



# Content

## 01 | The AI Opportunity for CPG Companies

Room for significant first mover competitive advantage

## 02 | Key Requirements for Successful AI

Must win capabilities to be fit for the future

## 03 | Applied AI at visualfabriq

Client driven models, embedded in the solution

## 04 | The ROI of AI

AI is not an endpoint but a journey; but it can really pay off



# The AI Opportunity for CPG Companies

The average TPM system is  
not replaced before  
**10 to 15 years** of service





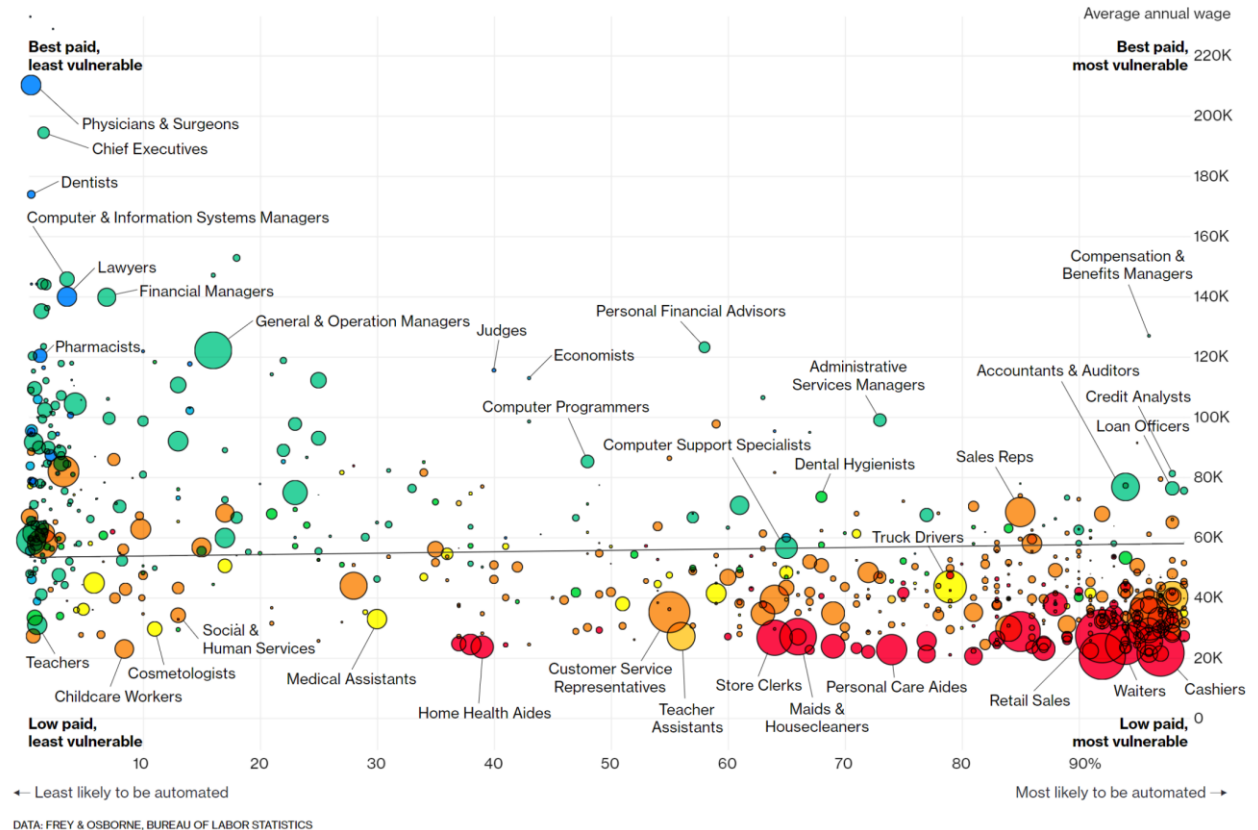
# SO, HAVE YOU THOUGHT ABOUT 2030?





By 2030 your work will have been radically transformed by AI

With many CPG functions affected





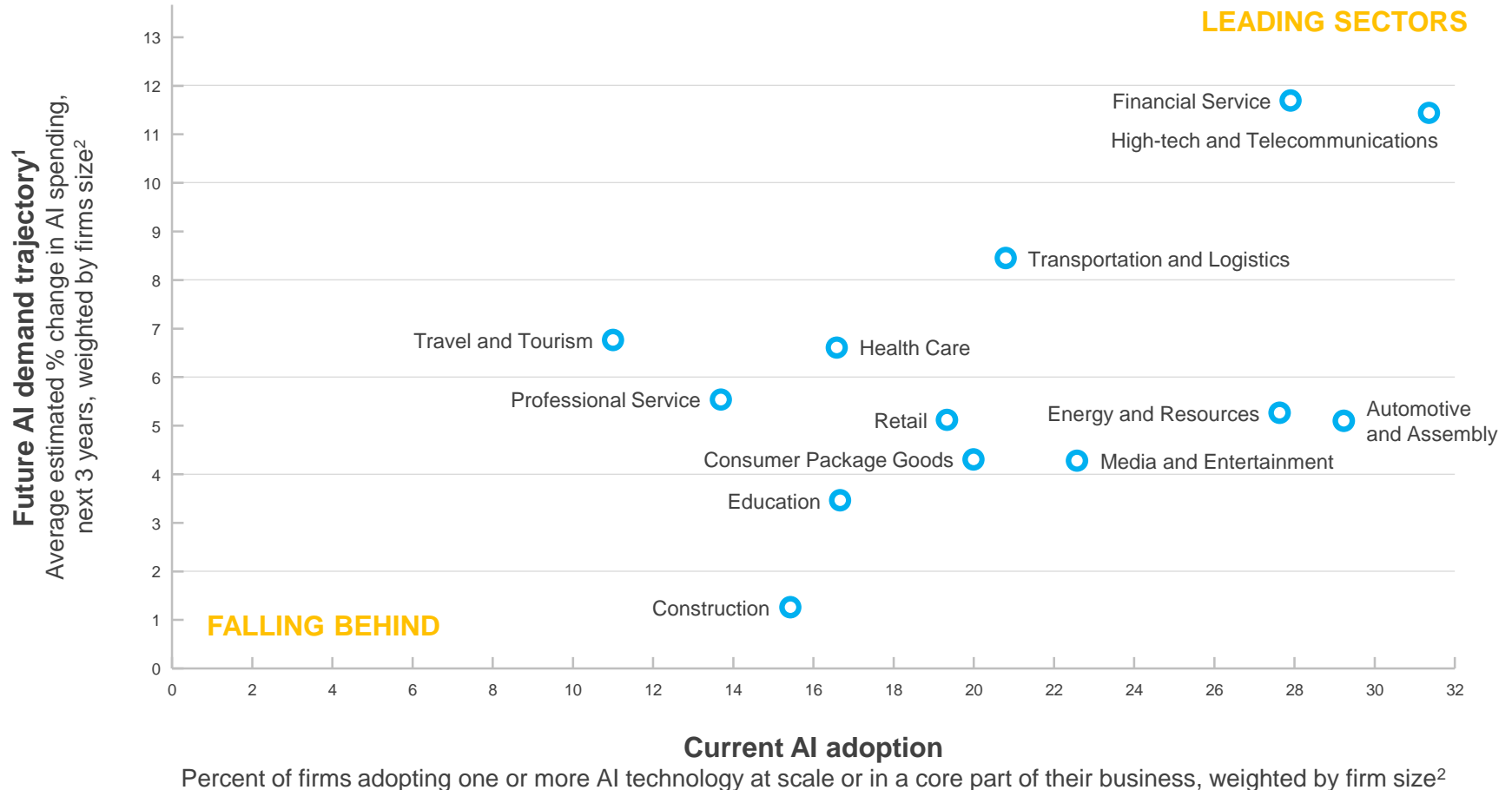
It's 2018.  
Most capabilities  
are **already** here.



