



## Employee Attrition Prevention

Select an employee, and learn how best to prevent their attrition.

Denise Mooney, 71.8%

Attrition Risk  
71.8%

Prediction  
Leave

# 6 Reasons Why is Great for Business

## An Introductory Webinar

Job Role

Sales Executive

StockOptionLevel = 0

Incentivize Specialization

Professional Development Recommendation

Matt Dancho & David Curry  
*Business Science Intro Webinars*

2 Application Demos

Machine Learning

And more!





# Learning Lab Structure

- **Intro**  
(10 min)
- **2 Demo's**  
(20 min)
- **6 Reasons**  
(30 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**.

# My Journey

Why I chose R

# About Me

- Consultant & Trainer

- Creator

(250K+ Downloads)

- Founder

(Business Science)



**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.



**S&P Global**

**MRM // McCANN**



**Business Science University**

600+ Students

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

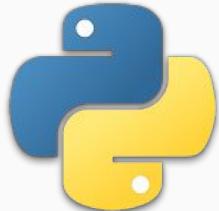




# 2 Types Of Data Scientists

“I Learned **Software First**”

- **Degrees:** Computer Science, Information Technology
- **Object-Oriented Programming:** Comfortable with OOP



“I Learned **Business Analysis First**”

- **Degrees:** Business Analytics, Mechanical Engineering, Statistics
- **Business Analysis:** Power User of Excel, Tableau, PowerBI

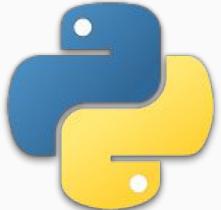




# 2 Types Of Data Scientists

“I Learned **Software First**”

- **Degrees:** Computer Science, Information Technology
- **Object-Oriented Programming:** Comfortable with OOP



“I Learned **Business Analysis First**”

- **Degrees:** Business Analytics, **Mechanical Engineering**, Statistics
- **Business Analysis:** Power User of Excel, Tableau, PowerBI



My Journey

# Becoming A Data Scientist

I struggled with my own learning path

- Initially started with Python
- Felt wrong
- Switched to R, felt way better





# What worked for me?

When I switched to R, I began  
**focusing on problems** rather than  
syntax.

- Designed for **statistics**
- **Tidyverse** helped shorten time from thought to code
- Now has amazing **Machine Learning & Web App** Capabilities





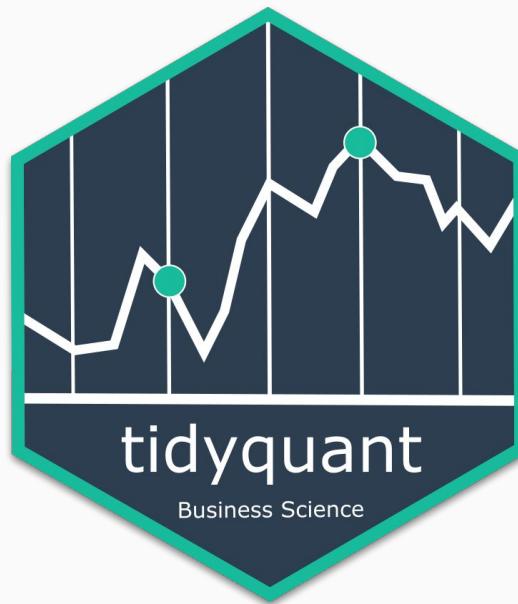
# What worked for me?

Learning R took a while.

I took courses, I read books.

Nothing worked.

When I created **tidyquant**, I learned more than any course would ever teach me.





# What worked for me?

I began **consulting** for organizations, which gave me **real-world experience** proving that R was the right choice.



**S&P Global**

**MRM // McCANN**



Elite Organizations Care About One Thing  
**Results**



# My Tool of Choice



# **Businesses Needs**

## Why learn R to begin with?





# HR Analytics: Employee Attrition

## \$60 Billion Dollar Problem

One of the world's largest mining companies projects that **employee attrition will be a \$60 Billion Dollar Problem** over the next 30 years.

We can solve this with



# Demo #1

## [NEW] HR Analytics App

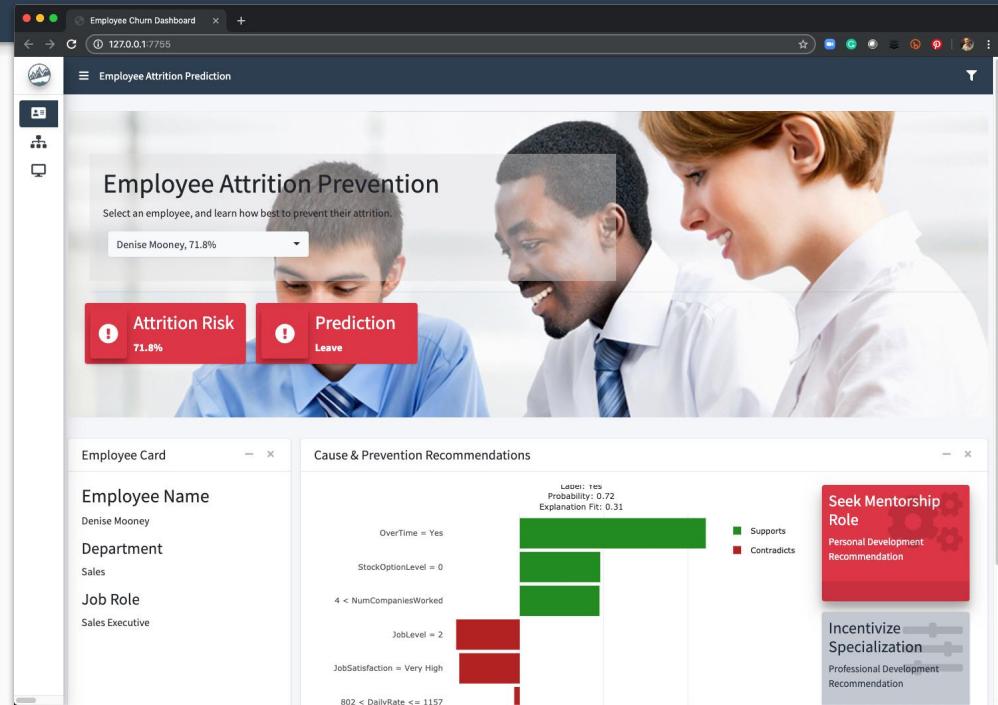


## Employee Attrition Prevention

How does the app work?

Why is it useful to an organization?

What technologies make this?





Employee Churn Dashboard 127.0.0.1:7755

## Employee Attrition Prediction

Select an employee, and learn how best to prevent their attrition.

Denise Mooney, 71.8%

**Attrition Risk** 71.8% **Prediction** Leave

### Employee Attrition Prevention

Employee Card Cause & Prevention Recommendations

**Employee Name:** Denise Mooney  
**Department:** Sales  
**Job Role:** Sales Executive

**Cause & Prevention Recommendations:**

Feature	Value	Probability	Explanation Fit
Overtime	Yes	0.72	0.31
StockOptionLevel	0		
NumCompaniesWorked	4 <		
JobLevel	2		
JobSatisfaction	Very High		
DailyRate	802 < DailyRate <= 1157		

**Seek Mentorship Role:** Personal Development Recommendation

**Incentivize Specialization:** Professional Development Recommendation

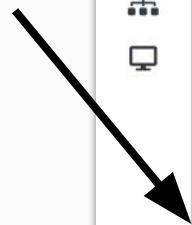


**H<sub>2</sub>O.ai**

Automated  
Machine Learning

50+ ML Models in  
2 Minutes

Predicts Attrition with  
Highest Accuracy  
(AUC / LogLoss)



The screenshot shows a web browser window titled "Employee Churn Dashboard" with the URL "127.0.0.1:7755". The main content is titled "Employee Attrition Prediction" and features a sub-section titled "Employee Attrition Prevention". It displays a dropdown menu showing "Denise Mooney, 71.8%". Below this is a callout box with a green border containing two red buttons. The left button is labeled "Attrition Risk" with "71.8%" underneath, and the right button is labeled "Prediction" with "Leave" underneath. In the background, there is a blurred image of three people in professional attire.

