

# Product Pricing Algorithm

Business Science

3/19/2019

## Problem Statement

Research and Development wants help to determine new product ideas and pricing using existing product line as a benchmark.

## Solution Summary

We've identified several product gaps in the existing product line including:



Making a  
Business Report  
that integrates  
Machine Learning

# Business Reporting in R

Communicating Machine Learning With  
The **RMarkdown** package



1 6070 Jekyll Carbon 2

Over Mountain

Carbon

Matt Dancho & David Curry  
**Business Science Learning Lab**

# Accelerating Your Career

## True Story

I received 3 promotions in 2 years

How did I do it?

I started making reports that my boss valued

- PDF Format
- Taking Existing Process & Making Better



## Product Pricing Algorithm

Business Science

3/19/2019

### Problem Statement

Research and Development wants help to determine new product ideas and pricing using existing product line as a benchmark.

### Solution Summary

We've identified several product gaps in the existing product line including:

1. Aluminum Over Mountain
2. Aluminum Triathalon

The Data Science Team has developed a pricing model that uses predictive analytics to estimate the price of the new bicycle models based on the existing fleet. This ensures that new models are priced comparatively to other similar bicycles.

**Next Steps:** Integrate the model into a proof-of-concept web application that can be deployed to the R&D department.

### Gap Analysis

#### Bike List

Our current product portfolio consists of 97 bike models.

```
## # A tibble: 97 x 15
##   id price model category_1 category_2 frame_material model_base
##   <int> <dbl> <chr> <chr> <chr> <chr> <chr>
## 1 1     6070 Jeky-Mountain Over Moun- Carbon Jekyll
## 2 2     5970 Trig-Mountain Over Moun- Carbon Trigger
## 3 3     2770 Beas-Mountain Trail   Aluminum Beast of -
## 4 4     10660 Supe-Road    Elite Road Carbon Supersix -
## 5 5     3200 Jeky-Mountain Over Moun- Carbon Jekyll
## 6 6     12790 Supe-Road   Elite Road Carbon Supersix -
## 7 7     5330 Supe-Road   Elite Road Carbon Supersix -
## 8 8     1570 Syna-Road   Endurance- Aluminum Synapse
## 9 9     4800 Syna-Road   Endurance- Carbon Synapse
## 10 10    480 Cata-Mountain Sport   Aluminum Catalyst
## # ... with 87 more rows, and 8 more variables: model_tier <chr>,
## #   black <dbl>, hi_mod <dbl>, team <dbl>, red <dbl>, ultegra <dbl>,
## #   dura_ace <dbl>, disc <dbl>
```

1



# Learning Lab Structure

- **Presentation**  
(15 & 15 min)
- **Demo's**  
(30 min)
- **Q&A**  
(10 mins)

