

# Forecasting Google Analytics

With the Facebook Prophet



Google Analytics

PROPHET

Matt Dancho & David Curry  
*Business Science Learning Lab*





# Learning Lab Structure

- **Presentation**  
(20 min)
- **Demo's**  
(30 min)
- **Pro-Tips**  
(15 mins)



**Matt Dancho**

Founder of Business Science, Matt designs and executes educational courses and workshops that deliver immediate value to organizations. His passion is up-leveling future data scientists coming from untraditional backgrounds.



**David Curry**

Founder of Sure Optimize, David works with businesses to help improve website performance and SEO using data science. His passion is **ethical Machine Learning initiatives**.

# Shiny API Series

- **Lab 28 - Shiny Real Estate App**

- Zillow API



- **Lab 29 - Shiny Oil & Gas App**

- Quandl API



- **Lab 30 - Shiny Finance App**

- Tidyquant API



- **Lab 31 - Shiny Marketing App**

- Google Analytics API



- **Lab 32 - Shiny Twitter App**

- Twitter API



 Learning Labs Pro  
Community-Driven Data Science Courses

 Matt Dancho 

# Learning Labs PRO

Every 2-Weeks

1-Hour Course

Recordings + Code + Slack

**\$19/month**

[university.business-science.io](http://university.business-science.io)

**Lab 31 - Shiny API Series, Pt 4**  
**Shiny + Google Analytics + Prophet**

**Lab 30 - Shiny API Series, Pt 3**  
**Shiny + Finance + Excel + R + Tidyquant**

**Lab 29 - Shiny API Series, Pt 2**  
**Shiny + Quandl + ARIMA for Energy Forecasting**

**Lab 28 - Shiny API Series, Pt 1**  
**Shiny + Zillow for Real Estate**

**Lab 27 - Marketing Series, Pt 4**  
**Google Trends Automation with Shiny**

**Lab 26 - Marketing Series, Pt 3**  
**Customer Journey with Machine Learning**

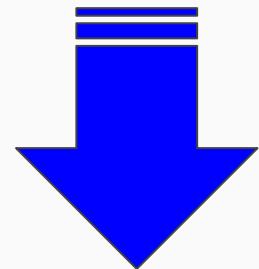
**Lab 25 - Marketing Series, Pt 2**  
**Attribution with ChannelAttribution**

**Lab 24 - Marketing Series, Pt 1**  
**A/B Testing with Infer**

**Lab 23 - SQL Series**



**Continuous Learning**  
Advanced Topics



**Learning Labs Pro**

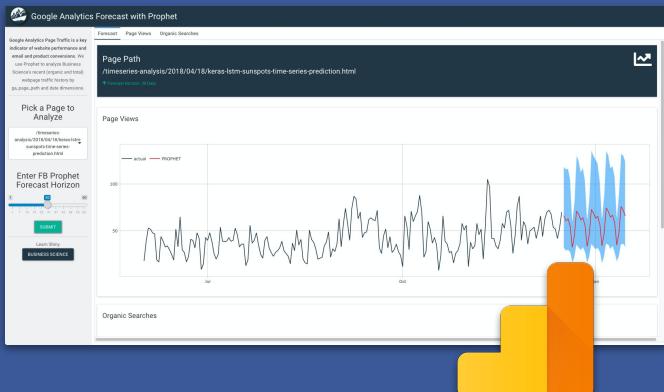
Community-Driven Data Science Courses

 Matt Dancho

**\$19/m**

# Agenda

## PROPHET



Google Analytics

- **Demo**
  - Google Analytics
  - Page Views
  - Forecast Automation
- **30-Min Demo**
  - Google Analytics API
  - Prophet
  - **Shiny Page View Forecast Automation [LL PRO]**
- **Business Case**
  - Marketing
- **Pro-Tips & Learning Guide**
  - **Build your career**
- **Google Analytics API**
  - How GA works (80/20)
- **Facebook Prophet**
  - What is it?
  - Fable Implementation





Forecast

Page Views

Organic Searches

Google Analytics Page Traffic is a key indicator of website performance and email and product conversions. We use Prophet to analyze Business Science's recent (organic and total) webpage traffic history by ga\_page\_path and date dimensions.