# Our industry is still way behind in adoption

And perhaps lacks some imagination

**Fig. 30. High-tech and Financial Service Firms Lead Demand for AI**





## BUT AI WILL CHANGE THE WAY OF WORKING

User interactions will be more  
and more voice based

Users will be able plan and review predicted  
results directly together with clients





**Anyone working in CPG Sales and Operations will have  
AI Assistants augmenting their day to day work and  
perform routine operational tasks**



“

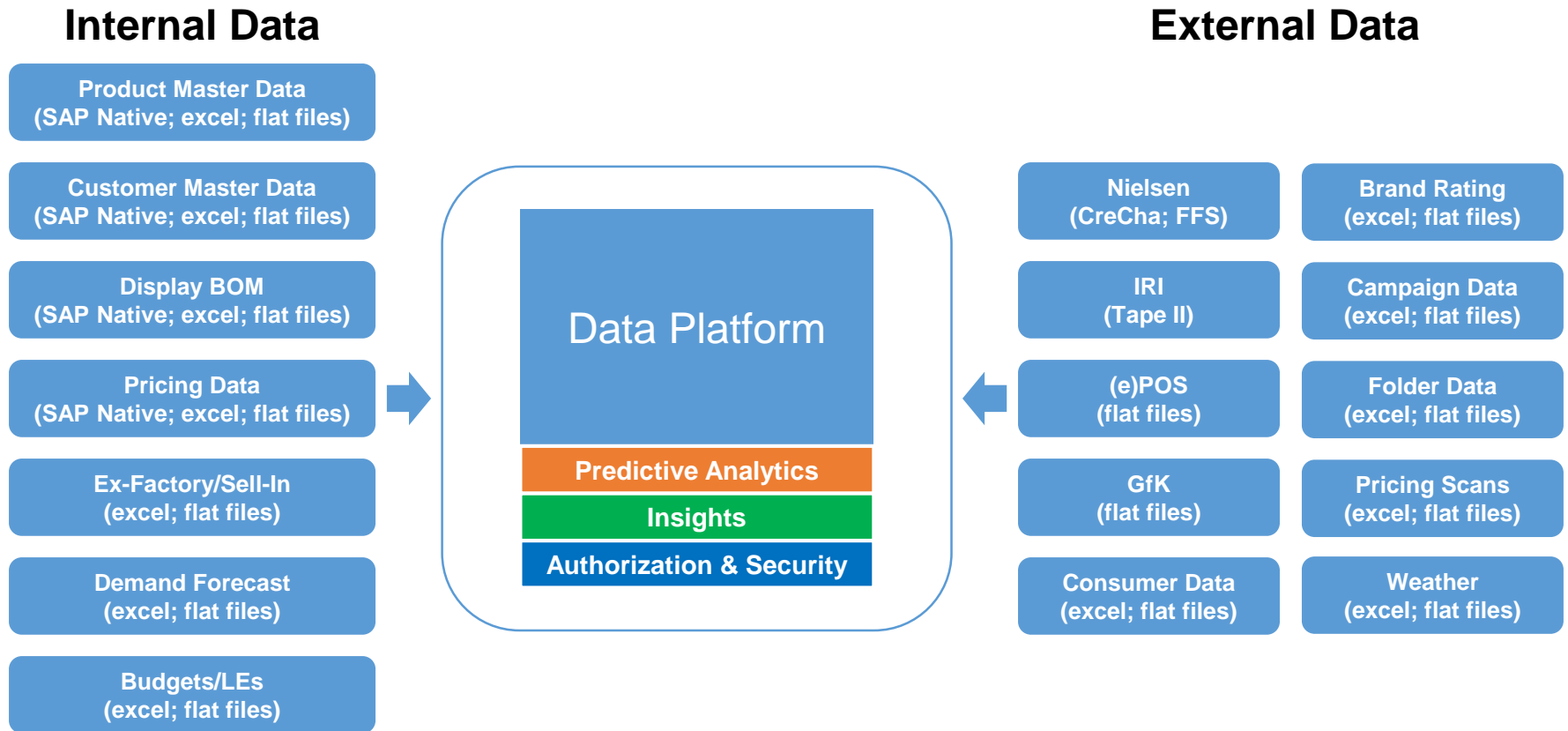
# Key Requirements for Successful AI



# Data is the fuel for AI



And a CPG organization can have a lot of diverse data available



Lack of external data does not make AI impossible, only less accurate



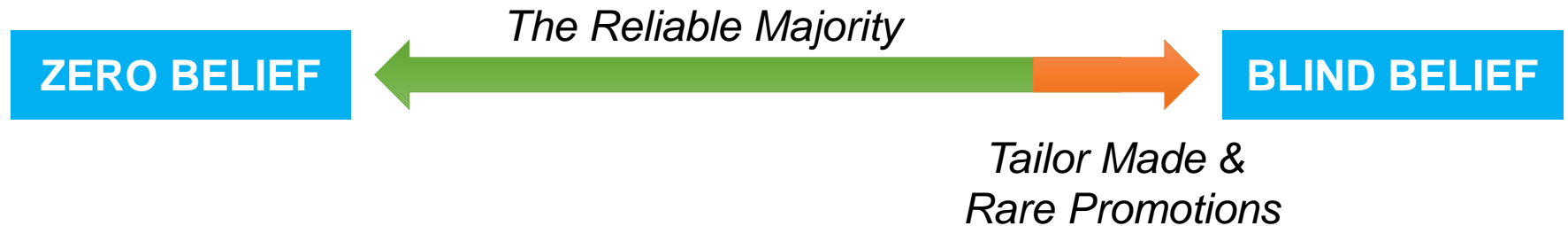
You need to be able to scale to any amount of data

Handle billions of records directly on-line for any client without performance degradation

	 NL	 SA	 NL	 DE
# Own Active Products	4,000	2,000	1,000	120
# Market Products	140,000	200,000	90,000	12,000
# Customers	4,000	12,000	6,000	2,000
# Plan Accounts	200	1,200	120	80
# Promotions	70,000	10,000	20,000	6,000
Avg # of Products per Promo	500	100	100	10
# Nielsen/IRI DBs	30	95	9	6
# ePOS Customers	2	0	2	0
Folder	Yes	Yes	Yes	Yes
Pricing Scans	Yes	No	Yes	Yes
# External Data Records	200 mln	500 mln	120 mln	60 mln
Data Allocation	Distributor	No	Distributor	Region / Banner

