



Forecasting Using Domo

Lunch & Learn | April 6, 2023

Florencia Silveira, Bryan Powell, & Steven Guitron

Data Scientists

Data Science Professional Services at Domo, Inc.

What is forecasting?

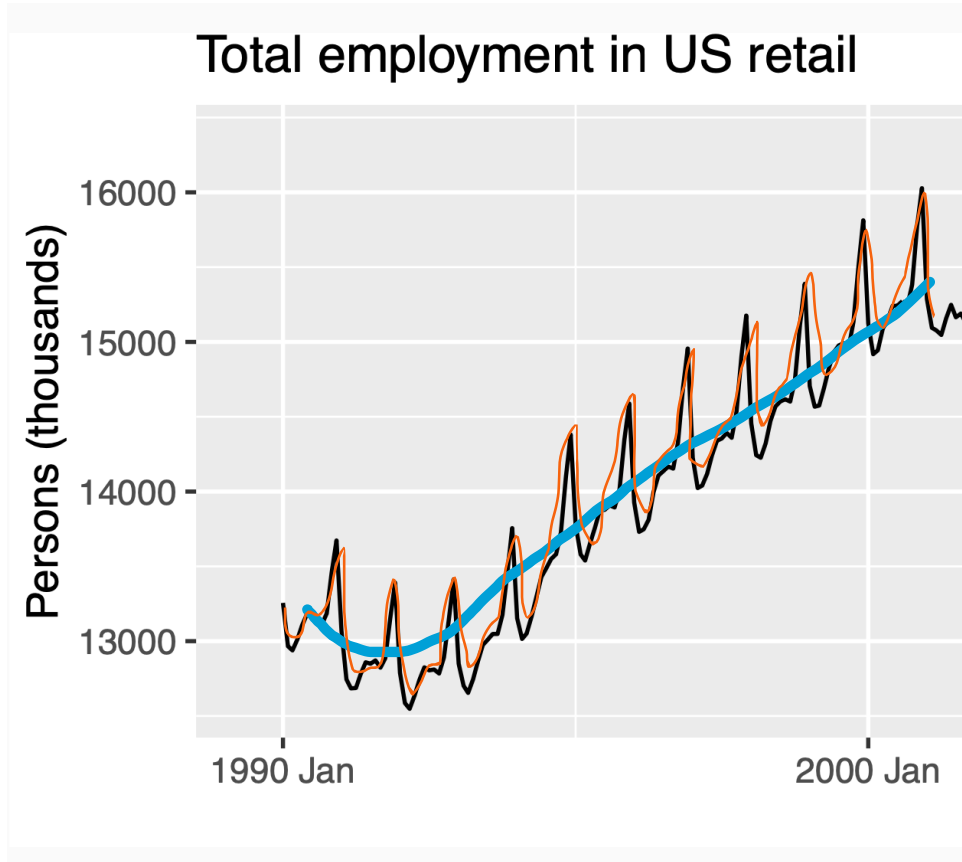
Forecasting involves predicting future outcomes using historical data

- Specific focus of this presentation: forecasting using historical *outcome* data



Why use data on the past to predict the future?

