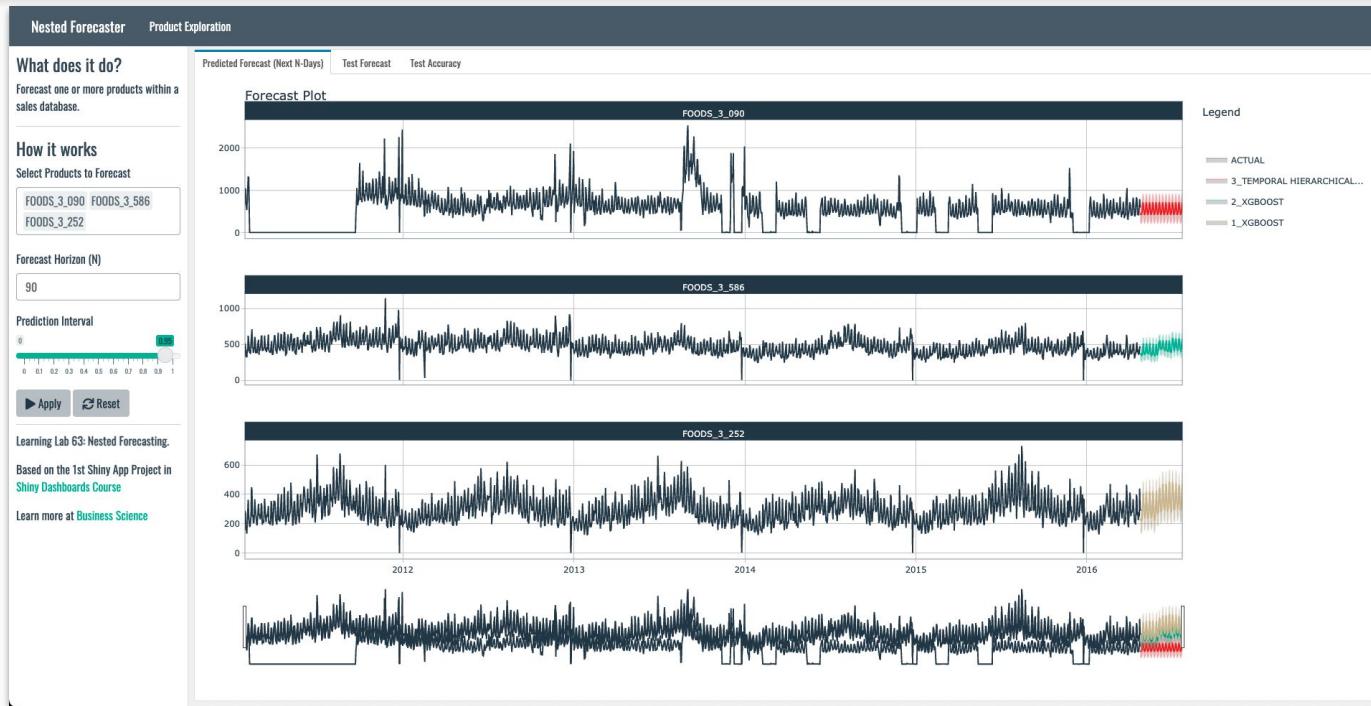
Problem: So. Many. Products.



Walmart carries **thousands** of products.



Solution: What we build today



3-Pro Tips

When dealing with large scale.



Pro-Tip 1: Start with 1 Time Series

Forecasting 1000's of time series is hard.

- Many Time Series
- Many Models

Simplify your problem.





Pro-Tip 2: Scale Efficiently

Time is your greatest resource. Use it wisely.

Strategy:

- **Test Many Models.** Hurts Time,
Increases Accuracy
- **Parallel Processing.** Helps Time.
- **Spark.** Many companies have adopted
Spark to use 1000's of clusters.





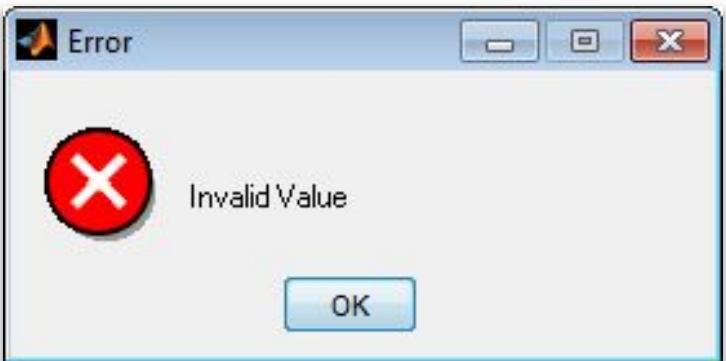
Pro-Tip 3: Anticipate errors.

