Breakout Sessions 2:15 pm – 3:00 pm

- ***Process and Technical Infrastructure for Data Science ROI. Case***
- ***Are You Vaccinated Against Risk? Cultural, Technological, and Process Tools,***

Process and technical infrastructure for Data Science ROI Case Study: Predicting and preventing clinic no shows

Tristan Markwell
Lindsay Mico
17 May 2018

AIRBNB ENGINEERING & DATA SCIENCE

Creative engineers and data scientists building a world where you can belong anywhere.

Featured Projects

ReAir

Easy-to-use tools for replicating tables and partitions between Hive data warehouses.

View Project

Apache Superset (incubating)

A data exploration platform designed to be visual, intuitive, and interactive.

View Project

Enzyme

JavaScript Testing utilities for React.

View Project

Aerosolve

A machine learning package built for humans.

View Project

Apache Airflow (incubating)

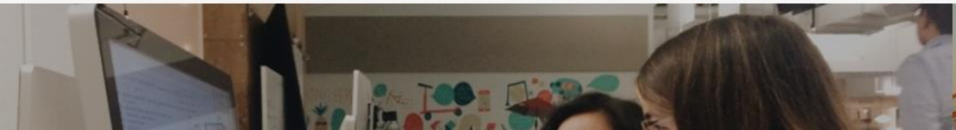
Use Apache Airflow (incubating) to author workflows as directed acyclic graphs (DAGs) of tasks.

View Project

Airpal

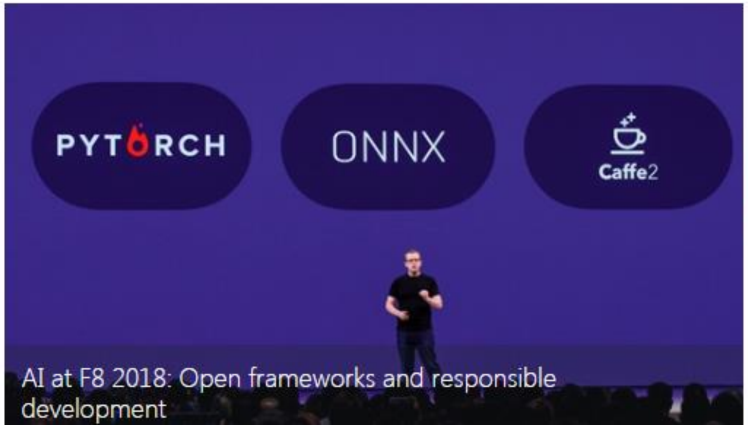
A web-based, query execution tool for Facebook's PrestoDB.

View Project



Applied Machine Learning

Our Applied Machine Learning team improves Facebook products and services through artificial intelligence. We develop and advance algorithms that rank feeds and search results, create new text understanding algorithms that keep spam and misleading content at bay, and automatically caption videos in your news feed through our speech recognition systems. We're also responsible for displaying billions of translated stories every day, developing computer vision algorithms that make images and videos accessible to the blind, and creating magical visual experiences such as turning panorama photos into fully interactive 360 images. Our efforts form the glue between science, research, and Facebook experiences.



AI at F8 2018: Open frameworks and responsible development

Announcing PyTorch 1.0 for both research and production


Advancing state-of-the-art image recognition with deep learning on hashtags

ONNX expansion speeds AI development

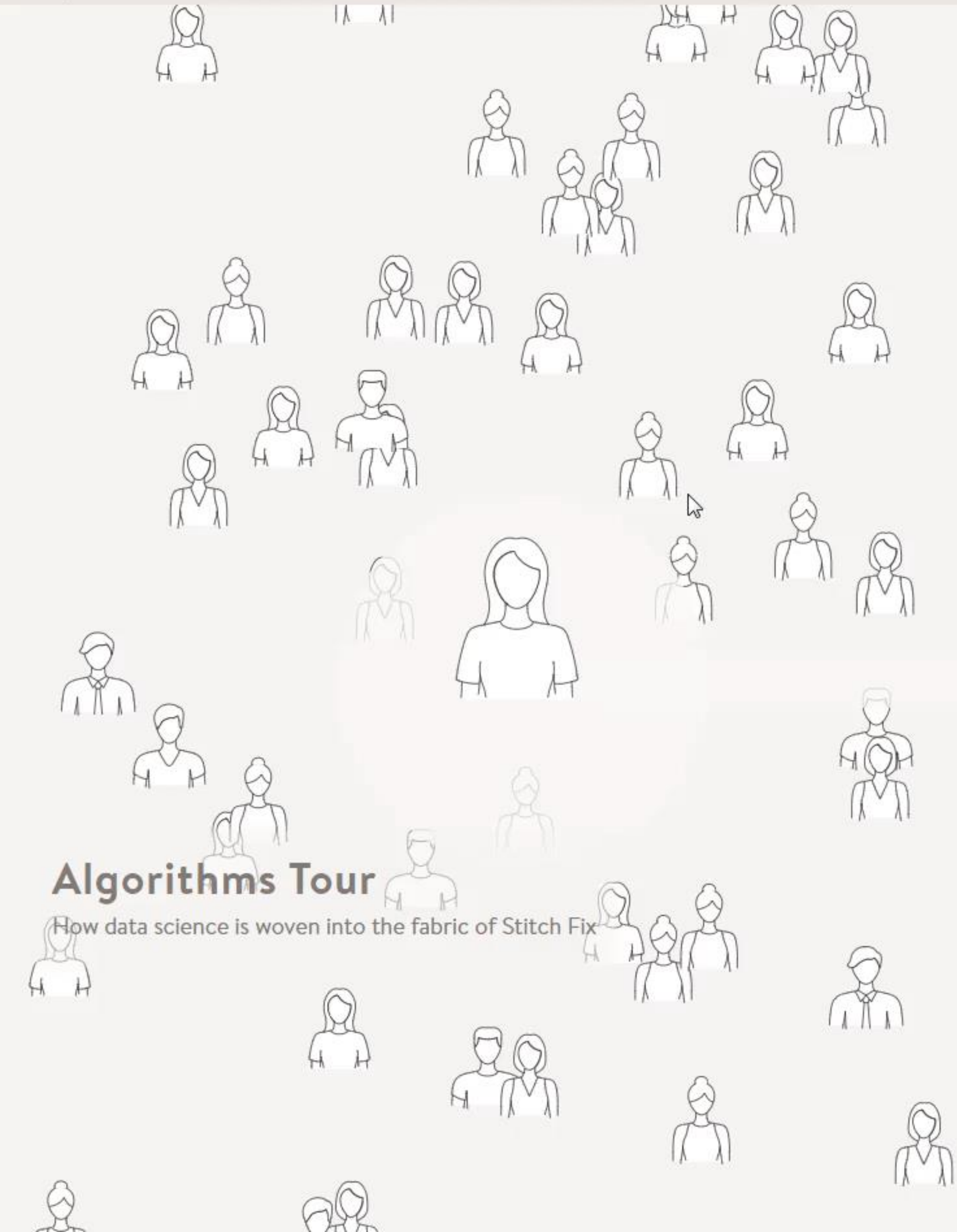

Embodied Question Answering: A goal-driven approach to autonomous agents

Tasks | Trello | Rocker: R and Docker (a | LIME - Local Interpretab | What are the steps / too | Cabin: Thoroughly Learn | Hosted Search as a Serv | Cabin - React & Redux E | Togaware:

algorithms-tour.stitchfix.com

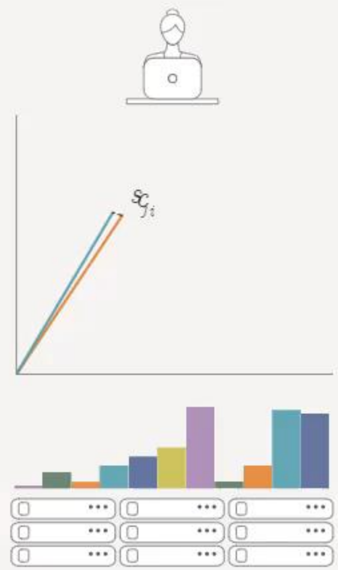
Apps | Travel | Conference rooms | Hue Dev | Archives - Data Scien | The Data Science Rou | Data Science Weekly | Top Stories | Get the O'Reilly Data | Analytics Dispatch Ar | Healthcare News | Hc | No Show De

MultiThreaded   Tweet the Tour!   Eng



Algorithms Tour

How data science is woven into the fabric of Stitch Fix



$$\log \frac{p}{1-p} = a + X\beta + Zb$$

...

$$\min_{a} \sum_{i} \sum_{j} a_{ij} q_{ij}$$

$$\text{s.t.} \quad a_{ij} \in \{0,1\},\ \forall i,j$$

$$\sum_{j} a_{ij} = 1\ \forall i$$

$$\sum_{i} a_{ij} < k_j\ \forall j$$

...

