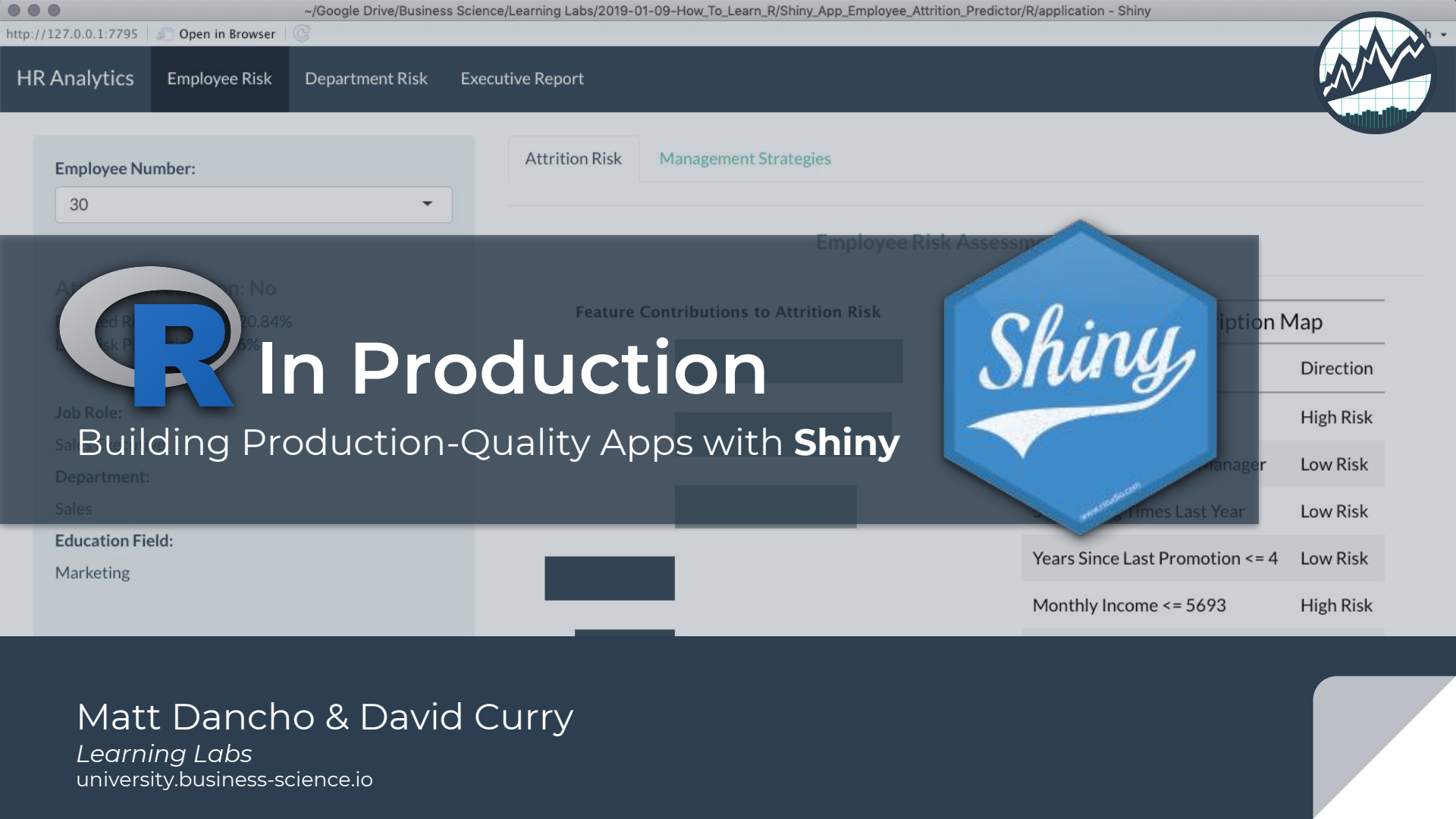


In Production

Building Production-Quality Apps with **Shiny**

Matt Dancho & David Curry
Learning Labs
university.business-science.io

Learning Lab Structure

- **Presentation** (30 Minutes)
- **Q&A** (15 Minutes)

university.business-science.io

Your Hosts!

Matt Dancho

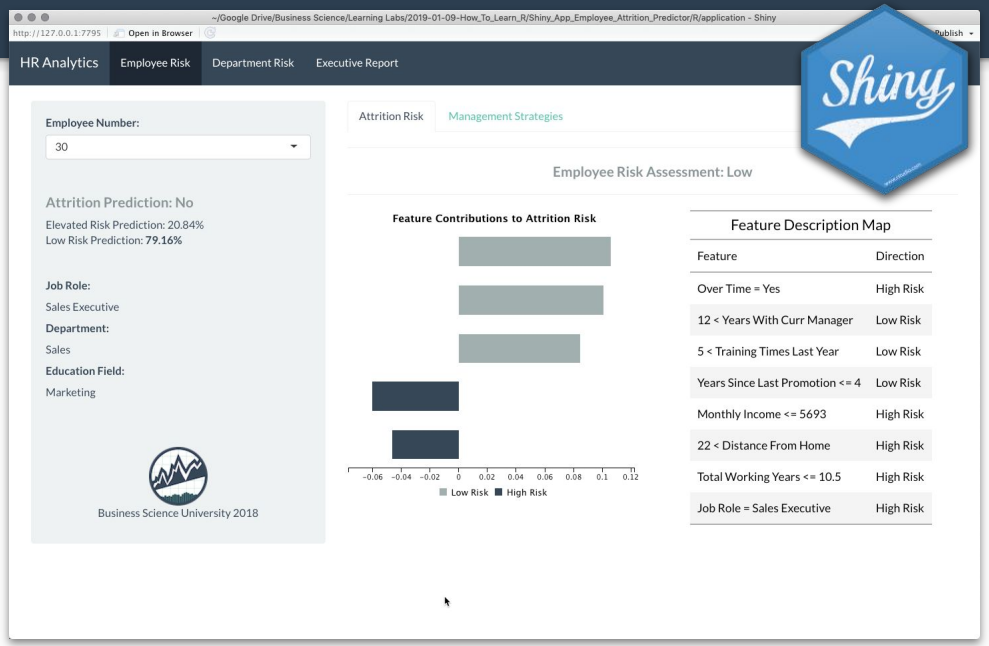
Founder of Business Science, Matt designs and executes educational courses and workshops that deliver immediate value to organizations. His passion is **up-leveling future data scientists** coming from **untraditional backgrounds**.

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: R In Production

- R in Production
  - Demo
  - New Tools
- Why Production?
- Workflow for Business Value
- Accelerated Learning Plan



in Production

Can R Be Used In Production?

What is Production?

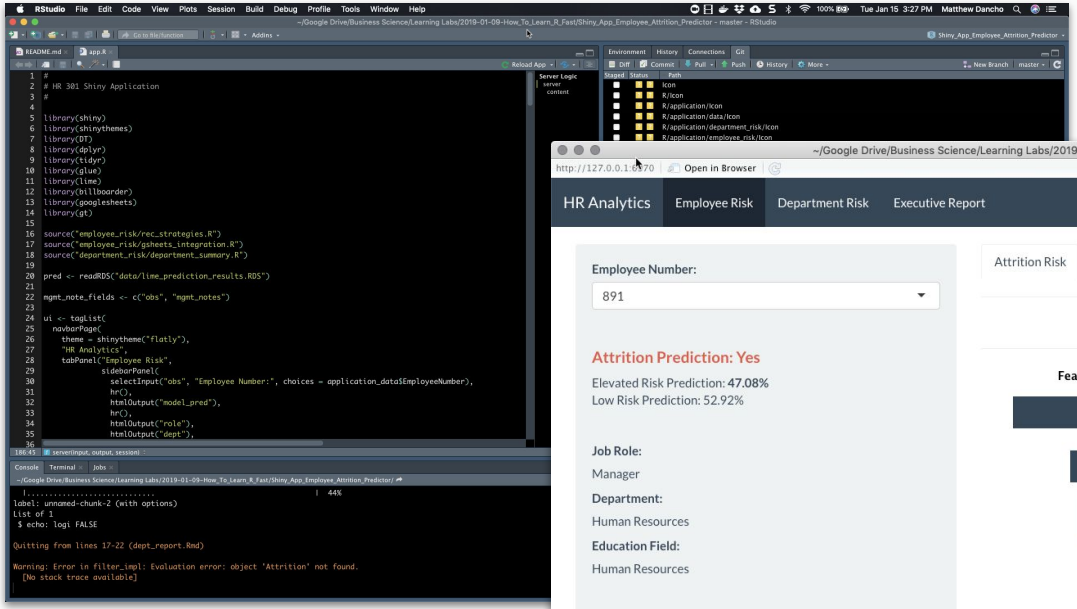
Putting a data science product **in the hands of decision makers**

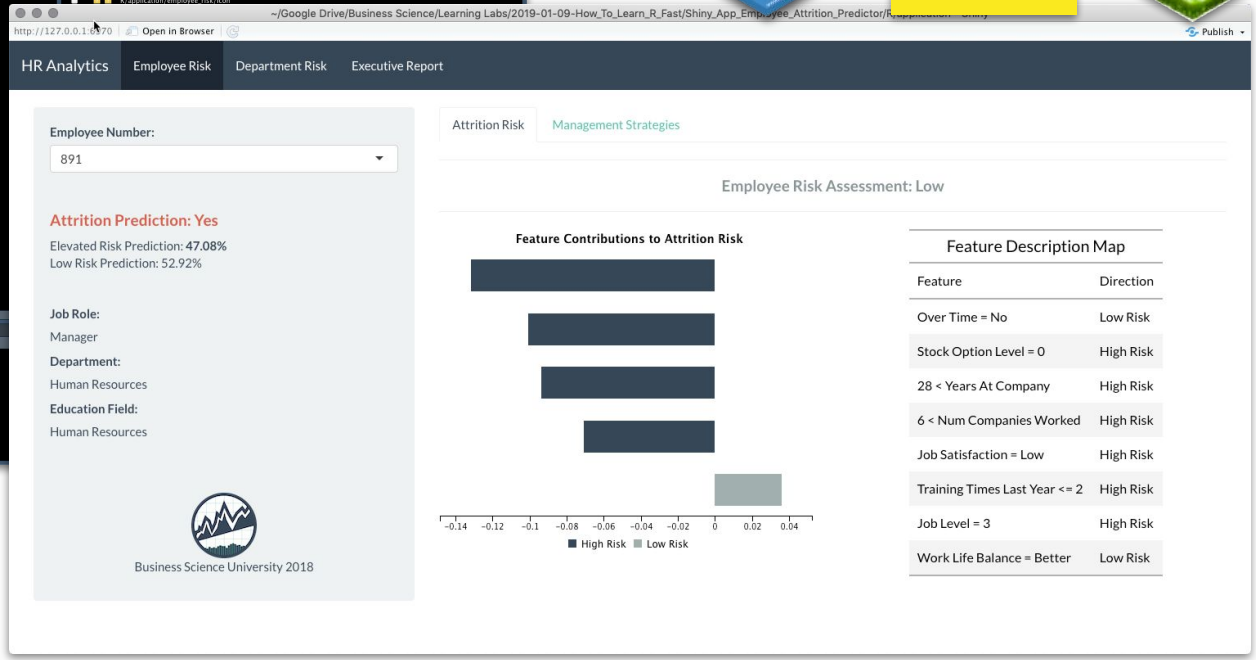
What is Shiny?

A web framework for building data science products that **go into the hands of decision makers**

Makes going from **concept to production fast**

Demo Time!