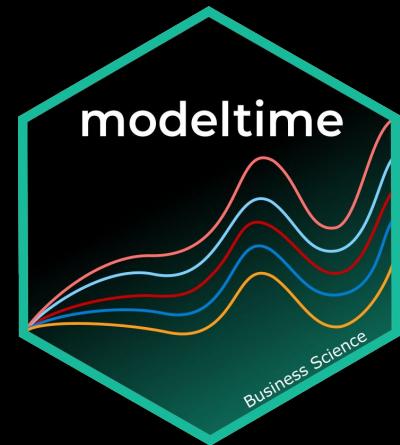
People don't plan to fail. They fail to plan.

Anticipate that something will go wrong.

Have a plan.

- Implement error reports
- Review all errors



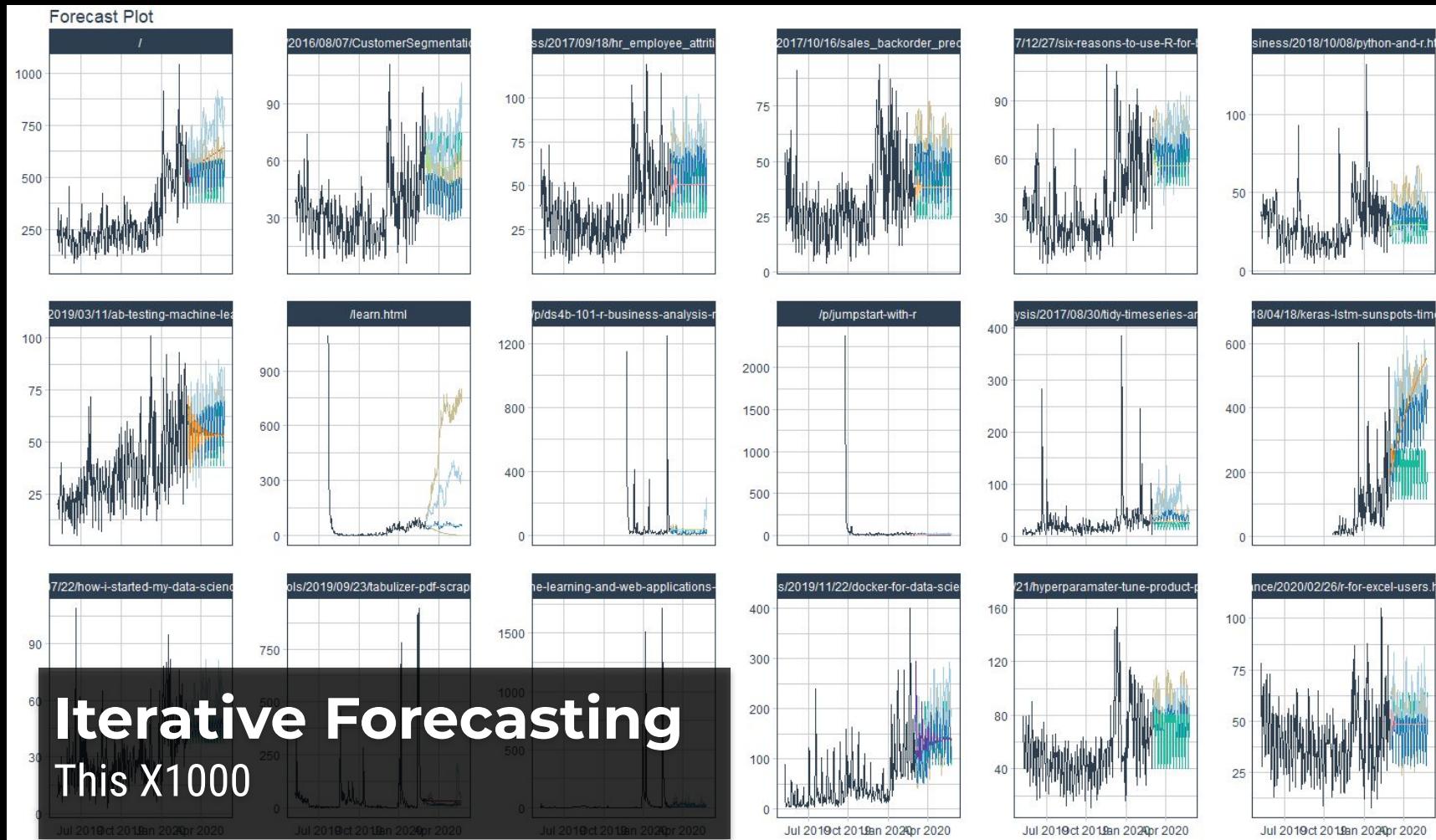


Modeltimer Nested Forecasting with Iteration



Businesses are Shifting To Automation Time Series Forecasting on Massive Scale

#TimeSeries
#AtScale



About Modeltime

R Package that I developed for Forecasting

Works with Tidymodels

Integrates ARIMA, Prophet, Tidymodels, & many time series algorithms

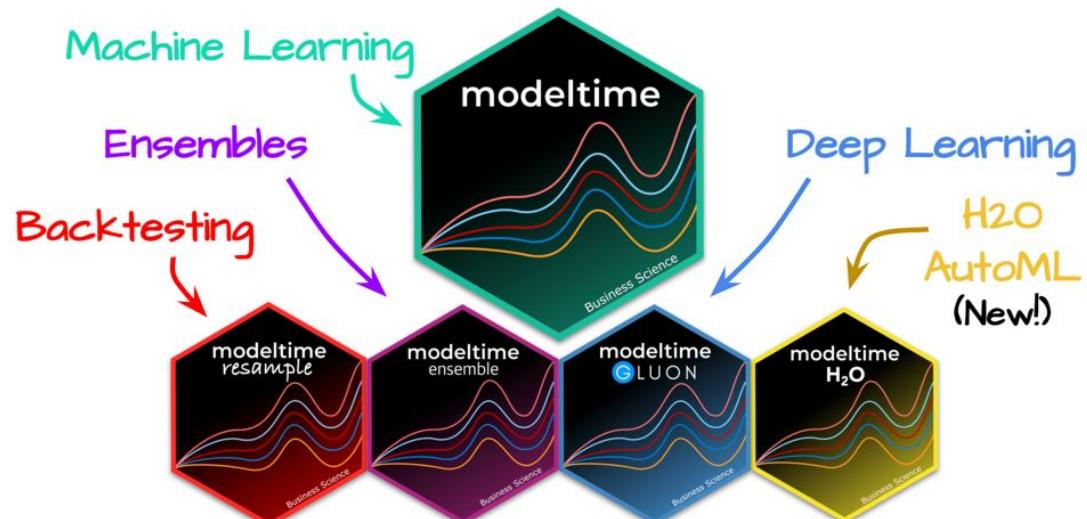
Is an Ecosystem too!

- Modeltime GluonTS
- Modeltime H2O
