



文部科学省

MINISTRY OF EDUCATION,
CULTURE, SPORTS,
SCIENCE AND TECHNOLOGY·JAPAN



大学共同利用機関法人 情報・システム研究機構
統計数理研究所
The Institute of Statistical Mathematics

文部科学省委託事業

「データサイエンティスト育成ネットワークの形成」

平成 27 年度事業報告書

平成 28 年 3 月



情報・システム研究機構
統計数理研究所

【再委託先】

東京大学大学院
情報理工学系研究科

事業期間／平成 25 年 7 月 8 日～平成 28 年 3 月 31 日

文部科学省委託事業
「データサイエンティスト育成ネットワークの形成」
平成 27 年度事業報告書

目 次

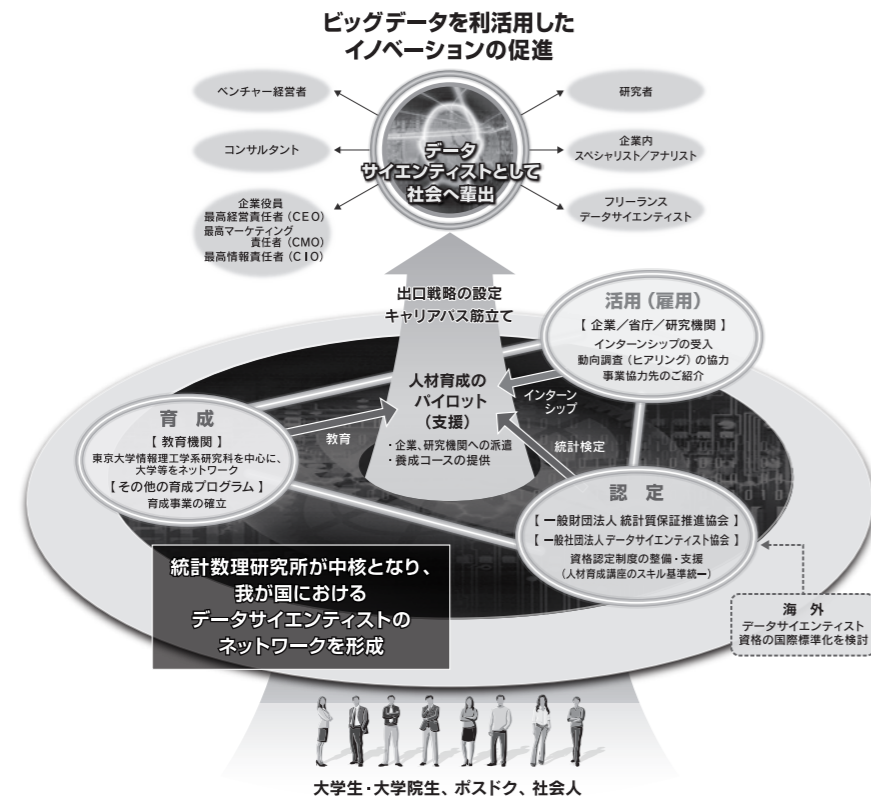
1. エグゼクティブ・サマリ	01
2. 平成 27 年度の主要成果と知見	02
2.1. 事業の基本理念	02
2.2. 平成 27 年度の主要成果と知見	02
2.2.1. 認知度向上・啓発活動	
2.2.2. 人材のローテーション	
2.2.3. ベスト・プラクティスの調査	
2.2.4. 育成教材の展開	
2.2.5. 海外との連携	
2.2.6. プロジェクトの総合推進	
2.3. 事業全体の成果と知見	04
2.4. 今後の展開	05
3. 事業実施内容の詳細	06
3.1. 認知度向上・啓発活動	06
3.2. 人材のローテーション	13
3.3. ベスト・プラクティスの調査	20
3.4. 育成教材の開発	22
3.5. 海外との連携	25
3.6. プロジェクトの総合推進	25
付録 A: データサイエンティスト協会 スキルチェックリスト ...	27
付録 B: 異分野異業種交流会 学生・若手研究者参加アンケート ...	80
付録 C: 提供されている関連講座・教育プログラムのリスト (平成 27 年 12 月 14 日現在)	81
付録 D: データサイエンス関連コンテストのリスト (平成 27 年 12 月 10 日現在)	116

1. エクゼクティブ・サマリ

ビッグデータをビジネス上の価値に変えることのできる人材、いわゆるデータサイエンティストの育成が我が国の競争力のために必須と考え、産官学のネットワークを通してデータサイエンティスト育成の土壌を醸成すべく、平成 25 年 7 月より 2 年 9 ヶ月に渡って本事業を推進してきた。この 3 年弱の間に、データサイエンティスト育成を取り巻く状況も大きく変化した。滋賀大学においては、我が国初のデータサイエンス学部が、平成 29 年度から設置されることになった。新しい職種であるデータサイエンティストの育成と業界の健全な発展を目指す民間団体であるデータサイエンティスト協会は平成 28 年 2 月現在 3,000 名以上の個人会員を擁するなど、その活動を拡充している。より広く世の中を見回してみても、データサイエンティスト育成のための多くの有料・無料のコースが提供され、勉強会やコンテストなども多く見られるようになってきた。規模的にはまだまだ不十分ではあるが、着実な変化を感じることができる。これらの動きの多くには、直接・間接に本事業で形成されたネットワークが関わっていて、その意味で本事業は一定の成果を挙げてきたと考える。

これらの情勢の変化を踏まえ、事業最終年度にあたる平成 27 年度では、今までネットワークを通して得られた知見をさらにスケールアップし、事業終了後の道筋も見据えて事業を推進した。平成 25 年度にパイロットを行い、平成 26 年度に 1 つの形をなしたインターンシップ・プログラムは、本年度においては、民間会社における営利事業として継続できる見通しを得た。データサイエンティスト協会は、データサイエンティストが身に付けるべきスキルを具体的に定義し、またそのためのチェックリストを公開した。平成 27 年 7 月には、今後の我が国におけるビッグデータ利活用人材育成の青写真とも言える提言を、産官学の有識者を集め、情報・システム研究機構が中心になって作成し公開した。また、その青写真に基づき、将来の棟梁クラス人材の育成と認定を視野に入れて、データ分析ハッカソンを企画し、試行した。

本報告書では、27 年単年度の成果と共に、事業全体の成果を振り返り、今後の方向性について考える。



2. 平成 27 年度の主要成果と知見

2.1. 事業の基本理念

データをビジネス上の価値に変えることのできる人材、いわゆるデータサイエンティストの必要性が叫ばれている。我々は、本事業を通して、「我が国におけるデータ利活用人材のあるべき姿」を明らかにし、それに基づいて戦略的に人材育成のエコシステムを創りだす、という目標を設定した。このため、本事業においては以下の 3 点を基本理念とした。

スケラビリティ

最終的に数万人規模のデータサイエンティストの育成が必要であったとしても、一つの機関ないしはプロジェクトがこれらの人材を輩出することはできない。必要なのはスケールする仕組みである。このため、本プロジェクトではデータサイエンティスト育成に熱意を持つ教育機関と、データサイエンティストのスキルを利用したい企業・組織を広くネットワークし、それらの間で知識・経験を共有することで、多くのデータサイエンティストが育成され、有効に活用されることを狙う。

出口戦略

人材育成の取り組みにおいて、もう一つ大事なことは出口戦略である。いくら大人数のデータサイエンティストが育成されたとしても、彼らの就職先が無ければ意味のないものになってしまう。データサイエンティストとしてのキャリアパスは何か、その青写真まで含めて、データサイエンティストのエコシステム全体を考える。

我が国の実情に即した方策

米国と日本では、産業構造、労働市場、教育の仕組みや国民のメンタリティなどが異なる。米国の事情は参考にするが、単に米国発のトレンドの後追いをするのではなく、我が国の実情に即した「あるべき姿」を追求する。

2.2. 平成 27 年度の主要成果と知見

上記の基本理念を元に、事業初年度の平成 25 年度は、特に「我が国におけるデータ利活用人材のあるべき姿」を明らかにすることに注力した¹。事業 2 年目にあたる平成 26 年度は、初年度の成果をベースに、実際の人材育成につながる取り組みに力を入れた²。特に、データサイエンティスト育成インターンシップ・プログラムは、初年度のパイロットの知見をふまえて、民間のノウハウを取り入れつつ一つの成功パターンを作ったと言える。また、教材の観点からも、現在民間・大学等で提供されているデータサイエンティスト育成のデータベースを作成したほか、初年度に作成した「データサイエンティスト・クラッシュコース」を無料動画サイトへアップロードしたことで、広く使ってもらえる環境が整った。最終年度にあたる平成 27 年度は、これまでに得られた知見に基づいてその成果をさらにスケールアップすべく、特に

- ①データサイエンティストのスキルレベル定義、
- ②データサイエンティスト育成関連データベースの更新・公開、
- ③我が国におけるデータサイエンティスト育成施策に関する提言、
- ④海外連携の検討、
- ⑤インターンシップ・プログラムの展開

といった点に取り組んだ。

¹ 初年度の成果については、初年度の報告書(<http://datascientist.ism.ac.jp/pdf/H25DSTN.pdf>)を参照されたい。
² 平成 26 年度の成果については、平成 26 年度の報告書(<http://datascientist.ism.ac.jp/pdf/H26DSTN.pdf>)を参照されたい。

2.2.1. 認知度向上・啓発活動

これまでに引き続き、多くの講演、出版物、Web サイト等を通して認知度向上・啓発に務めた。平成 28 年 3 月には滋賀大学佐和学長、情報・システム研究機構北川機構長、データサイエンティスト協会スキル委員会安宅委員長などを招いて「データサイエンティスト育成ネットワークの形成」最終年度報告シンポジウムを開催し、多くの参加者を得た。また、慶應義塾大学ビジネスデータ創造コンテスト、データサイエンティスト協会シンポジウム、情報処理学会ソフトウェアジャパンなどに共催・協力し、本事業ネットワークの認知度向上に務めた。

2.2.2. 人材のローテーション

平成 25 年度にトライアルを試行し、その知見を入れて設計しなおして一応の完成をみたインターンシップ・プログラムは、本年度においては、以下の 3 点での展開を行った。

1. 民間移転
インターンシップ・プログラムの事業終了後の継続をふまえ、本年度は株式会社アカリクが前年度のフォーマットに従ってデータサイエンティストインターンシップ説明会を平成 27 年 5 月に開催した。
2. 業種の拡大
これまでのインターンシップ・プログラムに参加していた企業は、情報処理関係やマーケティング関係に集中していたが、平成 28 年 1 月に開催し「ビッグデータやデータサイエンスに関わる学生／研究者、企業の異分野異業種交流会」においては、10 を超える学協会の共催・協賛を得たほか、製造業、通信業、シンクタンク、国立研究機関などの参加があり、データサイエンティストのキャリアを考える選択肢が大きく広がった。
3. ハッカソンの実施
平成 28 年 2 月には、統計数理研究所においてデータ分析ハッカソンを行った。このハッカソンは、クラウド環境を活かしたものであり、データ分析の PBL のモデルケースとして利用できるとともに、今後の棟梁クラスの人材育成における、質保証のためにも利用できるを考える。このハッカソンの設計とその実施経験から得られた知見は、論文として投稿中である。

2.2.3. ベスト・プラクティスの調査

前年度に続き、データサイエンティスト育成・利活用に関するベスト・プラクティスの調査を継続した。その一部は論文として情報処理学会デジタルプラクティス誌上で平成 27 年 7 月に刊行した。データサイエンティストの利活用に関しては特に、佐賀県と株式会社松竹の事例を追跡調査し、それぞれの特質を活かしたデータサイエンティストの利活用のパターンと、それによるメリット・デメリットに関する知見を得た。

2.2.4. 育成教材の開発

育成教材としては、初年度に開発した「データサイエンティスト・クラッシュコース」を YouTube 上に展開した。また、民間などのデータサイエンティスト教育プログラムのデータベースを拡充したほか、新たに、データ分析に関するコンテストの一覧を整備した。なお、これらのデータベースは、事業終了後に、データサイ

エンティスト協会に移管する予定である。

育成教材そのものではないが、データサイエンス協会が平成 27 年に発表したデータサイエンティストのスキル定義は、データサイエンティストの育成や認定において 1 つの指標となるものであり、重要な成果と言える。特に、本報告書の付録 A にあるスキルチェックリストは、データサイエンティスト協会スキル委員会における長時間の議論の成果物であり、世界に類をみない、詳細なものである。

2.2.5. 海外との連携

本年度においては、英国 Warwick 大学、米国オハイオ州立大学、ロチェスター大学等を訪問・調査し、海外におけるデータサイエンティスト育成プログラムの動向把握に務めた。これらの知見は、滋賀大学のデータサイエンス学部の設立に活かされている。

2.2.6. プロジェクトの総合推進

プロジェクトの推進においては、初年度と同様、運営委員会を開催し識者の意見を取り入れながら行った。また、事業の発注元である文部科学省・日本科学技術振興機構とも密接に連携しながら、正しい事業の遂行・予算執行に最大限の留意をした。平成 27 年 7 月には、今後の我が国におけるビッグデータ利活用人材育成の青写真とも言える提言を、産官学の有識者を集め、情報・システム研究機構が中心になって作成し公開した。

2.3. 事業全体の成果と知見

本事業期間を通して、我が国のデータサイエンティストを取り巻く環境は大きく変化した。データサイエンティストという言葉が広く浸透し、データサイエンティストに関する多くの書籍が刊行され、データサイエンティストを育成する多くの教材が現れた。滋賀大学においては、本邦初のデータサイエンス学部が平成 29 年度より設置されることが決まり、進学を考える高校生と、人材を求める企業の双方から注目されている。新しい職種であるデータサイエンティストの育成と業界の健全な発展を目指す民間団体であるデータサイエンティスト協会は平成 28 年 2 月現在 44 社の法人会員と、3,000 名以上の個人会員を擁するようになった。データサイエンティストとしての実務を経験する、データサイエンティスト向けのインターンシップ・プログラムも、民間の営利事業として軌道に乗り始めている。産官学の有識者からなる「ビッグデータの利活用に係る専門人材育成に向けた産学官懇談会」は、我が国におけるビッグデータ利活用人材育成の青写真とも言える提言を平成 27 年 7 月に公開し、その提言に沿った施策の検討が始まっている。これらの動きの多くには、直接・間接に本事業で形成されたネットワークが関わっていて、その意味で本事業は一定の成果を挙げてきたと考える。

より具体的な成果としては、多くの講演・出版物、オンライン教材、教材データベース、インターンシップ・プログラムの設計、データ分析ハッカソン、などがあり、それらの詳細は、本報告書を含む 3 年間の報告書に記載したとおりである。

本事業期間全体を通して得られた主要な知見は以下の 3 点にまとめることができる。

1. データサイエンティスト像の多様性
データサイエンティストという概念は米国で発生したものであるが、「ビッグデータをビジネス上の価値に変えることのできるプロフェッショナル」という観点で見ると、共通に求められる資質はあるものの、

3. 事業実施内容の詳細

3.1. 認知度向上・啓発活動

我が国にとってのデータサイエンティスト育成やデータサイエンティストの「あるべき姿」を発信し、啓発活動を行った。同時に、以下の活動のうち、特に講演による発信、啓発活動は、必然的に産官学の主要ステークホルダーとの意見交換を伴ったので、ネットワークづくりと密接に関わった。

講演等

- ・丸山 宏 情報サービス産業はどこへ？, 株式会社リンクレア (平成 27 年 4 月 3 日)
- ・樋口知之 分野の壁を突き崩すビッグデータ, 平成 27 年度総合研究大学院大学フレッシュマン・コース“知のフロンティア” (平成 27 年 4 月 9 日)
- ・丸山 宏 「研究が正しいかどうか」は、なぜわかる？, 情報・システム研究機構 第 2 回サイエンスカフェ (平成 27 年 5 月 11 日)
- ・竹村彰通 統計学分野の参照基準案について. 日本学術会議公開シンポジウム「学士課程教育における統計学分野の参照基準を考える」 (平成 27 年 7 月 9 日)
- ・樋口知之 ビッグデータを取り巻く課題と公共インフラ構築・社会サービスのイノベーション, 中国地域 ICT 利活用研究会 (平成 27 年 7 月 13 日)
- ・丸山 宏 ビッグデータと IoT, 日本アイ・ビー・エム株式会社 (平成 27 年 7 月 21 日)
- ・竹村彰通・佐和隆光・吉川英治 データサイエンスを専門的に学ぶ大学教育の展望：滋賀大学の目指すもの (招待講演), 日本行動計量学会第 43 回大会 (平成 27 年 9 月 2 日)
- ・丸山 宏 ビジネスに活かすデータサイエンス, 日本行動計量学会第 43 回大会 (平成 27 年 9 月 4 日)
- ・丸山 宏 汎用ヒト型ロボット活用に向けたデータサイエンス 「汎用ヒト型ロボット活用研究会」ワークショップ (平成 27 年 9 月 5 日)
- ・田栗正章・竹村彰通 日本学術会議における統計学分野の参照基準の検討状況について, 2015 年度統計関連学会連合大会 (平成 27 年 9 月 7 日)
- ・中西寛子・竹村彰通・岩崎学 統計検定準 1 級受験者の動向と試験結果, 2015 年度統計関連学会連合大会 (平成 27 年 9 月 9 日)
- ・樋口知之 人工知能とビッグデータ 文系の仕事は無くなる!?, 宮崎大宮高校ハローワーク (平成 27 年 9 月 19 日)
- ・樋口知之 ビッグデータを生かすデータサイエンス, リコー中央研究所講演会 (平成 27 年 10 月 26 日)
- ・丸山 宏 IoT 時代のビッグデータ活用方法と具体例, IoT イノベーション講座 株式会社半導体理工学研究センター (平成 27 年 10 月 26 日)
- ・丸山 宏 パネル討論「若手科学者の拡がる未来」, サイエнтиフィック・システム研究会 合同分科会「オープンデータとセキュリティ」 (平成 27 年 10 月 29 日)
- ・樋口知之 統計科学とビッグデータの利活用, 国東市偉人顕彰プロジェクト・武蔵フォーラム (平成 27 年 11 月 14 日)
- ・竹村彰通・佐和隆光・吉川英治・姫野哲人 データサイエンスの大学専門教育－滋賀大学の構想. 日本品質管理学会 (平成 27 年 11 月 14 日)
- ・丸山 宏 IoT 時代のビッグデータ活用方法と具体例, 「ソーシャル・センシングソリューション技術分科会」電子情報技術産業協会 (平成 27 年 11 月 17 日)
- ・丸山 宏 IoT 時代のビッグデータ活用方法と具体例, Embedded Technology 2015/IoT Technology 2015 カンファレンス (平成 27 年 11 月 18 日)

その専門性のレベル (見習い、独り立ち、棟梁クラス、業界代表レベル、など)、そもそものバックグラウンド (自然科学、情報科学、統計学、経済学、経営学など) や働き方 (サービスプロフェッショナル、部署内での専門家、フリーランスなど)、キャリアの形成には多くのバリエーションがあることがわかった。これらの多様性をサポートする育成を考える必要がある。

2. 現場体験の重要性

データサイエンティストには、統計学や機械学習を中心とするデータサイエンス力、情報科学やソフトウェア工学を中心とするデータエンジニアリング力、さらにはビジネスを理解し推進するビジネス力という分野横断型のスキルが求められる。特に後者の 2 つは、現場での経験から学ぶ割合の大きい分野であり、このため PBL など現場体験型の育成が不可欠になる。本事業でも現場体験を推進するため「人材ローテーション」を柱の 1 つに掲げてインターンシップ・プログラムなどを実施したが、そのフィードバックからも、現場体験の重要さは確認することができた。

3. データサイエンティスト利用側のリテラシー・洞察

初年度におけるデータサイエンティスト現状調査や、昨年度から本年度にかけて実施した利活用ベストプラクティス調査などを含め、本事業全体を通して繰り返し感じられたのは、データサイエンティストを育成するだけでなく、データサイエンティストを雇用し利用する側のリテラシーの重要性である。そのためには、社会全体のデータ・リテラシーも向上させる必要がある。

2.4. 今後の展開

ビッグデータがバズワードであったのは本事業開始の平成 25 年ころであったが、その後社会の関心は人工知能や IoT (Internet of Things) にシフトしつつあるように見える。現在騒がれているいわゆる「人工知能」のうち実体を持つものの多くは、統計的機械学習の様々な応用であり、その意味で現在の人工知能ブームの担い手はまさにデータサイエンティストそのものと言ってよい。また、IoT において、「ものづくり」に傾倒しがちな我が国ではセンサーなどのデバイスやロボットなどのハードウェアに目が行きがちだが、IoT がもたらす最大のチャレンジは、膨大な数のセンサーから生成されるビッグデータをどのようにビジネスに活かしていくかであり、ここでもデータサイエンティストが中心的な役割を果たしていくのは間違いがない。ビッグデータと呼ぶか、人工知能と呼ぶか、IoT と呼ぶかに関わらず、今後の社会においてデータサイエンティストの重要性はますます大きくなっていくだろう。

今後我が国におけるデータサイエンティストの育成は大きくスケールアップしていく必要があるのは明らかである。そのための戦略として、「ビッグデータの利活用に係る専門人材育成に向けた産学官懇談会」にもあるように、複数の分野を俯瞰的に見てデータ分析の活用を戦略的に考えることができる「棟梁クラス」の育成が急務であり、そのために、本事業で得られたネットワークやスキル定義、インターンシップ・プログラム、ハッカソン・プログラムなどが利用できるものとする。

- ・樋口知之 人工知能とデータサイエンティストの役回り, 数理システムユーザーコンファレンス 2015 (平成 27 年 11 月 20 日)
- ・樋口知之 データ駆動型イノベーションの神髄, 第 9 回 The KAITEKI 講演会 (平成 27 年 12 月 16 日)
- ・丸山 宏 情報技術は科学の営みをいかに変えるか, 日本学術会議公開シンポジウム「全細胞・マルチモーダル計測が拓く統合バイオサイエンス」理化学研究所 (平成 28 年 1 月 15 日)
- ・丸山 宏 機械学習とデータマイニングの最前線と人口知能への応用展開パネル討論 ソフトウェアジャパン 2016 人口知能が 2020 年を変える 情報処理学会 (平成 28 年 2 月 4 日)
- ・丸山 宏 パネル討論:「研究現場におけるオープンデータの進め方」, 情報・システム研究機構シンポジウム「オープンサイエンスにおける研究データのオープン化～大学のデータ駆動型学術研究を加速するために～」(平成 28 年 2 月 8 日)
- ・Maruyama, H. Big Data, Machine Learning, and Resilience, Tokyo Roundtable on Existential Risks and AI organized by Research Organization of Information and Systems (平成 28 年 2 月 22 日)
- ・樋口知之 日本におけるビッグデータ時代の人材ビッグバン計画, 第 10 回日本統計学会春季集会 (平成 28 年 3 月 5 日)
- ・佐和隆光・竹村彰通 滋賀大学のデータサイエンス学部創設について, 文部科学省委託事業「データサイエンティスト育成ネットワークの形成」最終年度報告シンポジウム (平成 28 年 3 月 7 日)
- ・丸山 宏 「データサイエンティスト育成ネットワークの形成」事業の報告, 文部科学省委託事業「データサイエンティスト育成ネットワークの形成」最終年度報告シンポジウム (平成 28 年 3 月 7 日)
- ・樋口知之 今後の展望, 文部科学省委託事業「データサイエンティスト育成ネットワークの形成」最終年度報告シンポジウム (平成 28 年 3 月 7 日)

著書・論文・記事等

- ・丸山 宏・神谷直樹・樋口知之・竹村彰通・大西立頭 データサイエンティスト育成と人材利活用のベスト・プラクティス. 情報処理学会デジタルプラクティス, 6(3), 223-229. (平成 27 年 7 月)
- ・丸山 宏・山田 敦・神谷直樹 データサイエンティスト・ハンドブック. 近代科学社 (平成 27 年 9 月)
- ・樋口知之 日経 BigData にインタビュー記事とコメント記事 (平成 27 年 9 月)
- ・樋口知之 機械エンジニアのためのビッグデータの基礎知識. 日本機械学会誌, 118(1163), 610-615. (平成 27 年 10 月)
- ・樋口知之 産経新聞『数学の時代』にコメント, 産経新聞 (平成 27 年 12 月 28 日)
- ・樋口知之 ビッグデータが変える日常と非日常の境目. 電子情報通信学会誌, 99(1), 30-35. (平成 28 年 1 月)
- ・樋口知之 スモールデータ、ビッグデータ、そしてスマートデーター人工知能ブームの中での統計学ー. 統計, 67(1), 9-14. (平成 28 年 1 月)
- ・樋口知之 企業の中核部隊にデータサイエンティスト、“棟梁級”人材の育成を. 日経 BigData (平成 28 年 1 月)
- ・樋口知之 産経 WEST『数学の時代 (3)』にコメント, 産経 WEST (平成 28 年 1 月 11 日)
- ・樋口知之 週刊ダイヤモンド特集「使える！数学」にインタビュー記事 (平成 28 年 1 月 23 日)
- ・樋口知之 日経 BigData にコメント記事 (平成 28 年 2 月)
- ・丸山 宏・神谷直樹・宮園法明 クラウド環境を利用したデータ分析ハッカソンの計画と実施. 情報処理学会デジタルプラクティス. (査読中)

共催・後援・協力等

- ・審査協力 (丸山 宏) 第 3 回データビジネス創造コンテスト (主催: 慶應義塾大学 SFC 研究所データビジネス創造・ラボ、ビジネスパートナー: 株式会社ブレインパッド)
- ・スキル委員会への協力 一般社団法人データサイエンティスト協会、一般社団法人データサイエンティスト

協会 2nd シンポジウム～実務者が集うデータサイエンスの最前線～

- ・共同主催 ビッグデータやデータサイエンスに関わる学生／研究者、企業の異分野異業種交流会 (主催: 知的人材連携センター)
- ・サポーター 情報処理学会ソフトウェアジャパン 2016
- ・審査協力 (丸山 宏) 第 4 回データビジネス創造コンテスト (主催: 慶應義塾大学 SFC 研究所データビジネス創造・ラボ、ビジネスパートナー: 日本マイクロソフト株式会社)

今年度は事業最終年度にあたるため、平成 28 年 3 月 7 日 (月) に「データサイエンティスト育成ネットワークの形成」最終年度報告シンポジウムを開催した。このシンポジウムでは、「ビッグデータを活用しイノベーションを促進する人材を育成するネットワークの形成」を目標に事業を推進してきた、約 3 年間に及ぶ活動成果と今後の展望について報告した。講演者ならびに演題は、以下の通りであった。

- ・佐和隆光 (滋賀大学・学長)・竹村彰通 (東京大学・教授)『滋賀大学のデータサイエンス学部創設について』
- ・北川源四郎 (情報・システム研究機構・機構長)『データサイエンティスト育成懇談会報告書について』
- ・安宅和人 (データサイエンティスト協会スキル委員会・委員長)『データサイエンティストのスキル定義』
- ・丸山 宏『「データサイエンティスト育成ネットワークの形成」事業の報告』
- ・樋口知之『今後の展望』

データサイエンティスト育成やデータサイエンティストの「あるべき姿」

この事業を通じて得られた知見は、①データサイエンティストのスキルレベル定義と、③我が国におけるデータサイエンティスト育成施策に関する提言に生かされた。①データサイエンティストのスキルレベル定義に関連して、平成 27 年 11 月 20 日に、一般社団法人データサイエンティスト協会は、データサイエンティスト スキルチェックリストを公開した (付録 A: データサイエンティスト協会 スキルチェックリスト)。この公開に先立ち、統計数理研究所の樋口所長と丸山宏教授は、その内容の確認作業に従事した。

また、③我が国におけるデータサイエンティスト育成施策は、報告書『ビッグデータの利活用のための専門人材育成について³⁾』(ビッグデータ利活用に係る専門人材育成に向けた産学官懇談会、座長: 情報・システム研究機構 北川機構長)に纏めた。情報通信技術の発展により、ビッグデータが登場し、社会の価値がビッグデータとそれを処理する情報技術へと大きくシフトしているにもかかわらず、我が国ではデータサイエンティストは大幅に不足しているだけでなく、データサイエンティストを育てる体制もできていない。そのため、本事業開始のきっかけにもなっているデータサイエンティスト育成が喫緊の課題となっている。産業界やアカデミアからの要請を把握し、育成すべきデータサイエンティストが持つべきスキルや能力を特定し、その育成のための方策を検討した。

1) 我が国におけるデータサイエンティストへの要請

産業界がデータサイエンティストに求めているスキルは、単にデータを収集してその統計分析をするだけでなく、むしろ「データに基づき予測し、その結果をビジネスに活用する」ことにある。統計学に加えて、機械学習や離散的最適化など最新の手法に精通すると共に、それらを用いたビジネス的な課題解決能力、すなわちデータと現実のビジネスをつなぐことのできる人材を求めている。このため、データサイエンティストには、①顧客と建設的な会話をする能力、あるいは顧客の課題をデータ分析や情報技術に落としこむ翻訳力、②課題領域を見通して本質的な問題 (イシュー) を見抜く能力、③課題解決のため各分野のエキスパートを動員でき

³⁾ http://www.rois.ac.jp/open/pdf/bd_houkokusho.pdf

る能力の3つが必要である。

一方、アカデミアで考えられているデータサイエンティスト像は、企業が求めるものとは異なっている。アカデミアにおけるデータサイエンティストとは、いわゆる「第4の科学」と呼ばれる方法論を実践する研究者をさすことが多い。その意味で、これからの研究者は、すべからくビッグデータ利活用の専門家、すなわちデータサイエンティストでもある必要があると考えられている。したがって、ビジネスで価値を創造し、その対価を報酬として得るタイプのデータサイエンティストとは区別される。

以上のような産業界とアカデミアそれぞれが要請するスキルをふまえると、データサイエンティストに必要なスキルとしては、

- ・データサイエンス力（統計学、機械学習、最適化などを理解し、使える力）
- ・データエンジニアリング（データサイエンスを実装する力）
- ・ビジネス力（課題を理解し、問題設定し、解決する力）

などが考えられる。結果として、データサイエンティストは方法と領域の双方に通暁したT型・II型人間であり、さらに、コミュニケーション能力や研究倫理も持ち合わせた人といえる。

産業界におけるイノベーションでもそうであるが、科学技術の世界でも複数の分野にまたがって議論するところに大きな発見や発明がありうる。データサイエンスは複数の分野の研究者が視野を共有するための共通言語になりうるし、データサイエンティストは分野を結ぶ本質的な媒介者になりうる。このため、分野と手法の様々な組み合わせを議論できるマルチディシプリンの研究機関もイノベーションには欠かせない。

データサイエンティストがこのような役割を達成するためには、彼らが求められるスキルを獲得していること以外に、国民一般の側でも、ビッグデータ分析がもたらす価値やデータサイエンティストの活動とその意味を理解するリテラシーが必要である。特に、データサイエンティストの成果をビジネスの中で活かしていくためには、その受け手となるミドルマネジメント層のデータリテラシーが不可欠であり、学校教育の成果を待つのではなく、ただちに再教育する場が必要である。一方、学校教育においては、現状の統計教育では学生に継続して興味を持たせることが難しいため、教育方法を変えていく必要がある。ビッグデータ分析が社会にどのようなインパクトを与え社会を変えつつあるか、またデータサイエンティストがどのように活躍できるのかを実例を通して学べる機会を与えるべきである。特に、データ分析の実習においては、今までの統計教育にありがちな小規模の人工（サンプル的な）データではなく、リアル感に満ちたビジネスで得られた実データを産業界等から提供してもらえれば、迫力ある実習ができ、学生の興味を掻き立てることができるのではないかとと思われる。

2) データサイエンティスト育成のあるべき姿と実現に向けた仮説

我が国におけるデータサイエンティスト育成のあるべき姿は、図1のようになると考えられる。スキルレベルとしては、対数スケールで6段階⁴としたが、それと前出のデータサイエンティスト協会スキル定義委員会が発表した4つのスキルレベル⁵との対応を図に示した。これらの人材が、大学や企業において継続的に育成され、かつ効果的に利活用されている姿を、あるべき姿と考える。なお、毎年の育成目標数は概数であり、桁としてこの程度必要であると解釈されたい⁶。

⁴ すべての大学生を想定したリテラシーレベル（50万人）から、見習いレベル（5万人）、独り立ちレベル（5000人）、棟梁レベル（500人）、業界代表レベル（50人）、世界トップレベル（5人）までの、6つのレベルで考えることにし、それぞれの育成方法を検討した。

⁵ <http://www.datascientist.or.jp/news/2014/pdf/1210.pdf>

⁶ 参考としてデータサイエンティストの必要数に関する推測（公表されたもの）を下記に記す。

- ・McKinsey Global Institute Report on Bigdata (2011): "deep analytical talent"が2018年までにUSで14万人から19万人不足。また、"data-savvy manager and analyst"については150万人不足。
- ・ガートナー(2011):「日本ではビッグデータ関連雇用が36万5000人分増える見込みだが、実際に雇用条件を満たせる人材は11万人程度しかいない」
- ・日経(2013.7.7):「国内ではデータサイエンティストが約25万人不足する」

これらの概数は、我が国の大学入学者数がおおよそ60万人であること、理系の修士入学者数がおおよそ5万人であること（見習いレベル）、資本金10億円以上の企業はおおよそ6000社であり、その企業の中で毎年1名程度ずつ採用ことを仮定（独り立ちレベル）、組織論の観点から独り立ちレベルのデータサイエンティストを指導統括リーダーは6~15人につき1人程度必要であること（棟梁レベル）⁷に基づいている。なお、指導的データサイエンティストとは、学術においてはデータサイエンスの最先端を切開くワールドクラスの研究者・開発者として指導的な能力を発揮する者、また産業界においては、業界におけるビッグデータ・データサイエンスに基づくイノベーションを牽引できる者をさし、年間数名から数十名が現実的な目標育成数と考えられる。

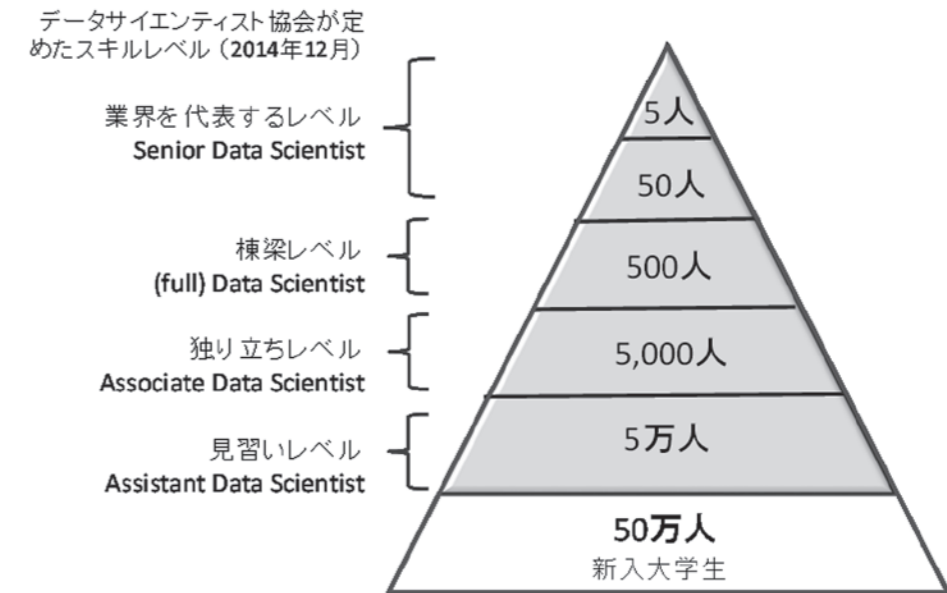


図1. 育成レベルと、毎年の育成目標人数

さて、これまでの活動を通して、我が国のデータサイエンティスト育成の現状を俯瞰すると、「棟梁レベル」の人材が育っていないことが問題と考えられる（「中抜き」仮説）。すなわち、独り立ちレベルまでは、現状でも一定程度の育成の道筋がつけられる。見習いレベルでは、データリテラシーに加え、適切な指導の下でビッグデータ利活用プロジェクトの一部を担当できるスキルが要求されることが想定される。現時点で我が国には、データサイエンスに関する学部・学科は存在しないが、滋賀大学に平成29年度設置予定のデータサイエンス学部卒業生はそのスキルを確実に習得しているであろう。加えて、一般に公開されている様々な教育プログラムでも受講者次第でスキル獲得が可能である。独り立ちレベルでは、データサイエンティストとしての専門能力、すなわちビジネス、データサイエンス、データエンジニアリングのいずれかの分野で専門的な能力を持ち、自らのイニシアチブで高度なデータ分析・問題解決能力を発揮できる人材が想定される。具体的なPBL (Project-Based Learning)がカリキュラムに含まれる予定の滋賀大学データサイエンス学部卒業生には、その素地が確実にできているであろう。ただし、高度なデータ分析・問題解決能力を身に着けるためには、博士課程修了程度の訓練が必要といえよう。しかしながら、棟梁レベルでは、データサイエンティストのチームを率いて、組織におけるビッグデータ利活用を先導できる能力をもっていることが想定される（付録A：データサイエンティスト協会スキルチェックリストも参照されたい）。つまり、複数の応用分野を俯瞰的にマネージすることができ、データサイエンスの観点から全体最適の戦略を策定し実行するリーダーシップが求められる。

この「中抜き」の状況の原因は、棟梁レベルの人材が活躍できる場がこれまで我が国になかったことと考え

⁷ Span of Control（1人の管理職がマネージする直接の部下の数）は業種や組織によって異なるが、平均6~7人と考えられている。GE社のガイドラインは、10~15人とされている。

られる。アカデミアにおいては最先端科学技術に関わる研究プロジェクト推進におけるデータサイエンスの重要性は、かねてからある程度認識されていたと思われるが、これまでの研究プロジェクトでは、データサイエンティストは分野を専門とするプロジェクトリーダーの指示に従いつつ研究開発を行うことが通常であり、棟梁レベルのデータサイエンティストがリーダーシップを発揮して、複数の分野の専門家とともに革新的な研究開発をすすめる機会が十分に存在しなかった。同様の問題は、部門縦割り主義の弊害が叫ばれている産業界においても多く見られ、データサイエンスを横串としたイノベーションを妨げる壁となっている。この状況は、企業におけるビッグデータ利活用に責任を持つ CAO (Chief Analytics Officer) が急速に普及しつつある米国⁸とは大きく異なる。

我が国におけるこの「中抜き」の状況は、すみやかに解消しなければならない。棟梁レベルの人材が年 500 人規模で育てば、その中から世界的なトップタレントが現れてこの分野全体を引っ張っていくことが期待でき、また同時にこの棟梁レベル人材が、実務の傍ら随時大学や企業で人材育成を支援することで、独り立ちレベル以下の人材育成が促進され、スケールアウトが進むことを期待することができる (図 2)。

