

Make Faster, Smarter Decisions by Combining IBM Machine Learning and IBM Decision Optimization

—
Susara van den Heever
IBM Data Science Practice Leader



Please note

IBM's statements regarding its plans, directions, and intent are subject to change or withdrawal without notice and at IBM's sole discretion.

Information regarding potential future products is intended to outline our general product direction and it should not be relied on in making a purchasing decision.

The information mentioned regarding potential future products is not a commitment, promise, or legal obligation to deliver any material, code or functionality. Information about potential future products may not be incorporated into any contract.

The development, release, and timing of any future features or functionality described for our products remains at our sole discretion.

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors, including considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve results similar to those stated here.

Notices and disclaimers

© 2018 International Business Machines Corporation. No part of this document may be reproduced or transmitted in any form without written permission from IBM.

U.S. Government Users Restricted Rights – use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM.

Information in these presentations (including information relating to products that have not yet been announced by IBM) has been reviewed for accuracy as of the date of initial publication and could include unintentional technical or typographical errors. IBM shall have no responsibility to update this information. **This document is distributed “as is” without any warranty, either express or implied. In no event, shall IBM be liable for any damage arising from the use of this information, including but not limited to, loss of data, business interruption, loss of profit or loss of opportunity.** IBM products and services are warranted per the terms and conditions of the agreements under which they are provided.

IBM products are manufactured from new parts or new and used parts. In some cases, a product may not be new and may have been previously installed. Regardless, our warranty terms apply.”

Any statements regarding IBM's future direction, intent or product plans are subject to change or withdrawal without notice.

Performance data contained herein was generally obtained in a controlled, isolated environments. Customer examples are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual performance, cost, savings or other results in other operating environments may vary.

References in this document to IBM products, programs, or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business.

Workshops, sessions and associated materials may have been prepared by independent session speakers, and do not necessarily reflect the views of IBM. All materials and discussions are provided for informational purposes only, and are neither intended to, nor shall constitute legal or other guidance or advice to any individual participant or their specific situation.

It is the customer's responsibility to insure its own compliance with legal requirements and to obtain advice of competent legal counsel as to the identification and interpretation of any relevant laws and regulatory requirements that may affect the customer's business and any actions the customer may need to take to comply with such laws. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the customer follows any law.

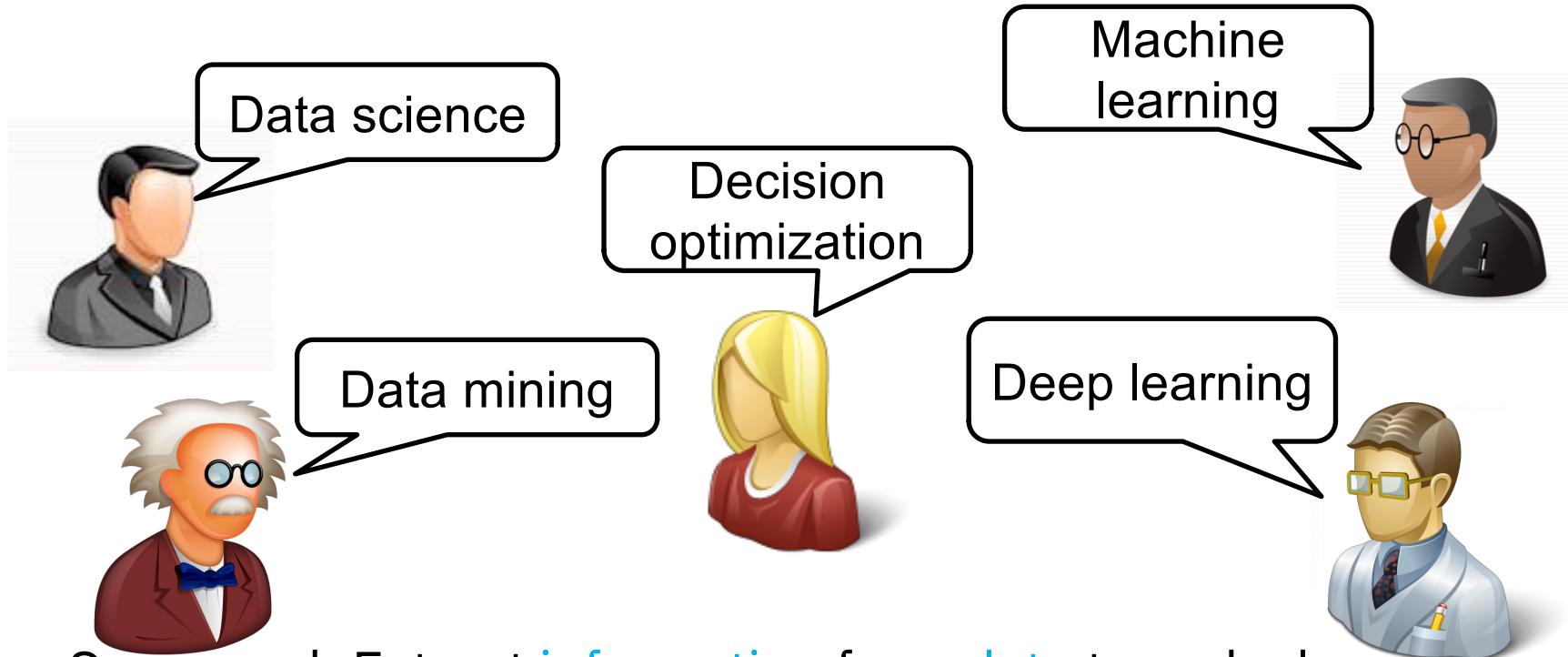
Notices and disclaimers continued

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products about this publication and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products. IBM does not warrant the quality of any third-party products, or the ability of any such third-party products to interoperate with IBM's products. **IBM expressly disclaims all warranties, expressed or implied, including but not limited to, the implied warranties of merchantability and fitness for a purpose.** The provision of the information contained herein is not intended to, and does not, grant any right or license under any IBM patents, copyrights, trademarks or other intellectual property right.