$$\frac{\partial x}{\partial t} = f(x_t,\ u_t,\ w_t)$$

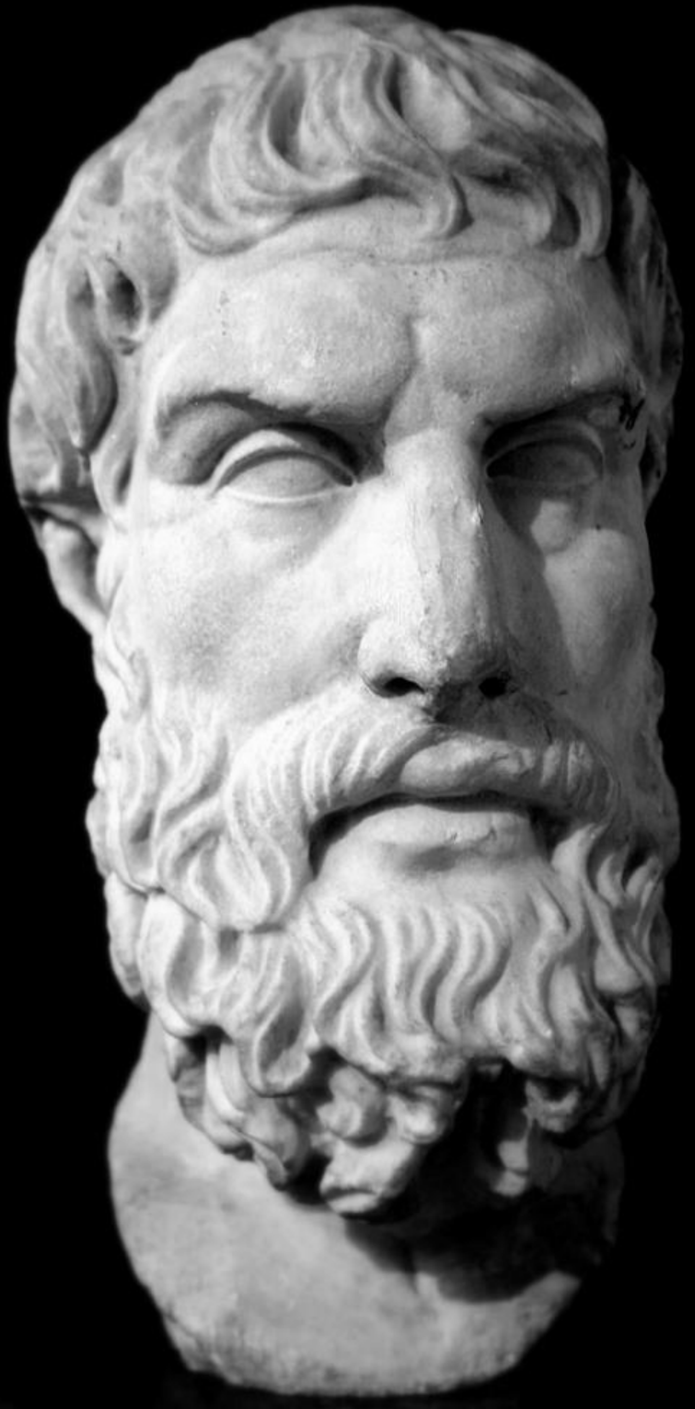
Awareness of the gap

Do we know

- How many patients will admit tomorrow?
- Which providers will retire next year?
- Which patients in the facility/community need the most attention?
- Which surgery method is best?

Photo by Dave Killen for the Oregonian


data science is a journey
Photo by Fabien Wl on Unsplash



If you want to improve, be content to be thought foolish and stupid.

Epictetus

Level 1: Business Intelligence

Photo by Aaron Burden on Unsplash

50 HOSPITALS

829 CLINICS

90 NON-ACUTE SERVICES

14 SUPPORTIVE HOUSING PROGRAMS

111k CAREGIVERS

38k NURSES

20k PHYSICIANS

HIGH SCHOOL, NURSING SCHOOLS AND UNIVERSITY

2 HEALTH PLANS

1.9m COVERED LIVES

$1.6b COMMUNITY BENEFIT

Picking a Great First Project

- Problem with lots of examples and a clear target
- Demonstrable return from Data Science
  - Concrete costs
  - Are you trying to demonstrate that something isn’t happening?
- Minimal integration needed
  - Likely not clinical
- Do you have an area where X intervention works but you don’t have enough of it for everyone?
- Do you have a situation where you’re already using a predictive model (LACE for readmission, etc)?

PMG No-Show Analysis

Lists no-shows as a count and a percentage of no-shows and arrived/completed appointments. Example: 2 (20%)
Immediate Care / Urgent Care departments are excluded.

Clinic(s) selected: All
Hospital clinic(s) selected: None

No-Show Rates 04/01/2018 to 04/30/2018: Enterprise

Percentages are based on the specific payor class, confirmation status, or appointment lag sub totals and not the overall total.

			Appointment Type					
		Total	OVR	OVE	NP	WE	PX	C
	Total	[redacted]	**10,885 (6.3%)**	**3,236 (7.1%)**	[redacted]	**2,037 (6.8%)**	**297 (5.1%)**	**19,046 (7**
Payor Class	Capitation	**302 (9.4%)**	86 (5.3%)	15 (5.7%)	[redacted]	10 (5.0%)	1 (3.6%)	181 (20
	Commercial	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Managed Care	**829 (7.4%)**	406 (6.9%)	46 (6.8%)	[redacted]	115 (7.8%)	4 (5.0%)	195 (9
	Medicaid	**1,689 (13.0%)**	398 (12.3%)	151 (14.5%)	[redacted]	64 (13.8%)	16 (8.3%)	906 (12
	Medicaid HMO	**7,938 (11.9%)**	3,154 (11.4%)	863 (14.1%)	[redacted]	789 (14.3%)	119 (10.9%)	2,193 (10
	Medicare	**4,673 (4.5%)**	1,589 (4.1%)	574 (4.9%)	[redacted]	64 (2.3%)	35 (3.1%)	2,128 (4
	Medicare HMO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Other	**31 (3.9%)**	12 (5.4%)	3 (2.8%)	[redacted]	2 (2.4%)	0 (0.0%)	11 (3
	OTHER GOVERNMENT	**566 (6.3%)**	180 (5.9%)	62 (7.8%)	[redacted]	24 (4.9%)	3 (2.1%)	236 (6
	Uninsured	**7,526 (12.5%)**	536 (19.8%)	242 (17.3%)	[redacted]	209 (17.4%)	12 (9.2%)	6,170 (11
	Worker's Comp	**390 (5.2%)**	40 (2.2%)	6 (2.9%)	[redacted]	1 (25.0%)	2 (6.3%)	337 (6
Confirmation Status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Confirmed	**4,109 (4.3%)**	1,422 (4.1%)	500 (4.5%)	[redacted]	260 (3.8%)	38 (2.3%)	1,390 (4
	Removed	**1,753 (11.3%)**	644 (10.4%)	226 (12.0%)	[redacted]	236 (9.3%)	18 (5.6%)	338 (14
Appointment Lag	0	**3,281 (2.7%)**	671 (1.9%)	116 (2.2%)	[redacted]	54 (3.4%)	19 (1.2%)	2,327 (3
	1	**2,151 (6.1%)**	741 (4.8%)	189 (5.1%)	[redacted]	63 (6.1%)	12 (5.6%)	1,043 (8
	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	8-14	**5,809 (8.4%)**	1,652 (8.0%)	573 (8.2%)	[redacted]	292 (6.6%)	51 (5.9%)	2,741 (9
	15-30	**7,354 (8.2%)**	2,044 (7.9%)	662 (8.3%)	[redacted]	461 (6.8%)	82 (7.0%)	3,379 (8
	31+	**11,886 (10.0%)**	3,373 (8.2%)	877 (8.2%)	[redacted]	859 (7.4%)	90 (7.2%)	5,656 (13

Level 2: Building Models

Photo by Iulian Pana on Unsplash

JOHNS HOPKINS UNIVERSITY #1 Specialization

Data Science Specialization

Launch Your Career in Data Science. A nine-course introduction to data science, developed and taught by leading professors.