Employee Attrition Prevention

Select an employee, and learn how best to prevent their attrition.

Denise Mooney, 71.8%

Attrition Risk  
71.8%

Prediction  
Leave

# Tech Talk



**LIME**

Explains Individual Observations

**Business Insights**  
Complex Models  
Can be Explained



# Financial Institutions: Stock Portfolio Optimization

## Risky bets lose customers

When a **hedge fund loses 5%** more than the market, the hedge fund loses customers 10% of customers.

**Managing Risk is Critical.**

We can manage risk with



# Demo #2

## Stock Portfolio Optimization App

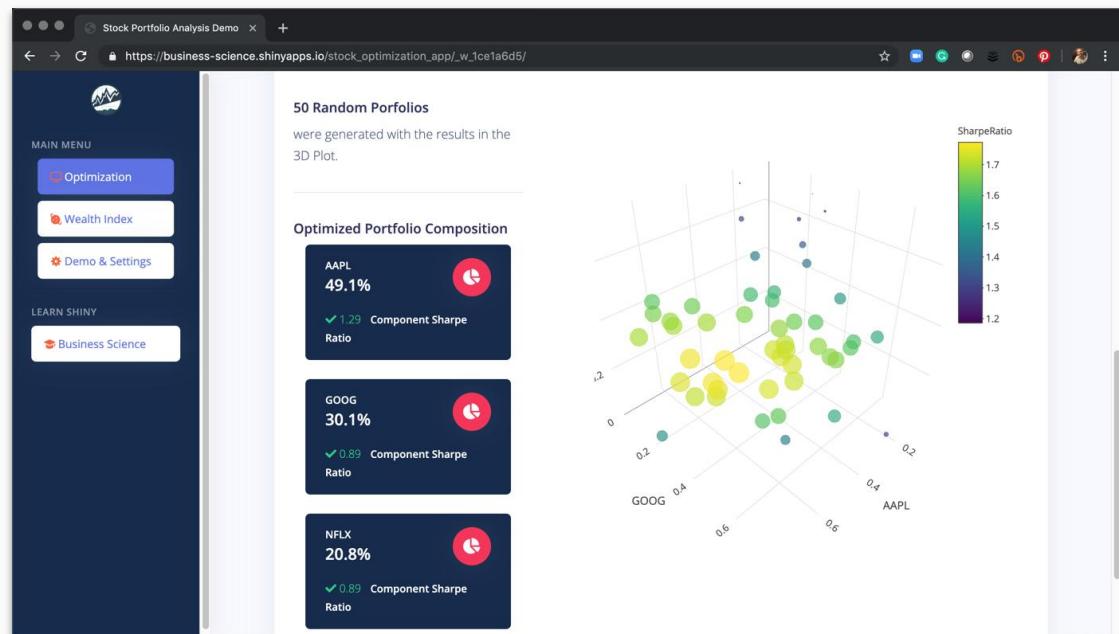


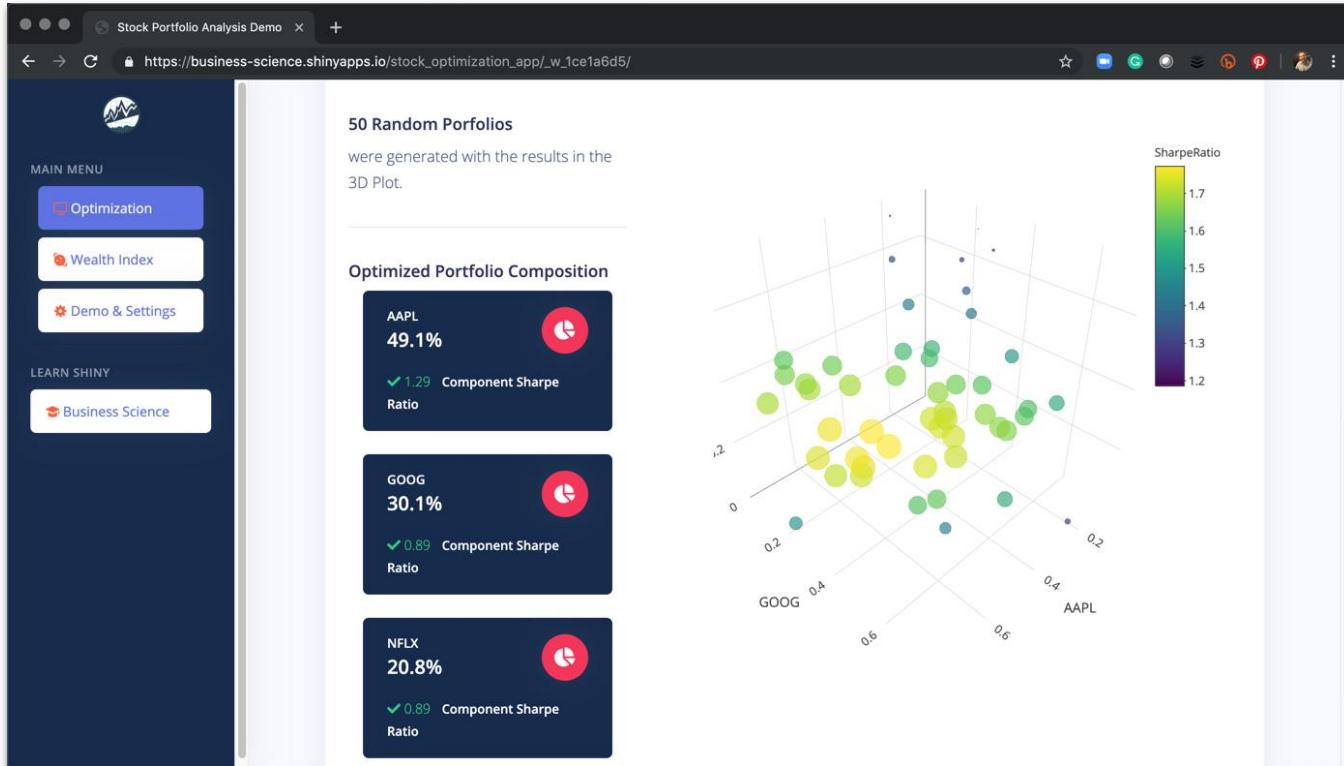
### Reduce Risk While Increasing Returns

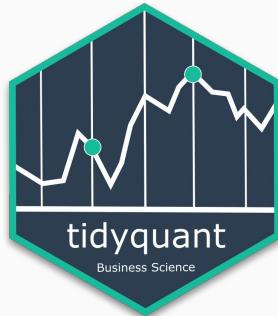
How does the app work?

Why is it useful to an organization?

What technologies make this?







## Stocks & Portfolios

50+ Random Portfolios in  
Seconds

Calculate the Optimal Mix

MAIN MENU

Optimization

Wealth Index

Demo & Settings

LEARN SHINY

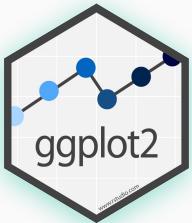
Business Science

50 Random Portfolios

were generated with the results in the 3D Plot.