IBM, the IBM logo, ibm.com and [names of other referenced IBM products and services used in the presentation] are trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at: www.ibm.com/legal/copytrade.shtml.

Everyone's talking about it!

What is the Buzz about ?



Same goal: Extract **information** from **data** to make better,
faster, smarter **decisions** & provide better **client experiences**

You use data science & machine learning every day while driving your car

Descriptive analytics

*How fast am I driving?
How far have I driven*



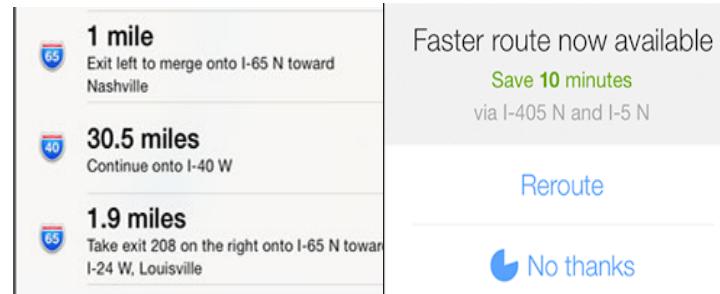
Predictive analytics

How long before I run out of gas?



Prescriptive analytics

What is the fastest route to my destination?



1 mile
Exit left to merge onto I-65 N toward Nashville

30.5 miles
Continue onto I-40 W

1.9 miles
Take exit 208 on the right onto I-65 N toward I-24 W, Louisville

Faster route now available
Save 10 minutes
via I-405 N and I-5 N

Reroute

No thanks

Since 1936, Data Science and Machine Learning delivered massive ROI

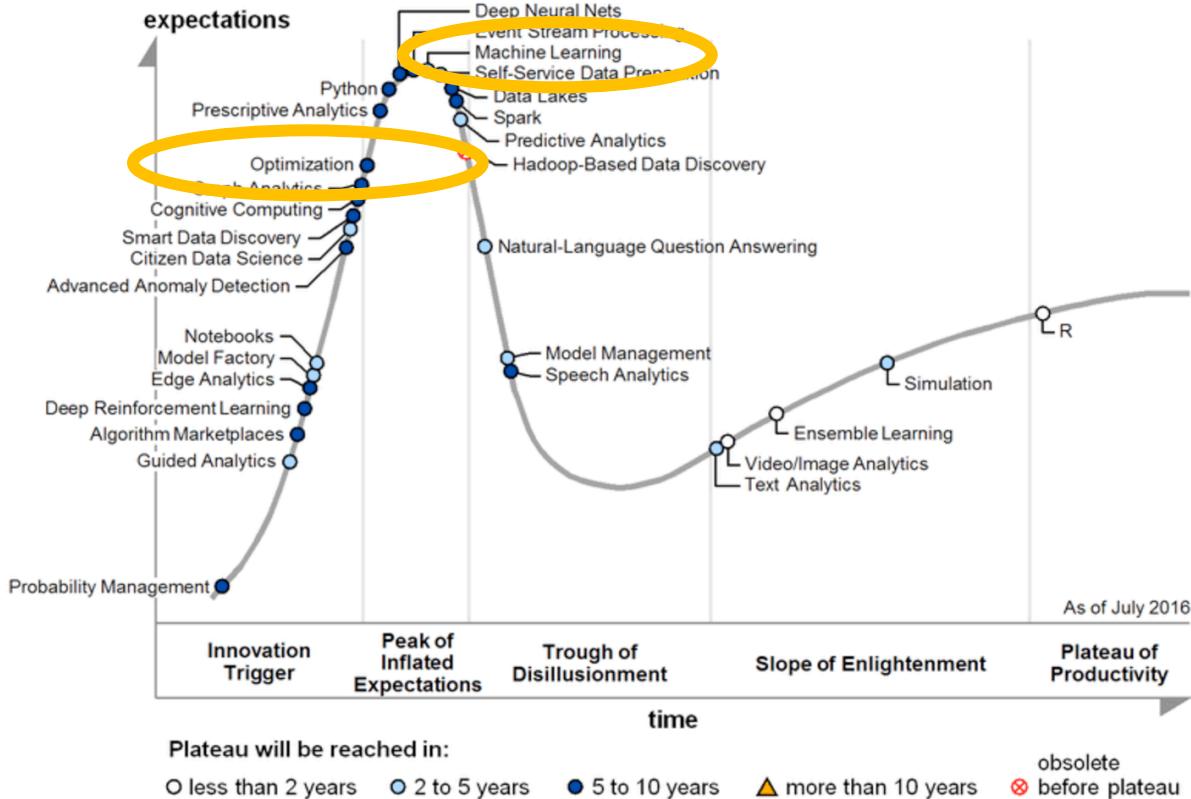
2 Chilean Forestry firms	Timber Harvesting	\$20M/yr + 30% fewer trucks
UPS	Air Network Design	\$40M/yr + 10% fewer planes
South African Defence	Force/Equip Planning	\$1.1B/yr
Motorola	Procurement Management	\$100M-150M/yr
Samsung Electronics	Semiconductor Manufacturing	50% reduction in cycle times
SNCF (French Railroad)	Scheduling & Pricing	\$16M/yr rev + 2% lower op ex
Continental Airlines	Crew Re-scheduling	\$40M/yr
AT&T	Network Recovery	35% reduction spare capacity
Grantham Mayo van Otterloo	Portfolio Optimization	\$4M/yr