Because the future is often like the past



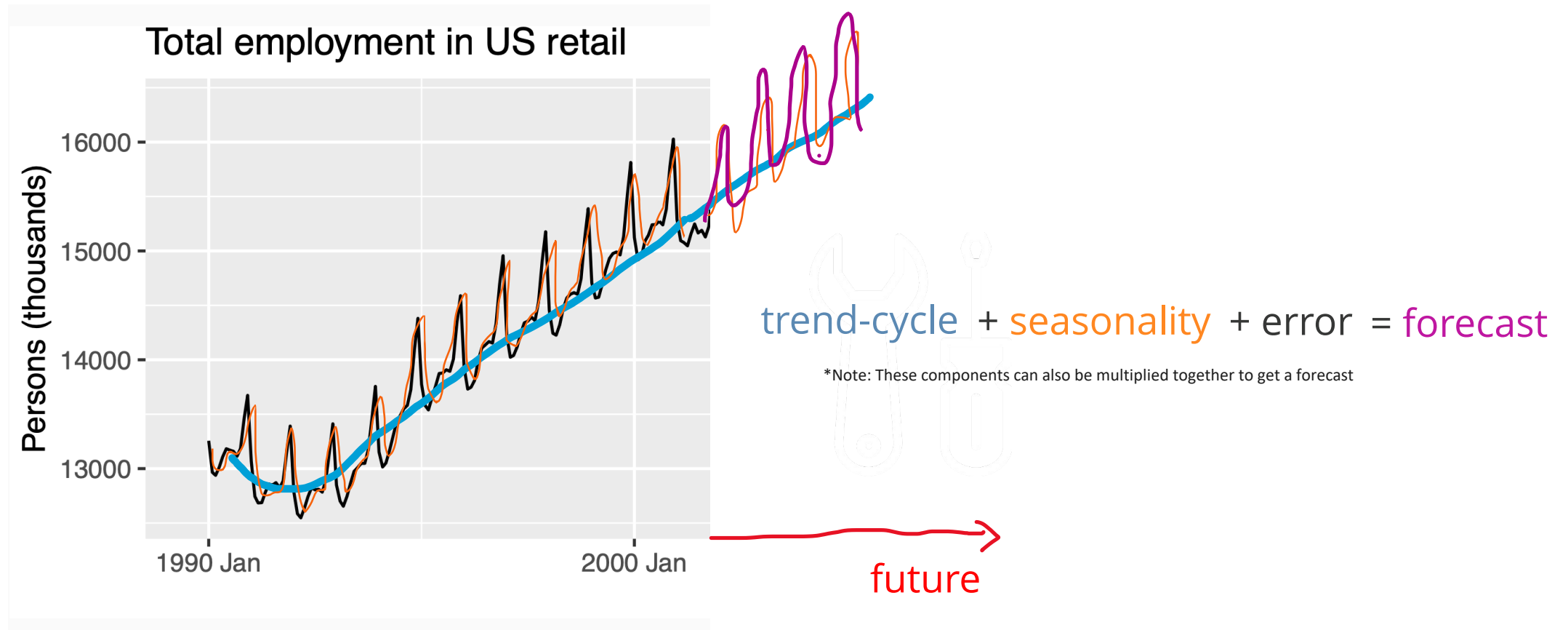
Retail employment tends to follow the same pattern over time

Components of the pattern:

- **Trend-cycle**: long-term increases or decreases in the data
- **Seasonality**: repeating cycle in the series with fixed frequencies (by hour of the day, day of the week, month, etc.)
- **Error/Noise**: leftover fluctuation (variation) in the data