- Modeltime Ensemble
- Modeltime Resample
- Timetk

What is Modeltime?



The Modeltime Ecosystem is Growing





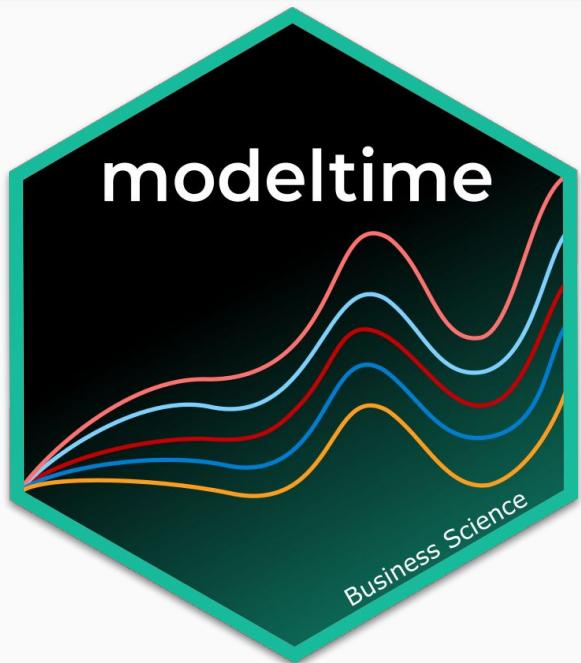
Developed for Scale

We have 2 Systems to Scale:

1. **Global Modeling:** 100X faster
2. **Nested Modeling:** Slower, More Accurate

Focus today: Nested Modeling

(See Course for Global Modeling)



Nested Forecasting

- Iterative Analysis
- Built for Scale
(Parallel Processing)
- Handles Errors
Gracefully
- Key Concepts: Logging

Nested Forecasting



Many Time Series



Many Models



Best Forecasts

Iterative
forecasting
using **Nested
Modeltime
Tables**





Extremely easy to run in parallel.