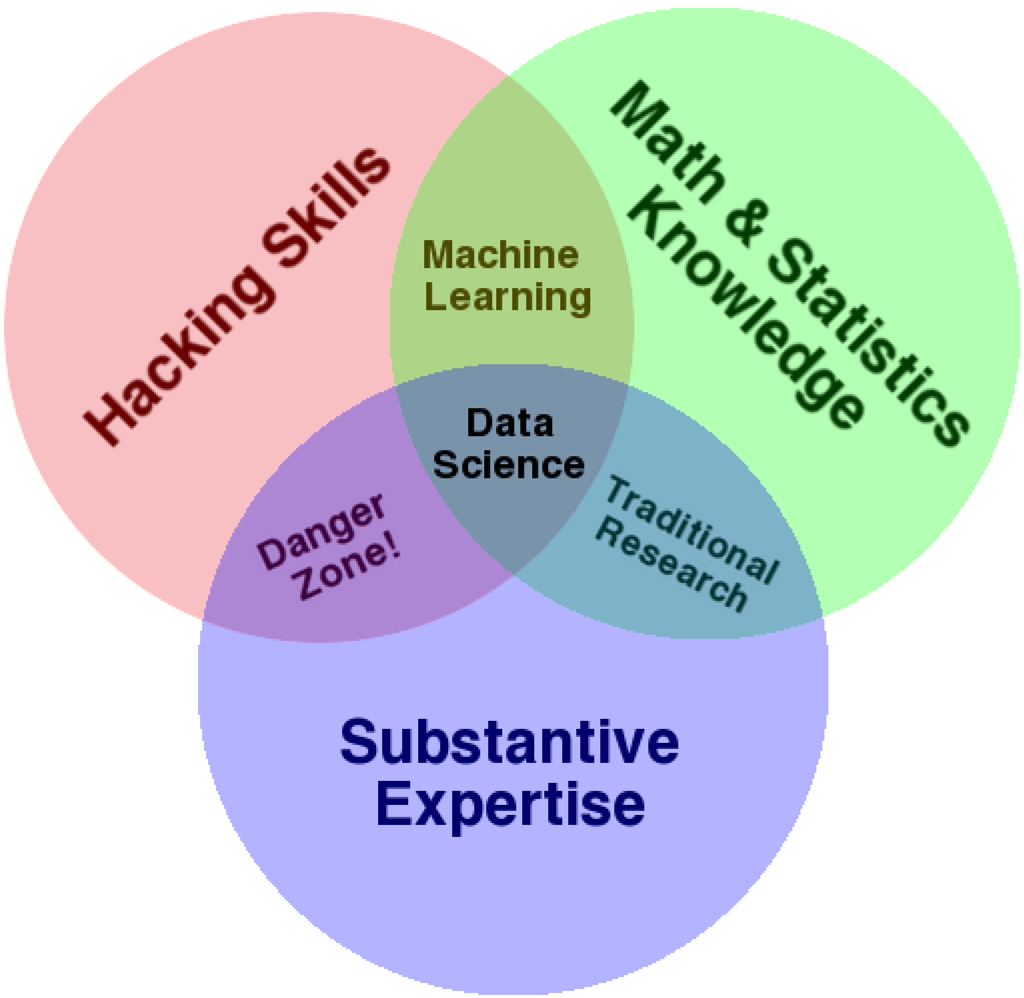
Enroll Financial aid available

About Data Science Specialization | Course 1 The Data Scientist's Toolbox | Course 2 R Programming | Course 3 Getting and Cleaning Data | Course 4 Explorat

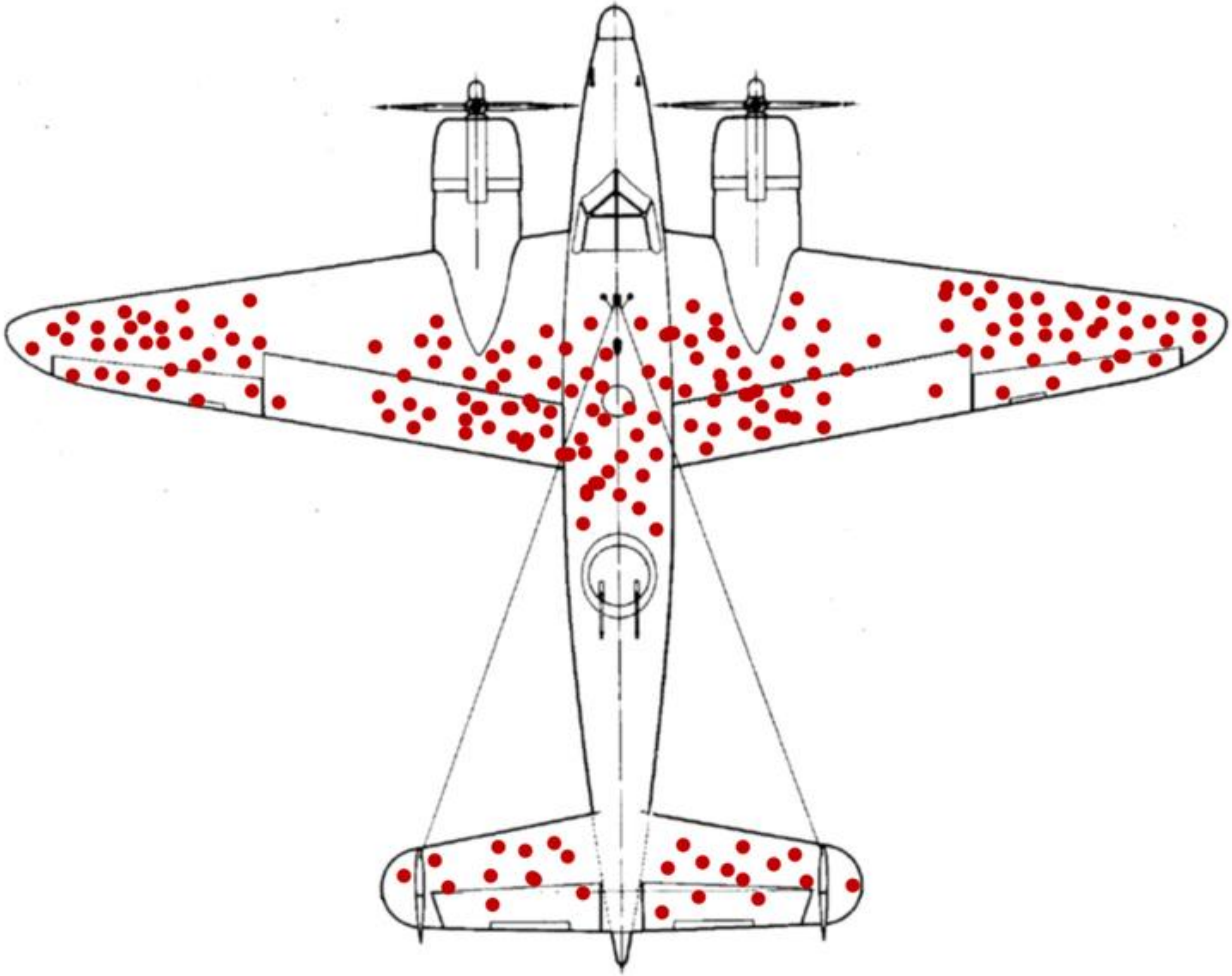
About this Specialization

Ask the right questions, manipulate data sets, and create visualizations to communicate results.

This Specialization covers the concepts and tools you'll need throughout the entire data science pipeline, from asking the right kinds of questions to making inferences and publishing results. In the final Capstone Project, you'll apply the skills learned by building a data product using real-world data. At completion, students will have a portfolio demonstrating their mastery of the material.