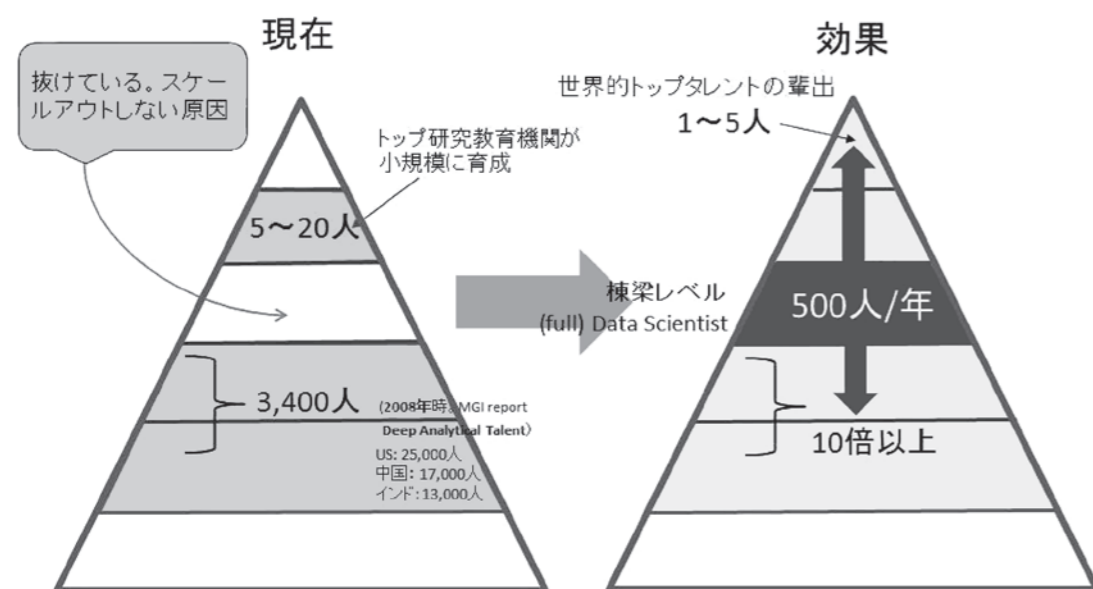


図 2. 育成が遅れ、最も欠けている棟梁レベル(左)と育成実現時の波及効果(右)

3) データサイエンティスト育成の具体的施策

前項で述べた「あるべき姿」を実現するための具体的施策の例を示す。これらの施策は個別に実施すべき性質のものではなく、適切に組み合わせることで量と質の双方を担保した適切なポートフォリオを考えるべきである。

- 社会全体のデータリテラシーやアウェアネスの向上
 - ・ 教養教育の実施、コンテストの開催、MOOC (Massive Open Online Course) などのオンライン教材や映像素材の充実などに取り組む。
 - ・ 高校教育・大学教養での講義において、世界がデータとその利活用で大きく変わっている重大性を教える。
 - ・ 大学基礎教育にデータサイエンスを取り入れる。
- 見習いレベルの育成 (5 万人規模)

- MOOC (Massive Open Online Course) を積極的に利用すると共に、コンテンツを充実させ、コース終了証を発行する。
- 社会人に対しては、特に機械学習など新技術に関して、ミドルマネジメント層を含む広い対象に、全国的に再教育のプログラムを提供する。
- 学部・大学院におけるデータサイエンスの教育カリキュラムに関しては、統計教育参照基準・情報教育参照基準に相当するものを早急に作成する⁹。
- 大学院に、ダブルディグリー・ジョイントディグリー¹⁰の考え方を導入し、専門科目と共にデータサイエンスを副専攻などの形で学べるようにする。
- 独り立ちレベルの育成 (5000 人規模)
 - ・ 企業との連携を通して、企業の事例を PBL 化したり、インターンシップ・プログラムを通して実務を学べる機会を与える。産学連携では経団連主体の CcFIL (高度情報通信人材育成支援センター)¹¹が参考になる。
 - ・ 問題設定能力、問題解決のための戦略立案能力、データの収集・キュレーション能力、データ分析結果の業務や事業への実装能力、異分野研究者や事業者との連携・コミュニケーション能力、研究倫理、情報セキュリティの能力を備え、独立してデータサイエンスを推進できるレベルを目指す¹²。
 - ・ 大学院において PBL に基づく専門育成プログラムを推進する。「分野・地域を越えた実践的情報教育協働ネットワーク(enPiT)¹³」、「先導的 IT スペシャリスト育成推進プログラム¹⁴」などにおける PBL の取組が参考になる。
- 棟梁レベルの育成 (500 人規模)
 - ・ データサイエンティストとしての実務経験のある社会人を、棟梁レベルに育成する集中的プログラムを提供する¹⁵。ここでは、最先端の手法を PBL を通じて実地で経験し、各応用領域での最新の成果をケーススタディとして学ぶと共に、拠点の特質を活かして棟梁レベルデータサイエンティスト間の人的ネットワークも形成する。
 - ・ 科学の諸分野において、既にビッグデータを活用した研究方法論をある程度身につけた人材 (博士号取得者等) に対して、産業界やアカデミアにおいて活躍できるキャリアパスを見据えて再教育する。
- 指導的データサイエンティストの育成 (50 人規模)
 - ・ 指導的データサイエンティストは、ワールドクラスの人材であり、系統的に育成するのは難しい。むしろ施策としては、世界最先端の手法・応用の研究・開発を推進することによって、このような指導的データサイエンティストが生まれてくる土壌を醸成することが大切である。
 - ・ 同時に、このような才能のポテンシャルを持つ者を若いうちに、発掘し、十分な機会を与える必要がある。このための施策として、以下の 2 点を提案する。一つは、定期的なデータサイエンス・ハッカソンを実施することである。IT セキュリティの分野では、同様の試みが IPA 「セキュリティ・キャンプ¹⁶」として行われている。これを参考に、データサイエンスにおけるプログラムを企画する。二つ目は、産業界で活躍する指導的データサイエンティストになれる人材を育成するために、才能のある

⁸ 米国における CAO の現状については、例えば Computerworld の以下の記事を参照。
<http://www.computerworld.com/article/2688352/chief-analytics-officer-the-ultimate-big-data-job.html>

⁹ HP の一覧で確認できるだけでも現在 279 のデータサイエンス教育プログラムが存在する (付録 E: 提供されている関連講座・教育プログラムのリスト (平成 27 年 12 月 14 日現在))。そのうち、修士レベルが 214 ある。

¹⁰ Double Degree は 2 つの大学が 1 つの学位論文に対して異なる学位記を発行すること、Joint Degree は、1 つの学位記に対して 2 つの専門分野を認定すること、である。これらに限らず、学位記とは別に、修了書を発行する形でもよい。

¹¹ <http://www.cefil.jp/>

¹² 日本学術会議 提言「ビッグデータ時代に対応する人材の育成」に、米国の修士プログラムのカリキュラム例が掲載されている (<http://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-22-t198-2.pdf>)。

¹³ <http://www.enpit.jp/>

¹⁴ http://www.mext.go.jp/a_menu/koutou/it/

¹⁵ 社会人に 1 年間の集中的な教育をほどこす情報学研究所「トップエスイープログラム」が参考になる。

¹⁶ <https://www.ipa.go.jp/jinzai/camp/index.html>

若い者にメンターをつけ、資金等の援助を与えると共に人脈を形成する機会を提供する。例えば IPA 「未踏 IT 人材発掘・育成事業¹⁷」や JST 「さきがけ」を参考に、指導的データサイエンティスト候補者向けのプログラムをつくる。

- トリクルダウン効果とスケールアウト
上記の具体的な施策の中で、最も重要なのが、現在わが国には決定的に不足し、ボトルネックとなっている棟梁レベルの育成であり、このレベルの人材が育ってくれば、独り立ちレベルのための PBL の指導者、見習いレベルのための専門科目の講義、リテラシーのための講師派遣などを通して、データサイエンティスト育成の全体像がスケールしていくものと考えられる。このためには産業界とアカデミアの密接な連携が欠かせない。
- 出口戦略
育成された人材は、民間・アカデミアを問わず広く活用されなければならない。このためには、ビッグデータ利活用とデータサイエンティストの重要性を広く社会に発信していくことが重要である。また、データサイエンティスト教育を受けた者、あるいはある一定のスキル基準を満たす者に対してスキル認定を行い、雇用する側とデータサイエンティストのスキルのミスマッチが起こらないようにすることも重要と考えられる。

3.2. 人材のローテーション

平成 26 年度事業では、以下の 3 種類のインターンシップをデザインするとともに、インターンシップにおける知的財産権の扱い等のノウハウを整理した。その結果、一つの成功パターンを作ることができたと考えられたため、実際に民間企業（株式会社アカリク）へ事業移管した。

タイプⅠ：特定のデータの深い分析を行う。研究開発等における高度な計算に基づいたデータ分析の現場を経験することによって、学生は専門教育で身につけた知識やスキルが実際にはどのように役立って社会に還元されていくのか学ぶ。

タイプⅡ：データ分析プロセスをひと通り経験する。例えば、CRISP-DM（SPSS、NCR、クライスラー、OHRA などにより確立されたデータ活用の標準プロセス）では、データサイエンティストのワークフローとして、1) ビジネス要件定義、2) データの理解、3) データの準備／加工、4) モデル作成、5) モデル評価、6) モデルの展開（改善方策の提案や効果検証など）が挙げられている。このタイプでは、こうした一連のワークフローを学ぶことによって、最終的にはデータ分析とビジネスを橋渡しするための知識や経験を積むことができる。

タイプⅢ：データに基づくビジネス上の意思決定現場を経験する。このタイプは、タイプⅡよりビジネスサイドに特化している。例えば、経営意思決定や顧客が抱える課題の解決に、データ分析結果がどのように役立てられているのか学ぶ。

株式会社アカリクは、本事業で作上げたフォーマットにしたがって、インターンシップ・プログラム合同説明会を平成 27 年 5 月 30 日（土）に実施した。この合同説明会には受入予定企業 6 社に対して、約 50 名の学生が参加し、ビジネスとして十分成立しうることが推測された。平成 26 年度事業では、受入予定企業 11 社に対して、約 70 名の学生が参加した。受入予定企業が前年度よりも少なかったのは、企業側の採用スケジュールの変更が影響したと思われる。

一方、本事業においては、インターンシップ・プログラムを新たに 2 つの方向へ拡大した。一つは、データサイエンティストのインターンシップが IT ベンダーに集中しがちであったため、異分野異業種交流会と銘打っ

¹⁷ <https://www.ipa.go.jp/jinzai/mitou/>

て 10 を超える学協会の共催・協賛を得て、製造業、通信業、小売業、シンクタンク、国立研究機関等に参加してもらい、データ分析に関わるインターンシップの説明会を行った。もう一つは、データ分析ハッカソンである。前節で述べたように、IT セキュリティの分野では、同様の試みが IPA 「セキュリティ・キャンプ」として行われている。これらを参考に、データサイエンスにおけるプログラムを企画した。

異分野異業種交流会

「ビッグデータやデータサイエンスに関わる学生／研究者、企業の異分野異業種交流会」を平成 28 年 1 月 31 日（日）に開催した（於 東京大学本郷キャンパス 山上会館）。プログラムの概要は以下の通りであった。「企業と学生／研究者の座談会」において、各企業・機関には、インターンシップ・プログラムの説明を要請した。

基調挨拶	文部科学省科学技術・学術政策局 坂本修一
基調講演 1	情報通信研究機構 統合データシステム研究開発室 村山泰啓
基調講演 2	ドワンゴ人工知能研究所 山川 宏
参加企業ショートプレゼンテーション	イオンマーケティング株式会社 株式会社インテージ 自然科学研究機構 国立天文台 ソフトバンクグループ株式会社 株式会社デンソー 株式会社ブレインパッド みずほ情報総研株式会社 株式会社リクルートコミュニケーションズ
パネルディスカッション	「大学でのデータ取り扱いの経験や知見は活かせるか？ー研究とデータサイエンス関連の仕事の関係性」 モデレーター：日本統計学会 会長 岩崎 学 パネリスト：NTT コミュニケーションズ株式会社 泉谷知範 株式会社構造計画研究所 滝 勇太 日本アイ・ビー・エム株式会社 辻 智 日本電気株式会社 石川 開
企業と学生／研究者の座談会	

以下の通り、民間企業 11 社、研究機関 1 機関が参加した。

日本アイ・ビー・エム株式会社
株式会社ブレインパッド
株式会社構造計画研究所
ソフトバンクグループ株式会社
株式会社デンソー
イオンマーケティング株式会社
NTT コミュニケーションズ株式会社
株式会社リクルートコミュニケーションズ
株式会社インテージ
日本電気株式会社
みずほ情報総研株式会社
自然科学研究機構 国立天文台

また、参加学生（社会人含む）は、79名であった（ただし申込者は124名）。参加学生（社会人含む）の内訳を図3に示す。半数以上が修士課程の学生で、博士課程の学生を含めると全体の80%に相当する63名が大学院生であった。わずかではあるが、PDや社会人の参加もあった。

参加学生（社会人含む）の専攻の内訳を図4に示す。情報学を専攻する学生が最も多く、全体の90%以上の参加者は理工学系・医歯学系であった。また、社会科学を専攻する参加者は、全体の10%程度であった。

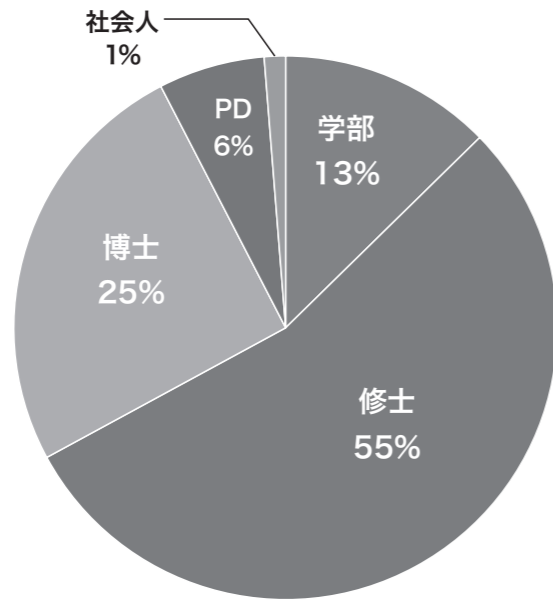


図3. 参加学生（社会人含む）の内訳（n=79）

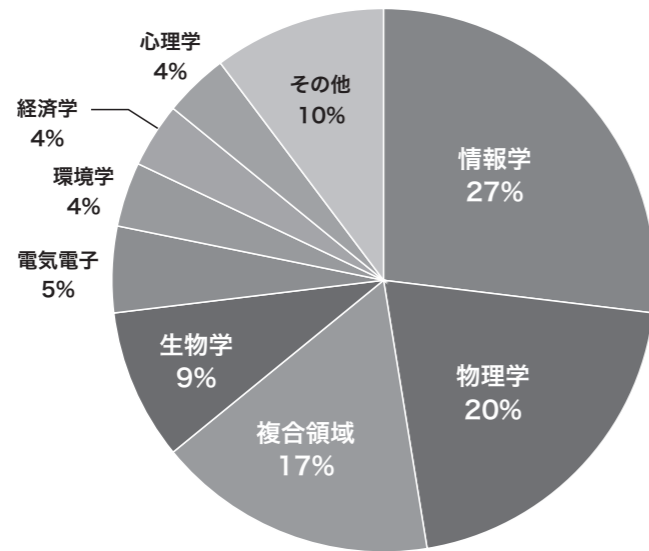


図4. 参加学生（社会人含む）の専攻（n=79）

インターンシップの実施時期は、各企業・機関で異なった。インターンシップをこの春に実施予定なのは2機関、夏に実施予定なのは4機関であった。残りの6機関は、随時行っているか、受入部門次第で実施のため

未定であった。2016年卒の就職活動時期が後ろ倒しになったことにより、採用側の対応が流動的になっていることが影響していると推測される。

79名の参加学生（社会人含む）に対して、簡単なアンケートを実施したところ（付録B：異分野異業種交流会 学生・若手研究者参加アンケート）、57名から回答を得た。この交流会では、データサイエンティストのインターンシップがITベンダーに集中しがちであることから、製造業や小売業等の企業も含めて幅広い業種におけるデータ分析に関わるインターンシップの説明会を行った。この試みに対し、今後も同様な取り組みを継続したほうが良いかどうか尋ねたところ、96%の参加者から「継続したほうが良い」という回答を得た（図5）。

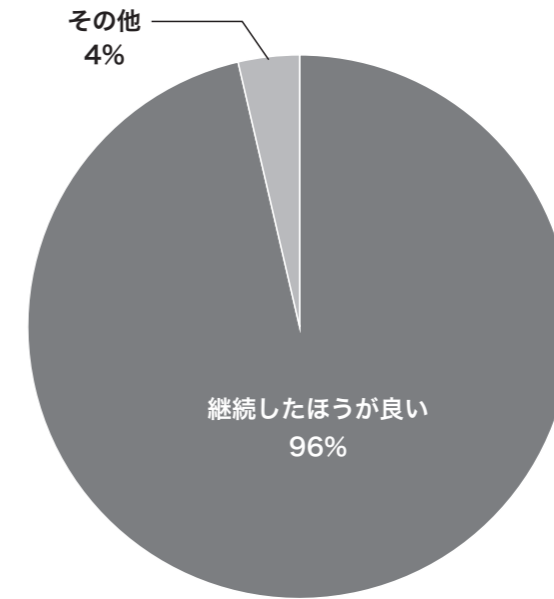


図5. 異分野異業種交流会という取り組みについて（n=57）

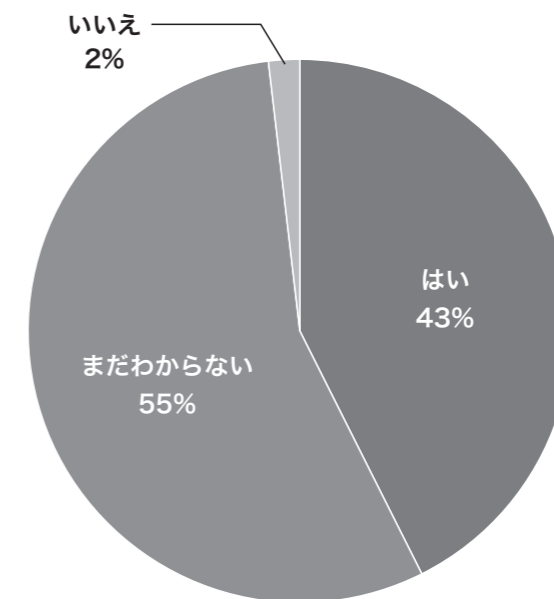


図6. データサイエンティストという職業に就きたいか（n=57）

また、データサイエンティストという仕事に就くことを考えているかどうか尋ねたところ（図 6）、43%の参加者が職業としてデータサイエンティストを選ぶ予定であると回答した。続けて、そのために何をしているか尋ねたところ、

- ・ 統計学や機械学習の勉強
- ・ R の学習、利用
- ・ Python の勉強
- ・ プログラミングの学習

というように座学が中心になっていることが伺えたが、中には、

- ・ プログラミングコンテストへの参加
- ・ POS データ等様々なデータに触れる
- ・ 「分析」にかかわるインターンシップ参加

というように座学の先の準備をしている学生もいた。さらに、

- ・ 自分の特性を見つめ、それに合ったデータサイエンティストの職種を探している

という回答もあった。これは、様々なデータサイエンティスト像がありうるという現状をふまえたコメントと解釈できる。

データ分析ハッカソン¹⁸

データサイエンティスト育成には、座学で学べる知識だけではなく、実際にデータに触って分析する経験を通してしか学べないことも多い。このため、我々はインターンシップ・プログラムを活用して、学生に企業におけるデータ分析の実際を経験してもらう試みを推進してきた。これに加えて、我々自身が実際に学生にデータ分析の機会を与える方法の1つとして、平成 28 年 2 月 20 日(土) から 21 日(日)にかけての 2 日間、統計数理研究所においてデータ分析ハッカソンを実施した。

近年では、多くのデータ分析コンテストが行われている。現在我々の事業で把握しているデータ分析コンテストは 37 にのぼり（付録 D：データサイエンス関連コンテストのリスト（平成 27 年 12 月 10 日現在））、コンテストに参加して腕を磨いたり、上位入賞することで就職を有利にしようという参加者に人気を博している。これらのコンテストには、予測精度を競うもの¹⁹のように、予め正解が決められているものと、データ分析の結果から新しい提案を求めるもののように、問題設定がオープンなものに分類することができる。また、コンテストのもう一つの分類軸として、統制された環境で分析・開発を行うのか、参加者に分析・開発環境を自由に選んで良いのか、の違いがある。

これら様々なコンテストの形態を表 1 にまとめた。今回のデータ分析ハッカソンにおいては、

表 1. コンテストの分類

分析環境	正解のある問題設定	オープンな問題設定
統制されている	世界学生プログラミングコンテスト	enPIT「クラウドアプリケーション開発演習」
オープン	The Data Science Bowl, Netflix Prize	ビジネスデータ創造コンテスト, スポーツデータ解析コンペティション等

- ・ データサイエンティスト育成の観点からは、創造性を重視するオープンな問題設定が望ましいこと
- ・ 2 日間という短期間で行うハッカソン形式とすること

¹⁸ 本項の詳細な内容は、情報処理学会デジタルプラクティスに投稿している（丸山 宏・神谷直樹・宮園法明クラウド環境を利用したデータ分析ハッカソンの計画と実施. 情報処理学会デジタルプラクティス. (査読中))

¹⁹ <http://www.netflixprize.com/>

- ・ データ提供者の強い要望により、データのセキュリティを重視したこと

から、enPIT「クラウドアプリケーション開発演習」に見られるような、オープンな問題設定かつ、統制された分析環境の形態を採用することとした。

参加対象者はデータ分析を学んだ学生（情報系、自然科学系などで、大学院・学部は問わない）とし、1 チーム 3 名、6 チームを想定して準備を開始した。データとしては、あるエンターテインメント系企業からおおよそ 1.5 年分、合計約 1 千万レコードのデータを提供していただいた（企業名と、データの詳細については、提供元の要請により公開していない）。実施は、平成 28 年 2 月 20 日(土) 9:00 から翌日 21 日(日) 18:00 まで、場所は統計数理研究所のセミナー室 1（スクール形式で 100 名が着席できる広さの教室）で行った。

1) 分析環境

分析環境としては、統計数理研究所共用クラウド計算システムを利用した。このインフラの上に、各チームに 1 つずつのクラウドインスタンスを提供することとした。また、運営側のテストとバックアップも含めて、全部で 8 個の同一環境インスタンスを用意し、これらをセキュリティ上の観点から 1 つの仮想サブネットに配置した（図 7）。この仮想サブネットは、インターネットをはじめとする外部のネットワークとは、論理的に切り離されている。参加者は自分で持ち込んだ PC を利用して、WiFi ルータ上に構成された仮想ルータから VNC 仮想デスクトップ経由で各インスタンスにアクセスした（通常の SSH によるログインは許さない）。

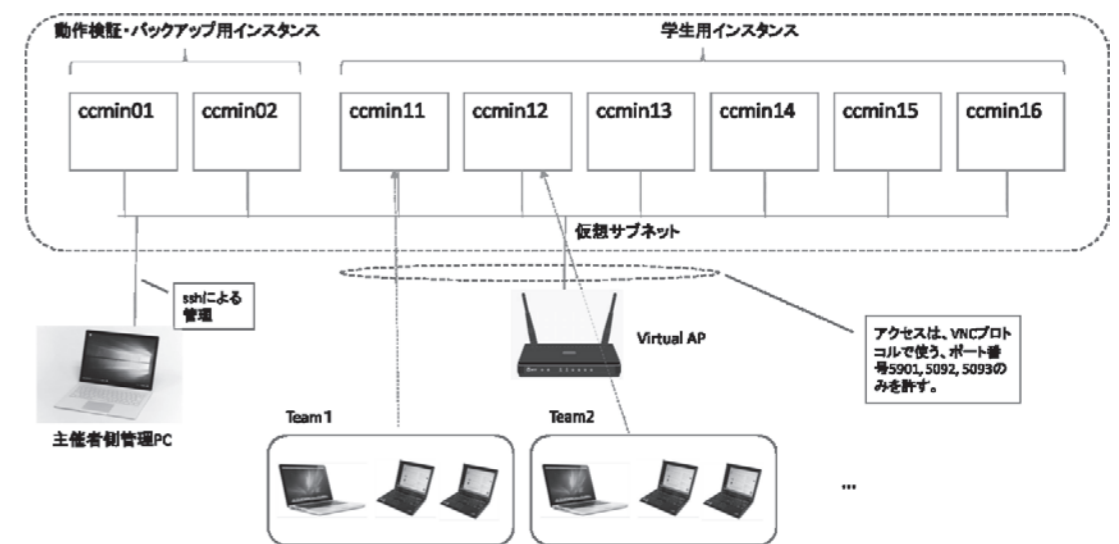


図 7. データ分析ハッカソンのネットワーク環境

各インスタンスのソフトウェア構成を図 8 に示す。OS は、Fedora23 の上に Spins+MATE デスクトップが動く、標準的な Linux 環境である。分析対象データは、MySQL 互換のデータベースである、MariaDB のテーブルとして格納した。分析用のツールとしては、下記のものを用意した。

- ・ MySQL Workbench（データベース用の GUI）
- ・ R（統計分析パッケージ）/ RStudio（その GUI）
- ・ Python 2.7 / Anaconda（Python 用の統計分析ライブラリをパッケージしたもの）/ Spyder（その統合開発環境）

さらに、プレゼンテーション用として、オフィス統合環境である LibreOffice を提供した。いずれもオープンソースであり、GUI ベースで動く。

2) AWS による事前テスト

ハッカソン当日に初めてこの環境にアクセスするのであれば、環境に慣れるまでに、無駄に時間を消費する

可能性がある。このため、アマゾン AWS 上に同等の環境を構築し、ハッカソン実施の 2 週間前に参加者に通知して、事前テストを行ってもらった。

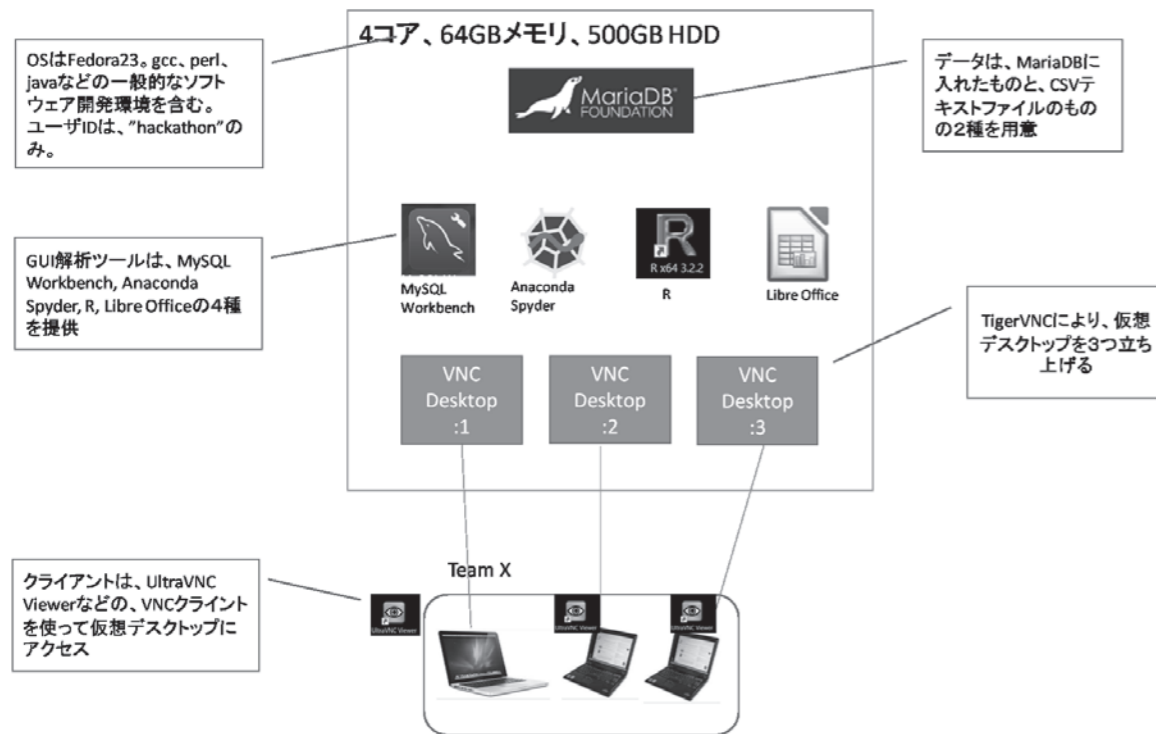


図 8. クラウドインスタンスの概要

3) 当日の進行

ハッカソンは、2 日に渡って行われた。はじめにデータ提供会社の担当者から、その会社におけるビジネスの概要と、データの詳細について解説してもらった。エンタテインメント系ビジネスということもあり、参加者の学生はビジネスの内容について、比較的スムーズに理解できたものとする。その後、計算環境についての説明を行い、実際に参加者のパソコンをネットワークに接続してもらった。Windows の PC においては、エンタープライズ WPA2 に基づく WiFi の設定に若干戸惑う場面も見られたが、20 分ほどで全員が VNC を通してインスタンスに接続することができ、10:00 にはすべてのチームがデータの分析を開始することができた(図 9)。



図 9. ハッカソン実施の様子

参加者に対しては、データを分析し、売り上げを向上させる施策を提案することを求めた。2 日目の午後 3

時から、各チームに対して 10 分のプレゼンテーションを求め、それを産業界 2 名、アカデミア 2 名計 4 名からなる審査員が聴いて審査を行った。審査の基準としては、創造性、有用性、技術力、表現力の 4 つの観点について、0-5 の範囲で採点し、それを集計した後、審査員全員の討議を経て、最優秀賞、優秀賞、審査員特別賞の 3 賞を決定した。

4) 参加者からのフィードバック

ハッカソン終了後、参加者からのフィードバックを依頼した。多くの参加者には、このハッカソンを有意義なものとして捉えていただけたようだった。いただいたポジティブな意見としては、以下のようなものがあった。

- ・ 他のチームのアイデアや手法を知ることによって、いろいろな見方があり参考になった
- ・ 実務とデータ分析の関わりを体験することができた
- ・ チーム作業を通して、リーダーシップの重要性を実感した
- ・ 限られた時間で結果を出すことの難しさを学んだ

一方、いくつか改善につながる点をご指摘いただいたので、それらを記す。

- ・ 他チームとの交流の時間がもっと欲しかった
- ・ 夜間にもデータ分析を続けさせて欲しかった
- ・ ツールがクラッシュするなど不安定なものがあった

3.3. ベスト・プラクティスの調査

佐賀県の取り組み

平成 26 年度に引き続き、組織の中でデータサイエンティストをどのようにに活用するかについて調査を行った。特に、データ分析を政策決定に活かそうという試みを始めた佐賀県について、以下の日時で Skype による追跡調査を行った。

- 平成 27 年 6 月 25 日 11 時 00 分～12 時 00 分
- 平成 27 年 10 月 7 日 13 時 30 分～14 時 20 分
- 平成 27 年 11 月 25 日 15 時 30 分～16 時 10 分

佐賀県では、平成 26 年の夏に、経験や勘だけではなく、データ分析に基づいた効果的な施策を打ち出したいとの思いからデータ活用プロジェクトがスタートした。このプロジェクトは、佐賀県が持つ公共データ等を収集・分析し、政策立案・評価等へ活用するとともに、オープンデータを推進し、新たな県民サービスの創出等の促進を目標としている。このプロジェクトの推進にあたっては、アクセンチュア株式会社をパートナーとしている。県が持つ公共データ等を収集・分析し、政策立案・評価等へ活用していく手始めとして、医療、消防・防災、商工・観光などの多岐にわたる領域で実データの分析と利活用の検討が進められている。

平成 26 年度の調査では市内全体の意識改革が課題として残されていた(「点と点が繋がるのではなく、面的な広がりを持ってデータとデータ分析の利活用が浸透するのが主な目標の一つ」)が、平成 27 年度は具体的な研修がスタートした。佐賀県での研修は、データに基づく意思決定を進めるためにはデータ分析担当者はもちろんのこと、広い視野を持ち、データをもとにした課題定義や意思決定を行おうとする各事業担当課の意思決定者の醸成も必要と考え、研修を 3 回に分け、回毎に主なターゲットを設定し、実施されていた。第 1 回は各事業担当課の意思決定者及びデータ分析担当をマネジメントする職員、第 2 回以降は実際にデータ分析を担当する職員や彼らをマネジメントする職員をターゲットに行われていた。教材は、県とアクセンチュア株式会社が共同で作成した。また、自治体で利用するデータや扱うケースはビジネス現場でのそれらとは明らかに異なるため、アクセンチュアと県庁側が掛け合いをするような形(アクセンチュアと県庁担当者が共に講師)で行っていた。この研修の効果が現れるのはこれからと考えられるが、職員には大いに反響があったことから、プ

プロジェクト担当者は職員のリテラシーの変化に一定の手応えを感じているようである。特に、統計分析課、健康増進課などの担当者は、データ分析を「自分の課題」として認識しているようだ。

一方、プロジェクト自体の進捗として注目すべきなのは、肝炎プロジェクトである。佐賀県は、肝がんによる死亡率が全国でワースト1である。そのため、データ利活用による打開策が検討され、データ分析結果に基づく啓発パンフレットを2種類作成し（県民用と医療機関用）配布した²⁰。この施策の評価は今後の課題として残されているが、「99さがネット」に続いて具体的な施策に結びついた成功例といえるだろう。

このプロジェクトを主導してきた戦略コンサルティング経験者の県庁担当者は、いわゆる「データサイエンティスト」ではないかもしれないが、ビッグデータ利活用人材育成懇談会のいうところの「棟梁レベル」の仕事をしているといえる。彼は複数の応用分野を俯瞰的にマネージし、データサイエンスの観点から全体最適の戦略を策定し実行するリーダーシップを発揮していた。懇談会報告書の「中抜き仮説」が正しいとすれば、この担当者のような人材を育てることが急務かもしれない。

エンターテインメント系民間企業の取り組み

データサイエンティスト利活用側のベスト・プラクティスとして、エンターテインメント系民間企業における取り組みについても調査した。この企業では、データサイエンティストを雇用するのでもなく、佐賀県のようにデータ分析に基づくコンサルティングファームをパートナーとするのでもなく、インターンを受け入れるところから、データ利活用が自社のビジネスにどのようなイノベーションをもたらしうのか検証している。

平成26年8月初め～8月末まで、2名のインターンを受け入れ、自社データの分析から施策の提案までデータ分析プロセスをひと通り経験させた(本事業でデザインしたインターンシップのタイプⅡに相当する)。また、インターンは、3日に1回程度、複数のメンターとの打ち合わせを行い、データ解釈のフィードバックを受けていた。また、データが生み出される現場も体験させてもらい、データからは得られない気づきを得ることができた。

これから初めて、データ分析をビジネスプロセスの中に組み込んでいこうとする企業にとって、その意義を社内で共有することは難しいことなのかもしれない。佐賀県でそうであったように、データリテラシーの向上が必然的に問われることになるからである。しかしながら、この企業では、インターンからの施策提案がそれまでのKKD（勘と経験と度胸）による経営では得られなかった示唆が得られたとして、インターンを引き続きアルバイトとして雇用して、データ分析の意義を引き続きビジネスプロセスに取り入れようと模索している。

本事業インターンシップ参加者の就業状況

平成25年度のインターンシップ・プログラム参加者のうち、この4月からデータサイエンティストとして就職した方のフォローアップを10月に行った。この方は現在、レコメンドエンジンを構築するチームの一員として、民間企業に勤務されている。実務においては最新の知見を実際に実装していくスキルが求められるため、学生時代にあまり必要でなかった、開発環境まわりのことで苦労されているようであった。ただし、こうした知識は必要な時に身につければよく、むしろ論文の読み方、確率・統計の基本的な部分の知識、研究を通して大いに鍛えられた物事を整理して考える能力が、現在の仕事に大いに役立っているようであった。

現時点では、データサイエンティスト像に流動的な側面があることを考慮すると、スキルレベルの継続的な見直しでは、実務におけるより多くの事例を集めて参考にすることが必要かもしれない。実際に、この例では、見習いレベルと独り立ちレベルの間にあるような人は、実際の業務を通じてスキルアップしており、外部からの教育プログラムの提供を必要としていない。しかしながら、業種その他の要因によって、産業界全てで同じことがいえるとも限らない。データサイエンティストを資格化する場合には、実務におけるスキルアップと

何らかの教育プログラムで得られたスキル等を、どのように評価していくのか検討できる枠組みが必要になるかもしれない。

3.4. 育成教材の開発

YouTube版データサイエンティスト・クラッシュ・コース

データサイエンティスト・クラッシュ・コースは、平成27年2月19日から無料ビデオコンテンツとして、ネット上で視聴できるようにした（YouTube版データサイエンティスト・クラッシュ・コース）。そして、内容の見直し等を経て、平成27年12月10日から改めて公開した。

公開開始から約1ヵ月毎の視聴履歴をみると、新規の1,000回以上の視聴回数が常に維持されていた(図10)。YouTubeでは、特定ユーザーが繰り返し視聴した場合、その履歴はカウントから除外される機械学習のアルゴリズムにしたがったデータをオーナーに提供しているため、新規ユーザーによる視聴が毎月されているとみなせる。

データサイエンティスト・クラッシュ・コースでは、以下に示すような9つのコンテンツを提供しているが、コンテンツによって視聴回数割合の変動に違いがあった。

- ・ コース概要
- ・ データサイエンティストとは
- ・ データ解析基礎
- ・ データ可視化とツール
- ・ 統計的モデリングと機械学習
- ・ 統計的時系列モデリング
- ・ 最適化
- ・ データ分析と意思決定
- ・ データ分析の知的財産権

中でも特に、「統計的時系列モデリング」と「コース概要」は視聴回数割合が増加傾向にあった。「コース概要」の視聴回数割合が常に高いことは、このデータサイエンティスト・クラッシュ・コースが多くの方から関心が寄せられていることの表れといえよう。

また、データサイエンティスト・クラッシュ・コースへのトラフィック・ソースをみると(図11)、「関連動画」からのトラフィックだけが増加傾向にあった。「関連動画」とは、コンテンツ視聴時にレコメンド表示される動画のことである。「関連動画」からのトラフィックが増加傾向にあるということは、データ分析関連の他のコンテンツを経てデータサイエンティスト・クラッシュ・コースを視聴しているユーザーが増加していることを意味している。したがって、データサイエンティスト・クラッシュ・コースは、データサイエンティストに求められるであろう内容を1つのパッケージにしているコンテンツとして、着実に有効利用されてきているといえるだろう。

