

ノーコード・ローコード データマイニングツール SPSS Modelerについて

日本アイ・ビー・エム株式会社
クラウド事業本部
アナリティクス
SPSS ITスペシャリスト
西牧洋一郎

2023年3月3日



自己紹介

日本アイ・ビー・エム株式会社
テクノロジー業本部
データ・AI・オートメーション事業部
Data & AI 第一テクニカルセールス
SPSS IT スペシャリスト

西牧 洋一郎



[https://www.linkedin.com/in/416nishimaki/](https://www.linkedin.com/in/416nishimaki/e51782@jp.ibm.com)

e51782@jp.ibm.com

- 1999年 エス・ピー・エス・エス株式会社入社
セールスコンサルタント
- 2006年 プリセールエンジニア
- 2009年- IBMのSPSS統合に伴いIBM ソフトウェア事業
SPSS ITスペシャリストとして数多くのデータ分析
プロジェクトの提案とサポートを手掛ける
- 2018年- データサイエンティスト協会企画員副委員長
- 2019年- 電気通信大学 非常勤講師

愛犬 小町（こまち）



日本 IBM

データサイエンティストジャパン2019 Review

Data Scientist J

多様化するデータ活用人材 AI統合環境で業務実装を目指す

データ活用には様々な課題がある。例えば、データやAIなどの技術にばかり目を向けてしまうと、これでは業務をうまく進められない。また、どのようなデータサイエンティストチームを組織し、その活動をどうやって実行していくかを考える。さらに、成果を確実に自動化や標準化が実現可能になるまで、これらデータ活用の課題をクリアするためのアプローチを「何を」「誰が」「どうやって」といふ3つのテーマに分けて解説する。

スモールスタートによって着実に成果を積み上げる

データ活用を担う人材が育たない。PoC（概念実証）までは進むが実装には至らない。AIや機械学習といった言葉ばかりが飛び交って、役員が完全に理解しているなど、データ活用が様々な課題で停滞している企業は多いようだ。このような状況に対して、日本IBMの西牧洋一郎氏は次のように指摘する。



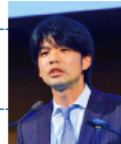
「着実に大きな成果を上げようとするのではなく、まずは小さな成功を実践することが重要。小さな成功を積み重ねて、データの価値や有効性を社内や役員に認知してもらい、取り組みに対する理解を深める。その上で実際の業務に実装し、段階的に成果を上げていくことが重要だ。かつ最適なアプローチのために、」

経営者・管理職のための データ活用実線フォーラム2016 REVIEW

日本IBM

アナリティクスの勝ち組企業が実践している3つの成功要因

「いくつか試練された予測モデルを作ったとしても、それを業務で商売でなければ1円のもうけにもなりません。こう強調するのは、日本IBMの西牧洋一郎氏だ。最初は高い精度の予測モデルではなくても、それを実行に移してモデルを磨いていけば、最終的には収益に結びつく予測モデルにたどり着けると指摘する。責任は、IBMが注力するコグニティブ(認知)技術とアナリティクスの役割の違いを解明した上で、業務上でデータ活用を成功に導くための3つのルールを解説した。



日本IBMアナリティクス事業部長 西牧 洋一郎



実践 IBM SPSS Modeler 顧客価値を引き上げるアナリティクス

西牧 洋一郎 著 日本IBM株式会社

予測モデルを素早く構築し、有益な情報をもたらす 待望のModeler入門書。 よりよい意思決定をはかるために、顧客軸マーケティングにAI×機械学習×統計解析を!