Pick a Page to Analyze

 / 

Enter FB Prophet Forecast Horizon

1 7 13 19 25 31 37 43 49 55 60

Learn Shiny

BUSINESS SCIENCE

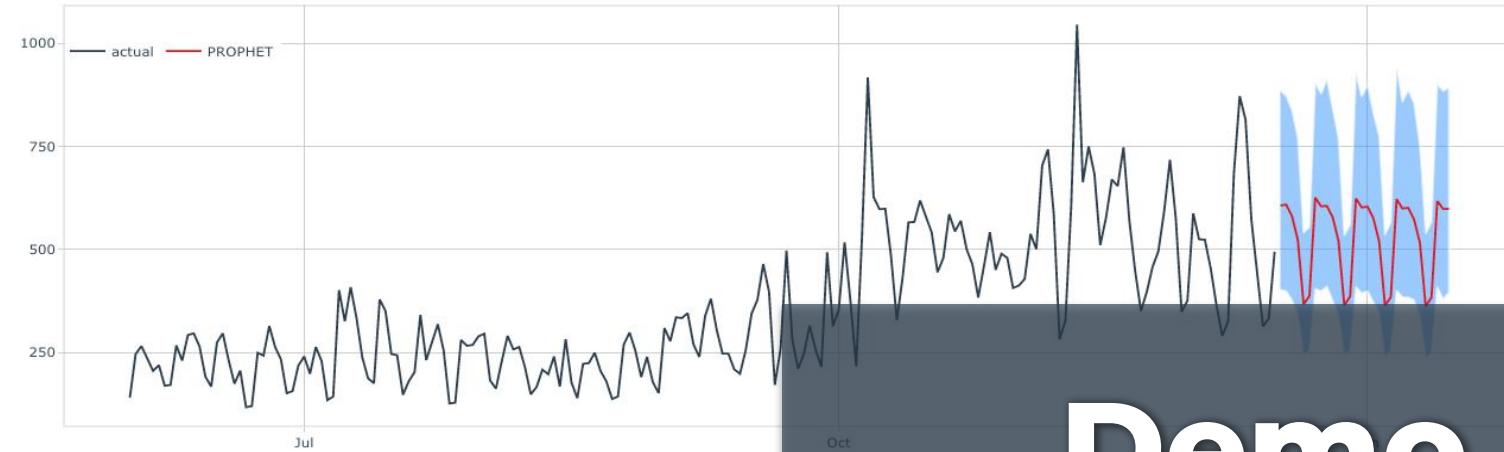
Page Path

/

↑ Forecast Horizon: 30 Days



Page Views



# Demo

Organic Searches

# **Business Case Study**

## Marketing



## Why web traffic is important

**Build  
Relationship** ②

Email enables Nurturing



**Sales**

Nurturing leads to Purchasing

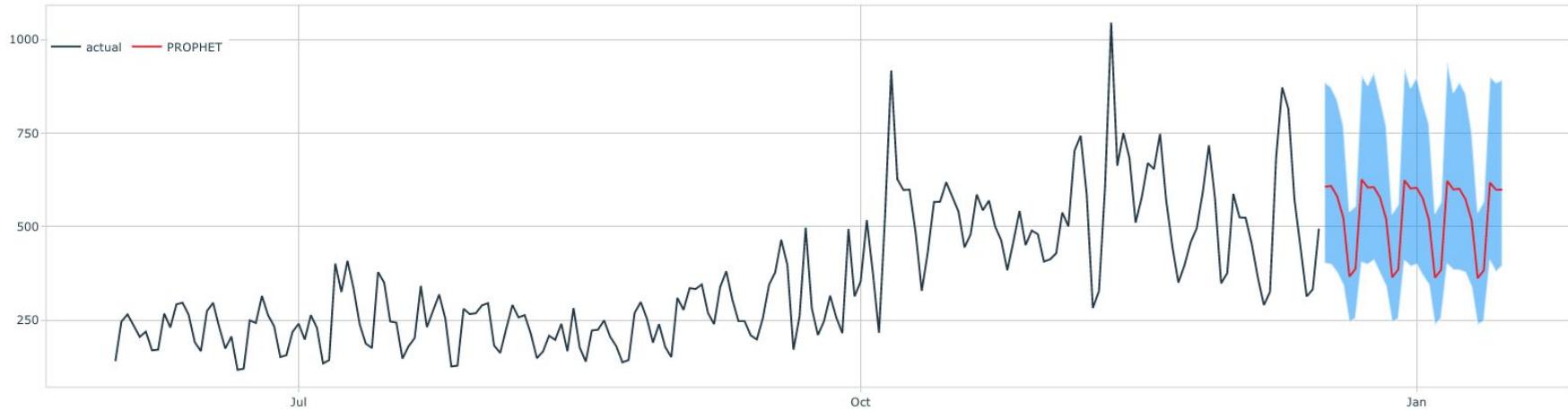
③ **Establish  
Relationship**

Web Traffic leads to email collection

# Forecasting is critical



Page Views



**The better we can predict**

- The better we can estimate email conversions
- The better we can estimate revenue
- The better we can **drive results**

**A 20% reduction in  
forecast error**

Can save an organization \$10M annually

# Google Analytics

## API in R



Google Analytics

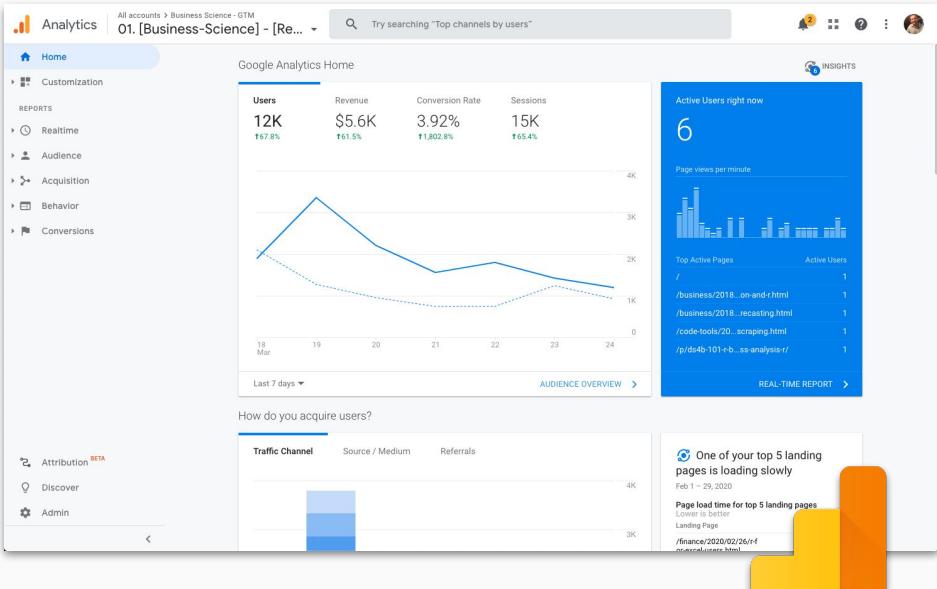


# Google Analytics

Access to **A TON** of important website data

Key Concepts:

- Data Organization
- Metrics
- Dimensions



Google Analytics



## Demos &amp; Tools

[Autotrack](#)[Account Explorer](#)[Campaign URL Builder](#)

## Dimensions &amp; Metrics Explorer

[Updated!](#)

## Embed API

[Basic Dashboard](#)[Multiple Views](#)[Interactive Charts](#)[Working with Custom Components](#)[Third Party Visualizations](#)[Server-side Authorization](#)Enhanced Ecommerce [Updated!](#)

## Hit Builder

## Polymer Elements

## Query Explorer

## Request Composer

## Spreadsheet Add-on

## Tag Assistant

## Usage Trends

## Resources

# Dimensions & Metrics Explorer

The Dimensions & Metrics Explorer lists and describes all of the dimensions and metrics available through the [Core Reporting API](#).

The Dimensions & Metrics Explorer has the following features:

**Explore all of the dimensions and metrics** – Search or browse by group. Select a dimension or metric for additional details such as descriptions and attributes.

**Identify valid combinations** – Not all dimensions and metrics can be queried together. Only certain dimensions and metrics can be used together to create valid combinations. Select a dimension or metric checkbox to see all the other values that can be combined in the same query.

Search 

X

 Only show fields that are allowed in segments Include deprecated fields[+ Expand All](#)[- Hide All](#)**- User**

## Dimensions

 User Type [ga:userType](#) Count of Sessions [ga:sessionCount](#) Days Since Last Session [ga:daysSinceLastSession](#) User Defined Value [ga:userDefinedValue](#) User Bucket [ga:userBucket](#)

## Metrics



# googleAnalyticsR

Welcome to the website for `googleAnalyticsR`, an R library for working with [Google Analytics](#) data.



Follow development on the project's [Github development site](#).

The Slack group `googleAuthRverse` includes a `#googleAnalyticsR` channel. For news, chat and support join via this [request form](#).

Collaboration is welcomed and encouraged, if you are interested get in touch.

## Features

- First [Google Analytics Reporting v4 API](#) library for R
- Automatic [anti-sampling techniques](#) to return more detailed data
- v4 features include: metric expressions, pivots, date comparisons, batching.
- Auto-paging, [auto-authentication](#) options
- API metadata of possible metrics and dimensions
- GA360 support for features such as [resource quotas](#).
- Multi-user login to enable [Google Analytics powered Shiny Apps](#)
- Integration with [BigQuery Google Analytics Premium/360 exports](#).
- Single authentication flow can be used with other `googleAuthR` apps like `searchConsoleR`
- Automatic batching, multi-account fetching, multi-channel funnels

## Links

Download from CRAN at

[https://cloud.r-project.org/  
package=googleAnalyticsR](https://cloud.r-project.org/package=googleAnalyticsR)

Report a bug at

[https://github.com/MarkEdmondson1234/  
googleAnalyticsR/issues](https://github.com/MarkEdmondson1234/googleAnalyticsR/issues)

## License

[MIT](#) + file [LICENSE](#)

## Developers

Mark Edmondson

Author, maintainer

[All authors...](#)

## Dev status

[downloads 5565/month](#)

[CRAN 0.7.1](#)

[build error](#)

# Facebook Prophet

## Forecasting Algorithm

PROPHET



# Facebook Prophet

<https://facebook.github.io/prophet/>

## Claims:

- Accurate & Fast
- Fully automatic
- Tunable forecasts
- Available in R & Python

**My experience**  
Prophet works well until it doesn't.  
  
Key Point: No silver bullet.

The screenshot shows the official website for Facebook Prophet. At the top right are links for "Docs" and "GitHub". Below the header is a large "PROPHET" logo consisting of three white dots above a white circular arrow icon. The main heading is "Forecasting at scale." followed by a brief description: "Prophet is a forecasting procedure implemented in R and Python. It is fast and provides completely automated forecasts that can be tuned by hand by data scientists and analysts." Below this are four buttons: "INSTALL PROPHET", "GET STARTED IN R", "GET STARTED IN PYTHON", and "READ THE PAPER". A detailed description of the Prophet procedure follows, along with four sections: "Accurate and fast.", "Fully automatic.", "Tunable forecasts.", and "Available in R or Python.". The bottom right corner features a large "PROPHET" logo.

PROPHET

Forecasting at scale.

Prophet is a forecasting procedure implemented in R and Python. It is fast and provides completely automated forecasts that can be tuned by hand by data scientists and analysts.

INSTALL PROPHET GET STARTED IN R GET STARTED IN PYTHON READ THE PAPER

Prophet is a procedure for forecasting time series data based on an additive model where non-linear trends are fit with yearly, weekly, and daily seasonality, plus holiday effects. It works best with time series that have strong seasonal effects and several seasons of historical data. Prophet is robust to missing data and shifts in the trend, and typically handles outliers well.

Prophet is open source software released by Facebook's Core Data Science team. It is available for download on CRAN and PyPI.

**Accurate and fast.**

Prophet is used in many applications across Facebook for producing reliable forecasts for planning and goal setting. We've found it to perform better than any other approach in the majority of cases. We fit models in Stan so that you get forecasts in just a few seconds.