Source: Edelman Finalists, <http://www.informs.org> or <http://www.scienceofbetter.org>

37 years, 37 Edelman winners, upwards of \$200 billion cumulative ROI

Gartner on Machine Learning & Optimization

Figure 1. Hype Cycle for Data Science, 2016

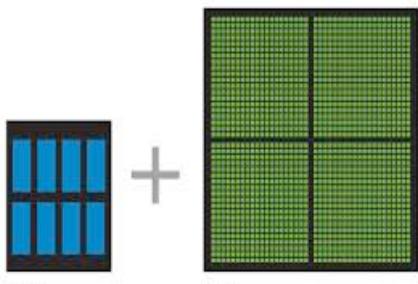


*"Even though the **optimization solutions market** is rapidly gaining momentum ...it is still in the emerging phase of maturity with a **market penetration of less than 10% of the possible target audience...**"*

*Gartner Market Guide for Optimization Solutions,
G00273712*

Researchers overcame 3 major challenges to data science adoption in the recent past

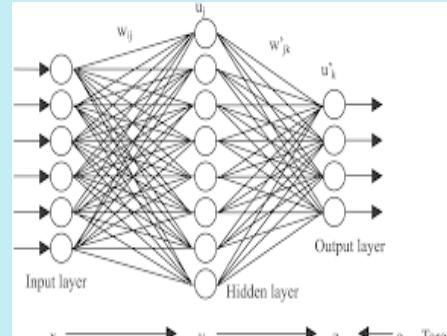
Scale



Falling hardware prices & GPU accelerators led to dramatic speed-ups

Pay-as-you-go cloud means infrequent expensive computation is now affordable

Models that learn & adapt



Algorithmic ML advances vastly improved accuracy & reliability

Optimization algorithms can solve millions of simultaneous decisions

“Analytics for everyone”



Karen Zack/@teenybiscuit

Analytics is all around us:

Facebook recognizes friends

Cars drive themselves

Cosmetic apps analyze your selfie & recommend treatments

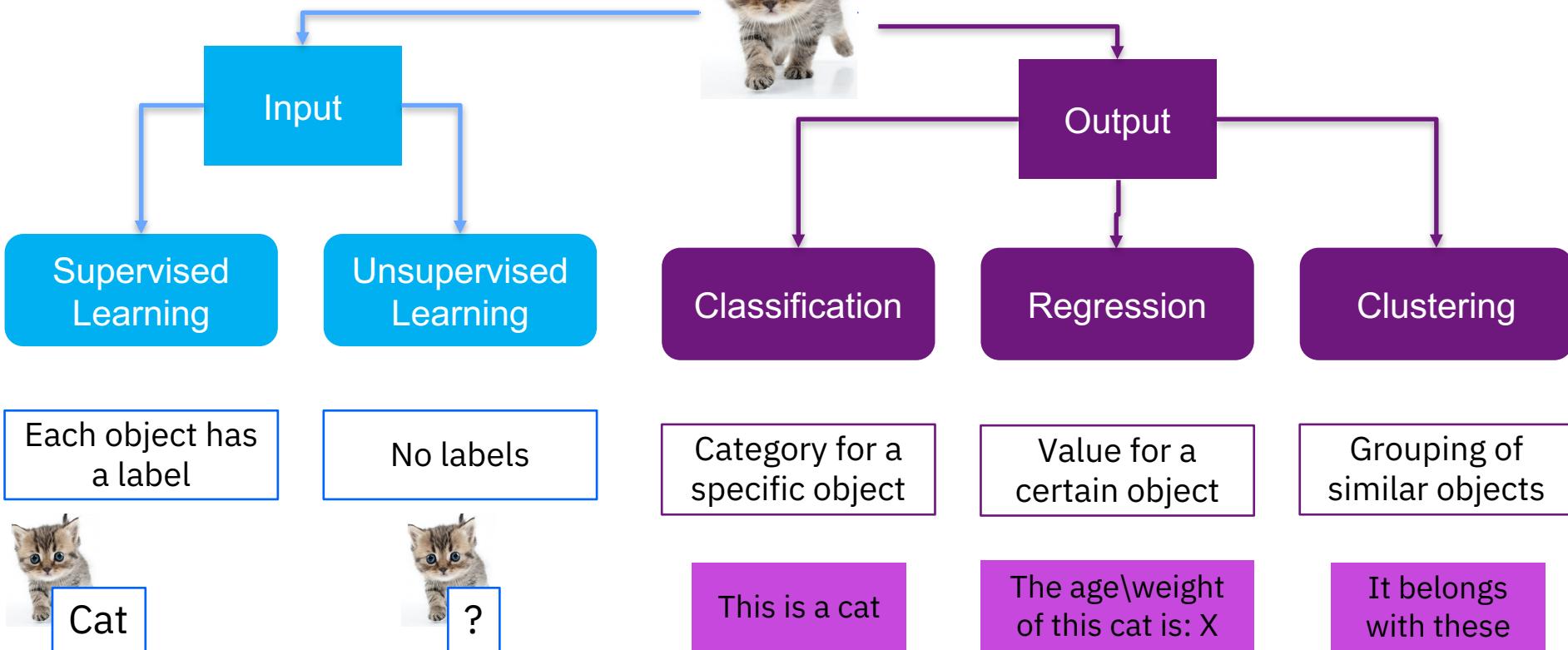
Machine Learning & Decision Optimization

- how does it work?