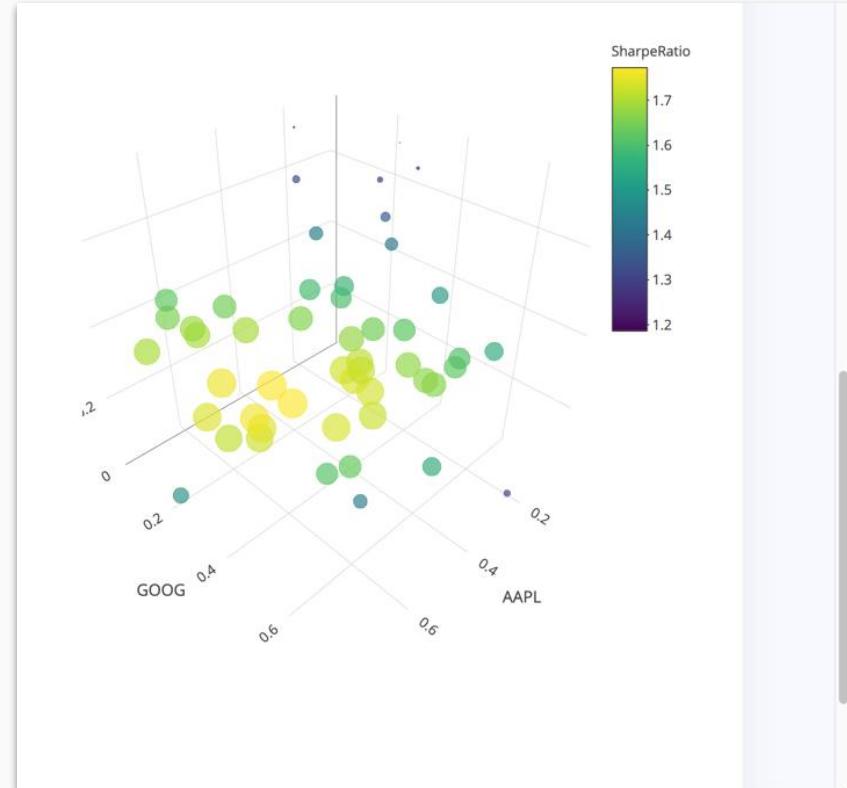
Optimized Portfolio Composition

Symbol	Weight (%)	Component Sharpe Ratio
AAPL	49.1%	✓ 1.29
GOOG	30.1%	✓ 0.89
NFLX	20.8%	✓ 0.89



## Data Visualization

tidyverse packages  
Plotly



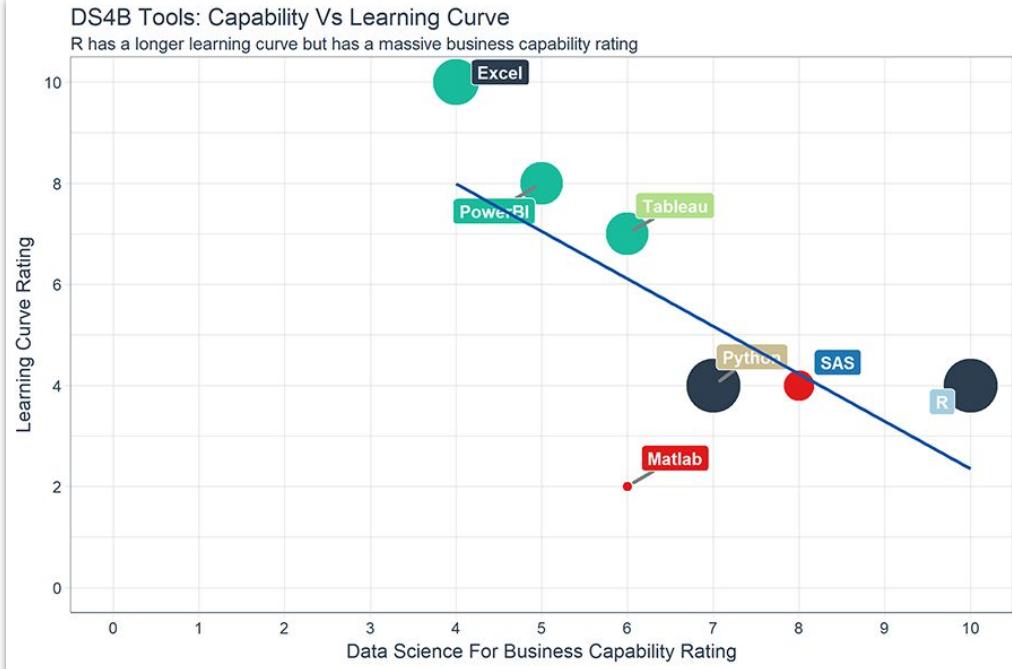
# **Reason #1**

## **Best Qualities for People Like Me**



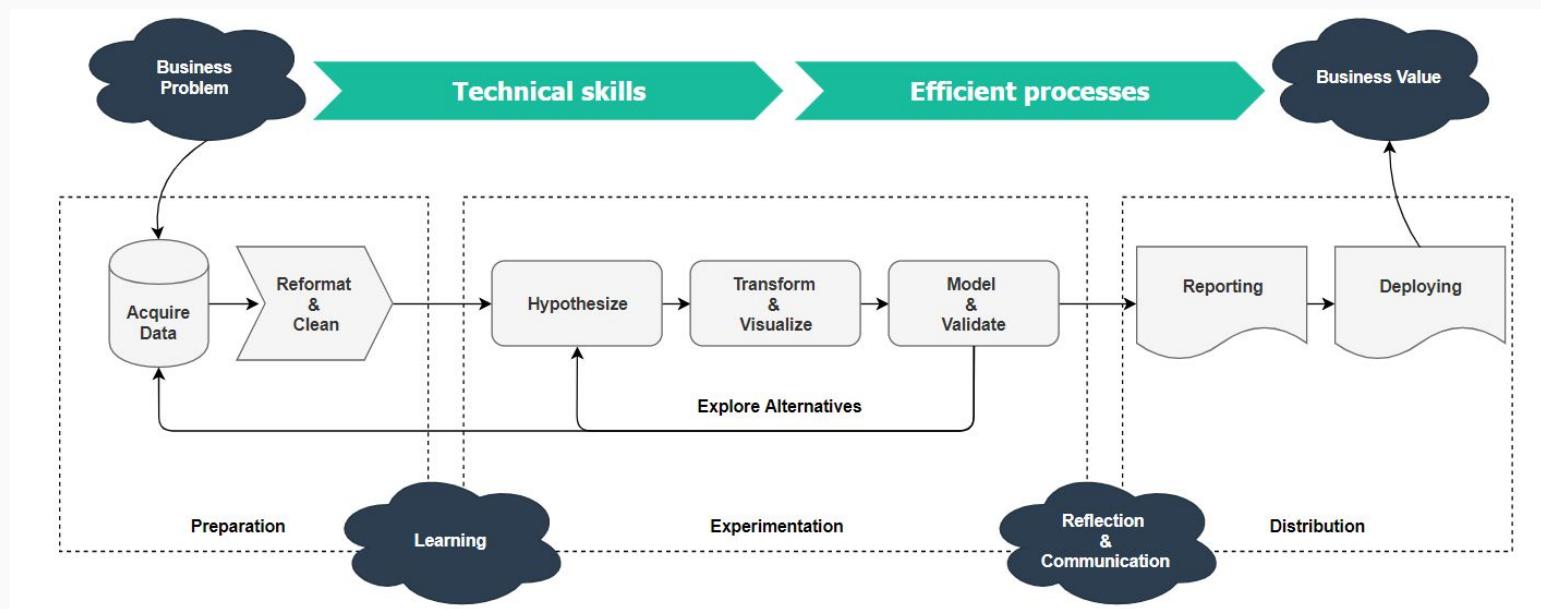
# Capability

- Reporting Tools
- Machine Learning
- Time Series
- Statistics
- Data Manipulation
- Easy to Learn

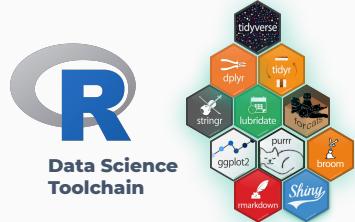


Why R? Tools like Excel, Tableau, PowerBI are easier to learn, but have lower Business Capability. Tools like Python, SAS, and Matlab have high Data Science Capability, but lack the visualization and interactive application tools needed for business. R has the best data science, visualization, and interactive tools plus it's free!