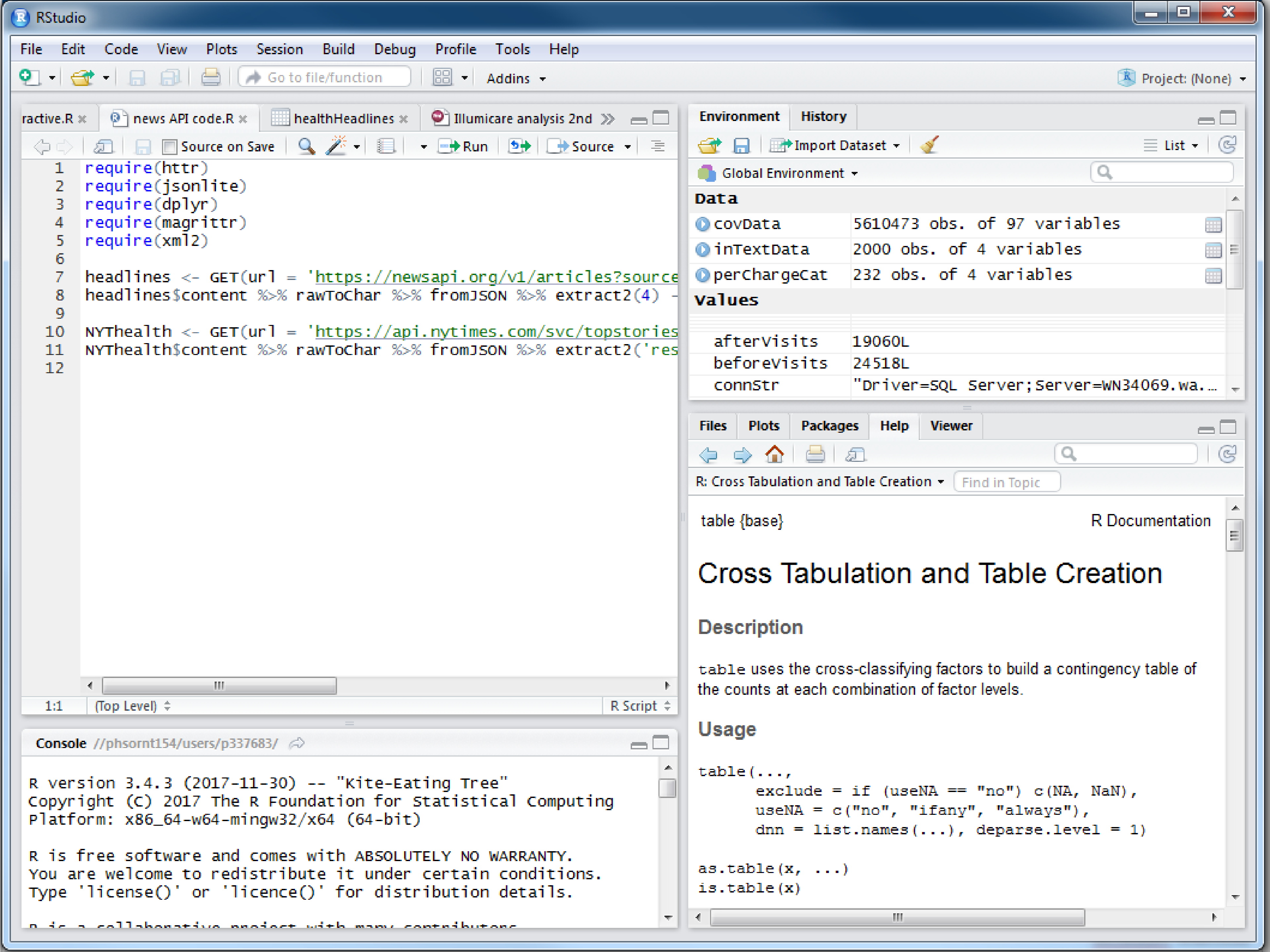
Drew Conway 2010

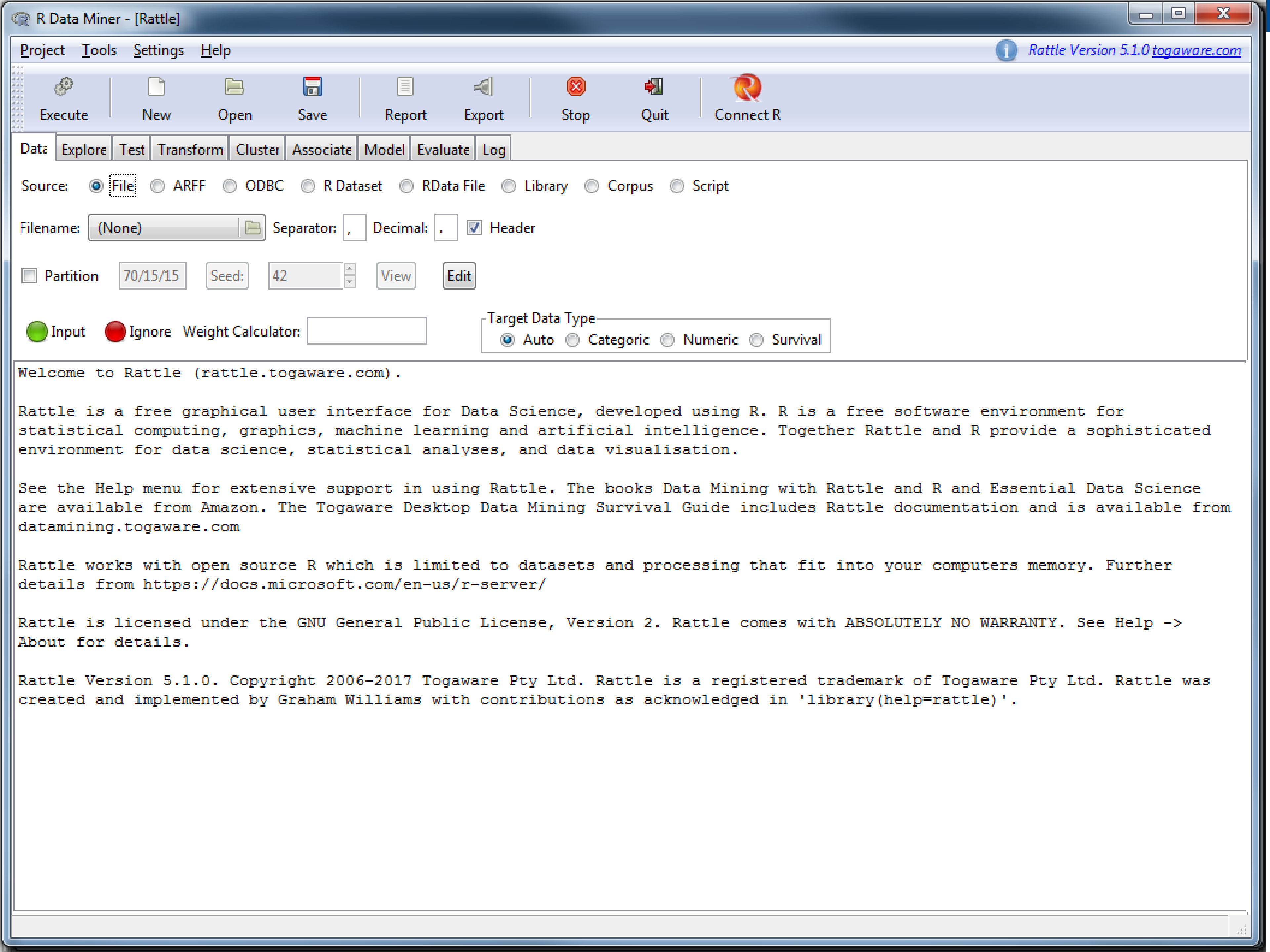


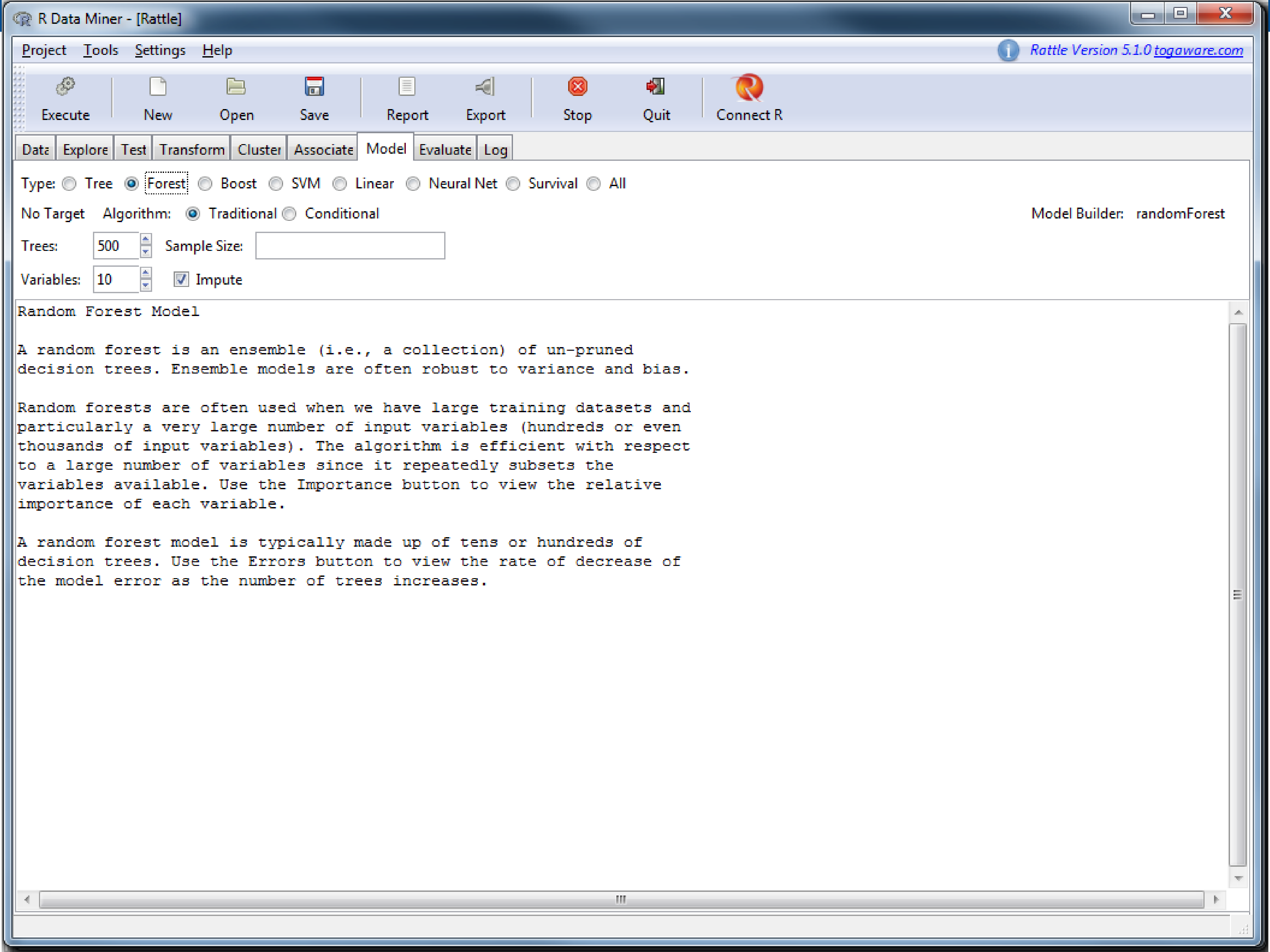
McGedden, Wikimedia

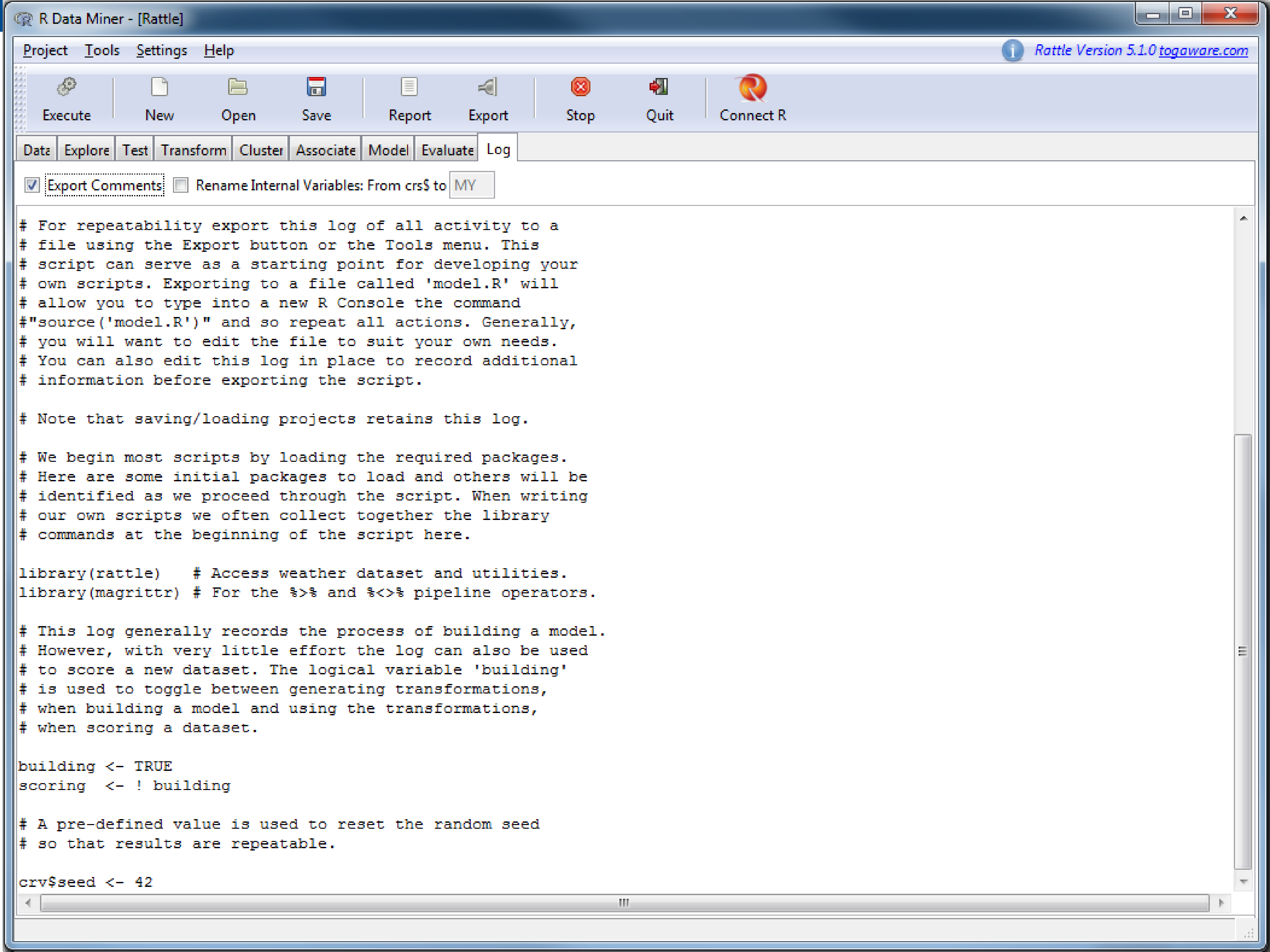
Computers are useless;
they can only give answers.

Pablo Picasso


RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
Project: (None)
ractive.R
news API code.R
healthHeadlines
Illumicare analysis 2nd
Source on Save
Run
Source
require(httr)
require(jsonlite)
require(dplyr)
require(magrittr)
require(xml2)
headlines <- GET(url = 'https://newsapi.org/v1/articles?source
headlines$content %>% rawToChar %>% fromJSON %>% extract2(4) -
NYThealth <- GET(url = 'https://api.nytimes.com/svc/topstories
NYThealth$content %>% rawToChar %>% fromJSON %>% extract2('res
1:1
(Top Level)
R Script
Console //phsornt154/users/p337683/
R version 3.4.3 (2017-11-30) -- "Kite-Eating Tree"
Copyright (C) 2017 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.
Environment
History
Import Dataset
List
Global Environment
Data
covData 5610473 obs. of 97 variables
inTextData 2000 obs. of 4 variables
perChargeCat 232 obs. of 4 variables
Values
afterVisits 19060L
beforeVisits 24518L
connStr "Driver=SQL Server;Server=WN34069.wa...
Files
Plots
Packages
Help
Viewer
R: Cross Tabulation and Table Creation
Find in Topic
table {base}
R Documentation
Cross Tabulation and Table Creation
Description
table uses the cross-classifying factors to build a contingency table of the counts at each combination of factor levels.
Usage
table(...,
exclude = if (useNA == "no") c(NA, NaN),
useNA = c("no", "ifany", "always"),
dnn = list.names(...), deparse.level = 1)
as.table(x, ...)
is.table(x)


R Data Miner - [Rattle]
Project Tools Settings Help
Rattle Version 5.1.0 togaware.com
Execute
New
Open
Save
Report
Export
Stop
Quit
Connect R
Data Explore Test Transform Cluster Associate Model Evaluate Log
Source: File ARFF ODBC R Dataset RData File Library Corpus Script
Filename: (None) Separator: , Decimal: . Header
Partition 70/15/15 Seed: 42 View Edit
Input Ignore Weight Calculator:
Target Data Type
Auto Categoric Numeric Survival
Welcome to Rattle (rattle.togaware.com).