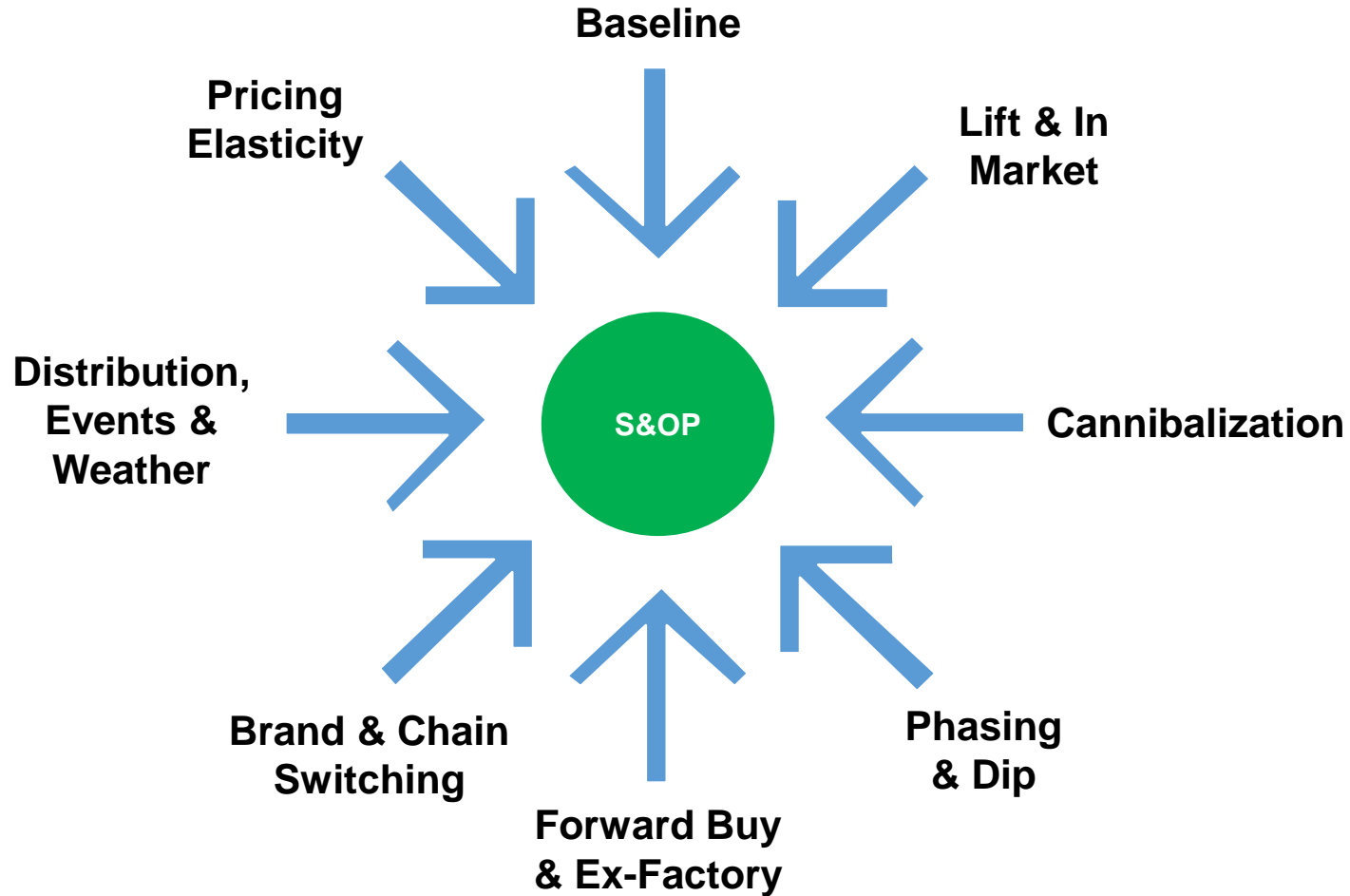
Users tend to take the extreme point of view with AI

# Trust the Machine?



There is a clear necessity to guide users on likely prediction reliability  
When is the user planning one of the 10% manual promotions?

There is no magic bullet...



...but an entire series of AI models working together...





# Prepare for **different** ways of working



# CPG leadership needs to make the difference...

...are you prepared to make a fundamental shift in how you work and on which platforms and enabling technologies

## The Opportunity

**First movers** can now sustain years of competitive advantage.

Better promotions, more control & insight, less resources



**visualfabriq**  
Let's unleash your excellence

Data mining

Algorithms

Artificial intelligence

**MACHINE LEARNING**

### Trust the machine

It does know it better

On a personal level we faithfully follow Google Maps without any concern. We marvel at voice recognition capability and happily depend on our tech assistants. Driverless vehicles make more and more sense to us. Over the past years, we've all been personally handing over more and more to the machine because it has become increasingly evident that is not just perfectly safe to do so, but better for us.

*Written by: Jaco Brusse, CEO visualfabriq*

Disclaimer  
The information in this document is intended solely for personal reference. Such information is subject to change without notice, and although we make all reasonable efforts to ensure the correctness of the presentation, its accuracy is not guaranteed and it may not contain all material information concerning Visualfabriq (the "company"). The company makes no representation regarding, and assumes no responsibility or liability for, the accuracy or completeness of, or any errors or omission in, any information contained herein.

## Summarizing...



### THE WATCH OUTS FOR SUCCESSFUL APPLIED AI



Be enthusiastic but realistic



Do not believe claims with external data and predictions too quickly, ask for proof



If evaluation data and planning are in separate systems, you will have issues with input consistency and direct application of models in planning



If the system cannot directly integrate predictive models, it will not be able to react to users and support them directly



There is no magic (yet); good AI models need persons with data knowledge and market experience



Change Management and Leadership will be key to really embrace a new way of working built on AI

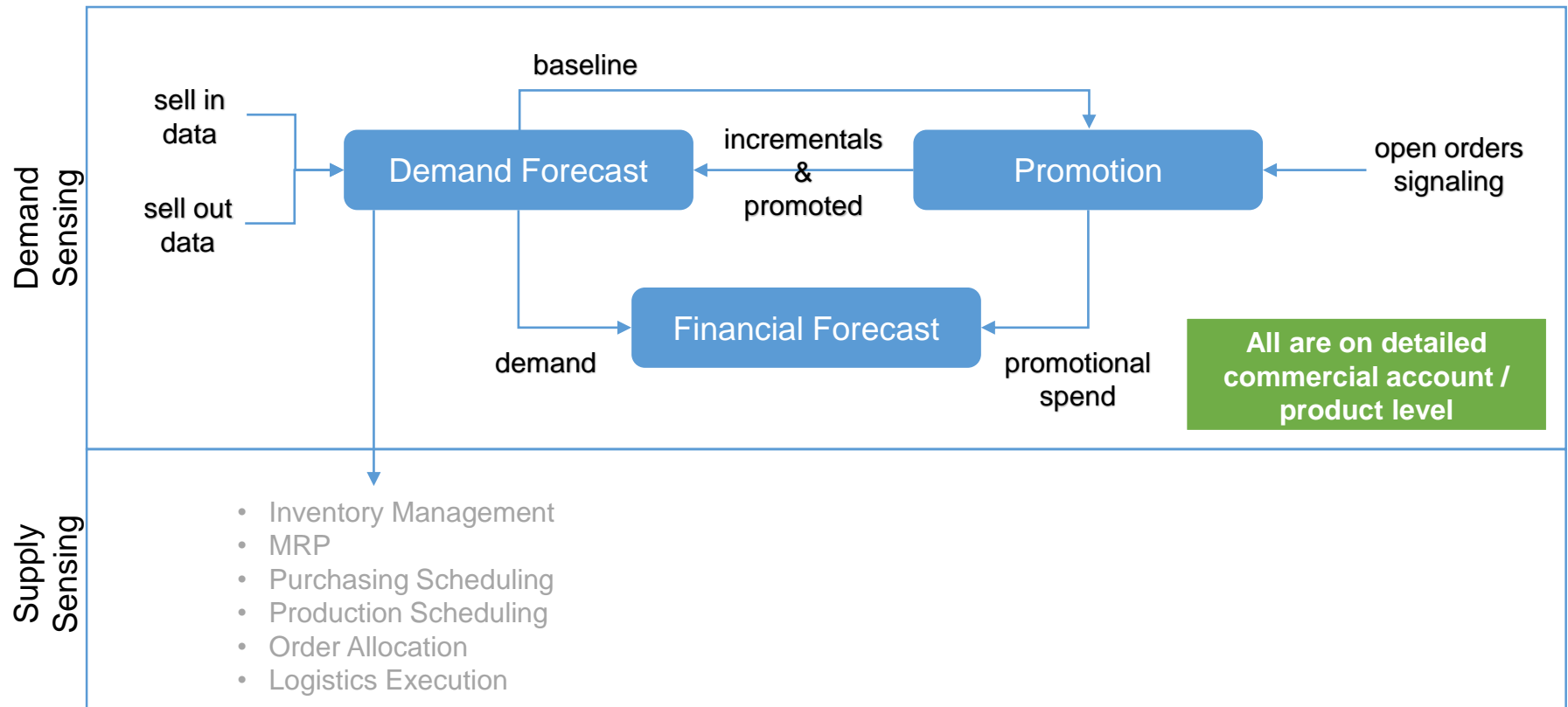
Any of these can prevent you from fully benefiting from AI in your processes