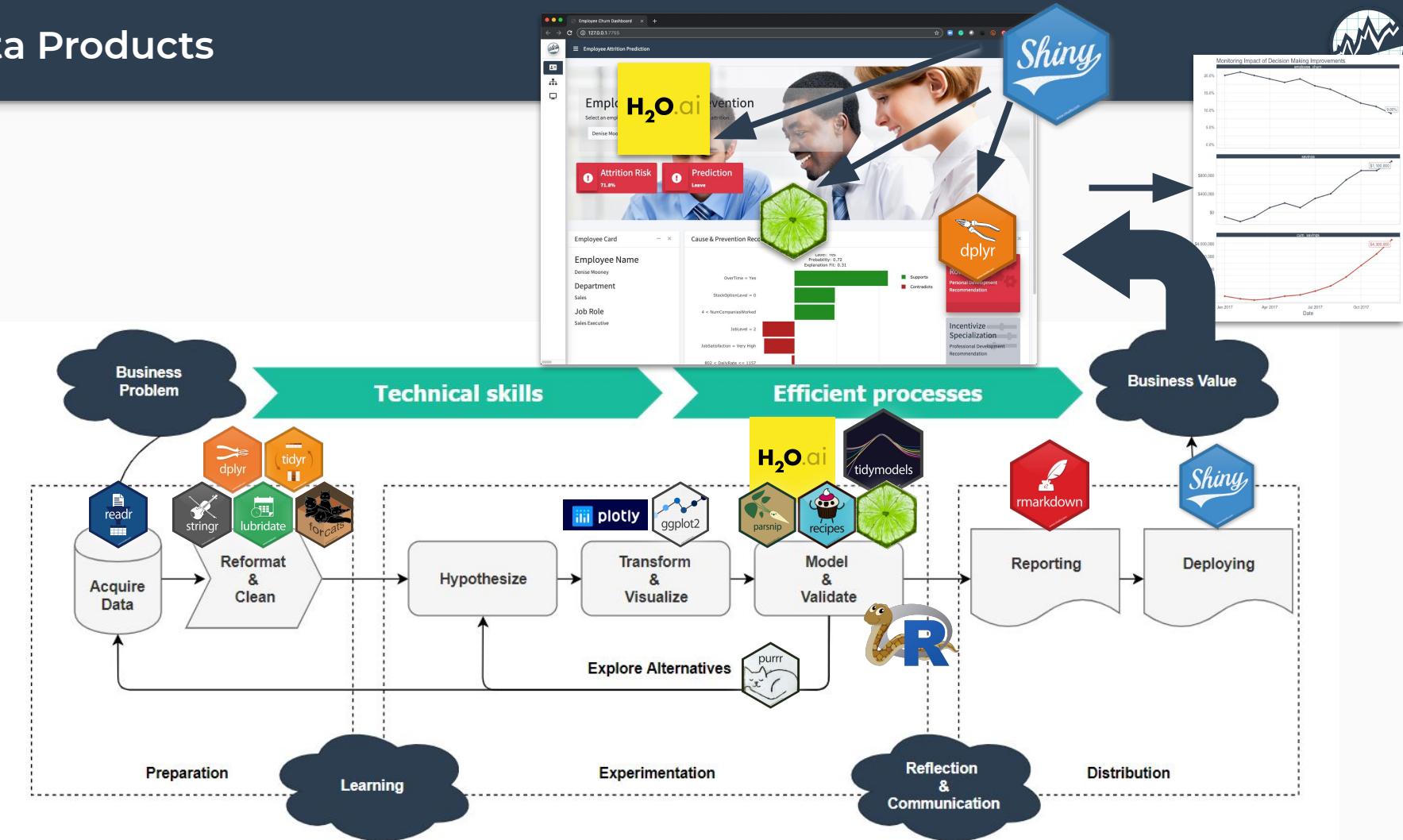
# Data Science Workflow



# Preparation



# Data Products



# Reason #2

Data Science for Non-Computer  
Scientists





# What are you trying to do?

I Need to:

- “Build A Self-Driving Car”
- “Make the Next Uber App”

I Need to:

- “Communicate results to an organization”
- “Make Web Apps that empower business decision-making”



# Reason #3

## Learning is FAST

RStudio IDE + tidyverse



# RStudio IDE



DS4B 101-R FOR BUSINESS ANALYSIS  
ITERATION WITH PURRR  
1.0 PRIMER ON PURRR  
2.0 MAPPING DATA FRAMES  
2.1 Column-wise Map  
2.2 Map Values  
2.3 Row-wise Map  
3.0 NESTED DATA  
select\_non\_na\_columns  
4.0 MODELING WITH PURRR  
4.1 Time Series Plot  
4.2 Modeling Primer  
4.3 Step 1: Function To Return Fitted Res...  
tidy\_loess  
4.4 Step 2: Test Function on Single Element  
4.5 Step 3: Map Function to All Categories

```
176 ungroup() %>%
177
178   mutate(category_2 = as_factor(category_2) %>% fct_reorder2(month_end, total_price))
179
180 rolling_avg_3_tbl %>%
181
182   ggplot(aes(month_end, total_price, color = category_2)) +
183
184   # Geometries
185   geom_point() +
186   geom_line(aes(y = rolling_avg_3), color = "blue", linetype = 1) +
187   facet_wrap(~ category_2, scales = "free_y") +
188
189   # Add Loess Smoother
190   geom_smooth(method = "loess", se = FALSE, span = 0.2, color = "black") +
191
192   # Formatting
193   theme_tq() +
194   scale_color_tq() +
195   scale_y_continuous(labels = scales::dollar_format(scale = 1e-3, suffix = "K"))
196
197
198
199 # 4.2 Modeling Primer ----
200
201 # Data Preparation
202
203 sales_by_m_cross_country_tbl <- rolling_avg_3_tbl %>%
204
205 4.2 Modeling Primer :
```

Console Terminal Jobs ~/Desktop/DS4B\_101\_R\_FINAL/ ↗



# What are you trying to do?



tidyverse

- Making Data Manipulations are like Reading A Book (**verbs**)
- Fast to go from thought to results
- **Visualization** capabilities unmatched

```
> bike_orderlines_tbl %>%
+   select(order_date, total_price) %>%
+
+   # lubridate
+   mutate(order_date = ymd(order_date)) %>%
+   mutate(year_month = floor_date(order_date, unit = "month")) %>%
+
+   # group_by + summarize
+   group_by(year_month) %>%
+   summarize(sales = sum(total_price))
# A tibble: 60 x 2
  year_month     sales
  <date>     <dbl>
1 2011-01-01  483015
2 2011-02-01  1162075
3 2011-03-01   659975
4 2011-04-01  1827140
5 2011-05-01   844170
6 2011-06-01  1413445
7 2011-07-01  1194430
8 2011-08-01   679790
9 2011-09-01   814720
10 2011-10-01   734920
# ... with 50 more rows
```



# Reason #4

## Powerful Ecosystem

Machine Learning & Big Data



# Data Science Toolchain



Data Science Toolchain



Amazing Toolchain for End-to-End Data Science



# Resource #1: Ultimate R Cheat Sheet

**Page 1: Core**

Click the links for Documentation

**Data Science with R Workflow**

If you want to learn R and this workflow for business analysis, take the [R for Business Analysis \(DS4B 101-R\)](#) course through Business Science University.

**Important Resources**

- R For Data Science Book: <http://r4ds.had.co.nz/>
- cheatsheets: <https://www.rstudio.com/resources/cheatsheets/>
- Data Visualization Book: <https://corybrunson.com/>
- More Cheatsheets: <https://www.rstudio.com/resources/cheatsheets/>
- Connecting to databases: <https://db.rstudio.com/>
- Shiny web applications website: <http://shiny.rstudio.com/>
- Jenny Bryan's purrr tutorial: <https://jennybryan.org/purrr.html>

**Page 2: Shiny**

**Page 3: ML, Time Series, Big Data**

**Data Science with R Special Topics**

**Time Series Analysis**

- Time-aware thinks: `tibbles & table`
- Convert between classes: `ltsm & tsbox`
- Time Series Index Summary: `ltsm`
- Generating Future Series: `ltsm`

**Forecasting**

- ARIMA, ETS, etc. Forecast & table
- Tidy, glance, augment for forecast models: `forecast`
- Converting forecast prediction to table: `forecast`

**Network Analysis**

- Static: `igraph`
- Interactive (JavaScript): `networkD3`, `D3 Networks` in R, `networkDS` (`js` network graphs) in R