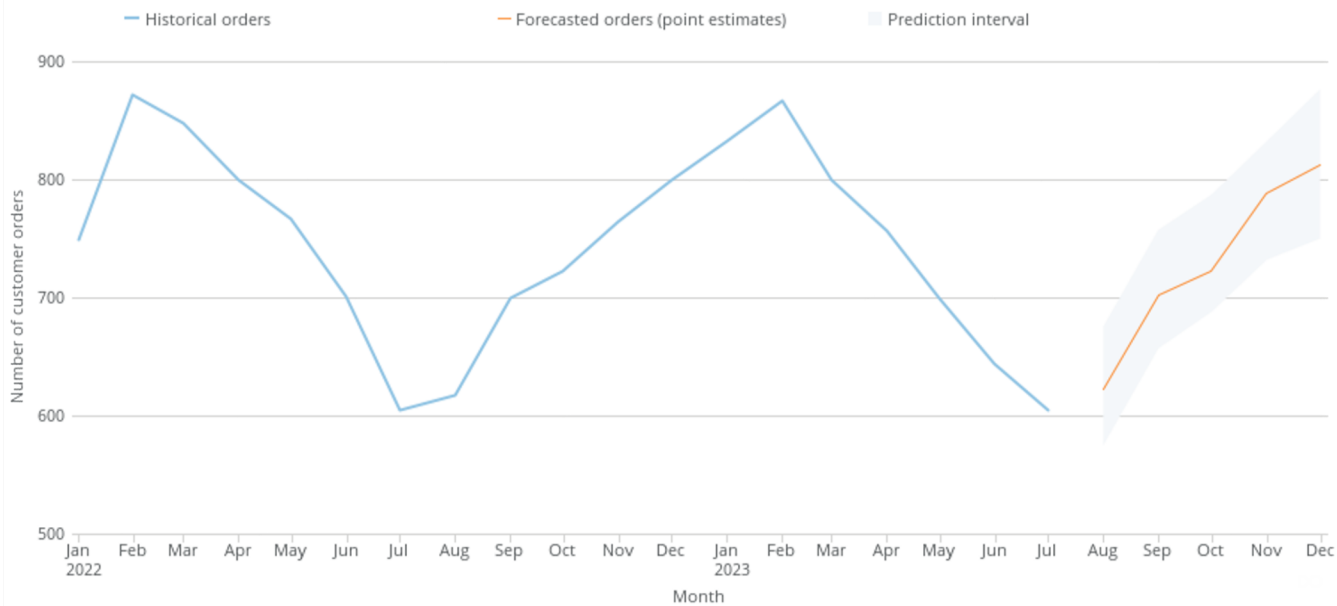
How Forecasting Works

Forecasting predicts future outcomes by extending historical patterns into the future (using statistics)



What information do I get from a forecasting model?

Number of Customer Orders by Month - Historical + Forecast



- *Prediction intervals*: Range of possible values that future outcome could take with relatively high probability
 - Typical width/range of prediction interval: 95%
- *Point estimates*: Average of possible values that future outcome could take
 - Measured in the same unit as the historical outcome data
- How far in the future you forecast is up to you
 - The farther out you forecast, the more uncertain the future is. Thus, the wider the prediction intervals are

What kinds of outcomes can be forecast?

Anything that can be observed sequentially over time & has a strong pattern

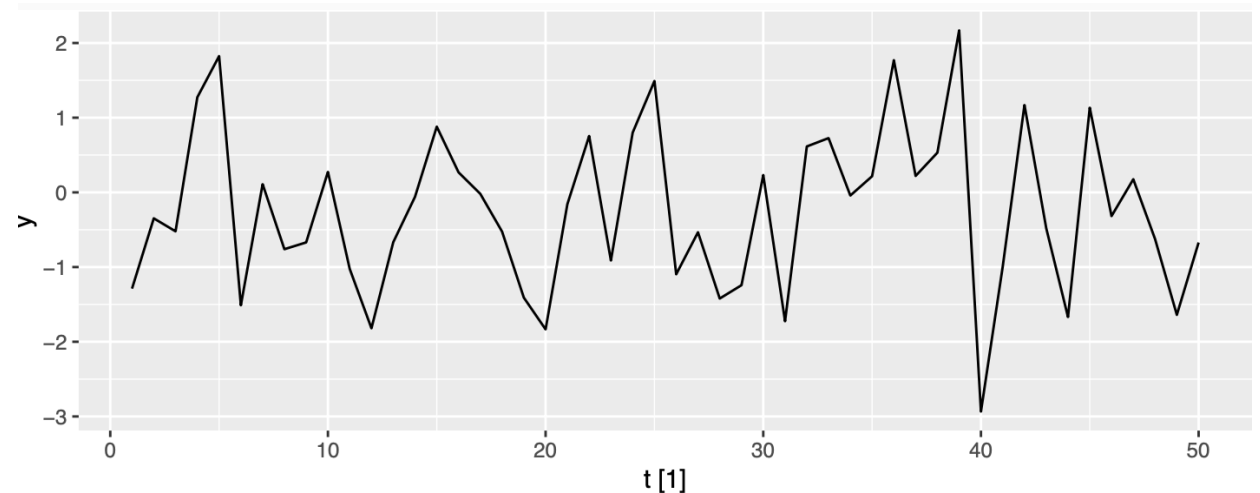
YES

- Annual company revenue
- Quarterly customer purchase amounts
- Weekly course enrollment
- Daily resort attendance

NO

- Random events (winning lotto numbers, fashion trends, natural disasters, etc.)


Outcome with no pattern over time



What data do I need in order to forecast?

You need historical data on your outcome that is.....


- *Sequential*, meaning observed at regular intervals of time
 - Format of data: 1 column with time variable, 1 column with outcome variable
 - Example time intervals: hourly, daily, weekly, monthly, quarterly, annually
 - The time interval of your forecast = time interval of your historical outcome data
 - Can't have lots of missing data!!

Month	Outcome
 24 unique values	123 17 unique values
2022-01-01	749
2022-02-01	871
2022-03-01	848
2022-04-01	799
2022-05-01	767
2022-06-01	701
2022-07-01	605
2022-08-01	617
2022-09-01	700
2022-10-01	722
2022-11-01	764
2022-12-01	800
2023-01-01	832
2023-02-01	866
2023-03-01	800

What data do I need in order to forecast?