東京図書

実践! 異常検知と故障予測 IBM SPSS ModelerによるIoT時系列データ活用

小川 勇(HONDA) 竹村 玄(JALエンジニアリング) 堀 恵治(JFE東日本) 本田智剛(産総研) 河田 大・西牧洋一郎・牧野春江(IBM) 著

東京図書



前半 37.5% 貴乃花 62.5% 白鵬

どすこいAI解析結果 貴乃花 - 白鵬

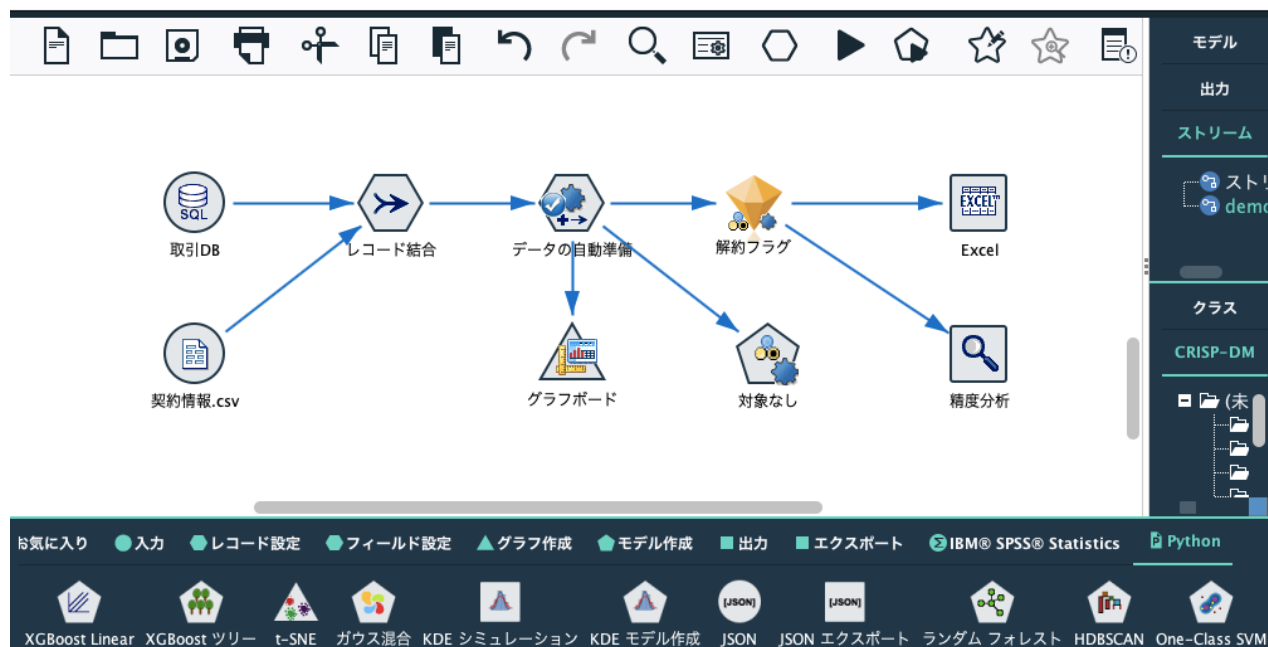
史上初! 横綱を作りだす どすこいAIの仕組み

どすこいAIの特徴は?

日本IBM 西牧 洋一郎

過去にその力士がどんな相手にどんな勝ち方をしたのか

データマイニングツール SPSS Modeler



①アイコン操作

⇒すぐはじめられる・試行錯誤

②プロセス可視化

⇒組織で知識共有

③業務適用

⇒予測を業務実装

①アイコンで簡単に始められて試行錯誤しやすく

②プロセスが資産として残るので組織で知識を共有でき

③素早く既存業務に予測プロセスを展開できる

JFEスチール様、三井化学様、日本ガイシ様のDX人材育成とノーコード・ローコード分析ツール

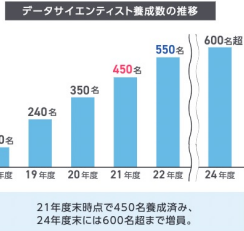


データサイエンティスト養成

DX推進に向けて、社内データサイエンティストの養成と活躍をさらに強化する仕組みや、多くの社員がDX推進を担えるよう使いやすいDS活用環境と仕組みを整備しています。

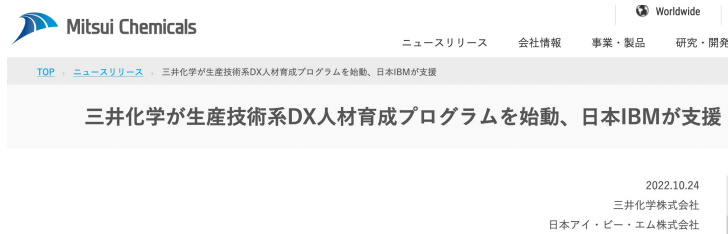
<階層>	<対象>	<養成方法>
1 DS先駆者	・研究所 研究員 ・製造エンジニア データサイエンティスト	・学術機関への派遣研修 ・研究所-JDXC*での研修
2 DS伝達者	・電気保全エンジニア ・製造エンジニア	・高度DSツール活用教育 ・OJT主体 ・DS全般教育
3 DS活用者	・技術系社員全員	・DSツール導入教育
4 DS利用者	・事務系含む社員	・リテラシー教育 (e-Learning)

