

The Making of a Real-World Moneyball

Finding Undervalued Players with H₂O, LIME and Shiny



Jo-fai (Joe) Chow

Data Science Evangelist at H2O.ai

joe@h2o.ai / @matlabulous

+

David Kearns

Product Manager at IBM

KEARNSD@ie.ibm.com / @DaithiOCiaran

Dublin Artificial Intelligence & Deep Learning
H₂O.ai + Zalando Joint Meetup

About Me



Last Week – Paris



- **Before H₂O**

- Water Engineer / EngD Researcher / Matlab Fan Boy
(wonder why @matlabulous?)
- Discovered R, Python, H₂O ... never look back again
- Data Scientist at Virgin Media (UK), Domino Data Lab (US)

- **At H₂O ...**

- Data Scientist / Evangelist /
- Sales Engineer / Solution Architect /
- Community Manager
... The harsh reality of startup life ...

Reminder: #360Selfie

H₂O.ai

H2O.ai Overview

Company	Founded in Silicon Valley in 2012 Funded: \$75M Investors: Wells Fargo, NVIDIA, Nexus Ventures, Paxion Ventures
Products	<ul style="list-style-type: none">• H2O Open Source Machine Learning (14,000 organizations)• H2O Driverless AI – Automatic Machine Learning
Leadership	Leader in Gartner MQ Machine Learning and Data Science Platform
Team	120 AI experts (Kaggle Grandmasters, Distributed Computing, Visualization)
Global	Mountain View, London, Prague, India



Worldwide Recognition in the H2O.ai Community

Open source
community

222 OF FORTUNE
THE 500
 H₂O

8 OF TOP 10
BANKS

7 OF TOP 10
INSURANCE COMPANIES

4 OF TOP 10
HEALTHCARE COMPANIES

Paying Customers



"H2O.ai's reference customers gave it the highest overall score for sales relationship and overall service and support" - Gartner MQ 2018

H₂O.ai

H2O.ai is a **Leader** in the 2018 Gartner Data Science and Machine Learning Platforms Magic Quadrant

- Technology leader with most completeness of vision
- Recognized for the mindshare, partner network and status as a **quasi-industry standard** for machine learning and AI
- H2O.ai customers gave the highest overall score among all the vendors for sales relationship and account management, customer support (onboarding, troubleshooting, etc.) and overall service and support

Figure 1. Magic Quadrant for Data Science and Machine-Learning Platforms



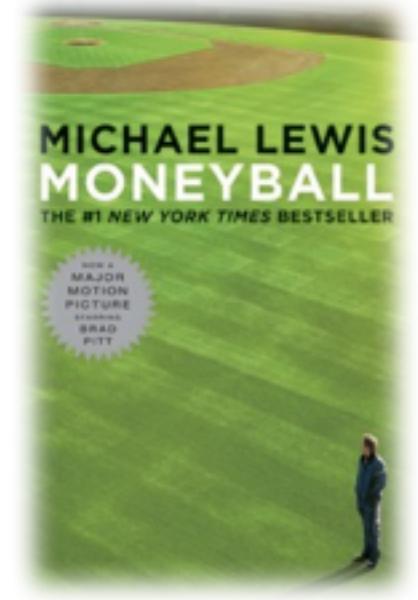
Get the
Gartner
Magic
Quadrant
[here](#)

Moneyball: The Multimillion-Dollar Business Problem

The quest to find the most undervalued players
(before other teams notice them)



Source: Moneyball, 2011 Columbia Pictures



The Real Business Problem in Major League Baseball (MLB)

- Existing Forecasts (e.g. ESPN) are usually projections for the **next year only**.
- MLB players usually consider terms for 3 to 5 years when they sign a new contract.
- MLB teams need to consider players' **long-term performance** (i.e. > 1 year).

The screenshot shows the ESPN Fantasy Baseball interface. At the top, there's a navigation bar with the ESPN logo, NFL, NBA, MLB, NCAAF, Soccer, and other links. Below that is a sub-navigation for Fantasy Baseball with links to Home, Top 300 Rankings, Forecaster: Hitting Matchups, and More. The main content area is titled "Sortable 2018 Projections" and "Position: Batters". It includes filters for By Name, Go, Team: All, and View: 2018 Season. A large orange arrow points from the text "Existing Forecasts (e.g. ESPN) are usually projections for the next year only." to this section. Below the filters is a table titled "PLAYERS" with columns for RANK, PLAYER, TEAM POS, and various statistics like R, HR, RBI, SB, and AVG. A second orange arrow points from the text "MLB teams need to consider players' long-term performance (i.e. > 1 year)." to the rightmost column of the table, which is labeled "2018 SEASON BATTING PROJECTIONS". A blue banner at the bottom of the table area reads "2018 SEASON BATTING PROJECTIONS".

RNK	PLAYER, TEAM POS	2018 SEASON BATTING PROJECTIONS				
		R	HR	RBI	SB	AVG
1	Mike Trout, LAA OF	119	40	98	22	.308
2	Jose Altuve, Hou 2B	106	24	83	32	.329
3	Nolan Arenado, Col 3B	105	38	132	3	.300
4	Mookie Betts, Bos OF	107	24	84	29	.294
5	Bryce Harper, Wash OF	109	35	102	12	.309
6	Trea Turner, Wash SS	97	15	59	57	.287
7	Charlie Blackmon, Col OF	116	30	84	14	.315
8	Paul Goldschmidt, Ari 1B	102	28	102	19	.296
9	Carlos Correa, Hou SS	99	28	107	12	.301
10	Giancarlo Stanton, NYY OF, DH	107	52	118	2	.269
11	Kris Bryant, Chi 3B	110	32	94	10	.296
12	Manny Machado, Bal 3B, SS	97	34	98	10	.294

The Moneyball Team



IBM

David Kearns
PM @ IBM Data Science

A portrait of David Kearns, a man with dark hair and a beard, wearing a pink striped shirt. He is standing in front of a wooden wall with the letters "TH" and "NC" visible. The IBM logo is in the bottom left corner of the photo frame.

H₂O

Jo-Fai Chow
Data Scientist @ H₂O.ai

A portrait of Jo-Fai Chow, a man with short dark hair, wearing a red and white checkered shirt. He is standing in front of a blue and white graphic background featuring stylized buildings and data points. The H₂O.ai logo is in the bottom left corner of the photo frame.