## Your Hosts!



**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**.

# Agenda



## • Why Machine Learning?

- Key Benefits
- High Impact Problems
- Data Science Workflow

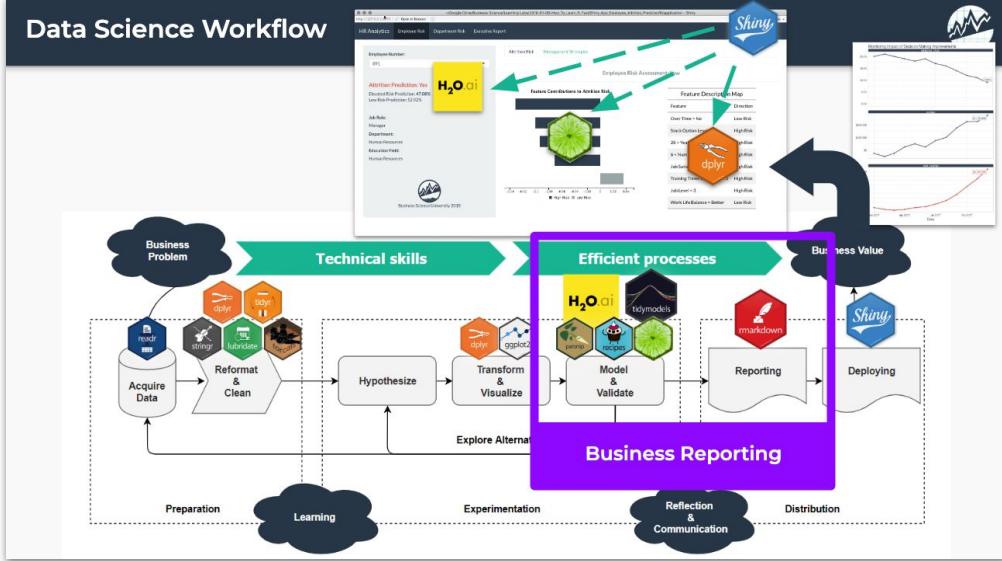
## • Effective Communication

- How to communicate
- What Tools Are Available
- **RMarkdown** R Package

## • Demonstration

- Report 1: **Product Pricing Algo**
- Report 2: **Customer Segmentation**

## Data Science Workflow



## • Bonus & Recap

- A little fun at the end :)

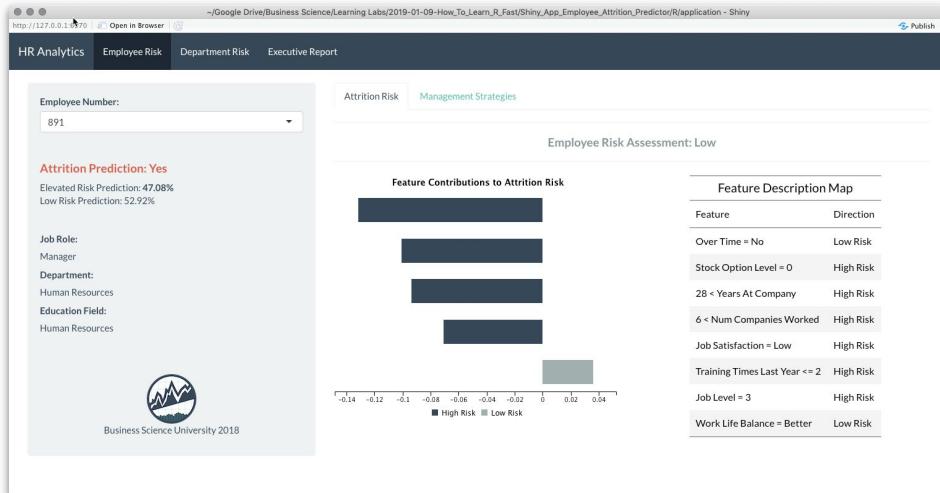


# Why Learn Machine Learning?

Most **in-demand** skill of the century

**Prediction** - Developing a model that learns a complex system, generalizes to new information

**Explanation** - Using the model to understand which features are important



# Machine Learning Solves High Impact Problems



## Sales

- Customer Churn
- Demand Forecasting
- Product Backorders
- Pricing Optimization