RStudio File Edit Code View Plots Session Build Debug Profile Tools Window Help
README.md
app.R
Server Logic
server
content
# HR 301 Shiny Application
library(shiny)
library(shinythemes)
library(DT)
library(dplyr)
library(tidyr)
library(glue)
library(lime)
library(billboarder)
library(googlesheets)
library(gt)
source("employee_risk/rec_strategies.R")
source("employee_risk/gsheets_integration.R")
source("department_risk/department_summary.R")
pred <- readRDS("data/lime_prediction_results.RDS")
mgmt_note_fields <- c("obs", "mgmt_notes")
ui <- tagList(
navbarPage(
theme = shinytheme("flatly"),
"HR Analytics",
tabPanel("Employee Risk",
sidebarPanel(
selectInput("obs", "Employee Number:", choices = application_data$EmployeeNumber),
hr(),
htmlOutput("model_pred"),
hr(),
htmlOutput("role"),
htmlOutput("dept"),
Console
Terminal
Jobs
label: unnamed-chunk-2 (with options)
List of 1
$ echo: logi FALSE
Quitting from lines 17-22 (dept_report.Rmd)
Warning: Error in filter_impl: Evaluation error: object 'Attrition' not found.
[No stack trace available]


~/Google Drive/Business Science/Learning Labs/2019-01-09-How_To_Learn_R_Fast/Shiny_App_Employee_Attrition_Predictor/R_application/Shiny
http://127.0.0.1:6570
Open in Browser
Publish
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
Business Science University 2018
Attrition Risk
Management Strategies
Employee Risk Assessment: Low
Feature Contributions to Attrition Risk
−0.14 −0.12 −0.1 −0.08 −0.06 −0.04 −0.02 0 0.02 0.04
High Risk
Low Risk
Feature Description Map
Feature
Direction
Over Time = No
Low Risk
Stock Option Level = 0
High Risk
28 < Years At Company
High Risk
6 < Num Companies Worked
High Risk
Job Satisfaction = Low
High Risk
Training Times Last Year <= 2
High Risk
Job Level = 3
High Risk
Work Life Balance = Better
Low Risk

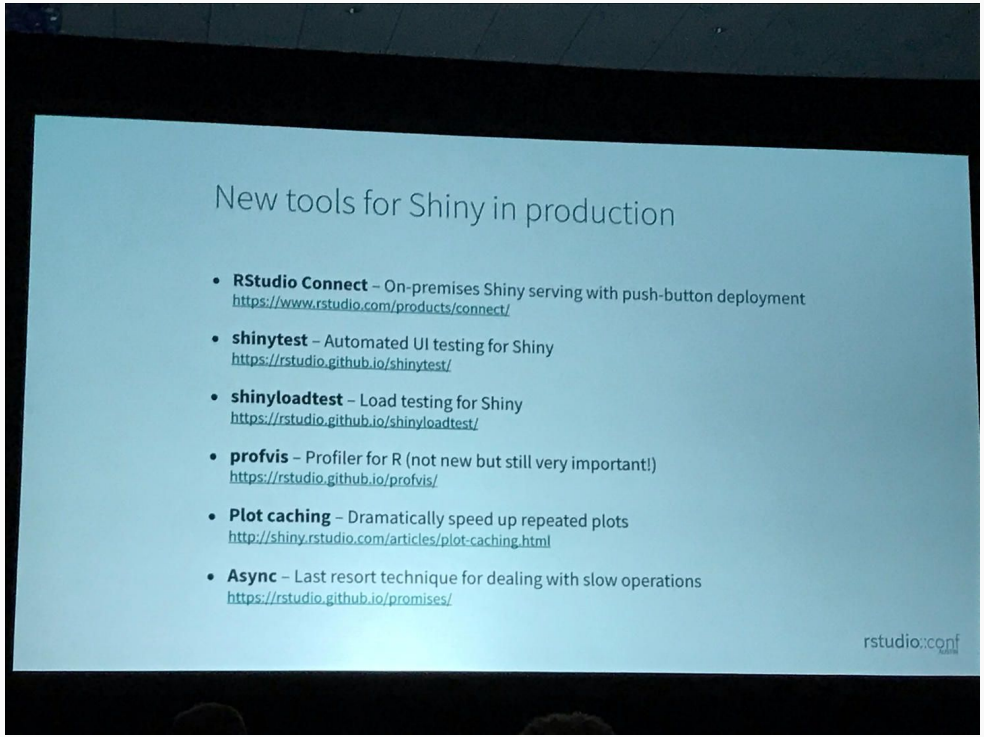
Can R Be Used In Production?

“Yes, it’s quite easy.”

-Joe Cheng, CTO RStudio
Keynote at Rstudio::conf 2019

New Tools

- shinytest
- shinyloadtest
- profvis



“We have new tools.”

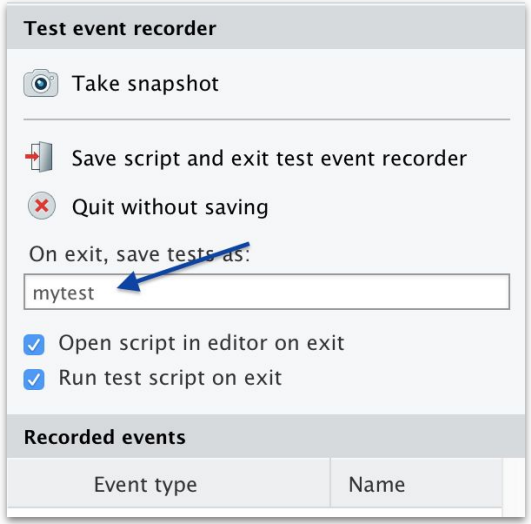
-Joe Cheng, CTO RStudio
Keynote at Rstudio::conf 2019

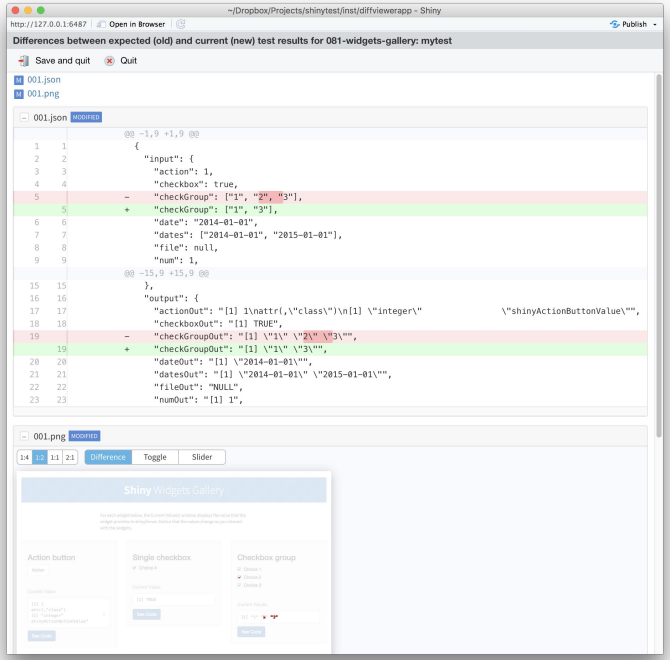
New Tools

shinytest

What has broken as your app has evolved?

https://rstudio.github.io/shinytest/articles/shinytest.html





New Tools

shinyloadtest

How well does your application scale?

https://rstudio.github.io/shinyloadtest/

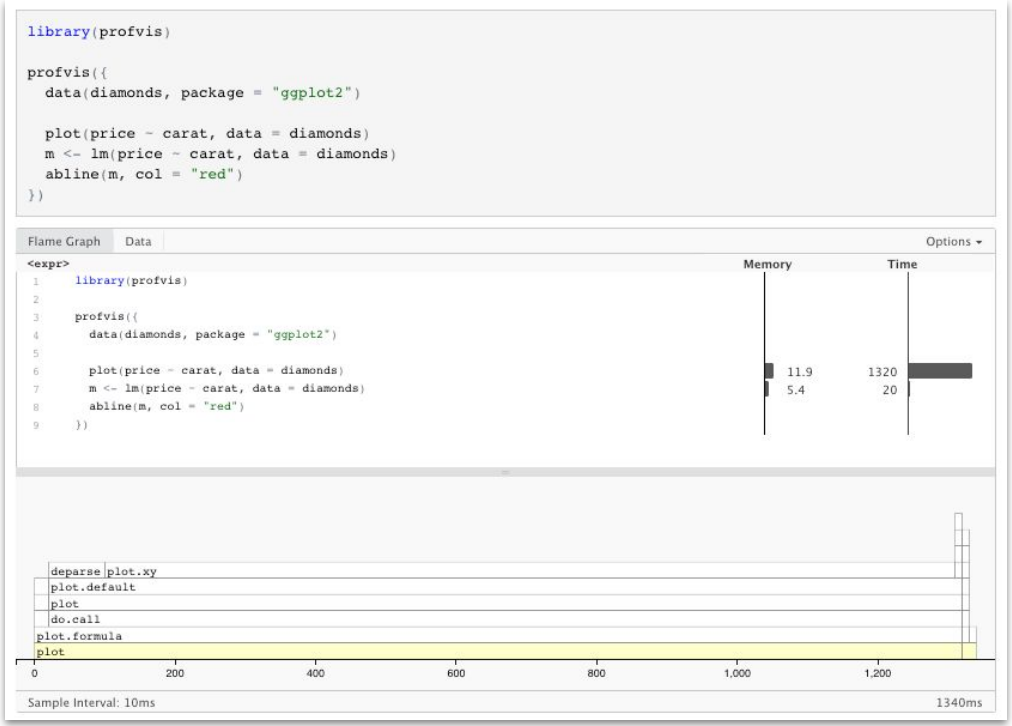
New Tools

profvis

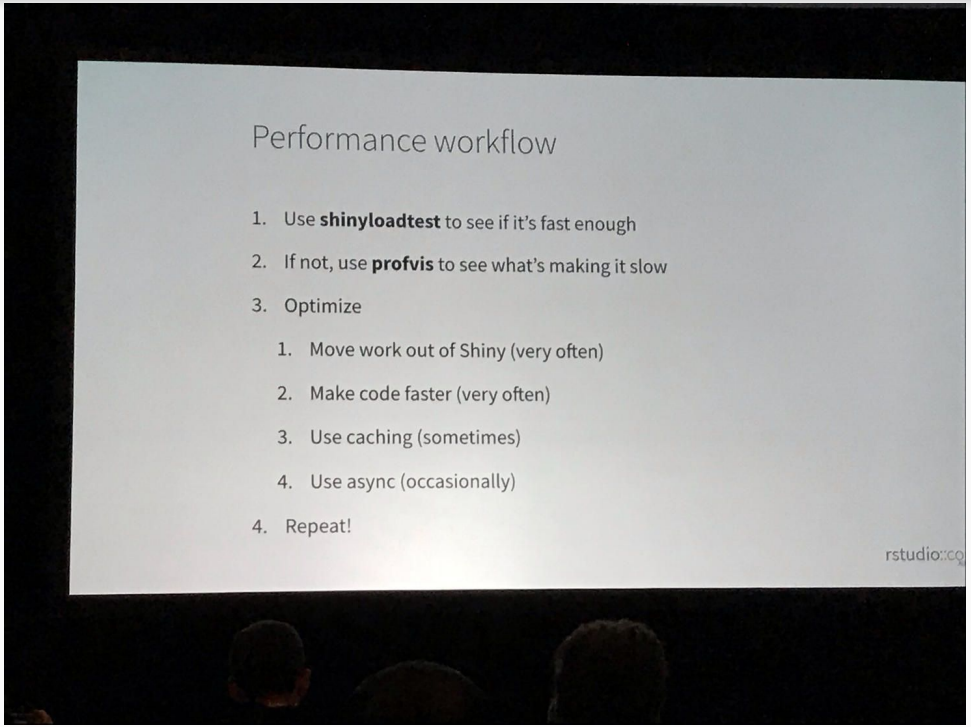
How efficient is your code?

https://rstudio.github.io/profvis/


```
library(profvis)

profvis({
  data(diamonds, package = "ggplot2")

  plot(price ~ carat, data = diamonds)
  m <- lm(price ~ carat, data = diamonds)
  abline(m, col = "red")
})
```



Can R Be Used In Production?



"Recipe for performance."

-Joe Cheng, CTO RStudio
Keynote at Rstudio::conf 2019

Why Production?