Machine Learning – Fundamentals

Input

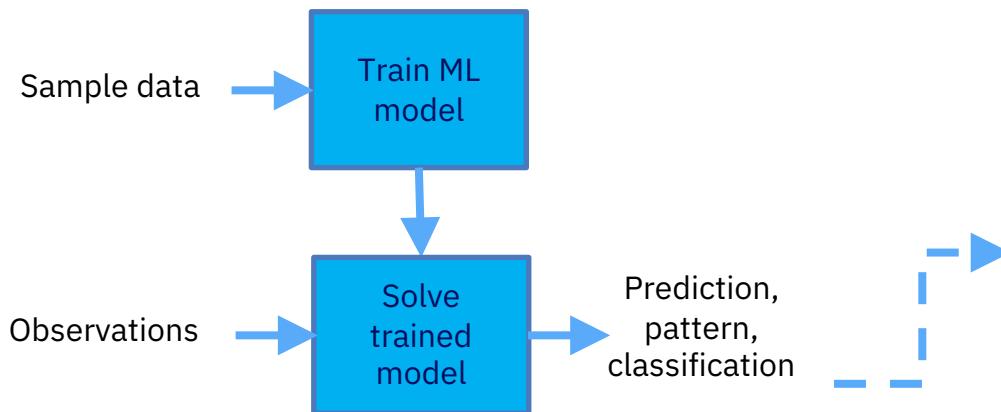
Image Audio Text DB Table



Machine Learning for Predictive vs Optimization for Prescriptive

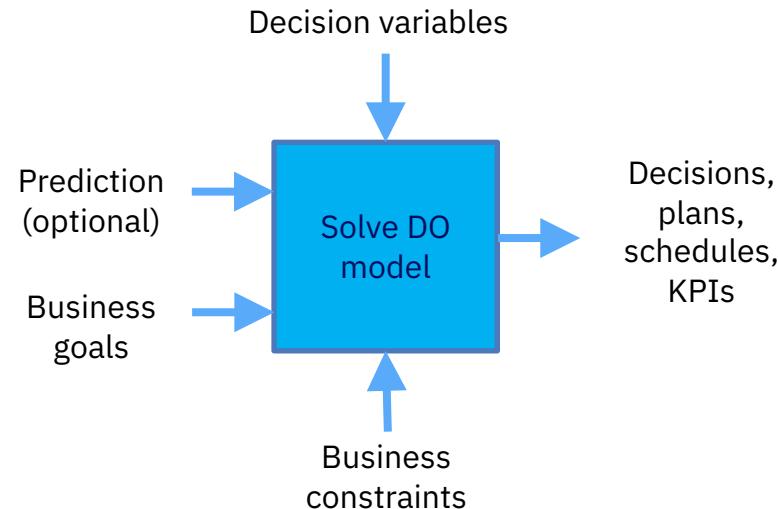
Machine Learning 101

- Basic (supervised): you **know the answer** and you **train the machine how to find it**
- More advanced – unsupervised, reinforcement, & deep learning

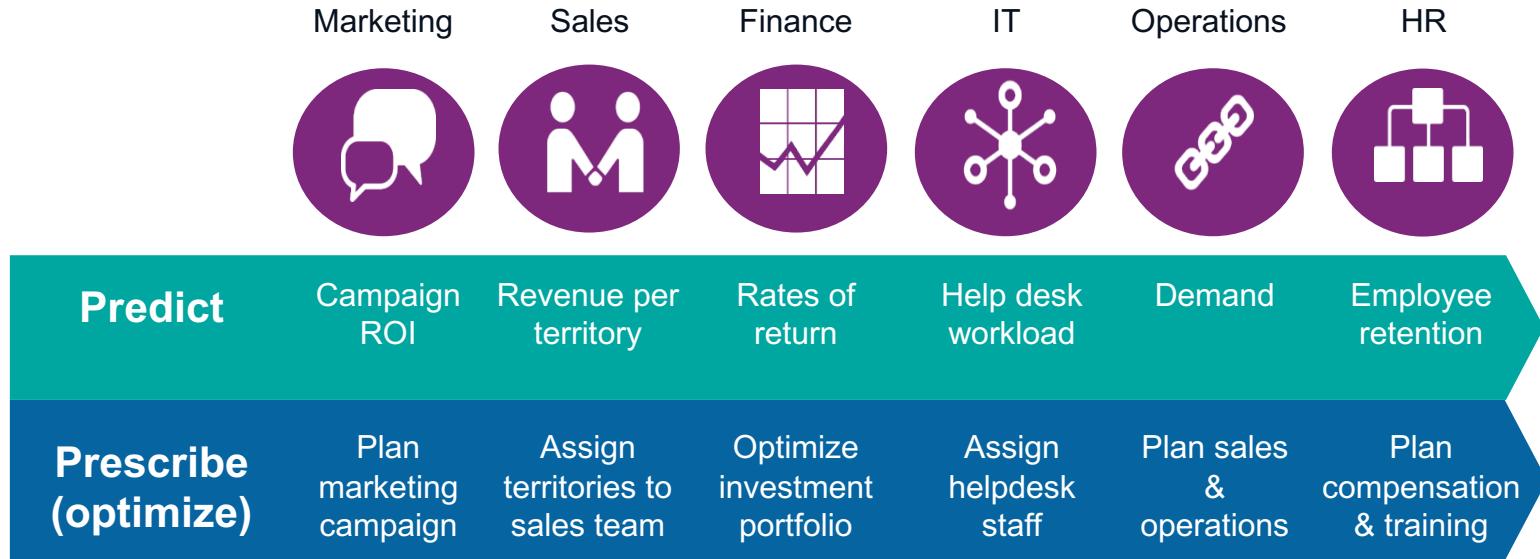


Decision Optimization 101

- You **don't know the answer**, and you **teach the machine the logic to find it**
- More advanced – robust/stochastic/...



Predictive vs Prescriptive: if you want to predict something, it's likely because you want to decide (prescribe) something



Important: many prescriptive problems don't involve a prediction

Use case: Marketing Campaign Optimization



Which product/campaign ?



Hundreds of products/campaigns Combinations with incompatibilities



Through which channel?