“

# Applied AI at visualfabriq

# Our Platform addresses a combination of key CPG issues

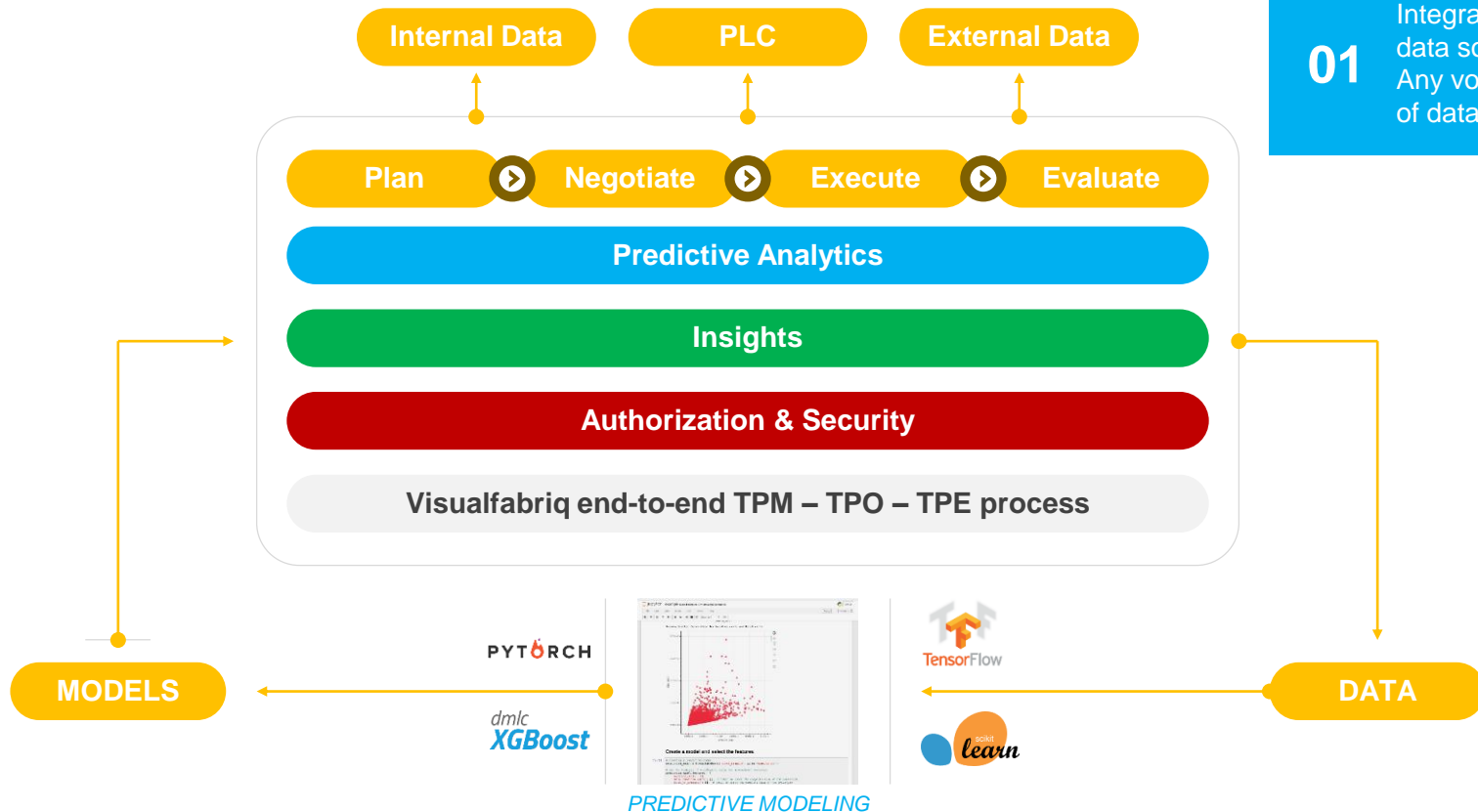


# And we are the only vendor with direct, applied AI

## BENEFITS

01

Integrate any data source  
Any volume of data



04

Industrialize models with one click Available to users directly

03

Use leading AI models

02

No manual data preparation  
No manual cleansing  
Save 90% of the work

# With JupyterLab our clients get the golden standard in data sciences for developing models

Leading. Powerful. Liked and Appreciated by Data Scientists.



## About Project Jupyter

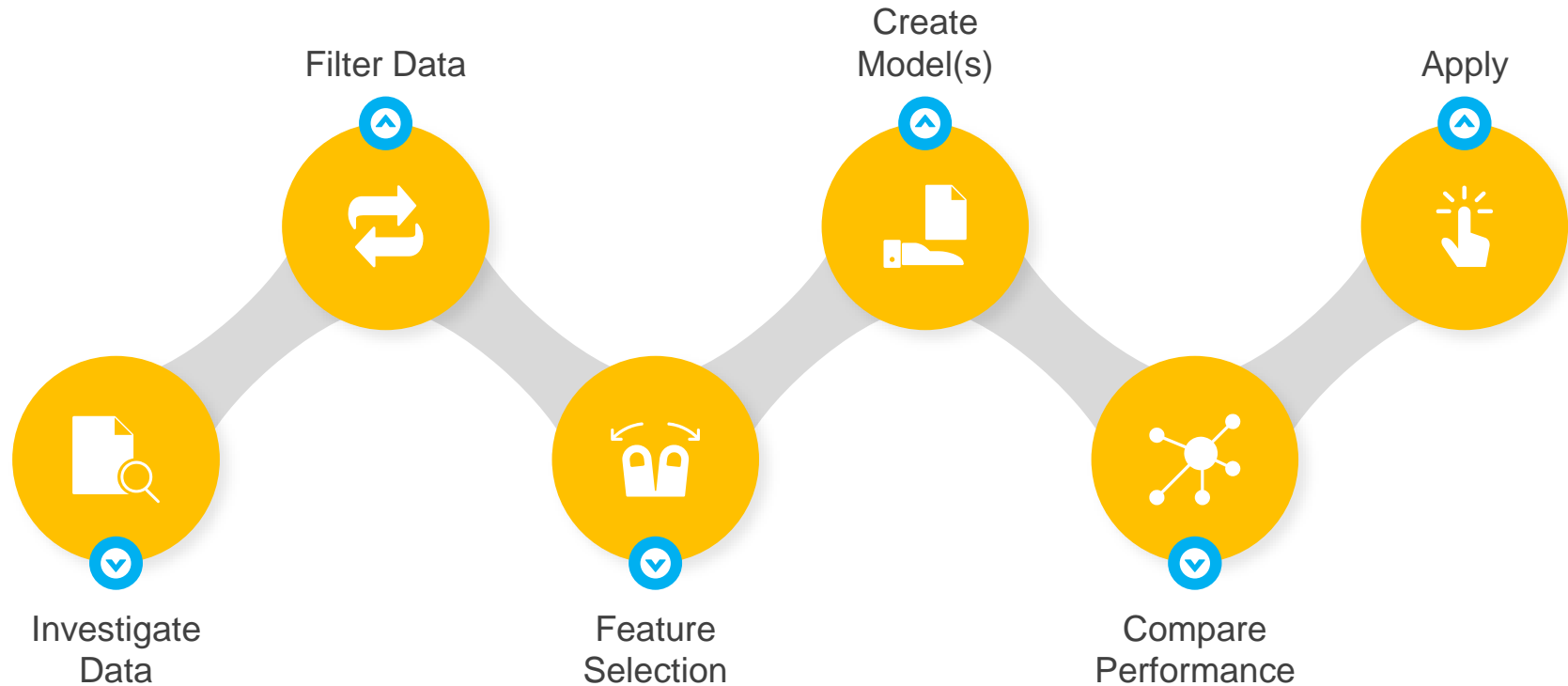
Project Jupyter is a non-profit, open-source project, born out of the IPython Project in 2014 as it evolved to support interactive data science and scientific computing across all programming languages. Jupyter will always be 100% open-source software, free for all to use and released under the liberal terms of the modified BDS license.