²⁰ http://www.pref.saga.lg.jp/web/kurashi/_1019/kanen/kanentaisaku.html

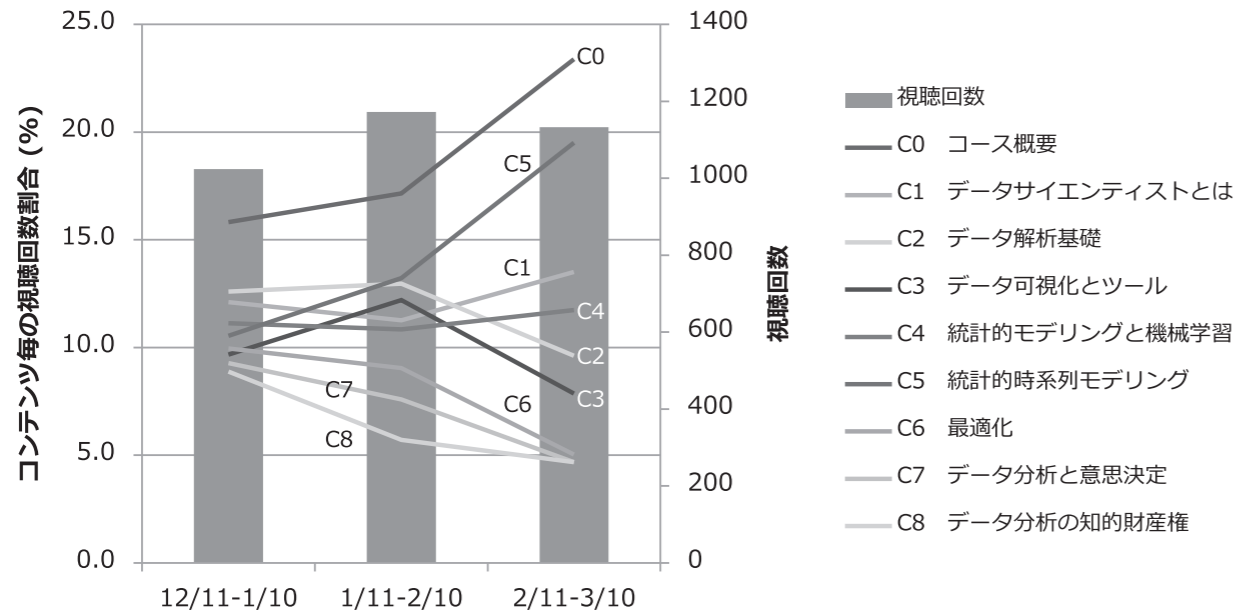


図 10. リニューアル後の視聴回数

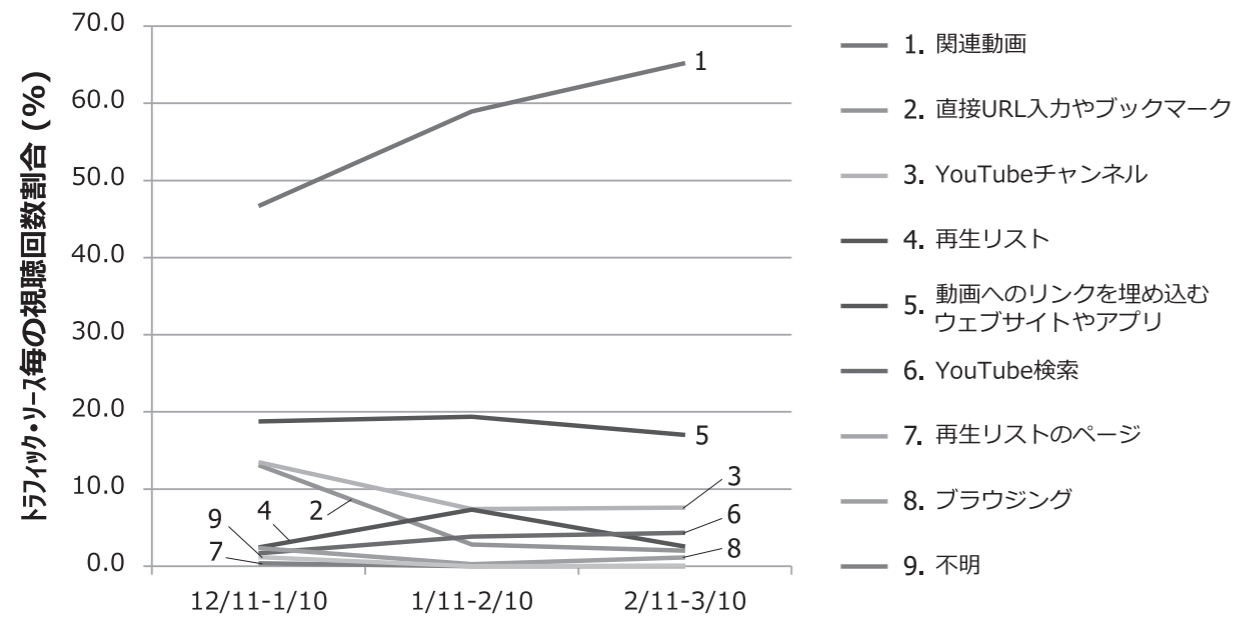


図 11. トラフィック・ソース毎の視聴回数割合（リニューアル後）

提供されている関連講座・教育プログラムのリストと、データサイエンス関連のコンテストのリスト

大学・民間を問わず多くのデータサイエンティスト育成コースが開講されているが、271 の教育プログラムについて一覧（付録 C：提供されている関連講座・教育プログラムのリスト（平成 27 年 12 月 14 日現在））を作り、事業のホームページで公開した²¹。このデータベースには、以下の情報を纏めてある。平成 26 年度版の情報に加え、利用者の利便性を考慮し、「教育プログラムの目標」、「対象者」と、可能な限り詳細な「教育プロ

²¹ <http://datascientist.ism.ac.jp/trainingaids.html>

この報告書の付録は、事業ホームページで公開しているリストの概略になっている。

グラムの内容」を追加して纏めた。

- ・ 教育プログラム（講座／教材名）
- ・ 教育プログラム提供元
- ・ 教育プログラムの問い合わせ先
- ・ 教育プログラム URL
- ・ 教育プログラム概要
- ・ 教育プログラムの目標
- ・ 対象者
- ・ 教育プログラムの内容
- ・ オンラインで受講可能か否か
- ・ 実習等の有無
- ・ 受講前提条件の有無
- ・ 学習期間
- ・ 受講料金
- ・ 事業インターンシップ類型との対応（分析者向け（タイプⅠ）／分析サービス提供者向け（タイプⅡ）／意思決定者向け（タイプⅢ））
- ・ 使用言語
- ・ その他備考

今年度調査では、昨年度から関連教育プログラムが 100 程度増加していることから、社会的な需要の高まりが示唆された。この傾向は国内と国外で同じというわけではなく、国外のプログラムのほうが爆発的に増加していることから、国外での需要のほうがより高まっているとみなせる。

データサイエンティストには座学で学べる知識だけではなく、実際にデータに触って分析する経験を通してしか学べないことも多い。そこで、教材の一つとして、データサイエンス関連コンテストのリストを纏めた（付録 D：データサイエンス関連コンテストのリスト（平成 27 年 12 月 10 日現在））。このデータベースには、以下の情報を纏めてある。

- ・ コンテスト名称
- ・ コンテスト主催団体
- ・ コンテスト URL
- ・ コンテスト概要
- ・ コンテスト参加資格
- ・ コンテスト参加費
- ・ コンテスト開催頻度
- ・ コンテスト問い合わせ先

データサイエンス関連のコンテストは、全部で 37 あるが（平成 27 年 12 月 10 日現在）、そのうち 29 が国外で開催されている。我が国でも、実情に即したコンテストが多く開催されることが望まれる。

これらのデータベースは、データサイエンティストを目指す者、あるいはデータサイエンティストの育成を考えている組織にとって有用な情報となるだろう。

3.5. 海外との連携

海外との連携を目指し、英国 Warwick 大学²²の事例調査を行った。Warwick 大学では、データサイエンスの学士号を取得することができる。平成 27 年 9 月 11 日に、プログラム全体の責任者である Prof. David Firth とカリキュラムの責任者である Dr. Anthony Lee を竹村教授が訪問した。

Warwick 大学のデータサイエンスの学士号は、英国におけるデータサイエンスの学士号として最初のものである。しかしながら、特定の学部が設置されているのではなく、統計学部とコンピューター・サイエンス学部が共同でこのプログラムを運営している。3 年間のプログラムになっており、学生は、1 年次では数学の基礎をかなり重点的に、2 年次は統計と情報の標準的な内容を学び、そして 3 年次に応用分野のプロジェクトに従事する。このプログラムへの入学は難関とされ、一学年 10 名強程度しか許されていない。ちなみに、このプログラムにおける 1 年時の必須講義は、以下の通りである。

- ・ CS118 Programming for Computer Scientists
- ・ CS126 Design of Information Structures
- ・ IB104 Mathematical Programming I
- ・ MA106 Linear Algebra
- ・ MA137 Mathematical Analysis
- ・ MA138 Sets and Numbers
- ・ ST104 Statistical Laboratory
- ・ ST115 Introduction to Probability
- ・ ST116 Mathematical Techniques

一方、米国では、統計学専攻のプログラムがデータサイエンスを意識した内容に変更されていて、統計学専攻は近年の理系学位 (STEM) の一番人気で、卒業後の給与も一番高いと言われている。したがって、英国 Warwick 大学よりも柔軟で、既存領域横断型の取り組みがなされていることが推測される。例えば、米国オハイオ州立大学では、大学独自で 150 億円投資、50 名の教員を新規採用し、2013 年から Data Analytics major を開始している。そして、様々な分野への応用（実データを使った解析、各種企業（Nation wide、JPMorgan Chase、Microsoft、IBM、Google、TATA Group、Infosys 等）との共同研究、様々なプロジェクト研究）が行われている。また、ロチェスター大学では、大学独自でデータサイエンス・プログラムに 100 億円投資、生命情報学、医用生体工学、脳科学、認知科学、ビジネス、経済学、コンピューター科学、物理学、数学、統計学、政治学等の多分野の研究者を採用し、様々な応用研究を可能にしている。そして、学士号として Bachelor of Sciences と Bachelor of Arts のどちらかを取得可能な体制が整えられている。ただし、これらの学位を取得するためには、3 つのコア科目（数学、コンピューター・サイエンス、統計学）を取得する必要がある。

以上のような調査をふまえると、我が国の実情に即した教育や資格化の在り方はどのようなものになるだろうか。その一つの形が、平成 29 年度に開設される滋賀大学データサイエンス学部といえよう。

3.6. プロジェクトの総合推進

プロジェクトの推進においては、これまでと同様、運営委員会を開催し（平成 27 年 1 月 14 日）識者の意見を取り入れながら行った。また、事業の発注元である文部科学省・日本科学技術振興機構とも密接に連携しながら、正しい事業の遂行・予算執行に最大限の留意をした。今年度は、これまでに形成してきた大学間ネットワーク、産学ネットワーク、関連団体や関連企業間ネットワークを維持した。

なお、今年度の運営委員会参加者は以下の通りであった。

伊藤 聡	統計数理研究所 教授
井川 甲作	松竹株式会社
井上 亮	株式会社アカリク
大西 立顕	東京大学大学院 情報理工学系研究科 ソーシャル ICT 研究センター 准教授／文部科学省委託事業「データサイエンティスト育成ネットワークの形成」 再委託先業務参加者
柏木 宣久	統計数理研究所 教授
神谷 直樹	統計数理研究所 特任研究員／文部科学省委託事業「データサイエンティスト育成ネットワークの形成」 委託先業務参加者
木村 正一	統計検定センター
草野 隆史	株式会社ブレインパッド
佐藤 俊哉	京都大学大学院 教授
清水 信夫	統計数理研究所 助教／文部科学省委託事業「データサイエンティスト育成ネットワークの形成」 委託先業務参加者
鈴木 律夫	情報サービス産業協会
高田 唯史	自然科学研究機構 国立天文台 天文データセンター 准教授
竹村 彰通	東京大学大学院 教授／文部科学省委託事業「データサイエンティスト育成ネットワークの形成」 再委託先業務主任者
田村 義保	統計数理研究所 教授
坪井 誠司	独立行政法人海洋研究開発機構 地球情報センター 部長
中野 純司	統計数理研究所 教授
橋本 武彦	一般財団法人データサイエンティスト協会
樋口 知之	統計数理研究所 所長／文部科学省委託事業「データサイエンティスト育成ネットワークの形成」事業 委託先業務主任者
藤山 秋佐夫	国立遺伝研究所 教授
堀 宗朗	東京大学地震研究所 教授
丸山 宏	統計数理研究所 教授／文部科学省委託事業「データサイエンティスト育成ネットワークの形成」 委託先実施責任者
三國 純一	日本経済新聞社
山島 大知	日本経済新聞社

(敬称略・50 音順)

²² <http://www2.warwick.ac.uk/fac/sci/statistics/courses/datsci/>

2015年11月20日

プレスリリース

報道関係各位

一般社団法人データサイエンティスト協会

**データサイエンティストに必要とされるスキルをまとめた
スキルチェックリストを初公開
総合人材サービス会社での活用が決定**

一般社団法人データサイエンティスト協会（所在地：東京都港区、代表理事 草野 隆史、以下データサイエンティスト協会）およびスキル委員会（委員長 安宅 和人、副委員長 佐伯 諭）は、データサイエンティストに必要とされるスキルを「データサイエンティストスキルチェックリスト（以下スキルチェックリスト）」としてまとめ、第1版を公開いたしました。また、昨年12月に公開の「データサイエンティストのミッション、スキルセット、定義、スキルレベル」について2015年版を発表いたしました。

データサイエンティストは、ビッグデータやIoTを活用する上で不可欠となる職種といわれており、国内においてもその不足が課題となっております。データサイエンティスト協会は、この課題に対し、データサイエンティストという新しい職業が健全に成長をさせるための環境整備を目的に2013年より活動を行っております。

活動の一環として昨年12月に発表した、「データサイエンティストのミッション、スキルセット、定義、スキルレベル」は、データサイエンティストとして目指すべき職種像を明らかにしたものでしたが、今回、新たに作成したスキルチェックリストは、各スキルレベルで求められるスキルの内容を具体的に整理したものととなります。

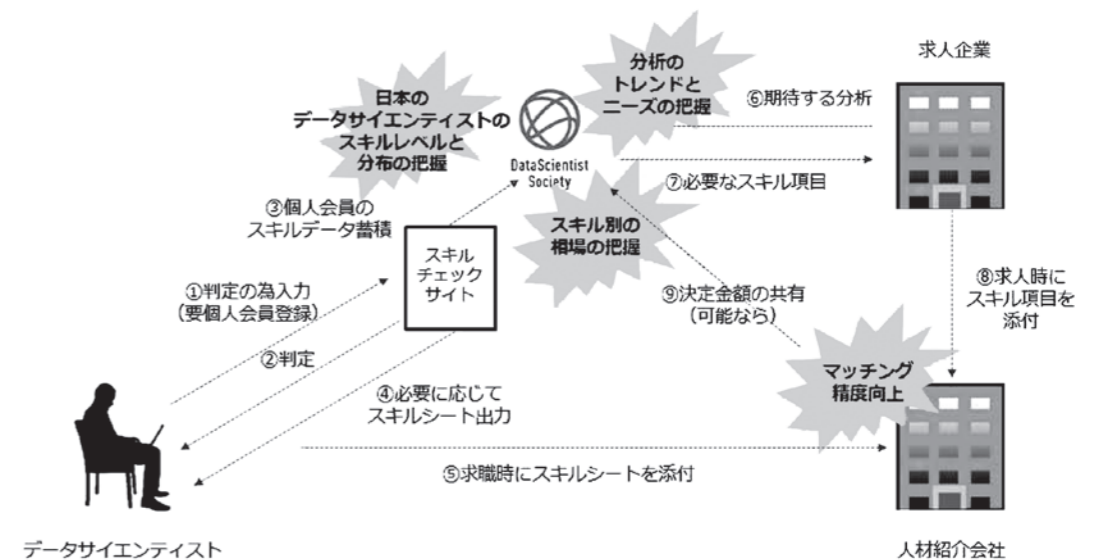
データサイエンティストが有すべきスキルを明確にしたスキルチェックリストがもたらす効果と対象は、以下のとおりです。今後、国内におけるデータサイエンティストの不足と、その定義が曖昧なことにより発生しているさまざまなミスマッチによる社会的損失を解消し、業界の健全な発展に寄与することを期待しております。

■想定する効果と対象

想定する対象	想定する効果
データサイエンティストを目指す人材	習得すべきスキルの明確化
データサイエンティストを採用する企業	獲得したい人材のスキルの明確化・明文化 同スキルの人材の待遇のバラつき解消
データサイエンティストの求職者	自分のスキルの明示とともに、 期待される仕事のレベルの事前把握
データサイエンティストを紹介する企業	求職者と求人内容のスキルによる マッチング
データサイエンティストを育成する 教育機関・企業	カリキュラムの明確化

データサイエンティスト協会は、スキルチェックリストを利用した自己診断サービス（データサイエンティスト協会のウェブサイト上でチェックすることで自身のスキルレベルを判定したチェックシートが出力される）の提供を来春に予定しており、総合人材サービス会社の株式会社インテリジェンスの人材紹介部門より、データサイエンティスト職の求人求職において同フォーマットを活用したマッチングに賛同の意を表明いただいております。

■「データサイエンティストスキルチェックリスト」の展開イメージ



■本スキルチェックリストに寄せられた声

株式会社インテリジェンス

キャリア Div.&メディア Div. マーケティング企画統括部
エグゼクティブマネジャー 木下 学 様

株式会社インテリジェンス

キャリアディビジョン マーケティング企画統括部 データアナリティクス部
マネジャー 大江 信明 様

「通常、人材紹介会社では、企業の人事部様に採用要件をヒアリングし、その要件から求人票を作成します。そして転職希望者様とマッチングの上、求人を案内していきます。今回のスキルチェックリストの利用により、企業への採用要件のヒアリングや、転職希望者のスキルのヒアリングにおいて、項目が統一されることにより、マッチングの向上に役立てることができることを期待しております。また、どのように使用したらより有効に機能するのか、引き続き検討したいと考えております。」

■スキルチェックリスト公開の背景

データサイエンティスト協会の設立の背景と目的

データサイエンティスト協会設立時の状況としては、以下の主な課題がありました。

- ・バズワードである「データサイエンティスト」という言葉の定義が欠落
- ・雇い主側の期待に雇われる側のスキルセットが一致しないケースが増加
- ・若い才能達が、自分をどう訓練し、スキルを身に着けていくべきかわからない
- ・企業としてこの新しい職種の適切な処遇や活用の方法が未確立
- ・その他、全般的に情報が不足している

この課題認識を踏まえ、データサイエンティスト協会は、以下の主な目的を果たすために設立されました。

1. この新しいデータプロフェッショナル（「データサイエンティスト」）に必要とされるスキルセットを定義する
2. データサイエンティストの育成と評価のための軸、基準を作る
3. データサイエンティストが相互に接し、情報をやりとりできる環境を提供する
4. 社会に対して、データサイエンティスト及びその仕事を啓蒙するとともに、必要な情報を発信する
5. その他、必要な仕組みづくりを推進する

スキル委員会は、昨年公開した上記の「1. スキルセット」に基づき、今回、「2. データサイエンティストの育成と評価のための軸、基準」の作成検討を進め、得られた成果をスキルチェックリストとしてまとめました。

■スキルチェックリストの概要

データサイエンティストのスキルレベル (2015年版)

データサイエンティストのスキルセット（ビジネス力/データサイエンス力/データエンジニアリング力）別に4段階のスキルレベルを定義しています。この1年間の技術・業界の進展を受けて、スキルレベル表を更新しました。（詳細は、図1を参照）

1. 業界を代表するレベル : Senior Data Scientist (★★★★)
2. 棟梁レベル : Full Data Scientist (★★★)
3. 独り立ちレベル : Associate Data Scientist (★★)
4. 見習いレベル : Assistant Data Scientist (★)

※★の数はスキルレベルを示します。

※独り立ちレベルより上位レベルは、必ずしも一人が全てのスキルを有する必要はないと考えています。全体を俯瞰してコーディネートする人材は必要ですが、個別のスキルに秀でた人とのチームを組成し、実務にあたることも現実的には多いと思っています。

スキルチェックリストの見方と活用方法

データサイエンティストのスキルセット（ビジネス力/データサイエンス力/データエンジニアリング力）と4段階のスキルレベルにあわせ、各象限におけるデータサイエンティストに必要なとされるスキルセットについて、チェックリストを整理しています。

- ・3つの領域別に判定します。
- ・”必須”に○がついているものは、「判定基準を満たしていても、これができないとそのレベルとは認められないもの」として設定しています。
- ・各レベルの判定基準の目安は、「Full Data Scientist」が50%、「Associate Data Scientist」が60%、「Assistant Data Scientist」が70%を想定しています。

スキルチェックリストは、人材育成や組織マネジメントなどで活用可能です。

- ・データサイエンティストのスキルレベルの現状把握や、現状から見た強み・弱みの把握
- ・今後のキャリアアップに向けて、どのような点を強めていけばよいかの把握など
- ・データサイエンス組織のマネジメントにおける上記育成観点や、採用など

■データサイエンティストのスキル定義の概要

(2014年12月10日発表 プレスリリース「データサイエンティスト協会、データサイエンティストのミッション、スキルセット、定義、スキルレベルを発表 (<http://www.datascientist.or.jp/news/2014/pdf/1210.pdf>)」より)

データサイエンティストに求められるスキルセット：「ビジネス (business problem solving) 力」、「データサイエンス (data science) 力」、「データエンジニアリング (data engineering) 力」

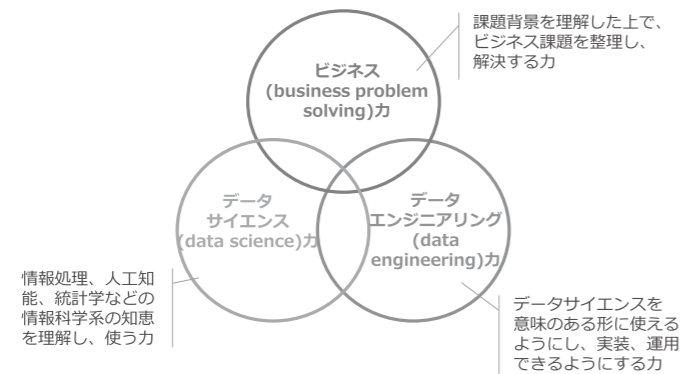
「データの持つ力を解き放つ」というデータサイエンティストのミッションを踏まえ、データサイエンティストに求められるスキルセットは3つあると考えています。

- ・ビジネス (business problem solving) 力：課題背景を理解した上で、ビジネス課題を整理し、解決する力
- ・データサイエンス (data science) 力：情報処理、人工知能、統計学などの情報科学系の知恵を理解し、使う力
- ・データエンジニアリング (data engineering) 力：データサイエンスを意味のある形に使えるようにし、実装、運用できるようにする力

図1：データサイエンティストに求められるスキルセット



データサイエンティストに求められるスキルセット



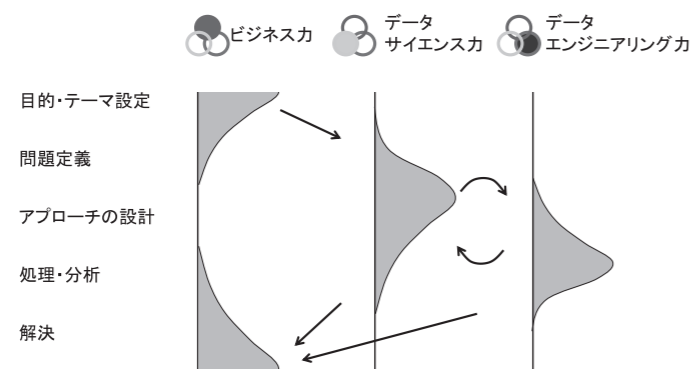
Source: The Japan Data Scientist Society discussions

※この3つのスキルはどの一つが欠けてもいけません。また、この3つのスキルは課題解決のフェーズによって中心となるスキルが変化します。(図2参照)

図2：課題解決の各フェーズで要求されるスキルセットのイメージ



課題解決の各フェーズで要求されるスキルセットのイメージ



Source: The Japan Data Scientist Society discussions

データサイエンティストの定義

必要なスキルセットの検討の結果、これからの時代に求められるデータサイエンティストを次のように定義しました。

「データサイエンティストとはデータサイエンス力、データエンジニアリング力をベースにデータから価値を創出し、ビジネス課題に答えを出すプロフェッショナル」

※ここで「ビジネス」とは社会に役に立つ意味のある活動全般を指します。

※プロフェッショナルとは、体系的にトレーニングされた専門性を持つスキルを持ち、それをベースに顧客（お客様、クライアント）にコミットした価値を提供し、その結果に対して認識された価値の対価として報酬を得る人を示します。

●スキル委員会 委員一覧 (2015年11月現在、敬称略)

	企業名	部署・役職	氏名
委員長	ヤフー株式会社	CSO (チーフストラテジーオフィサー)	安宅 和人
副委員長	株式会社電通	統合データ・ソリューションセンター データ・マネジメント部部长	佐伯 諭
委員	株式会社 ALBERT	取締役会長	山川 義介
	SAS Institute Japan 株式会社	カスタマーサービス本部エデュケーション グループ エデュケーション マネージャ	村上 元親
	株式会社インテリジェンス	キャリアディビジョンマーケティング 企画統括部 データアナリティクス部 マネージャー	大江 信明
	株式会社インテリジェンス	キャリアディビジョンマーケティング 企画統括部 データアナリティクス部	山之下 拓仁
	株式会社タクミインフォメーションテクノロジー	システム開発推進部	原茂 恵美子
	株式会社チェンジ	取締役	高橋 範光
	株式会社大和総研	スキルアップ推進部 次長	松島 寛
	データ・フォアビジョン株式会社	データ・サイエンス本部 本部長補佐	福本 信吾
	デジタル・アドバタイジング・コンソーシアム株式会社	プロダクト開発本部 ビッグデータ解析部 部長	薩摩 定壮
	株式会社電通	統合データ・ソリューションセンター データ・マネジメント部	猪飼 嘉男
	日本アイ・ビー・エム株式会社	テクニカルリーダーシップ 成長イニシアチブ推進 ビジネス・アナリティクス コンサルタント	守谷 昌久
	日本サード・パーティ株式会社	教育コンサルティング部 テクニカルマネージャ	関口 大五郎
	株式会社日立インフォメーションアカデミー	サービス企画部 主任技師	田中 貴博
	株式会社日立インフォメーションアカデミー	ビジネス研修部 技師	大黒 健一
	三菱電機インフォメーションシステムズ株式会社	産業・サービス事業本部 IT コンサルティング部	尾崎 隆
	ヤフー株式会社	データ&サイエンスソリューション統括本部 データ&サイエンスディレクター	小間 基裕
ヤフー株式会社	データ&サイエンスソリューション統括本部 データインフラ本部 本部長	佐々木 潔	

●一般社団法人データサイエンティスト協会について <http://www.datascientist.or.jp/>
データサイエンティスト協会は、新しい職種であるデータサイエンティストに必要となるスキル・知識を定義し、育成のカリキュラム作成、評価制度の構築など、高度 IT 人材の育成と業界の健全な発展への貢献、啓蒙活動を行っています。また、所属を超えてデータ分析に関わる人材が開かれた環境で交流や議論をし、自由に情報共有や意見発信ができる場を提供しています。2015年11月現在、50社6団体の法人会員と約3,050名の一般(個人)会員が参画しています。代表理事：草野 隆史(株式会社プレインパッド 代表取締役会長)、所在地：東京都港区、設立：2013年5月。

以上

<本件に関するお問い合わせ先>

●本件に関するお問い合わせ
一般社団法人データサイエンティスト協会
事務局
TEL: 03-6721-9001 e-mail: info@datascientist.or.jp

●報道関係の方からのお問い合わせ
一般社団法人データサイエンティスト協会
事務局 広報担当
TEL: 03-6721-9001 e-mail: pr@datascientist.or.jp

*本プレスリリースに記載されている会社名・商品名は、それぞれの権利者の商標または登録商標です。

*本プレスリリースに掲載されている情報は、発表日現在の情報です。

データサイエンティストのスキルレベル

	ビジネス (business problem solving) カ	データサイエンス (data science) カ	データエンジニアリング (data engineering) カ
①Senior Data Scientist 業界を代表するレベル	<ul style="list-style-type: none"> 業界を代表するデータプロフェッショナルとして、組織全体や市場全体レベルでのインパクトを産み出すことができる -対象とする事業全体、産業領域における課題の切り分け、論点の明確化・構造化 -新たなデータ分析、解析、利活用領域の開拓 -組織・会社・産業を横断したデータコンソーシアムの構築、推進 -事業や産業全体に対するデータ分析を核としたバリューチェーン創出など 	<ul style="list-style-type: none"> データサイエンスにおける既存手法の限界を打ち破り、新たに課題解決可能な領域を切り拓いている -既存手法では対応困難な課題に対する新規の分析アプローチの開発・実践・横展開 -高難度の分析プロジェクトのアプローチ設計、推進、完遂能力など 	<ul style="list-style-type: none"> データサイエンス領域で行おうとしている分析アプローチを、挑戦的な課題であっても安定的に実現できる -複数のデータソースを統合した例外的規模のデータシステム、もしくはデータプロダクトの構築、全体最適化 -技術的限界を熟知し、これまでにない代案の提示・実行 など
②Full Data Scientist 棟梁レベル	<ul style="list-style-type: none"> 生み出す価値にコミットするプロフェッショナルとして、データサイエンティストとは何かを体現したビジネス判断、課題解決ができる 初見の事業領域に向かい合う場合や、スコープが複数の事業にまたがる場合であっても本質的な課題や変数を見出し、構造化し、適切な分析・解析アプローチを設計、実行することができる 解決に必要な結果を総合した上で、説得力ある形で共有し、関連する組織、人を動かし、知見の横展開、組織を超えるつなぎ込みができる プロフェッショナルからなる複数のチームによるプロジェクトを推進し、全体としてのアウトプットにコミットできると共に、メンバーを育成、さらには持続的な育成システムを作りだすことができる 	<ul style="list-style-type: none"> 予測、グルーピング、機械学習、大量データの可視化、言語処理などの応用的なデータサイエンス関連のスキルを活かし、データ分析プロジェクトの技能的な軸を担うことができる 複数もしくは高度な分析プロジェクトを持つチームにおいて、Associate Data Scientist (独り立ちレベル) 以下のメンバーの技能を育成することができる 	<ul style="list-style-type: none"> 数十億レコード程度の分析環境の要件定義・設計、データ収集/蓄積/加工/共有プロセスやITセキュリティに関するデータエンジニアリング関連のスキルを活かし、データ分析プロジェクトを中核的に推進することができる 複数もしくは高度な分析プロジェクトを持つチームにおいて、Associate Data Scientist (独り立ちレベル) 以下のメンバーの技能を育成することができる
③Associate Data Scientist 独り立ちレベル	<ul style="list-style-type: none"> 大半のケースで自立したプロフェッショナルとしてビジネス判断、課題解決ができる 既知の領域、テーマであれば、新規課題であっても解くべき問題の見極めや構造化、その上での分析・解析アプローチの設計、実行、深掘りができる データ、分析結果に対する表面的な意味合いを超えた洞察力を持ち、担当プロジェクトの検討結果を取りまとめ、現場への説明、実装を自律的に行うことができる 5名前後のプロフェッショナルによるチームでのプロジェクトを推進しアウトプットにコミットできる -イシュードリブンでスピード感のある判断 -プロジェクトマネジメントと個別メンバーの育成 -異なるスキル分野の専門家、事業者との協働など 	<ul style="list-style-type: none"> 単一プロジェクトにおけるデータ分析をFull Data Scientist (棟梁レベル) に相談しつつ推進できる Assistant Data Scientist (見習いレベル) の日々の活動に適切な指示ができる 基礎的な分析活動については、自律的に実施できる -調査目的に合った標本抽出計画の立案 -多重 (質問間) クロス表を駆使した洞察 -パラメトリックな2群の検定の活用 (t検定) -主成分分析や因子分析 -適切な学習データとテストデータの作成 -重回帰分析を用いた予測モデルの構築 -適切な初期値設定を行った非階層クラスター分析 -形態素解析などを用いた基本的文書構造解析など 	<ul style="list-style-type: none"> 単一プロジェクトにおけるデータ処理・環境構築をFull Data Scientist (棟梁レベル) に相談しつつ推進できる Assistant Data Scientist (見習いレベル) の日々の活動に適切な指示ができる 数千レコード程度のデータ処理・環境構築については自律的に実施できる -データの重要性や分析要件に則したシステム要件定義 -適切なデータフロー図、論理データモデル作成 -Hadoopでの管理対象データ選定 -SQLの構文理解と実行 -分析プログラムのロジック理解と分析結果検証など データ匿名化方法の理解と加工処理の設計ができる
④Assistant Data Scientist 見習いレベル	<ul style="list-style-type: none"> ビジネスにおける論理とデータの重要性を理解したデータプロフェッショナルとして行動規範と判断が身についている -データを取り扱う倫理と法令の理解 -引き受けたことは逃げずにやり切るコミットメントなど データドリブンな分析的アプローチの基本が身についており、仮説や既知の問題が与えられた中で、必要なデータを入手し、分析、取りまとめることができる -データや事象のタブリとモレの判断力 -分析前の目的、ゴール設定 -目的に即したデータ入手と集計、分析 -分析の意味合い出しと言語化、取りまとめなど 担当する検討領域についての基本的な課題の枠組みを理解、説明できる -ビジネスモデルと主要プレーヤー -基本的なビジネスフレームワークなど 	<ul style="list-style-type: none"> 統計数理の基礎知識を有している (代表値、分散、標準偏差、正規分布、条件付き確率、母集団、相関など) データ分析の基礎知識を有している -検定 (帰無仮説、対立仮説・) -分類 (教師あり分類、教師なし分類・) -予測 (回帰係数、標準誤差・) など 適切な指示のもとに、以下を実施できる -基本統計量や分布の確認、および前処理 (外れ値・異常値・欠損値の除去・変換や標準化など) -前処理後のデータに対し、抽出、集計、可視化 (ヒストグラム、散布図など) 	<ul style="list-style-type: none"> データやデータベースに関する基礎知識を有している -構造化/非構造化データの判別、論理モデル作成 -ER図やテーブル定義書の理解 -SDKやAPIの概要理解など 数十万件程度のデータ加工技術を有している -ソート、クレンジング、集計、フィルタリングができる -SQLで簡単なSELECT文を記述・実行できる -設計書に基づき、プログラム実装できる 適切な指示のもとに、以下を実施できる -同様のデータを統合するシステムの設計 -インポート、レコード挿入、エクスポート セキュリティの基礎知識を有している (機密性、可用性、完全性の3要素など)
DS以前の方	<ul style="list-style-type: none"> ビジネスは勘と経験だけで回すものだと思っている 課題を解決する際に、そもそも定量化する意識が無い 	<ul style="list-style-type: none"> 基本統計量の意味を正しく理解していない 指数を指数で割り算したりする 「平均年収」をそのまま割合呑みにしたりする グラフ・チャートの使い方が不適切 	<ul style="list-style-type: none"> レポートされてくる数値サマリに目は通すが、特に記憶には残らない アクセサ解析システムを使っていない ExcelやAccessは数字しか入れない

データサイエンティスト スキルチェックリスト

「データサイエンティスト スキルチェックリスト」は、データサイエンティストに必要とされるスキルをチェックリスト化したものです。チェックリスト内の各項目のスキルレベルは、★の数によって示し、見習いレベル (Assistant Data Scientist) ~ 棟梁レベル (Full Data Scientist) までを判定します。

スキルレベル	判定基準
① Senior Data Scientist (業界を代表するレベル) ★★★★★	-
② Full Data Scientist (棟梁レベル) ★★★	★★★の全項目のうち、50%を満たしている。
③ Associate Data Scientist (独り立ちレベル) ★★	★★の全項目のうち、60%を満たしている。
④ Assistant Data Scientist (見習いレベル) ★	★の全項目のうち、70%を満たしている。

- ※ 「必須スキル」に○がついている項目は、判定基準を満たしていても、この項目が達成されていないとそのレベルとは認められない項目として設定しています。
- ※ 独り立ちレベル以上のレベルは、下位のレベルを満たしていることが前提となります。

*引用・改変時の注意事項

- ・チェックリストを引用される場合「一般社団法人データサイエンティスト協会 スキルチェックリストより引用」と引用元を明示願います。その他については著作権法に従っての引用を願います。
- ・チェックリストを改変された場合「一般社団法人データサイエンティスト協会 スキルチェックリストを改変」と改変した旨、明示願います。

スキルカテゴリー一覧							
		項目数					
データサイエンスカ (項目数：180)	1	統計数理基礎	14	データエンジニアリングカ (項目数：119)	1	環境構築	19
	2	予測	16	2	データ収集	12	
	3	検定/判断	11	3	データ構造	11	
	4	グルーピング	13	4	データ蓄積	16	
	5	性質・関係性の把握	11	5	データ加工	13	
	6	サンプリング	4	6	データ共有	13	
	7	データ加工	8	7	プログラミング	20	
	8	Data visualization	36	8	ITセキュリティ	15	
	9	機械学習	19	ビジネスカ (項目数：123)	1	行動規範	11
	10	時系列分析	7		2	論理的思考	18
	11	言語処理	10		3	プロセス	20
	12	画像処理	6		4	データの理解・検証	26
	13	音声処理	5		5	データ入手	6
	14	パターン発見	3		6	意味合いの抽出、洞察	10
	16	グラフィカルモデル	3		7	解決	4
	17	統計数理応用	4		8	事業に実装する	8
	18	シミュレーション/データ同化	3		9	活動マネジメント	20
	19	最適化	7		項目数合計		422

データサイエンシカ

Total: 180
★: 54
★★: 68
★★★: 58

No	Sub No	スキルカテゴリ	スキルレベル	サブカテゴリ	データサイエンシカ: チェック項目	必須スキル
1	1	統計数理基礎	★	統計数理基礎	1+4+9+16+25+36を2を用いて表せる	○
2	2	統計数理基礎	★	統計数理基礎	loga (x) の逆関数を説明できる	○
3	3	統計数理基礎	★	統計数理基礎	順列と組合せの式をP、C、m、nを用いて表せる	○
4	4	統計数理基礎	★	統計数理基礎	条件付き確率の意味を説明できる	○
5	5	統計数理基礎	★	統計数理基礎	平均(相対平均)、中央値、最頻値の算出方法の違いを説明できる	○
6	6	統計数理基礎	★	統計数理基礎	母集団データ(3,4,5,5,7,8)の分散と標準偏差を電卓を用いて計算できる	○
7	7	統計数理基礎	★	統計数理基礎	母(集団)平均が標本平均とは異なることを説明できる	○
8	8	統計数理基礎	★	統計数理基礎	標準正規分布の分散と平均がいくつかわかる	○
9	9	統計数理基礎	★	統計数理基礎	相関関係と因果関係の違いを説明できる	○
10	10	統計数理基礎	★	統計数理基礎	名義尺度、順序尺度、間隔尺度、比例尺度の違いを説明できる	○
11	11	統計数理基礎	★	統計数理基礎	一般的な相関係数(ピアソン)の分母と分子を説明できる	○
12	12	統計数理基礎	★	統計数理基礎	代表的な確率分布の特徴を5つ以上説明できる	○
13	13	統計数理基礎	★	統計数理基礎	変数が量的、質的どちらの場合の関係の強さも算出できる	○
14	1	予測	★	予測	単回帰分析について最小二乗法、回帰係数、標準誤差の説明ができる	
15	2	予測	★★	予測	重回帰分析において偏回帰係数と標準偏回帰係数、重回帰係数について説明できる	
16	3	予測	★★	予測	重回帰や判別を実行する際に変数選択手法の特徴を理解し、適用できる	
17	4	予測	★★	予測	ニューラルネットワークの基本的な考え方を理解し、出力される「ダイアグラム」の入力層、隠れ層、出力層の概要を説明できる	
18	5	予測	★★	予測	重回帰分析において多重共線性の対応ができ、適切に変数を評価・除去して予測モデルが構築できる	○
19	6	予測	★★	予測	決定木分析においてCHAID、C5.0などのデータ分割のアルゴリズムの特徴を理解し、適切な方式を選定できる	
20	7	予測	★★	予測	線形回帰分析が量的な予測するのに対して、ロジスティック回帰分析は何を予測する手法か(発生確率予測など)を説明でき、実際に使用できる	
21	8	予測	★★★	予測	予測アルゴリズムに応じ、目的変数と説明変数に対する必要な変数加工処理が設計、実施できる	
22	9	予測	★★★	予測	予測モデル構築において頑健性を維持するための具体的な方法を設計、実施できる	
23	10	予測	★★★	予測	過学習(正則化)の意味を説明ができ、回避する方法の設計・実施ができる	
24	11	予測	★★★	予測	尤度と最尤推定についての説明できる(尤度関数、ネイマンの分解定理、順序統計量)	
25	12	予測	★★★	予測	(モデルに交互作用が含まれていないとき)回帰パラメータとオッズ比の関係について説明できる	
26	13	予測	★★★	予測	ロジスティック回帰分析を行う際に、最小2乗法ではなく最尤法を使う際の利点(回帰誤差が近似的に正規分布しなくても適用できるなど)を説明し、適用することができる	
27	14	予測	★★★	予測	目的(予測・真のモデル推定など)に応じて、適切な損失関数とモデル選択基準(AIC(赤池情報量規準)、BIC(ベイズ情報量規準)、MDL(最小記述長)など)を選択し、モデル評価ができる	
28	15	予測	★★★	予測	データと分析要件から、モデル精度のモニタリング設計・実施と劣化が見込まれるモデルに対するリモデルの設計ができる	
29	16	予測	★★★	予測	分析要件に応じ、量的予測のためのモデリング手法(重回帰、決定木、ニューラルネットワークなど)の選択とパラメータ設定、結果の評価、チューニングを適切に設計・実施・指示できる	
30	1	検定/判断	★	検定/判断	点推定・区間推定の意味を理解している	