## Finance

- Fraud Detection
- Credit Default Risk
- Loan Delinquency

## HR

- Employee Attrition
- Promotion Readiness

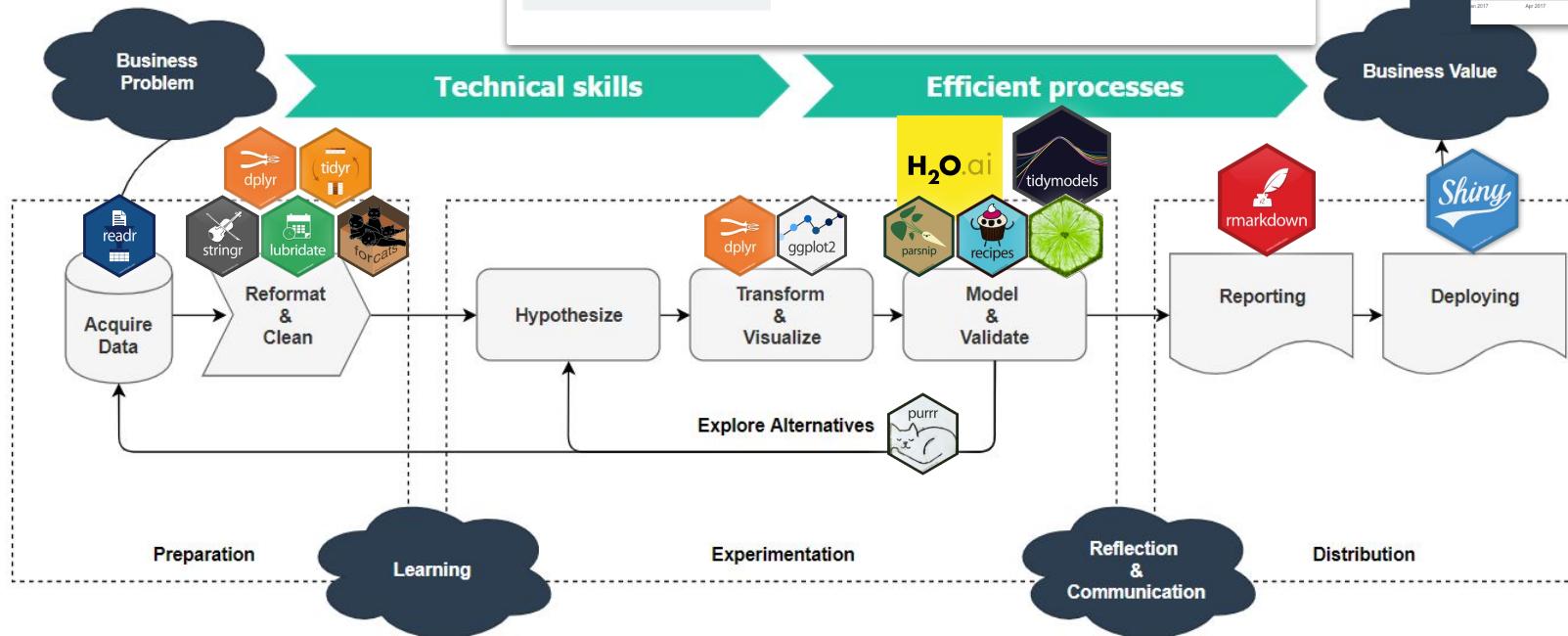
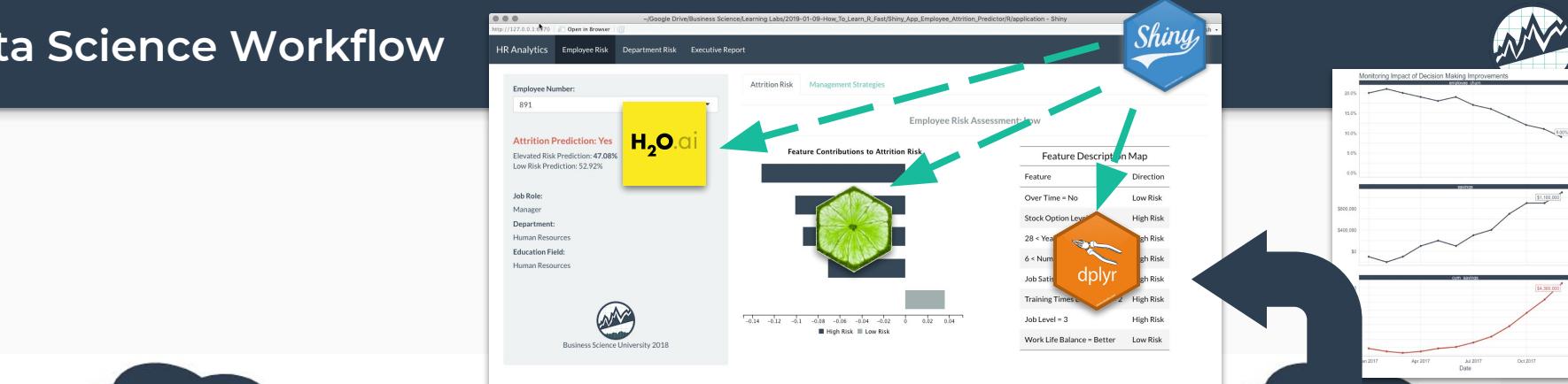
## Marketing

- Customer Journey
- Customer Segmentation
- Spend & Conversion Optimization
- Web Traffic Anomalies