**Network Viz**

- Static: `igraph`, `ggplot2`
- Interactive (JavaScript): `networkD3`

**Geospatial Analysis**

- Static: `sf` (`sp`), `sfheaders`
- Geocoding (getting latlong, bboxes, s, sf): `geocode`
- Google API (requires key): `google`
- OpenStreetOverpass API: `osmdata`
- Simple Features (sf objects): `sf` (`sf`)
- Spatial Objects (sp objects): `sp` (`monoly`)

**Geospatial Viz**

- Static: `ggmap`
- Interactive (JavaScript): `pkdx geom` (`sf`), `mapview` (`sf`)

**Speed & Scale**

- Faster than `sp` & `grid`: `data.table` (`CS`)
- Distributed Cluster (Spark): `sparklyr` (`CS`)
- Parallel Processing: `furrr`

**Machine Learning**

- Multi-Threaded Scalable/Production ML:
  - H2O (`CS`)
    - Extreme Gradient Boosting: `xgboost`
    - R + Spark: `sparklyr` (`CS`)
    - PyTorch Using Water (Spark + H2O): `sparkling`
  - ML (Tidy): `mlr` (`CS`)

**Deep Learning**

- R Interface to TensorFlow Homegrown:
  - Keras (`CS`)
    - TensorFlow
    - TensorFlow (Core)

**Interoperability**

- Python: `rpyc` (`CS`)
- C++: `RCpp`

**Miscellaneous Tools**

- Interactive Plotting: `htmlwidgets` for R
  - Building Interactive Dashboards: `dashr`
    - Pkg Development Tools: `devtools` (`CS`)
    - Code Coverage: `lcov`
    - Advanced Concepts (Advanced R Book)
      - Advanced R Book (`CS`)
  - Making Blogs & Books:
    - Blogdown (`CS`)
    - knitr (`CS`)
    - Plotly (`CS`)
  - Posting Code (GitHub, Stack Overflow): `ace2`

**Business Science University** [university.business-science.io](http://university.business-science.io)



# Resource #2: data.table Cheat Sheet

Page 3

2

[TensorFlow \(Core\)](#)

## Speed & Scale

- Faster than dplyr & pandas: [data.table \(CS\)](#)
- Distributed Cluster (Spark): [sparklyr \(CS\)](#)
- Parallel Processing: [furrr](#)

## Interoperability

- Python: [reticulate \(CS\)](#)
- Java: [rJava](#)
- C++: [Rcpp](#)

## Miscellaneous Tools

- Interactive Plotting: [htmlwidgets for R](#)
- Building R Packages: [R packages Book](#)
  - Pkg Development Tools: [devtools \(CS\)](#)
  - R Templates: [useRishtis](#)
  - Build Web Doc's: [pkdown](#)
- Advanced Concepts ([Advanced R Book](#))
  - [rlang & Tidy Evaluation \(CS\)](#)
- Making Blogs & Books:
  - [blogdown, bookdown](#)
- Posting Code (GitHub, Stack Overflow): [reprex](#)

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

## Data Transformation with data.table :: CHEAT SHEET

### Basics

data.table is an extremely fast and memory efficient package for transforming data in R. It works by converting R's native data frame objects into data.tables with new and enhanced functionality. The basics of working with data.tables are:

**dt[i, j, by]**

Take data.table **dt**,  
subset rows using **i**  
and manipulate columns with **j**,  
grouped according to **by**.

data.tables are also data frames - functions that work with data frames therefore also work with data.tables.

### Create a data.table

**data.table(a = c(1, 2), b = c("a", "b"))** - create a data.table from scratch. Analogous to data.frame().

**setDT(df)\*** or **as.data.table(df)** - convert a data frame or a list to a data.table.

### Subset rows using i

 **dt[1:2,]** - subset rows based on row numbers.

 **dt[a > 5,]** - subset rows based on values in one or more columns.

### LOGICAL OPERATORS TO USE IN i

<	=	is.na()	%in%		%like%
>	==	is.na() !		&	%between%

### Manipulate columns with j

#### EXTRACT

 **dt[, c(2)]** - extract columns by number. Prefix column numbers with "-" to drop.

 **dt[, .(b, c)]** - extract columns by name.

#### SUMMARIZE

 **dt[, .(x = sum(a))]** - create a data.table with new columns based on the summarized values of rows.

Summary functions like mean(), median(), min(), max(), etc. can be used to summarize rows.

#### COMPUTE COLUMNS\*

 **dt[, c := 1 + 2]** - compute a column based on an expression.

 **dt[, c == 1, c := 1 + 2]** - compute a column based on an expression but only for a subset of rows.

 **dt[, . := (c = 1, d = 2)]** - compute multiple columns based on separate expressions.

#### DELETE COLUMN

 **dt[, c := NULL]** - delete a column.

#### CONVERT COLUMN TYPE

 **dt[, b := as.integer(b)]** - convert the type of a column using as.integer(), as.numeric(), as.character(), as.Date(), etc.

### Group according to by

#### dt[, j, by = .(a)]

group rows by values in specified columns.

#### dt[, j, keyby = .(a)]

group and simultaneously sort rows by values in specified columns.

#### COMMON GROUPED OPERATIONS

**dt[, .c := sum(b), by = a]** - summarize rows within groups.

**dt[, .SD[1], by = a]** - extract first row of groups.

**dt[, .SD[N], by = a]** - extract last row of groups.

### Chaining

**dt[, ...][...]** - perform a sequence of data.table operations by chaining multiple "[ ]".

### Functions for data.tables

#### REORDER

 **setorder(dt, a, -b)** - reorder a data.table according to specified columns. Prefix column names with "-" for descending order.

#### \* SET FUNCTIONS AND %:=

data.table's functions prefixed with "set" and the operator "%=:" work without "<- to alter data without making copies in memory. E.g., the more efficient "setDT(df)" is analogous to "df <- as.data.table(df)".



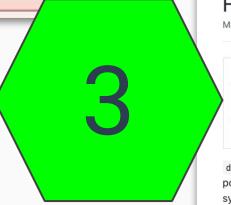
# Resource #3: data.table GitHub Wiki



## Page 3

### Speed & Scale

- Faster than dplyr & pandas: [data.table \(CS\)](#)
- Distributed Cluster (Spark): [sparklyr \(CS\)](#)
- Parallel Processing: [furr](#)



 / data.table

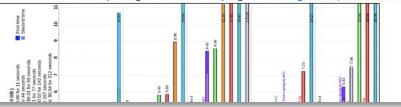




Latest news from dev: [NEWS](#)

`data.table` is one of the 13,000 add-on packages for the programming language R which is popular in [these fields](#). It provides a high-performance version of base R's `data.frame` with syntax and feature enhancements for ease of use, convenience and programming speed. As of Nov 2018, `data.table` was the 4th largest Stack Overflow tag about an R package with over 8,000 questions, the 10th most starred R package on GitHub and had over 650 CRAN and Bioconductor packages using it. Here are a further 6,000 accepted answers which use/mention `data.table` but where the question was not specifically about `data.table`.

We have updated the 2014 grouping benchmarks comparing `data.table` to pandas and dplyr, and included Spark and pydata.table. The benchmark is automated and runs regularly against the latest versions of these packages. It is a work in progress: [h2oai.github.io/db-benchmark](#)



<https://www.business-science.io/r-cheatsheet.html>

```
> require(data.table)
> example(data.table)

# basic row subset
DT[2]
DT[2:3]
w=2:3; DT[w]
DT[order(x)]
DT[order(x), ]
DT[y>2]
DT[y>2 & v>5]
DT[1:4]
DT[-(2:4)]

# select|compute columns
DT[, v]
DT[, list(v)]
DT[, .(v)]
DT[, sum(v)]
DT[, c(sum(v))]
DT[, .(sv=sum(v))]
DT[, .(v, v*2)]

# v column (as vector)
# v column (as data.table)
# same; .() is an alias for list()
# sum of column v, returned as vector
# same but return data.table
# same but name column "sv"
# return two column data.table

# subset rows and select|compute
DT[2:3, sum(v)]
DT[2:3, .(sum(v))]
DT[2:3, .(sv=sum(v))]
DT[2:5, cat(v, "\n")]

# select columns the data.frame way
DT[, 2]
colNum = 2
DT[, ...colNum]
DT[["v"]]

# grouping operations - j and by
DT[, sum(v), by=x]
DT[, sum(v), keyby=x]
DT[, sum(v), by=x][order(x)]

# 2nd column, a data.table always
# same, but return data.table
# same, but name column "sv"
# just for j's side effect

# appearance order of groups preserved
# order the result by group
# same by chaining expressions together
```

# Reason #5

## Designed for Data Products

Reporting & Web Applications



# Data Products



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

The data science team has identified 4 customer segments. The 4 customer segments were given descriptions based on the customer's top product purchases.