Jupyter is developed in the open GitHub, through the consensus of the community. For more information on our governance approach, please see our [Governance Document](#).

All online and in-person interactions and communications directly related to the project are covered by the [Jupyter Code of Conduct](#). This Code of Conduct sets expectations to enable a diverse community of users and contributors to participate in the project with respect and safety.

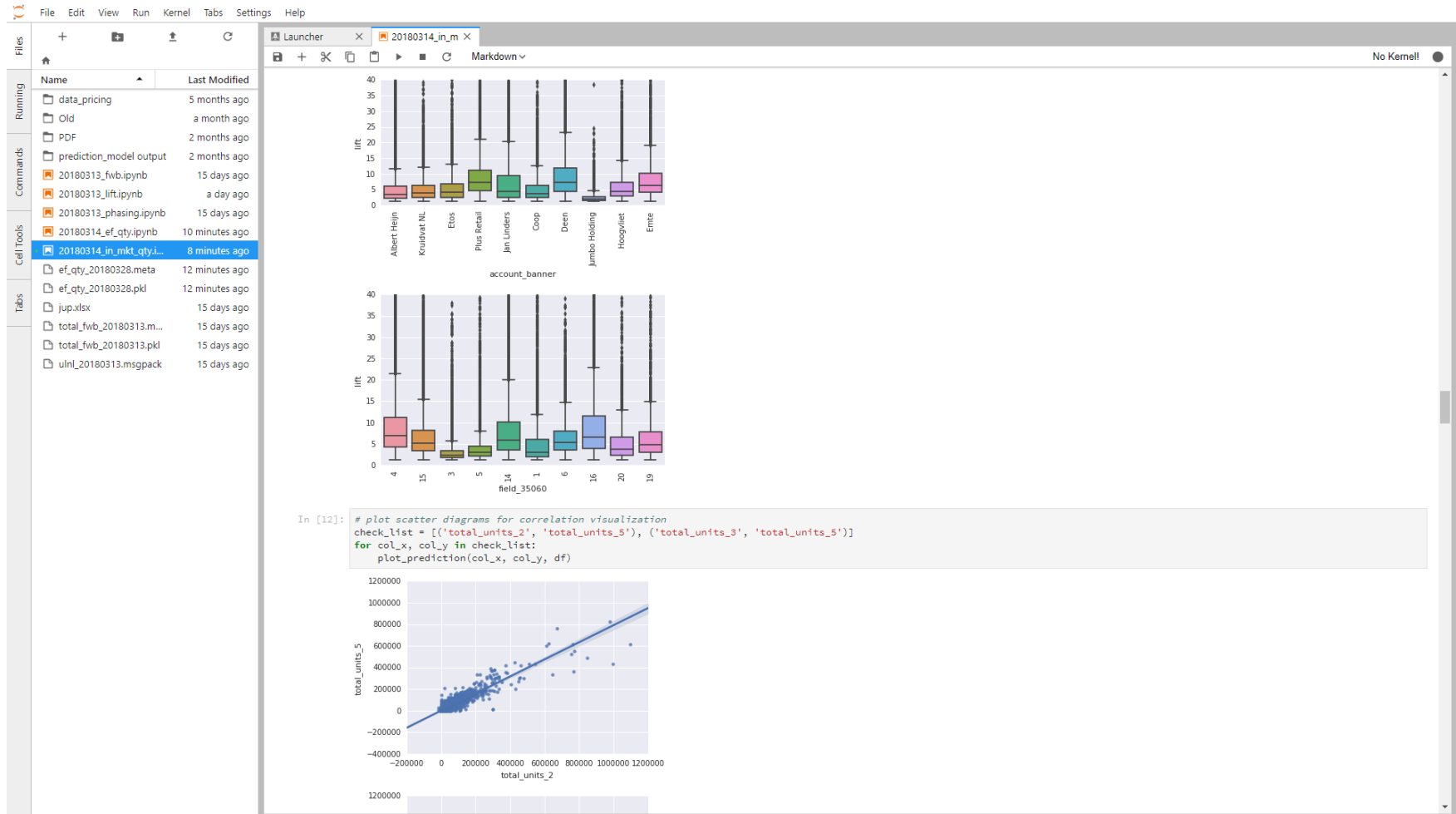
# We follow a - 6 step - proven and structured approach

We supply clients with standardized data sets and JupyterLab notebook templates





# The models are multi-variate and can handle many features



# Cleaning the data to relevant records is fast and straight forward

The screenshot shows a Jupyter Notebook with a file explorer on the left and a code editor on the right. The file explorer lists several files, including '20180313\_lift.ipynb' which is currently selected. The code editor displays a Python script for data filtering, and the output shows the results of each filter step.

```
future_df = future_df[baseline_mask]
print("\nNr of Future Records after baseline filter: %d rows with %d features." % (future_df.shape[0], future_df.shape[1]))

# filtering the contents
status_mask = df['promotion_status'] >= 112
df = df[status_mask]
print("\nAfter status filter: We have %d df with %d features." % (df.shape[0], df.shape[1]))

date_mask = df['yearweek'] >= 201612
df = df[date_mask]
print("\nAfter week filter: %d rows with %d features." % (df.shape[0], df.shape[1]))

small_df_mask = df['baseline_units'] > 0.0
df = df[small_df_mask]
print("\nAfter >0 baseline filter: %d rows with %d features." % (df.shape[0], df.shape[1]))

#huge_df_mask = df['baseline_units'] < 10.0**5
#df = df[huge_df_mask]
#print("\nAfter <10k baseline filter: %d rows with %d features." % (df.shape[0], df.shape[1]))

small_lift_mask = (df['lift'] > 1.2) & (df['lift'] < 40)
df = df[small_lift_mask]
print("\nAfter lift filter: %d rows with %d features." % (df.shape[0], df.shape[1]))

min_discount_mask = (df['discount_perc'] >= 5.0) & (df['discount_perc'] < 75.0)
df = df[min_discount_mask]
print("\nAfter discount filter: %d rows with %d features." % (df.shape[0], df.shape[1]))

mechanism_mask = df['mechanism'].notnull()
df = df[mechanism_mask]
print("\nAfter mechanism filter: %d rows with %d features." % (df.shape[0], df.shape[1]))

#distribution_normal_mask = df['weighted_distribution_normal'] > 5.0
#df = df[distribution_normal_mask]
#print("\nAfter distribution normal filter: %d rows with %d features." % (df.shape[0], df.shape[1]))

Nr of Excluded Records Removed: 692.

After promotion_type filter: 288812 rows with 139 features.

Nr of Future Records: 111408 rows with 139 features.

Nr of Future Records after week filter: 96379 rows with 139 features.

Nr of Future Records after baseline filter: 71232 rows with 139 features.

After status filter: We have 166667 df with 139 features.

After week filter: 166667 rows with 139 features.

After >0 baseline filter: 137232 rows with 139 features.

After lift filter: 112572 rows with 139 features.

After discount filter: 104225 rows with 139 features.

After mechanism filter: 104225 rows with 139 features.
```

This is about striking a balance:

- You want to filter out outliers (not optimize for the weird situations)
- But less data means you will be worse for the filtered data (that you will have to predict in real life)

# Visualfabriq helps clients with automated feature selection

The screenshot displays the Visualfabriq application interface. On the left, a file explorer shows a list of files and folders, including 'data\_pricing', 'Old', 'PDF', 'prediction\_model output', and several IPYNB files. The file '20180314\_in\_mkt\_qty.i...' is selected. The main window on the right shows the results of a feature selection process. It lists the 30 most important features in descending order of importance, with 'discount\_perc' being the most important feature (importance 100.0).