No	Sub No	スキルカテゴリ	スキルレベル	サブカテゴリ	データサイエンシカ: チェック項目	必須スキル
31	2	検定/判断	★	検定/判断	帰無仮説・対立仮説の意味を理解している	○
32	3	検定/判断	★	検定/判断	第1種の過誤、第2種の過誤、p値、有意水準の意味を理解している	
33	4	検定/判断	★	検定/判断	片側検定、両側検定の意味を理解している	
34	5	検定/判断	★	検定/判断	対応のあるデータ、対応のないデータの意味を理解している	
35	6	検定/判断	★	検定/判断	平均値、分散、平均値の差の検定手法を知っている	
36	7	検定/判断	★★	検定/判断	t検定を理解して、パラメトリックな2群の検定を活用することができる	○
37	8	検定/判断	★★	検定/判断	様々な分散分析の考え方(一元配置、多重比較、二元配置)を理解して、パラメトリックな多群の検定を活用することができる	
38	9	検定/判断	★★	検定/判断	ウィルコクソン検定(マンホイットニーのU検定)を理解して、ノンパラメトリックな2群の検定を活用することができる	
39	10	検定/判断	★★	検定/判断	クラスカル・ウォリス検定を理解して、ノンパラメトリックな多群の検定を活用することができる	
40	11	検定/判断	★★	検定/判断	カイ二乗検定、フィッシャーの直接確率検定を理解して、分割表における群間の関連性の検定を活用することができる	
41	1	グルーピング	★	グルーピング	教師あり分類と教師なし分類の違いを説明できる	○
42	2	グルーピング	★	グルーピング	判別分析とクラスター分析の概要や使い方を説明できる	
43	3	グルーピング	★	グルーピング	階層クラスター分析と非階層クラスター分析の違いを理解している	
44	4	グルーピング	★	グルーピング	クラスター分析のデンドログラムの見方が分かり、適切に解釈できる	
45	5	グルーピング	★★	グルーピング	非階層クラスター分析において、分析目的に合致したクラスター数を決定することができる	
46	6	グルーピング	★★	グルーピング	階層クラスター分析における代表的なクラスター間距離(群平均法、Ward法、最長一致法など)の概念を理解し、分析内容に応じて最適な手法を選択して、分析ができる	
47	7	グルーピング	★★	グルーピング	各種距離関数(ユークリッド距離、マンハッタン距離、cos類似度など)を理解し、分析内容に応じて、適切なものを選択し分析ができる	
48	8	グルーピング	★★	グルーピング	分析内容に応じて、線形か非線形を使い分けた判別分析ができる	
49	9	グルーピング	★★	グルーピング	k-means法は局所最適解であるため初期値問題があることを理解し、適切な初期値を設定して分析を行える	○
50	10	グルーピング	★★★	グルーピング	距離の公理を知っており、距離の公理を満たさない場合([1-cos類似度]など)のクラスター分析を適切に行える	
51	11	グルーピング	★★★	グルーピング	k-meansの派生手法(x-means、k-means++、ファジィk-meansなど)を理解し、分析内容に応じた適切な手法を選択して、分析ができる	
52	12	グルーピング	★★★	グルーピング	k-meansとカーネルk-means(非線形クラスターリング)、スペクトラルクラスターリングの違いを理解し、分析内容に応じて、適切な手法を選択して、分析ができる	
53	13	グルーピング	★★★	グルーピング	自己組織化マップ(SOM)、Affinity Propagation、混合分布モデル、ディリクレ過程混合モデルを理解し、分析内容に応じて適切な手法を選択して、分析ができる	
54	1	性質・関係性の把握	★	性質・関係性の把握	適切なデータ区間設定でヒストグラムを作成することができる	○
55	2	性質・関係性の把握	★	性質・関係性の把握	何でクロスするかを判断して、表側を設定しクロス集計表を作成することができる	○
56	3	性質・関係性の把握	★	性質・関係性の把握	量的変量の散布図を描き、2変数の関係性を把握することができる	○
57	4	性質・関係性の把握	★★	性質・関係性の把握	多重(質問間)クロス表などを駆使して、データから適切なインサイトを導くことができる	○
58	5	性質・関係性の把握	★★	性質・関係性の把握	主成分分析と因子分析の違いや使い分けを説明できる	○
59	6	性質・関係性の把握	★★	性質・関係性の把握	正準相関分析を説明、活用できる	
60	7	性質・関係性の把握	★★	性質・関係性の把握	正規分布を前提とした多変量解析において、分析の目的と用いるデータの種類から適切な手法を選択できる	○
61	8	性質・関係性の把握	★★	性質・関係性の把握	因子分析における、因子負荷量や因子軸の回転について説明できる	
62	9	性質・関係性の把握	★★	性質・関係性の把握	コレスポンデンス(対応)分析と数量化3類との関係を説明できる	

No	Sub No	スキルカテゴリ	スキルレベル	サブカテゴリ	データサイエンティスト: チェック項目	必須スキル
63	10	性質・関係性の把握	★★	性質・関係性の把握	コンジョイント分析を用いて効用値と寄与率のグラフを描くことができる	
64	11	性質・関係性の把握	★★	性質・関係性の把握	適切な類似度を設定した上で、多次元尺度構成法を用いてポジショニングマップを描くことができる	
65	1	サンプリング	★	サンプリング	標本誤差とは何かを説明できる	
66	2	サンプリング	★	サンプリング	実験計画法の概要を説明できる	
67	3	サンプリング	★★	サンプリング	分析に最低限必要な標本サイズを設定など、調査目的に合った標本抽出計画を立案できる	○
68	4	サンプリング	★★	サンプリング	属性数と水準数が決まれば適切な直交表を選択し実験計画ができる	
69	1	データ加工	★	データ加工	名義尺度の変数をダミー変数に変換できる	
70	2	データ加工	★	データ加工	標準化とは何かを知っていて、適切に標準化が行える	
71	3	データ加工	★	データ加工	外れ値・異常値・欠損値とは何かそれぞれ知っていて、指示のもと適切に検出と除去・変換などの対応ができる	○
72	4	データ加工	★	データ加工	指示のもと、加工済データに分析上の不具合がないか検証できる	
73	5	データ加工	★★	データ加工	各変数の分布・欠損率などを踏まえて、外れ値・異常値・欠損値の対応を決定できる	○
74	6	データ加工	★★	データ加工	分析要件や各変数の分布などを踏まえて、必要に応じて量的変数のカテゴリを設計・実行できる	
75	7	データ加工	★★	データ加工	加工データに不具合がないか自分でテストを設計し、検証できる	
76	8	データ加工	★★★	データ加工	加工データの統計的な前提によって不具合の早期発見ができるとともに、統計的観点で次ステップの解析に耐えうるデータであるか評価できる	
77	1	Data visualization	★	方向性定義	単に現場の作業支援する場合から、ビッグデータ中の要素間の関連性をダイナミックに表示する場合まで、可視化の目的の広がりについて理解している	
78	2	Data visualization	★★	方向性定義	特異点を明確にする、データ解析部門以外の方にデータの意味を正しく伝える、現場の作業を支援するといった可視化の役割・方向性を判別できる	
79	3	Data visualization	★★★	方向性定義	データ量が膨大で構造が捉えにくい場合や、アウトプットが想像しにくい場合であっても、可視化の役割・方向性を判断できる（ビッグデータ中の要素間の関連性をダイナミックに表現する、細部に入りきらずに問に対して答えを出すなど）	
80	4	Data visualization	★	軸だし	散布図などの軸だしにおいて、縦軸・横軸の候補を適切に洗い出せる	
81	5	Data visualization	★	軸だし	積み上げ棒グラフでの属性の選択など、適切な層化（比較軸）の候補を出せる	
82	6	Data visualization	★★	軸だし	抽出したい意味にふさわしい軸・層化の粒度、順番を考慮して軸のきざみや層化方法を選択できる	
83	7	Data visualization	★★	軸だし	膨大な属性を持つテーブルから目的に有用な属性を選択できる	
84	8	Data visualization	★★★	軸だし	非構造データから分析の軸になりうる候補を抽出し、付加すべき属性候補を適切に出せる	
85	9	Data visualization	★	データ加工	サンプリングやアンサンブル平均によってデータ量を減らして可視化できる	
86	10	Data visualization	★★	データ加工	データの持つ分散量の観点で、高次のデータを主成分分析（PCA）などにより1～3次元のデータに変換できる	
87	11	Data visualization	★★	データ加工	データの持つ特徴を統計量を使って効果的に可視化できる	
88	12	Data visualization	★★★	データ加工	非線形（高次の曲線、渦状の分布など）のデータであっても、高次のデータの次元を、次元圧縮（1～3次元のデータに変換）して、特徴（データの総分散および各データの位置関係）を損なわずに可視化できる	
89	13	Data visualization	★★★	データ加工	ネットワーク構造、グラフ構造などの表現において、ノードとエッジが増えすぎて特徴抽出が困難であっても、データの絞り込みや抽象度を上げることで適切に可視化できる	
90	14	Data visualization	★★★	データ加工	データ量が膨大（ペタバイト以上）なために、処理しきれず描画できない規模のデータに対しても、適度なデータや情報の抽出（間引き）、クラスタリングなどにより可視化しうる状態にデータを加工できる	
91	15	Data visualization	★★★	データ加工	大規模なデータへのリアルタイムな可視化が求められる場合であっても、特異点の抽出や次元圧縮を通じてデータを圧縮し、リアルタイム表示できる	
92	16	Data visualization	★★★	データ加工	大規模なデータへのリアルタイムな可視化が求められる場合であっても、データの分割転送、復元を通じて可視化できる	
93	17	Data visualization	★	表現・実装技法	適切な情報濃度（データインク比、データ濃度の概念）を理解し判断できる	

No	Sub No	スキルカテゴリ	スキルレベル	サブカテゴリ	データサイエンティスト: チェック項目	必須スキル
94	18	Data visualization	★	表現・実装技法	不必要な誇張をしないための軸表現の基礎を理解できている（コラムチャートのY軸の基準点は「0」からを原則とし軸を切らないなど）	
95	19	Data visualization	★	表現・実装技法	強調表現がもたらす効果を理解している（計量データに対しては位置やサイズ表現が色表現よりも効果的など）	
96	20	Data visualization	★	表現・実装技法	1～3次元の比較において目的（比較、構成、分布、変化など）に応じた図表化ができる	
97	21	Data visualization	★	表現・実装技法	端的に図表の変化をアニメーションで可視化できる（人口動態のヒストグラムが経年変化する様子を表現するなど）	
98	22	Data visualization	★	表現・実装技法	データ解析部門以外の方に、データの意味を伝えるサインとしての可視化ができる	
99	23	Data visualization	★★	表現・実装技法	1～3次元の図表を拡張した多変量の比較（平行座標、散布図行列、テーブルレンズ、ヒートマップなど）を適切に可視化できる	
100	24	Data visualization	★★	表現・実装技法	ネットワーク構造、グラフ構造、階層構造などの統計的な関係性の可視化ができる	
101	25	Data visualization	★★	表現・実装技法	GPSデータなどを平面地図上に重ね合わせた可視化ができる	
102	26	Data visualization	★★	表現・実装技法	挙動・軌跡の可視化ができる（店舗内でのユーザの回遊やEye trackingなど）	
103	27	Data visualization	★★	表現・実装技法	適切な情報（意味）を押しさえたうえで、デザイン性を高めるための要件提示ができる	
104	28	Data visualization	★★★	表現・実装技法	人体、標高を持つ地図、球面などの上にデータを重ね合わせた可視化ができる	
105	29	Data visualization	★★★	表現・実装技法	地図上で同時に動く数百以上のポイントにおける時間変化を動的に表現できる（多地点での風の動き、飛行物の軌跡など）	
106	30	Data visualization	★	意味抽出	データの性質を理解するために、データを可視化し眺めて考えることの重要性を理解している	
107	31	Data visualization	★	意味抽出	外れ値を見出すための適切な表現手法を選択できる	
108	32	Data visualization	★	意味抽出	データの可視化における基本的な視点を理解している（特異点、相違性、傾向性、関連性を見出すなど）	
109	33	Data visualization	★★	意味抽出	統計値（代表値の指標、バラツキの指標、有意性の指標、関係性）を正しく読み、回帰式や移動平均線に意味付けできる	
110	34	Data visualization	★★★	意味抽出	分類系の分析において、分布傾向から原因を追究、活用（分類に応じたDM発送による反応率の向上など）、ドリルダウンを計画し主導できる	
111	35	Data visualization	★★★	意味抽出	予測系の分析において、関連性、特異点、変曲点から原因を追究、活用（予測結果に基づく発注管理など）を計画し主導できる	
112	36	Data visualization	★★★	意味抽出	関連系の分析において関連が高い/低い原因、活用（リコメンドなど）、ドリルダウンを計画し主導できる	
113	1	機械学習	★	機械学習	機械学習にあたる解析手法（Random Forestなど）の名称を3つ以上知っており、手法の概要を理解している	
114	2	機械学習	★	機械学習	指示を受けて機械学習のモデルを使用したことがあり、どのような問題を解決することができるか理解している	
115	3	機械学習	★★	機械学習	「教師あり学習」「教師なし学習」「強化学習」の違いを理解しており、データの内容や学習手法に応じて適切な学習データとテストデータを作成できる	○
116	4	機械学習	★★	機械学習	決定木とRandomForestの違いを説明できる	
117	5	機械学習	★★	機械学習	ROCカーブを用いてモデルの精度を評価できる	
118	6	機械学習	★★	機械学習	混同行列（正誤分布のクロス表）を用いてモデルの精度を評価できる	
119	7	機械学習	★★	機械学習	行列分解（非負値行列因子分解、特異値分解）をツールを使って実行でき、その結果を正しく解釈できる	
120	8	機械学習	★★	機械学習	サポートベクターマシンによる分析を、統計解析ツールを使って実行でき、その結果を解釈できる	
121	9	機械学習	★★	機械学習	アンサンブル学習（Random Forest、勾配ブースティングマシン、バギング）による分析を、統計解析ツールを使って実行でき、その結果を正しく解釈できる	
122	10	機械学習	★★	機械学習	ニューラルネットワークによる分析を、統計解析ツールを使って実行でき、その結果を正しく解釈できる	
123	11	機械学習	★★★	機械学習	行列分解（非負値行列因子分解、特異値分解）を、目的に応じてパラメータを最適化して分析できる	
124	12	機械学習	★★★	機械学習	課題やデータ型に応じて、サポートベクターマシンの適切なモデルを選定し、目的に応じてアルゴリズムの調整や設計ができる	
125	13	機械学習	★★★	機械学習	課題やデータ型に応じて、アンサンブル学習（Random Forest、勾配ブースティングマシン、バギング）の適切なモデルを選定し、目的に応じてアルゴリズムの調整や設計ができる	

No	Sub No	スキルカテゴリ	スキルレベル	サブカテゴリ	データサイエンティスト: チェック項目	必須スキル
126	14	機械学習	★★★	機械学習	Deep Learning (深層学習) の適切なモデルを選定し、目的に応じてアルゴリズムの調整や設計ができる	
127	15	機械学習	★★★	機械学習	モデル選択のための汎化能力評価 (汎化誤差の推定) を、ホールドアウト法、交差確認法、一つ抜き法、フットストラップ法を用いて行える	
128	16	機械学習	★★★	機械学習	正則化 (誤差関数に正則化項を追加することで、モデルの複雑度・自由度を抑制する) を利用して、過学習を防ぐことができる	
129	17	機械学習	★★★	機械学習	高次元データの取り扱いについて、次元の呪いを考慮し適切に次元削減できる	
130	18	機械学習	★★★	機械学習	機械学習等の最新の論文を理解し、必要とあれば自分で実装し追試できる	
131	19	機械学習	★★★	機械学習	バイアスとバリエーションの関係を理解し、モデル選定を適切に行える	
132	1	時系列分析	★	時系列分析	時系列データについて説明ができる (時系列グラフ、周期性、移動平均など)	
133	2	時系列分析	★★	時系列分析	時系列データの時点差での相関関係を、系列相関やコレログラムを利用して評価ができる	
134	3	時系列分析	★★	時系列分析	時系列データに対し、ツールを使用して、分析結果の比較を行い、適切なモデルを選択できる (自己回帰モデル、統合モデル、移動平均モデル、ARIMA、SARIMAなど)	
135	4	時系列分析	★★	時系列分析	時系列分析の特徴 (通常は、トレンドおよび季節成分を省いたランダム部分を、定常時系列として分析する手法であること) を理解している	
136	5	時系列分析	★★★	時系列分析	状態空間モデルにおいて、カルマンフィルタを用いて、観測値から欠測値の補間をし、予測モデルを構築できる	
137	6	時系列分析	★★★	時系列分析	非線形・非ガウス型状態空間モデルにおいて、モンテカルロ・フィルタを用いて、複雑な時系列システムの予測モデルを構築できる	
138	7	時系列分析	★★★	時系列分析	ピリオドグラムにおいて、FFT (Fast Fourier Transform) など計算量を抑制する方法により、ピリオドグラムの計算ができる	
139	1	言語処理	★	言語処理	形態素解析や係り受け解析の概念を理解している	
140	2	言語処理	★★	言語処理	形態素解析や係り受け解析のツールを適切に使い、基本的な文書構造解析を行うことができる	○
141	3	言語処理	★★	言語処理	TF-IDFやcos類似度などの基本的なアルゴリズムを使い、単語ベクトルの作成や文書群の類似度計算を行うことができる	
142	4	言語処理	★★	言語処理	Precision、Recall、F値といった評価尺度を理解し、実際の精度評価を行うことができる	
143	5	言語処理	★★★	言語処理	形態素解析・構文解析・固有表現抽出のアルゴリズムを理解し、使いこなせる	
144	6	言語処理	★★★	言語処理	N-gram言語モデルの構築方法と代表的なスムージングアルゴリズムを理解し、使いこなせる	
145	7	言語処理	★★★	言語処理	索引型の全文検索の仕組み (転置インデックス、スコアリング、関連性フィードバック) を理解し、使いこなせる	
146	8	言語処理	★★★	言語処理	Trie、Suffix Arrayなどの代表的な高速文字列検索アルゴリズムを理解し、使いこなせる	
147	9	言語処理	★★★	言語処理	潜在的意味解析 (LSA) の仕組みを理解し、使いこなせる	
148	10	言語処理	★★★	言語処理	データの特性に合わせ、適切な言語処理アルゴリズムを選択し、誤り分析、辞書作成などを行い、成果を最大化することができる	
149	1	画像処理	★	画像処理	画像映像のデジタル表現の仕組みと代表的な画像映像フォーマットの概念を理解している	
150	2	画像処理	★★	画像処理	主要な画像映像フォーマットの種類および特性を理解し、適切に使い分けたり変換することができる	
151	3	画像処理	★★★	画像処理	画像に対して、目的に応じた適切な色変換や補正、フィルタ処理などを行うことができる	
152	4	画像処理	★★★	画像処理	問題に合わせてパターン検出や画像特徴抽出などの既存手法から適切なものを選択し、画像の処理や解析に用いることができる	
153	5	画像処理	★★★	画像処理	オブティカルフローやトラッキング手法などにより映像中の人や物体の移動を分析することができる	
154	6	画像処理	★★★	画像処理	目的に合わせてショット検出、物体検出、行動認識などの手法を用い、映像中の特定のイベントを効率よく検出することができる	
155	1	音声処理	★	音声処理	代表的な音声フォーマットの概要を理解している	
156	2	音声処理	★★	音声処理	代表的な音声フォーマットを理解し、波形データの抽出をすることができる	
157	3	音声処理	★★	音声処理	代表的な音声処理分野の概念を理解し、説明することができる	
158	4	音声処理	★★★	音声処理	代表的な音声信号分析手法を理解し、使いこなすことができる	

No	Sub No	スキルカテゴリ	スキルレベル	サブカテゴリ	データサイエンティスト: チェック項目	必須スキル
159	5	音声処理	★★★	音声処理	目的に合わせて、音声認識のパラメータ調整や手法変更、言語モデル・音響モデルの差し替えができ、音声認識の精度評価をすることができる	
160	1	パターン発見	★★	パターン発見	条件Xと事象Yの関係をリフト値を用いて評価できる	
161	2	パターン発見	★★	パターン発見	アプリアリアルゴリズムのアソシエーション分析の入カパラメータ値を調整して出力件数を調整できる	
162	3	パターン発見	★★★	パターン発見	レコメンデーション業務の要件から適切なフィルタリングロジック (ユーザベース、アイテムベースなど) の選定を行うことができる	
163	1	グラフィカルモデル	★★	グラフィカルモデル	パス解析において、変数間の因果関係をパス図を用いて説明できる	
164	2	グラフィカルモデル	★★	グラフィカルモデル	ベイジアンネットワーク分析結果のネットワーク図から目的事象の同時確率を算出できる	
165	3	グラフィカルモデル	★★	グラフィカルモデル	ツールを用いて共分散構造分析 (構造方程式モデリング: SEM) を行い、因果関係を説明できる	
166	1	統計数理応用	★	統計数理基礎	ベイズの定理を理解している	
167	2	統計数理応用	★	統計数理応用	ポロノイ図の概念と活用方法を説明できる	
168	3	統計数理応用	★★★	統計数理応用	空間的自己相関の手法を用いて空間的な類似性を数値化できる	
169	4	統計数理応用	★★★	統計数理応用	当該分野に則したベイズ統計に基づくアルゴリズムを理解し、モデルを構築できる	
170	5	統計数理応用	★★★	統計数理応用	ベイズ統計と頻度論による従来の統計との違いを、尤度、事前確率、事後確率などの用語を用いて説明できる	
171	1	シミュレーション/データ同化	★★	シミュレーション/データ同化	離散型・連続型シミュレーションについて、説明できる (モンテカルロ、ヒストリカル、Agentベースなど)	
172	2	シミュレーション/データ同化	★★★	シミュレーション/データ同化	データ同化 (データを用いてシミュレーション内の不確実性を減少させる計算技法) の概念を理解している	
173	3	シミュレーション/データ同化	★★★	シミュレーション/データ同化	シミュレーションにおける問題を理解し、対処を考慮することができる (初期条件・境界条件・パラメータの不確実性、データ分布の不均一性、実験計画の最適性など)	
174	1	最適化	★★	最適化	線形計画法について、説明することができる	
175	2	最適化	★★	最適化	連続最適化 (制約なし) において、アルゴリズムを説明することができる (ニュートン法、最急降下法など)	
176	3	最適化	★★	最適化	制約がある最適化問題において、代表的な組み合わせ解法の概念を説明することができる (一般的な手法 (Brute-force search、greedy algorithmなど)、ヒューリスティック (局所探索、山登りなど)、メタ解法 (GA、TS、SAなど))	
177	4	最適化	★★★	最適化	代表的な最適化問題に関して、モデリングを行い、ソルバーを使い、最適化できる (ナップザック問題、ネットワークフロー問題、巡回路問題など)	
178	5	最適化	★★★	最適化	ビジネス課題にあわせて、変数、目的関数、制約を定式化し、線形・非線形を問わず、最適化モデリングができる	
179	6	最適化	★★	最適化	凸関数とは何かを説明でき、目的関数が凸関数で実行可能な領域が凸集合である場合の数値計画法を凸計画法ということを理解している	
180	7	最適化	★★	最適化	複数のA/Bテストの統計的結果を踏まえ、デザイン等の最適化を行う手法を回すことができる	

データエンジニアリングカ

Total: 119
 ★: 39
 ★★: 43
 ★★★: 37

No	Sub No	スキルカテゴリ	スキルレベル	サブカテゴリ	データエンジニアリングカ:チェック項目	必須スキル
1	1	環境構築	★	システム運用	サーバー1~10台規模のシステム構築、システム運用を指示書があれば実行できる	
2	2	環境構築	★	システム運用	数十万レコードを持つデータベースのバックアップ・アーカイブ作成など定常運用ができる	
3	3	環境構築	★★	システム運用	数千レコードのデータを保持するシステムにおいてデータの重要性や分析要件に則したシステム構築、初期データ投入方法、システム運用の要件定義が行える	○
4	4	環境構築	★★	システム運用	顧客管理など分析システムの運用（異常検知、フェイルオーバー、バックアップ、リカバリ処理、開始・停止処理）の手順書作成や要件定義が行える	
5	5	環境構築	★★	システム運用	数千レコードのデータを保持するシステムのキャパシティ要件（データ容量）と必要処理性能（スループット）を定義できる	
6	6	環境構築	★★★	システム運用	数十億レコードのデータに対してリバランシングなども含めてシステム拡張方法や最適化の要件を整理できる	
7	7	環境構築	★★★	システム運用	扱うデータのデータ規模や機密性、分析要件を理解した上で、オンプレミスで構築するか、クラウド上で構築するかの要件を整理できる	
8	8	環境構築	★	システム企画	データベースから何らかのデータ抽出方法を活用し、小規模なExcelのデータセット作成ができる	○
9	9	環境構築	★	システム企画	オープンデータ活用目的でExcelを使った分析システムの要件定義が行える	
10	10	環境構築	★★	システム企画	HTTPを活用したオープンAPIと分析システムのサーバー環境及びデータベースの連携設計ができる	
11	11	環境構築	★★	システム企画	社内分析者向けのRDBMS、NoSQL、ETL、Visualizationなど単一コンポーネントのユーザー利用機能設計が行える	○
12	12	環境構築	★★	システム企画	ソフトウェア開発プロジェクトの管理方法、設計、テスト方法を理解した上で、データ管理・分析システムを要求定義することができる	
13	13	環境構築	★★★	システム企画	数十億レコードのデータに対してHiveとImpala、Sparkを組み合わせたなど処理可能なデータ量の違うシステムを連携させ、処理性能とのトレードオフを調整できる	
14	14	環境構築	★★★	システム企画	数十億規模の内部データ（CRM、webログ、ユーザー購買データ）、外部データ（購入しているデータ、オープンデータ）を理解し、複数のデータソースを統合する要件を整理できる	
15	15	環境構築	★★★	システム企画	サービスのそれぞれの機能がどのデータに関連があるか把握し、分析機能追加やシステム変更の要件を整理できる	○
16	16	環境構築	★★★	システム企画	KVS、カラム指向、ドキュメント指向などデータ構造の異なる複数のシステムからデータ取得と分析環境への連携が設計できる	
17	17	環境構築	★★★	システム企画	数十億規模のデータを持つ業務要件やリソース負荷に応じて、データフローや管理機構の統合、またバッチ実行スケジュールの最適化について要件定義が行える	
18	18	環境構築	★★★	アーキテクチャ設計	全体システム化計画及び個別システム化構想・計画を具体化するために、全体最適の観点を持ちながら、対象とするデータ分析システムの開発に必要な要件を整理することができる	
19	19	環境構築	★★★	アーキテクチャ設計	データ活用戦略を正しく理解し、各種業務プロセスについての専門知識とシステムに関する知識を有し、双方を活用して、適切な要求定義が行える	
20	1	データ収集	★	クライアント技術	対象プラットフォームが提供する機能（SDKやAPIなど）の概要を説明できる	○
21	2	データ収集	★★	クライアント技術	対象プラットフォーム（iOS、Android、HEMSなど）におけるデータ取得の仕様（精度など）を理解しており、システム要件を満たせるか判断できる	
22	3	データ収集	★★	クライアント技術	目的に適したログ取得項目を、対象プラットフォーム（iOS、Android、HEMSなど）で取得可能なデータを用いて設計できる	
23	4	データ収集	★★	クライアント技術	対象プラットフォームにおけるバッテリー消費や通信速度などを含めたシステム要件を作成できる	
24	5	データ収集	★★★	クライアント技術	取得するデータが増えることを想定し、拡張性を考慮したクライアントアプリケーション（ロガーなど）を設計できる	
25	6	データ収集	★★★	クライアント技術	プラットフォームの違いを吸収し、統一的にデータを取得するプログラムを実装できる	
26	7	データ収集	★	通信技術	対象プラットフォームに用意された機能（HTTP、FTP、SSHなど）を用い、データを収集先に格納するための機能を実装できる	
27	8	データ収集	★★	通信技術	ネットワークプロトコルや暗号化などの通信技術を用い、通信のポトルネットワークを考慮した上で、必要な通信機能を実装できる	
28	9	データ収集	★★★	通信技術	データ通信において、機能・性能問題に対し根本原因を特定できるだけでなく、必要に応じて新規技術の適用を検討できる	
29	10	データ収集	★	データ統合	同種のデータを統合するシステムを設計できる	○
30	11	データ収集	★★★	データ統合	異種フォーマットが混在するデータを統合するシステムを設計できる	

No	Sub No	スキルカテゴリ	スキルレベル	サブカテゴリ	データエンジニアリングカ:チェック項目	必須スキル
31	12	データ収集	★★★	データ統合	システム分析・業務分析を元に、必要なデータフロー管理やジョブ管理ツールを選定・評価できる	
32	1	データ構造	★	基礎知識	扱うデータが、構造化データ（顧客データ、商品データ、在庫データなど）なのか非構造化データ（雑多なテキスト、音声、画像、動画など）なのか判断できる	○
33	2	データ構造	★	基礎知識	ER図を読んでテーブル間のリレーションシップを理解できる	○
34	3	データ構造	★	要件定義	業務で使用するシステムのデータのライフサイクル（いつ、どんなデータが発生し、いつまで保持されているのかなど）を把握して、論理モデルを作成できる	○
35	4	データ構造	★★	要件定義	データ保持ルール（データアクセス、性能、保持期間、セキュリティなど）に基づき、データベース・DWHの運用ルールを定義できる	
36	5	データ構造	★	テーブル定義	正規化手法（第一正規化~第三正規化）を用いてテーブルを正規化できる	
37	6	データ構造	★★	テーブル定義	ビジネスプロセスを理解・整理して、データフロー図、論理データモデル、ER図、テーブル定義書を作成できる	○
38	7	データ構造	★★	テーブル定義	業務特性や基幹システムの特徴を元に、検索で頻繁に使用するデータのキー（顧客IDなど）を想定し、インデックスを作成・設定できる	
39	8	データ構造	★★	テーブル定義	データ集計を高速化またはSQLを単純化するため、スタースキーマ、スノーflakeスキーマなどを用いたデータモデルを設計できる	
40	9	データ構造	★★	テーブル設計	稼働中の複数のシステム間で発生するデータ項目の差異を、変換テーブルを活用して、埋めることができる	
41	10	データ構造	★★	テーブル設計	DWHに入れる元データ（基幹DBのデータなど）のキーに変更があった場合に、サロゲートキーやナビゲーションブリッジテーブルを用いて対応できる	
42	11	データ構造	★★★	テーブル設計	非正規化テーブルや一時テーブルなどを作成し、アプリケーションの処理速度を高速化できる	
43	1	データ蓄積	★	DWH	DWHアプライアンス（Oracle Exadata、IBM Pure Data/Netezza、Teradataなど）に接続し、複数テーブルを結合したデータを抽出できる	
44	2	データ蓄積	★★	DWH	DWHアプライアンス（Oracle Exadata、IBM Pure Data/Netezza、Teradataなど）の機能と特徴を理解し、適切な管理対象データを選定できる	
45	3	データ蓄積	★	分散技術	Hadoopに対応可能なインターフェース（MapReduce、Pig、Hiveなど）を用いてHadoop上の大量データから条件に合致するデータを抽出できる	
46	4	データ蓄積	★★	分散技術	Hadoopの得意な点、苦手な点を理解し、Hadoopにて管理すべきデータを選定できる	○
47	5	データ蓄積	★★	分散技術	KVSの特性（集計・ソートが苦手、データの一貫性保証など）を理解し、KVSがデータストア要件を満たすかを判断できる	
48	6	データ蓄積	★★★	分散技術	分散クラスタ構成が構築可能なRDBMS製品（Oracle RAC、IBM Pure Scaleなど）を用いてスケールアウト可能なオンプレミス構成を設計できる	
49	7	データ蓄積	★★★	分散技術	Hadoopの分散アーキテクチャを理解し、大容量データ処理のパフォーマンスチューニングができる	
50	8	データ蓄積	★	分散技術	NoSQLデータストア（HBase、Cassandra、Amazon DynamoDB、CouchDB、Azure DocumentDBなど）にAPIを介してアクセスし、新規データを登録できる	
51	9	データ蓄積	★	クラウド	クラウド上のストレージサービス（Amazon S3、Google Cloud Storage、Softlayer Object Storageなど）に接続しデータを格納できる	
52	10	データ蓄積	★★	クラウド	クラウド上のDWHサービス（Amazon Redshift、Google BigQueryなど）にデータをロードし公開できる	
53	11	データ蓄積	★★★	クラウド	クラウド上のデータストアサービスが機能面・非機能面で対象業務に合致するかの評価を行い、採用可否を判断できる	
54	12	データ蓄積	★★★	キャッシュ技術	基盤設計において、どこのシステム要素にmemcachedなどのキャッシュ機能を採用すると処理が高速化されるか判断できる	
55	13	データ蓄積	★★★	リアルタイムデータ分析	リアルタイムに入力されるストリームデータから指定条件のイベントを即時に抽出する複合イベント処理（CEP）を実現するサーバー環境・構成を設計できる	
56	14	データ蓄積	★★★	リアルタイムデータ分析	リアルタイムに連続して入力されるストリームデータの加工・集計処理を行うにあたり、Storm等のリアルタイム分散フレームワーク適用の有効性を判断できる	
57	15	データ蓄積	★★★	グリッドコンピューティング	分散処理のフレームワーク（Spark、Tezなど）を用いてアプリケーションの計算処理を複数サーバーに分散させる並列処理システムを設計できる	
58	16	データ蓄積	★★★	新規技術	データストアの技術動向に注目し、リレーショナルDBだけでなく、グラフDB・時系列DBなどの新規技術の検証・評価ができる	
59	1	データ加工	★	フィルタリング処理	数十万レコードのデータに対して、条件を指定してフィルタリングできる（特定値に合致する・もしくは合致しないデータの抽出、特定範囲のデータの抽出、部分文字列の抽出など）	○
60	2	データ加工	★★	フィルタリング処理	正規表現を活用して条件に合致するデータを抽出できる（メールアドレスの書式を満たしているか判定をするなど）	
61	3	データ加工	★	ソート処理	数十万レコードのデータに対して、レコード間で特定カラムでのソートができる。また、数千レコードのデータに対して、カラム間でソートできる	○

No	Sub No	スキルカテゴリ	スキルレベル	サブカテゴリ	データエンジニアリングカ:チェック項目	必須スキル
62	4	データ加工	★	結合処理	数十万レコードのデータに対して、単一条件による内部結合、外部結合、自己結合ができる。また、UNION処理ができる	
63	5	データ加工	★	クレンジング処理	数十万レコードのデータに対して、NULL値や想定外・範囲外のデータを持つレコードを取り除く、または既定値に変換できる	○
64	6	データ加工	★★	クレンジング処理	フラットファイルやバイナリファイルに対するデータロードの前処理（クレンジング操作、禁則処理やバイナリ処理）ができる	
65	7	データ加工	★★	クレンジング処理	線形補間など、複数のレコードを考慮したクレンジング処理ができる	
66	8	データ加工	★	マッピング処理	数十万レコードのデータに対して、規定されたリストと照合して変換する、都道府県名からジオコードに変換するなど、ある値を規定の別の値で表現できる	
67	9	データ加工	★★★	マッピング処理	データ定義や実際の観測データの状況をもとに、名寄せ処理を設計・実装できる	
68	10	データ加工	★	サンプリング処理	数十万レコードのデータに対して、ランダムまたは一定間隔にデータを抽出できる	
69	11	データ加工	★	集計処理	数十万レコードのデータを集計して、合計や最大値、最小値、レコード数を算出できる	○
70	12	データ加工	★	変換・演算処理	数十万レコードのデータに対する四則演算ができる。また、数値データを日時データに変換するなど別のデータ型に変換できる	
71	13	データ加工	★★	変換・演算処理	数千万レコードのデータに対して、カラムナー型のデータに変換できる	
72	1	データ共有	★	データ出力	加工・分析処理結果をCSV、XML、Excelなどの指定フォーマット形式に変換してエクスポートできる	○
73	2	データ共有	★	データ出力	加工・分析処理結果を、接続先DBのテーブル仕様に合わせてレコード挿入できる	○
74	3	データ共有	★	データ展開	データ取得用のWeb API (REST) やWebサービス (SOAP) などを用いて、必要なデータを取得できる	
75	4	データ共有	★★	データ展開	利用者の要件に合致したレポート (図、表) を、PDFやPostScriptなどの印刷用フォーマットで出力する変換機能を設計できる	
76	5	データ共有	★★	データ展開	BIツール (Tableau、QlikView、PowerBI、Cognos、SASなど) 用のサーバーを構築し、単一データソースのレポート用DBを設計・構築できる	
77	6	データ共有	★★	データ展開	データストア上のデータをメールやメッセージ (Webサービスなど) を用いてプッシュ配信するシステムのサーバー・ネットワーク・ソフトウェアの構成を設計できる	
78	7	データ共有	★★	データ展開	RSS、RDFや業界標準フォーマットなど要件に合致したデータ形式・配信形式で、情報提供するシステムのインターフェースを設計できる	
79	8	データ共有	★★	データ展開	Web API (REST) やWebサービス (SOAP) などを用いて、必要なデータを提供するシステムの公開インターフェースを設計できる	
80	9	データ共有	★	データ連携	FTPサーバー、ファイル共有サーバーから必要なデータファイルをダウンロードして、Excelなどの表計算ソフトに取り込み活用できる	○
81	10	データ共有	★	データ連携	BIツール (Tableau、QlikView、PowerBI、Cognos、SASなど) のレポート編集機能を用いて新規レポートを公開できる	
82	11	データ共有	★	データ連携	BIツールの自由検索機能を活用し、必要なデータを抽出して、グラフを作成できる	
83	12	データ共有	★★	データ連携	連携対象システムの仕様に合わせて、ETLツールを用いたデータ変換、ファイル転送処理を実装できる	
84	13	データ共有	★★★	データ連携	ESB・EAIなどのデータ連携基盤を活用してシステム間のデータ連携 (データ配信・交換) を行うインターフェースを設計できる	
85	1	プログラミング	★	基礎プログラミング	小規模な構造化データ (CSV、RDBなど) を扱うデータ処理 (抽出・加工・分析など) を、設計書に基づき、プログラム実装できる	○
86	2	プログラミング	★	データインタフェース	JSON、XMLなど標準的なフォーマットのデータを受け渡すために、APIを使用したプログラムを設計・実装できる	
87	3	プログラミング	★★	アルゴリズム	非効率なループ処理の性能向上などのために、計算量やメモリを意識したプログラム実装ができる	
88	4	プログラミング	★★★	アルゴリズム	最新の論文に発表された数式処理や機械学習ロジックをプログラム実装できる	
89	5	プログラミング	★★	拡張プログラミング	データ型を意識したプログラム実装・試験ができる (C言語での性能・誤差を意識したデータ型の実装、Pythonでのエラー時のデータ型を意識したデバッグなど)	
90	6	プログラミング	★★	拡張プログラミング	GPU (GPGPU) 環境を処理するライブラリを有効に活用できる	
91	7	プログラミング	★★	拡張プログラミング	異なるタイプの複数の処理を効率よく行うために、スクリプトを用いたプログラムを設計・実装できる (ハイブライン処理のluigiなど)	
92	8	プログラミング	★★★	拡張プログラミング	対象プラットフォーム (iOS、Android、HEMSなど) におけるデータ取得の業界標準を理解しており、今後の技術動向や規制についてのリスクを提示できる	