Rattle is a free graphical user interface for Data Science, developed using R. R is a free software environment for statistical computing, graphics, machine learning and artificial intelligence. Together Rattle and R provide a sophisticated environment for data science, statistical analyses, and data visualisation.
See the Help menu for extensive support in using Rattle. The books Data Mining with Rattle and R and Essential Data Science are available from Amazon. The Togaware Desktop Data Mining Survival Guide includes Rattle documentation and is available from datamining.togaware.com
Rattle works with open source R which is limited to datasets and processing that fit into your computers memory. Further details from https://docs.microsoft.com/en-us/r-server/
Rattle is licensed under the GNU General Public License, Version 2. Rattle comes with ABSOLUTELY NO WARRANTY. See Help -> About for details.
Rattle Version 5.1.0. Copyright 2006-2017 Togaware Pty Ltd. Rattle is a registered trademark of Togaware Pty Ltd. Rattle was created and implemented by Graham Williams with contributions as acknowledged in 'library(help=rattle)'.


R Data Miner - [Rattle]
Project Tools Settings Help
Rattle Version 5.1.0 togaware.com
Execute
New
Open
Save
Report
Export
Stop
Quit
Connect R
Data Explore Test Transform Cluster Associate Model Evaluate Log
Type: Tree Forest Boost SVM Linear Neural Net Survival All
No Target Algorithm: Traditional Conditional
Model Builder: randomForest
Trees: 500
Sample Size:
Variables: 10
Impute
Random Forest Model
A random forest is an ensemble (i.e., a collection) of un-pruned decision trees. Ensemble models are often robust to variance and bias.
Random forests are often used when we have large training datasets and particularly a very large number of input variables (hundreds or even thousands of input variables). The algorithm is efficient with respect to a large number of variables since it repeatedly subsets the variables available. Use the Importance button to view the relative importance of each variable.
A random forest model is typically made up of tens or hundreds of decision trees. Use the Errors button to view the rate of decrease of the model error as the number of trees increases.