$5M

Per Year

How much you can save your organization by solving a large business problem

Case Study: True Cost of Employee Churn

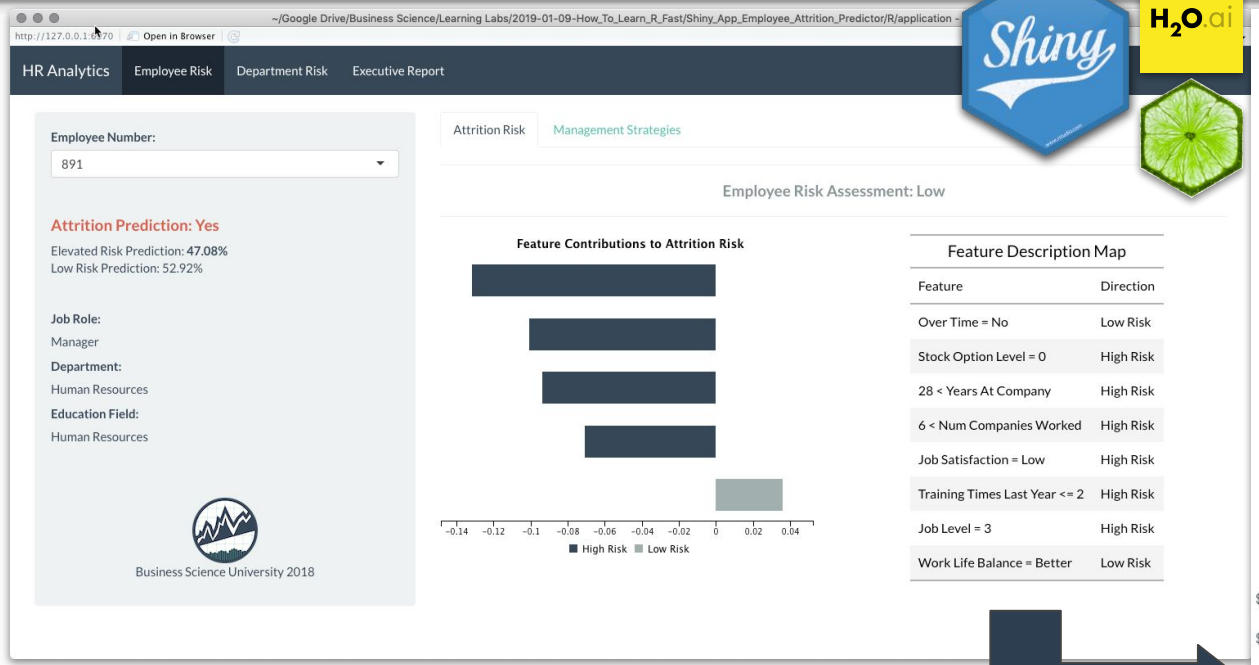
```r
calculate_attrition_cost <- function(

    # Employee
    n                    = 1,
    salary               = 80000,

    # Direct Costs
    separation_cost      = 500,
    vacancy_cost         = 10000,
    acquisition_cost     = 4900,
    placement_cost       = 3500,

    # Productivity Costs
    net_revenue_per_employee = 250000,
    workdays_per_year        = 240,
    workdays_position_open   = 40,
    workdays_onboarding      = 60,
    onboarding_efficiency    = 0.50

) {

    # Direct Costs
    direct_cost <- sum(separation_cost, vacancy_cost, acquisition_cost, placement_cost)

    # Lost Productivity Costs
    productivity_cost <- net_revenue_per_employee / workdays_per_year *
        (workdays_position_open + workdays_onboarding * onboarding_efficiency)

    # Savings of Salary & Benefits (Cost Reduction)
    salary_benefit_reduction <- salary / workdays_per_year * workdays_position_open

    # Estimated Turnover Per Employee
    cost_per_employee <- direct_cost + productivity_cost - salary_benefit_reduction

    # Total Cost of Employee Turnover
    total_cost <- n * cost_per_employee

    return(total_cost)

}
```

- **SIMPLE CALCULATION**
  - Direct costs
  - Lost Productivity
  - Savings (Salary & Benefits)
- **$78K COST / EMPLOYEE**
- **IF ORGANIZATION LOSES 200 HIGH PERFORMERS EACH YEAR...**