Aginity

Ari Kaplan
Mr. Moneyball @ Aginity

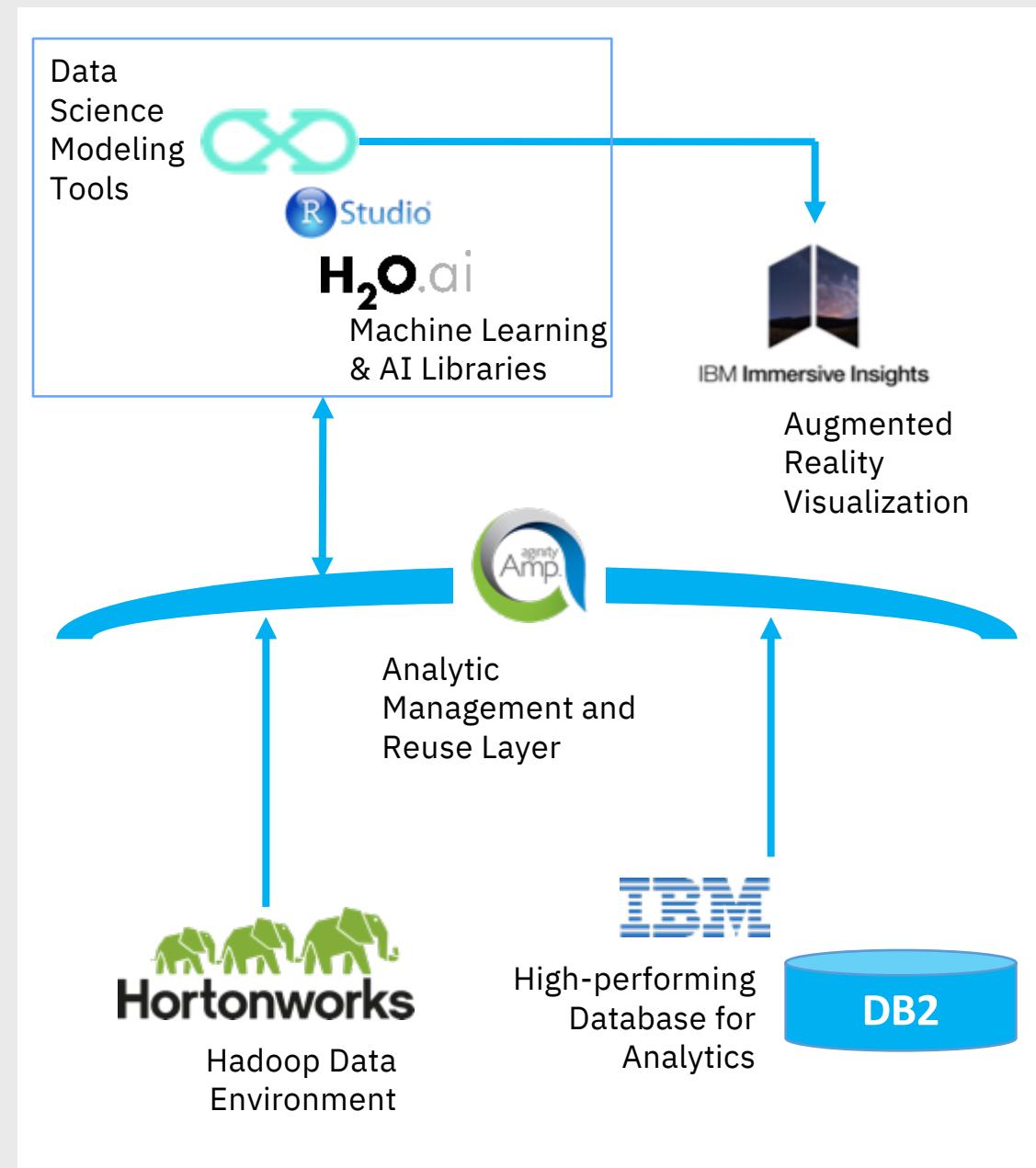
A portrait of Ari Kaplan, a bald man with glasses, wearing a dark suit and a yellow tie. He is smiling at the camera. The Aginity logo is in the bottom left corner of the photo frame.

IBM + Aginity + H₂O.ai

Enterprise Solution

The Workflow

1. Data loaded into the databases
2. Connected diverse data sources to Amp
3. Amp used to create derived attributes and publish them and data to DSX and H₂O
4. DSX and H₂O to build and tweak statistical and machine learning models
5. Visualizations tested in Immersive Insights
6. Steps 4 and 5 repeated to get settled data
7. Statistical and machine learning models saved in Amp
8. Data exported to Immersive Insights for final visualizations



In case you're wondering ... final project result

led to the signing of a
Major League Baseball (MLB) player

\$20M

multi-year contract

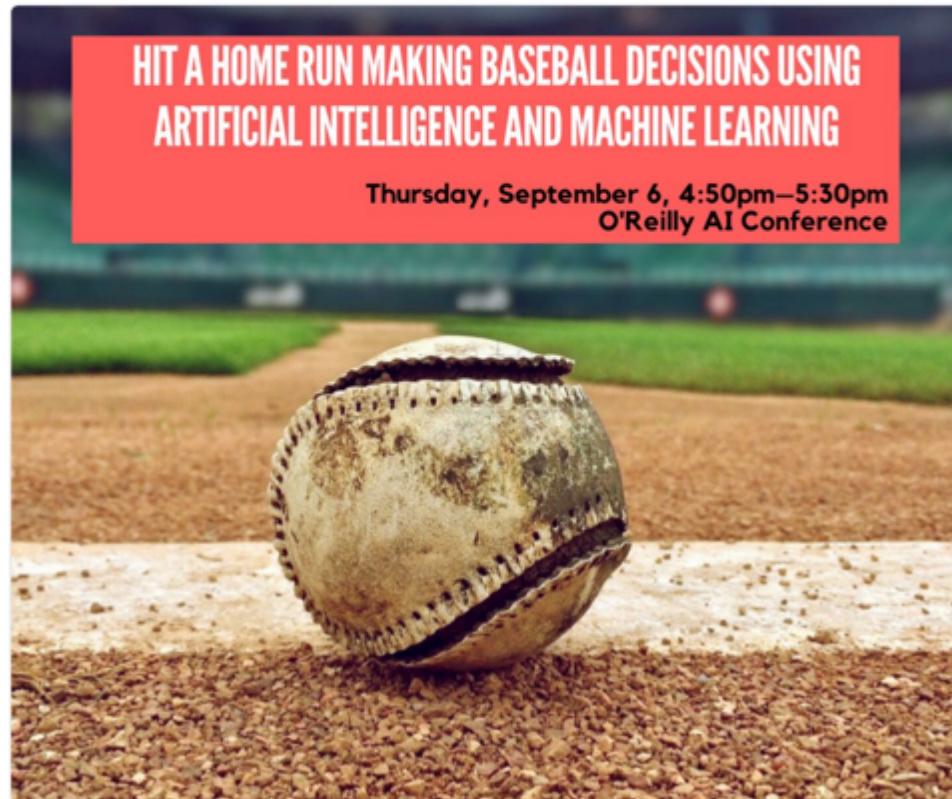
finalised two weeks
before the regular season





H2O.ai
@h2oai

Attending O'Reilly #AI Conference? Learn how to hit a home run using #artificialintelligence and #machinelearning from @Ledell and former MLB Moneyball analyst @arikaplan1: conferences.oreilly.com/artificial-int ...



8:39 PM - 24 Aug 2018

Following ▾



David Kearns
@DaithiOCiaran

Following ▾

Today is the day #TheAIConf. Want to hear how #MachineLearning can be applied to #Moneyball. Join myself, @arikaplan1 @chriscoad and @ledell @IBMDatascience @h2oai @Aginity 4.50pm Location: Imperial A



3:29 PM - 6 Sep 2018



Jo-fai (Joe) Chow

@matlabulous



Oh man, I am missing out. BIG TIME!!!



David Kearns @DaithiOCiaran

Today is the day #TheAIConf. Want to hear how #MachineLearning can be applied to #Moneyball. Join myself, @arikaplan1 @chriscoad and @ledell @IBMDatascience @h2oai @Aginity 4.50pm Location: Imperial A

3:35 PM - 6 Sep 2018

3 Likes



1



3



Add another Tweet



David Kearns @DaithiOCiaran · Sep 6



Replying to @matlabulous

I told the guys this pic would hurt ya. You are sorely missed.

Framing the Business Problem for Machine Learning

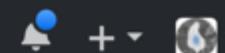
Code on GitHub (without Ari's proprietary data)

<https://github.com/woobe/moneyball>



Search or jump to...