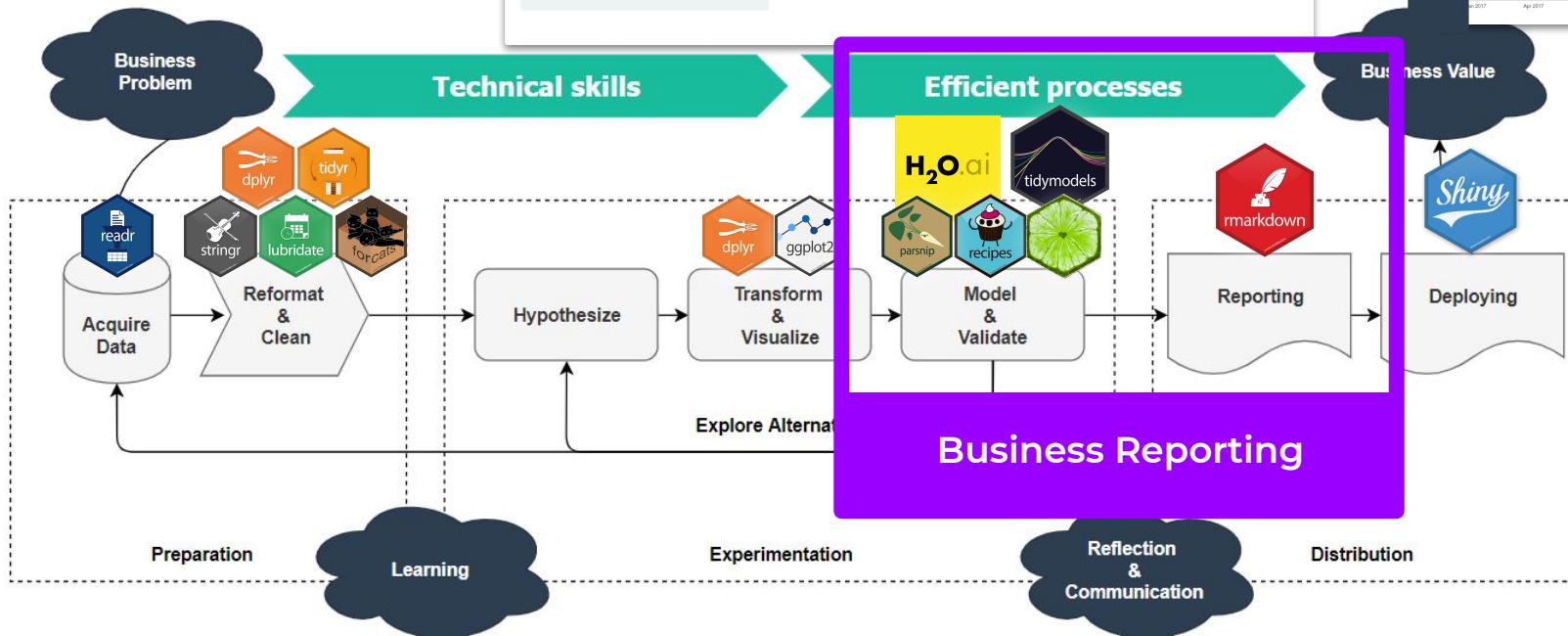
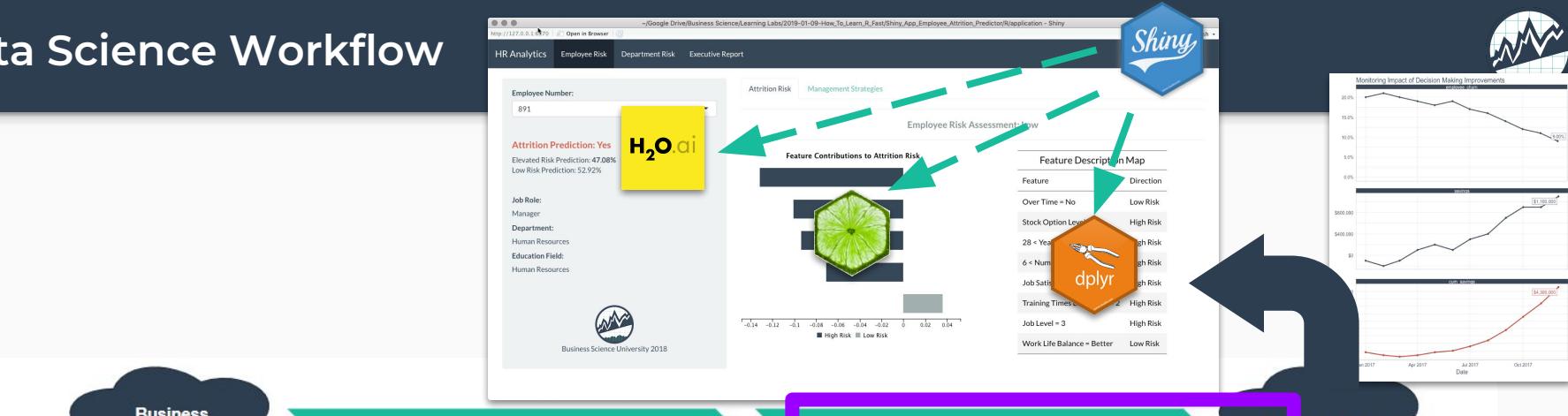
## Accounting