$15M / YEAR
PROBLEM

You Only Generate
Business Value When
You Effect Decision

Cause & Effect

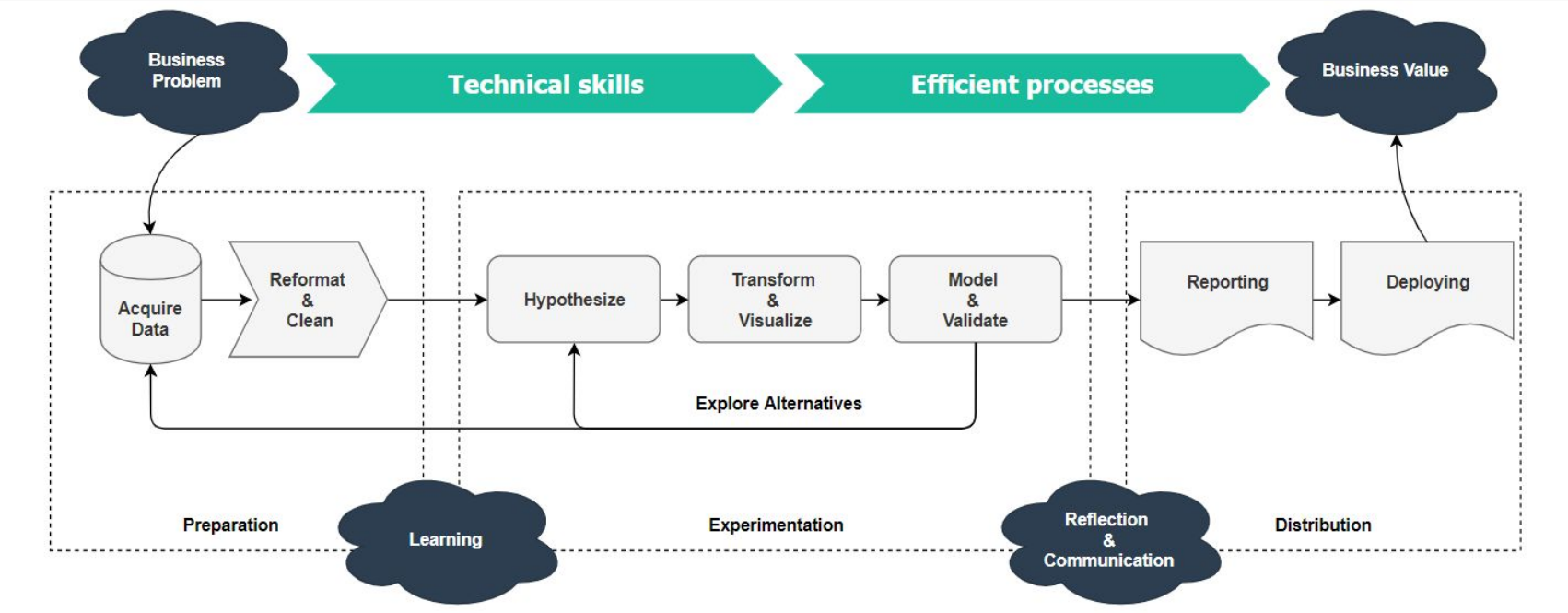


THE "BETTER DECISION-MAKING" EFFECT

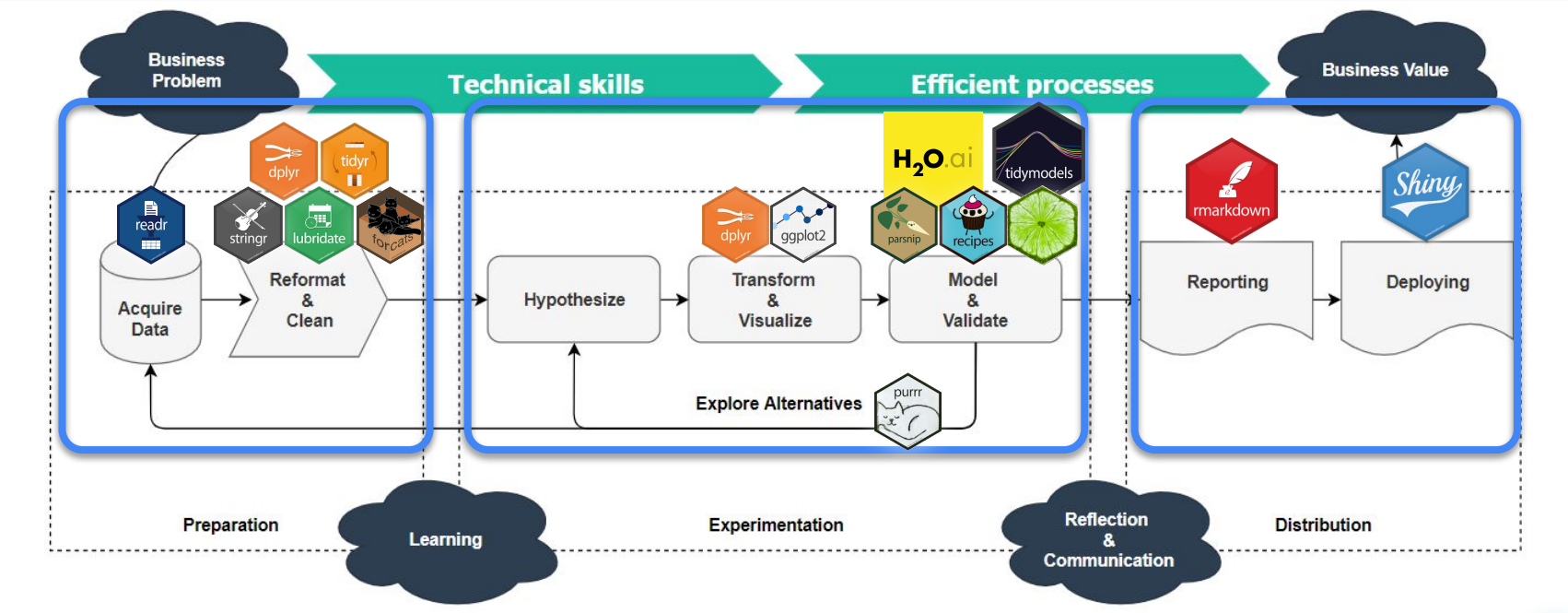
Workflow

For Business Value

Data Science Workflow



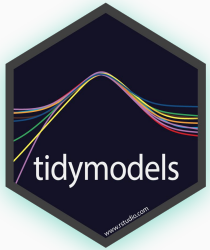
Data Science Workflow



Apply Tools To Problem

Data Science Toolchain

Data Science Workflow



Data Science Toolchain

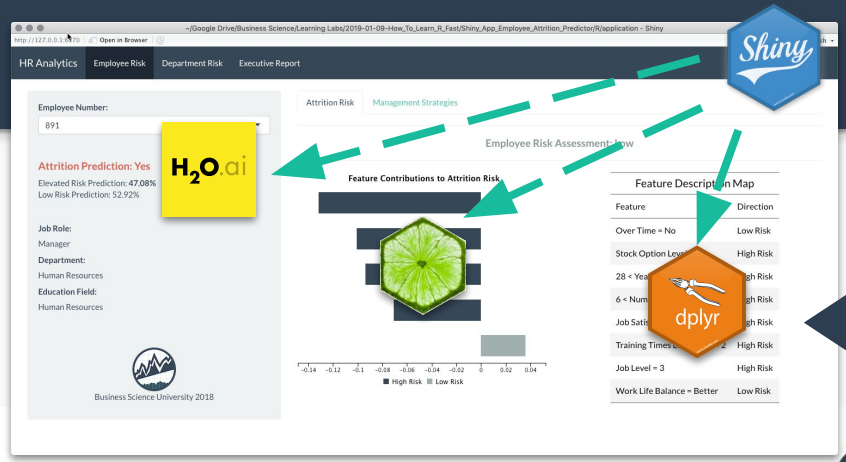
Building

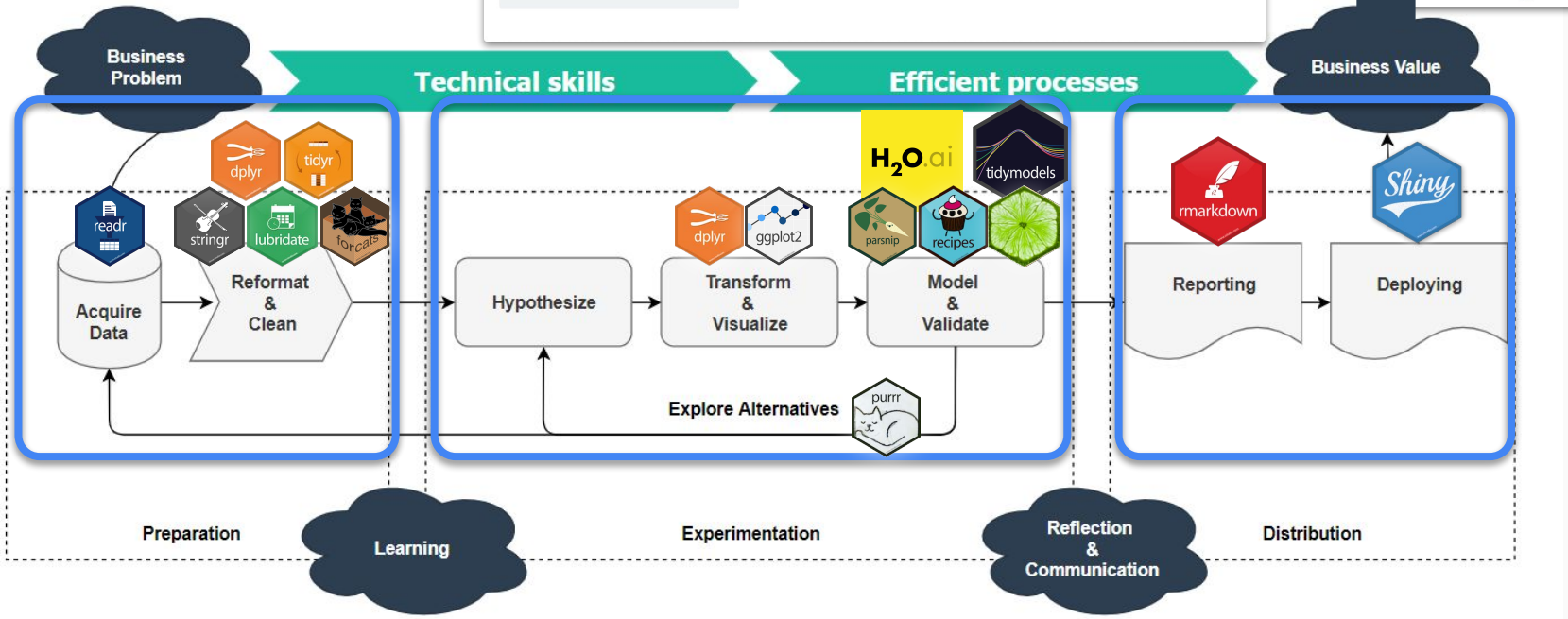
- HTML
- CSS

Testing

- shinytest
- shinyloadtest
- profvis

Data Science Workflow





Case Study

For Generating

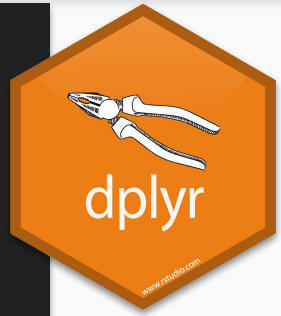
Business Value

Data Science Workflow

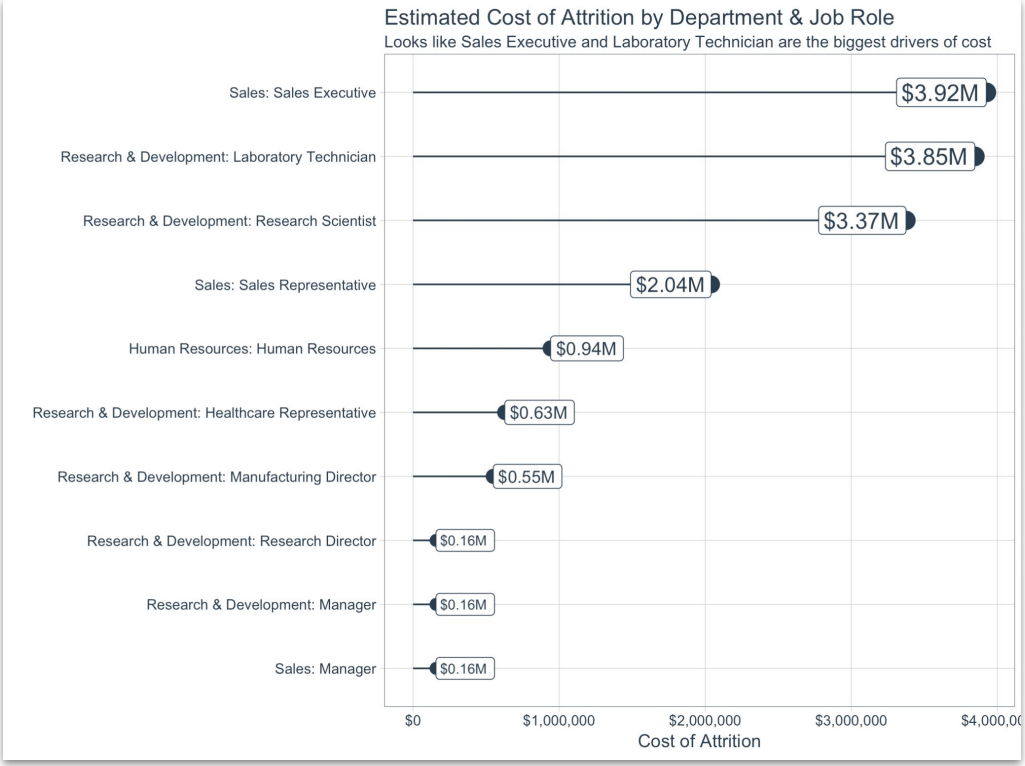
```
# 1C. Measure The Drivers ----

# Collect Information on Employee Attrition: On going

# Develop KPI's: Industry KPIs: 8.8%

dept_job_role_tbl %>%

    count(Department, JobRole, Attrition) %>%

    count_to_pct(Department, JobRole) %>%

    assess_attrition(Attrition, attrition_value = "Yes", baseline_pct = 0.088) %>%

    mutate(
        cost_of_attrition = calculate_attrition_cost(n = n, salary = 80000)
    )
```

Department <chr>	JobRole <chr>	Attrition <chr>	n <int>	pct <dbl>	above_industry_avg <chr>	cost_of_attrition <dbl>
Sales	Sales Representative	Yes	26	0.40000000	Yes	2040566.7
Human Resources	Human Resources	Yes	12	0.30769231	Yes	941800.0
Research & Development	Laboratory Technician	Yes	49	0.21875000	Yes	3845683.3
Sales	Sales Executive	Yes	50	0.18315018	Yes	3924166.7
Research & Development	Research Scientist	Yes	43	0.16602317	Yes	3374783.3
Research & Development	Healthcare Representative	Yes	8	0.07619048	No	627866.7
Sales	Manager	Yes	2	0.06451613	No	156966.7
Research & Development	Manufacturing Director	Yes	7	0.05691057	No	549383.3
Research & Development	Manager	Yes	2	0.04166667	No	156966.7
Research & Development	Research Director	Yes	2	0.02739726	No	156966.7

1–10 of 10 rows



Size the Problem



Visualize Employee Churn Cost

Data Science Workflow

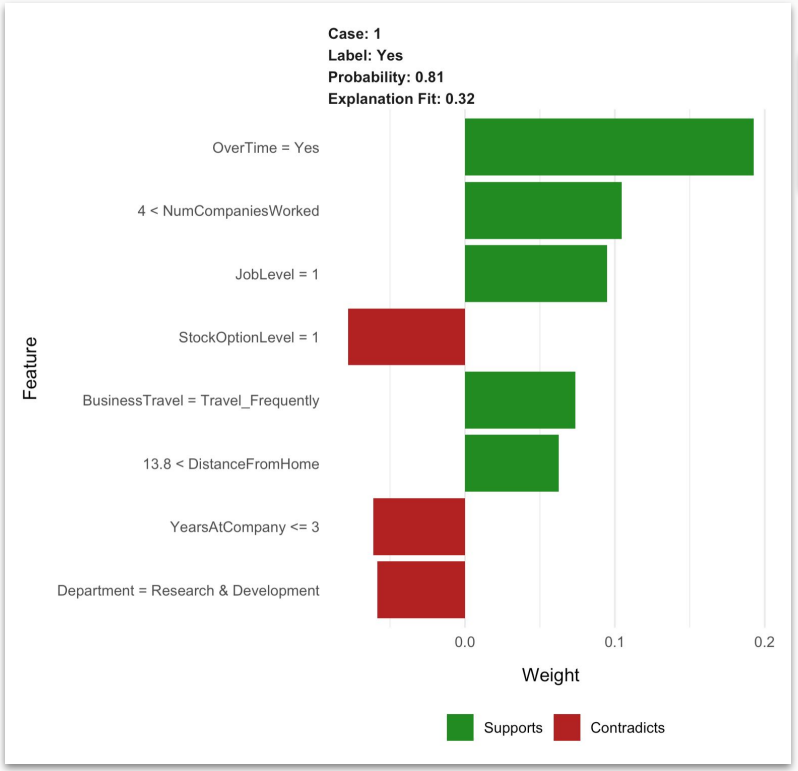
```
# 2. Modeling ----

h2o.init()

split_h2o <- h2o.splitFrame(as.h2o(train_tbl), ratios = c(0.85), seed = 1234)

train_h2o <- split_h2o[[1]]
valid_h2o <- split_h2o[[2]]
test_h2o  <- as.h2o(test_tbl)

y <- "Attrition"
x <- setdiff(names(train_h2o), y)

automl_models_h2o <- h2o.automl(
    x = x,
    y = y,
    training_frame = train_h2o,
    validation_frame = valid_h2o,
    leaderboard_frame = test_h2o,
    max_runtime_secs = 30,
    nfolds = 5
)
```

H2O.ai

Predict Employee Churn Risk

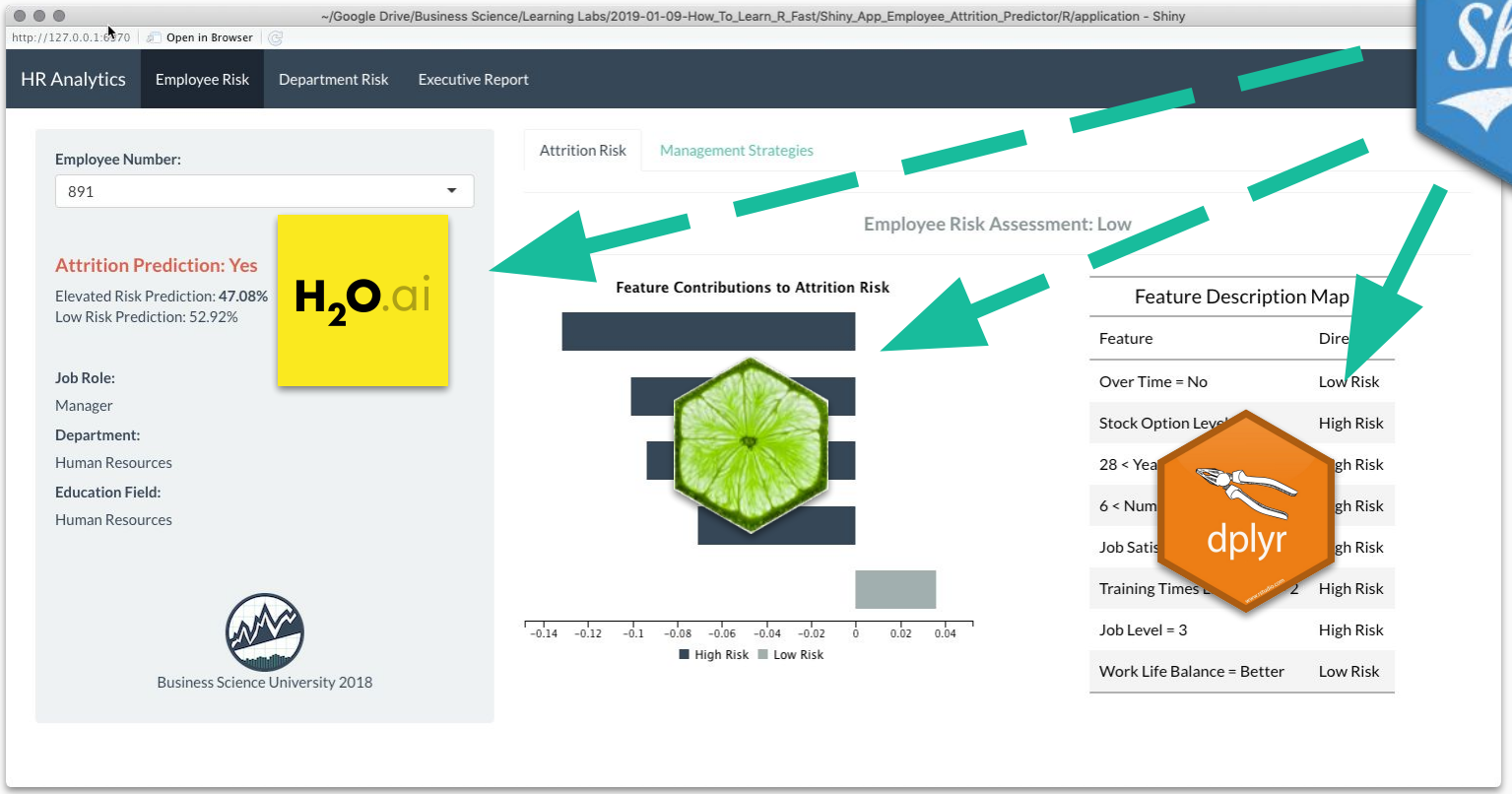
Data Science Workflow

```r
explainer <- train_tbl %>%
    select(-Attrition) %>%
    lime(
        model           = automl_leader,
        bin_continuous  = TRUE,
        n_bins          = 4,
        quantile_bins   = TRUE
    )

explainer

explanation <- test_tbl %>%
    slice(5) %>%
    select(-Attrition) %>%
    lime::explain(
        explainer = explainer,
        n_labels   = 1,
        n_features = 8,
        n_permutations = 5000,
        kernel_width   = 1
    )
```