Decide best campaigns. Explore alternatives of different objectives/requirements using what-if analysis.



Different customer segments

To which customers?

You are a marketing campaign planner

– decide which offers to extend to which customers to maximize expected revenue

- **Input**

- Predicted revenue per customer, per offer
(output from a machine learning model)

- **Decisions**

- Select up to one offer for each customer

- **Rules or Constraints**

- Each offer can be used at most 3 times

- **Objective**

- Maximize total expected revenue

Customers	Mortgage	Savings	Pension
1	70	60	60
2	80	80	70
3	90	90	80
4	50	50	50
5	100	100	100
6	110	130	150
7	20	20	90
8	10	40	80
9	0	50	60
10	40	40	80

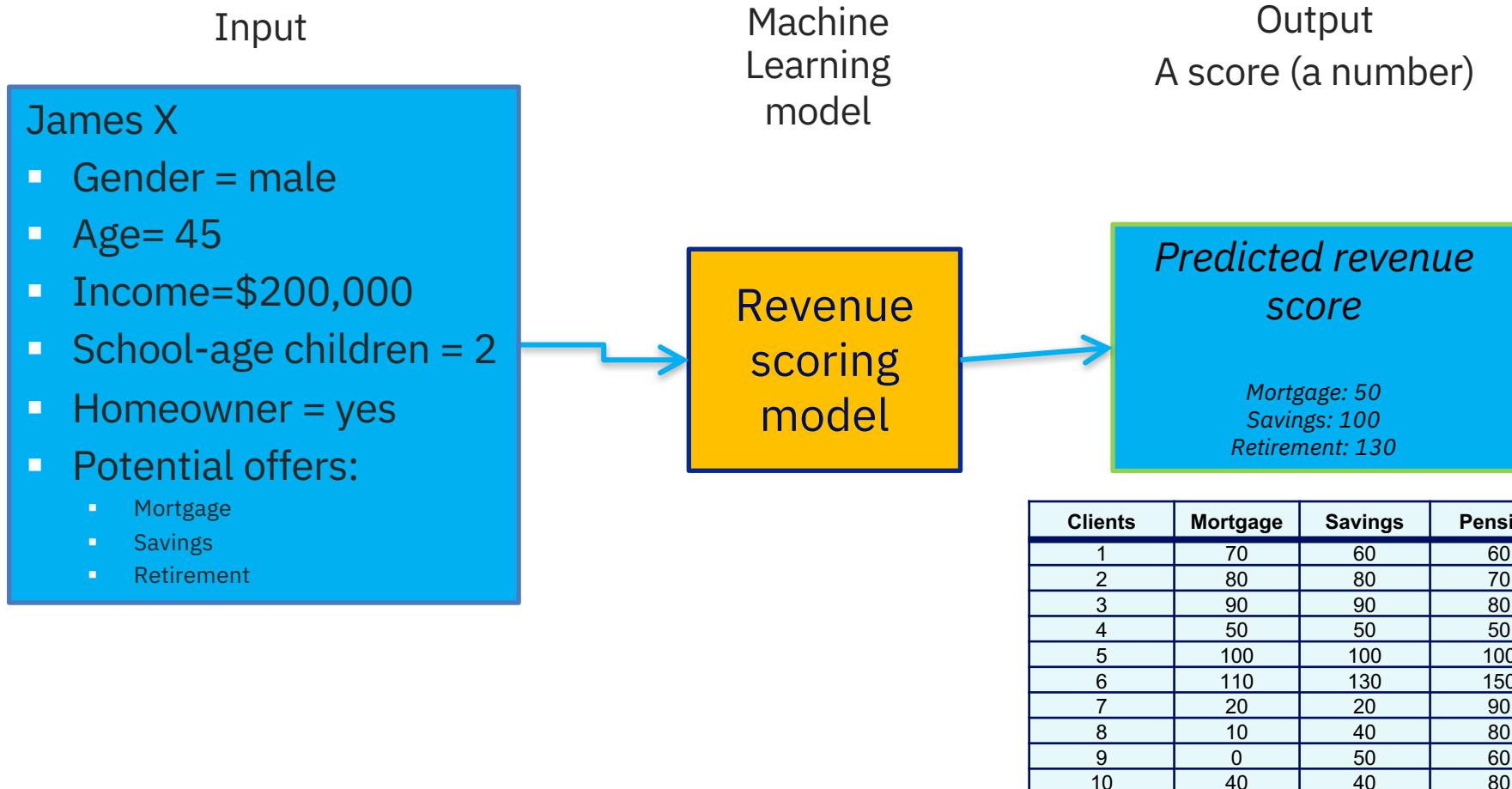
Expected revenue of offer("Savings") to customer("9") = 50

How did you do?