- Cash Flow Forecasting
- Payment Anomaly Detection

# Data Science Workflow



# Data Science Workflow



# **Effective Communication**

What needs to happen

# 3 components to effective Data Science communication



**Problem Statement**

**Solution Statement**

**Everything Else**

# Problem & solution statements



- **Problem Statement**
  - Goal
  - \$\$\$ (if possible)
- **Solution Statement**
  - Outcome
  - Savings/Benefit (if possible)
- **Everything Else**
  - Explain visually
  - Limit technical jargon

## Customer Segmentation

*Business Science*

3/19/2019

### Problem Statement

Marketing would like to increase email campaign engagement by segmenting the customer-base using their buying habits.

### Solution Summary

We've identified 4 customer segments that can be used to customize email engagement:

1. Segment 1 Preferences: Road Bikes, Below \$4000 (Economical Models)
2. Segment 2 Preferences: Mountain Bikes, Above \$3000 (Premium Models)
3. Segment 3 Preferences: Road Bikes, Above \$3000 (Premium Models)
4. Segment 4 Preferences: Below \$3000 (Economical Models)

### Customer Preferences

#### Heat Map

Our customer-base consists of 30 bike models. Several customers have purchasing preferences for Road or Mountain Bikes based on the proportion of bikes purchased by category\_1 and category\_2.

Heatmap of Purchasing Habits

|                  | Mountain           |      |      |       |       | Road |       |       |      |
|------------------|--------------------|------|------|-------|-------|------|-------|-------|------|
|                  | Albuquerque Cycles | 3.1% | 4.5% | 12.2% | 13.3% | 2.4% | 24.1% | 18.9% | 4.5% |
| Ann Arbor Speed  | 5.3%               | 0.7% | 3.2% | 7.6%  | 7.1%  | 6.6% | 37.0% | 23.6% | 8.8% |
| Austin Cruisers  | 8.5%               | 0.8% | 2.0% | 11.8% | 6.5%  | 3.7% | 37.4% | 19.9% | 9.3% |
| Cincinnati Speed | 15.3%              | 0.3% | 7.2% | 0.8%  | 6.1%  | 6.1% | 30.9% | 23.8% | 9.5% |



- **RMarkdown Uses:**

- Reports (HTML, PDF, Word)
- Slide Decks (HTML, PowerPoint)
- Web Apps
- Web Pages
- Books

- Part of **tidyverse**

- **Links**

- [rmarkdown.rstudio.com](https://rmarkdown.rstudio.com)
- [rmarkdown.rstudio.com/gallery](https://rmarkdown.rstudio.com/gallery)

The screenshot shows the R Markdown website. The top navigation bar includes links for "Get Started", "Gallery", "Formats", "Articles", and "Book". The main content area features a large image of a map of the United States with various data overlays, including density gradients and scatter plots. Below the image, the text "Analyze. Share. Reproduce." is displayed, followed by the tagline "Your data tells a story. Tell it with R Markdown. Turn your analyses into high quality documents, reports, presentations and dashboards." At the bottom of the page, a code editor window shows R code related to "Introduction to R Markdown".