Pull requests Issues Marketplace Explore



woobe / moneyball

Unwatch ▾ 2

Star 1

Fork 0

Code

Issues 0

Pull requests 0

Projects 0

Wiki

Insights

Settings

Moneyball Demo (Public Version)

Add topics

7 commits

More Info → [github.com/woobe/
moneyball](https://github.com/woobe/moneyball)

Apache-2.0

Branch: master ▾

New pull request

Create new file

Upload files

Find file

Clone or download ▾

woobe Added descriptions

Latest commit d630812 2 days ago

cache_data

Raw data from Lahman database

3 days ago

.gitignore

Initial commit

22 days ago

LICENSE

Initial commit

22 days ago

README.md

Added descriptions

2 days ago

step_1_data_munging.R

Data munging for Lahman data only

3 days ago

step_2_model_pitching.R

H2O AutoML Model Building Scripts

2 days ago

step_3_model_batting.R

H2O AutoML Model Building Scripts

2 days ago

README.md

Baseball Player Performance Data

- Open data – **Lahman** Database.
- Proprietary data (**AriDB**) from Ari Kaplan – our real Moneyball guy.
- Enriched Lahman data with Ari's Data – Final dataset for predictive modelling



Predictive Modelling – H₂O AutoML

- Framed data as regression problems for performance prediction.
- Historical player performance as features.
- Used H₂O AutoML to build ensembles (linear model, random forests, gradient boosting, and deep neural networks).



```
# Install 'h2o' from CRAN  
install.packages('h2o')
```

Lahman Database

<http://www.seanlahman.com/baseball-archive/statistics/>

Attribute	Description
playerID	Player ID code
yearID	Year player was born
G	Games
AB	At Bats
R	Runs
H	Hits
2B	Doubles
3B	Triples
HR	Homeruns
SO	Strike Outs
IBB	Intentional Walks
SF	Sacrifice flies

Ari's Database

- Private database containing 5 years of data
- Pitch-by-pitch play for each MLB game:
 - Pitch type, top speed, end speed, spin rate, x, y, z coordinates, batter result etc.

Attribute	Description
Pitch_Type	Two - character code of type of pitch. FF=fastball, CU=curveball, SL=slider, etc.
Spin_rate	Spin of the pitch in rotations per minute. One of the top fields for a feature...the theory is the more spin the harder it is to hit.
Start_speed	The velocity of the pitch in mph (when it leaves the hand, which is the measure used for tv).
End_speed	The velocity of the pitch when it arrives at the plate
Z0	Feet off the ground when the pitch is released.
Spray_x	When ball is hit into play, this is the x - coordinate of where it is hit/picked up by a fielder
Spray_y	When ball is hit into play, this is the y - coordinate of where it is hit/picked up by a fielder
Spray_des	Classification of type of hit: pop out, flyout, groundout, hit, error

Creating Consistent and Reusable Analytic Assets Managed by Aginity Amp

The screenshot shows the Aginity Amp application interface. On the left is a sidebar with navigation links: Home, New, Browse, Jobs, and Settings. The main workspace title is "Money Ball". A search bar is present. The main content area shows a file structure under "AriDB2012_batter_derived_attributes_sa". The "Select Analytic" path is set to "/AriDB2012_batter_derived_attributes/aridb2012_batter_derived_output". The "Frame Parameters" section contains a table with columns: NAME, DATATYPE, PROMPT FOR INPUT?, and VALUES. Below this is a preview tabular view with columns: SCHEMA, RELATIONSHIPS, PREVIEW, and Test. The preview table lists various derived attributes. On the right, the "OUTPUT SCHEMA" pane lists all the derived attributes.

Main Workspace: Current Workspace

Home

New

Browse

Jobs

Settings

Money Ball

Search

Analytics

Sources

AriDB2012

AriDB2012_batter_derived_attributes_sa

AriDB2012_pitcher_derived_attributes_sa

AriDB2013

AriDB2013_batter_derived_attributes_sa

AriDB2013_pitcher_derived_attributes_sa

AriDB2014

AriDB2014_batter_derived_attributes_sa

AriDB2014_pitcher_derived_attributes_sa

AriDB2015

AriDB2015_batter_derived_attributes_sa

AriDB2015_pitcher_derived_attributes_sa

AriDB2016

AriDB2012_batter_derived_attributes_sa

Select Analytic
/AriDB2012_batter_derived_attributes/aridb2012_batter_derived_output

\$ Frame Parameters

NAME	DATATYPE	PROMPT FOR INPUT?	VALUES
------	----------	-------------------	--------

SCHEMA	RELATIONSHIPS	PREVIEW	Test

COLUMN NAME	LABEL	DATA TYPE	DESCRIPTION	PRIMARY KE
batter	batter	INTEGER		<input type="checkbox"/>
BA_fb_over_93	BA_fb_over_93	DOUBLE		<input type="checkbox"/>
AB_fb_over_93	AB_fb_over_93	LONG		<input type="checkbox"/>
H_fb_over_93	H_fb_over_93	LONG		<input type="checkbox"/>
TB_fb_over_93	TB_fb_over_93	LONG		<input type="checkbox"/>
HR_fb_over_93	HR_fb_over_93	LONG		<input type="checkbox"/>
BA_fastball_under_93	BA_fastball_under_93	DOUBLE		<input type="checkbox"/>
AB_fastball_under_93	AB_fastball_under_93	DOUBLE		<input type="checkbox"/>
H_fb_under_93	H_fb_under_93	DOUBLE		<input type="checkbox"/>
TB_fb_under_93	TB_fb_under_93	DOUBLE		<input type="checkbox"/>
HR_fb_under_93	HR_fb_under_93	DOUBLE		<input type="checkbox"/>
BA_slider	BA_slider	DOUBLE		<input type="checkbox"/>
AB_slider	AB_slider	DOUBLE		<input type="checkbox"/>
H_slider	H_slider	DOUBLE		<input type="checkbox"/>
TB_slider	TB_slider	DOUBLE		<input type="checkbox"/>
HR_slider	HR_slider	DOUBLE		<input type="checkbox"/>
BA_curve	BA_curve	DOUBLE		<input type="checkbox"/>

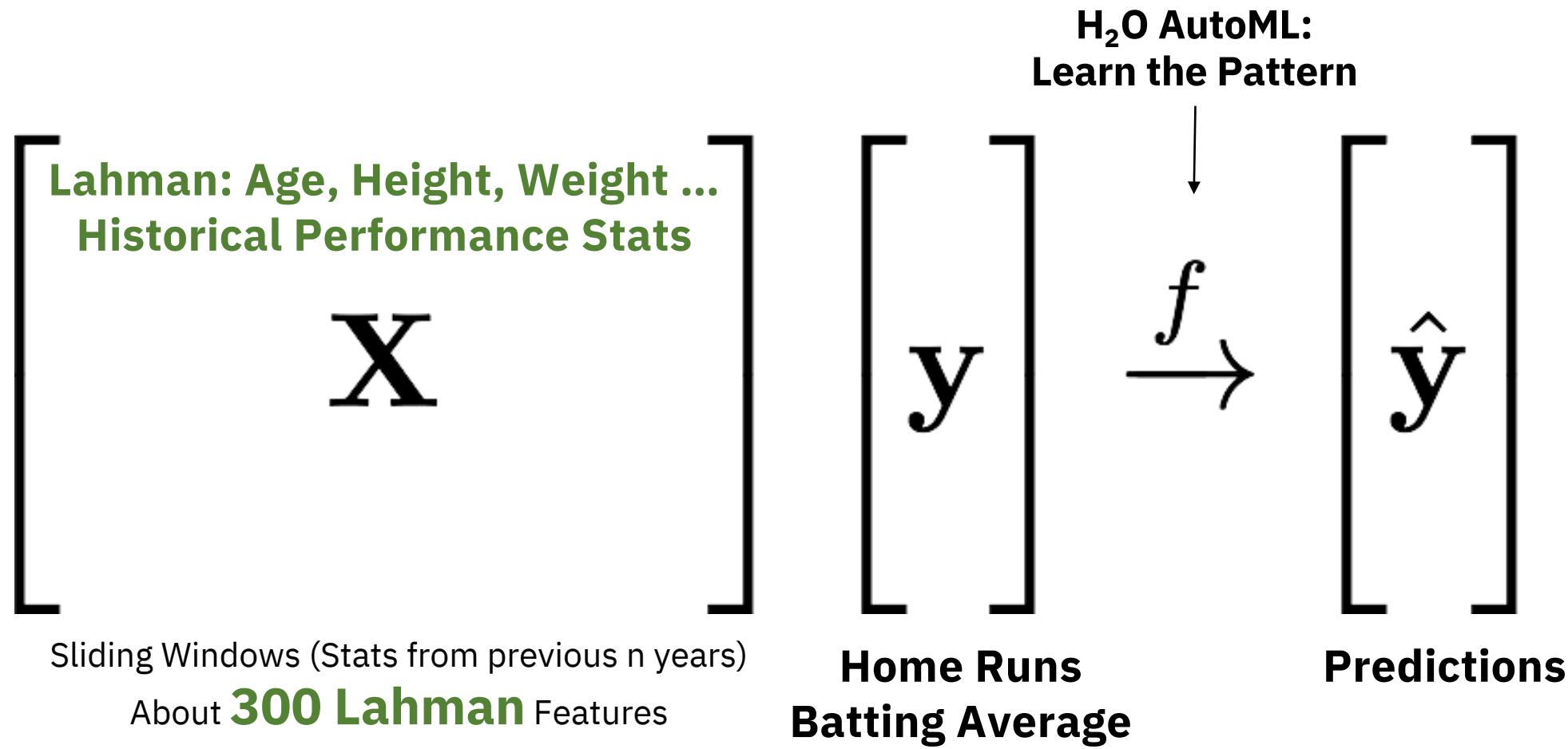
Creating Consistent and Reusable Analytic Assets Managed by Aginity Amp