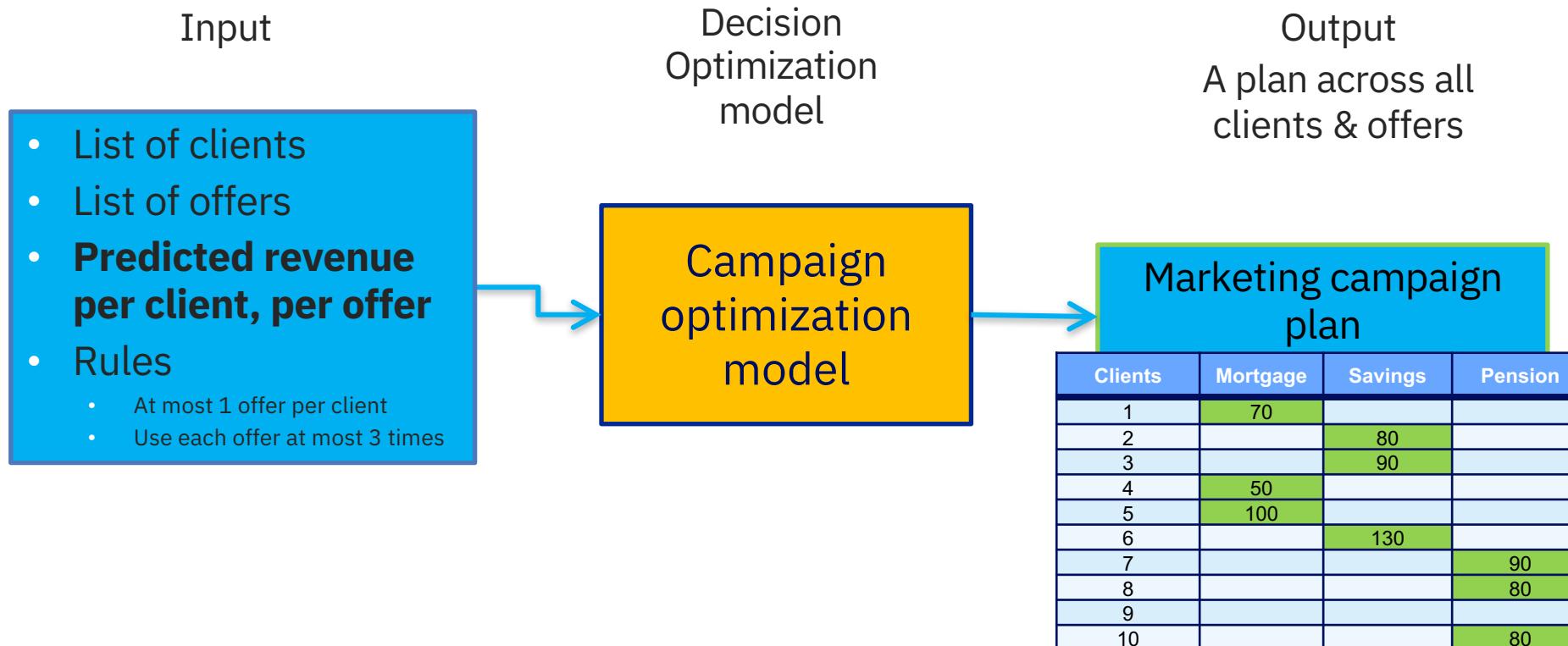
- **The best possible total revenue is 770**
- **Sequential rules would fail to find this solution**
 - For customer 6 we didn't pick the highest scoring campaign
- **Consider how difficult this would be if there were**
 - Millions of customers
 - 10s of campaigns
 - Additional rules or constraints

Customers	Mortgage	Savings	Pension
1	70	60	60
2	80	80	70
3	90	90	80
4	50	50	50
5	100	100	100
6	110	130	150
7	20	20	90
8	10	40	80
9	0	50	60
10	40	40	80

Machine Learning example: predict client revenue per offer



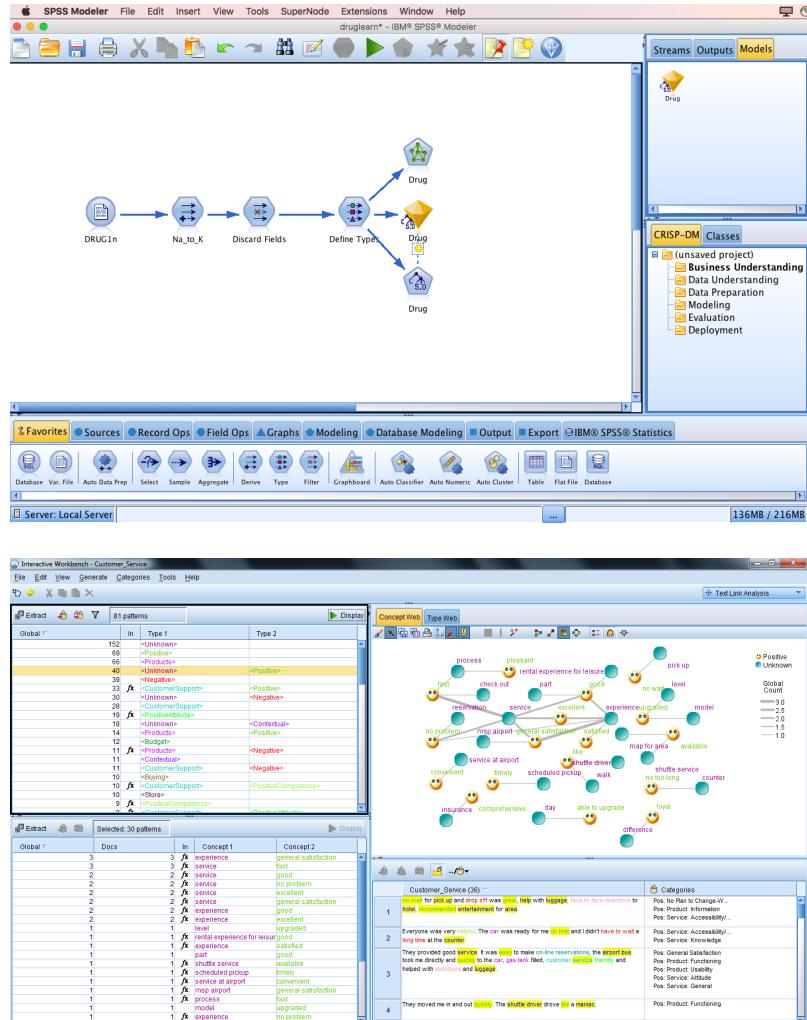
Decision Optimization example: Decide which offers to make to which clients