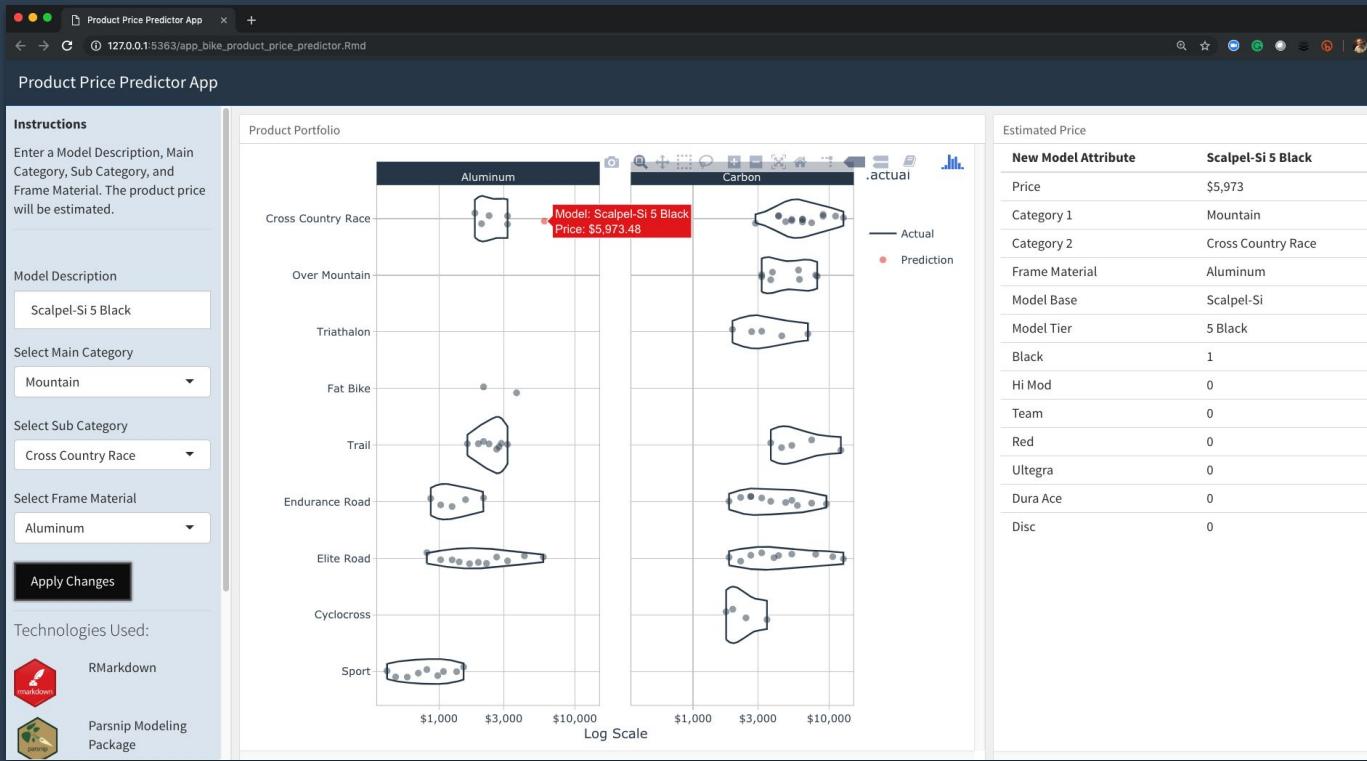
# Demo Time

RMarkdown in action

- *Product Pricing Algo*
- *Customer Segmentation*

# Bonus

You can do this with RMarkdown



# Learn Business Reporting



## Business Analysis with R

- **7 Week System**
  - Data Science Fundamentals
- **Week 7 (Business Reporting)**
  - RMarkdown
  - plotly interactive plotting



The thumbnail for the course "DS4B 101-R: Business Analysis With R" features a collage of various data visualization elements. It includes a line chart with numerical values (24,705, 67,0940, 86,590, 57,400, 807.5, 6780.70, 0.697), a bar chart showing "22.1 MS", several colored gears, a large blue "R" logo, a dark circular badge with the word "tidyverse", and small icons of people and charts.