**Fully automatic.**

Get a reasonable forecast on messy data with no manual effort. Prophet is robust to outliers, missing data, and dramatic changes in your time series.

**Tunable forecasts.**

The Prophet procedure includes many possibilities for users to tweak and adjust forecasts. You can use human-interpretable parameters to improve your

**Available in R or Python.**

We've implemented the Prophet procedure in R and Python, but they share the same underlying Stan code for fitting. Use whatever language you're

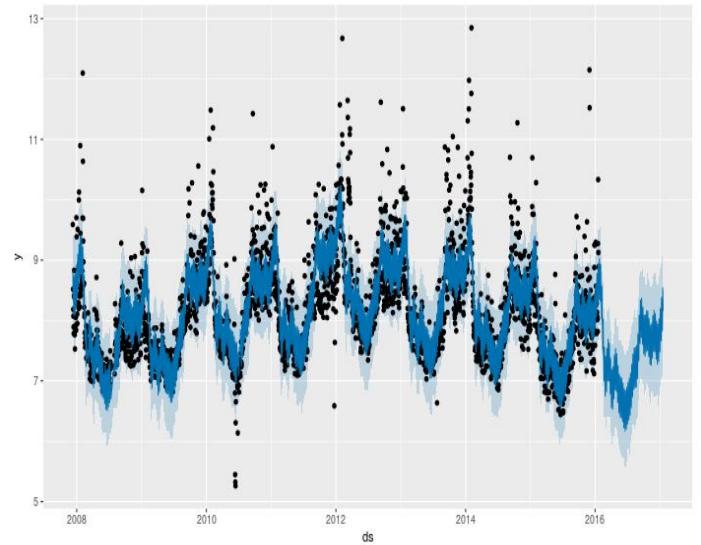
PROPHET

# How Prophet works



**...additive model where non-linear trends are fit with yearly, weekly, and daily seasonality, plus holiday effects**

**Prophet Model**

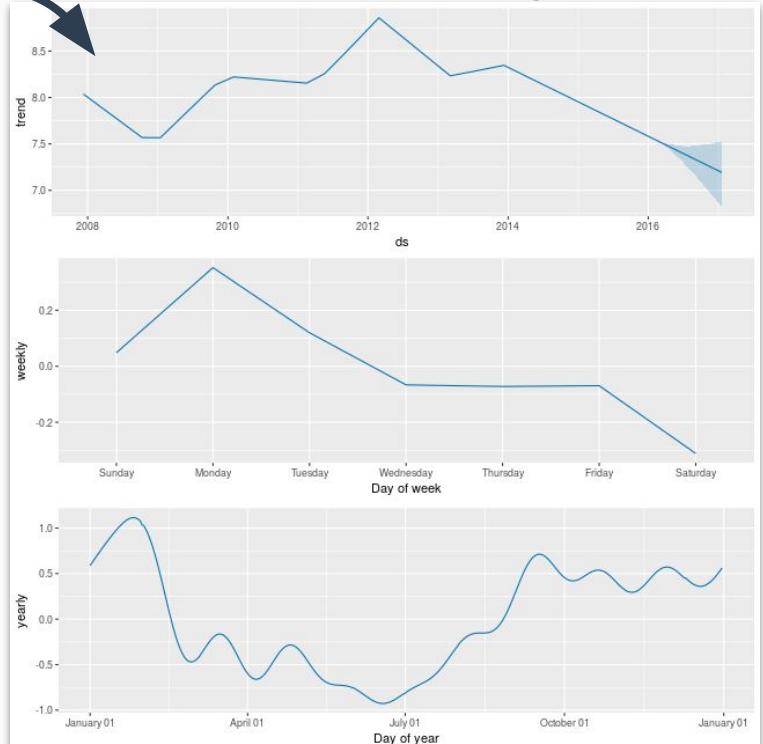


Trends modeled by selecting changepoints

Weeks modeled with dummy variables

Seasonal with Fourier Series

...is 3 models added together





# Fable Implementation

fable.prophet 0.1.0    [Introduction](#)    [Reference](#)    [Changelog](#)

## fable.prophet

This package provides a tidy R interface to the prophet forecasting procedure using [fable](#). This package makes use of the [prophet package](#) for R.

### Installation

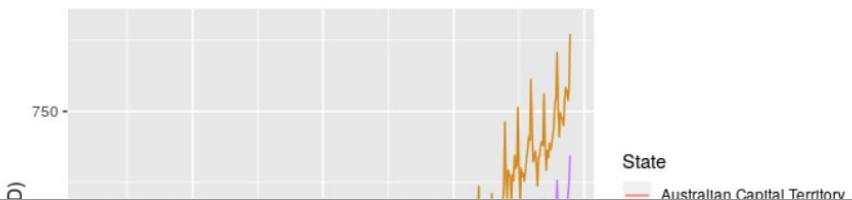
You can install the development version of fable.prophet from [Github](#) with:

```
# install.packages("remotes")
remotes::install_github("mitchelloharawild/fable.prophet")
```

### Example

Suppose we wanted to model Australia's monthly turnover for cafes, restaurants and catering services. The data is available from the Australian Bureau of Statistics catalogue 8501.0, and in the [tsibbledata](#) package.

```
library(tsibble)
library(dplyr)
cafe <- tsibbledata::aus_retail %>%
  filter(Industry == "Cafes, restaurants and catering services")
```



### Links

Report a bug at  
<https://github.com/mitchelloharawild/fable.prophet/issues>

### License

[GPL-3](#)

### Community

[Contributing guide](#)