Explain Features Employee Churn Risk

Data Science Workflow



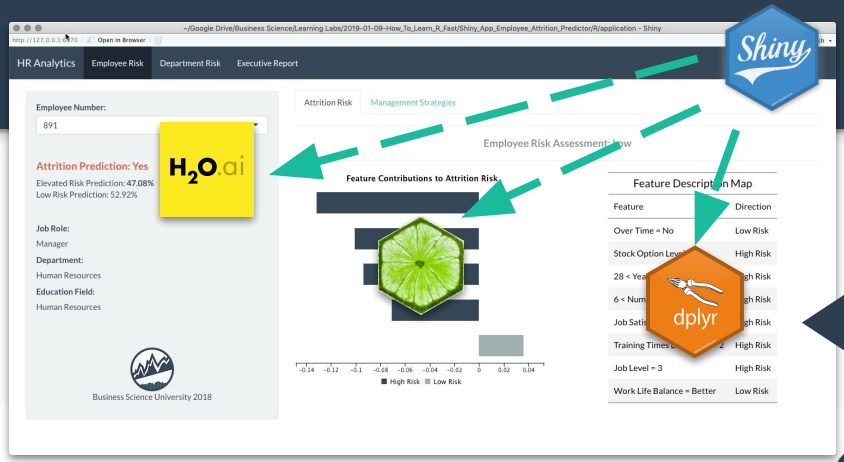
Data Science Workflow

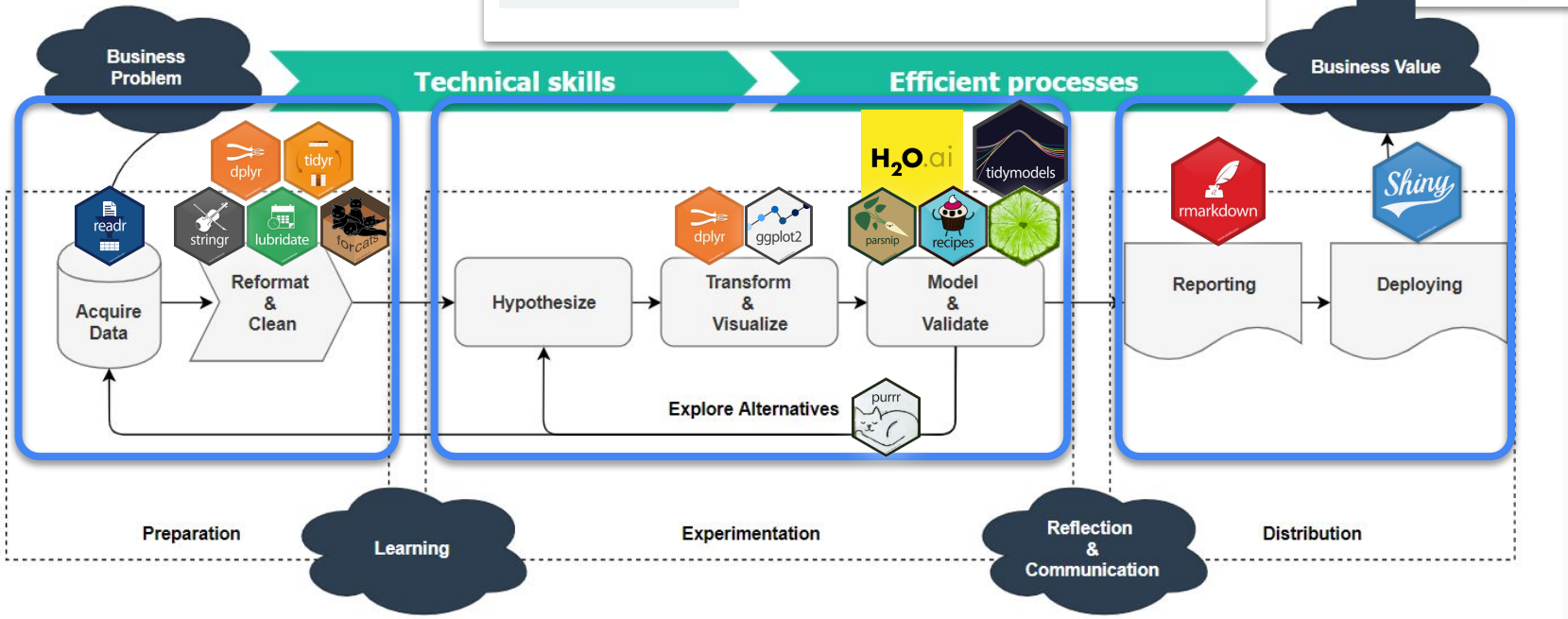
Better decisions benefits the organization financially