・2018年度に全社で階層別教育プログラムを開始
・データサイエンティストの8割以上がDS業務課題を遂行中



SPSS (IBM) 中級者用DS解析・モデリングツール
使用者 1000名 実用化 430件 準備中 60件

<https://www.jfe-holdings.co.jp/investor/library/dxreport/2022/pdf/all.pdf>



プログラムの特徴

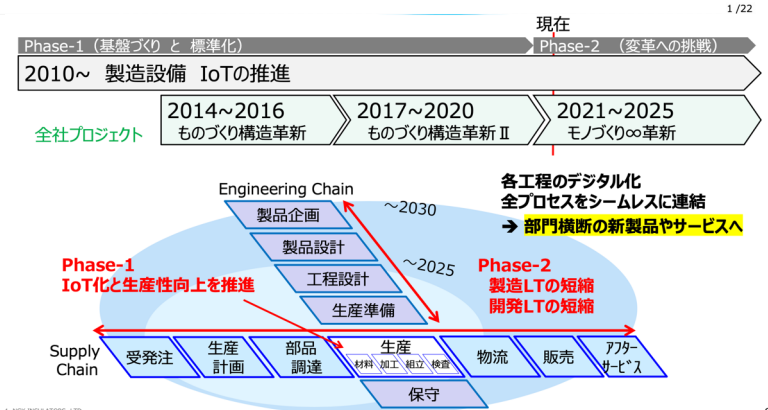
本プログラムの構成は、基礎・応用・データサイエンティスト育成の大きく3段階とし、知識やスキルの保有を評価するだけでなく、「できることベース」で更にレベルを6段階で定義し、ビジネスの貢献度合いも合わせて評価します。更に、本プログラムの研修・認定は、アルゴリズム、統計・分析、データ倫理、プロジェクトマネジメントなどをOJTとOFF-JTを組み合わせて実施します。なお、レベル1のデータアナリスト基礎教育は生産技術系エンジニア全員(約250名)を受講対象とし、既に120名の教育を完了しています。

	データアナリスト基礎 (レベル1)	データアナリスト応用 (レベル2)	データサイエンティスト (レベル3以上)
できること	作業に最低限必要な基礎知識を有する	上位者の指導の下、要求された作業を遂行できる	独力で解析作業を遂行できる
教育方法	講義の聴講でデータ解析概要を学び、ハンズオンで統計解析ソフトSPSSの使い方を学ぶ	品質悪化の要因解析と、品質予測モデル作成を演習形式で行う	自職場の技術課題解決をOJT形式で実施する

*当社生産技術系の教育体系では、レベル1相当の基礎編、レベル2相当の応用編までをデータアナリスト育成とし、レベル3からレベル6までをデータサイエンティスト育成教育とし、大きく3つの段階に分類しています。

https://jp.mitsuichemicals.com/jp/release/2022/2022_1024.htm

2021 IBM SPSS Modeler 秋のオンラインユーザー会 2021.11.05 データドリブンなモノづくりを支えるSPSSの利活用



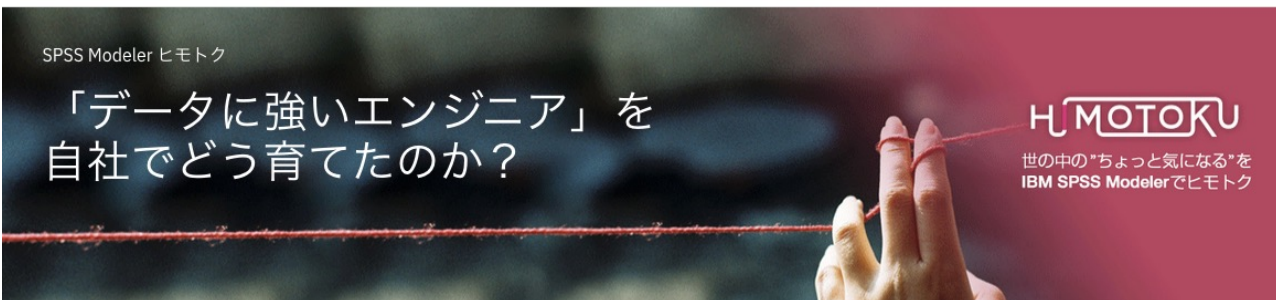
1: HOK INSILATORS, LTD.