[Code of conduct](#)

### Developers

[Mitchell O'Hara-Wild](#)

Author, maintainer

[All authors...](#)

### Dev status

[build](#) failing

[codecov](#) 64%

[lifecycle](#) experimental



# Workflow

From zero to  
Forecast automation

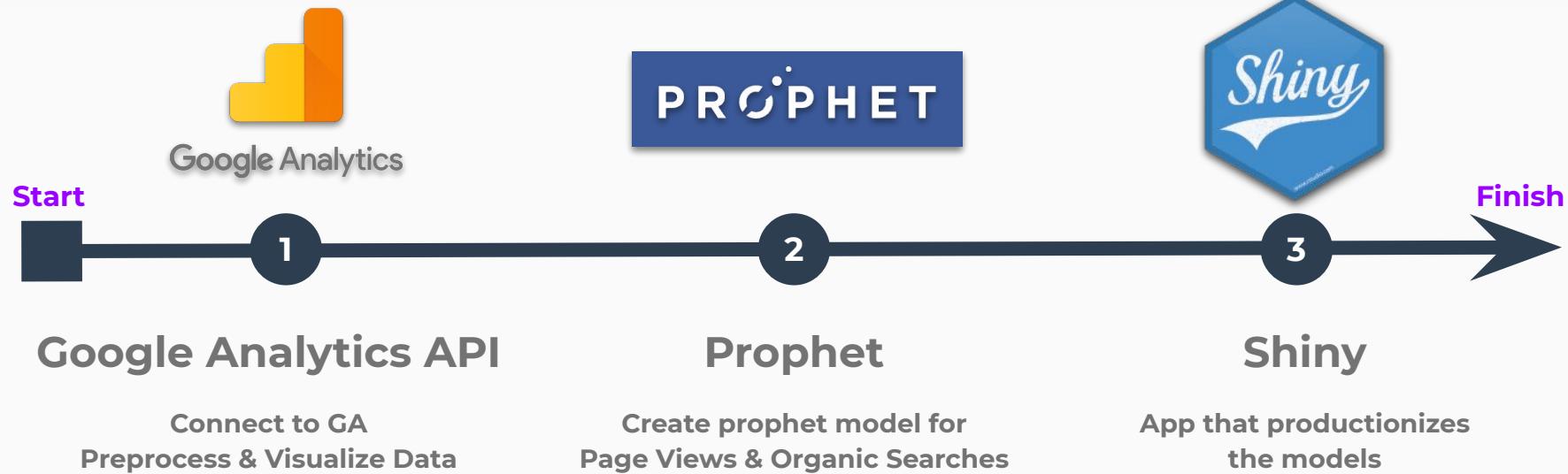


Google Analytics

PROPHET



# Workflow Step-By-Step



# 30-Min Demo

Let's do this!

# PRO-TIPS

Yeahhhhhh!

# Pro-Tip #1: Give Businesses Apps



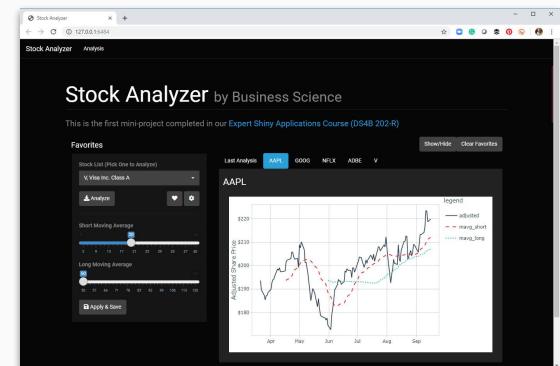
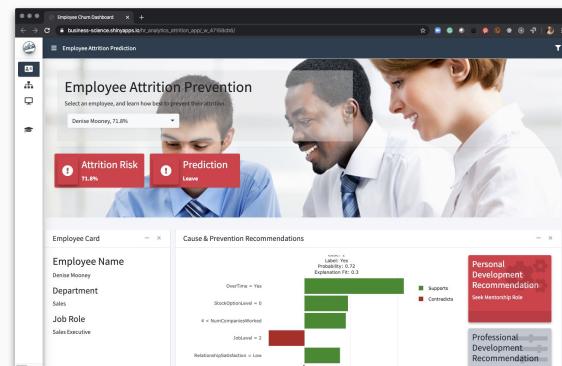
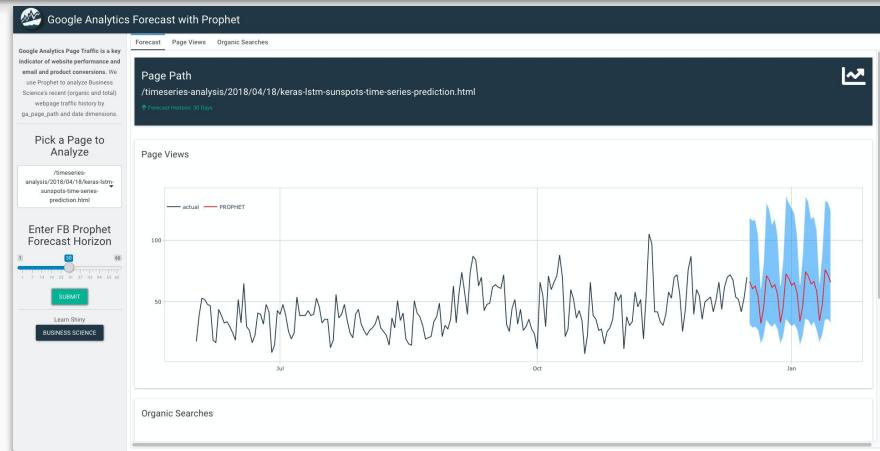
Apps are what businesses need