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

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

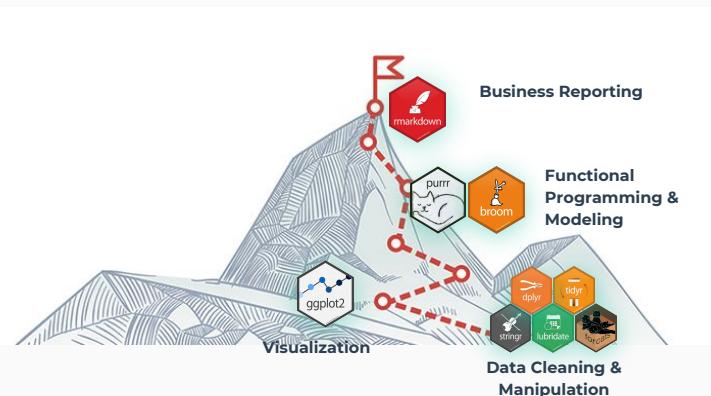
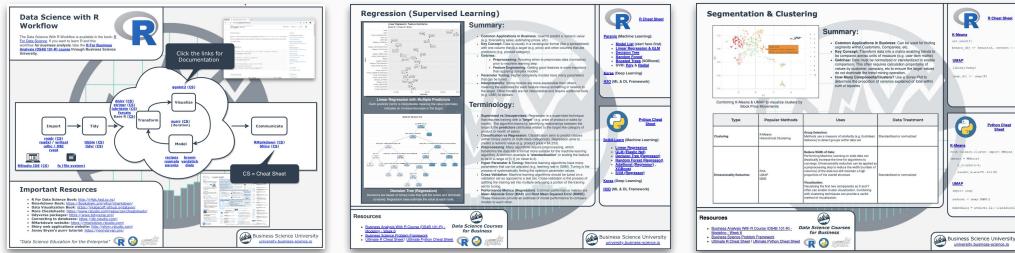
 Matt Dancho

# 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**



# Learn Machine Learning

*Most Critical Business Problems  
are Binary Classification (Yes/No)*

*Will this customer churn?*

*Is this employee going to leave?*

*Will this customer going to miss payment  
in next 90-days?*

## Data Science for Business with R

- **10 Week System**
  - Advanced Machine Learning
- **Churn with H2O AutoML**



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

# Data Science For Business with R

## (DS4B 201-R)



# **Weeks 5 & 6: H2O**

## Modeling & Performance

In **Week 5 (modeling)**, you learn generate 30+ models & visualize results using ggplot2

In **Week 6 (performance)**, you go in-depth learning ROC & AUC, Precision vs Recall, & Gain & Lift Plots

# Data Science For Business with R (DS4B 201-R)



## Week 7: Explaining Black-Box Models

*“The business won’t care how high your AUC is if you can’t explain your Machine Learning Model”*

In **Week 7 (Explaining Models)**, you use LIME to explain the “black-box” ensemble models & extract insights to answer **what is causing churn**

# Distribute Results

*Businesses need Web Applications*

*Interactive Visualizations*

*Dashboard Interface*

*Built-In Prediction & Explanation*

## Data Science for Business with R

- **6 Week System**
  - Integrate ML into
  - Web App

DS4B 301-R: Shiny Web Apps For Business

Build Distributed Web Applications using Machine Learning

Coming Soon!



# R Shiny Web Apps For Business

## (DS4B 301-R)

~/Google Drive/Business Science/Learning Labs/2019-01-09-How\_To\_Learn\_R\_Fast/Shiny\_App\_Employee\_Attrition\_Predictor/R/application - Shiny  
http://127.0.0.1:870 / Open in Browser

HR Analytics Employee Risk Department Risk Executive Report

Employee Number: 891

Attrition Prediction: Yes  
Elevated Risk Prediction: 47.08%  
Low Risk Prediction: 52.92%

Job Role: Manager  
Department: Human Resources  
Education Field: Human Resources

**H2O.ai**

Attrition Risk Management Strategies

Employee Risk Assessment: Low

Feature Contributions to Attrition Risk

Feature Description Map

Feature Direction

|                            |           |
|----------------------------|-----------|
| Over Time = No             | Low Risk  |
| Stock Option Level = 1     | High Risk |
| 28 < Years = No            | High Risk |
| 6 < Num of Children = No   | High Risk |
| Job Satisfaction = High    | High Risk |
| Training Times Less than 2 | High Risk |
| Job Level = 3              | High Risk |
| Work Life Balance = Better | Low Risk  |