<https://speakerdeck.com/jpsps>

HONDA様はデータに強いエンジニアを育てています

<https://www.ibm.com/blogs/solutions/jp-ja/spssmodeler-honda-story/>

IBM ソリューション ブログ



「作って、売って、終わり」ではなくなってきたビジネスモデル



ホンダのクルマ作りに訪れた変化とは、具体的には何を指しているのか。中川氏は次のように説明する。

「従来の我々のビジネスモデルは、よい自動車を開発・製造してお店に届け、お客さまにショールームに来ていただいて、お買い求めいただくというものでした。ところが、5年ほど前から、購入後の自動車の使われ方についても、我々への期待が徐々に大きくなっていったのです」

スマートフォンの普及によって、いまや人々は常にインターネットにつながっている。ところが、クルマに乗っているときだけ、インターネットに自由につながることはできない。

データに強いエンジニアは自社で育てる

こうしてホンダは、2012年に「ビッグデータプロジェクト」を立ち上げる。当初からチームに参加していた中川氏は、次のように振り返る。

「プロジェクトは立ち上がったものの、何をしたらよいのか、当初はまったくの手探り状態でした。そこで、研究開発に携わる社員たちに、データを使って何を分析したいのか聞いてみたのです。すると『高齢者の運転について研究したい』などのテーマが53個上がってきました」

次にチームは、そのテーマを研究するうえで何が必要なかを聞いた。返ってきたのは、データマイニングを実施する環境やスキルの不足を訴える声だった。

データのサイロ化を克服し、新しいアイデアを生み出す「デジタル砂場」を目指す

データ分析に対するホンダの取り組みは、着実に成果を出しつつある。たとえば、データサイエンティスト協会が、毎年、データ分析・活用による産業への貢献を表彰する「データサイエンスアワード」では、本田技術研究所の「ビッグデータを活用したリチウムイオン電池の性能設計・検証プロセスの構築」が、2017年度のファイナリストに選ばれた。

「興味深いのは、この取り組みの中心になったのが、タイプの異なる3名がコアとなるチームであったことです。1人はデータに強いエンジニアで、もう1人はMBAの資格を持つビジネスに強いエンジニア、そして3人目がデータサイエンティストです。このユニークな3名が組んで成果を出せたことは、非常に意義深いと思います」

ホンダ社内では、こうした「データサイエンスに強い社員」が着実に育ち、活躍の場を広げつつある。この動きをさらに加速するため、現在、中川氏が検討しているのが「デジタル砂場」の実現だ。

a
読み込み

b
データ表示

c
基礎統計

d
グラフ作成

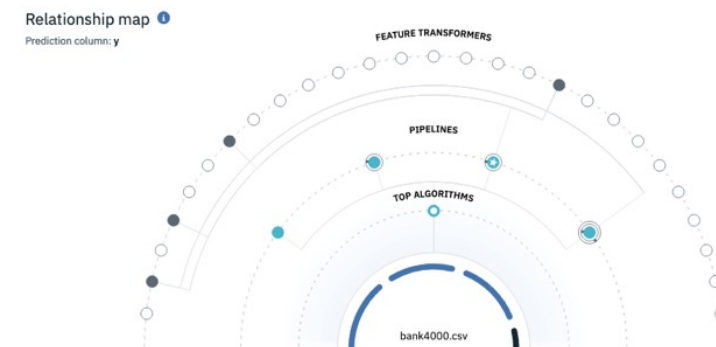
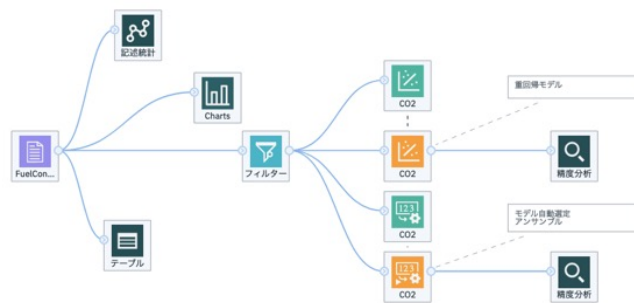
e
データ加工

f
モデル作成



```
#重回帰分析モデル
msk = np.random.rand(len(df)) < 0.8
train = cdf[msk]
test = cdf[~msk]

from sklearn import linear_model
regr = linear_model.LinearRegression()
x = np.asanyarray(train[['ENGINE_SIZE', 'CYLINDERS', 'FUEL_CONSUMPTION_COMB']])
y = np.asanyarray(train[['CO2EMISSIONS']])
regr.fit(x, y)
# The coefficients
print('Coefficients: ', regr.coef_)
```





IBM