The screenshot shows the Aginity Amp application interface. On the left, a sidebar navigation bar includes icons for Home, New, Browse, Jobs, and Settings, along with a search bar and a main workspace dropdown set to "Main Workspace". Below the sidebar, a list of "Analytics" assets is displayed, including "AriDB2012_batter_derived_attributes", "AriDB2012_pitcher_derived_attributes", "AriDB2013_batter_derived_attributes", "AriDB2013_pitcher_derived_attributes", "AriDB2014_batter_derived_attributes", "AriDB2014_pitcher_derived_attributes", "AriDB2015_batter_derived_attributes", "AriDB2015_pitcher_derived_attributes", "AriDB2016_batter_derived_attributes", "AriDB2016_pitcher_derived_attributes", and "AriDB2017_batter_derived_attributes".

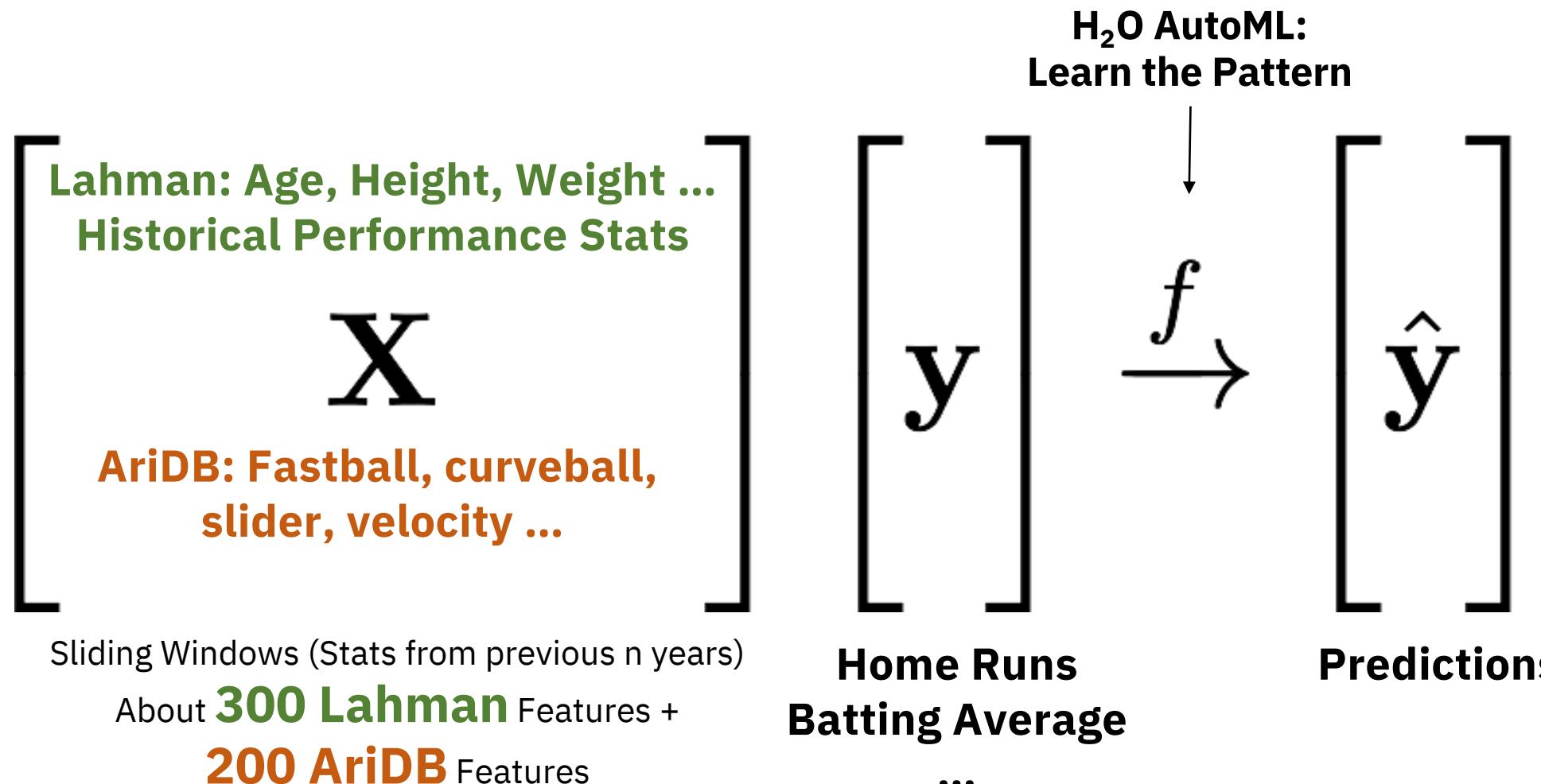
The main workspace area is titled "Money Ball" and contains a detailed view of the "AriDB2012_batter_derived_attributes" asset. This view includes an "INPUT SCHEMA" section with a "Parameters" dropdown set to "AriDB... ./AriDB20...". The right side of the workspace displays the SQL code for this asset:

```
%sql
create or replace temporary view AriDB2012_batter_derived_output as
  select batter,
  round(sum(if(Event in ('Single','Double','Triple','Home Run') and
  end_ab_flag='1' and start_speed>=93.0 and pitch_type in ('FF',
  'FT','FA'),1,0)) / sum(if(Event in ('Single','Double','Triple'
  , 'Home Run','Batter Interference','Bunt groundout','Bunt pop out',
  'Double Play','Fielders Choice','Fielders Choice Out','Flyout',
  'Forceout','Grounded Into DP','Groundout','Lineout','Pop Out',
  'Strikeout','Strikeout - DP','Triple Play') and end_ab_flag='1'
  and start_speed>=93.0 and pitch_type in ('FF','FT','FA'),1,0)),3)
  as BA_fb_over_93,
  sum(if(Event in ('Single','Double','Triple','Home Run','Batter
  Interference','Bunt groundout','Bunt pop out','Double Play',
  'Fielders Choice','Fielders Choice Out','Flyout','Forceout',
  'Grounded Into DP','Groundout','Lineout','Pop Out','Strikeout',
  'Strikeout - DP','Triple Play') and end_ab_flag='1' and
  start_speed>=93.0 and pitch_type in ('FF','FT','FA'),1,0)) as
  AB_fb_over_93,
  sum(if( spray_type='H' and end_ab_flag='1' and start_speed>=93.0 and
  pitch_type in ('FF','FT','FA'),1,0)) as H_fb_over_93,
  sum(if( spray_type='H' and end_ab_flag='1' and start_speed>=93.0 and
  pitch_type in ('FF','FT','FA'),if(event='Home Run',4,0)+if(event
  ='Triple',3,0)+if(event='Double',2,0)+if(event='Single',1,0),0))
  as TB_fb_over_93,
  sum(if(event='Home Run' and end_ab_flag='1' and start_speed>=93.0 and
  pitch_type in ('FF','FT','FA'),1,0)) as HR_fb_over_93,
  round(sum(if(Event in ('Single','Double','Triple','Home Run') and
  end_ab_flag='1' and start_speed<93.0 and pitch_type in ('FF','FT',
  'FA'),1,0)) / sum(if(Event in ('Single','Double','Triple','Home
  Run','Batter Interference','Bunt groundout','Bunt pop out',
  'Double Play','Fielders Choice','Fielders Choice Out','Flyout',
  'Forceout','Grounded Into DP','Groundout','Lineout','Pop Out',
  'Strikeout','Strikeout - DP','Triple Play') and end_ab_flag='1'
  and start_speed<93.0 and pitch_type in ('FF','FT','FA'),1,0)),3)
  as BA_fb_over_93,
```