Parallel Processing

```
111 # REFIT ----  
112  
113 parallel_start(2) ← Setup Clusters  
114  
115 nested_modeltime_refit_tbl <- best_nested_modeltime_tbl %>%  
116     modeltime_nested_refit(  
117         control = control_nested_refit(  
118             verbose    = TRUE,  
119             allow_par = TRUE ← Control  
120         )  
121     )  
122
```

Can scale to millions of time series
with Spark



Logging & Error Handling

Pushes **everything** you need to the most expensive operation. Proceeds even when errors occur.

Benefit 1: Can retrieve key information instantly.

Benefit 2: Can check error logs & isolate edge cases where your analysis fails

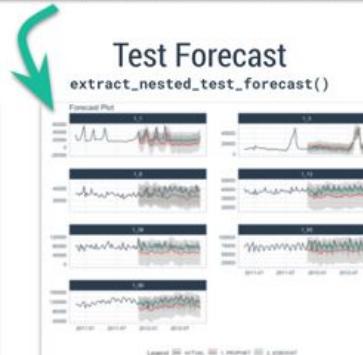
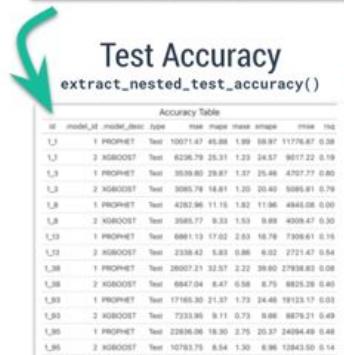
Extracting Nested Modeltime Table Logs



Nested Modeltime Table

```
# Nested Modeltime Table
Trained on: .actual_data | Model Errors: [0]
# A tibble: 7 x 5
  id   .actual_data    .future_data    .splits    .modeltime_tables
  <fct> <tibble [104 x 2]> <tibble [52 x 2]> <split [52/52]> <mdl_time_tbl [1 x 5]>
  1_1   <tibble [104 x 2]> <tibble [52 x 2]> <split [52/52]> <mdl_time_tbl [1 x 5]>
  1_3   <tibble [104 x 2]> <tibble [52 x 2]> <split [52/52]> <mdl_time_tbl [1 x 5]>
  1_8   <tibble [104 x 2]> <tibble [52 x 2]> <split [52/52]> <mdl_time_tbl [1 x 5]>
  1_13  <tibble [104 x 2]> <tibble [52 x 2]> <split [52/52]> <mdl_time_tbl [1 x 5]>
  1_38  <tibble [104 x 2]> <tibble [52 x 2]> <split [52/52]> <mdl_time_tbl [1 x 5]>
  1_93  <tibble [104 x 2]> <tibble [52 x 2]> <split [52/52]> <mdl_time_tbl [1 x 5]>
  1_95  <tibble [104 x 2]> <tibble [52 x 2]> <split [52/52]> <mdl_time_tbl [1 x 5]>
```

Contains
Logged
Attributes
for fast
extraction



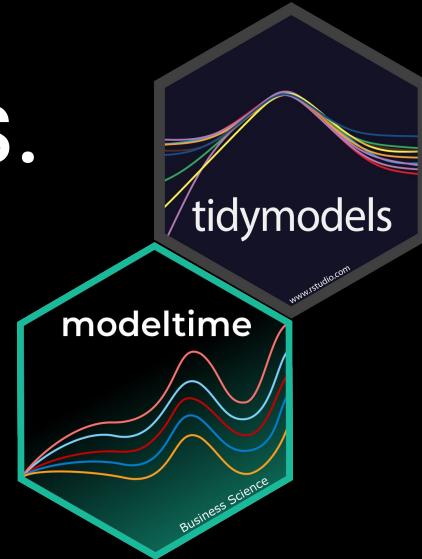
Error Reporting extract_nested_error_report()

```
# A tibble: 4 x 4
  id   .model_id .model_desc .error_desc
  <fct> <int> <chr> <chr>
  1_1   2 BOOST_TREE  "data" has class 'character' and length 52.\n  'data...
  1_3   2 BOOST_TREE  "data" has class 'character' and length 52.\n  'data...
  1_8   1 BOOST_TREE  "'x' should be an `rsplit` object"
  1_8   2 BOOST_TREE  "'x' should be an `rsplit` object"
```

30-Min Demo

Forecasting 100 time series.