$4.3M

Per Year Savings

Path to Business Value

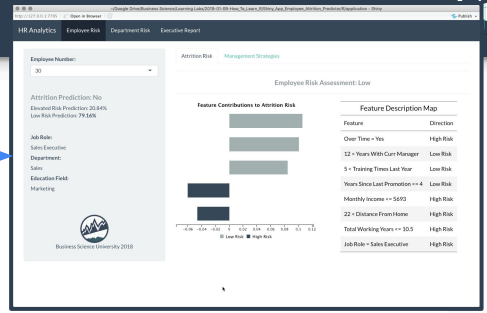


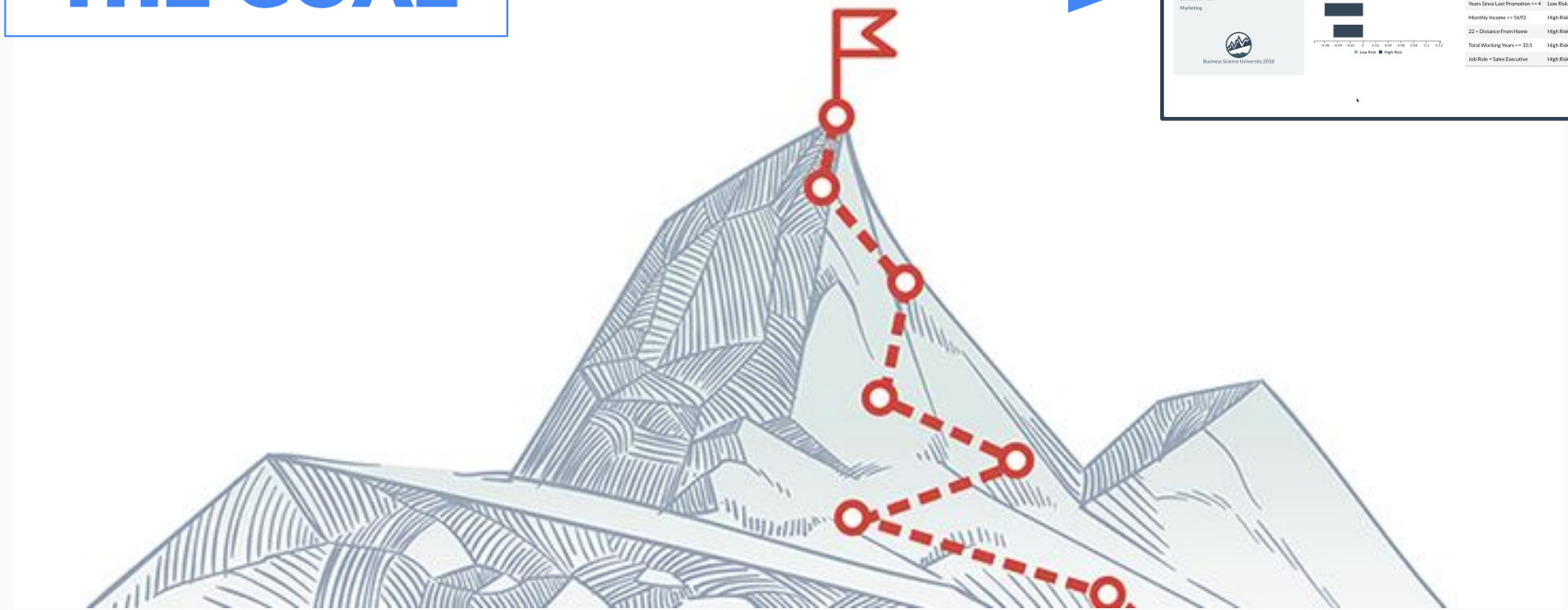


How to get to
Production

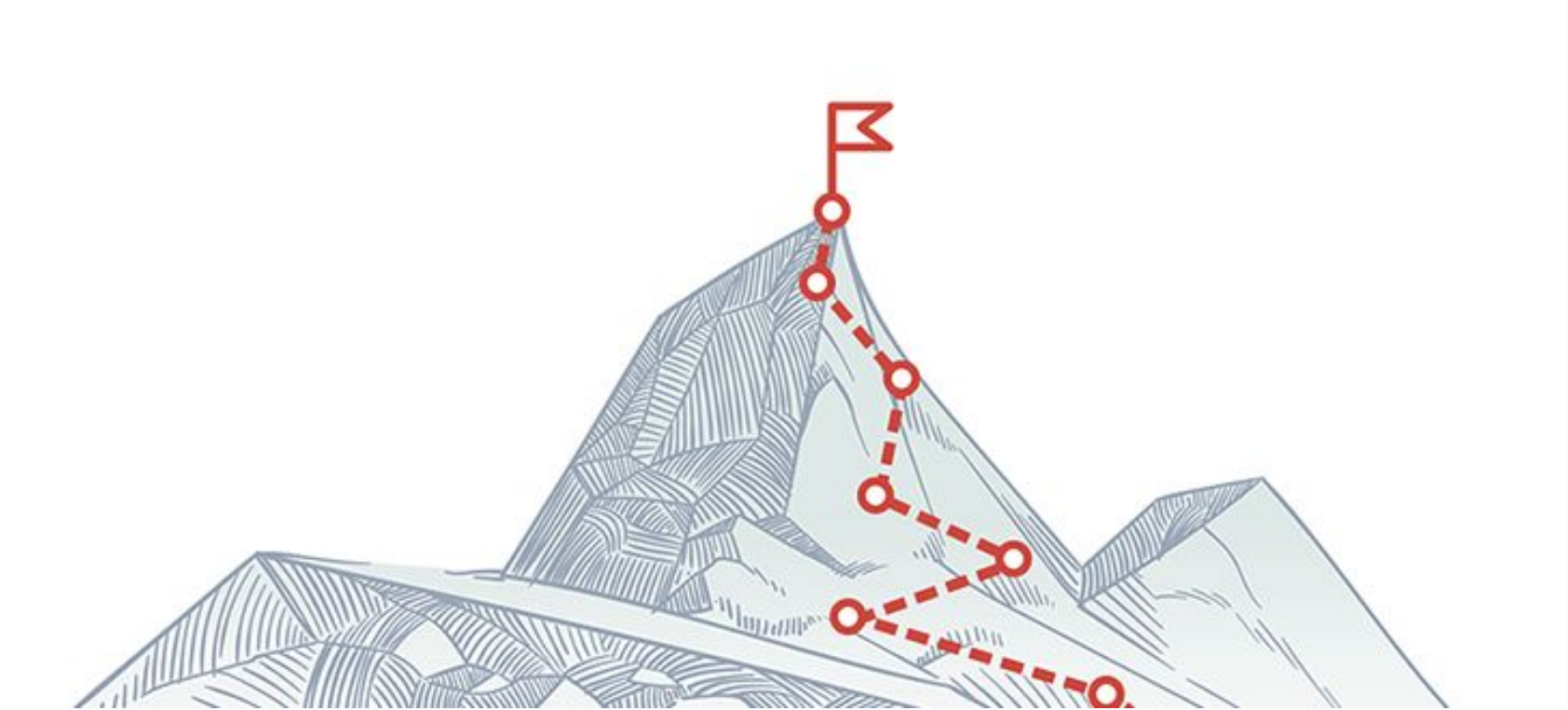
Learning R is a Hill Climb

THE GOAL

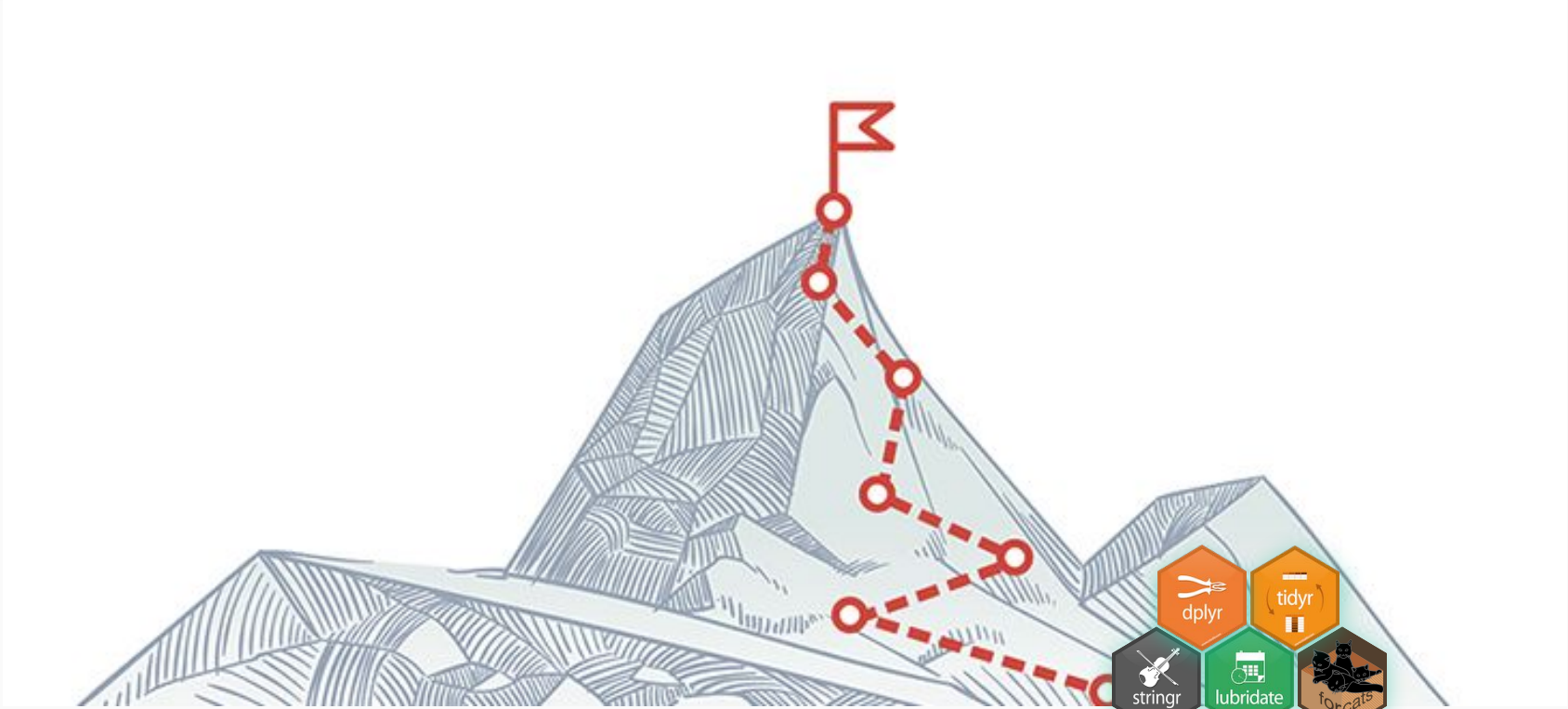




Learning R is a Hill Climb

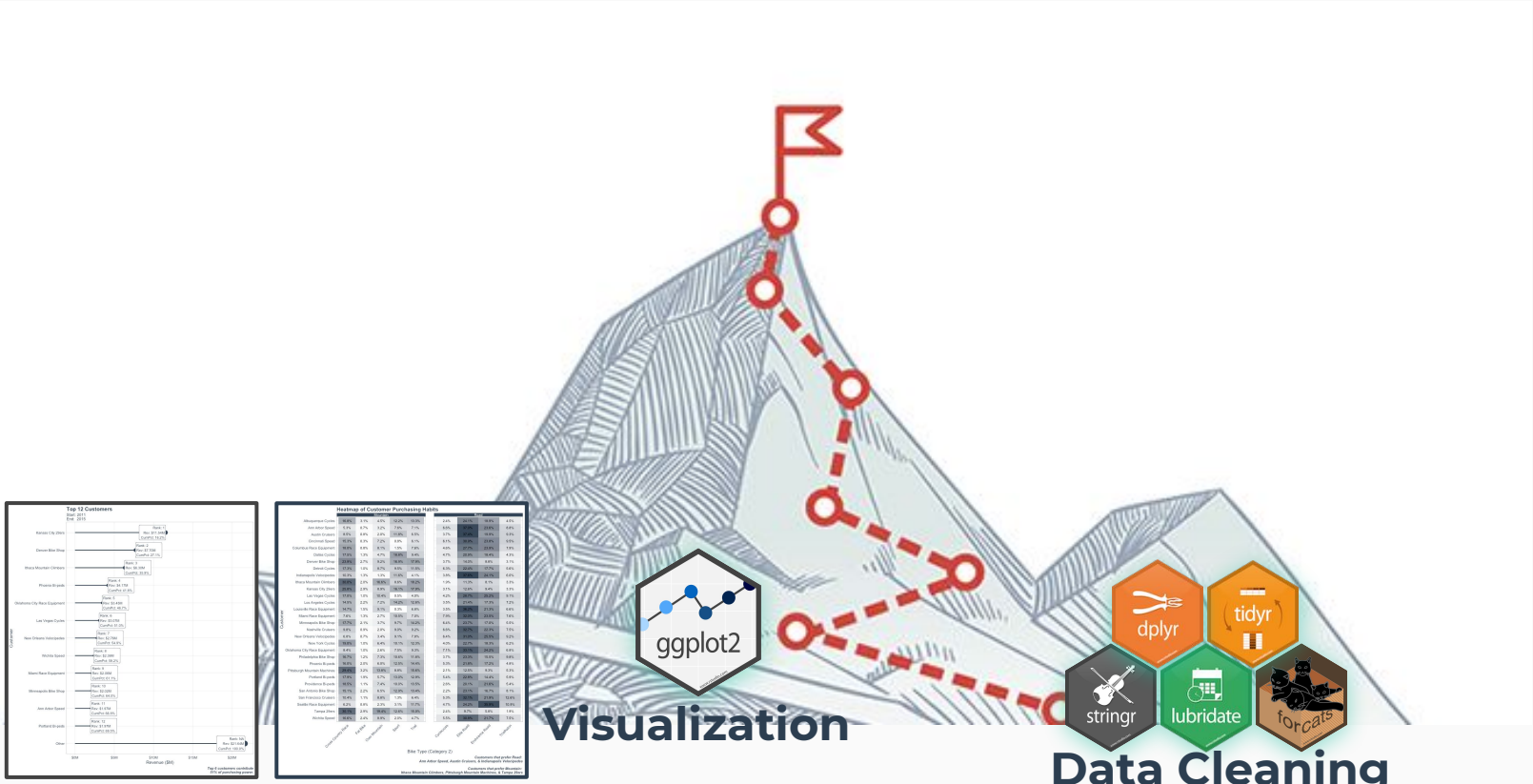


Learning R is a Hill Climb

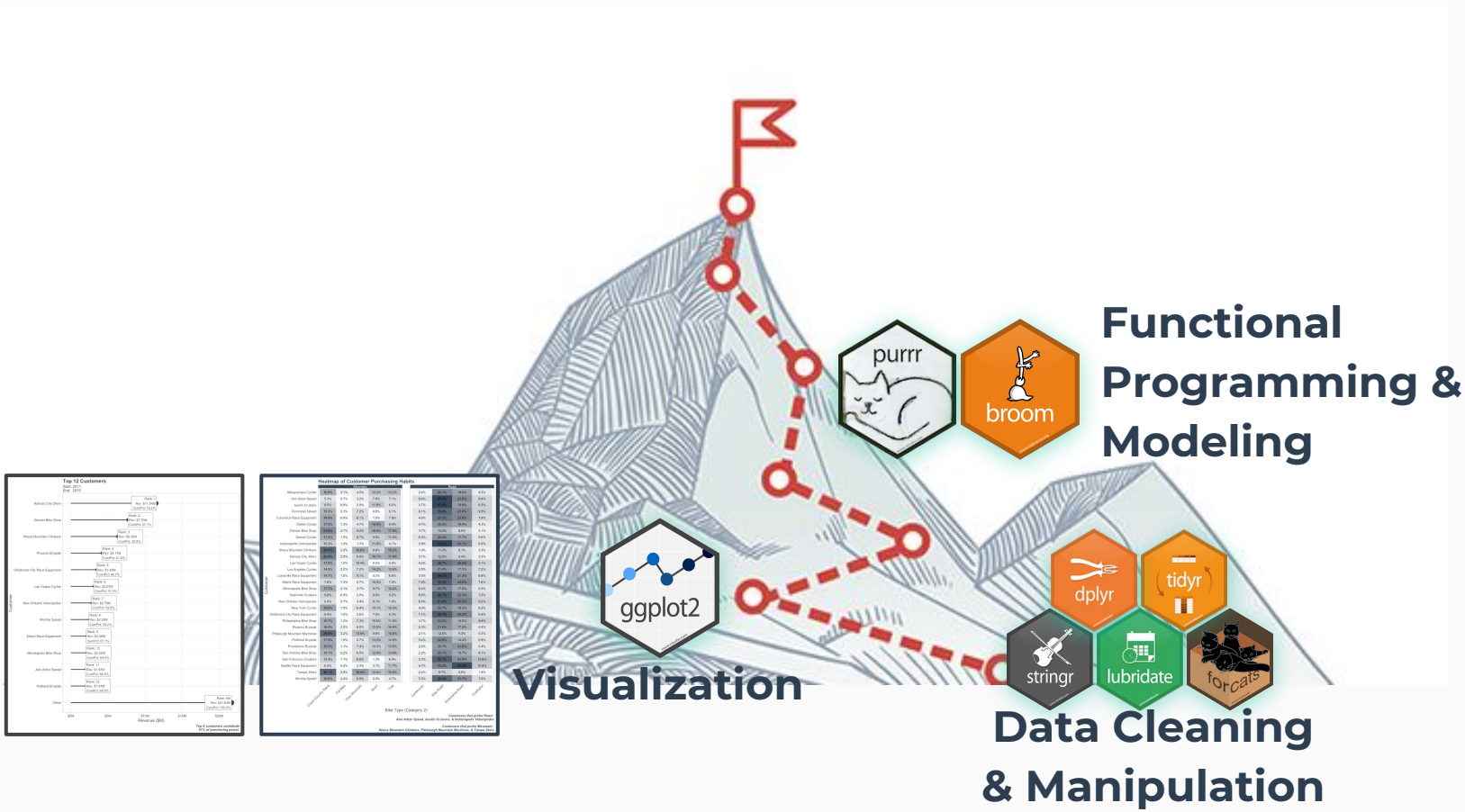


Data Cleaning & Manipulation

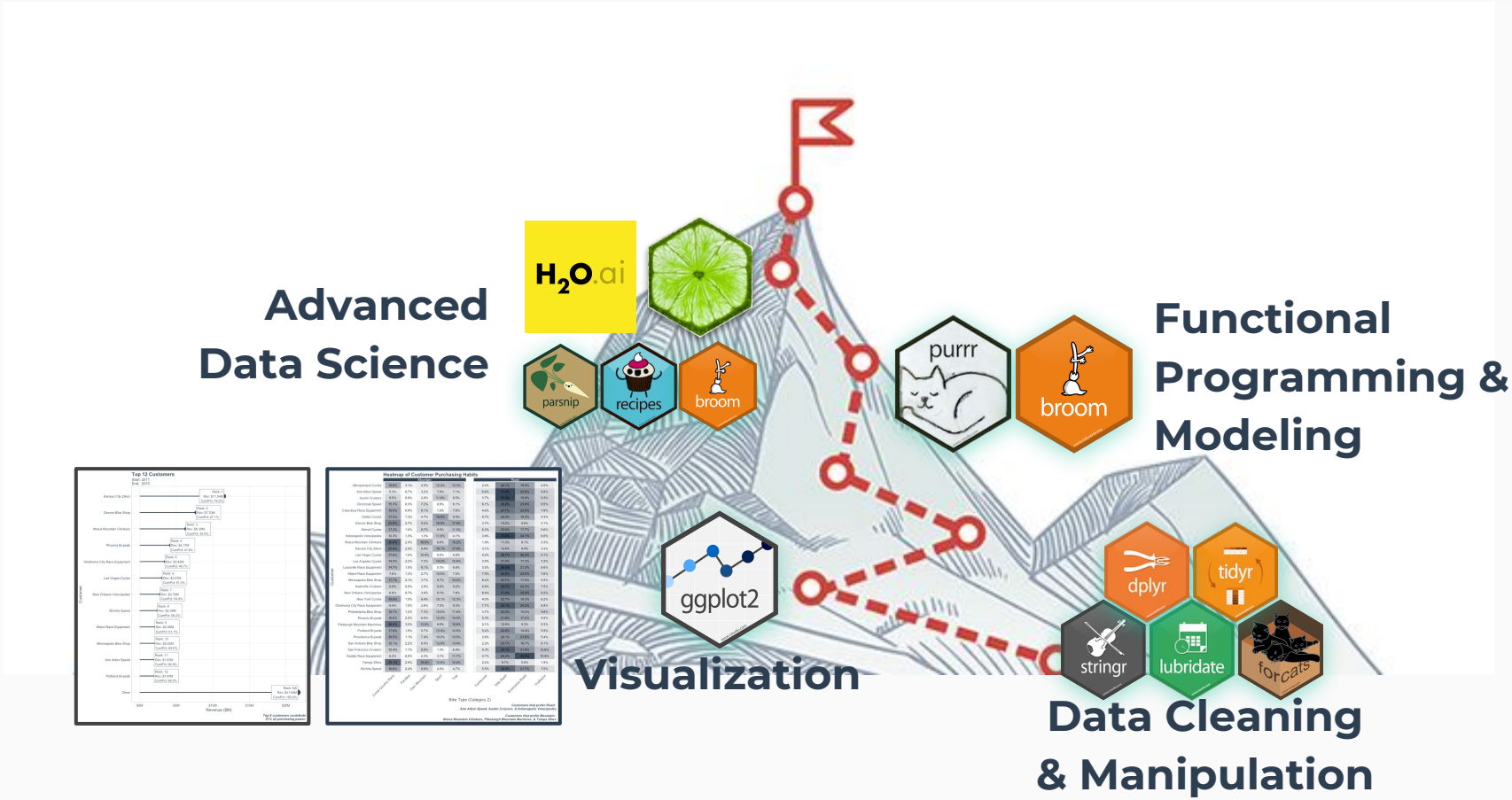
Learning R is a Hill Climb



Learning R is a Hill Climb

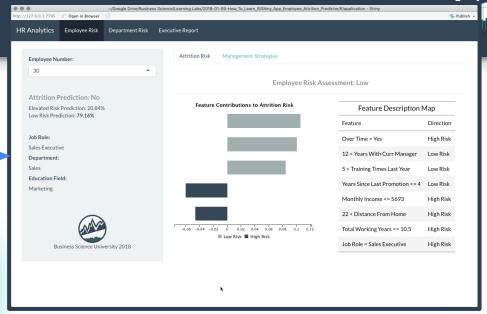


Learning R is a Hill Climb



Learning R is a Hill Climb

THE GOAL



Advanced Data Science

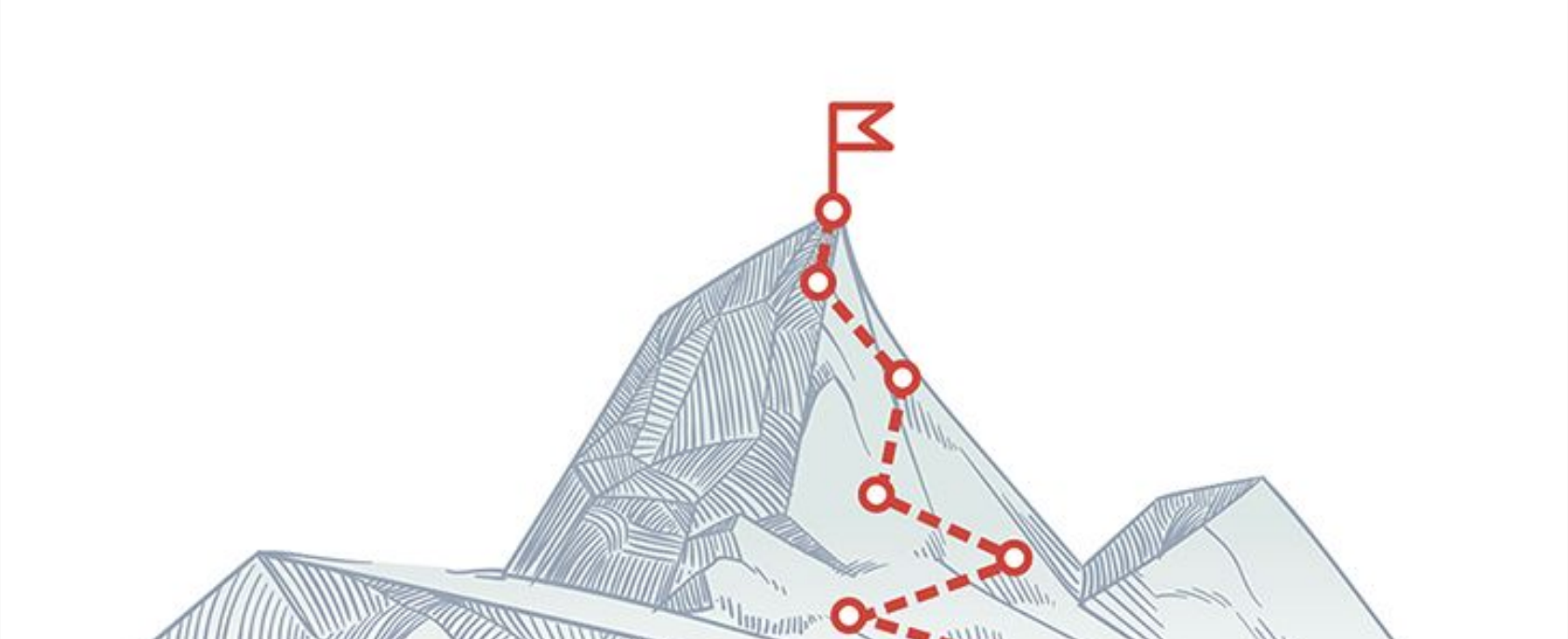
Functional Programming & Modeling

Visualization

Data Cleaning & Manipulation

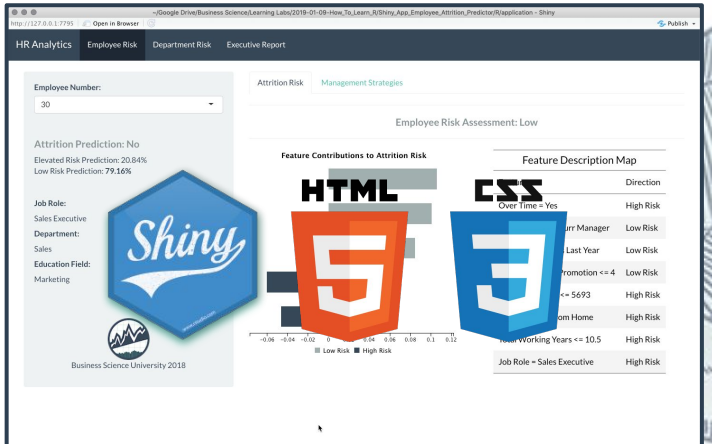
Last Mile
is Challenging

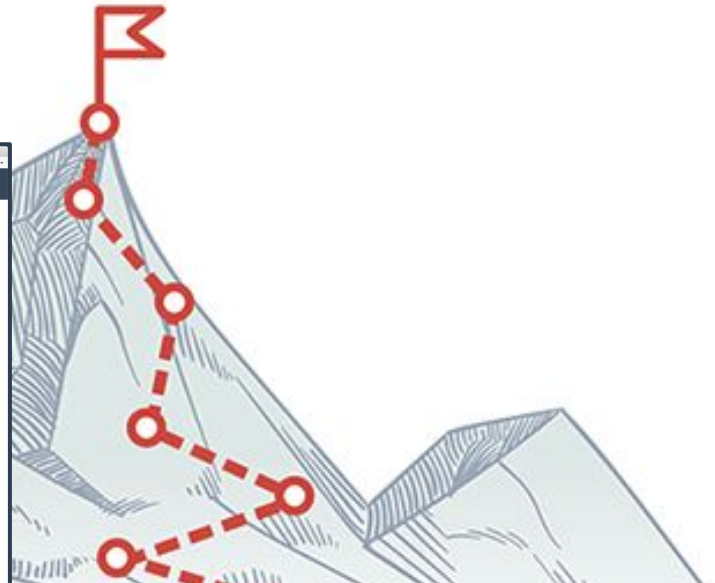
Last Mile is Challenging



Last Mile is Challenging

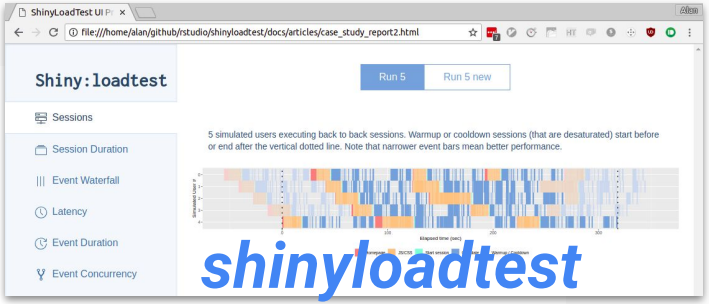
Build App



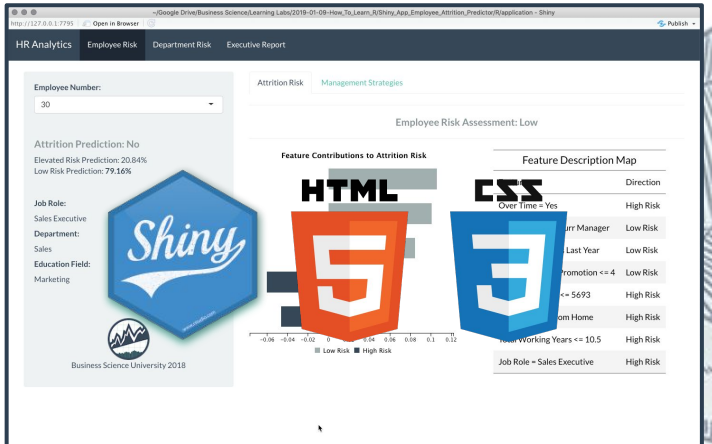


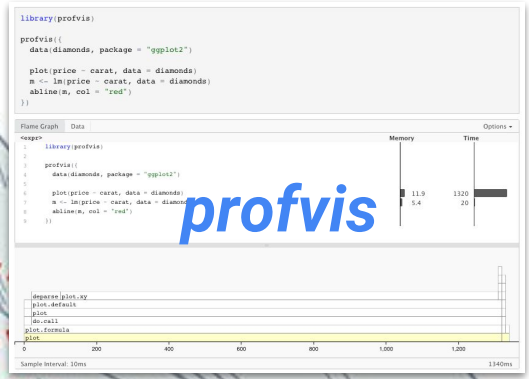