Approach One: Learning from **Lahman** only



Approach Two: Learning from **Lahman** & **AriDB**



Lahman Data

Player's information

birthYear	birthMonth	birthDay	birthCountry	birthState	birthCity					
1991	8	7	USA	NJ	Vineland					
nameFirst	nameLast	nameGiven	weight	height	bats	throws	debut	finalGame	retroID	bbrefID
Mike	Trout	Michael Nelson	235	74	R	R	2011-07-08	2017-10-01	troum001	troutmi01

Player's past performance (batting in this case)

	playerID	yearID	stint	teamID	lgID	G	AB	R	H	2B	3B	HR	RBI	SB	CS	BB	SO	IBB	HBP	SH	SF	GIDP
95484	troutmi01	2011	1	LAA	AL	40	123	20	27	6	0	5	16	4	0	9	30	0	2	0	1	2
96904	troutmi01	2012	1	LAA	AL	139	559	129	182	27	8	30	83	49	5	67	139	4	6	0	7	7
98308	troutmi01	2013	1	LAA	AL	157	589	109	190	39	9	27	97	33	7	110	136	10	9	0	8	8
99744	troutmi01	2014	1	LAA	AL	157	602	115	173	39	9	36	111	16	2	83	184	6	10	0	10	6
101226	troutmi01	2015	1	LAA	AL	159	575	104	172	32	6	41	90	11	7	92	158	14	10	0	5	11
102712	troutmi01	2016	1	LAA	AL	159	549	123	173	32	5	29	100	30	7	116	137	12	11	0	5	5
104195	troutmi01	2017	1	LAA	AL	114	402	92	123	25	3	33	72	22	4	94	90	15	7	0	4	8

Lahman Data Framed as a ML problem

yearID	teamID	lgID
2011	LAA	AL
2012	LAA	AL
2013	LAA	AL
2014	LAA	AL
2015	LAA	AL
2016	LAA	AL
2017	LAA	AL
2018	LAA	AL
2019	LAA	AL
2020	LAA	AL

yearID	teamID	lgID	weight	height	bats	throws	birthYear	birthCountry	birthState	birthCity
2011	LAA	AL	235	74	R	R	1991	USA	NJ	Vineland
2012	LAA	AL	235	74	R	R	1991	USA	NJ	Vineland
2013	LAA	AL	235	74	R	R	1991	USA	NJ	Vineland
2014	LAA	AL	235	74	R	R	1991	USA	NJ	Vineland
2015	LAA	AL	235	74	R	R	1991	USA	NJ	Vineland
2016	LAA	AL	235	74	R	R	1991	USA	NJ	Vineland
2017	LAA	AL	235	74	R	R	1991	USA	NJ	Vineland
2018	LAA	AL	235	74	R	R	1991	USA	NJ	Vineland
2019	LAA	AL	235	74	R	R	1991	USA	NJ	Vineland
2020	LAA	AL	235	74	R	R	1991	USA	NJ	Vineland

yearID	age	career_year
2011	20	1
2012	21	2
2013	22	3
2014	23	4
2015	24	5
2016	25	6
2017	26	7
2018	27	8
2019	28	9
2020	29	10

Player
Attributes

Past
Performance
Sliding
Windows
+
Other
Stats

last1_HR	last2_HR	last3_HR	last4_HR	last5_HR	avg_last2_HR	avg_last3_HR	avg_last4_HR	avg_last5_HR
NA	NA	NA	NA	NA	Nan	Nan	Nan	Nan
5	NA	NA	NA	NA	5.0	5.00000	5.00000	5.00000
30	5	NA	NA	NA	17.5	17.50000	17.50000	17.50000
27	30	5	NA	NA	28.5	20.66667	20.66667	20.66667
36	27	30	5	NA	31.5	31.00000	24.50000	24.50000
41	36	27	30	5	38.5	34.66667	33.50000	27.80000
29	41	36	27	30	35.0	35.33333	33.25000	32.60000
33	29	41	36	27	31.0	34.33333	34.75000	33.20000
33	33	29	41	36	33.0	31.66667	34.00000	34.40000
33	33	33	29	41	33.0	33.00000	32.00000	33.80000

No data. Used 2017 value. Not perfect (a quick hack).

One of the Targets

yearID	HR
2011	5
2012	30
2013	27
2014	36
2015	41
2016	29
2017	33
2018	NA
2019	NA
2020	NA

Training

Validation

Forecast

H₂O AutoML Code

```
# H2O AutoML with Lahman only
automl_lahman = h2o.automl(x = features,
                            y = targets[n_target],
                            training_frame = h_train,
                            validation_frame = h_valid,
                            max_models = 10, # increase this to allow more models
                            max_runtime_secs = 120, # increase this to allow more time
                            stopping_metric = "RMSE",
                            stopping_rounds = 3,
                            seed = n_seed,
                            exclude_algos = c("DeepLearning"), # you can exclude any algo
                            project_name = paste0("AutoML_Lahman", targets[n_target]))
```

H₂O AutoML Results

```
H2OResgressionMetrics: stackeddenseensemble
** Reported on cross-validation data. **
** 5-fold cross-validation on training data (Metrics computed for combined holdout predictions) **

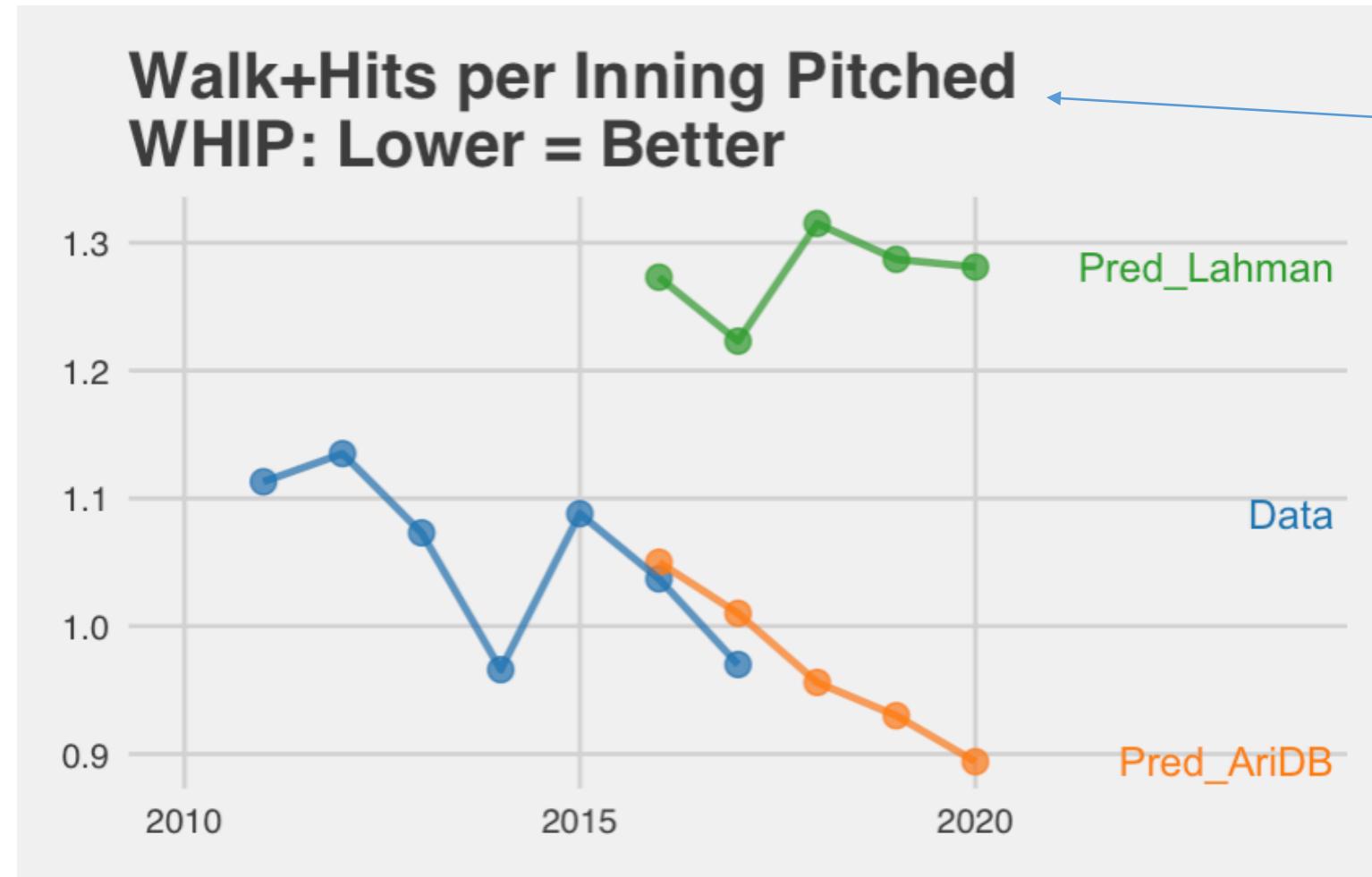
MSE: 0.00246453
RMSE: 0.04964404
MAE: 0.03335875
RMSLE: 0.04124294
Mean Residual Deviance : 0.00246453
```