R Data Miner - [Rattle]
Project Tools Settings Help
Rattle Version 5.1.0 togaware.com
Execute New Open Save Report Export Stop Quit Connect R
Data Explore Test Transform Cluster Associate Model Evaluate Log
Export Comments
Rename Internal Variables: From crs$ to MY
For repeatability export this log of all activity to a
file using the Export button or the Tools menu. This
script can serve as a starting point for developing your
own scripts. Exporting to a file called 'model.R' will
allow you to type into a new R Console the command
#"source('model.R')" and so repeat all actions. Generally,
you will want to edit the file to suit your own needs.
You can also edit this log in place to record additional
information before exporting the script.
Note that saving/loading projects retains this log.
We begin most scripts by loading the required packages.
Here are some initial packages to load and others will be
identified as we proceed through the script. When writing
our own scripts we often collect together the library
commands at the beginning of the script here.
library(rattle)   # Access weather dataset and utilities.
library(magrittr) # For the %>% and %<>% pipeline operators.
# This log generally records the process of building a model.
# However, with very little effort the log can also be used
# to score a new dataset. The logical variable 'building'
# is used to toggle between generating transformations,
# when building a model and using the transformations,
# when scoring a dataset.
building <- TRUE
scoring <- ! building
# A pre-defined value is used to reset the random seed
# so that results are repeatable.
crv$seed <- 42

Level 3: Deploying to Users

Photo by Tom Ezzatkhah on Unsplash

Medford Pilot Reduces No-Shows by 49%

Tristan Markwell, 22 September 2015

Background

From early July through early September 2015, two Providence clinics (Doctors Clinic Internal Medicine and Medford Neurology) called patients at 6% or higher risk of no-show as determined by a Providence-wide machine learning model built by the Providence Data Science team. Lists were distributed each morning to staff at the clinics, who called patients up to three times to confirm the appointments. 3,235 calls were made over the course of the 8-week pilot. Results were compared to those for the same clinics in May and June, before the pilot had started.

Raw Results

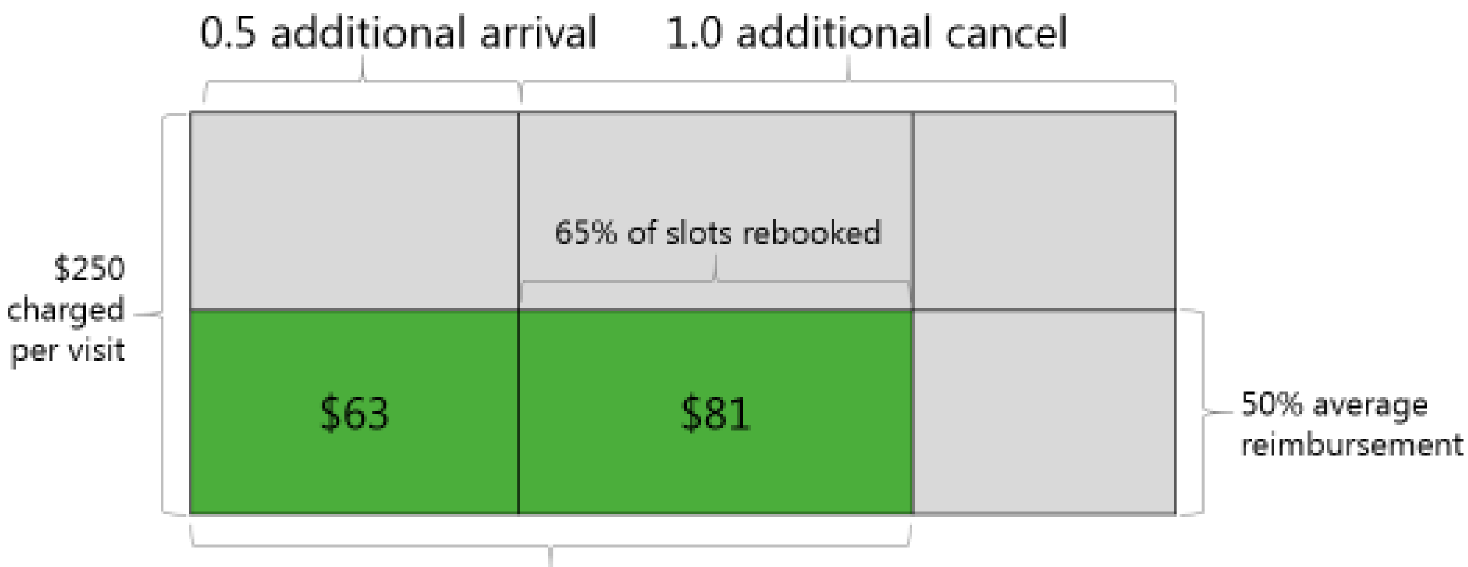
Usually the no-show rate is computed as the number of no-shows out of the number of appointments that weren't cancelled (that is, arrived and no-show). By this measure, **no-shows went down sharply – 8.6% for May-June but only 5.6% for the pilot, a 35% drop**. Considering cancellations within two days in both denominators, the rates are 7.6% for May-June and 4.7% for the pilot, a drop of 38%.

Calibration

terms of their likelihood to no-show. Indeed, the average no-show risk for the May-June patients was 10.2%, but 12.2% for the pilot group. That means that **the number of no-shows went down, even in an environment where we would have expected them to increase by 20%**. 46% of each group was classified as high-risk, but even with the reduction in no-shows, the high-risk group accounts for 77% of the no-shows for May-June and 75% of the no-shows for the pilot period. One way to better account for this mismatch between the pilot and control group is to calibrate the two time periods for specified bins of risk. Because no-shows are relatively rare events, for the bins to be similar in size they are much narrower at the low end of risk and wider at the high end. The cut points selected were 0%, 2%, 3%, 4%, 6%, 8%, 10%, 15%, 20%, 30%, 40%, and 100%. Comparing the two groups within each of these bins, the effects of the calls for patients at a given level of risk are clearer.

From Figure 1, it is clear that the group that was called had a much lower no-show rate (that is, the intervention pretty clearly bends the curve), but rather than those patients coming in it appears that most are being

Per hour of calls

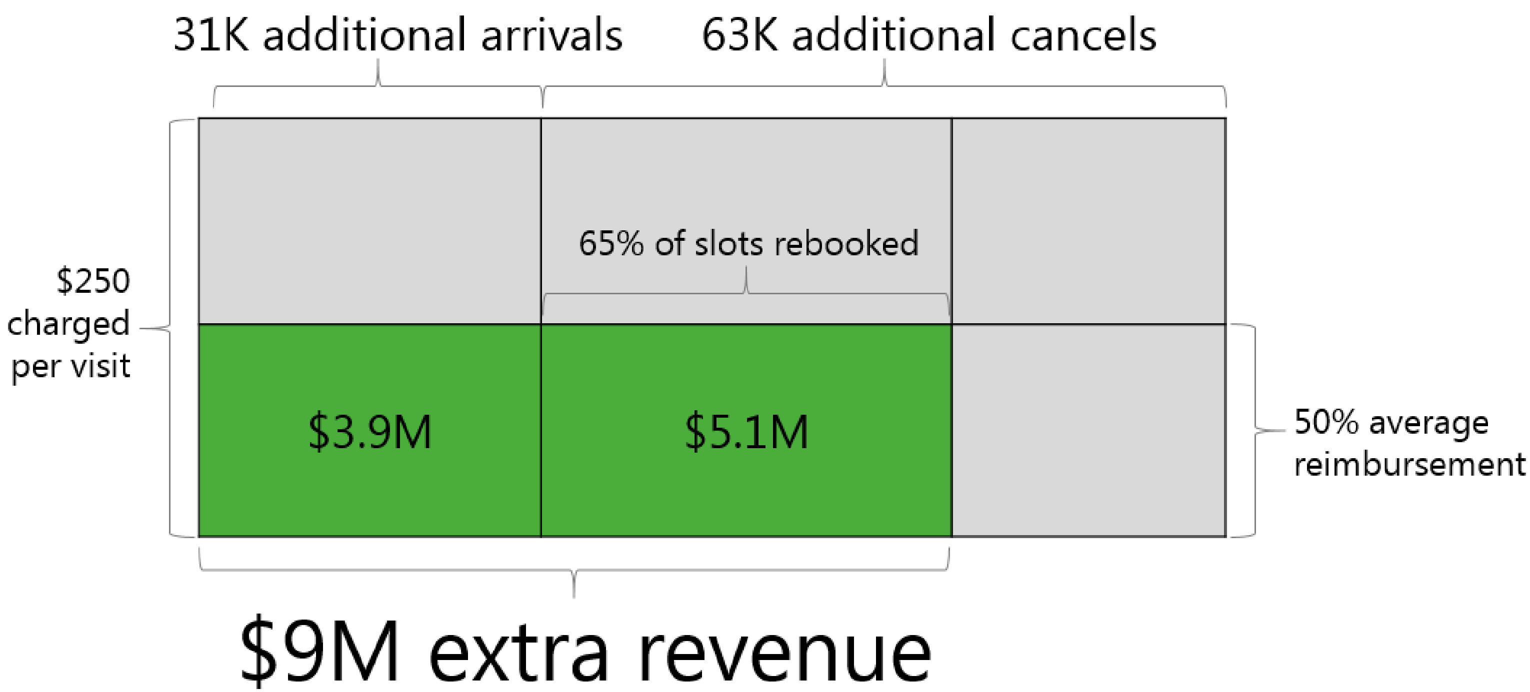


$144 extra revenue

Calling Patients at Risk for No-Show:

Overall Providence Per Year

60-65K hours of calls (31 FTE, ≈$1.5M wages and benefits?)