Walmart Product Sales



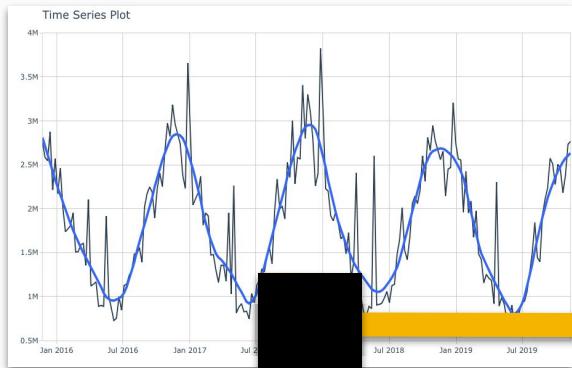
How do I learn what we just did?

Become the Time Series Expert for your organization. 

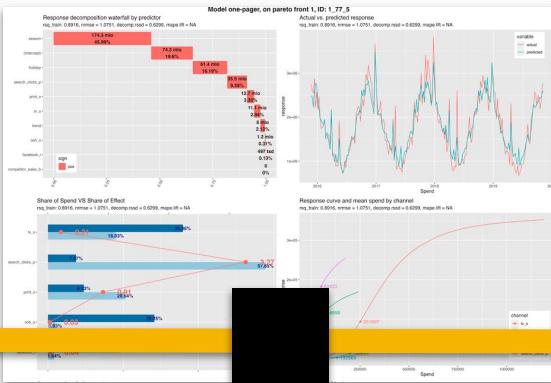
A lot to learn to become a Senior Data Scientist. But not to get started.



Data Preparation & Visualization



Modeling



Shiny Apps



Career Growth Curve

Learn DS Foundations
(101)

Add Machine Learning
(201 & 203)

Learn Shiny Apps,
Cloud, & Automation
(102 & 202A)

Senior Data Scientist

Data Scientist

Business Analyst

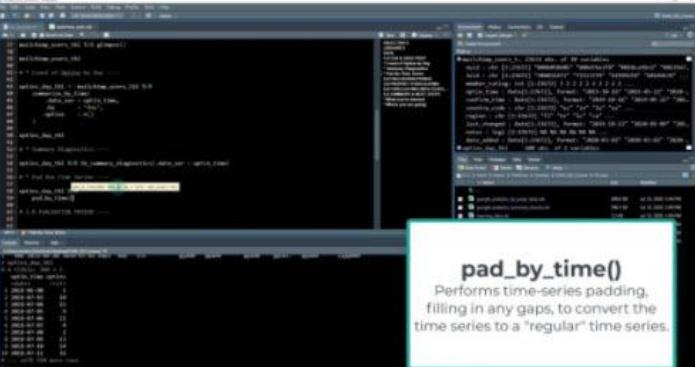


 Jonathan Regenstein • 1st
Head of Data and Quantamental Research at Truist Securities
1d • 

Loving the new High Performance Time Series course (<https://lnkd.in/dfaBCTJ>) from Matt Dancho.

I'm only about 10% through the material and already picked up a couple super useful new moves like the one below - pad_by_time!

#rstats #rstudio #python #forecasting



pad_by_time()
Performs time-series padding, filling in any gaps, to convert the time series to a "regular" time series.

103 + 10 Comments

Reactions



+95

 Love  Comment  Share  Send

"I'm only 10% through and I already picked up a couple super useful new moves."

-Jonathan Regenstein

Author of Reproducible Finance with R
Head of Data & Quantamental Research at Truist Securities



Thursday, August 5th ▾



Chris Selig 1:19 PM

Hey folks, just wanted to let you know that because of all the R courses [@Matt Dancho](#) has provided, I landed my first Data Scientist role at a global energy company.

Setting up shiny dashboards in AWS was extremely helpful during my interview as the people interviewing me were looking for someone that has some knowledge of cloud technologies. Even though I haven't finished the python course yet, that was also a huge boost because although where I'm going is an R company, there are some python people that I'm to help support. I ticked off ALOT of the boxes for the person they were looking for.

Thanks Matt!

P.S. Although I haven't gotten around to updating my LinkedIn yet, feel free to add me!
<https://www.linkedin.com/in/chris-selig/>



 3 replies Last reply 6 days ago

*“Because of the **R courses** Matt has provided, I landed my **first Data Scientist Role.**”*

-Chris Selig

Brand New Data Scientist!