Slot "leaderboard":

		model_id	mean_residual_deviance	rmse	mae	rmsle
1	StackedEnsemble_BestOfFamily_0_AutoML_20180615_040834		0.002465	0.049644	0.033359	0.041243
2	StackedEnsemble_AllModels_0_AutoML_20180615_040834		0.002467	0.049669	0.033367	0.041265
3	GLM_grid_0_AutoML_20180615_040834_model_0		0.002480	0.049802	0.033560	0.041401
4	GBM_grid_0_AutoML_20180615_040834_model_4		0.002486	0.049856	0.033707	0.041373
5	GBM_grid_0_AutoML_20180615_040834_model_2		0.002564	0.050638	0.034346	0.042008
6	GBM_grid_0_AutoML_20180615_040834_model_1		0.002569	0.050684	0.034261	0.042022

[12 rows x 5 columns]

Predictive Modelling – H₂O AutoML

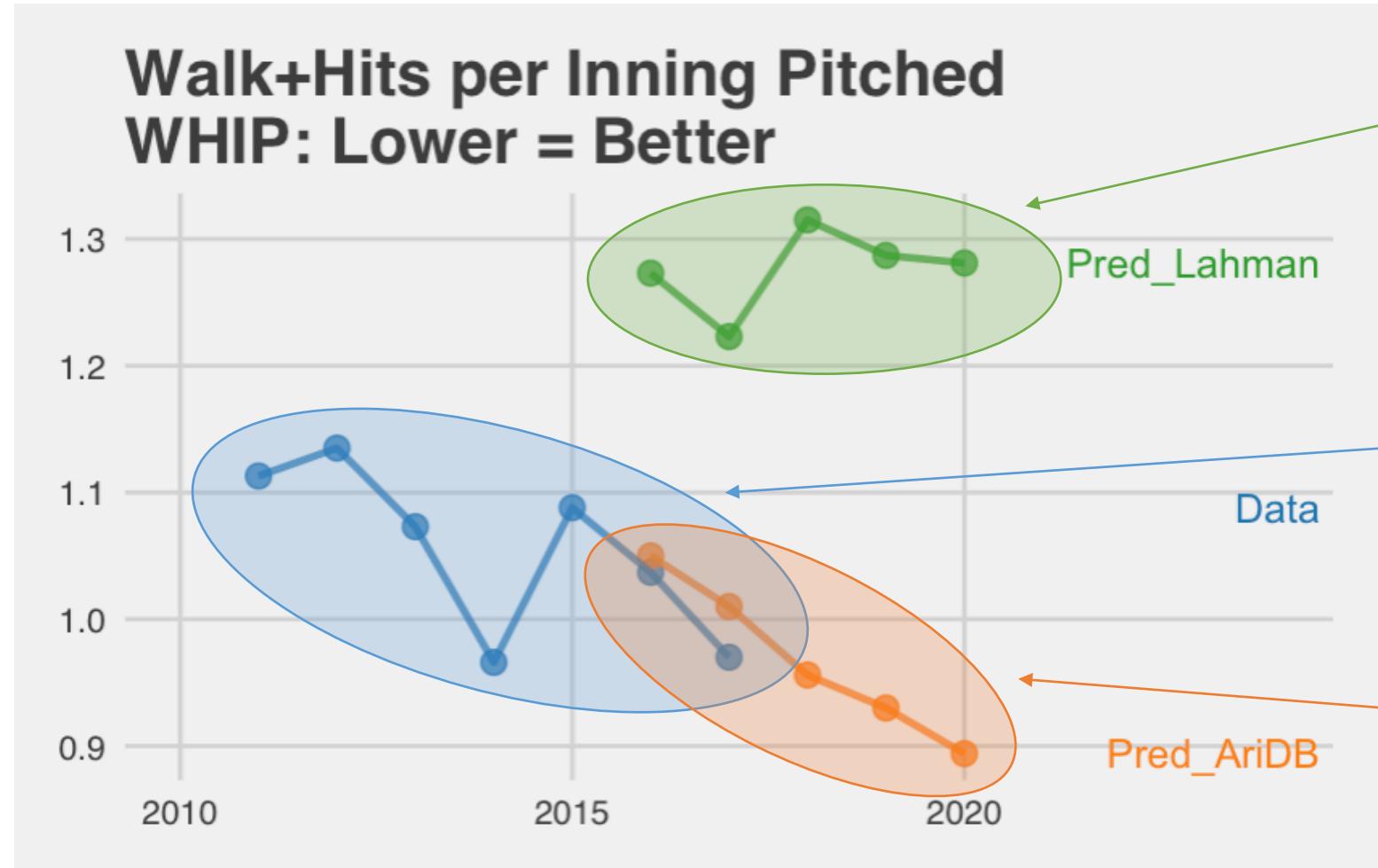


One of Many Targets
(e.g. Home Runs, Batting Average)



```
# Install 'h2o' from CRAN  
install.packages('h2o')
```