No	Sub No	スキルカテゴリ	スキルレベル	サブカテゴリ	データエンジニアリングカ:チェック項目	必須スキル
93	9	プログラミング	★★	データ規模	SNSから抽出した非構造化データを、適切な言語やライブラリを選んでプログラム実装できる	
94	10	プログラミング	★★★	データ規模	単一サーバーの物理メモリを超える複数のデータソースを組合わせたデータ処理において、分散処理アーキテクチャやデータのインメモリ処理の特性を意識してプログラム設計ができる	
95	11	プログラミング	★★	分析プログラム	分析プログラムのロジックと処理手順を理解した上で正しい分析結果を出力しているか検証ができる	○
96	12	プログラミング	★★★	リアルタイム処理	ストリーミング処理や複合イベント処理 (CEP) などを設計し、適切な言語やライブラリを選んでプログラム実装できる	
97	13	プログラミング	★	SQL	SQLで簡単なSELECT文を記述・実行できる (副問合せを含まない、2テーブル程度の結合と簡単なWHERE条件を含むSELECT文)	○
98	14	プログラミング	★★	SQL	SQLの構文を一通り知っていて、記述・実行できる (DML・DDLの理解、各種JOINの使い分け、集計関数とGROUP BY、CASE文を使用した縦横変換、副問合せやEXISTSの活用など)	○
99	15	プログラミング	★★★	SQL	RDBにおける分析関数の構文と挙動を理解し、分析関数を用いて複雑な副問合せや自己結合を解消できる	
100	16	プログラミング	★★★	SQL	N:Nの結合や完全外部結合の危険性 (計算量の増大、結果の不完全性)、暗黙の型変換の危険性 (インデックス不使用、小数点以下の切り捨てなど) を考慮したSQLを記述できる	
101	17	プログラミング	★★★	SQL	記述したSQLの実行計画の確認と判断ができ、SQLの修正やインデックス作成により、処理時間を大幅に改善するようなパフォーマンスチューニングができる	
102	18	プログラミング	★★	Pig/Hive	Hiveを記述して、パーティションが切られているデータを適切に処理できる	
103	19	プログラミング	★★	Pig/Hive	Pigを記述して行数・内容の異なる行が混在しているデータセットやネスト構造を持つデータセットを処理できる	
104	20	プログラミング	★★★	Pig/Hive	HiveまたはPigで使用するためのUDFが実装できる	
105	1	ITセキュリティ	★	基礎知識	セキュリティの3要素 (機密性、可用性、完全性) について具体的な事例を用いて説明できる	○
106	2	ITセキュリティ	★★	プライバシー	データ匿名化の方法を理解し、匿名化方法 (ハッシュ化、マスキング、k-匿名化など) に応じた加工処理を設計できる	○
107	3	ITセキュリティ	★	攻撃と防御手法	コンピュータウイルスとワームの違いを説明でき、データやシステムを防御する方法を理解している	
108	4	ITセキュリティ	★	攻撃と防御手法	ファイルシステム、ネットワーク、アプリケーション、データに対するユーザーごとのアクセスレベルを手順に従い設定できる	
109	5	ITセキュリティ	★	攻撃と防御手法	DoS攻撃、不正アクセス、マルウェアなど外部からの攻撃が発覚した場合に既存のルールに基づき対応できる	
110	6	ITセキュリティ	★★	攻撃と防御手法	ファイルシステム、ネットワーク、アプリケーション、データに対するユーザーごとのアクセスレベルを設計できる	
111	7	ITセキュリティ	★★	攻撃と防御手法	SQLインジェクションやバッファオーバーフローなどの素通りを防止する対策を判断できる	
112	8	ITセキュリティ	★★★	攻撃と防御手法	なりすまし、改ざん、中間者攻撃による盗聴などを防御するための対策を特定できる	
113	9	ITセキュリティ	★★★	攻撃と防御手法	侵入検知システム (IDS) やファイアウォールなど外部からの不正アクセスを検知、防御する環境を設計できる	
114	10	ITセキュリティ	★★★	攻撃と防御手法	個別の案件ごとに、依頼元との契約約款、依頼元がデータをどのように保持し利用するかに応じて、適切な匿名化の手法を選択し適用できる	
115	11	ITセキュリティ	★	暗号化技術	暗号化されていないデータは盗聴される恐れがあることを理解し、データの機密度合いに応じてソフトウェアを使用した暗号化と復号化ができる	○
116	12	ITセキュリティ	★	暗号化技術	なりすましや偽造された文書でないことを証明するために電子署名が必要であることを理解している	
117	13	ITセキュリティ	★	暗号化技術	一方ハッシュ関数を用いて、データの改ざんを検出できる	
118	14	ITセキュリティ	★★	暗号化技術	SSHやSSL/TLSなどのセキュアプロトコルの概要と必要性を説明できる	
119	15	ITセキュリティ	★★★	暗号化技術	DES、AES、RC4を用いたKerberos認証が使われる事例と仕組みを説明できる	

ビジネスカ

Total: 123
★: 30
★★: 49
★★★: 44

No	Sub No	スキルカテゴリ	スキルレベル	サブカテゴリ	ビジネスカ: チェック項目	必須スキル
1	1	行動規範	★	ビジネスマインド	ビジネスにおける論理とデータの重要性を認識し、分析的でデータドリブンな考え方に基づき行動できる	○
2	2	行動規範	★★	ビジネスマインド	ビジネスではスピード感がより重要であることを認識し、時間と情報が限られた状況下でも、言わば「ザックリ感」を持って素早く意思決定を行うことができる	
3	3	行動規範	★★	ビジネスマインド	作業ありきではなく、本質的な問題（イシュー）ありきで行動できる	○
4	4	行動規範	★★	ビジネスマインド	最終的な結論に関わる部分や、ストーリーラインの骨格に大きな影響を持つ部分から着手するなど、取り組むべき分析上のタスクの優先度を判断できる	
5	5	行動規範	★★★	ビジネスマインド	プロフェッショナルとして、作業量ではなく、生み出す価値視点で常に判断、行動でき、依頼元にとって真に価値あるアウトプットを生み出すことをコミットできる	○
6	6	行動規範	★	データ倫理	データを取り扱う人間として相応しい倫理を身に着けている（データのねつ造、改ざん、盗用を行わないなど）	○
7	7	行動規範	★★	データ倫理	チーム全員がデータを取り扱う人間として相応しい倫理を持てるよう、適切にチームを管理できる	
8	8	行動規範	★★★	データ倫理	データの取り扱いに関する、会社や組織全体の倫理を維持、向上させるために、必要な制度や仕組みを策定し、その運営を主導することができる	
9	9	行動規範	★	法令	個人情報に関する法令の概要を理解している	○
10	10	行動規範	★★	法令	担当するビジネスや業界に関する法令を理解しており、データの保持期間や運用ルールに活かすことができる	○
11	11	行動規範	★★	法令	個人情報の扱いに関する法令、その他のプライバシーの問題、依頼元との契約約款に基づき、匿名化するべきデータを選別できる（名寄せにより個人を特定できるもの、依頼元がデータ処理の結果をどのように保持し利用するのかなども考慮して）	
12	1	論理的思考	★	MECE	データや事象の重複に気づくことができる	○
13	2	論理的思考	★★	MECE	初見の領域に対して、抜け漏れや重複をなくすことができる	
14	3	論理的思考	★★★	MECE	未知の領域であっても、類似する事象の推測などを活用し、抜け漏れや重複をなくすことができる	
15	4	論理的思考	★★		様々なデータや事象を、階層やグルーピングによって、構造化できる（ピラミッド構造）	
16	5	論理的思考	★★		データをもとに論理的に捉える	
17	6	論理的思考	★★★		仮説思考を用いて、論点毎に分析すべき点を識別できる	
18	7	論理的思考	★		データをもとに論理的に捉える	○
19	8	論理的思考	★★		論理的な整理にとらわれず、批判的・複合的な視点で課題を識別できる	
20	9	論理的思考	★★★		分析結果の意味合いを適切に言語化できる	○
21	10	論理的思考	★		近い概念がある場合、データ表現に対して適切な言葉を選択できる	
22	11	論理的思考	★★		データ表現に適した言葉がない場合でも、適切な言葉を新たに作り出すことができる	
23	12	論理的思考	★★★		一般的な論文構成について理解している（序論⇒アプローチ⇒検討結果⇒考察や、序論⇒本論⇒結論 など）	
24	13	論理的思考	★		因果関係に基づいて、ストーリーラインを作れる（観察⇒気づき⇒打ち手、So What?, Why So?など）	
25	14	論理的思考	★★		相手や内容に応じて、自在にストーリーラインを組み上げることができる	
26	15	論理的思考	★★★		1つの図表〜数枚程度のドキュメントを論理立ててまとめることができる（課題背景、アプローチ、検討結果、意味合い、ネクストステップ）	
27	16	論理的思考	★		10〜20枚程度のミニパッケージ（テキスト&図表）、もしくは5ページ程度の図表込みのビジネスレポートを論理立てて作成できる	
28	17	論理的思考	★★		30〜50枚程度のフルパッケージ（テキスト&図表）、もしくは10ページ以上のビジネスレポートを論理立てて作成できる	
29	18	論理的思考	★★★		プレゼンテーションの論拠不足や論理破綻を指摘された際に、すみやかに理解できる	
30	1	プロセス	★		論理的なプレゼンテーションができる	
31	2	プロセス	★		プレゼンテーションの相手からの質問や反論に対して、説得力のある形で回答できる	
32	3	プロセス	★		目的やゴールの設定がないままデータを分析しても、意味合いが出ないことを理解している	○

No	Sub No	スキルカテゴリ	スキルレベル	サブカテゴリ	ビジネスカ: チェック項目	必須スキル
31	2	プロセス	★	基礎知識	課題や仮説を言語化することの重要性を理解している	
32	3	プロセス	★	KPI	一般的な収益方程式、バランスシートなどに加え、自らが担当する業務の主要な変数（KPI）を理解している	
33	4	プロセス	★★	KPI	自らが関連する事業領域であれば、複数の課題レイヤーにまたがっていても、KPIを整理・構造化できる	
34	5	プロセス	★★★	KPI	初見の事業領域であっても、KPIを構造化し、重要なKPIを見極められる	
35	6	プロセス	★		自分の担当する業界について、市場規模、主要なプレーヤー、支配的なビジネスモデル、課題と機会について説明できる	○
36	7	プロセス	★★		事業モデルやバリューチェーンなどの特徴や事業の主たる課題を自力で構造的に理解でき、問題の大枠を整理できる	
37	8	プロセス	★★★		領域の主要課題を他領域の課題との関連も含めて構造的に理解でき、問題の大枠を定義できる	
38	9	プロセス	★		担当する業務領域であれば、基本的な課題の枠組みが理解できる（調達活動を5フォースで整理する、CRM課題をRFMで整理するなど）	
39	10	プロセス	★★		取り扱う課題領域に対して、新規課題の切り分けや枠組み化ができる	
40	11	プロセス	★★★		仮説や可視化された問題がなくとも、解くべき課題を構造的に整理でき、見極めるべき論点を特定できる	○
41	12	プロセス	★		ウォーターフォール開発とアジャイル開発の違いを理解している	
42	13	プロセス	★★		アジャイル開発体制のポイントを理解した上で、アジャイルな開発チームを立ち上げ、推進することができる	
43	14	プロセス	★		スコープ、検討範囲・内容が明快に設定されていれば、必要なデータ、分析手法、可視化などを適切に選択できる	
44	15	プロセス	★★		解くべき課題がフレーミングされていれば、必要なデータ、分析手法、可視化などを適切に選択できる	○
45	16	プロセス	★★★		複数の事業や課題にまたがっていても、必要なデータ、分析手法、可視化などを適切に選択し作業手順に落とし込める	
46	17	プロセス	★★★		複数のアプローチの組み合わせでしか解けない課題であっても、その解決までの道筋を設計できる	
47	18	プロセス	★★★		他社による模倣を防ぐなど、競争力を保つ観点でアプローチの設計ができる（機械化や標準化の範囲を絞るなど）	
48	19	プロセス	★★		分析で解くべき課題か否かを判断できる	
49	20	プロセス	★★★		プロジェクトの開始時点で、入手可能なデータ、分析手法、インフラ、ツールの生み出すビジネス価値を適切に見積りができる	
50	1	データの理解・検証	★		ニュース記事などで統計情報に接したときに、数字やグラフの持つメッセージを理解できる	
51	2	データの理解・検証	★★		積極的に統計情報を収集しているとともに、表現に惑わされず数字を正当に評価できる（原点が0ではないグラフ、不当に誇張されたグラフなど）	
52	3	データの理解・検証	★★★		数字やデータの検証のために、何と比較するべきかすみやかに把握し、収集・利用できる（業務データや過去に接触した統計情報の想起・活用を含む）	
53	4	データの理解・検証	★		単独のグラフに対して、集計ミスなどがないかチェックできる	○
54	5	データの理解・検証	★		データ項目やデータの量・質について、指示のもと正しく検証し、結果を説明できる	
55	6	データの理解・検証	★★		複数のグラフや集計表で構成されているレポートに対して、全体として集計ミスや不整合が起きていないかチェックできる	○
56	7	データの理解・検証	★★		データ項目やデータの量・質の検証方法を計画・実行し、その結果をもとにその後の分析プロセスを立案・修正できる	
57	8	データの理解・検証	★★★		多数のグラフ、集計表、外部の統計情報、高度なデータ解析手法を用いた解析結果などを含むレポートに対して、不整合が起きていないか、妥当性の高い論理構造であるかチェックできる	
58	9	データの理解・検証	★★★		分析に必要なデータを想定し、現在取得可能なデータで量・質ともに分析に耐える内容であるか判断できる。また、目的が達成可能であるか見込みが立てられる	○
59	10	データの理解・検証	★		データの裏を読むことの重要性を理解している	
60	11	データの理解・検証	★★		データを俯瞰して、変化をすみやかに察知できる。また、変化が誤差の範囲かどうか判断できる	○
61	12	データの理解・検証	★★★		複数のデータを多元的かつ大局的に俯瞰して、大きな動きや本質的な事実を見抜くことができる	○
62	13	データの理解・検証	★		データから事実を正しく浮き彫りにするために、集計の切り口や比較対象の設定が重要であることを理解している	

No	Sub No	スキルカテゴリ	スキルレベル	サブカテゴリ	ビジネス力: チェック項目	必須スキル
63	14	データの理解・検証	★	データ理解	普段業務で扱っているデータの発生トリガー・タイミング・頻度などを説明でき、また基本統計量を把握している	
64	15	データの理解・検証	★	データ理解	何のために集計しているか、どのような知見を得たいのか、目的に即して集計できる	○
65	16	データの理解・検証	★★	データ理解	生データを眺めて、どのような切り口で集計・比較すればデータの理解や事実の把握につながるか検討できる	○
66	17	データの理解・検証	★★	データ理解	扱ったことのない新たなデータに内容の不明な項目があっても、生データの閲覧や集計を通して何の項目かあたりをつけられる	
67	18	データの理解・検証	★★	データ理解	扱っているデータの関連業務の知識と分析目的を踏まえて、どんな説明変数が効きそうか、あたりをつけて洗い出し、構造的に整理できる	
68	19	データの理解・検証	★★★	データ理解	データの変化の裏で起きている事実を読み取り、仮説を立て検証方法を立案できる	
69	20	データの理解・検証	★★★	データ理解	データを入手する前に、存在するであろうデータとその分布を想定して基礎俯瞰の方向性やその結果の想定ができ、それを前提とした解析方法の検討・ラフ設計をすることができる	
70	21	データの理解・検証	★★★	データ理解	扱ったことのない新たなデータであっても、ER図やテーブル定義、生データなどを見ることによってデータの発生源や欠損値の意味などのあたりをつけられる	
71	22	データの理解・検証	★★	データ粒度	分析目的とデータの量・質を踏まえて、想定されるメッセージと統計的観点から適切な集計単位とサンプリング率を決定できる	
72	23	データの理解・検証	★★★	データ粒度	分析目的とデータの量・質に加えて、想定しているメッセージ、深掘りの方向性・可能性、処理負荷、データ処理フローなども総合的に踏まえた最適な集計単位とサンプリング率を決定できる	
73	24	データの理解・検証	★	ビジネス観点	ビジネス観点で仮説を持ってデータをみることの重要性和、仮説と異なる結果となったときにそれが重大な知見である可能性を理解している	
74	25	データの理解・検証	★★	ビジネス観点	特定の分析プロセスにおいて、統計手法による間値や係数（量的変数の統計的カテゴリなど）に対して商品単価や購入周期を考慮するなど、ビジネス観点で納得感のある調整ができる	
75	26	データの理解・検証	★★★	ビジネス観点	分析プロセス全体を通して常時、ビジネス観点での妥当性をチェックしている。また、データから得られた示唆がビジネス観点で新しい知見であるか都度判断できる	○
76	1	データ入手	★	データ入手	仮説や既知の問題が与えられた中で、必要なデータにあたりをつけ、入手できる	○
77	2	データ入手	★	データ入手	現場に向いてヒアリングするなど、一次情報に接することの重要性を理解している	
78	3	データ入手	★★	データ入手	自身が担当するプロジェクトやサービスを越えて、必要なデータのあたりをつけることができる	○
79	4	データ入手	★★	データ入手	既存のサービスやアプリケーションに対して、有効な分析をするためのログ出力の追加仕様を整理することができる	
80	5	データ入手	★★★	データ入手	組織全体及び関連する社外のデータを見渡して、必要なデータのあたりをつけることができる	
81	6	データ入手	★★★	データ入手	入手可能なデータに加え、分析結果の品質・効果を向上させる新たなデータ入手方法を提案できる（IoTでの新設センサーの種類・配置場所・データ入手間隔など）	
82	1	意味合いの抽出、洞察	★★★	開示	担当および関連する分析プロジェクトのデータ、分析結果を顧客、外部に開示すべきか判断できる	
83	2	意味合いの抽出、洞察	★	洞察	分析、図表から直接的な意味合いを抽出できる（バラツキ、有意性、分布傾向、特異性、関連性、変曲点、関連度の高低など）	○
84	3	意味合いの抽出、洞察	★	洞察	想定に影響されず、分析結果の数値を客観的に解釈できる	
85	4	意味合いの抽出、洞察	★★	洞察	ビジネス上の意味を捉えるために、特異点、相違性、傾向性、関連性を見出したうえで、ドメイン知識を持つ人に適切な質問を投げかけられる	
86	5	意味合いの抽出、洞察	★★	洞察	データの特徴を見て意味合いの明確化に向けた分析の深掘り、データ見直しの方向性を設計できる	
87	6	意味合いの抽出、洞察	★★★	洞察	担当および関連する分析プロジェクトのデータを眺め、裏にある構造や意味合い（真実）を見ぬくことができる	
88	7	意味合いの抽出、洞察	★★	評価	担当する分析プロジェクトの分析結果を見て検討目的と合っているか再評価できる	○
89	8	意味合いの抽出、洞察	★★	評価	各種の解析手法（主成分分析、クラスター分析、決定木分析など）の結果を解釈し、意味合いを適切に表現・説明できる	
90	9	意味合いの抽出、洞察	★★	評価	分析結果が当初の目的を満たしていない場合に、必要な分析手順を追加できる	
91	10	意味合いの抽出、洞察	★★★	評価	分析結果が当初の目的を満たしていない場合に、必要に応じてプロジェクト全体を再設計できる	
92	1	解決	★★	総合的評価	分析内容を総合的に評価し、モデルの定性的な評価ができる	

No	Sub No	スキルカテゴリ	スキルレベル	サブカテゴリ	ビジネス力: チェック項目	必須スキル
93	2	解決	★★	提案	分析的検討に基づき、担当業務に対する必要なアクション、改革案を整理して結論を導くことができる	
94	3	解決	★★★	提案	分析的検討に基づき、経営レベルに必要なアクション、改革案を整理して結論を導くことができる	○
95	4	解決	★★★	提案	データを利活用した持続性のある事業モデルを設計できる	
96	1	事業に実装する	★★	実装	現場に実装する際、現場での実行可能性を考慮し適切に対応できる	
97	2	事業に実装する	★★	実装	担当する案件が自分の持っている予算内で解決するように取り組みをデザインし、実装できる	
98	3	事業に実装する	★★★	実装	費用対効果、実行可能性、業務負荷を考慮した実装ができる	
99	4	事業に実装する	★★	異分野とのコミュニケーション	異なるスキル分野の専門家や事業者と適切なコミュニケーションをとりながら実装を進めることができる	○
100	5	事業に実装する	★	モニタリング	やりっぱなしではなく結果、改善の度合いをモニタリングする重要性を理解している	
101	6	事業に実装する	★★	モニタリング	事業・現場へ実装するにあたりモニタリングの仕組みを適切に組み込むことができる	
102	7	事業に実装する	★★★	モニタリング	既存のPDCAに新たな仕込みを行い、次の改善的な取り組みにつなげることができる	
103	8	事業に実装する	★★★	横展開	特定のビジネス課題に向けた新しいソリューションを個別の現場の特性を考慮し横展開できる	
104	1	活動マネジメント	★★★	データ倫理	ウイルス、DDoS攻撃などの深刻なセキュリティ攻撃を受けた場合に対応する最新の技術を熟知し、対応する専門組織（CSIRT）の構成を責任者にすみやかに提案できる	
105	2	活動マネジメント	★	リソースマネジメント	指示に従ってスケジュールを守り、チームリーダーに頼まれた自分の仕事を完遂できる	○
106	3	活動マネジメント	★★	リソースマネジメント	自身とチームメンバーのスキルを把握し、適切なプロジェクト管理ができる	○
107	4	活動マネジメント	★★	リソースマネジメント	担当するプロジェクトで、設定された予算やツール、システム環境を適切に活用し、プロジェクト進行できる	
108	5	活動マネジメント	★★	リソースマネジメント	5名前後のチームを設計し、スケジュール通りに進行させ、ステークホルダーに対して期待値に見合うアウトプットを安定的に生み出せる	
109	6	活動マネジメント	★★★	リソースマネジメント	プロジェクトに求められるスキル要件と各メンバーのスキル・成長目標・性格をふまえ、現実的にトレードオフ解消とシナジーを狙ったリソースマネジメントができる	
110	7	活動マネジメント	★★★	リソースマネジメント	プロジェクトメンバーの技量を把握した上で、プロジェクト完遂に必要なツール選定、予算策定、スコープ設定、またはアウトソーシング体制を検討・構築できる	
111	8	活動マネジメント	★★★	リソースマネジメント	複数のチームを設計し（総合的なプロジェクトマネジメント）、スケジュール通りに進行させ、複合的なステークホルダーに対し、期待値を超えたアウトプットを安定的に生み出せる	
112	9	活動マネジメント	★	リスクマネジメント	担当するタスクの遅延や障害などを発見した場合、プロジェクトオーナーに迅速かつ適切に報告ができる	
113	10	活動マネジメント	★★	リスクマネジメント	担当するプロジェクトでの遅延や障害などの発生を検知し、リカバリーするための提案・設計ができる	
114	11	活動マネジメント	★★★	リスクマネジメント	プロジェクトに何らかの遅延・障害などが発生した場合、適切なリカバリー手順の判断、リカバリー体制構築、プロジェクトオーナーに対する迅速な対応ができる	○
115	12	活動マネジメント	★★	育成/ナレッジ共有	自身とチームメンバーのスキルを大まかに把握し、担当するプロジェクトを通してチームメンバーへのスキルアドバイスやスキル成長のための目標管理ができる	
116	13	活動マネジメント	★★	育成/ナレッジ共有	チームメンバーのスキルに応じ、研修参加や情報収集への適切なアドバイスやチーム内でのナレッジ共有を推進できる	
117	14	活動マネジメント	★★★	育成/ナレッジ共有	チームの各メンバーに対し、データサイエンティストとしてのスキル目標の設定、到達させるためのプロジェクトを通した適切なアドバイスができる	
118	15	活動マネジメント	★★★	育成/ナレッジ共有	チーム育成の上で、データサイエンティストに求められるスキルについて、研修制度の設計やナレッジ共有の仕組み構築と運営ができる	
119	16	活動マネジメント	★★★	育成/ナレッジ共有	チームに必要な情報やデータサイエンスの新しい技術・手法に関する情報収集戦略やラーニング方法を主導し、自ら情報を取捨選択し、チームにフィードバックできる	
120	17	活動マネジメント	★★	営業マネジメント	ビジネス要件を整理し、分析・データ活用のプロジェクトを企画・提案することができる	
121	18	活動マネジメント	★★★	営業マネジメント	依頼元やステークホルダーのビジネスをデータ面から理解し、分析・データ活用のプロジェクトを立ち上げ、プロジェクトにかかるコストと依頼元の利益を説明できる	
122	19	活動マネジメント	★★	組織マネジメント	データサイエンスチームの役割を認識し、担当するプロジェクトにおいて、組織内や他部門・他社間でのタスク設定や調整ができる	

No	Sub No	スキルカテゴリ	スキルレベル	サブカテゴリ	ビジネスカ: チェック項目	必須スキル
123	20	活動マネジメント	★★★	組織マネジメント	データサイエンスチームを自社・他社の様々な組織と関連付け、組織内での役割の規定、組織的な目標設定を行う	

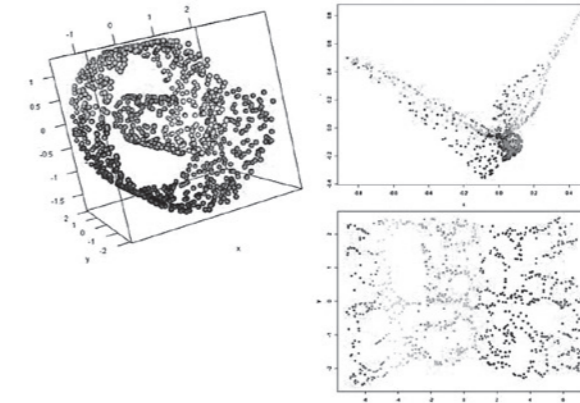
【添付別紙】

スキルチェックリスト：データサイエンスカ

Data visualizationに関するイメージ共有のための参考資料（例示）

項目NO

10 データの持つ分散量の観点で、高次のデータを主成分分析(PCA)などにより1~3次元のデータに変換できる

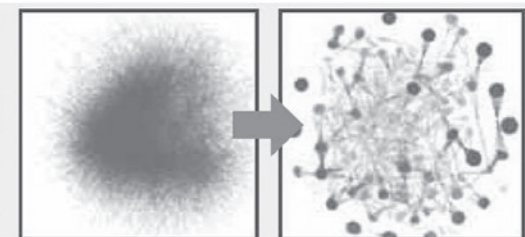


引用元：株式会社ALBERT WEBサイト
http://www.albert2005.co.jp/analyst_blog/?p=1214

13 ネットワーク構造、グラフ構造などの表現において、ノードとエッジが増えすぎて特徴抽出が困難であっても、データの絞り込みや抽象度を上げることで適切に可視化できる

図2 企業間取引データをクラスター化して把握しやすくした例

	研磨前	研磨後
頂点数	3,282	3,282
枝数	35,168	73,132
クリーク数	32,953	343



引用元：（出版物）大学共同利用機関法人 情報・システム研究機構 国立情報学研究所 発行
 「国立情報学研究所ニュース [NII Today] 第66号 平成26年12月」 6ページ
http://www.nii.ac.jp/userdata/results/pr_data/NII_Today/66/all.pdf

19 強調表現がもたらす効果を理解している(計量データに対しては位置やサイズ表現が色表現よりも効果的など)

表現例	量を表現する	順序を表現する	分類を表現する	関係を表現する
位置 	✓	✓	✓	✓
長さ 	✓	✓		
明度/彩度・色 		✓		✓

引用元：「solutions-2」WEBサイト
<http://www.solutions2.be/Books.aspx> をもとに独自に修正加筆

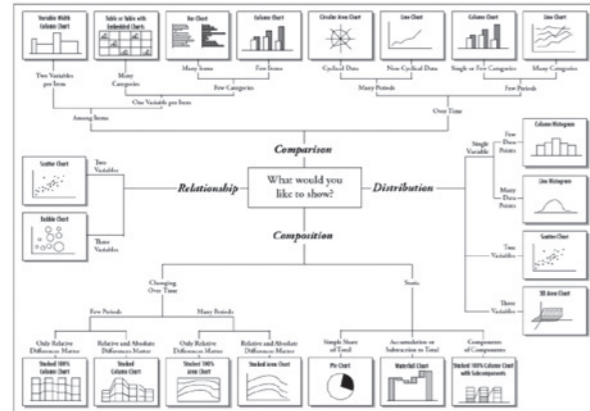
【添付別紙】

スキルチェックリスト：データサイエンサ

Data visualizationに関するイメージ共有のための参考資料（例示）

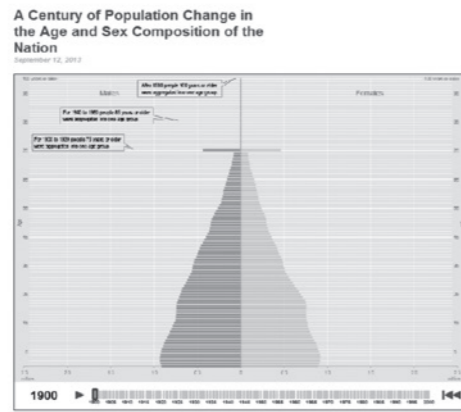
項目NO

20 1〜3次元の比較において目的(比較、構成、分布、変化など)に応じた図表化ができる



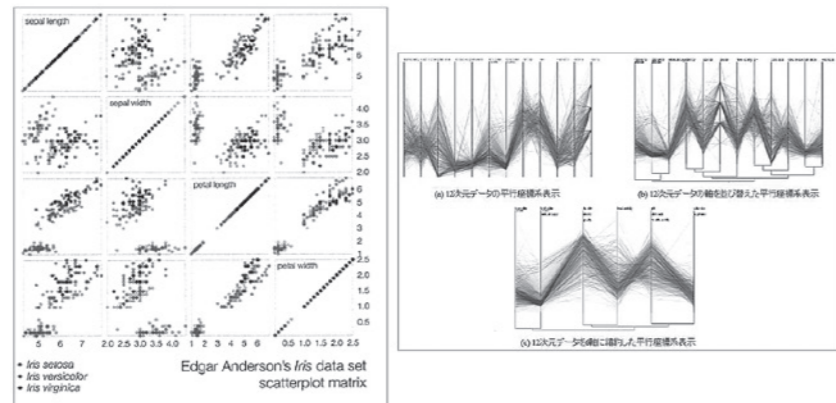
引用元： Amit Agarwal 「Digital Inspiration」 WEBサイト
http://www.labnol.org/software/find-right-chart-type-for-your-data/6523/

21 端的に図表の変化をアニメーションで可視化できる(人口動態のヒストグラムが経年変化する様子を表現するなど)



引用元： アメリカ国税調査局 WEBサイト
http://www.census.gov/dataviz/visualizations/055/

23 1〜3次元の図表を拡張した多変量の比較(平行座標、散佈図行列、テーブルレンズ、ヒートマップなど)を適切に可視化できる



引用元： [左図] 「Data-Driven Documents」 WEBサイト http://mboostock.github.io/d3/talk/20111116/iris-splom.html
[右図] 「グラフスペクトル解析を用いた軸縮約可能平行座標系」 WEBサイト
http://user.keio.ac.jp/~yun/projects/cpc/index-j.html

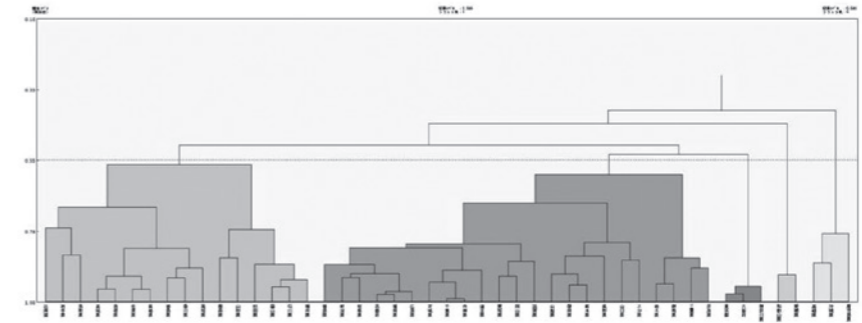
【添付別紙】

スキルチェックリスト：データサイエンサ

Data visualizationに関するイメージ共有のための参考資料（例示）

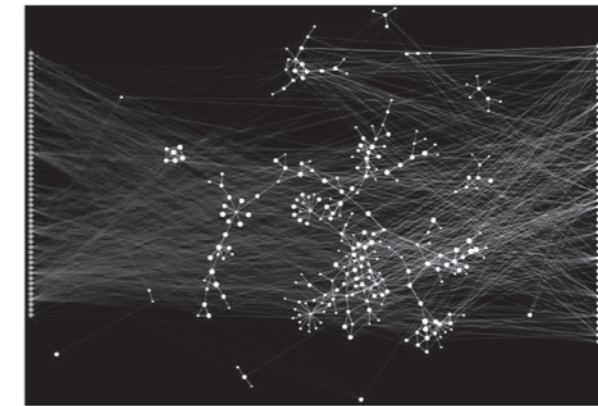
項目NO

24 ネットワーク構造、グラフ構造、階層構造などの統計的な関係性の可視化ができる



引用元： 株式会社ALBERT WEBサイト
http://www.albert2005.co.jp/technology/images/tech_mining_img106.jpg

24 ネットワーク構造、グラフ構造、階層構造などの統計的な関係性の可視化ができる



赤・白点：Wine
黄点：Cheese

引用元： 「Cytoscape.js」 WEBサイト (Wine & cheese)
http://cytoscape.github.io/cytoscape.js/demos/cde4db55e581d10405f5/

25 GPSデータなどを平面地図上に重ね合わせた可視化ができる



引用元： 「A WORLD OF TWEETS」 WEBサイト
http://aworldoftweets.frogdesign.com/

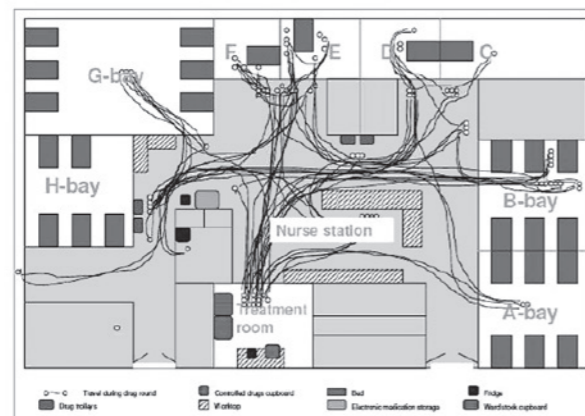
【添付別紙】

スキルチェックリスト：データサイエンサカ

Data visualizationに関するイメージ共有のための参考資料（例示）

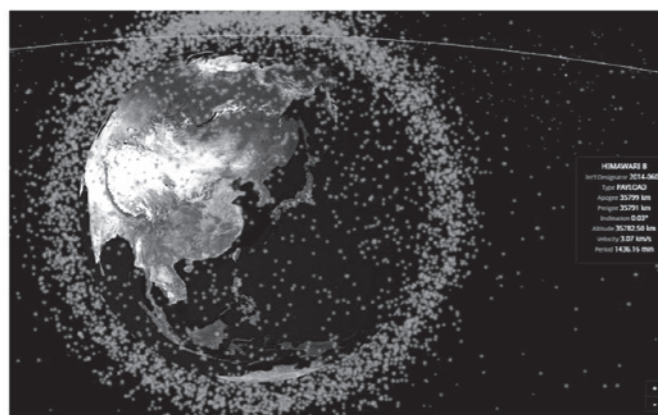
項目NO

26 挙動・軌跡の可視化ができる(店舗内でのユーザの回遊やEye trackingなど)



引用元：「Imperial College London」WEBサイト
<https://www1.imperial.ac.uk/cpsq/ms/m/>

28 人体、標高を持つ地図、球面などの上にデータを重ね合わせた可視化ができる



引用元：James Yoder「Stuff in Space」WEBサイト
<http://stuffin.space/>

29 地図上で同時に動く数百以上のポイントにおける時間変化を動的に表現できる(多地点での風の動き、飛行物の軌跡など)



引用元：Kaspersky Lab「CYBERTHREAT REAL-TIME MAP」WEBサイト
<https://cybermap.kaspersky.com/>

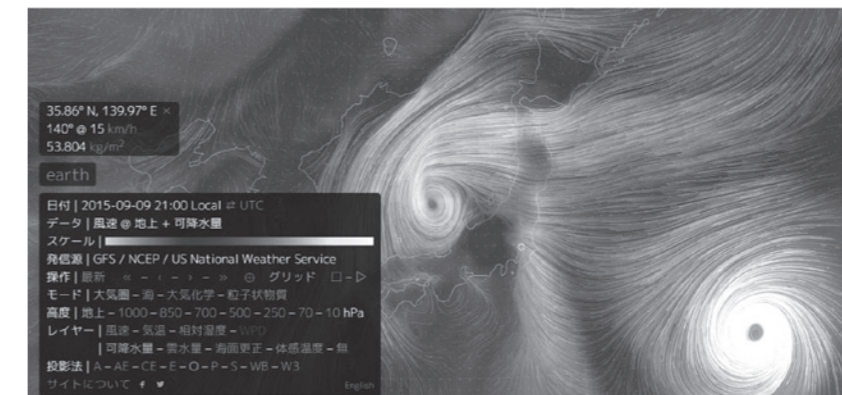
【添付別紙】

スキルチェックリスト：データサイエンサカ

Data visualizationに関するイメージ共有のための参考資料（例示）


項目NO

29 地図上で同時に動く数百以上のポイントにおける時間変化を動的に表現できる(多地点での風の動き、飛行物の軌跡など)



引用元：Cameron Beccario「earth」WEBサイト
http://earth.nullschool.net/jp/#2015/09/09/1200Z/wind/surface/level/overlay=total_precipitable_water/orthographic=-224.70,38.05,3000/grid=on