Businesses can't scale reports

Businesses have cloud & servers

Businesses need data democratized

Businesses can solve these problems with shiny apps



Apps solve these challenges

# Pro-Tip #2: Say no to reports



# No

arXiv:1811.08963v1 [cs.LG] 21 Nov 2018

## Multivariate Forecasting of Crude Oil Spot Prices using Neural Networks

Ganapathy S. Natarajan

School of Mechanical, Industrial, and Manufacturing Engineering  
Oregon State University  
Corvallis, OR, USA  
Email: gana.natarajan@oregonstate.edu

Aishwarya Ashok

Department of Computer Science and Engineering  
University of Texas at Arlington  
Arlington, TX, USA  
Email: aishwarya.ashok@mavs.uta.edu

**Abstract**—Crude oil is a major component in most advanced economies of the world. Accurately predicting and understanding the behavior of crude oil price is important for the energy analysts, forecasters, and traders, to name a few. The price of crude oil has declined in the past decade and is seeing a phase of stability; but will this stability last? This work is an empirical study on how multivariate analysis may be employed to predict crude oil spot prices using neural networks. The concept of using neural networks showed promising potential. A very simple neural network model was able to perform on par with ARIMA models - the state-of-the-art model in time-series forecasting. Advanced neural network models using larger datasets may be used in the future to extend this proof-of-concept to a full scale framework.

**Keywords:**crude oil; multivariate forecasting; neural networks; ARIMA; regression

### I. INTRODUCTION

Crude oil spot prices saw a tremendous up-tick in the first decade of the 21<sup>st</sup> century. Since 2014, crude oil prices have fallen and may have stabilized now. However, there has always been a constant interest in accurately predicting crude oil prices; given that crude oil drives a major portion of the economy. Economists, scientists, data analysts, and traders are all interested in models that give them the best accuracy. In the last decade, advances in machine learning has enabled everyday data analysts to use techniques like neural networks; case-in-point is Google's TensorFlow™.

Crude oil prices have mostly been forecasted using time-series methods. There is some work in dealing with crude oil price forecasting as an econometric problem; however, there is very limited work in using multivariate techniques that move away from traditional regression modeling. This paper combines the lack of multivariate forecasting and advances in machine learning to provide a proof-of-concept for using Neural Networks in multivariate forecasting of crude oil prices. The results presented in this study are mostly empirical and lays the foundation for more in-depth studies in this direction.

### II. RELATED WORK

We submitted a literature review paper to a conference [1]. This work looked at the different types of models and methods used in forecasting oil prices in general. Time series

models were predominantly used. Time series models use just the oil price over time to predict future prices. ARIMA group of methods were the most commonly used time series forecasting methods [2]–[5] either for model building or as a benchmark or for a part of the model building. Artificial intelligence and machine learning based models [6]–[10] including neural network models [11] were also used to just predict the time series or components of the time series.

The second type of approach to forecasting oil prices was econometric models. Econometric models usually perform linear regression based analysis and take into account economic factors. Some of the common independent variables used were income growth, price increase, and population growth [12], gas tax exemption [13], OECD stocks, and OPEC spare capacity [14]. OECD is the Organization for Economic Co-operation and Development - a group of developed economies in the world, and OPEC is the Organization of the Petroleum Exporting Countries. Some of the variables used in econometric studies historically, will be used in this study too, viz. CPI, population, petroleum stocks, to name a few.

The third commonly used methods are multivariate models, which are of interest in this paper. Multivariate models use multiple variables akin to econometric models, but do not always assume a linear relationship or use just regression. Some of the commonly used multivariate methods include multivariate time series models [15], neural networks [16], [17], and multi-criteria decision models [18]. Reference [16] used neural networks to predict commercial oil price using country of origin, sulfur content, and density, and [17] used wavelet neural network to predict monthly crude oil price using just the inventory levels - not truly a multivariate methodology.

Recent works have continued with the trend of using different methods but still mostly perform time-series forecasting. Although the recent works use hybrid models [19], ensemble neural network models [20], gray wave model [21], and wavelet neural network models [22], they still are time-series forecasting models. Reference [23] used network science and predicted parts of a time series to deal with specific non-linear patterns.

The common pattern in all the related work is that most of

# Pro-Tip #3: ...and say YES to apps.



# Yes

The screenshot shows a web browser displaying the "App Gallery" section of the "Apps by Business Science" website at [apps.business-science.io](https://apps.business-science.io). The page features a dark header with tabs for "App Gallery", "Search Apps", "Submit", and "Clear". Below the header is a navigation bar with categories: ALL, BUSINESS, FINANCE, HUMAN RESOURCES, MARKETING, and SALES. The "ALL" tab is currently selected. Three app cards are displayed:

- Application Library** (Business, Shiny, AWS)  
with Full-Text Search  
A meta-app for hosting multiple business applications for an organization. Includes Full-Text Search and advanced filtering capability. Hosted on AWS.  
[Launch App](#) | [Build It in DS4B 202A-R](#)
- Stock Analyzer** (Finance, Shiny, AWS, MongoDB, Auth)  
Multi-User App (MongoDB)  
A financial application for analyzing trends in your favorite SP 500 stocks. Leverages AWS EC2 and MongoDB Atlas Cloud for managing multiple users.  
[Launch App](#) | [Build It in DS4B 202A-R](#)
- HR Employee Attrition Prevention** (Human Resources, Shiny, H2O, Bootstrap 4)  
Apply H2O machine learning to predict and explain employee churn. Identify which departments, job roles, job levels, and people have the highest expected attrition risk.  
[Launch App](#) | [Coming Soon](#)