Two days out patients most likely to no-show

Minimum Risk

10%

Market(s)

OR Portland

Department

PMG GLISAN FAMILY ...

Department	Patient Name	Phone 1	Phone 2	Provider	Appointment Time	No-show Risk
PMG GLISAN FAMILY MEDICINE	[illegible]	[illegible]	[illegible]	ROSENFELD, SHAI DAN	3/22/2017 15:00	38%
	[illegible]	[illegible]	[illegible]	WHITE, CHARLES THAYER	3/22/2017 13:40	36%
	[illegible]	[illegible]	[illegible]	ESKESEN, STACI JOANNA	3/22/2017 18:20	31%
	[illegible]	[illegible]	[illegible]	GLISAN MA	3/22/2017 15:40	29%
	[illegible]	[illegible]	[illegible]	RODGERS-ROBEY, MELISSA D..	3/22/2017 13:20	28%
	[illegible]	[illegible]	[illegible]	ROSENFELD, SHAI DAN	3/22/2017 15:20	28%
	[illegible]	[illegible]	[illegible]	WHITE, CHARLES THAYER	3/22/2017 14:20	27%
	[illegible]	[illegible]	[illegible]	GLISAN FM RN	3/22/2017 13:00	24%
	[illegible]	[illegible]	[illegible]	RODGERS-ROBEY, MELISSA D..	3/22/2017 14:40	21%
	[illegible]	[illegible]	[illegible]	WHITE, CHARLES THAYER	3/22/2017 11:20	21%
	[illegible]	[illegible]	[illegible]	GRAVES, RACHEL SINEX	3/22/2017 11:00	18%
	[illegible]	[illegible]	[illegible]	ROSENFELD, SHAI DAN	3/22/2017 10:40	17%
	[illegible]	[illegible]	[illegible]	ESKESEN, STACI JOANNA	3/22/2017 17:40	16%
	[illegible]	[illegible]	[illegible]	RODGERS-ROBEY, MELISSA D..	3/22/2017 17:00	15%
	[illegible]	[illegible]	[illegible]	WHITE, CHARLES THAYER	3/22/2017 13:00	15%
	[illegible]	[illegible]	[illegible]	GRAVES, RACHEL SINEX	3/22/2017 09:00	15%
	[illegible]	[illegible]	[illegible]	RODGERS-ROBEY, MELISSA D..	3/22/2017 15:20	14%
	[illegible]	[illegible]	[illegible]	WHITE, CHARLES THAYER	3/22/2017 10:00	13%
	[illegible]	[illegible]	[illegible]	ESKESEN, STACI JOANNA	3/22/2017 17:20	12%
	[illegible]	[illegible]	[illegible]	WHITE, CHARLES THAYER	3/22/2017 15:00	12%
	[illegible]	[illegible]	[illegible]	WHITE, CHARLES THAYER	3/22/2017 08:20	12%
	[illegible]	[illegible]	[illegible]	ESKESEN, STACI JOANNA	3/22/2017 15:20	12%
	[illegible]	[illegible]	[illegible]	ROSENFELD, SHAI DAN	3/22/2017 11:40	11%
	[illegible]	[illegible]	[illegible]	WHITE, CHARLES THAYER	3/22/2017 15:40	11%
	[illegible]	[illegible]	[illegible]	ESKESEN, STACI JOANNA	3/22/2017 13:00	10%

Level 4: App Development

Photo by Tim Gouw on Unsplash

Providence St. Joseph Health

IBM

Vijay Venkatesan
PSJH Chief Data Officer

myHIway

Search...

Search

At a glance | Search | Galleries | About

Logout

myHIway

My Saved Searches

Click the heart in the Search Box, and save it with a Search name to see it here...

My Bookmarks

Bookmark a search result to see it here ...

PROVIDENCE Health & Services
SWEDISH
KADLEC
pacific medical centers
Facey MEDICAL GROUP
St.JosephHealth
hoag
CovenantHealth
UNIVERSITY OF PROVIDENCE
INHS

myHIway

breast cancer

Search

At a glance | Search | Galleries | About | Logout

Showing 1 - 50 of 486 (372)ms

Sort by relevance

breast cancer ×

All 486 | APIs 0 | Apps 0 | Data 2 | Metrics 58 | Reports 425 | Terms 1 | Training

TOPICS

- Business Operations 7
- Clinical Performance Groups 4
- Institutes 9
- Quality 2

ENVIRONMENT

- PHS Epic Hyperspace 204
- Tableau 78
- SHS Epic Hyperspace 52
- Providence InfoView 44
- Population Health 32

Show more

BUSINESS ENTITY

- Providence Health & Services 240
- Swedish 59
- Kadlec 38

PRODUCT TYPE

- Reporting Workbench Report 283
- Crystal Report 62
- Tableau View 49

Breast Cancer Workbook (Reports)
https://tableauserver.providence.org/#/workbooks/34265
System-wide Breast Cancer Volume, Treatment, and Quality Indicators
More

Vantage

Breast Cancer Screening (Metrics)
This measure is available in the report(s) listed in the "Measure", "Measure Alias", "Measure(s)" or "You Have Selected (one of three dropdowns)" dropdown menus.
Definition: ...
More

Metric

Breast Cancer Screening (Metrics)
Breast cancer screening is checking for cancer before there are signs or symptoms of the disease. All women need to be informed by their health care provider about the best screening options for them...
More

Vantage

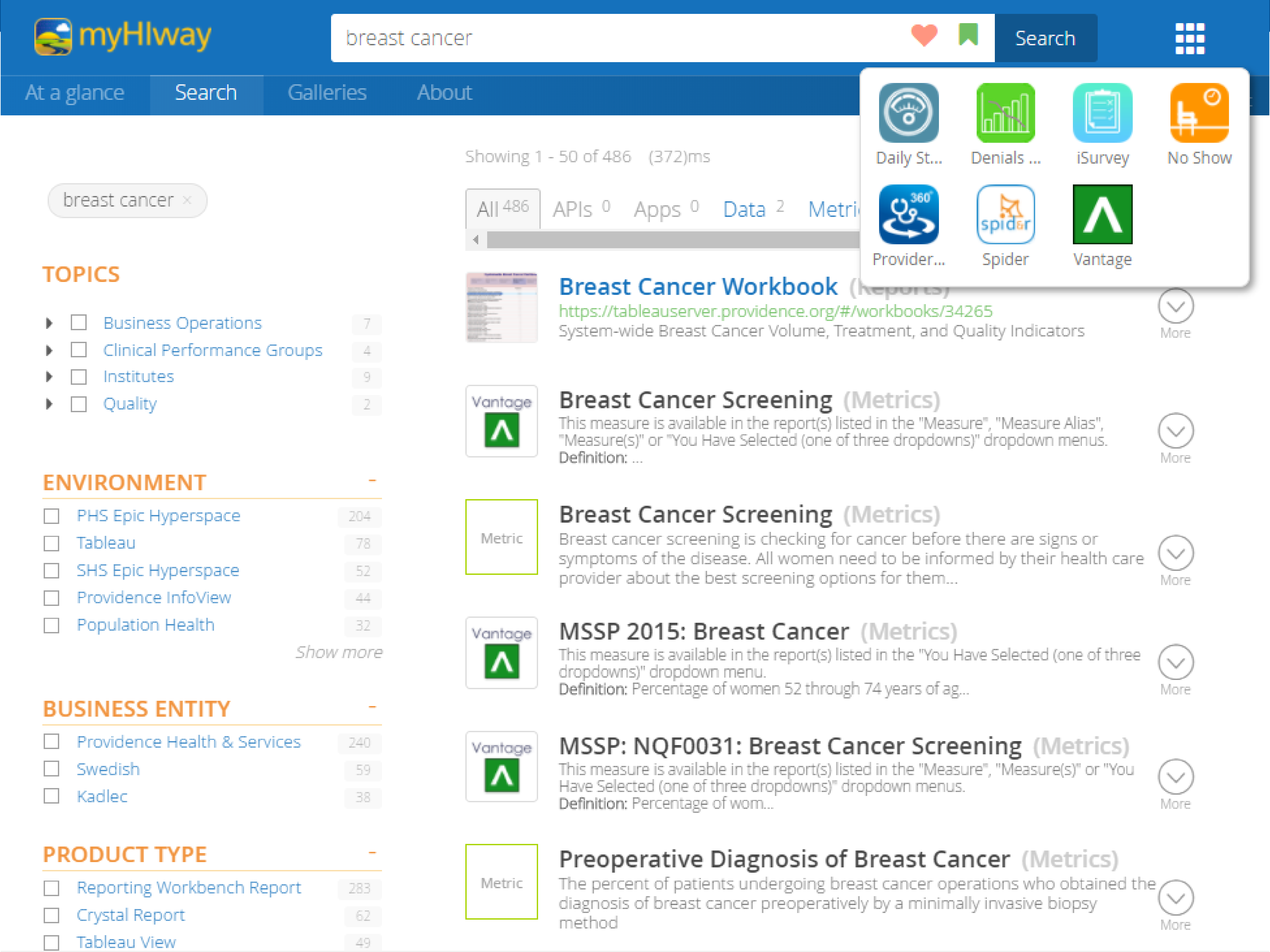
MSSP 2015: Breast Cancer (Metrics)
This measure is available in the report(s) listed in the "You Have Selected (one of three dropdowns)" dropdown menu.
Definition: Percentage of women 52 through 74 years of ag...
More