CNOOC International | Global Energy Company

Leo Timmermans 7:17 AM

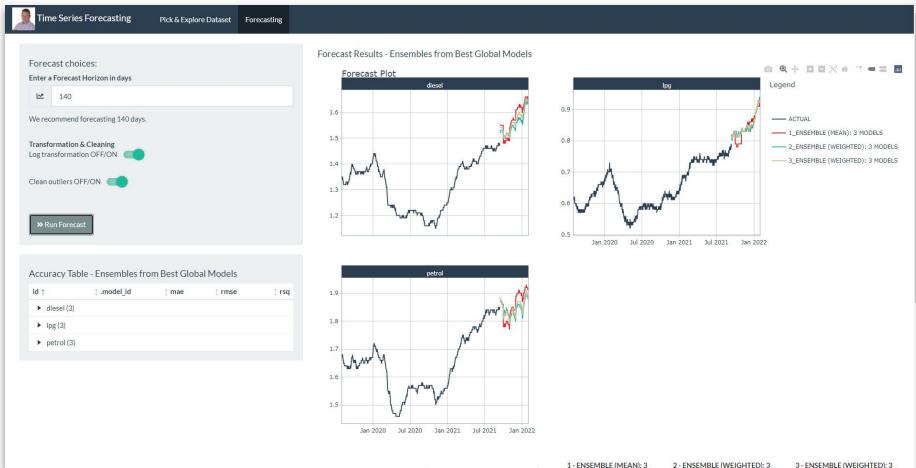
Hi everybody,

Been working on a time series forecasting shiny app, which is a bit like Matt's Nostrodamus app (<https://business-science.shinyapps.io/nostradamus/>) from Learning Lab 46.

This app, the time series course and the shiny courses were my inspiration.

You can find the result here: https://leotimmermans.shinyapps.io/time_series/

I will be adding this to my portfolio shiny app: <https://portfolio.leotimmermans.nl/>



“The time series courses and shiny courses were my inspiration”

-**Leo Timmermans**
Data Scientist
OTTO Work Force

5-Course R-Track System



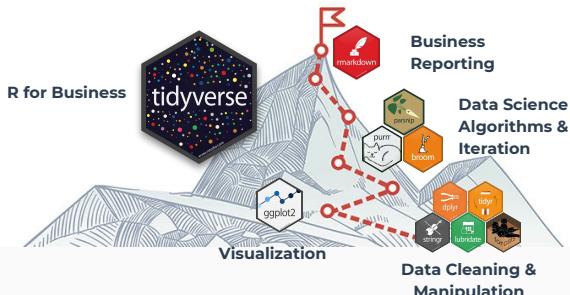
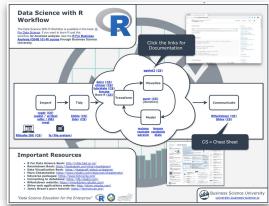
Business Analysis with R (DS4B 101-R)

Advanced ML & Time Series (DS4B 201-R + DS4B 203-R)

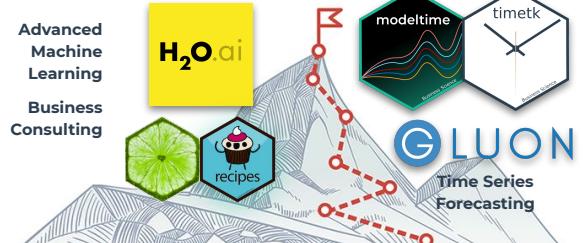
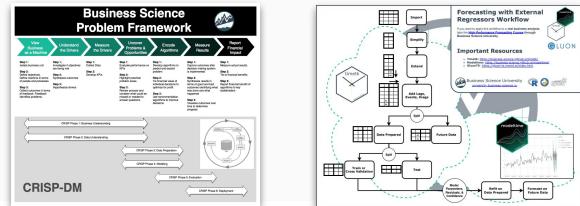
Web Apps & Shiny Developer (DS4B 102-R + DS4B 202A-R)