[apps.business-science.io](https://apps.business-science.io)



# 4-Course R-Track System



# Business Analysis with R (DS4B 101-R)

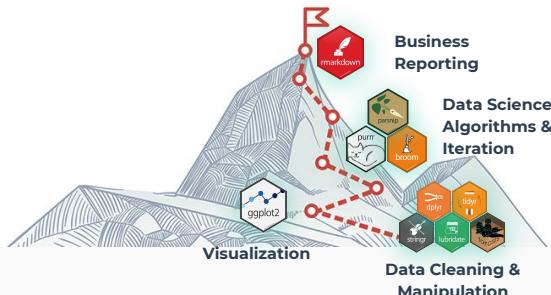
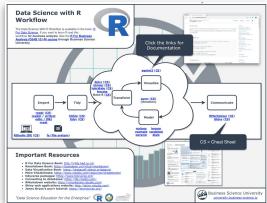
# Data Science For Business with R

## (DS4B 201-R)

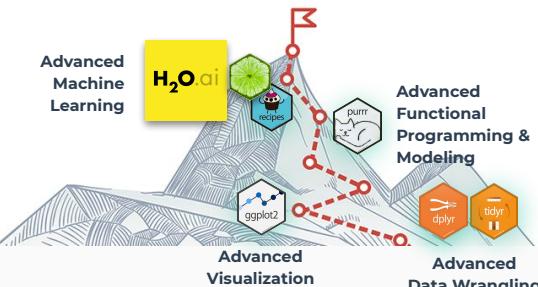
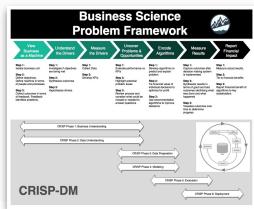
Web Apps & Shiny Developer  
(DS4B 102-R + DS4B 202A-R)

## Project-Based Courses with Business Application

## Data Science Foundations **7 Weeks**

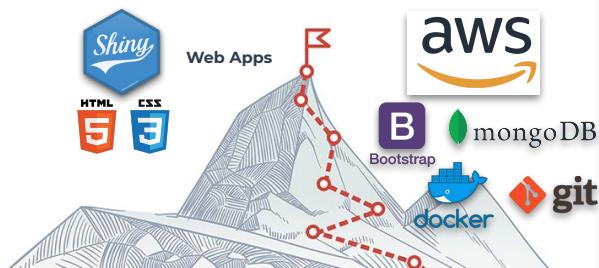
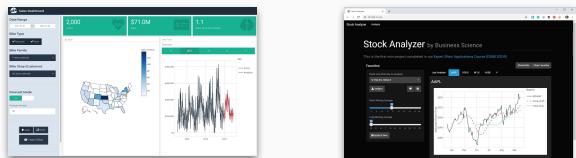


Machine Learning & Business Consulting  
**10 Weeks**



# Web Application Development

## 12 Weeks

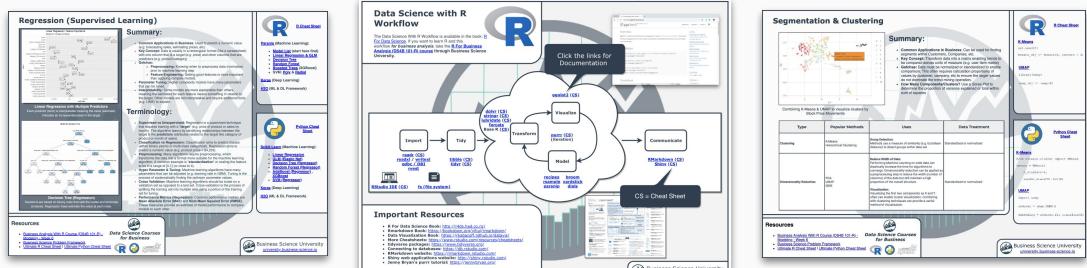


# Key Benefits

- Fundamentals - Weeks 1-5 (25 hours of Video Lessons)
  - Data Manipulation (dplyr)
  - Time series (lubridate)
  - Text (stringr)
  - Categorical (forcats)
  - Visualization (ggplot2)
  - Programming & Iteration (purrr)
  - 3 Challenges
- **Machine Learning - Week 6 (8 hours of Video Lessons)**
  - Clustering (3 hours)
  - Regression (5 hours)
  - 2 Challenges
- Learn Business Reporting - Week 7
  - RMarkdown & plotly
  - 2 Project Reports:
    1. Product Pricing Algo
    2. Customer Segmentation

# Business Analysis with R (DS4B 101-R)

Data Science Foundations  
**7 Weeks**



# Key Benefits

## End-to-End Churn Project

Understanding the Problem & Preparing Data - Weeks 1-4

- Project Setup & Framework
- Business Understanding / Sizing Problem
- Tidy Evaluation - rlang
- EDA - Exploring Data -GGally, skimr
- Data Preparation - recipes
- Correlation Analysis
- 3 Challenges