1. Segment 1 Preferences: Road Bikes, Below \$3200 (Economical Models)
2. Segment 2 Preferences: Mountain Bikes, Above \$3200 (Premium Models)
3. Segment 3 Preferences: Road Bikes, Above \$3200 (Premium Models)
4. Segment 4 Preferences: Both Road and Mountain, Below \$3200 (Economical Models)

## Customer Preferences

### Heat Map

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

		Mountain					Road			
		16.8%	3.1%	4.5%	12.2%	13.3%	2.4%	24.1%	18.9%	4.5%
Albuquerque Cycles	5.3%	0.7%	3.2%	7.6%	7.1%	6.6%	37.0%	23.6%	8.8%	
Ann Arbor Speed	8.5%	0.8%	2.0%	11.8%	6.5%	3.7%	37.4%	19.9%	9.3%	
Austin Cruisers										



## Business Reporting

PDF & HTML Reports in Seconds

Way better than Jupyter Notebooks



# Data Products

## 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

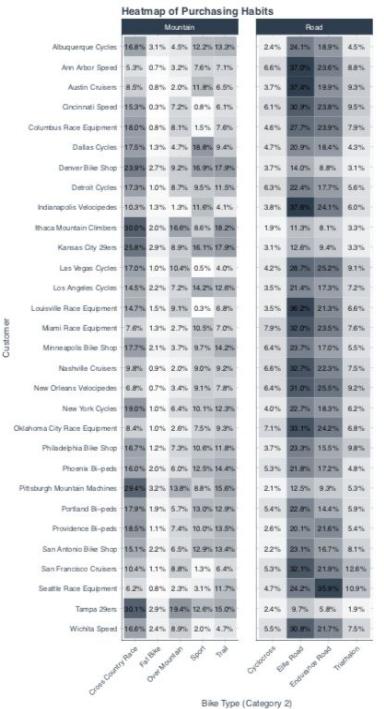
The data science team has identified 4 customer segments. The 4 customer segments were given descriptions based on the customer's top product purchases.

- Segment 1 Preferences:** Road Bikes, Below \$3200 (Economical Models)
- Segment 2 Preferences:** Mountain Bikes, Above \$3200 (Premium Models)
- Segment 3 Preferences:** Road Bikes, Above \$3200 (Premium Models)
- Segment 4 Preferences:** Both Road and Mountain, Below \$3200 (Economical Models)

### Customer Preferences

#### Heat Map

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



1

2



## Business Reporting

PDF & HTML Reports in Seconds

Way better than Jupyter Notebooks

# Reason #6

## Strong Community Support





# Community

## Social/Web



R-Ladies DC  
@RLadiesDC

For your #rstats holiday wish list consideration 🎁🍪

[etsy.com/listing/562883...](https://etsy.com/listing/562883...)

♡ 119 1:21 AM - Dec 2, 2017

21 people are talking about this >

Twitter: #Rstats

The screenshot shows a list of R conferences and meetings. The page title is "Events | A list of R conferences". The main content area is titled "Events" and contains sections for "2020", "2019", "2018", "2017", and "2016". Each section lists specific events with their dates, locations, and Twitter handles.

Year	Event	Date	Location	Twitter Handle
2020	useR! 2020	July 7-10	St Louis, USA	@useR2020stl
2019	rstatsconf	January 15-18	Austin, USA	
2019	ConnectR 2019	January 24-26	San Jose, Costa Rica	
2019	satRdays Paris	February 23	Paris, France	
2019	Chicago R Unconference	March 9-10	Chicago IL, USA	
2019	satRday Los Angeles	April 6	Los Angeles CA, USA	
2019	satRday Johannesburg	April 6	Johannesburg, South Africa	
2019	satRdays Newcastle	April 6	Newcastle, UK	

Conferences & Meetups



# Community



**Business Science**

*R-Track*

400+ Students, Consultants,  
Business Analysts

Data Scientists with odd  
paths just like me!

The screenshot shows a messaging interface with a dark theme. On the left is a sidebar with a list of topics and a direct messages section. The main area shows a conversation between two users.

**#ask\_the\_professor**

**Thread** #ask\_the\_professor

**Matt Dancho** 23 hours ago It's just what level of responsibility to take on.

**Jochem Donkers** 11 hours ago Building a business case with the right argument is key. This should not be a big problem if the organization can afford to pay you to spend time on building models. The last 20k for licensing rstudio products is not going to kill the case.

Once you have the case ensure buy-in from key (business leadership) stakeholders in the business. They have the influence to shift the focus of any IT department.

Good luck!

**Matt Dancho** < 1 minute ago Words of Wisdom! @Jochem Donkers

**Message #ask\_the\_professor**

**Also send to #ask\_the\_professor**

# Learning Plan

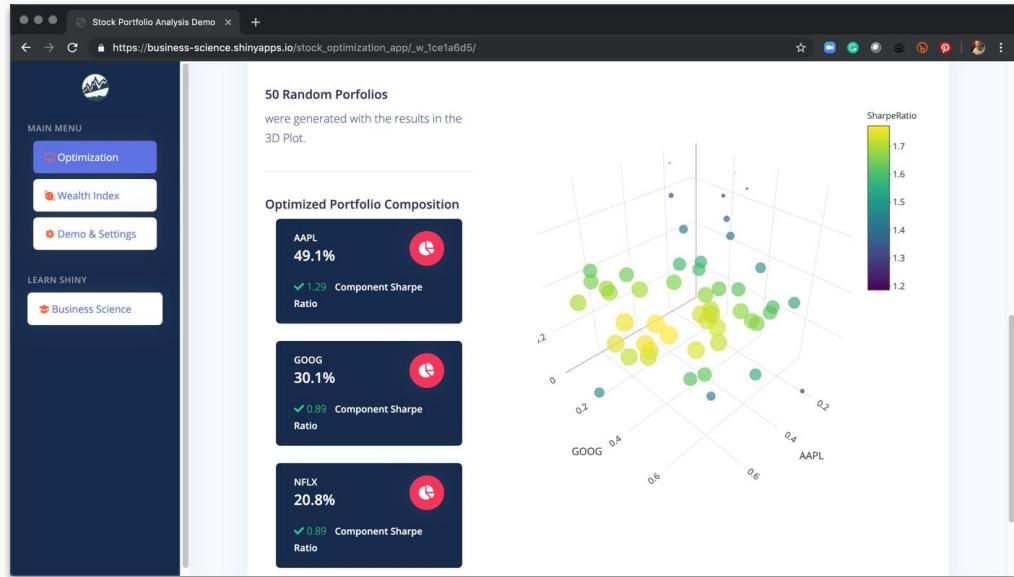
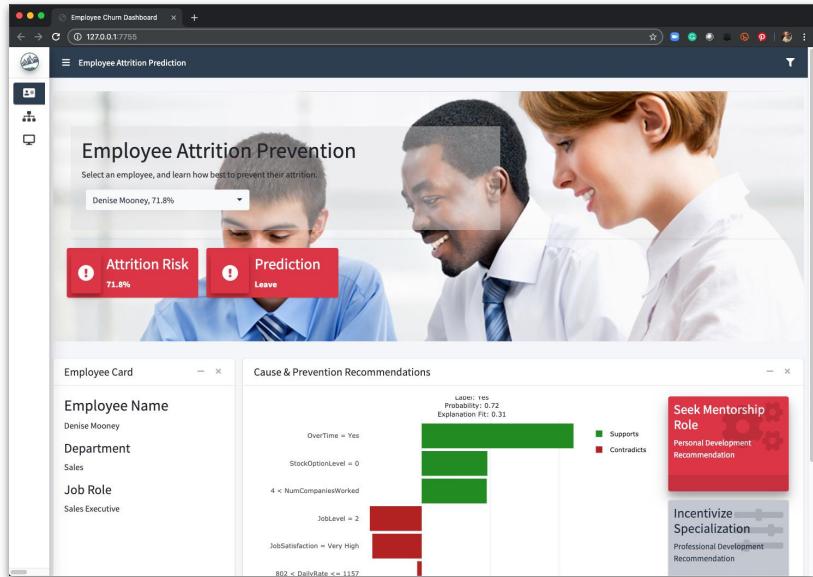
I will teach you