The 30 most important features are (in descending order):

Feature	Importance
discount_perc	100.0
total_baseline_units	29.87501583207251
multi_buy_y	29.46558234781471
second_placement_perc	25.02379538743049
second_placement_yn	16.22992105219977
total_nr_products	16.08932080933273
discount_amt	11.380897469192636
consumer_length	9.909287488324262
mechanism_type	9.24644573017625
planned_discount_amt	8.286220836617153
field_35060	7.322786255314786
planned_discount_perc	6.181457476187908
planned_base_price	5.670938135929294
template	5.406082056730605
baseline_vol	4.7212168404079655
baseline_units	4.457171339598108
ef_base_price	3.9924639489548213
planned_ef_base_price	3.988548072663035
ef_promoted_price	3.526254915544128
pre_fwb_qty	3.4179304540022044
total_ef_qty	3.3236260102866515
baseline_units_ext	3.3109994008917214
planned_ef_promoted_price	3.034419876345116
baseline_units_int	2.994586988704439
planned_ef_discount_amt	2.8748155645228106
yearweek	2.033470507683652
post_fwb_qty	1.5721439097698762
week	1.4666142277183578
week_agg_2	1.4626132971495318
week_agg_4	1.4515423187708307

Please do not forget to manually check for co-correlation (such as baseline\_units and baseline\_vol)

**Try Automated Feature Selection**

# All leading AI models can be used

Files

Name	Last Modified
data_pricing	5 months ago
Old	a month ago
PDF	2 months ago
prediction_model output	2 months ago
20180313_fwbp.py	15 days ago
20180313_lift.py	a day ago
20180313_phasing.py	15 days ago
20180314_ef_qty.py	24 minutes ago
20180314_in_mkt_qty...	7 minutes ago
ef_qty_20180328.meta	25 minutes ago
ef_qty_20180328.pkl	25 minutes ago
jup.xlsx	15 days ago
total_fwbp_20180313.m...	15 days ago
total_fwbp_20180313.pkl	15 days ago
u1n1_20180313.msppack	15 days ago

Running

Commands

Cell Tools

Tabs

Launcher

20180313\_lift.py × 20180314\_in\_m × 20180314\_ef\_qt ×

Markdown

```

# )
# regressor = ensemble.AdaBoostRegressor(ensemble.ExtraTreesRegressor(n_estimators=20, n_jobs=-1), n_estimators=79)
regressor = ensemble.GradientBoostingRegressor(
    n_estimators=150, # 300 how many optimizations do we do
    subsample=0.3, # 0.3 how much data should we use in every optimization run (standard 1.0, but lower values can prevent overfitting)
    learning_rate=0.1, # 0.1 how much should each additional tree shrink in weight (standard 0.1)
    max_depth=20, # 20 what is the max depth (standard 5)
    min_samples_split=100, # 100 what nr (int) or % (float, 100% = 1.0) of observations do we need for a split in the forest
    min_samples_leaf=20, # 20 what nr (int) or % (float, 100% = 1.0) of observations do we need for a leaf in the forest
    max_features=len(prediction_model.features), # do this if you want to limit the amount of features relating to the data
)

# prepare the input
prediction_model.model = regressor

# possibly check: https://github.com/autml/auto-sklearn/blob/master/example/example_regression.py
# Use LightGBM GradientBoostingRegressor https://github.com/Microsoft/LightGBM

In [83]: scores = cross_val_score(prediction_model.model, predict_df, target_df, cv=5)
print("Accuracy: %.2f (+/- %.2f)" % (scores.mean(), scores.std() * 2))

Accuracy: 0.91 (+/- 0.03)

In [84]: # now actually train the model on the entire data set
prediction_model.model.fit(predict_df, target_df)
# save the model
prediction_model.save()

In [76]: # predict the lift and check additional scoring
predict_ef_qty = prediction_model.model.predict(predict_df)

if log:
    # if it was logarithmic, expand the lift again
    predict_ef_qty = np.exp(predict_lm_qty)

predict_ef_qty[predict_ef_qty == np.inf] = np.nan
# score the model
score_model(predict_ef_qty, target_df)
print('Forecast Error: ' + str(100.0 * np.sum(np.abs(target_df - predict_ef_qty)) / len(target_df)))

print('\nCompared to the version 3 actuals:')
score_model(df['total_ef_qty_3'], target_df)
print('Forecast Error Actuals: ' + str(100.0 * np.sum(np.abs(target_df - predict_ef_qty_3)) / len(target_df)))

R^2: 0.9419453238680193
MAE: 1338.3663649540438
MAPE: inf
Forecast Error: 22.20202283582475

Compared to the version 3 actuals:
R^2: 0.8788368801090273
MAE: 2394.5608679315646
MAPE: inf
Forecast Error Actuals: 38.12635545306076

```

**Predict and score the model**

Price vs Size

$\theta_0 + \theta_1 x$

High bias  
(underfit)

Price vs Size

$\theta_0 + \theta_1 x + \theta_2 x^2$

"Just right"

Price vs Size

$\theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 + \theta_4 x^4$

High variance  
(overfit)