Last Mile is Challenging

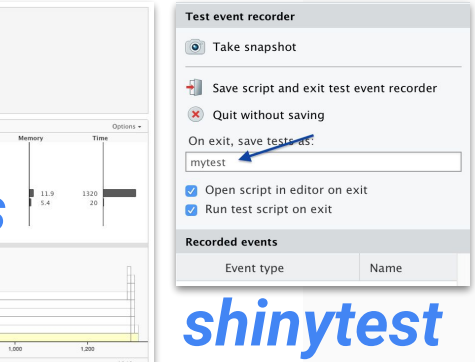
Test App



Build App



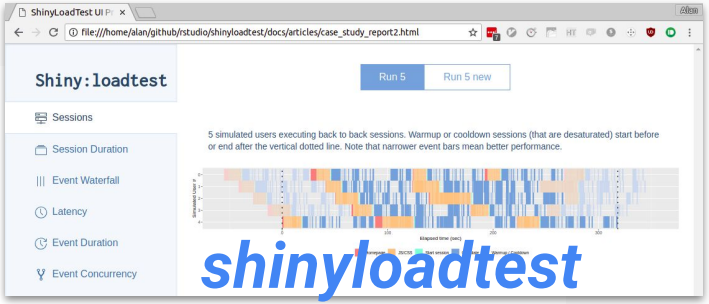




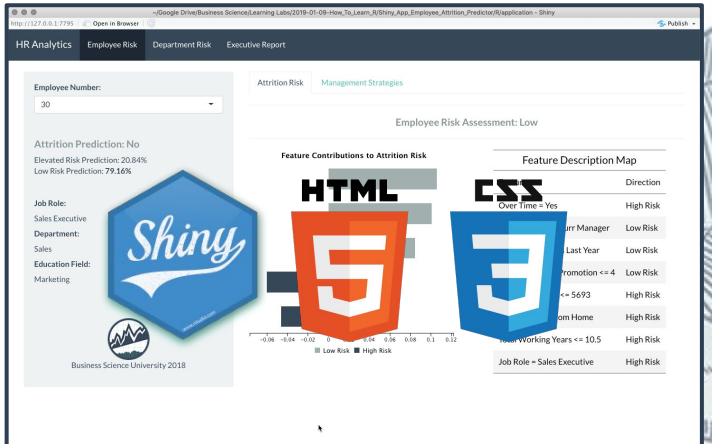
Last Mile is Challenging

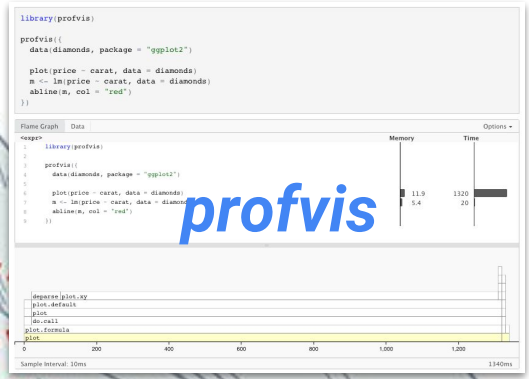
Production Quality

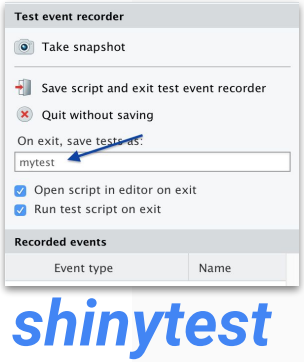
Test App



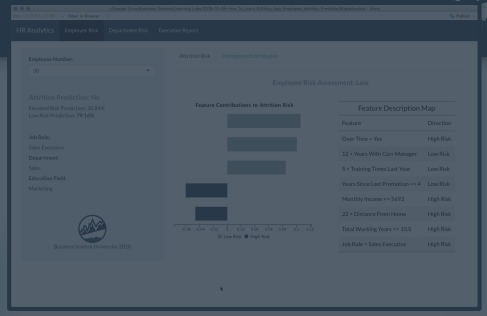
Build App







Path To Top Can Be Accomplished FAST

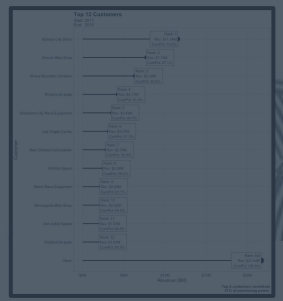


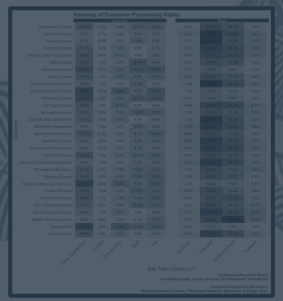
Advanced Data Science

Functional Programming & Modeling

Visualization

Data Cleaning & Manipulation





The Plan

R-TRACK

THE PLAN

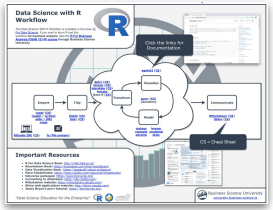
Business Analysis with R
(DS4B 101-R)

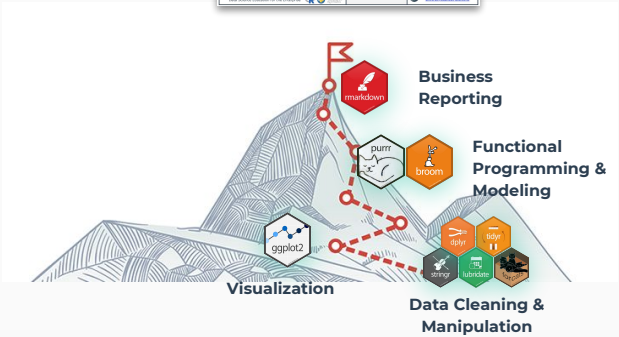
Data Science For Business with R
(DS4B 201-R)

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

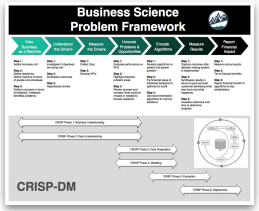
Project-Based Courses with Business Application

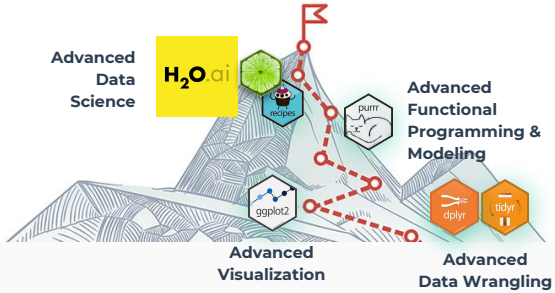
Data Science Foundations
7 Weeks



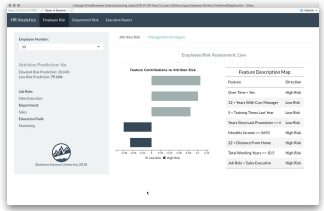


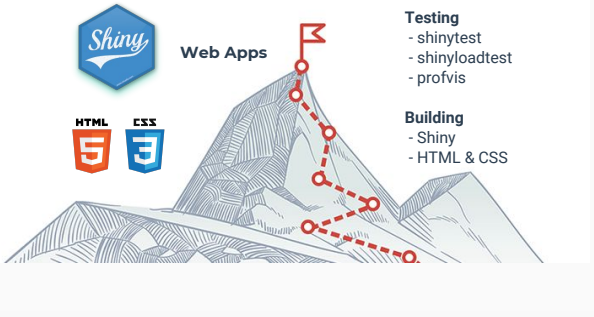