## Machine Learning - Weeks 5, 6, 7

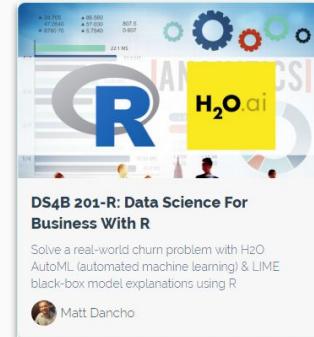
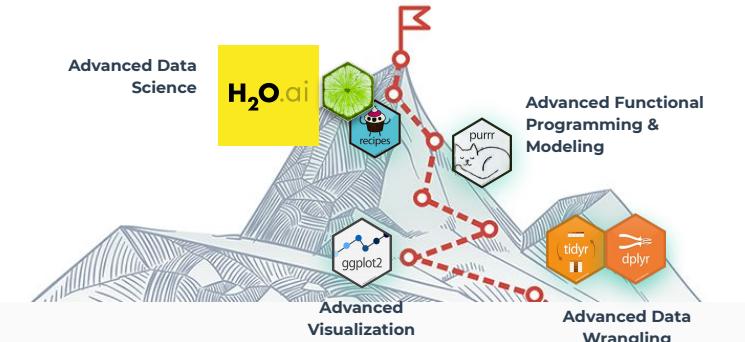
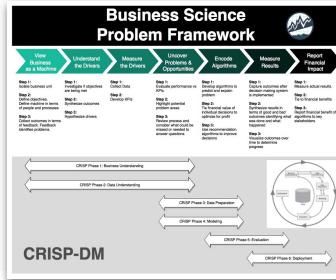
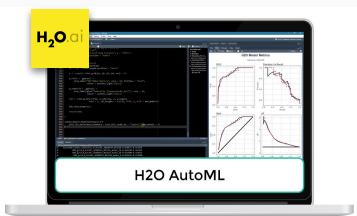
- H2O AutoML - Modeling Churn
- ML Performance
- LIME Feature Explanation

## Return-On-Investment - Weeks 7, 8, 9

- Expected Value Framework
- Threshold Optimization
- Sensitivity Analysis
- Recommendation Algorithm

# Data Science For Business (DS4B 201-R)

Machine Learning & Business Consulting  
**10 Weeks**



# Key Benefits

## Learn Shiny & Flexdashboard

- Build Applications
- Learn Reactive Programming
- Integrate Machine Learning

## App #1: Predictive Pricing App

- Model Product Portfolio
- XGBoost Pricing Prediction
- Generate new products instantly

## App #2: Sales Dashboard with Demand Forecasting

- Model Demand History
- Segment Forecasts by Product & Customer
- XGBoost Time Series Forecast
- Generate new forecasts instantly

# Shiny Apps for Business (DS4B 102-R)



Web Application Development  
**4 Weeks**

The collage includes:

- A "Data Science with R" course page featuring a "Predictive Pricing App" dashboard.
- A "Flexdashboard Apps" section showing a dashboard with a map of the US and time series plots.
- A "Shiny Apps" section showing a dashboard with a scatter plot and a histogram.
- A "Themes, Dashboards, & Examples" section showing a dashboard with multiple panels and a sidebar.
- A "Business Analytics" section showing a dashboard with a map and a bar chart.
- A "Machine Learning" section showing a dashboard with a scatter plot and a sidebar.
- A "Data Science with R" course page featuring a "Sales Dashboard with Demand Forecasting" dashboard.



The collage includes:

- A "Shiny" logo and a bar chart.
- A "DATA ANALYTICS" section with a large blue "R" icon.
- A "Machine Learning" section with a green gear icon.
- A "Shiny" logo and a bar chart.
- A "DS4B 102-R: Shiny Web Applications for Business (Level 1)" course page.
- A "Build a predictive web application using Shiny, Flexdashboard, and XGBoost" section.
- A photo of Matt Dancho.

# Key Benefits

Frontend + Backend + Production Deployment

## Frontend for Shiny

- Bootstrap

## Backend for Shiny

- MongoDB Atlas Cloud
- Dynamic UI
- User Authentication & Security

## Production Deployment

- AWS
- EC2 Server
- SSL & HTTPS Encryption

# Shiny Apps for Business (DS4B 202A-R)



Web Application Development  
**6 Weeks**



# 15% OFF PROMO Code: **learninglabs**



## R-TRACK BUNDLE

**4-Course Bundle - Machine Learning + Expert Web Applications (R-Track)**

Go from Beginner to Expert Data Scientist & Shiny Developer in Under 6-Months

4 Course Bundle ~~\$1,500~~

**\$127/mo  
Limited Time**

**DS4B 101-R: Business Analysis With R**

Your Data Science Journey Starts Now! Learn the fundamentals of data science for business with the tidyverse.

Matt Dancho

**DS4B 102-R: Shiny Web Applications For Business (Level 1)**

Build a predictive web application using Shiny, Flexdashboard, and XGBoost.

Matt Dancho

**DS4B 201-R: Data Science For Business With R**

Solve a real-world churn problem with H2O AutoML (automated machine learning) & LIME black-box model explanations using R.

Matt Dancho

**DS4B 202A-R: Expert Shiny Developer with AWS**

Learn how to build Scalable Data Science Applications using R, Shiny, and AWS Cloud Technology.

Matt Dancho

<input type="radio"/>	<b>Paid Course</b> 15% COUPON DISCOUNT	\$1,596 <b>\$2,356.60</b>
<input checked="" type="radio"/>	<b>12 Low Monthly Payments</b> 15% COUPON DISCOUNT	12 payments of \$149/m <b>12 payments of \$126.65/m</b>

# Career acceleration awaits

[university.business-science.io](http://university.business-science.io)