Project-Based Courses with Business Application

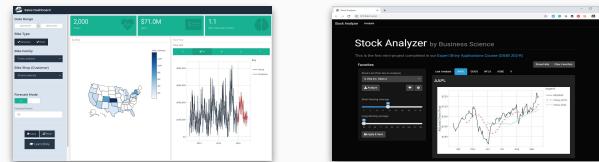
Data Science Foundations
8 Weeks



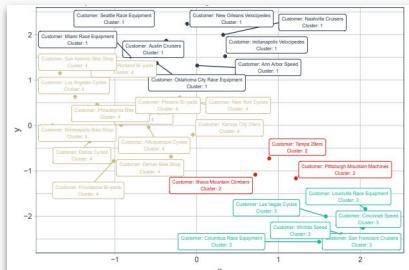
Machine Learning & Time Series
20 Weeks



Web Application Development
12 Weeks

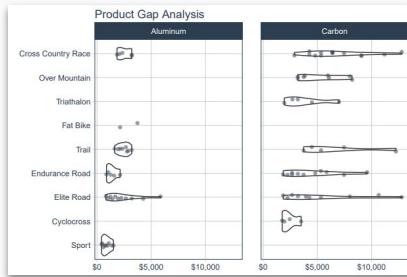


You complete 8 End-To-End Projects.



Customer Segmentation

DS4B 101-R



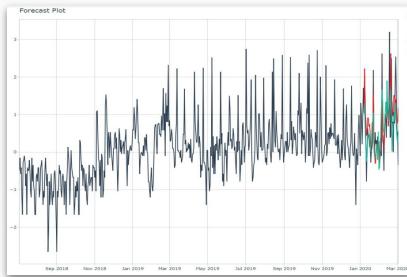
Product Price Prediction

DS4B 101-R



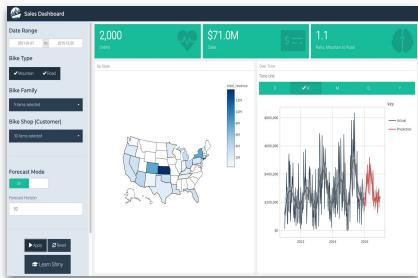
Churn Prediction

DS4B 201-R



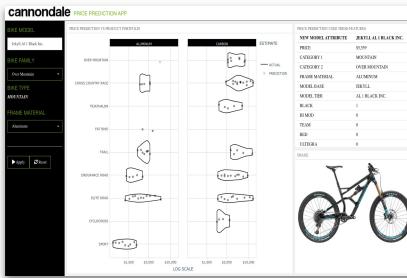
Demand Forecasting

DS4B 203-R



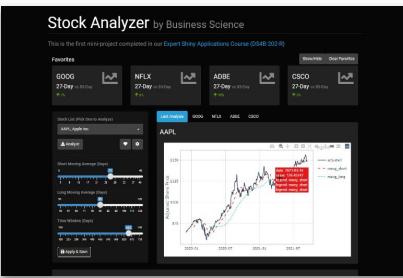
Demand Forecast App

DS4B 102-R



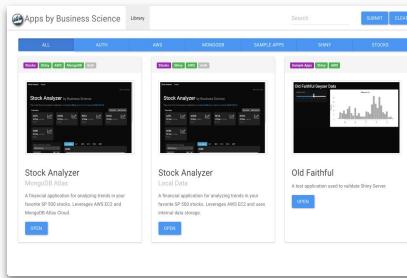
Product Price Recommender App

DS4B 102-R



Stock Analyzer App

DS4B 202A-R



Application Library

DS4B 202A-R



The R-Track Transformation



I got a
DS Job!

"Thanks to Matt and what I've learned so far, I was able to do an in-depth analysis of Consumer Financial Protection Bureau (CFPB) data, following his Business Science Process Framework and complete the project using RMarkdown. The polished, finished product impressed the hiring manager so much, he was willing to fast-track an offer."

-Jennifer Cooper, VP Strategic Analytics



How Jennifer Landed a VP, Strategic Analytics Role



The Transformation

Got a DS Job!



Masatake Hirono landed a job at a top Management Consulting Firm.

[Learn More](#)

Got a Senior DS Role!



Z. Du applied used data science to improve patient care while reducing costs.

[Learn More](#)

Got 3 raises in 1-year



Mohana Chittoor got a 40% raise with data science.