#ds_symp2nd




データ社会に求められる 新しい才能とスキル

第2回シンポジウム

2015年11月13日
一般社団法人 データサイエンティスト協会
スキル委員会

本協会設立の背景と目的

#ds_symp2nd

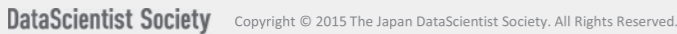


設立時の状況


- バズワードである「データサイエンティスト」という言葉の定義が欠落
- 雇い主側の期待に雇われる側のスキルセットが合致しないケースが増加
- 若い才能たちが、自分達をどう訓練し、スキルを身に付けていくべきかわからない

設立の目的

- この新しいデータプロフェッショナル（「データサイエンティスト」）に必要とされるスキルセットを定義する
- スキル育成と評価のための軸・基準を作る
- このデータプロフェッショナルの皆さんが相互に接し、やりとりできるオープンな環境を提供する
- 社会との対話を行なう


#ds_symp2nd
3


データサイエンティスト とは何か？


Copyright © 2015 The Japan DataScientist Society. All Rights Reserved.
#ds_symp2nd
4

コンピュータを持ち歩く時代に

#ds_symp2nd

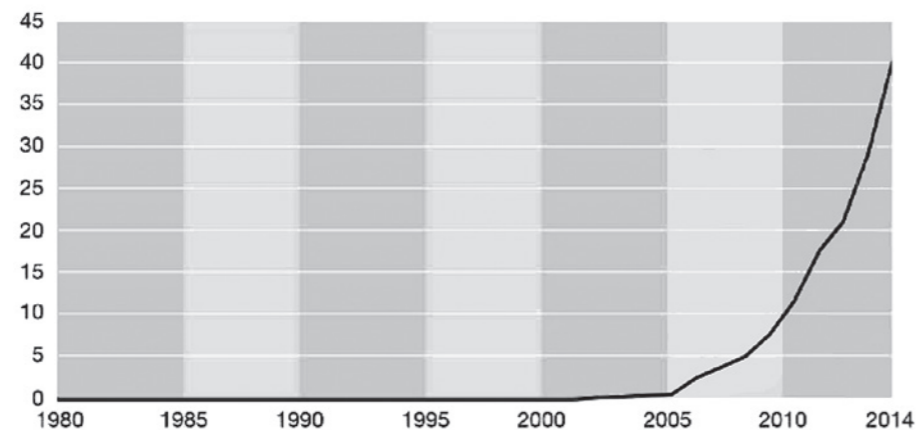
資料：安宅和人「データ時代に向けたビジネス課題とアカデミアに向けた期待」応用統計学セミナー2015.5.23
(<http://www.applstat.gr.jp/seminar/ataka.pdf>)


Copyright © 2015 The Japan DataScientist Society. All Rights Reserved.
#ds_symp2nd
5

モノからも含め
多種多様なログデータが
今後、爆発的に発生
… ビッグデータ

計算キャパシティの爆増

Annual additions to global business and consumer computing power, exaflops¹



¹An exaflop is 1 quintillion (10 to the 18th power) floating-point operations per second.
Source: William D. Nordhaus, "Two centuries of progress in computing," *Journal of Economic History*, 2007, Volume 67, Number 1, pp. 128-59; IDC; US Bureau of Economic Analysis; McKinsey analysis

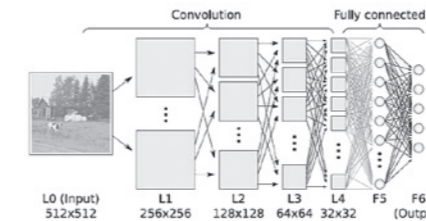
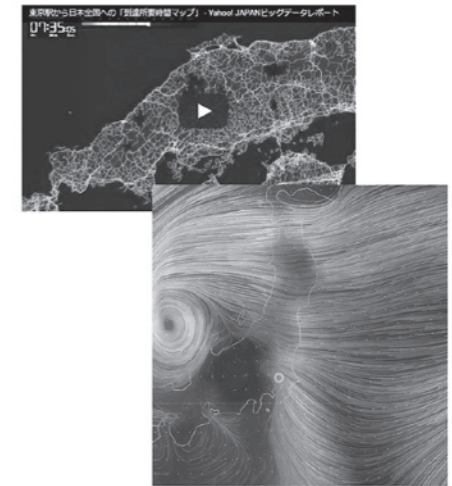
資料 : McKinsey Quarterly 2014 Number3 p.16

情報科学の進化

大量データ処理技術
深層学習の実用化



可視化 (Data visualization)
技術の劇的な進展



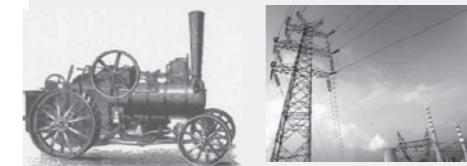
資料 : 東京大学 喜連川・豊田研HP、Universität Bonn, Institute for Computer Science HP、ヤフービッグデータレポートHP、Cameron Beccario 「earth」 WEBサイト

歴史的な局面

産業革命
(18~20世紀)

新しい
リソース

- 石炭と石油
- 蒸気機関
- 電気工学

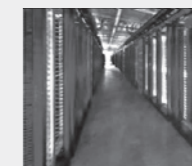


本質

- 人間と家畜を肉体労働、手作業から開放する

情報産業革命
(現在)

- ビッグデータ
- 高い計算能力
- 情報科学の進化



- 人間を退屈な数字入力、情報処理作業から開放する

資料 : データサイエンティスト協会スキル委員会ディスカッション:安宅和人「データ時代に向けたビジネス課題とアカデミアに向けた期待」応用統計学セミナー2015.5.23 (<http://www.applstat.gr.jp/seminar/ataka.pdf>)

Mission

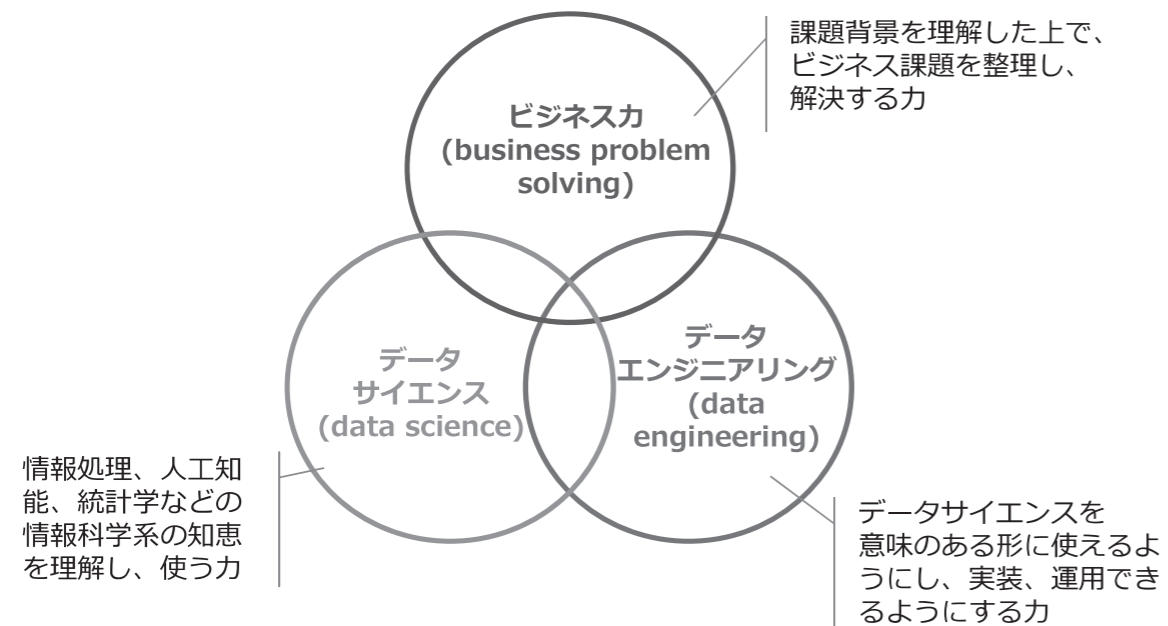
「データの持つ力を解き放つ」

DataScientist Society Copyright © 2015 The Japan DataScientist Society. All Rights Reserved.

#ds_symp2nd

10

3つのスキルセット

資料：データサイエンティスト協会プレスリリース (2014.12.10) <http://www.datascientist.or.jp/news/2014/pdf/1210.pdf>

DataScientist Society Copyright © 2015 The Japan DataScientist Society. All Rights Reserved.

#ds_symp2nd

11

データサイエンティストとは、

データサイエンスカ、
データエンジニアリング力をベースに
データから価値を創出し、
ビジネス課題に答えを出す
プロフェッショナル

* ここで「ビジネス」とは社会に役に立つ意味のある活動全般を指す

資料：データサイエンティスト協会プレスリリース (2014.12.10) <http://www.datascientist.or.jp/news/2014/pdf/1210.pdf>

DataScientist Society Copyright © 2015 The Japan DataScientist Society. All Rights Reserved.

#ds_symp2nd

12

単なるスペシャリストでもエキスパートでもない

スペシャリスト

- 何らかの分野に特化した人

エキスパート

- 何らかの分野について体系的で秀でた知識とスキルを持っている人

プロフェッショナル

- 体系的にトレーニングされた専門性を持つスキルを持ち、
- それをベースに顧客（お客様、クライアント）にコミットした価値を提供し、
- その結果に対し、認識された価値の対価として報酬を得る人

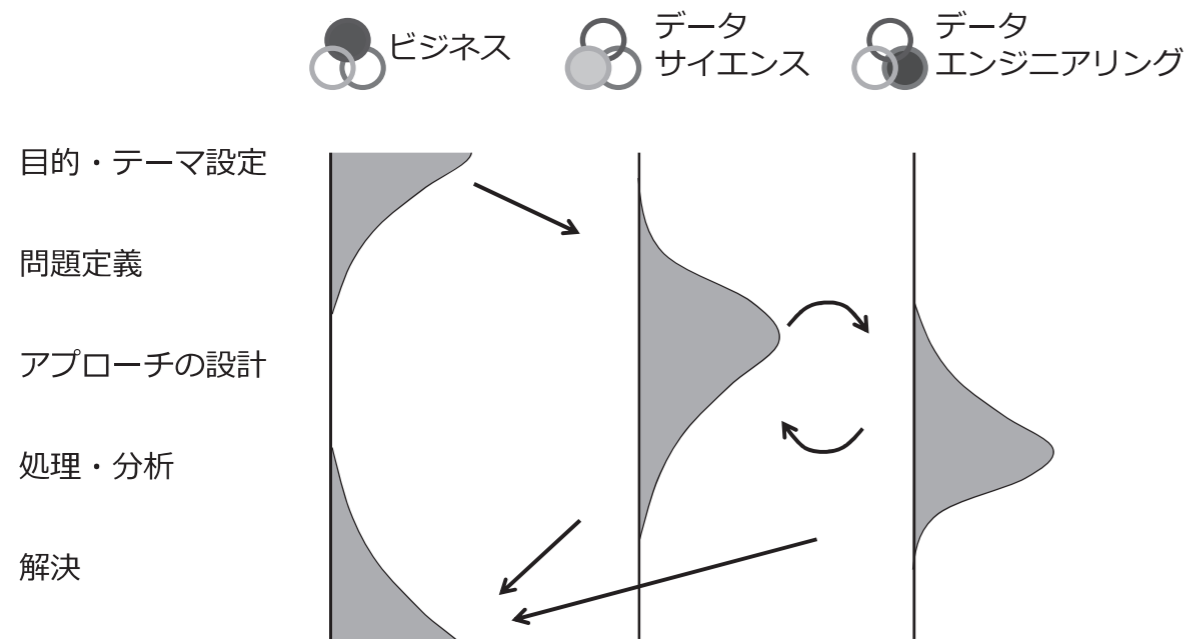
資料：データサイエンティスト協会スキル委員会ディスカッション

DataScientist Society Copyright © 2015 The Japan DataScientist Society. All Rights Reserved.

#ds_symp2nd

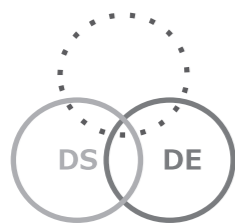
13

課題解決のフェーズによって核となるスキルが変わる



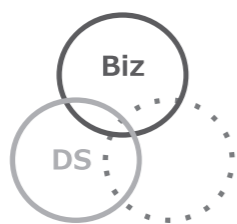
資料：データサイエンティスト協会プレスリリース (2014.12.10) <http://www.datascientist.or.jp/news/2014/pdf/1210.pdf>

どの一つが欠けてもダメ、、、優位はない



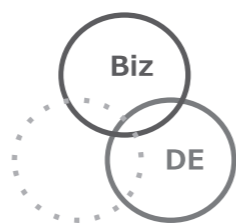
データスペシャリストではあるが、そもそも解決すべき問題が定義、整理できない

> プロフェッショナルではない



ビジネス課題もわかりそこでのサイエンスの利活用もわかるが実装できない

> 必要な変化を起こせない



ビジネス課題の上で、実装を用意できるが、かなめとなるサイエンスの知恵が足りない

> 賢いやり方を提供できない

資料：データサイエンティスト協会スキル委員会ディスカッション

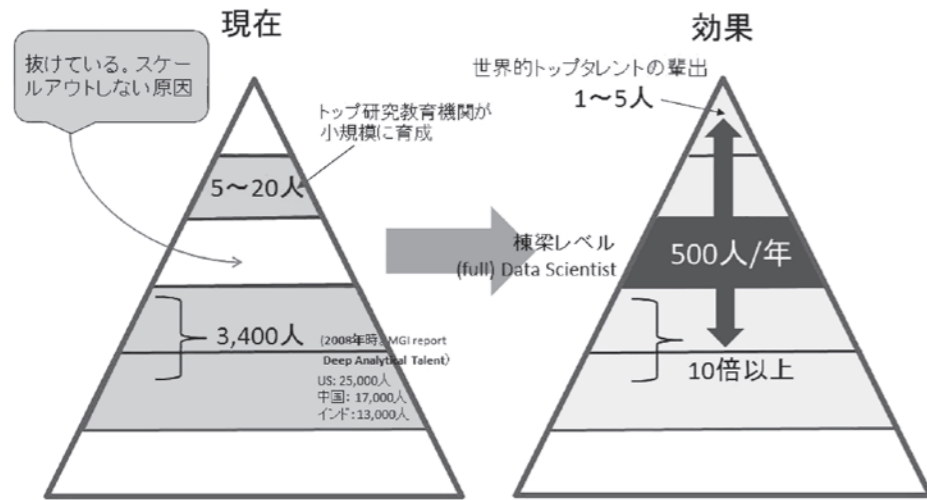
スキルのレベル感は？

スキルレベル

	スキルレベル	目安	対応できる課題
データサイエンティスト	シニア・データサイエンティスト	・ 業界を代表するレベル	・ 産業領域全体 ・ 複合的な事業全体
	フル・データサイエンティスト	・ 棟梁レベル	・ 対象組織全体
	アソシエート・データサイエンティスト	・ 独り立ちレベル	・ 担当プロジェクト全体 ・ 担当サービス全体
	アシスタント・データサイエンティスト	・ 見習いレベル	・ プロジェクトの担当テーマ
一般人	データ使い	・ 賢くデータを器用に使える人	・ 担当業務
	普通の人	・ 特になし	

資料：データサイエンティスト協会プレスリリース (2014.12.10) <http://www.datascientist.or.jp/news/2014/pdf/1210.pdf>

とりわけ棟梁レベルが足りていない



資料: 情報・システム研究機構 『ビッグデータの利活用のための専門人材育成について』 (ビッグデータの利活用に係る専門人材育成に向けた産学官懇談会) 2015.7.15

詳細検討から見えてきたこと

	気づき	ポイント
データサイエンス力	<ul style="list-style-type: none"> 情報科学も情報系、機械学習系、データ可視化系いずれかだけでは足りない 	<ul style="list-style-type: none"> 総合的な解決アプローチにつなげる力がより必要。分野別の専門性はnice to haveにすぎない
データエンジニアリング力	<ul style="list-style-type: none"> いわゆるプログラマー、SIer的なエンジニアの通常技術だけでは足りない 	<ul style="list-style-type: none"> 既存のエンジニアリング業務を超え、ビッグデータ処理を実現できる必要がある
ビジネス力	<ul style="list-style-type: none"> 問題解決力は単なる領域での知見やコンサル的なものを超える必要がある 	<ul style="list-style-type: none"> 他のスキル領域と話せる力を含めたデータ視点での課題解決力を持つ必要がある

資料: データサイエンティスト協会スキル委員会討議

内容

1. 昨年のrecap

- ① 本協会設立の背景と目的
- ② データサイエンティストとは何か?
- ③ 3つのスキルとスキルレベル

2. スキルチェックリスト

- ① 全体として (昨年 vs 今年)
- ② 個別領域の広がり
- ③ レベル感の考え方

3. 実行に向けて

- ① チームをどう作るか?
- ② 人をどう育てるか?
- ③ 目指す人はどう着手すべきか?

スキル表 (項目数)

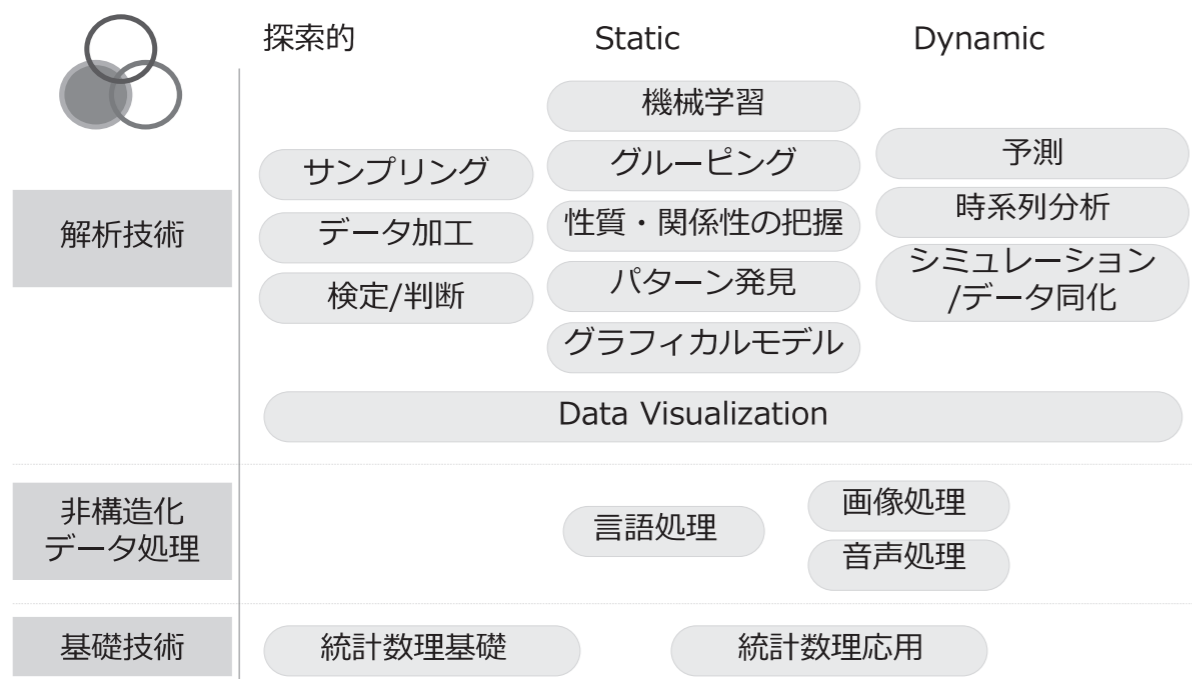
	データサイエンス	データエンジニアリング	ビジネス	レベル計
★★★ 棟梁レベル (full)	58	37	44	(139)
★★ 一人前レベル (Associate)	68	43	49	(160)
★ 見習いレベル (Assistant)	54	39	30	(123)
領域計	180	119	123	(422)

留意事項

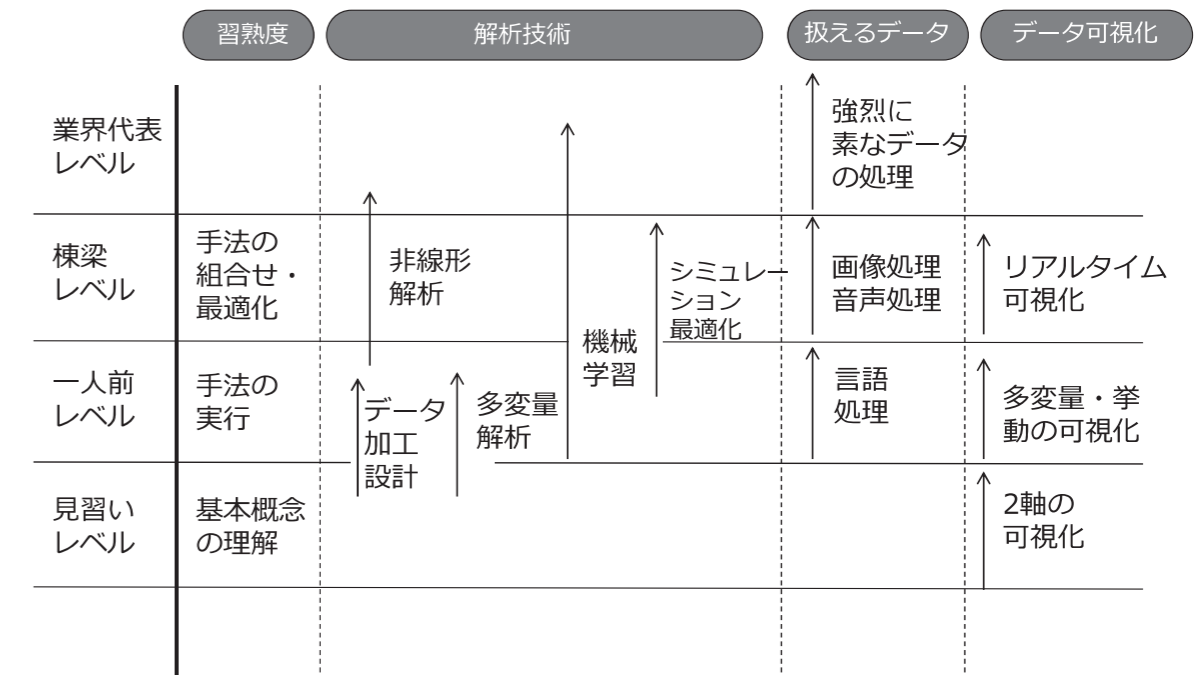
- データセキュリティはデータエンジニアリングとビジネスそれぞれの視点から扱っています
- 以下の項目は本チェックリストではカバーしていません
 - ✓ 特定のスキル項目ができるのと当然できると考えられるスキル
 - ✓ 特定の業界で重要であっても、一般性が低いスキル
 - ✓ 特定作業ができるために必要になるSAS、SPSSなど具体的ツールを使うスキル*
- 20名弱の実務家の協議による検討結果 (ver.1) であり、必要に応じ、適宜、見直していく位置づけです
- 個別の項目ができるできないの判断は、いまのところ各自のご判断で行っていただく前提です (後述)

* ただし具体名を入れないと理解が困難なケースでは固有名詞を入れています

領域の広がり (データサイエンス)



レベル感の目安 (データサイエンス) 2015.11.13 現在



資料：データサイエンティスト協会スキル委員会討議

領域の広がり (データエンジニアリング)



レベル感の目安 (データエンジニアリング)

2015.11.13 現在

	データ規模・タイプ	環境	その他
業界代表レベル			
棟梁レベル	数十億 (世界)	・クラウド ・アプライアンス	・リアルタイム処理 ・分散処理
一人前レベル	数千万 (国レベル) 非構造化データ	数台のサーバー	セキュリティ対策 Pig Hive
見習いレベル	数十万 (市レベル) 構造化データ	PC (Excel, BI ツール利用)	SQL

資料：データサイエンティスト協会スキル委員会討議

領域の広がり (ビジネス)



ビジネス課題解決	活動マネジメント 事業に実装する
データ課題解決	データ入手 データの理解・検証 意味合いの抽出・洞察 プロセス
基礎能力	論理的思考 行動規範

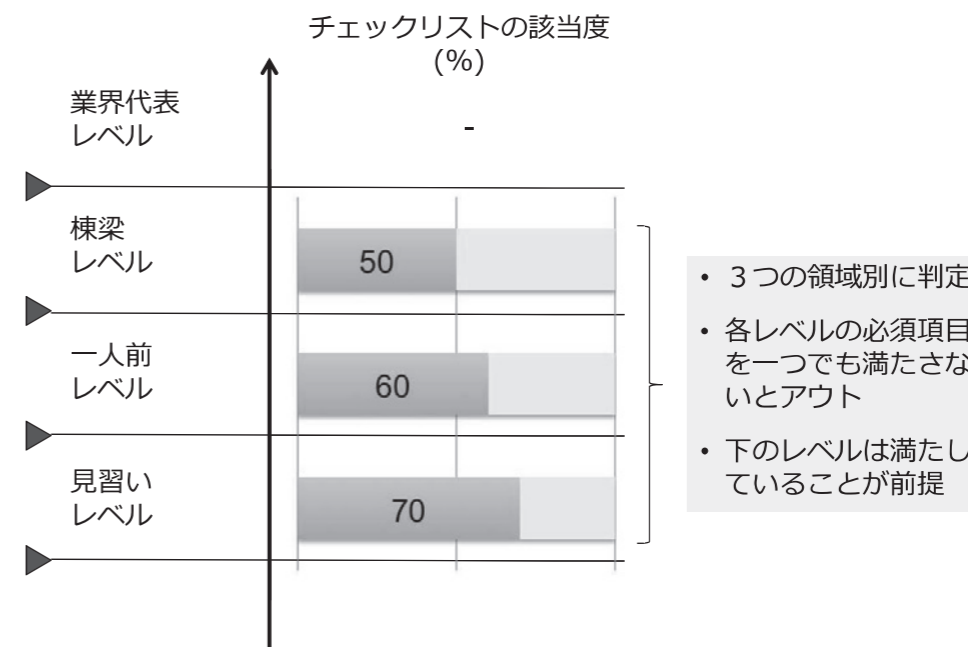
レベル感の目安 (ビジネス)

2015.11.13 現在

	対応力	問題解決	責任範囲	実装	育成
業界代表レベル			複合的体制		
棟梁レベル	新規の問題に対応	課題設定 新規データ解読	複数チーム 複数事業	横展開	育成システム
一人前レベル	既存の問題に対応	課題発見 既存データ解読	単一チーム 単一事業	事業実装	個別メンバー
見習いレベル	指示下で実行	データ確認	担当(タスク) 特定課題		

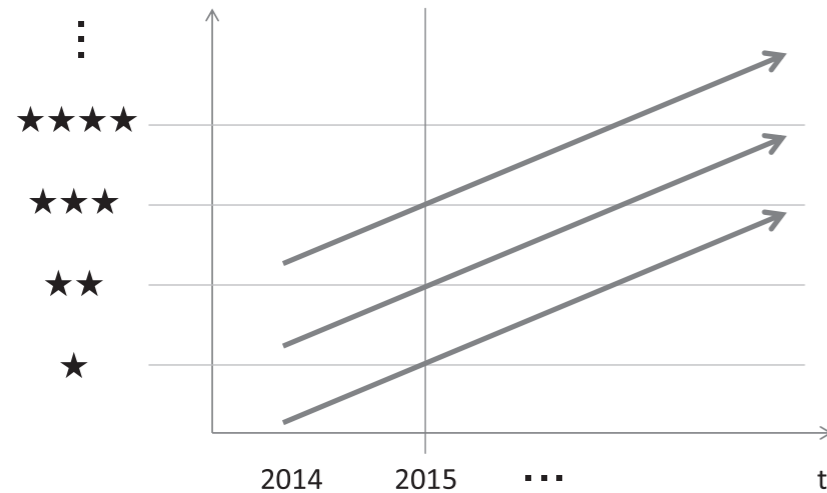
資料：データサイエンティスト協会スキル委員会討議

レベルの目安



スキルは本質的にmoving target

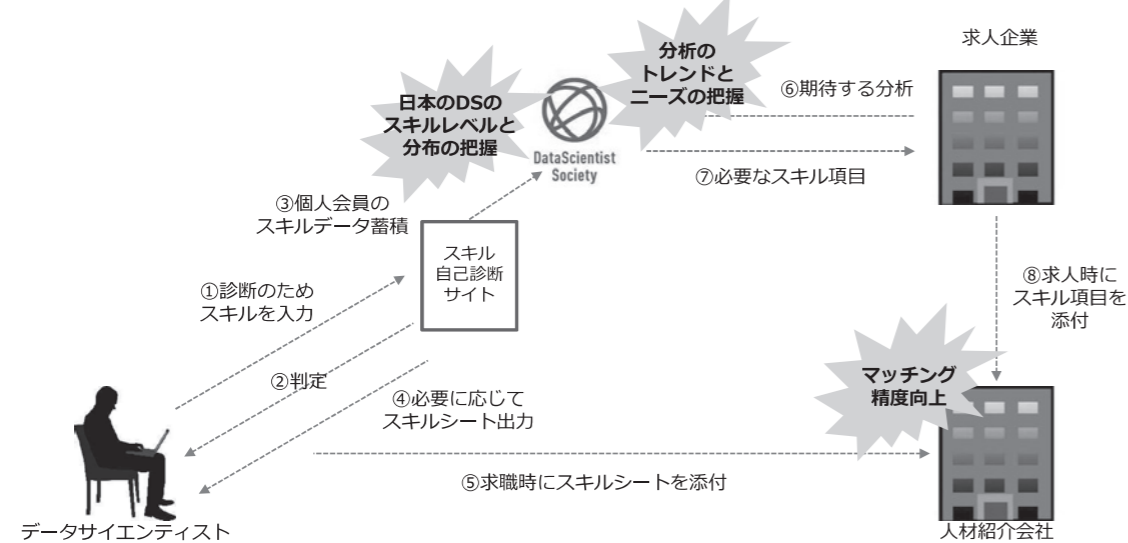
CONCEPTUAL



資料：データサイエンティスト協会スキル委員会討議

スキルチェックの仕組み (イメージ)

- 2016年度にデータサイエンティスト自身がスキルのチェックをサイト上で実施できる自己診断の仕組みを公開予定。
人材紹介業界でのスタンダードを目指し、データサイエンティストや企業間でのスキルに関する情報の標準化を通じて、ミスマッチの解消や取引の適正化を実現します。



内容

1. 昨年のrecap

- ① 本協会設立の背景と目的
- ② データサイエンティストとは何か？
- ③ 3つのスキルとスキルレベル

2. スキルチェックリスト

- ① 全体として (昨年 vs 今年)
- ② 個別領域の広がり
- ③ レベル感の考え方

3. 実行に向けて

- ① チームをどう作るか？
- ② 人をどう育てるか？
- ③ 目指す人はどう着手すべきか？

本日討議したいこと

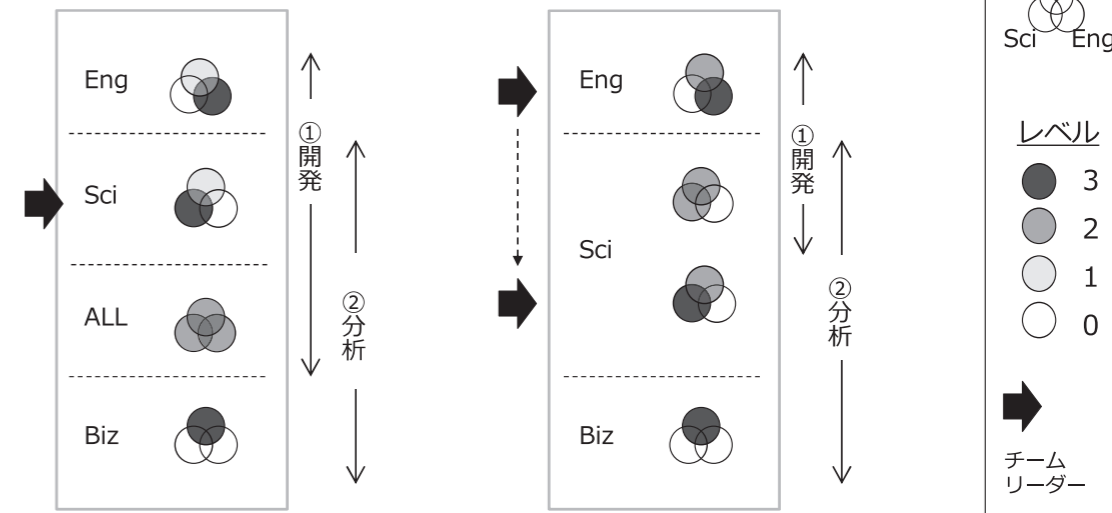
- 具体的にどう回すか？
- 具体的にどうやって育成していくべきか？
- 目指す人は何をどうしていったらいいのか？

実際のチームは どう組んだらいいのか？

チーム編成事例(2/3) : マーケティングソリューション

例3 : 行動データ解析

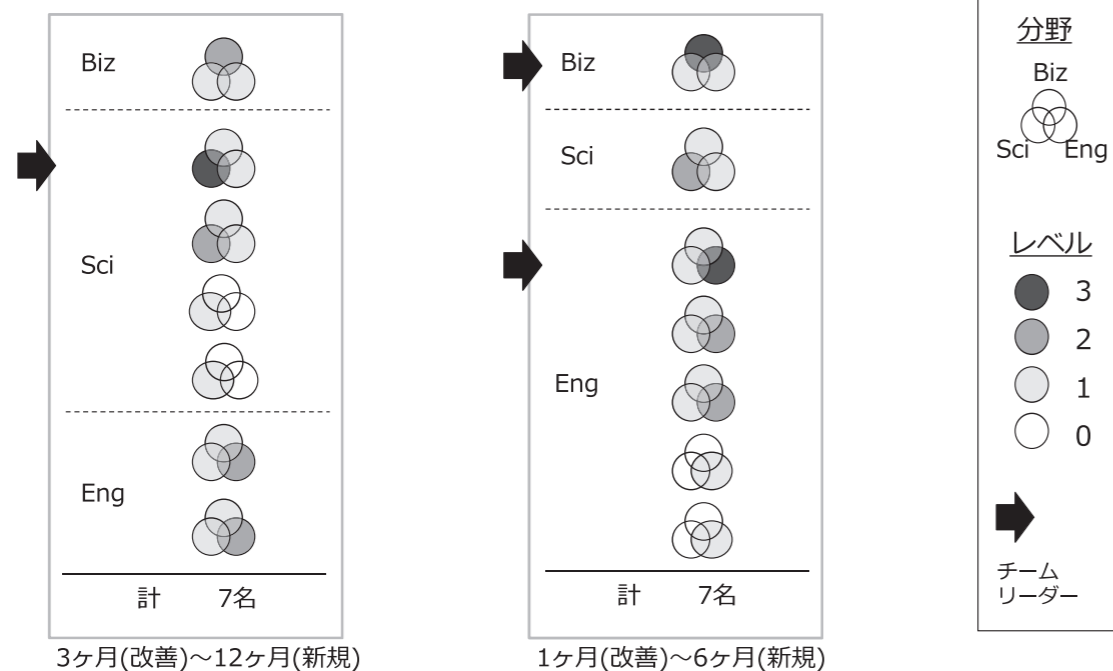
例4 : マーケティング投資最適化



チーム編成事例 (1/3) : ウェブサービス提供

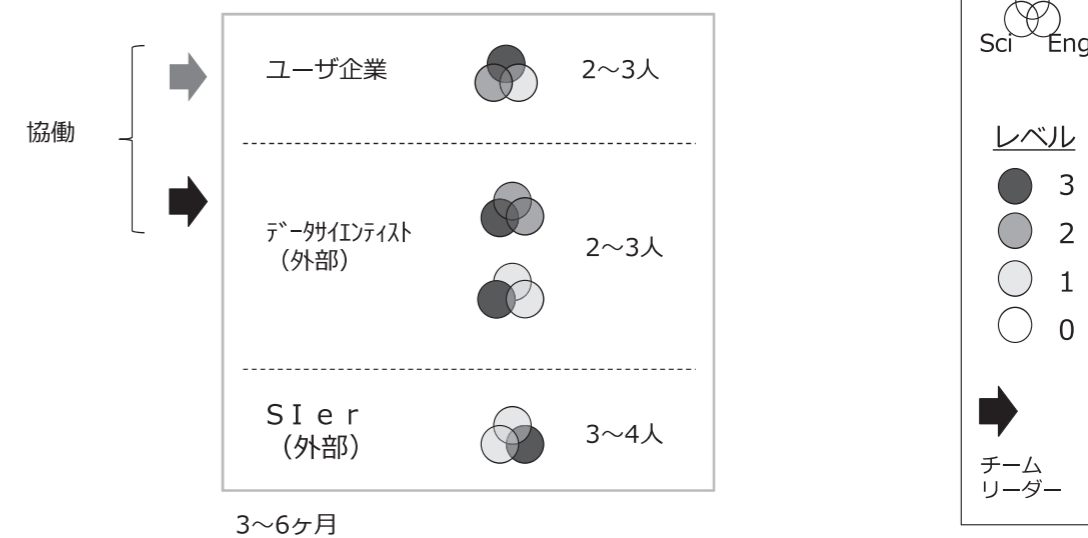
例1 : レコメンド (Sci寄り)

例2 : アクセス解析 (Eng寄り)



チーム編成事例(3/3) : ユーザ企業への実装

例5 : マーケティングオートメーション導入



Key takeaways (チーム作り)

- 単独の一人ではなく、チームとして必要なスキルセットを実現するのが現実解
- 課題解決の重さがどこにあるかで組成も変化 (サイエンスより、エンジニアよりなど)
- 開発、運用などのフェーズによっても異なる
- チームリーダーはコアになる課題解決と俯瞰できる力の有無でまる