**Managing overfitting vs robustness is the key to creating reliable and accurate predictions**

Typically used libraries

PyTorch

TensorFlow

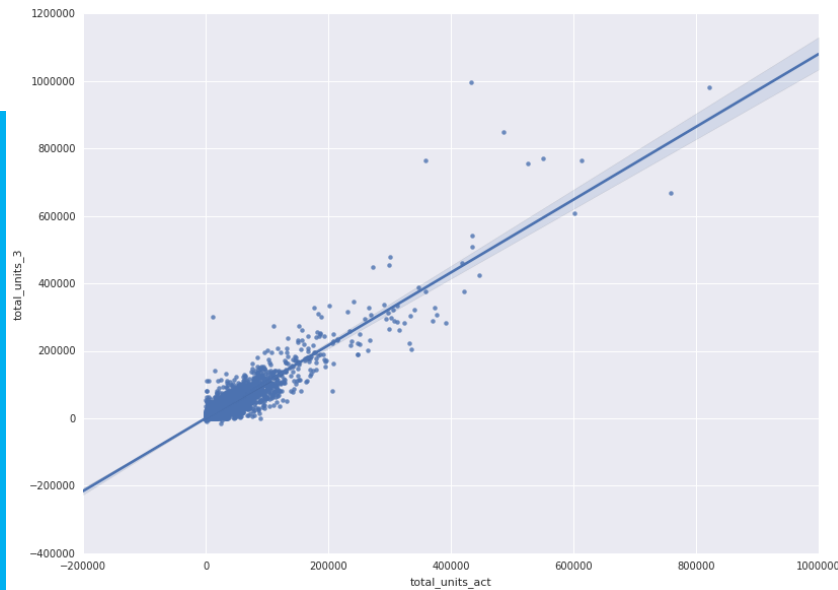
dmlc XGBoost

StatsModels

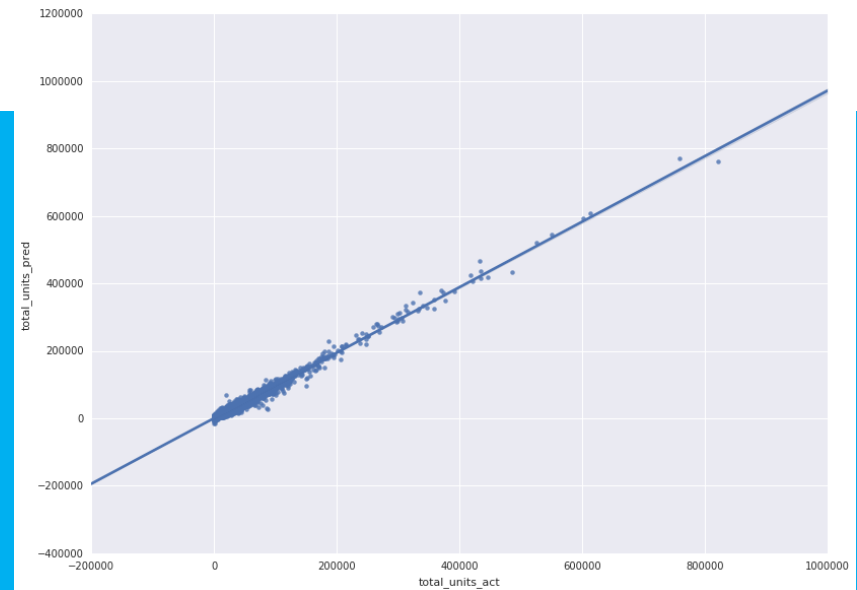
LightGBM

# Models can be compared to human estimations and other models

An example of in-market promotion effect estimations



Human Planning



AI Planning

# Start to move to AI based promotion planning

Always up-to-date, person-independent and reliable



## Enter Zero Touch Planning for Promotions

- The system looks at different scenarios for you and suggest the best options
- It will tell you how reliable the prediction will be
- All operational steps are fully automated and guided by AI, leaving only the negotiation and approval cycles for the users

**THIS IS ALL AVAILABLE NOW IN 2018**

# Views on AI in the TPx space from the leading analysts



Algorithms and machine learning will ultimately outpace traditional methods in defining business best practices and process models. Technologies already exist to provide complete traceability from raw material to consumption, but consumer-driven enterprises still struggle to fully leverage this capability. Technologies (e.g., blockchain) already exist to provide complete traceability from raw material to consumption, but consumer-driven enterprises still struggle to fully leverage this capability.

To qualify as a comprehensive TPO solution, the solution must enable key predictive or prescriptive capabilities. These include the specification of constraints, iterative scenario planning, recommendations for effective promotions and allowances for cannibalization and halo effects.

Substantial data integration will be required, as will modeling and domain expertise.



Deeper analytics, better user experience, and capabilities for optimizing promotions through predictive models are key differentiators that are highly sought after.

The ability to articulate and deliver against a vision of where the TPx space is going in the two-to-five-year time horizon and beyond for its targeted geographies and tiers of customers.

POI about Visualfabriq:

- › Very positive growth trend and a strong roadmap going forward. Investment in people and product have been steady. The company has consistently been able to win deals with companies that have announced other vendors as the global standard. This is particularly true in the Netherlands. We believe that the work it is doing on machine learning will pay dividends as an early mover advantage.
- › A very robust set of administrative tools that should allow the user organization to have full control of the solution. This includes the ability to get inside of the predictive models (like Python) using Jupyter to facilitate the modeling in a repeatable way. Ability to handle significant amounts of data. Everybody makes this claim, however we watched the system load 15,000 promotions in just 30 seconds. Visualfabriq acts like a very mature software company by offering 2 major and 2 minor releases per year.



With visualfabriq's  
applied AI we  
ensure the  
following:

Ability to handle up to large amounts  
of very diverse data for each market



Access to the leading machine  
learning & deep learning ai models



The ability to directly industrialize a  
predictive model directly in your  
processes