You need historical data on your outcome that is.....

- *Plentiful*, meaning enough data to be able to detect patterns (trends + seasonality)
 - What is considered “plentiful” is highly dependent on your outcome, the time interval of observation, and the complexity of the pattern
 - The more data the better!!

Month	Outcome
 24 unique values	123 17 unique values
2022-01-01	749
2022-02-01	871
2022-03-01	848
2022-04-01	799
2022-05-01	767
2022-06-01	701
2022-07-01	605
2022-08-01	617
2022-09-01	700
2022-10-01	722
2022-11-01	764
2022-12-01	800
2023-01-01	832
2023-02-01	866
2023-03-01	800

****This is not enough data to capture yearly trends (only have ~1 year of data) or monthly seasonality (only observe most months once)**

Forecasting using Domo

Bronze

Silver

Silver

Gold

Forecasting Tools	Availability	Ease of Use	Forecasting Output		Customization		Forecasting Visualization Options
			Point estimates	Prediction intervals	Available algorithms	Assess accuracy using train/test samples	
Line and bar chart cards: last value projection or multi-period projection option More info	Available to all customers	Very easy	Yes	No	Linear regression, forecasting, averages (you specify)	No	Forecast displayed as part of line or bar chart cards
Forecasting Magic ETL Data Science tile More info	Premium feature, available as part of Data Science toolset	Moderate	Yes	Yes	ARIMA	No	Forecasting chart card
AutoML More info	Premium feature, available as part of Data Science toolset	Moderate	Yes	No	Linear Learner, XGBoost, multilayer perceptron (MLP) (the tool chooses)	Yes (automatically provided, but not customizable)	Forecasting chart card
Jupyter Workspaces More info	Premium feature, available as part of Data Science toolset	More challenging (requires knowledge of how to code in R or Python)	Yes	Yes	Any forecasting algorithm you specify	Yes (fully customizable)	Forecasting chart card OR custom visualization created in Jupyter

Live Demos of Domo Forecasting Tools



Thank You



Jupyter Workspaces

Domo's integrated Jupyter Workspaces let's you explore your data, train & productionize models, and create custom visualizations - all in one place

FEATURES

- Read & write data directly to & from Domo
- Write code in notebooks using Python or R programming languages
- Choose your compute tier
- Productionize models in a notebook
- Run tasks programmatically (for developers)
- Schedule notebooks to run automatically at a set time or frequency
- Share workspaces to enable collaboration
- Integrate with GitHub for version control

The screenshot displays the Domo Jupyter Workspaces interface. On the left, a file explorer shows a directory with files: README.md (16 minutes ago), untitled.md (8 minutes ago), and untitled.txt (8 minutes ago). The main area is divided into a 'Launcher' sidebar with icons for Notebook, Console, and Terminal, and a central workspace. The workspace shows an open R notebook titled 'R_sentiment_analysis.ipynb'. The notebook content includes a title 'Sentiment Analysis for demo', a purpose statement, a 'Prep Work' section, and code for installing and loading R packages. The final cell shows the output of a data query, displaying a tibble with columns: date, review_id, stars, text, store, and city.

```
[1]: #required packages
packages <- c("SentimentAnalysis", "sentimentr", "corpus", "SnowballC")

#this code only installs the packages listed above if they're not already installed (which saves time)
install.packages(setdiff(packages, rownames(installed.packages())))
```

```
[2]: library("domojupyter")
library("SentimentAnalysis")
library("sentimentr")
library("SnowballC")
library("data.table")
library(tn)
library(dplyr)
library(tidyverse)
```