Fast, Business-First, Effective



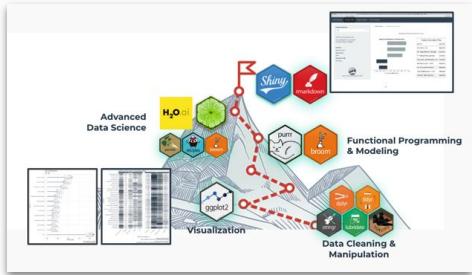


# How do you do this?

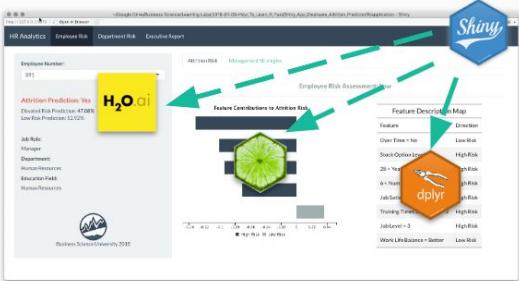


Coming soon to Business Science University

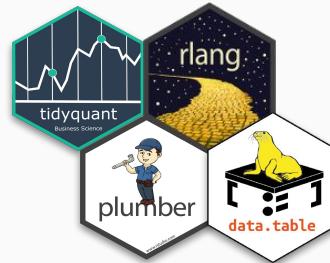
# YOUR Transformation



**Do Business Projects**  
Climb the Hill



**Build Production-Ready**  
Web Apps

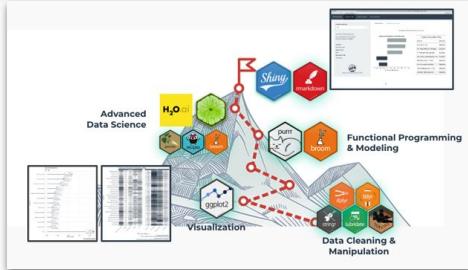


**Continuously Learn**  
Big Data, API's, Finance

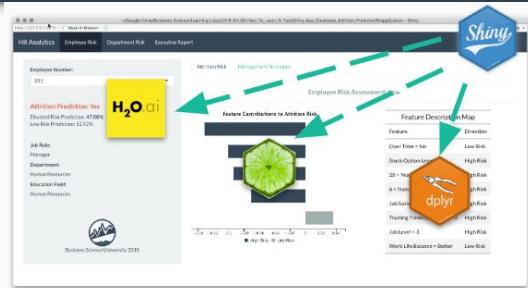
**Start**



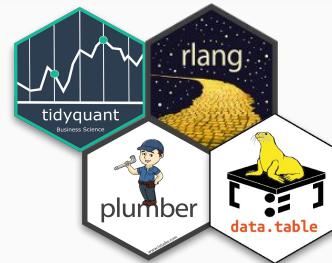
# YOUR Transformation



**Do Business Projects**  
Climb the Hill



**Build Production-Ready**  
Web Apps



**Continuously Learn**  
Big Data, API's, Finance

**Start**

**1**



Analysis Courses

**2**



App Development  
Courses

**Finish**

**3**



Learning Labs PRO

Everything is **Taken Care of** For You in Our Platform



# 3-Course R-Track System



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

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

## R Shiny Web Apps For Business (DS4B 102-R)

### Project-Based Courses with Business Application

Data Science Foundations  
**7 Weeks**



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

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



Machine Learning & Business Consulting  
**10 Weeks**



#### 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



Web Application Development  
**4 Weeks**

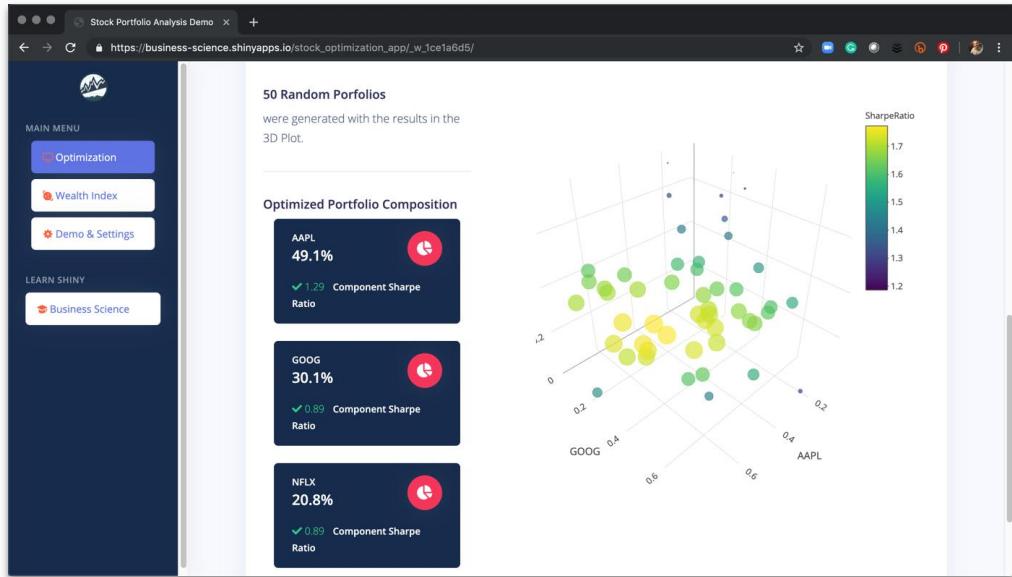
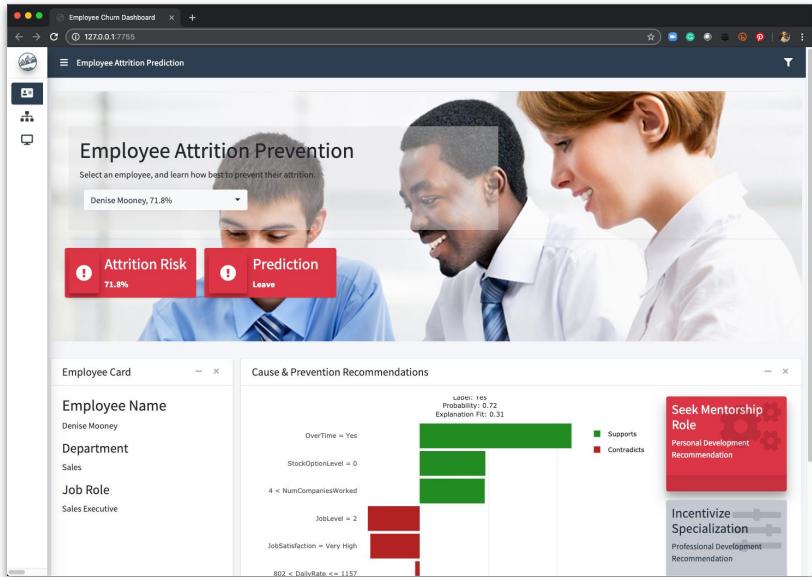