Predictive Modelling – H₂O AutoML

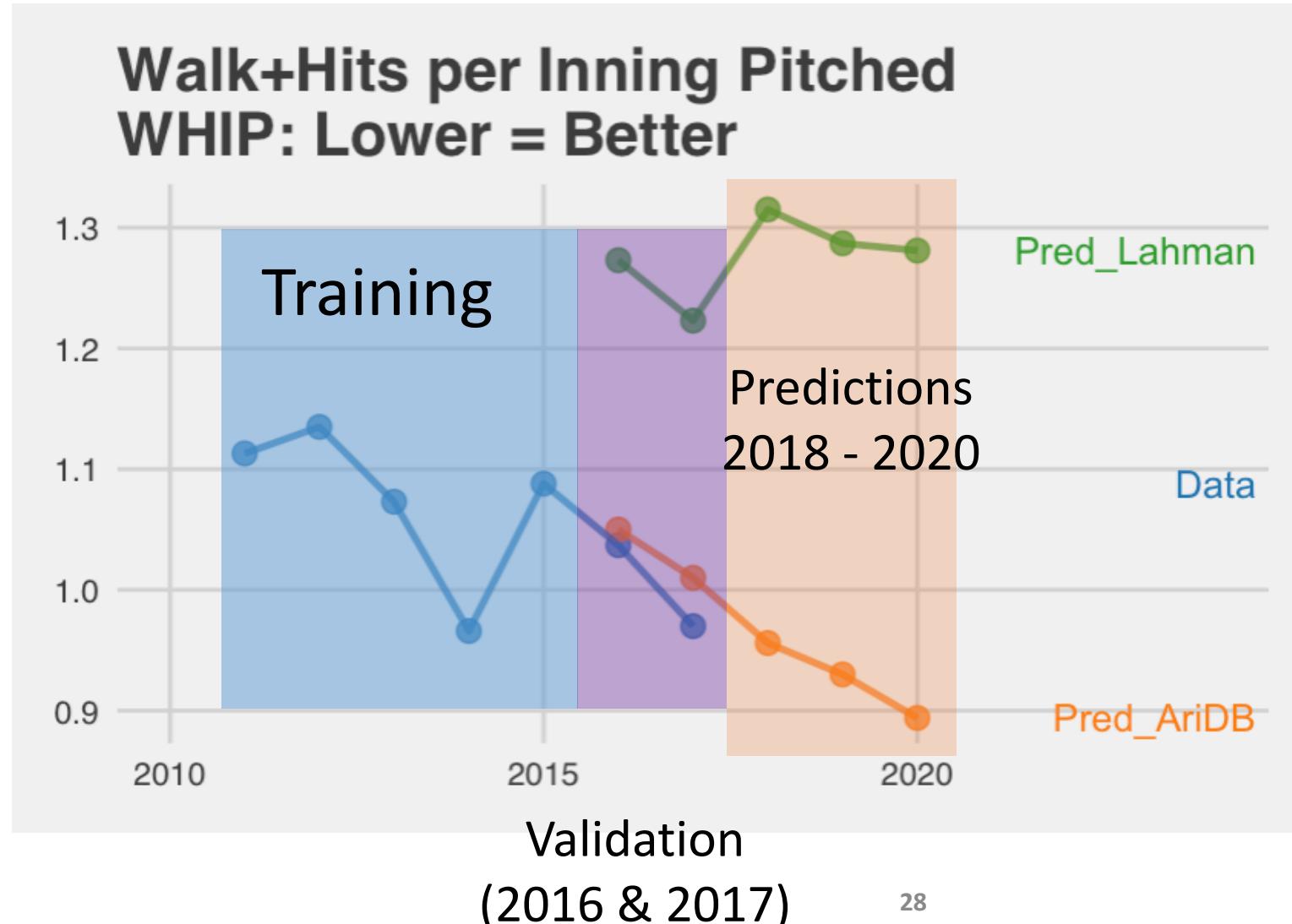


Results from models based on Lahman data only

Historical player performance data

Results from models based on final dataset (Lahman + AriDB)

Predictive Modelling – H₂O AutoML

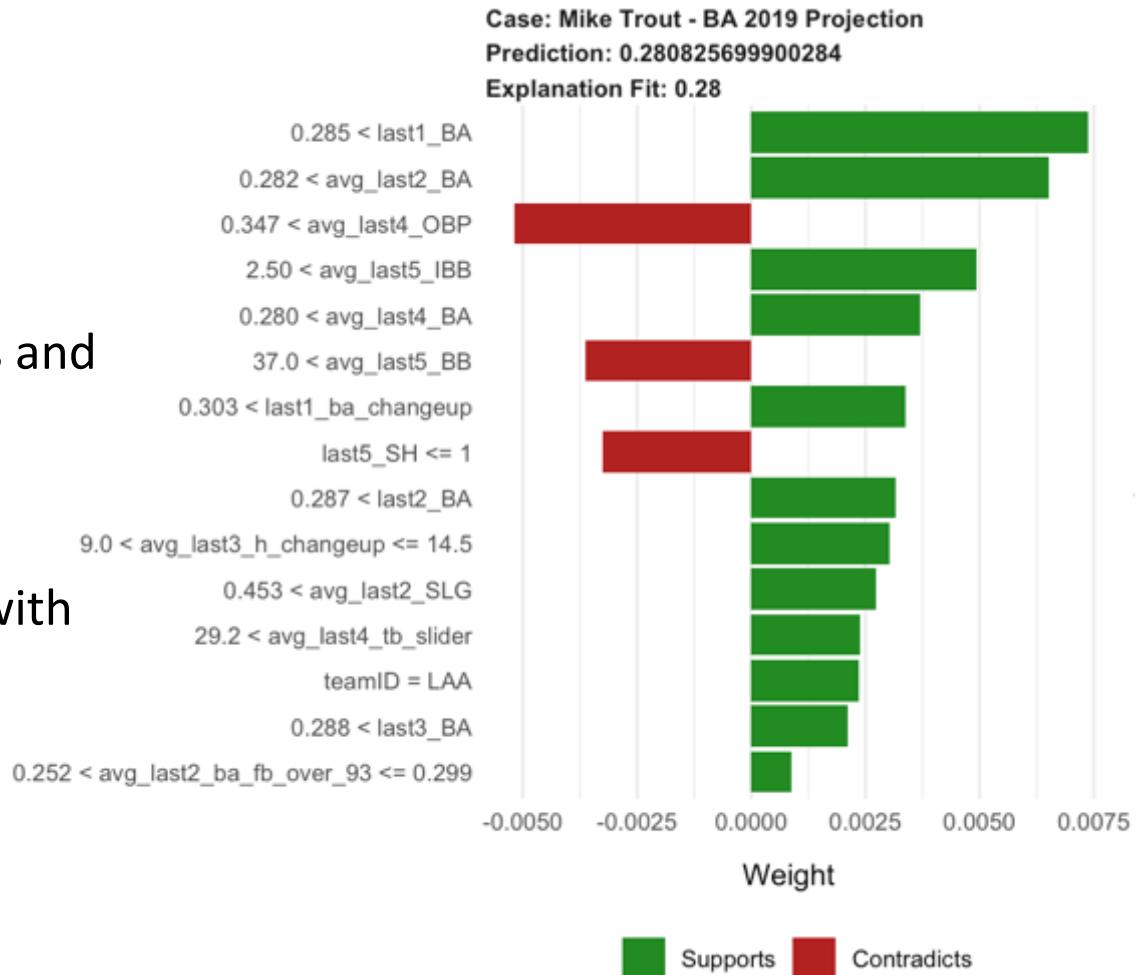


```
# Install 'h2o' from CRAN  
install.packages('h2o')
```

Explaining the Predictions

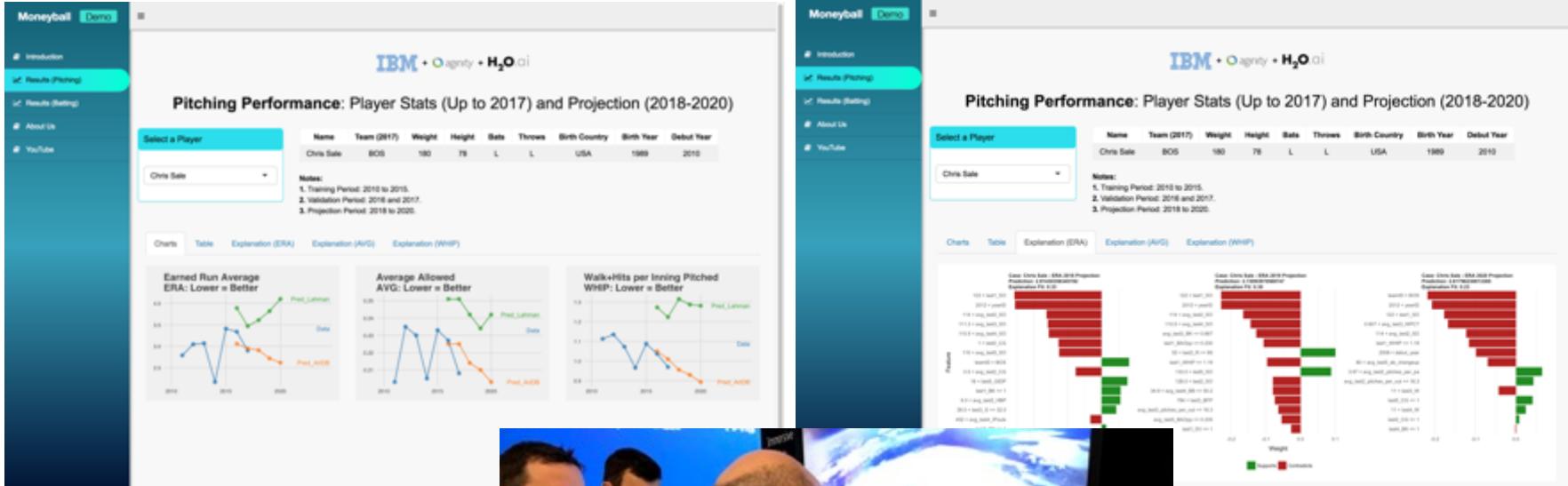
LIME – Local Interpretable Model-agnostic Explanations