# Business Science Learning System

## THE PLAN



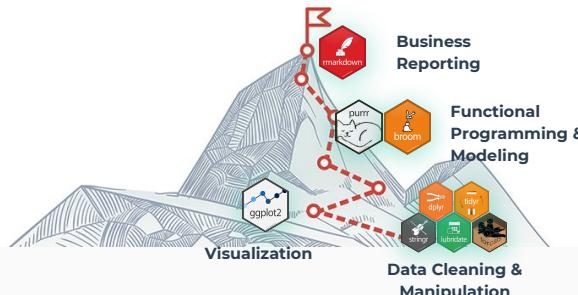
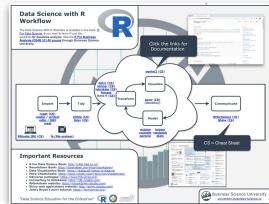
### Business Analysis with R (DS4B 101-R)

### Data Science For Business with R (DS4B 201-R)

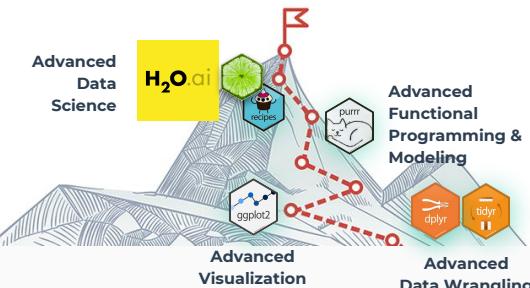
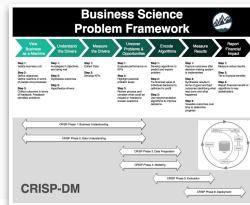
### R Shiny Web Apps For Business (DS4B 301-R)

#### Project-Based Courses with Business Application

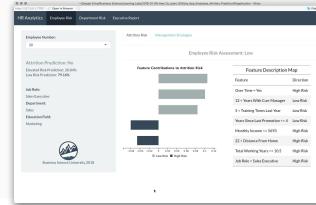
Data Science Foundations  
**7 Weeks**



Machine Learning & Business Consulting  
**10 Weeks**



Web Application Development  
**6 Weeks**





# THE RESULT



**Business Analysis with R**  
(DS4B 101-R)  
7 WEEKS



**Data Science For Business with R**  
(DS4B 201-R)  
10 WEEKS



**R Shiny Web Apps For Business**  
(DS4B 301-R)  
6 WEEKS

**Career Acceleration**  
**23 WEEKS!**  
**(OR LESS)**



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



**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

MSRP: ~~\$349~~

Your Price! **\$297**

Save: \$52

**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

MSRP: ~~\$499~~

Your Price! **\$421**

Save: \$78

Courses Included with Purchase

**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 **\$495**

**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 **\$349**

Original Price: \$844



MSRP: ~~\$844~~

Your Price! **\$637**

Save: \$207

# THE BONUS



Start your journey with **15% OFF**  
**PROMO Code: learninglabs**  
[university.business-science.io](http://university.business-science.io)

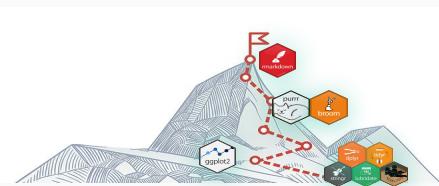
Business Analysis with R  
(DS4B 101-R)

Data Science For Business with R  
(DS4B 201-R)

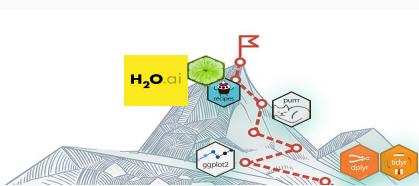
R Shiny Web Apps For Business  
(DS4B 301-R)

## Project-Based Courses with Business Application

Data Science Foundations  
**7 Weeks**



Machine Learning &  
Business Consulting  
**10 Weeks**



Web Application Development  
**6 Weeks**