Vantage

MSSP: NQF0031: Breast Cancer Screening (Metrics)
This measure is available in the report(s) listed in the "Measure", "Measure(s)" or "You Have Selected (one of three dropdowns)" dropdown menus.
Definition: Percentage of wom...
More

Metric

Preoperative Diagnosis of Breast Cancer (Metrics)
The percent of patients undergoing breast cancer operations who obtained the diagnosis of breast cancer preoperatively by a minimally invasive biopsy method
More


myHIway
breast cancer
Search
At a glance
Search
Galleries
About
Daily St...
Denials ...
iSurvey
No Show
Provider...
Spider
Vantage
Showing 1 - 50 of 486 (372)ms
breast cancer ×
All 486
APIs 0
Apps 0
Data 2
Metri
TOPICS
Business Operations 7
Clinical Performance Groups 4
Institutes 9
Quality 2
ENVIRONMENT
PHS Epic Hyperspace 204
Tableau 78
SHS Epic Hyperspace 52
Providence InfoView 44
Population Health 32
Show more
BUSINESS ENTITY
Providence Health & Services 240
Swedish 59
Kadlec 38
PRODUCT TYPE
Reporting Workbench Report 283
Crystal Report 62
Tableau View 49
Breast Cancer Workbook (Reports)
https://tableauserver.providence.org/#/workbooks/34265
System-wide Breast Cancer Volume, Treatment, and Quality Indicators
More
Vantage
Breast Cancer Screening (Metrics)
This measure is available in the report(s) listed in the "Measure", "Measure Alias", "Measure(s)" or "You Have Selected (one of three dropdowns)" dropdown menus.
Definition: ...
More
Metric
Breast Cancer Screening (Metrics)
Breast cancer screening is checking for cancer before there are signs or symptoms of the disease. All women need to be informed by their health care provider about the best screening options for them...
More
Vantage
MSSP 2015: Breast Cancer (Metrics)
This measure is available in the report(s) listed in the "You Have Selected (one of three dropdowns)" dropdown menu.
Definition: Percentage of women 52 through 74 years of ag...
More
Vantage
MSSP: NQF0031: Breast Cancer Screening (Metrics)
This measure is available in the report(s) listed in the "Measure", "Measure(s)" or "You Have Selected (one of three dropdowns)" dropdown menus.
Definition: Percentage of wom...
More
Metric
Preoperative Diagnosis of Breast Cancer (Metrics)
The percent of patients undergoing breast cancer operations who obtained the diagnosis of breast cancer preoperatively by a minimally invasive biopsy method
More

No Show

Data Refreshed: Tue 04/17/2018 6:06 AM

Cancelled appointments removed by MultiScale

KADLEC, ... | KC RICHLAND PRIMARY CARE | 2 Days Out (04/19/2018)

Sort by: No Show Risk (high to low)

Include all appointments with no show risk at least: 10%

☑ Hide confirmed/cancelled ☐ Exclude EPIC confirmation

Kobe Vandeberg — No Show 69% / Show 31%

MRN: **22700251**
DOB: **03/09/1996**
Provider: **HUYNH, TONY HUU**
Appt Type: **OB FIRST**
Appt Time: **04/19/2018 01:30 PM**

(478) 837-1226
(803) 899-1994

Pris Lasky — No Show 55% / Show 45%

MRN: **22461126**
DOB: **12/14/1996**
Provider: **SCHEVE, DAWN A**
Appt Type: **PULM FOLLOW UP**
Appt Time: **04/19/2018 09:00 AM**

(561) 712-2312
(650) 524-9323

Sheron Gotcher — No Show 45% / Show 55%

MRN: **91926828**

(281) 199-2016

+

a myHIway app

No Show

Cancelled appointments removed by MultiScale

Data Refreshed: Tue 04/17/2018 6:06 AM

KADLEC, ...

KC RICHLAND PRIMARY CARE

2 Days Out (04/19/2018)

Sort by:

No Show Risk (high to low)

Include all appointments with no show risk at least:

10%

☑ Hide confirmed/cancelled

☐ Exclude EPIC confirmation

Kobe Vandeberg

No Show 69%

Show 31%

MRN: **22700251**
DOB: **03/09/1996**
Provider: **HUYNH, TONY HUU**
Appt Type: **OB FIRST**
Appt Time: **04/19/2018 01:30 PM**

(478) 837-1226

(803) 899-1994

Call in progress: (478) 837-1226

Confirmed Patient

Cancelled Patient

Rescheduled Patient

Left Msg for Patient

Other (enter Note)

Confirmed Patient.

Done

Cancel

+

a myHIway app

No Show Value Estimate

Tristan Markwell

January 17, 2018

Executive Summary

Use of the No Show app by a few clinics in Oregon, fed by a predictive algorithm from the Data Science team, had a net positive impact of about $8K in December 2017 (annual rate of $96K), with a net value of nearly $5 per call. If scaled to all PHS, SHS and KHS, this equates to $8.5M annually due to reduced no shows and earlier cancels. The app already contains all these clinics, so implementation would be purely organizational.

Situation

The No Show app has been in production since March 2017. Various clinics in multiple regions have tried it in different iterations, but its use has never been required, and at present almost all use is concentrated in the clinics in Oregon, which has a history of and central support for initiatives supporting improved access, and which held optional trainings in 2017 for all the clinics.

Background

Healthcare Intelligence first took up the question of predicting clinic no shows using machine learning in 2013. Over time this has evolved from a technical demonstration, to a 2015 pilot in Southern Oregon, to a Tableau dashboard, and finally to a myHIway app that automatically tracks how users are interacting with it. The Medford pilot was used a two-month test period compared to the two months prior, and estimated that an hour of calls could generate $144 in additional revenue through reduced no-show rates and earlier cancels (enough time to rebook the slot); this scaled to about $9M per year for all of PHS, for something like $1.5M in wages for the callers. A more formal study of the effect of the calls on no show was undertaken in early 2017 and didn't replicate this finding; however, analysis showed that patients in the treatment and control arms were contacted at almost exactly the same rate (ie massive non-compliance), so there was no way to tease out the effect. Looking at the patients naturally contacted by clinics using the myHIway app is a way to estimate the effect of the calls in the wild, and a proper method of finding appropriate cases to match against should help mitigate issues with using a convenience sample.

Analysis

Exclusions

Since the app has gained traction in the clinics in the Oregon Region, the analysis is restricted to clinics and hospital clinics in the various Oregon markets in December 2017. Of the 107,068 encounters from that time period that were scored (the system automatically ignores new patient appointment types and weekend appointments), 87,314 (82%) had an associated two-business-day prediction, which is what the app would use to


Photo by Inês Pimentel on Unsplash

Level 5: Operationalizing

Photo by Inma Ibáñez on Unsplash

Providence St. Joseph Health

Data Science Infrastructure POC

SAP HANA

R Studio

Microsoft SQL Server

R

cloudera

Apache Spark

Key Lessons from the POC

- Need for an integrated team to develop shared vocabulary
  - Architects
  - System administrators
- Data Science team needs expansive database privileges
- Open-source without vendor solutions don't scale well with complexity
- Too easy for people to get pulled off for production support

tristan.markwell@providence.org
lindsay.mico@providence.org

Refreshment Break in Foyer 3:00 – 3:20

THANK YOU to our Conference Sponsors

Bloodworks Northwest

TRG clinical informatics advisors

Nordisk Systems, Inc.
Technology Focused, Committed Success, Visionary Results

IVOXY CONSULTING

rubrik

CDW·G PEOPLE WHO GET IT

NUANCE

NUTANIX