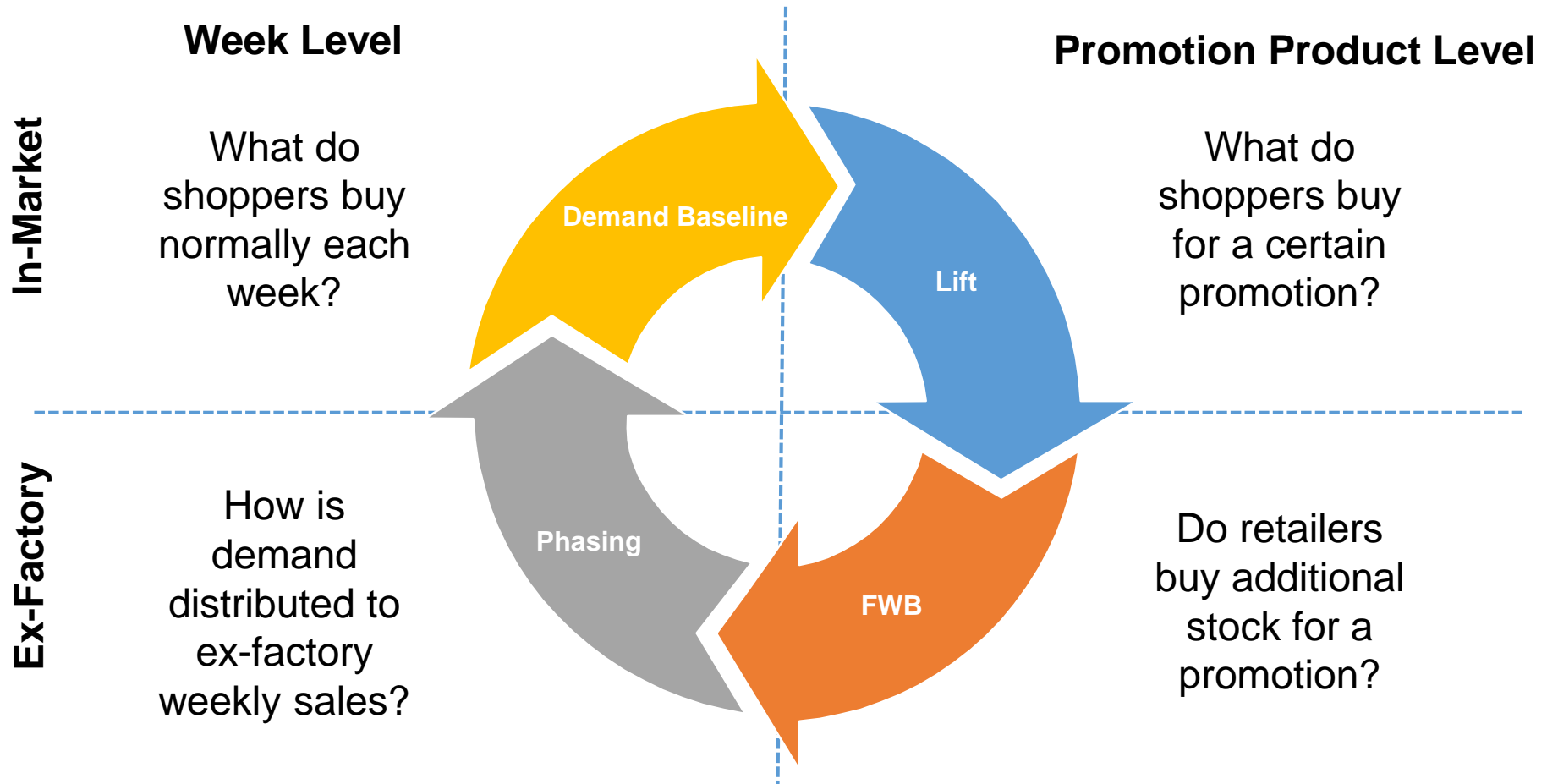
“

# The ROI of AI

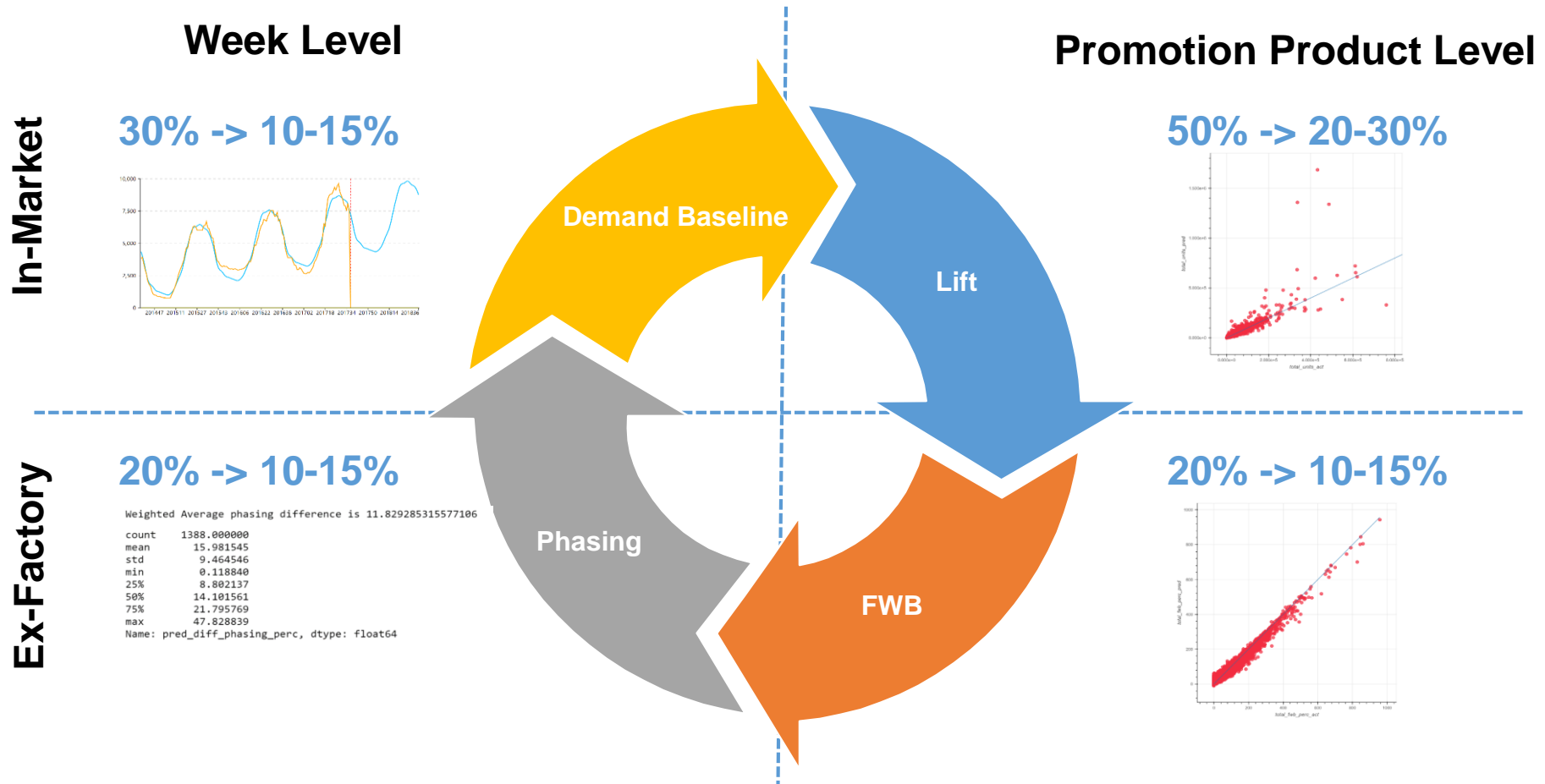
Before you start, set ambitious but feasible goals

	Baselines	Promotions
Automatic	95%	90%
Accurate	90%	75%
Improve ROI		10%

# AI can deliver enhancements to many parts of the process

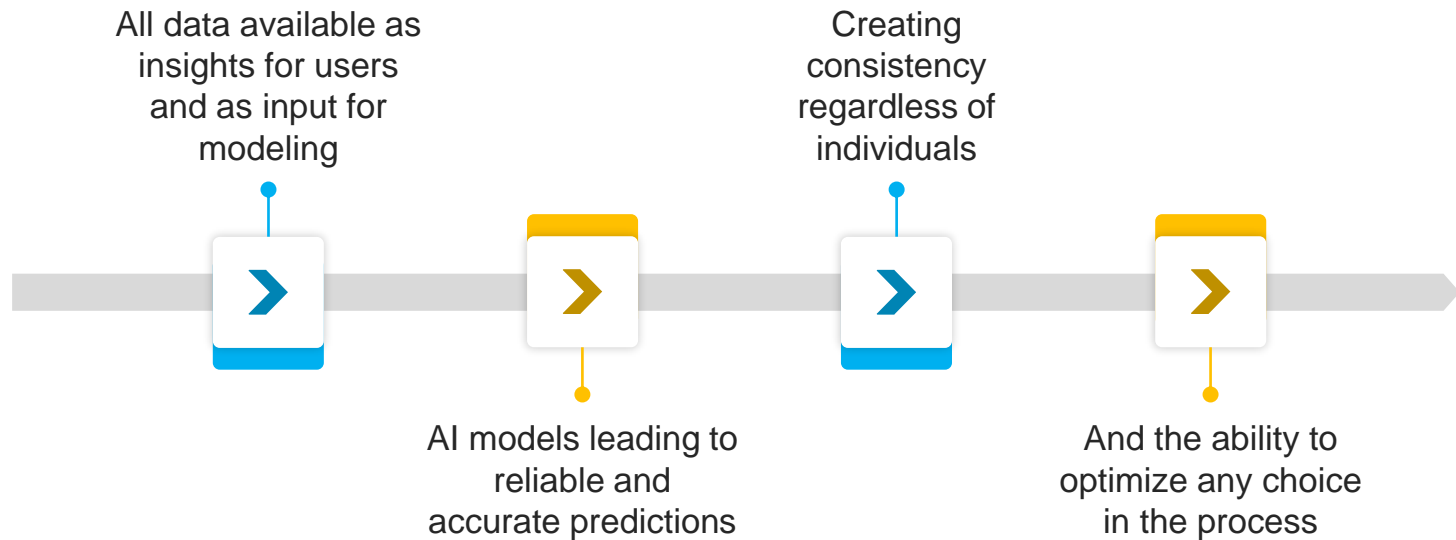


# AI models are consistently twice as accurate as human estimations\*



\*Human and prediction accuracies are based on averages of various categories and clients in over 10 markets

# Realize that AI is not a goal but a journey



But one that will deliver results along every step of the way



“

# Questions?





## ***Disclaimer***

The information contained in this presentation is intended solely for personal reference. Such information is subject to change without notice, and although we made all reasonable efforts to ensure the correctness of this presentation, its accuracy is not guaranteed and it may not contain all material information concerning Visualfabriq (the “Company”). The Company makes no representation regarding, and assumes no responsibility or liability for, the accuracy or completeness of, or any errors or omissions in, any information contained herein.



# visualfabriq

Let's unleash your excellence