資料：データサイエンティスト協会スキル委員会討議

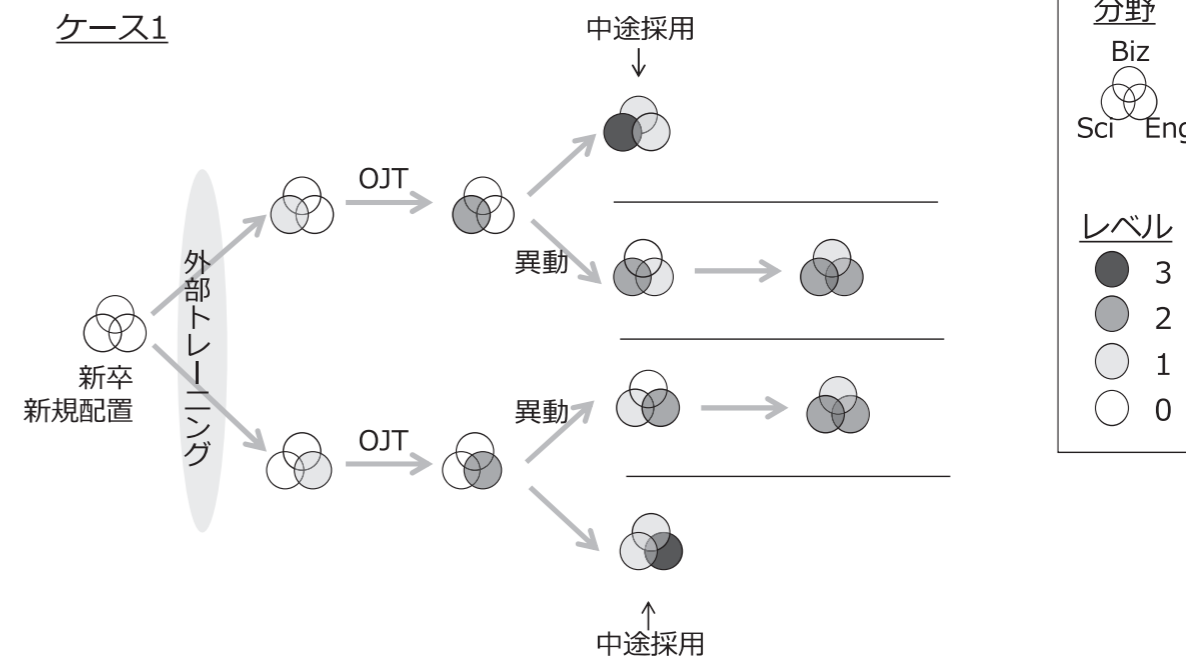
Discussion

- そもそもこのような体制を作れない会社はどうしたらいいのか？
- どういうヒトの組み合わせ、体制にしたらいいのかわからない会社はどうしたらいいのか？

ではスキル育成はどうしたらいいのか？

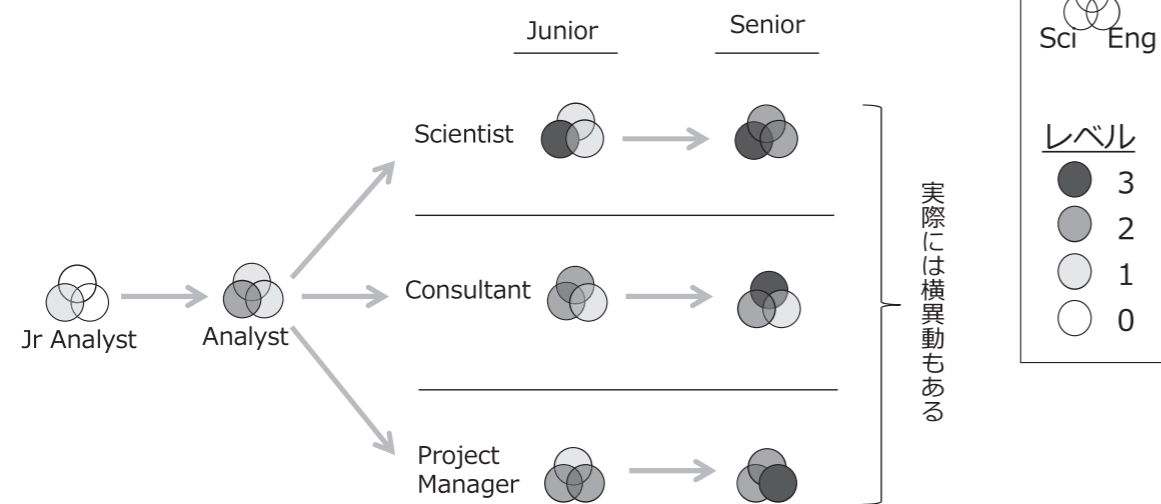
育成の実例(1/4)：マーケティングソリューション会社

ケース1



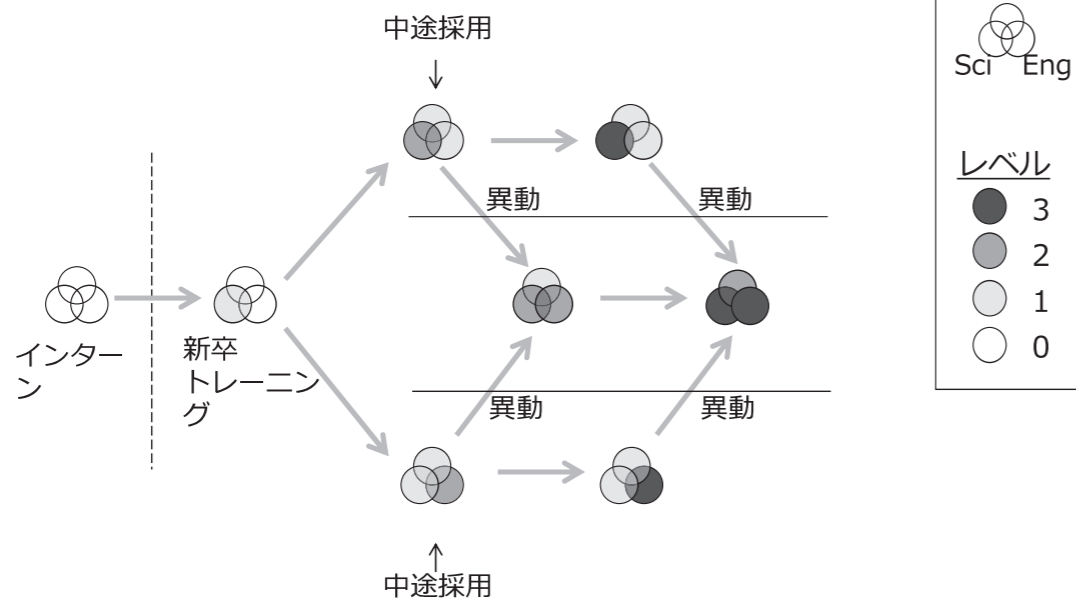
育成の実例(2/4)：データソリューション会社

ケース2



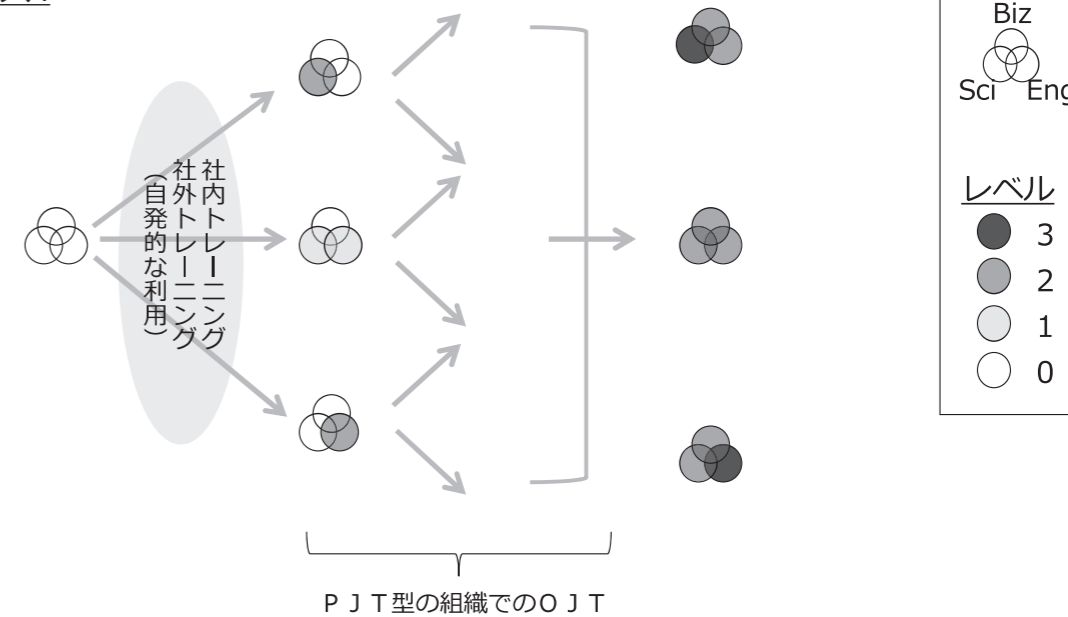
育成の実例(3/4)：分析ソリューション会社

ケース3



育成の実例(4/4)：データ分析ソリューション事業

ケース4



Key takeaways (育成)

- 見習いレベルまでは共通的なトレーニングで、そこから先は実践を通じてが基本
- サイエンス、エンジニア両方のスキルを同時というより、プロジェクトや部門異動によって育成していくのが現実解
- ビジネス力は複数の領域を超えて仕事を始め、全体を俯瞰してマネージする中で育成するケースが多い

資料：データサイエンティスト協会スキル委員会討議

Discussion

- 特定の領域の職人的な環境にいるひと（特にエンジニア上がりの人）はどうやってデータサイエンティストになるのか？
- より高度化していくスキルにどのように対応していくべきなのか？
- 異動を自分でコントロールしようがない時にどのように新しいスキルを身につけていくべきなのか？
- 高等教育機関（特に大学）はどう活用したら良いのか？

Mission

「データの持つ力を解き放つ」

付録 B：異分野異業種交流会 学生・若手研究者参加アンケート

- 1) 本日の交流会はいかがでしたか。感想を理由とともにご記入ください。
- 2) 本日のプログラムで印象に残った内容や話をご記入ください。
- 3) 本日の感想や、改善点、気づいた点がありましたらご記入ください。
- 4) 異分野異業種横断での交流会についてのご意見をお願いします。
(このような活動は継続した方が良い／継続しなくても良い／その他)
(上記の理由)
- 5) データサイエンティストという仕事に就くことを考えていますか？
(はい／まだわからない／いいえ)
- 6) 上記で「はい」の回答の方、データサイエンティストという仕事に就くためにしていることがあれば、ご記入ください。

付録 C : 提供されている関連講座・教育プログラムのリスト (平成 27 年 12 月 14 日現在)

注) 各講座・教育プログラムについて、本事業が推薦、あるいは支持するものではない。
より詳細な情報は、事業ホームページの教育プログラムのページ (http://datascientist.ism.ac.jp/trainingaids.html) を参照。

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	イカランツアア 類型との対応		使用言語	
										分析者向け	分析者向け		意思決定者向け
1	ITエンジニアのためのデータサイエンスエンジニア養成講座	アイティメディア株式会社 TEL : 03-6824-9393 (代表) 問い合わせフォーム有り 〒107-0052 東京都港区赤坂8-1-22 赤坂王子ビル	http://www.itmedia.co.jp/kyw/ords/ait_datascience.html http://www.atmarkt.co.jp/ait/kw/ait_datascience.html	○	○	x	x	-	無料	○	○	○	日本語
2	Advanced Methods in Data Science and Big Data Analytics	EMCジャパン株式会社 E-Mail : Japan_Education@emc.com Tel : 044-520-9830 Fax : 044-520-9839 〒212-0013 神奈川県川崎市幸区堀川町580-16 川崎テックセンター20階	http://japan.emc.com/collatera/support-training/training/customer-offerings/Advanced_Methods_in_DSBD.pdf	○	○	○	○	5日間	2,000ドル (最低価格)	○	○	○	日本語/英語
3	Data Science and Big Data Analytics	EMCジャパン株式会社 E-Mail : Japan_Education@emc.com Tel : 044-520-9830 Fax : 044-520-9839 〒212-0013 神奈川県川崎市幸区堀川町580-16 川崎テックセンター20階	http://japan.emc.com/collatera/support-training/training/customer-offerings/Data_Science_and_Big_Data_Analytics.pdf	○	○	○	○	5日間	2,000ドル (最低価格)	○	○	○	日本語
4	Data Science and Big Data Analytics for Business Transformation	EMCジャパン株式会社 E-Mail : Japan_Education@emc.com Tel : 044-520-9830 Fax : 044-520-9839 〒212-0013 神奈川県川崎市幸区堀川町580-16 川崎テックセンター20階	http://japan.emc.com/collatera/support-training/training/customer-offerings/1Day_Data_Science_and_Business_Transformation.pdf	○	○	○	○	1日間	2,000ドル (最低価格)	○	○	○	日本語
5	現代統計実務講座	一般財団法人 実務教育研究所 TEL : 03-3357-8153 FAX : 03-3358-7259 問い合わせフォーム有り 〒160-0022 東京都新宿区新宿1-1-12	http://www.jitsumu.or.jp/	x	○	x	○	標準受講期間 : 8カ月 (最長在籍期間 : 12カ月)	59,800円 (入学金 : 5,000円、受講料 : 54,800円)	○	○	○	日本語

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	イカランツアア 類型との対応		使用言語	
										分析者向け	分析者向け		意思決定者向け
6	多変量解析実務講座	一般財団法人 実務教育研究所 TEL : 03-3357-8153 FAX : 03-3358-7259 問い合わせフォーム有り 〒160-0022 東京都新宿区新宿1-1-12	http://www.jitsumu.or.jp/	x	○	○	○	標準受講期間 : 4カ月 (最長受講期間 : 8カ月)	54,500円 (入学金 : 5,000円、受講料 : 49,500円)	○	○	○	日本語
7	ウェブ解析士マスター認定講座	一般社団法人 ウェブ解析士協会 TEL : 03-6892-3182 問い合わせフォーム有り 〒160-0022 東京都新宿区新宿6-27-30 新宿イーストサイドスクエア17階	http://www.web-mining.jp/course/	x	○	○	○	5日間	324,000円 (税込/認定試験費用込) ※上記にはマクロレポート作成試験、ミクロレポート作成試験、講義実技試験3試験及び認定費用が含まれる。 【2日目】マクロレポート作成試験 (21,600円) 【3日目】ミクロレポート作成試験 (21,600円) 【4日目】講義実技試験 (21,600円)	○	○	○	日本語
8	上級Web解析士認定講座	一般社団法人 ウェブ解析士協会 TEL : 03-6892-3182 問い合わせフォーム有り 〒160-0022 東京都新宿区新宿6-27-30 新宿イーストサイドスクエア17階	http://www.web-mining.jp/course/	x	○	○	○	2日間	86,400円 (税込) (内 訳 : 受講・試験費用 75,600円、認定料 10,800円)	○	○	○	日本語
9	初級Web解析士認定講座	一般社団法人 ウェブ解析士協会 TEL : 03-6892-3182 問い合わせフォーム有り 〒160-0022 東京都新宿区新宿6-27-30 新宿イーストサイドスクエア17階	http://www.web-mining.jp/course/	x	x	x	x	2日間	【認定講座を受講して学習し受験する場合】 講座費用 10,800円 (税込) 試験費用 17,280円 (税込) (認定料込み) 公式テキスト 4,320円 (税込) 【公式教材で独自に学習し受験する場合】 試験費用 17,280円 (税込) (認定料込み) 公式テキスト 4,320円 (税込)	○	○	○	日本語

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	「カンファ」類型との対応		使用言語
										分析者向け	分析者向け	
10	今からでも遅くない基礎統計学講座 一般社団法人 日本マーケティング・リサーチ協会	一般社団法人 日本マーケティング・リサーチ協会 TEL : 03-3256-3101 FAX : 03-3256-3105 〒101-0044 東京都千代田区鍛冶町1-9-9 石川ビル2階	https://www.jmra-net.or.jp/seminar/detail.php?document_id=2355	x	o	x	5日間	JMRA 正・賛助会員社(者) : 72,360円(税込) 一般 : 87,480円(税込)	o	o	日本語	
11	ビッグデータ分析	インテル株式会社	https://www.intel.co.jp/content/www/jp/ja/big-data/analytcs-turning-big-data-into-intelligence.html?cid=sem69p69482g-c&gclid=CPG-wu27IsACFdd7vQodOisAMg#	o	x	x	-	無料	o	o	日本語	
12	情報を生かす、経営に活かす：ビッグデータ活用人材の育成	ウイングアーク1st株式会社	https://www.wingarc-support.com/training/cttechnic-al/bigdata	x	o	x	?	?	?	o	日本語	
13	ビッグデータ活用実践講座 (入門編)	ウイングアーク1st株式会社	https://www.wingarc-support.com/training/cttechnic-al/bigdata	?	?	?	?	?	?	?	?	
14	Data Science and Big Data Analytics (データ・サイエンスとビッグ・データ分析)	NECマネジメントパートナー株式会社	https://www.neclearning.jp/courseoutline/courseid/DB150/	x	o	o	5日間	356,400円(税込)	o	o	日本語	
15	OSSを活用したビッグデータ解析基礎構築体験	NECマネジメントパートナー株式会社	https://www.neclearning.jp/courseoutline/courseid/DB117/	x	x	x	1日間	43,200円(税込)	o	o	日本語	
16	Power BI入門 - Power Query, Power Pivot, Power View, Power Map -	NECマネジメントパートナー株式会社	https://www.neclearning.jp/courseoutline/courseid/B1039/	x	o	x	1日間	43,200円(税込)	o	o	日本語	

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	「カンファ」類型との対応		使用言語
										分析者向け	分析者向け	
17	R言語基礎	NECマネジメントパートナー株式会社	https://www.neclearning.jp/courseoutline/courseid/DB160/	x	o	o	1日間	43,200円(税込)	o	o	日本語	
18	R言語によるデータ分析演習 初級	NECマネジメントパートナー株式会社	https://www.neclearning.jp/courseoutline/courseid/DB282/	x	o	x	1日間	43,200円(税込)	o	o	日本語	
19	R言語プログラミング基礎1 (eトラーニング)	NECマネジメントパートナー株式会社	https://www.neclearning.jp/courseoutline/courseid/DB193/	o	x	x	1日間	21,600円(税込)	o	o	日本語	
20	R言語プログラミング基礎2 (eトラーニング)	NECマネジメントパートナー株式会社	https://www.neclearning.jp/courseoutline/courseid/DB195/	o	x	o	1日間	21,600円(税込)	o	o	日本語	
21	R言語入門	NECマネジメントパートナー株式会社	https://www.neclearning.jp/courseoutline/courseid/DB152/	x	o	x	1日間	43,200円(税込)	o	o	日本語	
22	データウェアハウス/ビジネスインテリジェンス入門	NECマネジメントパートナー株式会社	https://www.neclearning.jp/courseoutline/courseid/B1001/	x	x	o	1日間	37,800円(税込)	o	o	日本語	
23	データマイニング基礎 (eトラーニング)	NECマネジメントパートナー株式会社	https://www.neclearning.jp/courseoutline/courseid/DB230/	o	o	x	1日間	21,600円(税込)	o	o	日本語	

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	「カンファ」類型との対応	
										分析者向け	分析者向け
24	データ活用研修-データマイニングプロジェクトの推進と実践を学ぶ-	NECマシメントナー株式会社 TEL: 03-5232-3075 E-mail: dbg@educ.jp.nec.com (本社) 〒211-8601 神奈川県川崎市中原区下沼部1753	NECマシメントナー株式会社 TEL: 03-5232-3075 E-mail: dbg@educ.jp.nec.com (本社) 〒211-8601 神奈川県川崎市中原区下沼部1753	https://www.neclearning.jp/courseoutline/courseid/DB245/	x	o	x	2日間	108,000円 (税込)	o	o
25	データ分析のためのPython入門	NECマシメントナー株式会社 TEL: 03-5232-3075 E-mail: dbg@educ.jp.nec.com (本社) 〒211-8601 神奈川県川崎市中原区下沼部1753	NECマシメントナー株式会社 TEL: 03-5232-3075 E-mail: dbg@educ.jp.nec.com (本社) 〒211-8601 神奈川県川崎市中原区下沼部1753	https://www.neclearning.jp/courseoutline/courseid/DB192/	x	o	o	1日間	43,200円 (税込)	o	o
26	ビッグデータ概説	NECマシメントナー株式会社 TEL: 03-5232-3075 E-mail: dbg@educ.jp.nec.com (本社) 〒211-8601 神奈川県川崎市中原区下沼部1753	NECマシメントナー株式会社 TEL: 03-5232-3075 E-mail: dbg@educ.jp.nec.com (本社) 〒211-8601 神奈川県川崎市中原区下沼部1753	https://www.neclearning.jp/courseoutline/courseid/DB093/	x	x	x	1日間	43,200円 (税込)	o	o
27	ビッグデータ入門 (eトレニング)	NECマシメントナー株式会社 TEL: 03-5232-3075 E-mail: dbg@educ.jp.nec.com (本社) 〒211-8601 神奈川県川崎市中原区下沼部1753	NECマシメントナー株式会社 TEL: 03-5232-3075 E-mail: dbg@educ.jp.nec.com (本社) 〒211-8601 神奈川県川崎市中原区下沼部1753	https://www.neclearning.jp/courseoutline/courseid/DB201/	o	x	x	1日間	10,800円 (税込)	o	o
28	ビッグデータの分析と活用~データマイニング基礎編~	NECマシメントナー株式会社 TEL: 03-5232-3075 E-mail: dbg@educ.jp.nec.com (本社) 〒211-8601 神奈川県川崎市中原区下沼部1753	NECマシメントナー株式会社 TEL: 03-5232-3075 E-mail: dbg@educ.jp.nec.com (本社) 〒211-8601 神奈川県川崎市中原区下沼部1753	https://www.neclearning.jp/courseoutline/courseid/DB095/	x	o	x	1日間	48,600円 (税込)	o	o
29	ビッグデータの分析と活用~統計解析手法によるデータ分析入門~	NECマシメントナー株式会社 TEL: 03-5232-3075 E-mail: dbg@educ.jp.nec.com (本社) 〒211-8601 神奈川県川崎市中原区下沼部1753	NECマシメントナー株式会社 TEL: 03-5232-3075 E-mail: dbg@educ.jp.nec.com (本社) 〒211-8601 神奈川県川崎市中原区下沼部1753	https://www.neclearning.jp/courseoutline/courseid/DB101/	x	o	x	2日間	86,400円 (税込)	o	o
30	機械学習 概要 ~R言語で始める機械学習~	NECマシメントナー株式会社 TEL: 03-5232-3075 E-mail: dbg@educ.jp.nec.com (本社) 〒211-8601 神奈川県川崎市中原区下沼部1753	NECマシメントナー株式会社 TEL: 03-5232-3075 E-mail: dbg@educ.jp.nec.com (本社) 〒211-8601 神奈川県川崎市中原区下沼部1753	https://www.neclearning.jp/courseoutline/courseid/DB281/	x	o	o	1日間	43,200円 (税込)	o	o

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	「カンファ」類型との対応	
										分析者向け	分析者向け
31	enPIT-BizApp	enPIT (文部科学省「情報技術人材育成のための実践教育ネットワーク形成事業」分野・地域を越えた実践的情報教育協働ネットワーク)	大阪大学大学院情報科学研究科 enPIT事務局 お問い合わせフォーム有り	http://www.enpit.jp/fields/bizapp.html	o	o	☆	約1年間	無料	o	o
32	SPSS講座: データサイエンス入門	株式会社アイ・ラーニング TEL: 0120-623-629 お問い合わせフォーム有り 〒103-0015 東京都中央区日本橋箱崎町4-3 国際箱崎ビル	株式会社アイ・ラーニング TEL: 0120-623-629 お問い合わせフォーム有り 〒103-0015 東京都中央区日本橋箱崎町4-3 国際箱崎ビル	https://www.i-learning.jp/products/detail.php?course_code=0DA18	x	o	o	1日間	64,800円	o	o
33	SPSS講座: データサイエンス入門	株式会社アイ・ラーニング TEL: 0120-623-629 お問い合わせフォーム有り 〒103-0015 東京都中央区日本橋箱崎町4-3 国際箱崎ビル	株式会社アイ・ラーニング TEL: 0120-623-629 お問い合わせフォーム有り 〒103-0015 東京都中央区日本橋箱崎町4-3 国際箱崎ビル	https://www.i-learning.jp/products/detail.php?course_code=0D0P2	x	o	o	2日間	129,600円	o	o
34	SPSS講座: データサイエンス入門	株式会社アイ・ラーニング TEL: 0120-623-629 お問い合わせフォーム有り 〒103-0015 東京都中央区日本橋箱崎町4-3 国際箱崎ビル	株式会社アイ・ラーニング TEL: 0120-623-629 お問い合わせフォーム有り 〒103-0015 東京都中央区日本橋箱崎町4-3 国際箱崎ビル	https://www.i-learning.jp/products/detail.php?course_code=0D0Q2	x	o	o	1日間	64,800円	o	o
35	SPSS講座: データサイエンス入門	株式会社アイ・ラーニング TEL: 0120-623-629 お問い合わせフォーム有り 〒103-0015 東京都中央区日本橋箱崎町4-3 国際箱崎ビル	株式会社アイ・ラーニング TEL: 0120-623-629 お問い合わせフォーム有り 〒103-0015 東京都中央区日本橋箱崎町4-3 国際箱崎ビル	https://www.i-learning.jp/products/detail.php?course_code=0D0Q2	x	o	o	1日間	64,800円	o	o
36	SPSS講座: データサイエンス入門	株式会社アイ・ラーニング TEL: 0120-623-629 お問い合わせフォーム有り 〒103-0015 東京都中央区日本橋箱崎町4-3 国際箱崎ビル	株式会社アイ・ラーニング TEL: 0120-623-629 お問い合わせフォーム有り 〒103-0015 東京都中央区日本橋箱崎町4-3 国際箱崎ビル	https://www.i-learning.jp/products/detail.php?course_code=0DA19	x	o	o	1日間	64,800円	o	o
37	SPSS講座: データサイエンス入門	株式会社アイ・ラーニング TEL: 0120-623-629 お問い合わせフォーム有り 〒103-0015 東京都中央区日本橋箱崎町4-3 国際箱崎ビル	株式会社アイ・ラーニング TEL: 0120-623-629 お問い合わせフォーム有り 〒103-0015 東京都中央区日本橋箱崎町4-3 国際箱崎ビル	https://www.i-learning.jp/products/detail.php?course_code=0D0P2	x	o	o	2日間	129,600円	o	o

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	インテグリア 類型との対応		使用言語
										分析者向け	分析者向け	
38	SPSS講座：データサイエンティストを目指す方へ コースC：組合せ・順列/ターンの発見 IBM SPSS Modeler 観密分析【クラスターリング&アソシエーションモデル編】	株式会社アイ・ラーニング TEL：0120-623-629 問い合わせ先 〒103-0015 東京都中央区日本橋箱崎町4-3 国際箱崎ビル	https://www.i-learning.jp/products/detail.php?course_code=0D0R2	x	o	o	1日間	64,800円	o		日本語	
	SPSS講座：データサイエンティストを目指す方へ コースC：組合せ・順列/ターンの発見 IBM SPSS Modeler入門2 データ加工	株式会社アイ・ラーニング TEL：0120-623-629 問い合わせ先 〒103-0015 東京都中央区日本橋箱崎町4-3 国際箱崎ビル	https://www.i-learning.jp/products/detail.php?course_code=0D0Q2	x	o	o	1日間	64,800円	o		日本語	
40	SPSS講座：データサイエンティストを目指す方へ コースC：組合せ・順列/ターンの発見 IBM SPSS Modeler入門1	株式会社アイ・ラーニング TEL：0120-623-629 問い合わせ先 〒103-0015 東京都中央区日本橋箱崎町4-3 国際箱崎ビル	https://www.i-learning.jp/products/detail.php?course_code=0D0P2	x	o	o	2日間	129,600円	o		日本語	
41	データサイエンティスト入門	株式会社アイ・ラーニング TEL：0120-623-629 問い合わせ先 〒103-0015 東京都中央区日本橋箱崎町4-3 国際箱崎ビル	https://www.i-learning.jp/products/detail.php?course_code=DW800	x	x	x	2日間	86,400円		o	日本語	
42	データ分析の基礎(EXCEL実践編)	株式会社アイ・ラーニング TEL：0120-623-629 問い合わせ先 〒103-0015 東京都中央区日本橋箱崎町4-3 国際箱崎ビル	https://www.i-learning.jp/products/detail.php?course_code=DW801	x	x	x	1日間	43,200円 (税込)	o		日本語	
43	データ分析の基礎 (SPSS実践編)	株式会社アイ・ラーニング TEL：0120-623-629 問い合わせ先 〒103-0015 東京都中央区日本橋箱崎町4-3 国際箱崎ビル	https://www.i-learning.jp/products/detail.php?course_code=DW802	x	x	x	1日間	54,000円 (税込)	o	o	日本語	
44	データサイエンス基礎講座2015秋【機械学習実践編】	株式会社インプレス インプレス イベント事務局 E-mail: bc-event@impress.co.jp TEL: 03-6837-9820 〒102-0075 東京都千代田区三番町 20 番地	http://www.impressbm.co.jp/event/datascientist2014/	x	o	x	4日間	54,000円 (税込)	o		日本語	

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	インテグリア 類型との対応		使用言語
										分析者向け	分析者向け	
45	データサイエンティスト養成コース (バージョン)	株式会社チエンジ TEL：03-6303-1250 FAX：03-6303-1251 E-mail: info@change-jp.com 〒150-0011 東京都渋谷区東1-26-20 東京建物東渋谷ビル11階	http://www.change-jp.com/bigdata/	x	o	o	5日間	324,000円 (税込)	o	o	日本語	
46	ビッグデータ活用 - ビッグデータ無料勉強会	株式会社チエンジ TEL：03-6303-1250 FAX：03-6303-1251 E-mail: info@change-jp.com 〒150-0011 東京都渋谷区東1-26-20 東京建物東渋谷ビル11階	http://www.change-jp.com/bigdata/big_trial.html	x	x	x	1日間	無料		o	o	日本語
47	社会人のためのデータサイエンス入門	株式会社ドコモgacco 問い合わせ先 〒106-0047 東京都港区南麻布一丁目6番15号 アーバンネット麻布ビル	https://jms.gacco.org/courses/course-v1:gacco+ga031+2015_11/ab	o	o	x	4週間	補助教材：1,000円	o	o	o	日本語
48	統計学 I：データ分析の基礎	株式会社ドコモgacco 問い合わせ先 〒106-0047 東京都港区南麻布一丁目6番15号 アーバンネット麻布ビル	https://jms.gacco.org/courses/re_gacco/ga014/2015_05/ab	o	o	x	5週間	補助教材：1,000円 反転学習コースは有料 9,000円 (ただし補助教材代金含む)	o	o	o	日本語
49	統計学 II：推測統計の方法	株式会社ドコモgacco 問い合わせ先 〒106-0047 東京都港区南麻布一丁目6番15号 アーバンネット麻布ビル	https://jms.gacco.org/courses/course-v1:gacco+ga047+2015_10/ab	o	o	x	5週間	補助教材：1,000円 反転学習コースは有料 9,000円 (ただし補助教材代金含む)	o		日本語	
50	[マシン学習付き] <eラーニング> Hadoopプログラミング	株式会社日立インフォメーションアカデミー 問い合わせ先 〒140-0013 東京都品川区南大井6丁目26番3号 大森ベルポートD館13階	https://www.hitachia.co.jp/courses/view/DBE317/3	o	o	o	21日間	32,400円 (税込)	o		日本語	
51	<eラーニング> Exadata and Database Machine 管理ワークショップ	株式会社日立インフォメーションアカデミー 問い合わせ先 〒140-0013 東京都品川区南大井6丁目26番3号 大森ベルポートD館13階	https://www.hitachia.co.jp/courses/view/DBE312/3	o	o	o	21日間	385,560円 (税込)	o		日本語	

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	インテグリア 類型との対応	
										分析者 向け	分析者 提供 者向け
52	<eラーニング>>Oracle Database 11g R2 RAC管理	株式会社日立インフォメーションアカデミー	株式会社日立インフォメーションアカデミー 〒140-0013 東京都品川区南大井6丁目26番3号 大森ベルポートD館13階	https://www.hitachi-ia.co.jp/courses/view/DBE313/	○	○	○	21時間	289,008円 (税込)	○	日本語
53	<eラーニング>>Oracle Grid Infrastructure 11g R2 クラスタ&ASM管理	株式会社日立インフォメーションアカデミー	株式会社日立インフォメーションアカデミー 〒140-0013 東京都品川区南大井6丁目26番3号 大森ベルポートD館13階	https://www.hitachi-ia.co.jp/courses/view/DBE314/	○	○	○	28時間	385,344円 (税込)	○	日本語
54	<eラーニング>>Oracleではじめる統計入門	株式会社日立インフォメーションアカデミー	株式会社日立インフォメーションアカデミー 〒140-0013 東京都品川区南大井6丁目26番3号 大森ベルポートD館13階	https://www.hitachi-ia.co.jp/courses/view/DBE019/	x	○	○	1日間	77,112円 (税込)	○	日本語
55	<eラーニング>>データウェアハウスのためのデータモデリング解説	株式会社日立インフォメーションアカデミー	株式会社日立インフォメーションアカデミー 〒140-0013 東京都品川区南大井6丁目26番3号 大森ベルポートD館13階	https://www.hitachi-ia.co.jp/courses/view/DBE013/	○	○	○	6時間	受講料 18,360円 (税込) ヘルプデスク 5,400円 (税込)	○	日本語
56	<eラーニング>>データウェアハウスのためのデータモデリング解説	株式会社日立インフォメーションアカデミー	株式会社日立インフォメーションアカデミー 〒140-0013 東京都品川区南大井6丁目26番3号 大森ベルポートD館13階	https://www.hitachi-ia.co.jp/courses/view/DBE012/	○	○	○	6時間	受講料 18,360円 (税込) ヘルプデスク 5,400円 (税込)	○	日本語
57	<eラーニング>>データウェアハウスのためのデータモデリング解説	株式会社日立インフォメーションアカデミー	株式会社日立インフォメーションアカデミー 〒140-0013 東京都品川区南大井6丁目26番3号 大森ベルポートD館13階	https://www.hitachi-ia.co.jp/courses/view/SJE003/	○	○	x	8時間	受講料 11,016円 (税込) ヘルプデスク 5,400円 (税込)	○	日本語
58	Hadoop入門ーインストールと操作	株式会社日立インフォメーションアカデミー	株式会社日立インフォメーションアカデミー 〒140-0013 東京都品川区南大井6丁目26番3号 大森ベルポートD館13階	https://www.hitachi-ia.co.jp/courses/view/CLJ002/	x	○	○	1日間	37,800円 (税込)	○	日本語

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	インテグリア 類型との対応	
										分析者 向け	分析者 提供 者向け
59	Hitachi Advanced Data Binder プラットフォーム 1st Step	株式会社日立インフォメーションアカデミー	株式会社日立インフォメーションアカデミー 〒140-0013 東京都品川区南大井6丁目26番3号 大森ベルポートD館13階	https://www.hitachi-ia.co.jp/courses/view/DBJ091/	x	x	○	0.5日間	21,600円 (税込)	○	日本語
60	Hitachi Advanced Data Binder プラットフォーム 2nd Step	株式会社日立インフォメーションアカデミー	株式会社日立インフォメーションアカデミー 〒140-0013 東京都品川区南大井6丁目26番3号 大森ベルポートD館13階	https://www.hitachi-ia.co.jp/courses/view/DBJ098/	x	x	○	2日間	86,400円 (税込)	○	日本語
61	NoSQL入門ービッグデータ時代のデータベース	株式会社日立インフォメーションアカデミー	株式会社日立インフォメーションアカデミー 〒140-0013 東京都品川区南大井6丁目26番3号 大森ベルポートD館13階	https://www.hitachi-ia.co.jp/courses/view/DBJ089/	x	○	○	1日間	32,400円 (税込)	○	日本語
62	R入門ーRによるデータ分析	株式会社日立インフォメーションアカデミー	株式会社日立インフォメーションアカデミー 〒140-0013 東京都品川区南大井6丁目26番3号 大森ベルポートD館13階	https://www.hitachi-ia.co.jp/courses/view/DBJ099/	x	○	○	1日間	37,800円 (税込)	○	日本語
63	ケースメソッドで学ぶデータ分析ソリューションーデータ分析の基礎を学ぶ	株式会社日立インフォメーションアカデミー	株式会社日立インフォメーションアカデミー 〒140-0013 東京都品川区南大井6丁目26番3号 大森ベルポートD館13階	https://www.hitachi-ia.co.jp/courses/view/HSJ140/	x	○	○	2日間	81,000円 (税込)	○	日本語
64	データサイエンティスト入門ー統計・分析の基礎を学ぶ	株式会社日立インフォメーションアカデミー	株式会社日立インフォメーションアカデミー 〒140-0013 東京都品川区南大井6丁目26番3号 大森ベルポートD館13階	https://www.hitachi-ia.co.jp/courses/view/DBD046	x	○	○	2.5日間	216,000円 (税込)	○	日本語
65	データサイエンティスト入門ー統計・分析の基礎を学ぶ	株式会社日立インフォメーションアカデミー	株式会社日立インフォメーションアカデミー 〒140-0013 東京都品川区南大井6丁目26番3号 大森ベルポートD館13階	https://www.hitachi-ia.co.jp/courses/view/DBD047/	x	○	○	1.5日間	129,600円 (税込)	○	日本語

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	イカンファ 類型との対応		使用言語
										分析者 向け	分析者 ビズ提供 者向け	
66	データサイエンティスト入門研修 機械学習による問題解決実践 - 機 械学習で大量データを分析し、 データからビジネス価値を創出 -	株式会社日立インフォメーション アカデミー 問い合わせ先フォーム有り 〒140-0013 東京都品川区南大 井6丁目26番3号 大森ベルポート D館13階	https://www.hitachi- ia.co.jp/courses/view/DBD053/ /3	x	o	x	2日間	216,000円 (税込)	o	o	日本語	
	データベース入門 - 解説と操作体 験 -	株式会社日立インフォメーション アカデミー 問い合わせ先フォーム有り 〒140-0013 東京都品川区南大 井6丁目26番3号 大森ベルポート D館13階	https://www.hitachi- ia.co.jp/courses/view/DB1045/ 3	x	o	x	1日間	32,400円 (税込)	o		日本語	
68	データベース入門 - 解説と操作体 験 -	株式会社日立インフォメーション アカデミー 問い合わせ先フォーム有り 〒140-0013 東京都品川区南大 井6丁目26番3号 大森ベルポート D館13階	https://www.hitachi- ia.co.jp/courses/view/DB1045/ 3	x	o	x	1日間	32,400円 (税込)	o		日本語	
	データ分析に使えるSQL - SQLで ビッグデータに立ち向かう -	株式会社日立インフォメーション アカデミー 問い合わせ先フォーム有り 〒140-0013 東京都品川区南大 井6丁目26番3号 大森ベルポート D館13階	https://www.hitachi- ia.co.jp/courses/view/DB1090/ 3	x	o	o	1日間	43,200円 (税込)	o		日本語	
70	データ分析手法の理論と適用 - ビ ジネスにおける統計的手法活用の 広がり -	株式会社日立インフォメーション アカデミー 問い合わせ先フォーム有り 〒140-0013 東京都品川区南大 井6丁目26番3号 大森ベルポート D館13階	https://www.hitachi- ia.co.jp/courses/view/HSJ1116/ 3	x	o	o	2日間	81,000円 (税込)	o		日本語	
	ビッグデータ概説	株式会社日立インフォメーション アカデミー 問い合わせ先フォーム有り 〒140-0013 東京都品川区南大 井6丁目26番3号 大森ベルポート D館13階	https://www.hitachi- ia.co.jp/courses/view/DB1081/ 3	x	x	x	0.5日間	21,600円 (税込)	o	o	日本語	
72	ロジカルシンキング基礎 - 論理の 可視化と論理ツエックのポイント -	株式会社日立インフォメーション アカデミー 問い合わせ先フォーム有り 〒140-0013 東京都品川区南大 井6丁目26番3号 大森ベルポート D館13階	https://www.hitachi- ia.co.jp/courses/view/HSJ1116/ 3	x	o	x	2日間	81,000円 (税込)	o	o	日本語	