IBM Machine Learning & Decision Optimization for Enterprise Applications

IBM SPSS Modeler

- Tool for the **data scientist and business analyst** who focuses on time to value in data science: no-code data science
- Available on-premises or private cloud
- Key features include:
 - Built-in leading data science methodology (CRISP-DM)
 - Over 40 native algorithms, incl. text analytics
 - Easy access to open source algorithms and methods through extensions
 - Robust deployment, version control and sharing capability in Collaboration and Deployment Services
 - Can push data preparation and scoring into relational databases
 - Can (with Analytic Server) push model building into Hadoop/Spark
 - Batch, real-time & streaming deployment
 - New: embeds IBM Decision Optimization OPL node



IBM Decision Optimization

- Tools for the **operations research expert, application developer, and LoB user** to build, deploy, and use decision optimization
- Available on-premise, public, or private cloud
- Key features include:
 - Leading CPLEX & CP Optimizers
 - Optimization Programming Lanugauge (OPL) for non-coders OR experts
 - Application build & deploy environment for IT
 - Multi-scenario “what-if” analysis
 - User roles / permissions
 - Interactive planning & scheduling UI
 - Automated / batch solves
 - Decision-making under uncertainty
 - Extensive APIs – integrate anywhere
 - Built-in SPSS connector

The screenshot displays two main windows of the IBM Decision Optimization platform.

The top window is the "Developer Community" page, titled "Decision Optimization on Cloud". It features a sub-header "Use DOcloud APIs to integrate powerful online solvers with your application. Easier to develop, faster to get results." Below this is a "Register for Free Trial" button. To the right is a graphic of a green bar chart with an upward arrow and a white cloud icon.

The bottom window is the "IBM ILOG CPLEX Optimization Studio" interface. It shows an "OPL Projects" tree view with several files listed, including "vollcommun.dat", "warehouse (A warehouse location model, Mixed Integer Programming)", "Run Configurations", and various CPLEX and OPL files. Below the tree view is a detailed Gantt chart for tasks like "Tsk 1812 M305-14" and "Tsk 2171 M490-11" across dates from Nov 24 to Nov 25. At the bottom, there is a bar chart visualization with categories labeled "Mecanis[2]", "Mecanis[3]", "Mecanis[4]", "Mecanis[5]", "Mecanis[6]", "Mecanis[7]", and "Mecanis[8]".

IBM Machine Learning

- IBM Machine Learning

- A component of DSX Local
- Development and Deployment

- Watson Machine Learning

- Hosted managed service on IBM Cloud
- Deployment only service

- Machine Learning for z/OS

- Development and deployment on the mainframe

- Key features

- For experts: Python API to ingest & transform data, train & evaluate models
- For non-experts: Model building wizard with guided steps to select & transform data, and train & evaluate models
- Version & access control
- Performance monitoring

The screenshot shows two main sections of the IBM Machine Learning interface.

Top Section: This section is titled "Deep Learning for Image Classification". It displays three handwritten digit images with their corresponding numerical labels: "7 → 7", "5 → 5", "8 → 8", "3 → 3", "2 → 2", and "4 → 4". Below this, there is a brief description of the task: "Given a set of images of handwritten digits, we build a classification model that maps these images into its corresponding numerical digit." It also notes that the adapted code is from a PyImageSearch tutorial and was taken from Wolfram.com.

Bottom Section: This section is titled "Trained Model" under "My Projects > Demo project". It is a "Select a technique" wizard.

- Select Data:** Shows tabs for "Train" (selected) and "Evaluate".
- Column value to predict (Label Col):** Set to "PRODUCT_LINE".
- Feature columns:** Set to "All (default)".
- Suggested technique:** "Multiclass Classification" is selected (indicated by a green checkmark).
 - Binary Classification:** Described as classifying new data into defined categories based on existing data, choosing if the label column contains two distinct categories.
 - Multiclass Classification:** Described as classifying new data into defined categories based on existing data, choosing if the label column contains a discrete number of categories.
 - Regression:** Described as predicting values from a continuous set of values, choosing if the label column contains a large number of values.
- Validation Split:** A slider bar shows "Train: 60", "Test: 20", and "Holdout: 20".

IBM SPSS Modeler, IBM Machine Learning, and IBM Decision Optimization
are coming together in:

IBM Data Science Experience Local

Self-contained, collaborative, data science platform to build & deploy analytics applications, combining the best of open source with the best IBM proprietary tools

IBM Data Science Experience Local

Built-in **Jupyter Notebooks** for visualizing and coding on data science tasks using Python, R, & Scala.



Built-in **RStudio** and **Spark** framework parallelizes & accelerates data science tasks using **Sparklyr** and **dplyr** libraries.

DSX on Public Cloud

DSX Desktop

DSX Local on Private Cloud

- PayGo consumption with **as-a-service** delivery, up & running in seconds
- Integrated with IBM Spark-as-a-Service for compute, IBM Object Store for data, as well as other platform assets
- Immediate cloud collaboration via RStudio and Jupyter notebooks
- Easily installed on your **laptop or PC**
 - Won't scale beyond the hardware available on your machine
- Access to RStudio and Jupyter notebooks, powered by one small Spark worker operating locally on your machine
- Load CSV data files into Data Frames
- DSX Local can deploy with **Hortonworks Data Platform** on-premises
- Scalable DSX cluster deployed on your **private infrastructure** for total isolation
 - "Dockerized" containers via Kubernetes
- LDAP integration for user management
- Meaningful collaboration via notebooks, IDEs, community, and social features

SPSS Modeler & Decision Optimization are integrated in DSX
(MVP, continuously extending capabilities)

Demo – marketing campaign planning

Demo – Marketing Campaign Planning with CPLEX & DSX Local

- Bob, the marketing campaign planner, wants to create a campaign of Mortgage, Pension, & Savings offers
- Budget of \$25K
- Step 1: Use Machine Learning to predict customer behavior:



- Step 2: Use Decision Optimization (CPLEX) to decide which offers to make to which customers. Compare CPLEX with a Greedy (not optimized) algorithm.