#### DS4B 102-R: Shiny Web Applications For Business (Level 1)

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



# Coming soon

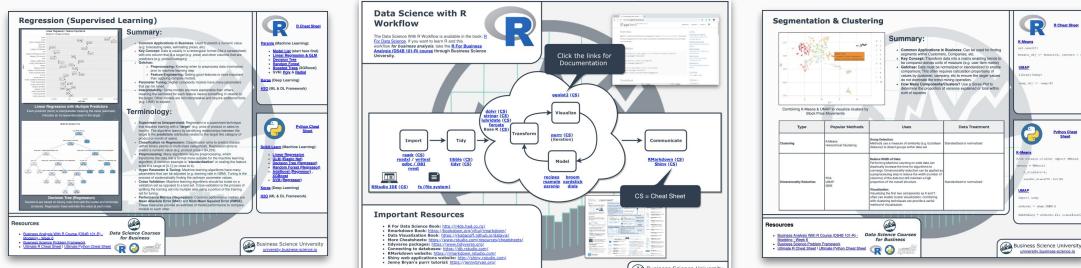


# Key Benefits

- 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

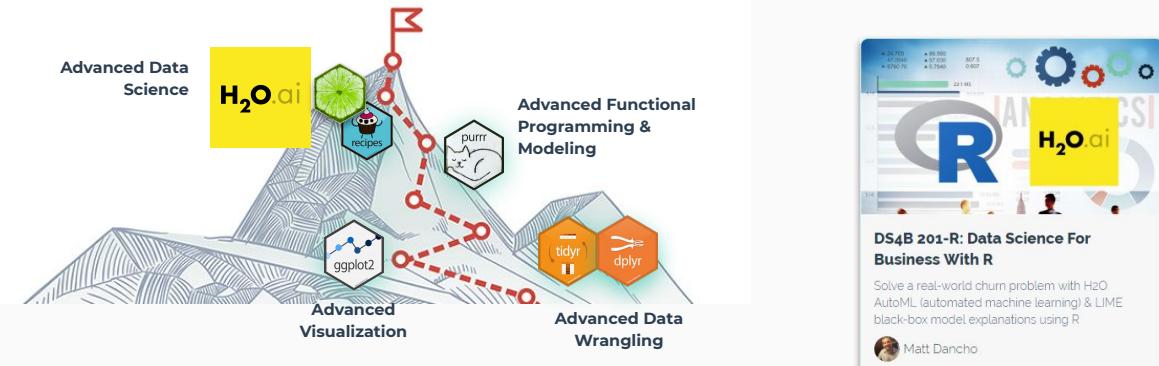
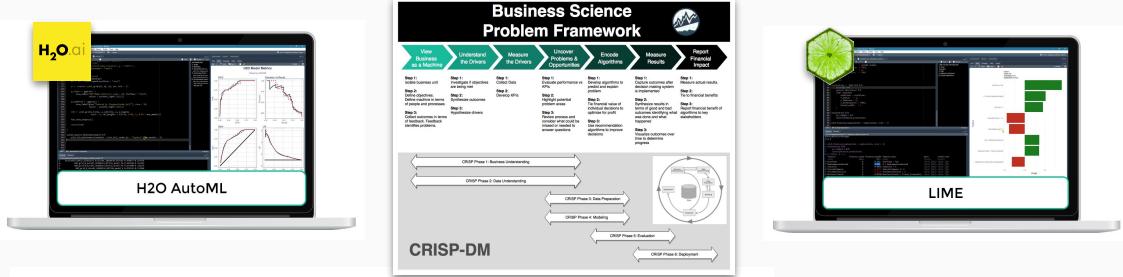
Return-On-Investment - Weeks 7, 8, 9

- 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 with data points.
- A "DATA SCIENCE" section showing a dashboard with a pie chart and a bar chart.
- A "ANALYTICS" section showing a dashboard with a scatter plot and a bar chart.
- A "Machine Learning" section showing a dashboard with a scatter plot and a sidebar.
- A "Business Analytics" section showing a dashboard with a map and a bar chart.
- A "Data Science with R" course page featuring a "Sales Dashboard with Demand Forecasting" dashboard.

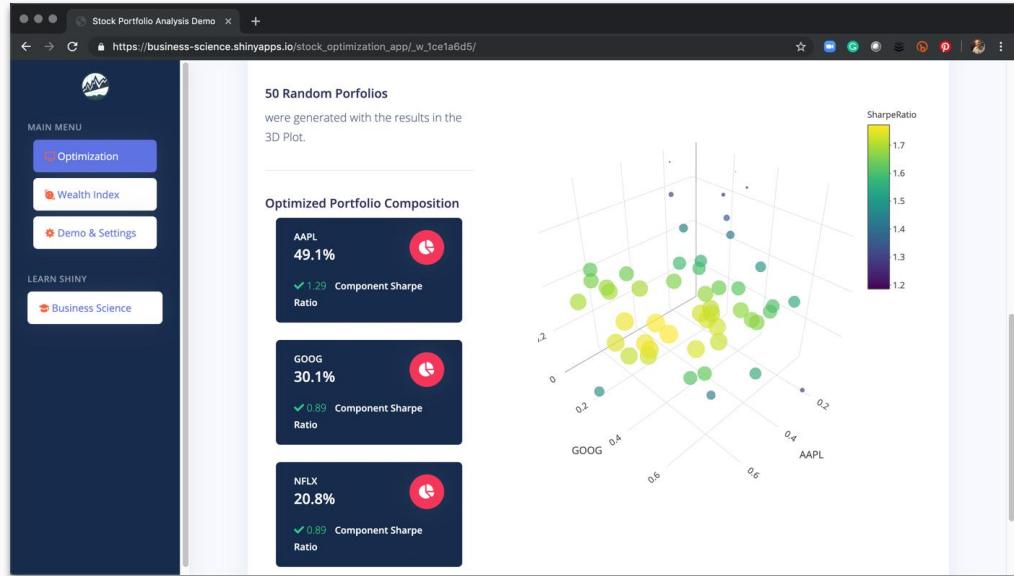
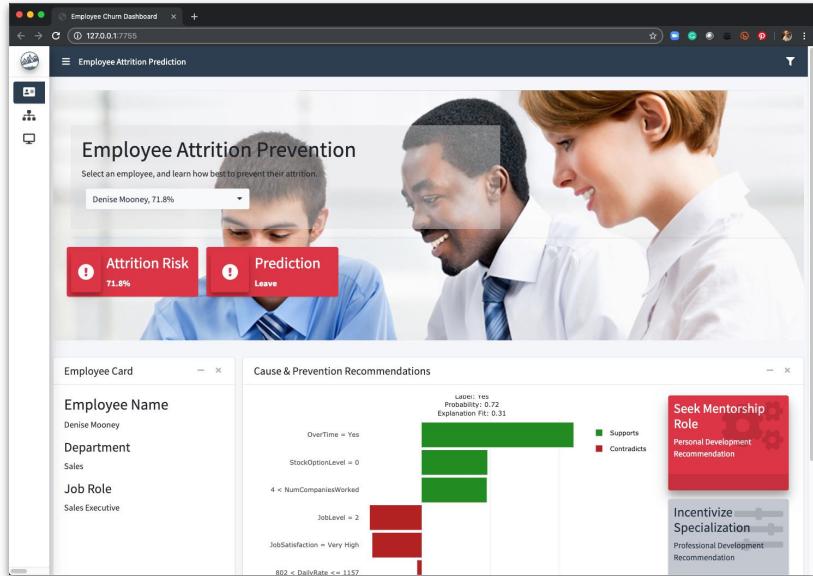
**DS4B 102-R: Shiny Web Applications for Business (Level 1)**

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

Matt Dancho



# Build these data products



Everything is **Taken Care of** For You in Our Platform



## Testimonials



*“Your program allowed me to cut down to **50% of the time** to deliver solutions to my clients.”*

*-Rodrigo Prado, Managing Partner Big Data Analytics & Strategy at Genesis Partners*



*“I can already **apply** a lot of the early gains from the course to current working projects.”*

*-Adam Mitchell, Data Analyst with Eurostar*



*“My work became **10X easier**. I can spend quality time asking questions rather than wasting time trying to figure out syntax.”*

*-Mohana Chittor, Data Scientist with Kabbage, Inc*

Achieve  
**Results** that  
Matter to  
the  
**Business**



# PROMO Code: learninglabs

**H<sub>2</sub>O.ai**

**tidyverse**

**Shiny**

**Bundle - DS For Business + Web Apps (Level 1): R-Track - Courses 101, 102,**

3 Course Bundle

0% COMPLETE

**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

## R-TRACK BUNDLE

MSRP: \$234/mo

**6 Low Monthly Payments**

**\$199/mo**

Save: \$35/mo

# Begin Learning Today

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