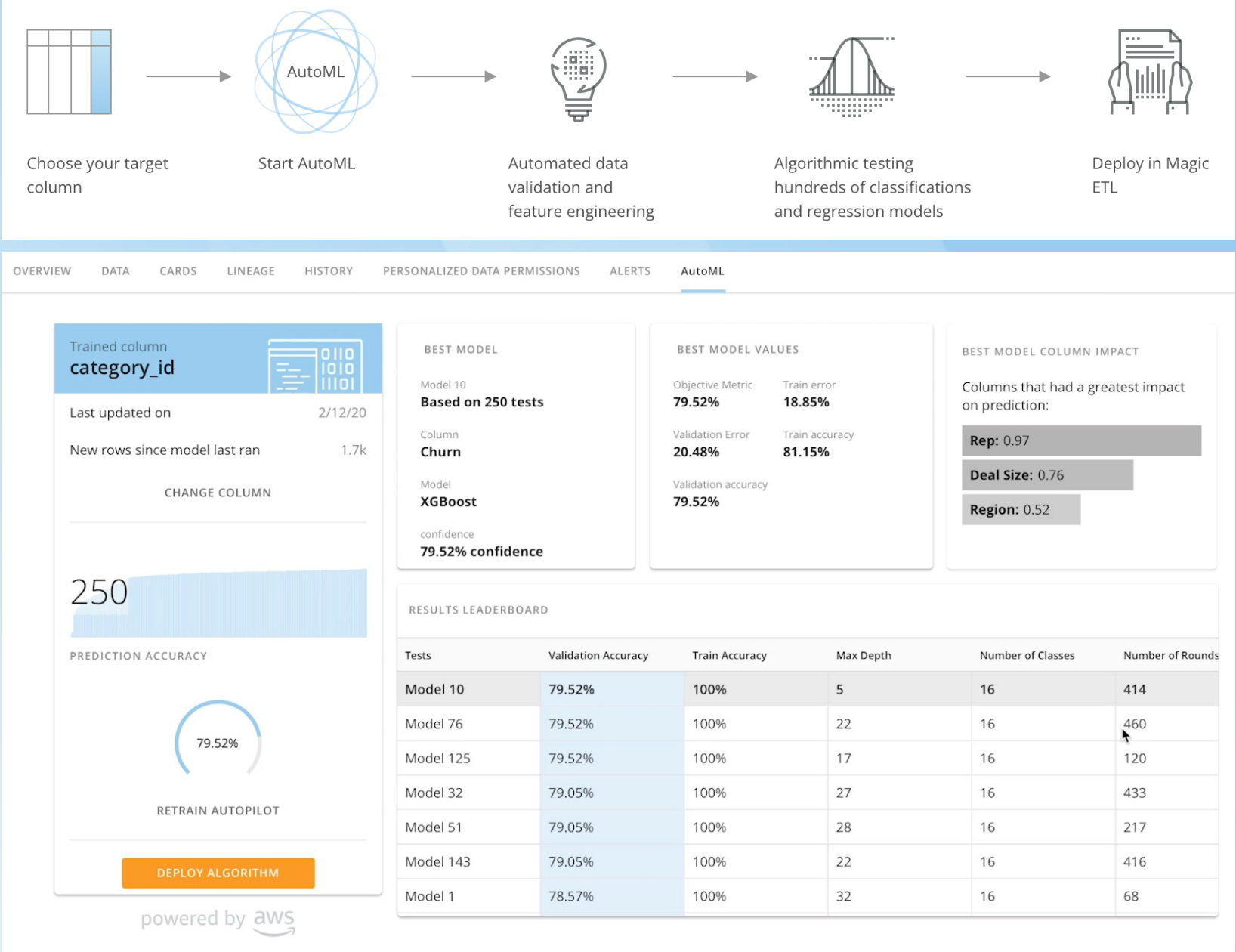
```
[3]: data <- read_dataframe("Reviews_for_sentiment_analysis", query = "SELECT * FROM table", show_col_types = FALSE)
head(data)
```

date	review_id	stars	text	store	city
<dtm>	<chr>	<dbl>	<chr>	<chr>	<chr>
2022-10-25 11:03:15	ChZDSuHNMG9nSOVJQ0FnSUMtc1B2SUnEAE	3	(Translated by Google) It's still New York! (Original) やはりニューヨークですね!	UNIQLO	New York
2022-10-19 04:52:57	ChdDSuHNMG9nSOVJQ0FnSURicZPvmI3RRAB	5	Bought a very nice hat from here.	UNIQLO	New York
2022-10-14 21:13:21	ChdDSuHNMG9nSOVJQ0FnSURiAGVYXzJnRRAB	5	Love UNIQLO. This store has a good selection of clothes. Prices are the same as the store in Brooklyn and NJ.	UNIQLO	New York
2022-10-11 18:29:11	ChdDSuHNMG9nSOVJQ0FnSURIZZVieHhBRAB	4	A store with huge selection but poor signage. Staff is helpful.	UNIQLO	New York
2022-10-11 10:16:24	ChdDSuHNMG9nSOVJQ0FnSURIM3UTZ5BRAB	5	(Translated by Google) Great store, special products (Original) Tolles Geschäft, besondere Produkte	UNIQLO	New York
2022-10-06 08:01:26	ChdDSuHNMG9nSOVJQ0FnSURiHRTnJnRRAB	3	(Translated by Google) It's smaller than the UNIQLO's in Soho or on Fifth Avenue. (Original) Es más pequeña que los UNIQLO del Soho o de la Quinta Avenida	UNIQLO	New York