Algorithm	Revenue	Number of clients	Mortgage offers	Pension offers	Savings offers	Budget Spent
Greedy	50800	1123	299	111	713	21700
CPLEX	72600	1218	381	117	691	25000

Today – Marketing Campaign Planning with CPLEX & DSX Local, using Notebooks

IBM Data Science Experience Local 46 Trial Days Left

Projects > tutorials > MachineLearning_and_CPLEX

File Edit View Insert Cell Kernel Widgets Help | Python 2

CellToolbar

Using IBM Decision Optimization CPLEX Modeling for Python

Let's create the optimization model to select the best ways to contact customers and stay within the limited budget.

In [31]:

```
import sys
import docplex.mp
```

Set up the prescriptive model

There are two ways to solve the model:

- Subscribe to our private cloud offer or Decision Optimization on Cloud solve service [here](#).
- Use a local solver, a licensed installation of [CPLEX Optimization Studio](#). Note that this model is too big to solve using the Community Edition, so if url and key are None, the script will fail.

In [35]:

```
url = None
key = None
```

docplex solve methods take various arguments: you can pass the api and key at each function call or you can put them at model declaration with a context.

In [36]:

```
from docplex.mp.context import Context
context = Context.make_default_context()
```

What's coming – Scenario Management for Marketing Campaign Planning with CPLEX & DSX Local

Projects > dods-advanced-examples-2 > CampaignM ... - CPLEX

Select Data

1table

Run Model

Python

Explore Solution

Processed

Dashboard

KPIs



Notes

Table

Chart

Gantt Frame

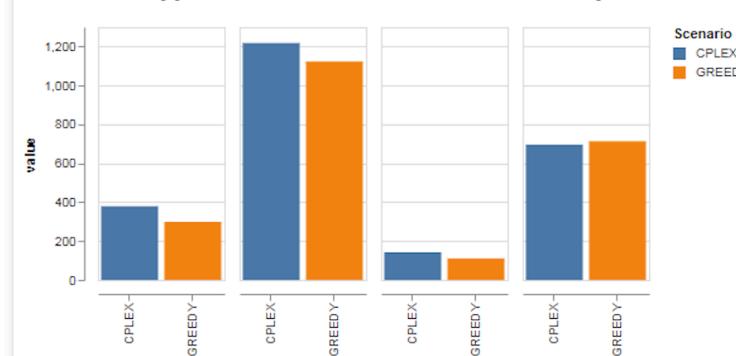
KPIs

@CPLEX | @GREEDY



Mortgage Offers Pension Savings

kpi



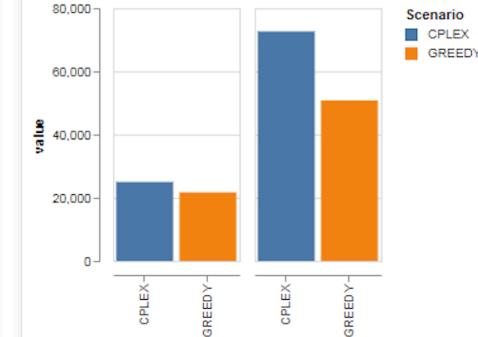
Scenario

- CPLEX
- GREEDY

KPIs

kpi

BudgetSpent Revenue



Scenario

- CPLEX
- GREEDY

Dashboard

Where to learn more

- Free learning: Cognitiveclass.ai – Decision Optimization and other Data Science courses
- Decision Optimization resources
 - Free 30-day cloud trial: ibm.co/docloudtrial
 - Free CPLEX Optimization Studio Community Edition: <http://www-01.ibm.com/software/websphere/products/optimization/cplex-studio-community-edition/>
 - CPLEX on cloud <https://onboarding-ooas.docloud.ibmcloud.com/>
 - Industry use cases / videos: IBM Decision Optimization MediaCenter
<https://developer.ibm.com/docloud/docs/category/about/>
- SPSS resources
 - Free trial: ibm.com/tryspss
 - Learn more: <https://www.ibm.com/us-en/marketplace/spss-modeler>
 - Developer community: developer.ibm.com/predictiveanalytics
- Twitter: #cplex, #orms, #datascience, #ibmoptimization, #ibmspss

Need help? The IBM Data Science Elite team is dedicated to your success

Seth Dobrin
VP & Chief Data Officer
Seth.Dobrin@ibm.com



John Thomas
DE & Director
jothomas@us.ibm.com



Susara van den Heever
Prescriptive Analytics
svdheever@fr.ibm.com



Carlo Appugliese
Machine Learning
Carlo.appugliese@us.ibm.com



Vishnu Alavur Kannan
Data Engineering
vishnu@ibm.com



Keith Carlsen
Operations
carlsenk@us.ibm.com

Who are we?

A team of **IBM Data Science experts**, with skills in:

- Descriptive, predictive & prescriptive analytics
- Industry-specific use cases
- Machine learning, deep learning, decision optimization, data engineering, data journalism

What do we offer?

- ✓ **Up to 3 months' free onsite** engagement
- ✓ **Identify use case(s)** and Minimal Viable Products via discovery and design workshops
- ✓ **Collaboratively build & evaluate** up to 4 sprints
- ✓ **Mentor and enable** client teams with hands-on learning

What do we ask of clients?

- ✓ **Dedicate team members** to match our resources
- ✓ **Install and use Data Science Experience** Local/Cloud
- ✓ **Use case defined** with Data Science Elite Team
- ✓ **Public reference** for Data Science Experience

Thank you

Susara van den Heever
IBM Data Science Practice Leader
—
svdheever@fr.ibm.com