- Approximate reasoning of complex ML models (ensembles).
- Most important attributes and their contributions to the predictions.
- Ari validated the models with his 30+ years of baseball domain knowledge.
- He trusted the models.



```
# Install 'lime' from CRAN  
install.packages('lime')
```

Putting Everything Together – Moneyball Shiny App



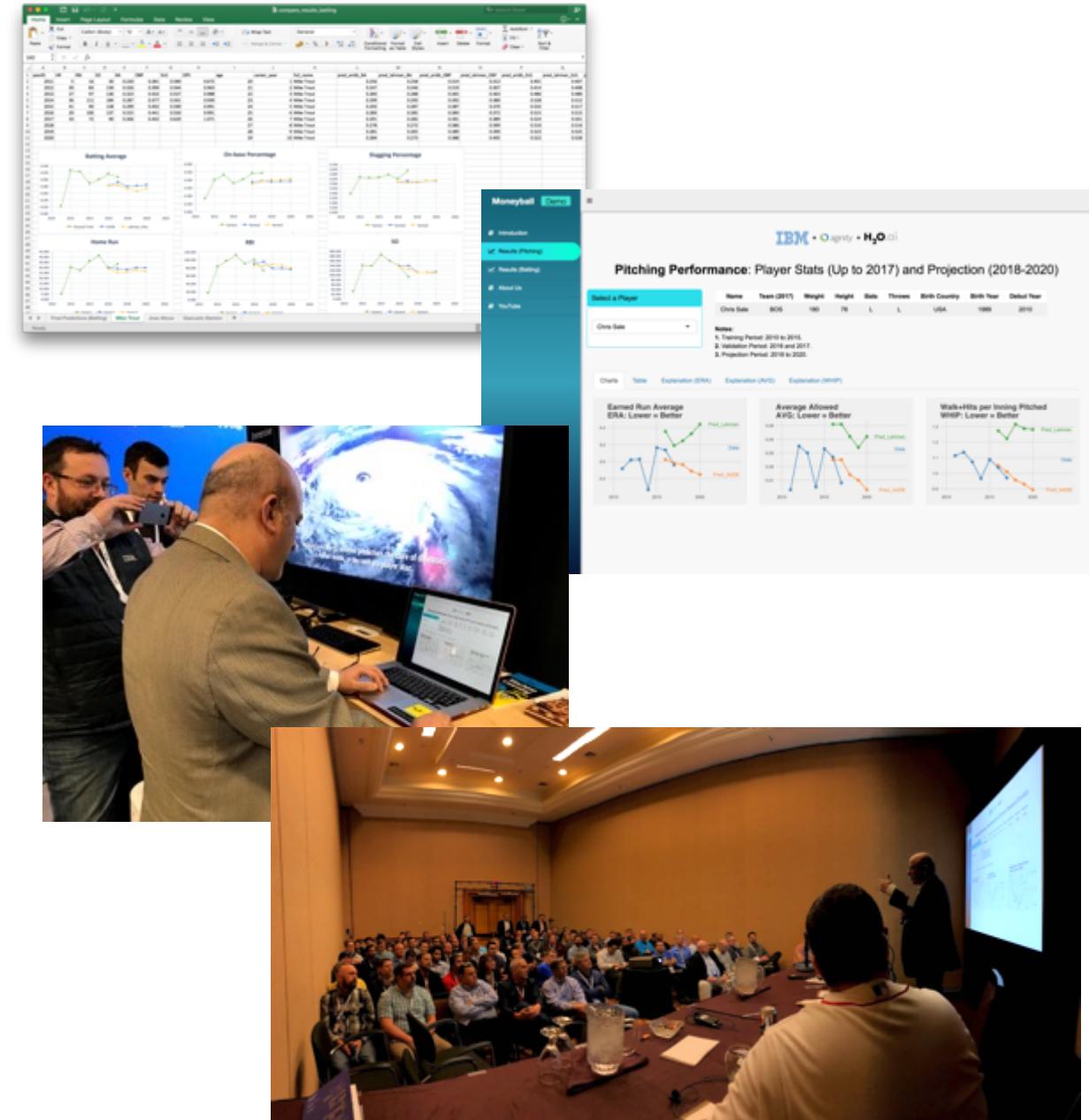
Live Demo
on my laptop



H₂O.ai

From Toy Demo to Real Moneyball

- **March 19** – AutoML Predictions finalized.
Initial presentation in Excel.
- **March 20** – Version 1 of Shiny app. Ari used the app to validate some players he had in mind and recommended one player to his team.
- **March 21** – Multimillion-dollar contract finalized.
- **March 22** – Moneyball presentation at IBM Think

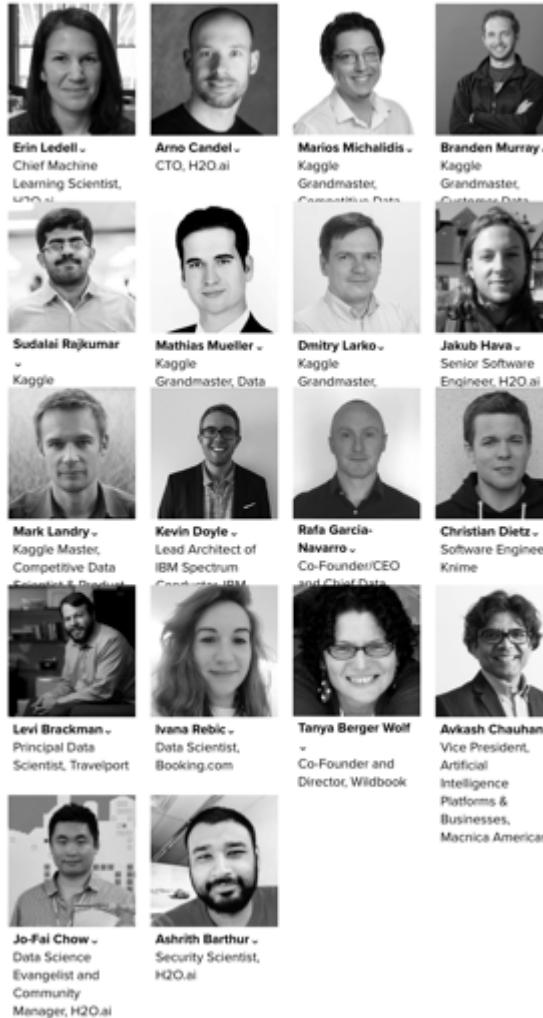


If you want to hear the Moneyball story from Ari ...



29th & 30th Oct, London

Session Speakers



More real-world use cases + All H₂O Kaggle Grandmasters + Hands-on Training

H₂O.ai

Thanks!



- More Info, Code, and Slides
 - [bit.ly/
h2o_meetups](https://bit.ly/h2o_meetups)
- Contact
 - joe@h2o.ai
 - [@matlabulous](https://twitter.com/matlabulous)
 - github.com/woobe