AutoML

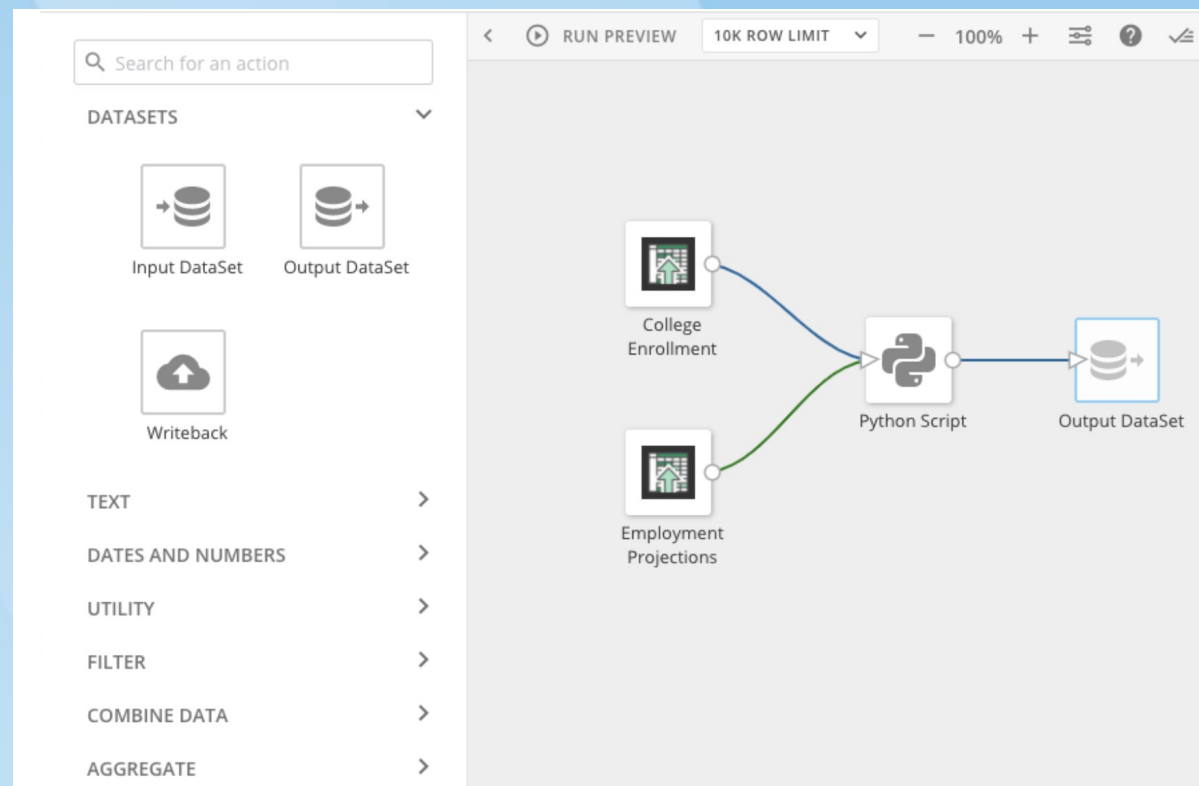
AutoML is a low-code option for automatically generating & testing multiple machine learning models





Scripting Tiles

Use Scripting Tiles to run brief R or Python code directly in an ETL dataflow



```
Python Script  CODE SCHEMA NOTES

1 # Import the domomagic package into the script
2 from domomagic import *
3
4 # read data from inputs into a data frame
5 input1 = read_dataframe('College Enrollment')
6
7 # write your script here
8 |
9
10 # write a data frame so it's available to the next action
11 write_dataframe(input1)
```