Machine Learning & Business Consulting
10 Weeks





Web Application Development
6 Weeks





THE RESULT

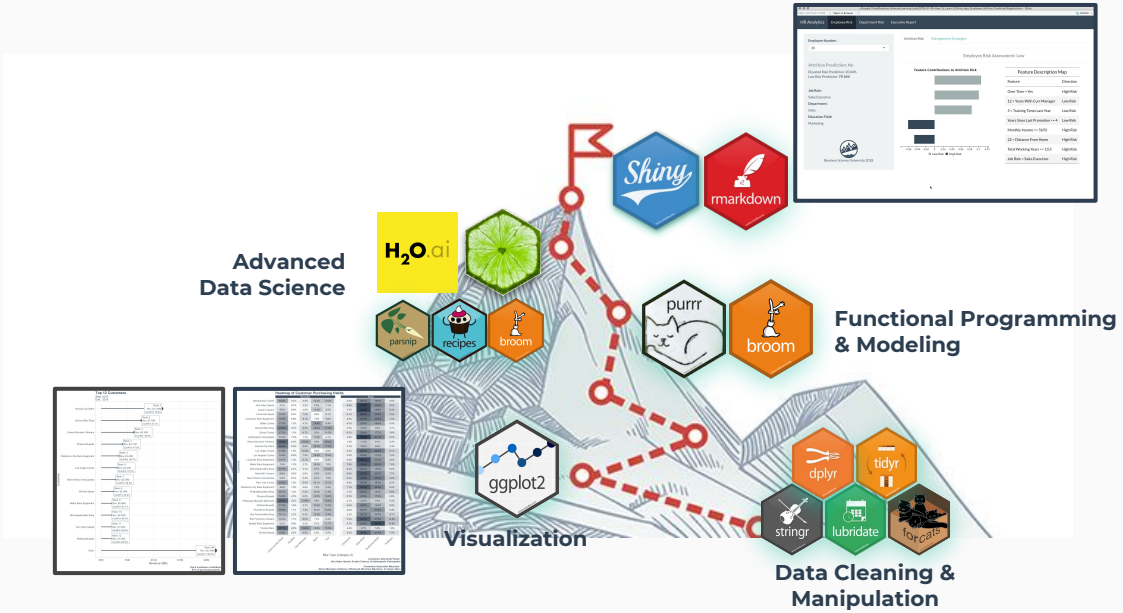
Business Analysis with R
(DS4B 101-R)
7 WEEKS

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

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

23 WEEKS!

(OR LESS)



15% OFF PROMO Code: learninglabs

DS4B 101-R: Business Analysis With R

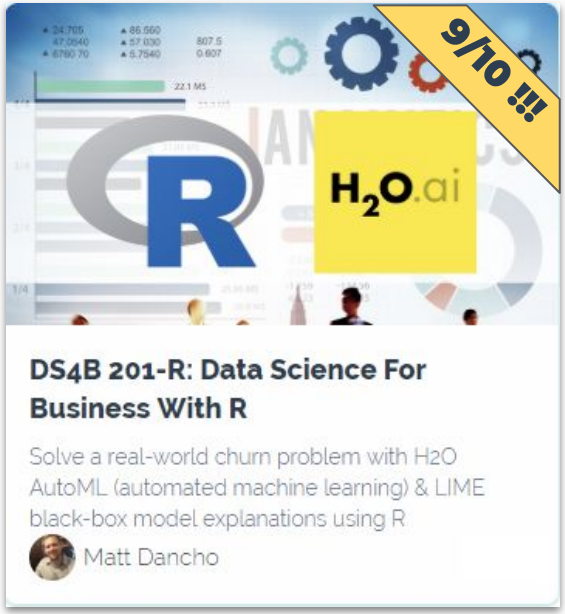
Your Data Science Journey Starts Now! Learn the fundamentals of data science for business with the tidyverse.

Matt Dancho

MSRP: ~~$349~~

Your Price! **$297**

Save: $52



DS4B 201-R: Data Science For Business With R

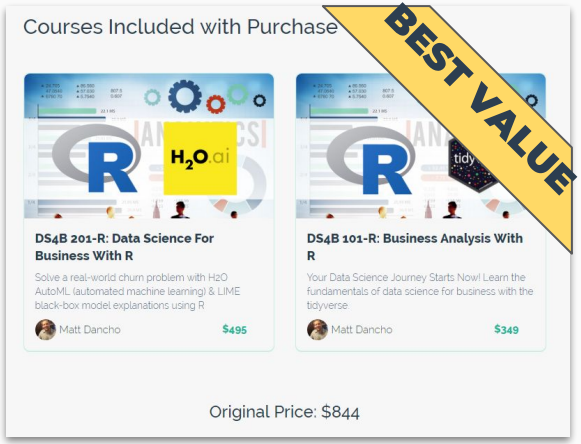
Solve a real-world churn problem with H2O AutoML (automated machine learning) & LIME black-box model explanations using R

Matt Dancho

MSRP: ~~$499~~

Your Price! **$421**

Save: $78



R-TRACK BUNDLE

MSRP: ~~$844~~

Your Price! **$637**

Save: $207

R-TRACK

THE BONUS

Start your journey with 15% OFF
PROMO Code: learninglabs

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**