[Learn More](#)

Built 1st shiny app in 3-months



Herb went from no experience to deploying his first app in 3 months

[Learn More](#)

Won the 2020 shiny competition



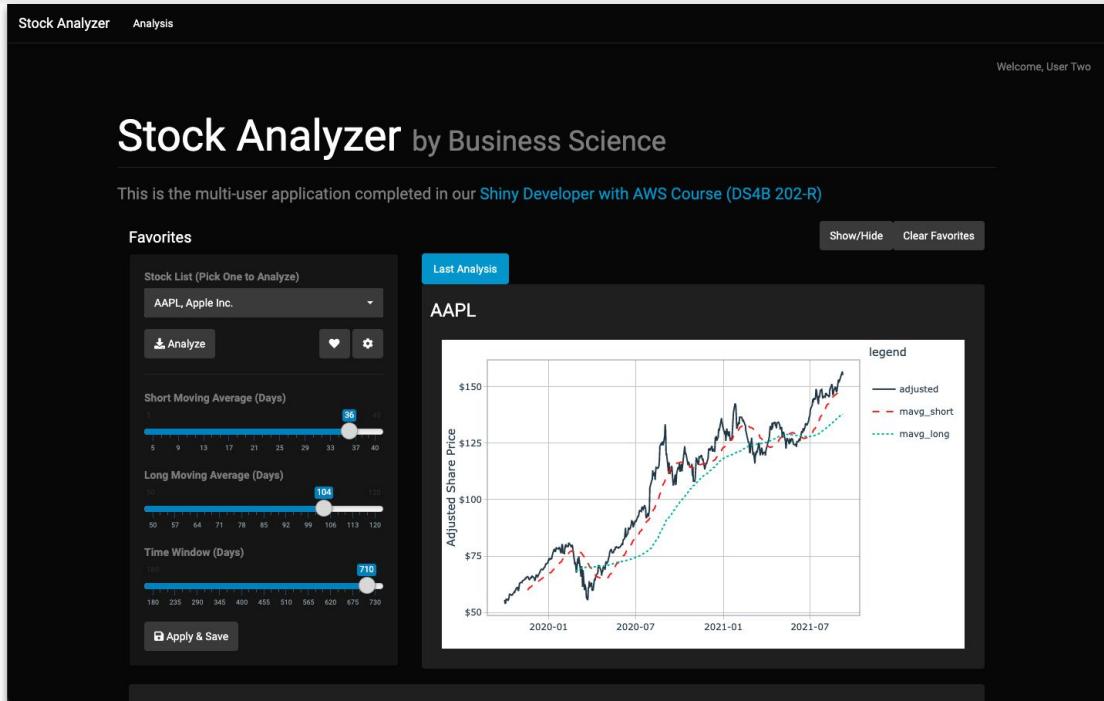
Raj used his new Shiny experience to win top 5 of a Competition.

[Learn More](#)



What you build.

1 of 8 projects





Summary: What's inside

1000+ Videos

130+ Hours of Content

8 End-to-End projects

Private Community Mentorship

Lifetime Access

30-Day Money Back Guarantee

\$5,000 value

The screenshot displays the R-Track system interface. On the left, there's a promotional banner for the "NEW 5-Course R-Track System" featuring courses like "Business Analysis with R (DS4B 101-R)" and "Advanced Machine Learning for Business". Below this is a diagram of the "Data Science Foundations" program, which spans 8 weeks and includes modules such as "R for Business", "R tidyverse", "Business Reporting", "Data Cleaning & Manipulation", "Machine Learning", and "Business Consulting". On the right, a forum thread titled "BSU DS4B R Tra..." is shown, with a post from Matt Dancho asking about zero variance features in R. Adam responds with a detailed explanation involving `map_dfc()`, `gather()`, and `filter()` functions. Other users like Matt and Adam are also visible in the conversation.

15% OFF PROMO Code: LEARNINGLABS



~~\$5,000+~~ value

~~\$1999~~

~~\$1699~~



R-TRACK NEW 5-Course R-Track System

The diagram illustrates the R-TRACK system, which consists of three project-based courses with business applications:

- Business Analysis with R (DS4B 101-R)**: 8 Weeks. Focuses on Data Science Foundations, R for Business, and various data manipulation and visualization packages like tidyverse, dplyr, and ggplot2.
- Advanced ML & Time Series (DS4B 201-R + DS4B 203-R)**: 20 Weeks. Focuses on Machine Learning & Time Series, Advanced Machine Learning, Business Consulting, and tools like H2O.ai, TensorFlow, and GLUON.
- Web Apps & Shiny Developer (DS4B 102-R + DS4B 202A-R)**: 12 Weeks. Focuses on Web Application Development, Web Apps & Dashboards, and Full Stack Development & Deployment, involving technologies like AWS, Shiny, and Docker.

Business Science University R-Track

https://university.business-science.io/p/5-course-bundle-machine-learning-web-apps-time-series/?coupon_code=LEARNINGLABS

Career acceleration awaits

university.business-science.io