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	イカンファ 類型との対応		使用言語
										分析者 向け	分析者 ビズ提供 者向け	
73	仮説構築と調査検証 - 仮説構築の 考え方を学び、仮説に沿った調査 手法を選択し、調査結果を読み解 く -	株式会社日立インフォメーション アカデミー 問い合わせ先フォーム有り 〒140-0013 東京都品川区南大 井6丁目26番3号 大森ベルポート D館13階	https://www.hitachi- ia.co.jp/courses/view/HSJ112/ 3	x	o	o	2日間	81,000円 (税込)	o	o	日本語	
	基礎から学ぶSQL - 現場で使える 力をつける -	株式会社日立インフォメーション アカデミー 問い合わせ先フォーム有り 〒140-0013 東京都品川区南大 井6丁目26番3号 大森ベルポート D館13階	https://www.hitachi- ia.co.jp/courses/view/DB1063/ 3	x	o	o	2日間	64,800円	o		日本語	
75	体感！機械学習 - ビッグデータ時 代のシステムを支える技術 -	株式会社日立インフォメーション アカデミー 問い合わせ先フォーム有り 〒140-0013 東京都品川区南大 井6丁目26番3号 大森ベルポート D館13階	https://www.hitachi- ia.co.jp/courses/view/DB1088/ 3	x	o	o	1日間	43,200円 (税込)	o		日本語	
	脱初心者のためのSQL - SQLでこ こまでできる -	株式会社日立インフォメーション アカデミー 問い合わせ先フォーム有り 〒140-0013 東京都品川区南大 井6丁目26番3号 大森ベルポート D館13階	https://www.hitachi- ia.co.jp/courses/view/DB1067/ 3	x	o	o	1日間	43,200円 (税込)	o		日本語	
77	定量的分析のスキル - 問題解決の精 度を高める -	株式会社日立インフォメーション アカデミー 問い合わせ先フォーム有り 〒140-0013 東京都品川区南大 井6丁目26番3号 大森ベルポート D館13階	https://www.hitachi- ia.co.jp/courses/view/HSJ108/ 3	x	o	o	2日間	81,000円 (税込)	o	o	日本語	
	RIによる統計解析	株式会社ブレインパッド 教育講 座事務局 TEL: 03-6721-7001 FAX: 03- 6721-7010 E-mail: edu_info@brainpad.co.jp 〒108-0071 東京都港区白金3- 2-10 白金台ビル	http://school.brainpad.co.jp/pr ogram/r?ss_ad_code=	x	o	x	3日間	216,000円 (税込)	o	o	日本語	

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	インターンシップ 類型との対応		
										分析者 向け	分析者 向け	分析者 向け
79	SQLによる集計・分析	株式会社ブレインパッド	株式会社ブレインパッド 教育講座 TEL : 03-6721-7001 FAX : 03-6721-7010 E-mail : edu_info@brainpad.co.jp 〒108-0071 東京都港区白金台3-2-10 白金台ビル	http://school.brainpad.co.jp/pr ogram/sql?ss_ad_code=	x	o	x	2日間	129,600円 (税込)	o	o	日本語
80	WPS Softwareプログラミング	株式会社ブレインパッド	株式会社ブレインパッド 教育講座 TEL : 03-6721-7001 FAX : 03-6721-7010 E-mail : edu_info@brainpad.co.jp 〒108-0071 東京都港区白金台3-2-10 白金台ビル	http://school.brainpad.co.jp/pr ogram/wps?ss_ad_code=	x	o	o	2日間	151,200円 (税込)	o	o	日本語
81	機械学習による問題解決実践	株式会社ブレインパッド	株式会社ブレインパッド 教育講座 TEL : 03-6721-7001 FAX : 03-6721-7010 E-mail : edu_info@brainpad.co.jp 〒108-0071 東京都港区白金台3-2-10 白金台ビル	http://school.brainpad.co.jp/pr ogram/mining?ss_ad_code=	x	o	x	2日間	216,000円 (税込)	o	o	日本語
82	統計・分析のためのデータ加工入門	株式会社ブレインパッド	株式会社ブレインパッド 教育講座 TEL : 03-6721-7001 FAX : 03-6721-7010 E-mail : edu_info@brainpad.co.jp 〒108-0071 東京都港区白金台3-2-10 白金台ビル	http://school.brainpad.co.jp/pr ogram/data?ss_ad_code=	x	o	x	1日間	75,600円 (税込)	o	o	日本語
83	データ活用の技術 [データ解析編]	産業能率大学	産業能率大学総合研究所 TEL : 03-5758-5110 FAX : 03-5758-5503 問い合わせフォーム有り 〒158-8630 東京都世田谷区等々力6-39-15 自由が丘キャンパス 1号館内	http://seminar.hj.sanno.ac.jp/s /3602/	x	o	x	1日間	38,880円 (税込)	o	o	日本語

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	インターンシップ 類型との対応		
										分析者 向け	分析者 向け	分析者 向け
84	ビジネス定量的分析講座	グロービス・マネジメント・スクール	グロービス・マネジメント・スクール 【東京校】 TEL : 03-5275-3806 FAX : 03-5275-3787 E-mail : tokyo@globis.co.jp 〒102-0084 東京都千代田区二番町5-1住友不動産麹町ビル【大阪校】 TEL : 06-6391-0201 FAX : 06-6391-0218 E-mail : osaka@globis.co.jp 〒532-0003 大阪府大阪市淀川区宮原1-1 新大阪阪急ビル4F【名古屋校】 TEL : 052-533-3790 FAX : 052-533-3782 E-mail : nagoya@globis.co.jp 〒450-0003 愛知県名古屋市中村区名駅南1-24-20 名古屋三井ビルディング新館4F	http://gms.globis.co.jp/curriculum/core/qab/outline.html	x	o	x	全6回 (隔週で3ヶ月間)	126,000円	o	o	日本語
85	データ活用の技術 [読み方・つくり方編]	産業能率大学	産業能率大学総合研究所 TEL : 03-5758-5110 FAX : 03-5758-5503 問い合わせフォーム有り 〒158-8630 東京都世田谷区等々力6-39-15 自由が丘キャンパス 1号館内	http://seminar.hj.sanno.ac.jp/s /3516/	x	o	x	1日間	37,800円 (税込)	o	o	日本語
86	ビッグデータ時代を生き抜くデータサイエンス入門	産業能率大学	産業能率大学総合研究所 TEL : 03-5758-5110 FAX : 03-5758-5503 問い合わせフォーム有り 〒158-8630 東京都世田谷区等々力6-39-15 自由が丘キャンパス 1号館内	http://seminar.hj.sanno.ac.jp/s /3569/	x	o	x	2日間	38,880円 (税込)	o	o	日本語
87	高度ICT利活用人材育成カリキュラム	総務省	総務省 情報流通行政情報通信利用促進課 ICT人材の育成 TEL : 03-5253-5743 〒100-8926 東京都千代田区霞が関2-1-2 中央合同庁舎第2号館	http://www.soumu.go.jp/main_sosiki/joho_tsusin/joho_jinzai/index.html	-	o	-	-	0円	o	o	日本語

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	「ナレッジ」類型との対応		使用言語
										分析者 向け	分析者 ビジネス提供 者向け	
88	データサイエンス・スクール	総務省統計局	電話 03-5273-2020 E-mail: Y-senryaku@soumu.go.jp 〒162-8668 東京都新宿区若松町19-1 総務省第2庁舎	http://www.stat.go.jp/dss/	○	×	×	not described	無料	○	○	日本語
89	PHPによる機械学習入門	学びing株式会社	学びing 株式会社 contact@manabing.jp TEL: 048-813-8207 FAX: 048-813-8210 〒330-0073 さいたま市浦和区元町2-1-3 元町シティー2-202	http://manabing.jp/seminar	?	?	?	?	無料	○	?	日本語
90	データアナリスト育成基礎研修	株式会社 うえじま企画	(株) うえじま企画 セミナー担当 TEL: 03-5692-5030 E-mail: info@uknet.co.jp 〒114-0013 東京都北区東田端1-13-10 ツインビル田端A棟7F	http://www.uknet.co.jp/school_top.html	×	?	?	2日間	34,560円 (税込)	○	○	日本語
91	データサイエンティスト育成研修	株式会社 うえじま企画	(株) うえじま企画 セミナー担当 TEL: 03-5692-5030 E-mail: info@uknet.co.jp 〒114-0013 東京都北区東田端1-13-10 ツインビル田端A棟7F	http://www.uknet.co.jp/school_top.html	×	?	?	1ヶ月	172,800円 (税込) *バツ受講料 *各テーマを単発で受講することもできる。	○	○	日本語
92	ビジネスアナリスト人材育成研修 (ERM)	株式会社 うえじま企画	(株) うえじま企画 セミナー担当 TEL: 03-5692-5030 E-mail: info@uknet.co.jp 〒114-0013 東京都北区東田端1-13-10 ツインビル田端A棟7F	http://www.uknet.co.jp/school_top.html	×	?	?	4日間	37,800円 (税込) *バツ受講料 *各テーマを単発で受講することもできる。	○	○	日本語
93	ビジネスアナリスト人材育成研修 (MBA基礎)	株式会社 うえじま企画	(株) うえじま企画 セミナー担当 TEL: 03-5692-5030 E-mail: info@uknet.co.jp 〒114-0013 東京都北区東田端1-13-10 ツインビル田端A棟7F	http://www.uknet.co.jp/school_top.html	×	?	?	4日間	37,800円 (税込)	○	○	日本語
94	データサイエンティストコース	株式会社 ナガセPCスクール	株式会社 ナガセPCスクール 問い合わせフォーム有り 〒163-1505 東京都新宿区西新宿1-6-1 新館エルクワー5F	http://www.nps.ne.jp/datas/	×	○	×	3ヶ月~9ヶ月 集中講義で10日間	237,600円 (税込) *入学金11,000円 (税込) 個別受講必要	○	○	日本語
95	データ分析の手法	株式会社 ナガセPCスクール	株式会社 ナガセPCスクール 問い合わせフォーム有り 〒163-1505 東京都新宿区西新宿1-6-1 新館エルクワー5F	http://www.nps.ne.jp/datas/	×	?	?	1日間	23,760円 (税込) *入学金11,000円 (税込) 個別受講必要	○	○	日本語

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	「ナレッジ」類型との対応		使用言語
										分析者 向け	分析者 ビジネス提供 者向け	
96	データ分析の手法、定量分析/バツク	株式会社 ナガセPCスクール	株式会社 ナガセPCスクール 問い合わせフォーム有り 〒163-1505 東京都新宿区西新宿1-6-1 新館エルクワー5F	http://www.nps.ne.jp/datas/	×	?	?	3日間	68,040円 (税込) *入学金11,000円 (税込) 個別受講必要	○	○	日本語
97	定量分析	株式会社 ナガセPCスクール	株式会社 ナガセPCスクール 問い合わせフォーム有り 〒163-1505 東京都新宿区西新宿1-6-1 新館エルクワー5F	http://www.nps.ne.jp/datas/	×	?	?	2日間	47,520円 (税込) *入学金11,000円 (税込) 個別受講必要	○	○	日本語
98	Python速習講座	株式会社ALBERT	株式会社ALBERT TEL: 03-5333-3747 問い合わせフォーム有り 〒151-0053 東京都渋谷区代々木2-22-17	http://www.albert2005.co.jp/D/ST/open_lecture.html	?	?	?	?	?	?	?	?
99	R速習講座 (R(Rstudio)入門)	株式会社ALBERT	株式会社ALBERT TEL: 03-5333-3747 問い合わせフォーム有り 〒151-0053 東京都渋谷区代々木2-22-17	http://www.albert2005.co.jp/D/ST/open_lecture.html	×	○	×	2日間	59,400円 (税込)	○	○	日本語
100	SQL速習講座 (分析のためのSQL入門)	株式会社ALBERT	株式会社ALBERT TEL: 03-5333-3747 問い合わせフォーム有り 〒151-0053 東京都渋谷区代々木2-22-17	http://www.albert2005.co.jp/D/ST/open_lecture.html	×	○	×	2日間	59,400円 (税込)	○	○	日本語
101	データサイエンティストのためのマーケティング基礎	株式会社ALBERT	株式会社ALBERT TEL: 03-5333-3747 問い合わせフォーム有り 〒151-0053 東京都渋谷区代々木2-22-17	http://www.albert2005.co.jp/D/ST/open_lecture.html	×	×	×	1日間	32,400円 (税込)	○	○	日本語
102	データサイエンティストのためのマーケティングリサーチ基礎	株式会社ALBERT	株式会社ALBERT TEL: 03-5333-3747 問い合わせフォーム有り 〒151-0053 東京都渋谷区代々木2-22-17	http://www.albert2005.co.jp/D/ST/open_lecture.html	×	×	×	1日間	32,400円 (税込)	○	○	日本語
103	データサイエンティストのためのロジカルシンキング	株式会社ALBERT	株式会社ALBERT TEL: 03-5333-3747 問い合わせフォーム有り 〒151-0053 東京都渋谷区代々木2-22-17	http://www.albert2005.co.jp/D/ST/open_lecture.html	×	×	×	2日間	59,400円 (税込)	○	○	日本語
104	データ分析の基礎	株式会社ALBERT	株式会社ALBERT TEL: 03-5333-3747 問い合わせフォーム有り 〒151-0053 東京都渋谷区代々木2-22-17	http://www.albert2005.co.jp/D/ST/open_lecture.html	×	○	×	1日間	32,400円 (税込)	○	○	日本語

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	「カンファ」類型との対応		使用言語
										分析者向け	分析者向け	
105	できる！クラウド分析の基礎	株式会社ALBERT	株式会社ALBERT TEL: 03-5333-3747 問い合わせフォーム有り 〒151-0053 東京都渋谷区代々木2-22-17	http://www.albert2005.co.jp/D/ST/open_lecture.html	x	o	x	1日間	54,000円 (税込)	o	o	日本語
106	わかる！ビッグデータ対応のクラウド分析	株式会社ALBERT	株式会社ALBERT TEL: 03-5333-3747 問い合わせフォーム有り 〒151-0053 東京都渋谷区代々木2-22-17	http://www.albert2005.co.jp/D/ST/open_lecture.html	x	o	x	1日間	54,000円 (税込)	o	o	日本語
107	黒箱編A [データマイニング]	株式会社eftax内 SOLEIL DATA DOJO 事務局	SOLEIL DATA DOJO事務局 (株式会社eftax内) TEL: 06-6423-8240 FAX: 06-6419-5601 E-mail: info@soleildataadojo.com 〒660-0881 尼崎市昭和通3丁目9-0番地1 尼崎K. Rビルディング602号	http://www.soleildataadojo.com/education.html	x	o	x	4日間	SDD会員: 25,920円 (税込) 非会員: 32,400円 (税込)	o	o	日本語
108	ツール編A [データ前処理]	株式会社eftax内 SOLEIL DATA DOJO 事務局	SOLEIL DATA DOJO事務局 (株式会社eftax内) TEL: 06-6423-8240 FAX: 06-6419-5601 E-mail: info@soleildataadojo.com 〒660-0881 尼崎市昭和通3丁目9-0番地1 尼崎K. Rビルディング602号	http://www.soleildataadojo.com/education.html	x	o	x	4日間	SDD会員: 25,920円 (税込) 非会員: 32,400円 (税込)	o	o	日本語
109	データ分析教育講座 白帯編A [統 計学]	株式会社eftax内 SOLEIL DATA DOJO 事務局	SOLEIL DATA DOJO事務局 (株式会社eftax内) TEL: 06-6423-8240 FAX: 06-6419-5601 E-mail: info@soleildataadojo.com 〒660-0881 尼崎市昭和通3丁目9-0番地1 尼崎K. Rビルディング602号	http://www.soleildataadojo.com/education.html	x	o	x	11日間	SDD会員: 77,760円 (税込) 非会員: 97,200円 (税込)	o	o	日本語

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	「カンファ」類型との対応		使用言語
										分析者向け	分析者向け	
110	データ分析教育講座 入門編A [統 計学]	株式会社eftax内 SOLEIL DATA DOJO 事務局	SOLEIL DATA DOJO事務局 (株式会社eftax内) TEL: 06-6423-8240 FAX: 06-6419-5601 E-mail: info@soleildataadojo.com 〒660-0881 尼崎市昭和通3丁目9-0番地1 尼崎K. Rビルディング602号	http://www.soleildataadojo.com/education.html	x	o	x	2日間	SDD会員: 無料 非会員: 5,400円 (税込)	o	o	日本語
111	アドテクスタジオ・データサイエンスアカデミー	株式会社サイバーエージェント アドテクスタジオ・データサイエンスアカデミー	株式会社サイバーエージェント アドテクスタジオ・データサイエンスアカデミー E-mail: adtech_seminar@cyberagent.co.jp 〒150-0043 東京都渋谷区道玄坂一丁目12番1号	http://www.cyberagent.co.jp/r/adtech_career/adtech_academy/	x	?	o	講座回数5回 (約1ヶ月間)	?	o	o	日本語
112	初級統計学	株式会社すらがくぶんか	株式会社すらがくぶんか TEL: 03-6276-9978 Email: suugakubunnka@gmail.com (問い合わせフォーム有り) 〒160-0023 東京都新宿区西新宿7-4-4 武蔵ビル5F(HAPON Shinjuku内)	http://sugakubunka.com/list/schedule/	x	o	x	5ヶ月間	97,500円 (税込)	o	o	日本語
113	統・初級統計学-多変量解析入門	株式会社すらがくぶんか	株式会社すらがくぶんか TEL: 03-6276-9978 Email: suugakubunnka@gmail.com (問い合わせフォーム有り) 〒160-0023 東京都新宿区西新宿7-4-4 武蔵ビル5F(HAPON Shinjuku内)	http://sugakubunka.com/list/schedule/	x	o	x	5ヶ月間	97,500円 (税込)	o	o	日本語
114	統計検定2級対策 問題演習ゼミ	株式会社すらがくぶんか	株式会社すらがくぶんか TEL: 03-6276-9978 Email: suugakubunnka@gmail.com (問い合わせフォーム有り) 〒160-0023 東京都新宿区西新宿7-4-4 武蔵ビル5F(HAPON Shinjuku内)	http://sugakubunka.com/list/schedule/	x	o	o	2ヶ月間	50,000円	o	o	日本語

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	イテラティブ型との対応				
										分析者 向け	分析者 へ提供 者向け	意思決定 者向け	使用 言語	
115	統計検定準1級対策講座	株式会社すちがくぶん か	株式会社すちがくぶん TEL : 03-6276-9978 Email : suugakubunnka@gmail.com (問い合わせフォーム有り) 〒160-0023 東京都新宿区西新 宿7-4-4 武蔵ビル5F(HAPON Shinjuku内)	http://sugakubunka.com/list/schedule/	x	?	?	1年間	228,000円 (税込)	○	○	○	日本語	
116	データモデリング入門ハンズオン	株式会社データアーキテクト	株式会社データアーキテクト TEL : 045-777-6055 E-mail : info@dataarch.co.jp 〒230-???? 神奈川県横浜市鶴見区	http://dataarch.co.jp/seminar.html	?	○	x	?	?	?	○	○	○	日本語
117	ビッグデータの分析活用入門	株式会社データアーキテクト	株式会社データアーキテクト TEL : 045-777-6055 E-mail : info@dataarch.co.jp 〒230-???? 神奈川県横浜市鶴見区	http://dataarch.co.jp/seminar.html	?	○	x	?	?	?	○	○	○	日本語
118	データサイエンティスト養成コース (ベシック:5日間)	株式会社ワイ・ディ・シー	株式会社ワイ・ディ・シー 関西支社 TEL : 06-4797-1611 FAX : 06-4797-1619 問い合わせフォーム有り 〒530-0003 大阪府大阪市北区 堂島2-4-27 新藤ビル6階	https://sonar.ydc.co.jp/blog/date/2015/?cat=2	x	○	○	5日間	324,000円 (税込)	○	○	○	○	日本語
119	KVS (Key Value Store) 入門	株式会社富士通ラーニングメディア	株式会社富士通ラーニングメディア TEL : 0120-55-9019 (平日9:00~17:30) 問い合わせフォーム有り 〒108-0075 東京都港区港南 2-13-34 NSS-IIビル	https://www.knowledgegwing.com/icm/srv/course-application/init-detail?cd=FLM&pcd=FLMC&cscd=UBS	x	○	○	1日間	38,880円 (税込)	○	○	○	○	日本語
120	R言語によるデータ分析応用編	株式会社富士通ラーニングメディア	株式会社富士通ラーニングメディア TEL : 0120-55-9019 (平日9:00~17:30) 問い合わせフォーム有り 〒108-0075 東京都港区港南 2-13-34 NSS-IIビル	https://www.knowledgegwing.com/icm/srv/course-application/init-detail?cd=FLM&pcd=FLMC&cscd=UBS	x	○	○	1日間	54,432円 (税込)	○	○	○	○	日本語

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	イテラティブ型との対応				
										分析者 向け	分析者 へ提供 者向け	意思決定 者向け	使用 言語	
121	R言語によるデータ分析入門	株式会社富士通ラーニングメディア	株式会社富士通ラーニングメディア TEL : 0120-55-9019 (平日9:00~17:30) 問い合わせフォーム有り 〒108-0075 東京都港区港南 2-13-34 NSS-IIビル	https://www.knowledgegwing.com/icm/srv/course-application/init-detail?cd=FLM&pcd=FLMC&cscd=UBS	x	○	x	1日間	45,360円 (税込)	○	○	○	○	日本語
122	データサイエンスの基礎~データの収集、分析、評価~	株式会社富士通ラーニングメディア	株式会社富士通ラーニングメディア TEL : 0120-55-9019 (平日9:00~17:30) 問い合わせフォーム有り 〒108-0075 東京都港区港南 2-13-34 NSS-IIビル	https://www.knowledgegwing.com/icm/srv/course-application/init-detail?cd=FLM&pcd=FLMC&cscd=UBS	x	○	x	2日間	108,000円 (税込)	○	○	○	○	日本語
123	ビジネス分析のための統計学入門	株式会社富士通ラーニングメディア	株式会社富士通ラーニングメディア TEL : 0120-55-9019 (平日9:00~17:30) 問い合わせフォーム有り 〒108-0075 東京都港区港南 2-13-34 NSS-IIビル	https://www.knowledgegwing.com/icm/srv/course-application/init-detail?cd=FLM&pcd=FLMC&cscd=UBS	x	○	x	1日間	45,360円 (税込)	○	○	○	○	日本語
124	ビッグデータの基礎	株式会社富士通ラーニングメディア	株式会社富士通ラーニングメディア TEL : 0120-55-9019 (平日9:00~17:30) 問い合わせフォーム有り 〒108-0075 東京都港区港南 2-13-34 NSS-IIビル	https://www.knowledgegwing.com/icm/srv/course-application/init-detail?cd=FLM&pcd=FLMC&cscd=UBS	x	x	x	1日間	45,360円 (税込)	○	○	○	○	日本語
125	ビッグデータ処理基盤 (Esperanto) を活用したアプリケーションの開発	株式会社富士通ラーニングメディア	株式会社富士通ラーニングメディア TEL : 0120-55-9019 (平日9:00~17:30) 問い合わせフォーム有り 〒108-0075 東京都港区港南 2-13-34 NSS-IIビル	https://www.knowledgegwing.com/icm/srv/course-application/init-detail?cd=FLM&pcd=FLMC&cscd=UBS	x	○	○	2日間	98,496円 (税込)	○	○	○	○	日本語
126	体験! Apache Hadoop	株式会社富士通ラーニングメディア	株式会社富士通ラーニングメディア TEL : 0120-55-9019 (平日9:00~17:30) 問い合わせフォーム有り 〒108-0075 東京都港区港南 2-13-34 NSS-IIビル	https://www.knowledgegwing.com/icm/srv/course-application/init-detail?cd=FLM&pcd=FLMC&cscd=UBS	x	○	○	1日間	45,360円 (税込)	○	○	○	○	日本語

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	イタキャンアタイプ 類型との対応		使用言語
										分析者向け	分析者向け	
127	体験!Apache Spark 体験!Apache Spark プログラム	株式会社富士通ラーニングメディア	株式会社富士通ラーニングメディア TEL: 0120-55-9019 (平日9:00~17:30) 問い合わせフォーム有り 〒108-0075 東京都港区港南 2-66L 13-34 NSS- IIビル	https://www.knowledgegrowing.com/icm/srv/course-application/init-detail?cd=FLM&pcd=FLMCC&ccd=UBS66L	x	o	o	1日間	46,656円 (税込)	o	o	日本語
128	BD利活用事始め	公益財団法人九州先端科学技術研究所	公益財団法人九州先端科学技術研究所 TEL: 092-850-3452 (担当: 栗森) Email: lab.cloud-mooc@isit.or.jp 〒814-0001 福岡県福岡市早良区百道浜2-1-22 福岡SPRセンタービル7階	http://www.isit.or.jp/blog/2014/08/22/2438/	o	o	x	6週間	無料	o	o	日本語
129	Rを用いたデータ分析 (応用編)	公益財団法人九州先端科学技術研究所	公益財団法人九州先端科学技術研究所 TEL: 092-850-3452 (担当: 栗森) Email: lab.cloud-mooc@isit.or.jp 〒814-0001 福岡県福岡市早良区百道浜2-1-22 福岡SPRセンタービル7階	not described	o	?	x	not described	無料	o	o	日本語
130	Rを用いたデータ分析 (基礎編)	公益財団法人九州先端科学技術研究所	公益財団法人九州先端科学技術研究所 TEL: 092-850-3452 (担当: 栗森) Email: lab.cloud-mooc@isit.or.jp 〒814-0001 福岡県福岡市早良区百道浜2-1-22 福岡SPRセンタービル7階	http://www.isit.or.jp/blog/2014/11/14/2743/	o	o	x	4週間	無料	o	o	日本語
131	データサイエンティスト講座	朝日インタラクティブ株式会社	朝日インタラクティブ株式会社 問い合わせフォーム有り 〒101-0051 東京都千代田区神田神保町3-29-1 住友不動産一ツ橋ビル2F	http://japan.zdnet.com/business-application/sp_14data_scientist/	o	x	x	-	無料	o	o	日本語
132	Rによる統計分析	日本サード・パーティ株式会社	日本サード・パーティ株式会社 TEL: 03-6408-2488 〒140-0001 東京都品川区北品川4-7-35 御殿山トラストタワー14階	http://edu.jtp.co.jp/course/266	x	o	x	3日間	216,000円 (税込)	o	o	日本語

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	イタキャンアタイプ 類型との対応		使用言語
										分析者向け	分析者向け	
133	SQLによる集計・分析	日本サード・パーティ株式会社	日本サード・パーティ株式会社 TEL: 03-6408-2488 〒140-0001 東京都品川区北品川4-7-35 御殿山トラストタワー14階	http://edu.jtp.co.jp/course/265	x	o	x	2日間	129,600円 (税込)	o	o	日本語
134	ビッグデータ分析入門	日本サード・パーティ株式会社	日本サード・パーティ株式会社 TEL: 03-6408-2488 〒140-0001 東京都品川区北品川4-7-35 御殿山トラストタワー14階	http://edu.jtp.co.jp/course/226	x	o	x	1日間	43,200円 (税込)	o	o	日本語
135	応用1: アンケート分析のための統計	和から株式会社	和からのための楽しい数学教室 (なごみ) TEL: 03-6868-3450 問い合わせフォーム有り 東京都渋谷区渋谷3-5-16 渋谷3丁目スクエアビル2F 〒105-0021 東京都港区東新橋2-10-10 東新橋ビル2F	http://toukeigaku.jp/p/?gclid=CkU2peC6j8ACFZCXvQodY5AAATA	x	?	?	?	?	?	o	日本語
136	応用1: 医師のための統計	和から株式会社	和からのための楽しい数学教室 (なごみ) TEL: 03-6868-3450 問い合わせフォーム有り 東京都渋谷区渋谷3-5-16 渋谷3丁目スクエアビル2F 〒105-0021 東京都港区東新橋2-10-10 東新橋ビル2F	http://toukeigaku.jp/p/?gclid=CkU2peC6j8ACFZCXvQodY5AAATA	x	?	?	?	?	?	o	日本語
137	応用1: 研究論文のための統計	和から株式会社	和からのための楽しい数学教室 (なごみ) TEL: 03-6868-3450 問い合わせフォーム有り 東京都渋谷区渋谷3-5-16 渋谷3丁目スクエアビル2F 〒105-0021 東京都港区東新橋2-10-10 東新橋ビル2F	http://toukeigaku.jp/p/?gclid=CkU2peC6j8ACFZCXvQodY5AAATA	x	?	?	?	?	?	o	日本語
138	応用1: 心理学のための統計	和から株式会社	和からのための楽しい数学教室 (なごみ) TEL: 03-6868-3450 問い合わせフォーム有り 東京都渋谷区渋谷3-5-16 渋谷3丁目スクエアビル2F 〒105-0021 東京都港区東新橋2-10-10 東新橋ビル2F	http://toukeigaku.jp/p/?gclid=CkU2peC6j8ACFZCXvQodY5AAATA	x	?	?	?	?	?	o	日本語

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	「カンファ」類型との対応		使用言語
										分析者向け	分析者向け	
139	応用 2：経理・財務分析のための統計	和から株式会社	和からのための楽しい数学教室 (なごみ) TEL：03-6868-3450 問い合わせフォーム有り 東京都渋谷区渋谷3-5-16 渋谷3丁目スクエアビル2F 〒105-0021 東京都港区東新橋 2-10-10 東新橋ビル2F	http://toukeigaku.jp/p/?gclid=Cku2peC6j8ACFZCXvQod YSAATA	x	?	?	?	?	?	○	日本語
140	応用 2：データサイエンティストのための統計	和から株式会社	和からのための楽しい数学教室 (なごみ) TEL：03-6868-3450 問い合わせフォーム有り 東京都渋谷区渋谷3-5-16 渋谷3丁目スクエアビル2F 〒105-0021 東京都港区東新橋 2-10-10 東新橋ビル2F	http://toukeigaku.jp/p/?gclid=Cku2peC6j8ACFZCXvQod YSAATA	x	?	?	?	?	?	○	日本語
141	応用 2：統計解析パッケージのための統計	和から株式会社	和からのための楽しい数学教室 (なごみ) TEL：03-6868-3450 問い合わせフォーム有り 東京都渋谷区渋谷3-5-16 渋谷3丁目スクエアビル2F 〒105-0021 東京都港区東新橋 2-10-10 東新橋ビル2F	http://toukeigaku.jp/p/?gclid=Cku2peC6j8ACFZCXvQod YSAATA	x	?	?	?	?	?	○	日本語
142	応用 2：品質管理のための統計	和から株式会社	和からのための楽しい数学教室 (なごみ) TEL：03-6868-3450 問い合わせフォーム有り 東京都渋谷区渋谷3-5-16 渋谷3丁目スクエアビル2F 〒105-0021 東京都港区東新橋 2-10-10 東新橋ビル2F	http://toukeigaku.jp/p/?gclid=Cku2peC6j8ACFZCXvQod YSAATA	x	?	?	?	?	?	○	日本語
143	基礎：エクセルを使った統計	和から株式会社	和からのための楽しい数学教室 (なごみ) TEL：03-6868-3450 問い合わせフォーム有り 東京都渋谷区渋谷3-5-16 渋谷3丁目スクエアビル2F 〒105-0021 東京都港区東新橋 2-10-10 東新橋ビル2F	http://toukeigaku.jp/p/?gclid=Cku2peC6j8ACFZCXvQod YSAATA	x	?	?	?	?	?	○	日本語

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	「カンファ」類型との対応		使用言語
										分析者向け	分析者向け	
144	基礎：統計基礎	和から株式会社	和からのための楽しい数学教室 (なごみ) TEL：03-6868-3450 問い合わせフォーム有り 東京都渋谷区渋谷3-5-16 渋谷3丁目スクエアビル2F 〒105-0021 東京都港区東新橋 2-10-10 東新橋ビル2F	http://toukeigaku.jp/p/?gclid=Cku2peC6j8ACFZCXvQod YSAATA	x	?	?	?	?	?	○	日本語
145	基礎：ビッグデータのための統計基礎	和から株式会社	和からのための楽しい数学教室 (なごみ) TEL：03-6868-3450 問い合わせフォーム有り 東京都渋谷区渋谷3-5-16 渋谷3丁目スクエアビル2F 〒105-0021 東京都港区東新橋 2-10-10 東新橋ビル2F	http://toukeigaku.jp/p/?gclid=Cku2peC6j8ACFZCXvQod YSAATA	x	?	?	?	?	?	○	日本語
146	経済：MBAのための統計	和から株式会社	和からのための楽しい数学教室 (なごみ) TEL：03-6868-3450 問い合わせフォーム有り 東京都渋谷区渋谷3-5-16 渋谷3丁目スクエアビル2F 〒105-0021 東京都港区東新橋 2-10-10 東新橋ビル2F	http://toukeigaku.jp/p/?gclid=Cku2peC6j8ACFZCXvQod YSAATA	x	?	?	?	?	?	○	日本語
147	経済：計量経済学のための統計	和から株式会社	和からのための楽しい数学教室 (なごみ) TEL：03-6868-3450 問い合わせフォーム有り 東京都渋谷区渋谷3-5-16 渋谷3丁目スクエアビル2F 〒105-0021 東京都港区東新橋 2-10-10 東新橋ビル2F	http://toukeigaku.jp/p/?gclid=Cku2peC6j8ACFZCXvQod YSAATA	x	?	?	?	?	?	○	日本語
148	経済：ファイナンスのための統計	和から株式会社	和からのための楽しい数学教室 (なごみ) TEL：03-6868-3450 問い合わせフォーム有り 東京都渋谷区渋谷3-5-16 渋谷3丁目スクエアビル2F 〒105-0021 東京都港区東新橋 2-10-10 東新橋ビル2F	http://toukeigaku.jp/p/?gclid=Cku2peC6j8ACFZCXvQod YSAATA	x	?	?	?	?	?	○	日本語

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	イカンファ 類型との対応		
										分析者 向け	分析者 ビス提供 者向け	意思決定 者向け
149	経済：マーケティング・市場分析のための統計	和から株式会社	大人のための楽しい数学教室 和 (なごみ) TEL：03-6868-3450 問い合わせフォーム有り 東京都渋谷区渋谷3-5-16 渋谷3丁目スクエアビル2F 〒105-0021 東京都港区東新橋2-10-10 東新橋ビル2F	http://toukeigaku.jp/p/?gclid=Cku2peC6j8ACFZCXvQodY5AAATA	x	? ? ?	? ? ?	? ? ?	?	○	○	日本語
150	資格対策：アクチュアリーのための統計	和から株式会社	大人のための楽しい数学教室 和 (なごみ) TEL：03-6868-3450 問い合わせフォーム有り 東京都渋谷区渋谷3-5-16 渋谷3丁目スクエアビル2F 〒105-0021 東京都港区東新橋2-10-10 東新橋ビル2F	http://toukeigaku.jp/p/?gclid=Cku2peC6j8ACFZCXvQodY5AAATA	x	? ? ?	? ? ?	? ? ?	?	○	○	日本語
151	資格対策：データ解析士のための統計	和から株式会社	大人のための楽しい数学教室 和 (なごみ) TEL：03-6868-3450 問い合わせフォーム有り 東京都渋谷区渋谷3-5-16 渋谷3丁目スクエアビル2F 〒105-0021 東京都港区東新橋2-10-10 東新橋ビル2F	http://toukeigaku.jp/p/?gclid=Cku2peC6j8ACFZCXvQodY5AAATA	x	? ? ?	? ? ?	? ? ?	?	○	○	日本語
152	資格対策：統計検定のための統計	和から株式会社	大人のための楽しい数学教室 和 (なごみ) TEL：03-6868-3450 問い合わせフォーム有り 東京都渋谷区渋谷3-5-16 渋谷3丁目スクエアビル2F 〒105-0021 東京都港区東新橋2-10-10 東新橋ビル2F	http://toukeigaku.jp/p/?gclid=Cku2peC6j8ACFZCXvQodY5AAATA	x	? ? ?	? ? ?	? ? ?	?	○	○	日本語
153	資格対策：統計士のための統計	和から株式会社	大人のための楽しい数学教室 和 (なごみ) TEL：03-6868-3450 問い合わせフォーム有り 東京都渋谷区渋谷3-5-16 渋谷3丁目スクエアビル2F 〒105-0021 東京都港区東新橋2-10-10 東新橋ビル2F	http://toukeigaku.jp/p/?gclid=Cku2peC6j8ACFZCXvQodY5AAATA	x	? ? ?	? ? ?	? ? ?	?	○	○	日本語
154	Applied Logistic Regression The Ohio State University	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/course/logisticregression	○	○	x	8週間	Free	○	○	英語

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	イカンファ 類型との対応		
										分析者 向け	分析者 ビス提供 者向け	意思決定 者向け
155	Applied Regression Analysis The Ohio State University	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/course/appliedregression	○	○	x	6週間	Free	○	○	英語
156	Big Data Analytics for Healthcare Georgia Institute of Technology	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/course/bigdataanalytics	○	?	? ?	? ?	Free	○	○	英語
157	Big Data Science with the BD2K-LINCS Data Coordination and Integration Center Icahn School of Medicine at Mount Sinai	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/course/bd2kllincs	○	○	x	7週間	Free	○	○	英語
158	Big Data The San Diego Supercomputer Center (SDSC) at UC San Diego	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/specializations/big-data	○	○	○	5週間~	\$49 (first course) + 5 (other courses) x \$79	○	○	英語
159	Business Analytics The Wharton School of the University of Pennsylvania	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/specializations/business-analytics	○	○	x	4週間~	\$95 x 4 (courses) + \$215 (Capstone Project)	○	○	英語
160	Case-Based Introduction to Biostatistics Johns Hopkins University	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/course/casebasedbiostat	○	?	x	5週間	Free	○	○	英語
161	Coding the Matrix: Linear Algebra through Computer Science Applications Brown University	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/course/matrix	○	?	○	10週間	Free	○	○	英語
162	Computational Methods for Data Analysis University of Washington	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/course/compmethods	○	?	○	10週間	Free	○	○	英語
163	Core Concepts in Data Analysis Higher School of Economics	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/course/datan	○	○	x	8週間	Free	○	○	英語
164	Data Analysis and Interpretation Wesleyan University	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/specializations/data-analysis	○	○	x	4週間~	\$79 x 4 (courses) + \$79 (Capstone Project)	○	○	英語
165	Data Mining University of Illinois at Urbana- Champaign	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/specializations/datamining	○	○	? ?	? ?	\$49 x 5 (courses) + \$49 (Capstone Project)	○	○	英語
166	Data Science at Scale The University of Washington	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/specializations/data-science	○	○	○	4週間~	\$79 x 3 (courses) + \$79 (Capstone Project)	○	○	英語
167	Data Science Johns Hopkins University	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/specializations/jhudata-science?utm_medium=spark_cdp_side_bar	○	○	○	4週間~	\$29 (first course) + 9 (other courses) x \$49	○	○	英語

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	イカランタイプ 類型との対応		使用言語
										分析者向け	分析者向け	
168	Data Warehousing for Business Intelligence University of Colorado System	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/specializations/data-warehousing	0	0	0	5週間~	\$79 x 4 (courses) + \$79 (Capstone Project)	0	0	英語
169	Digital Marketing University of Illinois at Urbana-Champaign	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/specializations/digital-marketing	0	0	x	4週間~	\$79 x 5 (courses) + \$79 (Capstone Project)	0	0	英語
170	Discrete Optimization The University of Melbourne	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/courses/optimization	0	0	0	9週間	Free	0	0	英語
171	Econometrics: Methods and Applications Erasmus University Rotterdam	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/learn/erasmus-econometrics	0	0	x	8週間	Free	0	0	英語
172	Excel to MySQL: Analytic Techniques for Business Duke University	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/learn/analytics-business-metrics	0	0	x	4週間~	\$79 x 4 (courses) + \$79 (Capstone Project)	0	0	英語
173	Executive Data Science Johns Hopkins University	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/specializations/executive-data-science	0	0	x	1週間~	\$49 x 4 (courses) + \$49 (Capstone Project)	0	0	英語
174	Experimentation for Improvement McMaster University	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/learn/experimentation	0	0	x	6週間	Free	0	0	英語
175	Genomic Data Science Johns Hopkins University	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/specializations/genomics	0	0	x	4週間~	\$49 x 7 (courses) + \$49 (Capstone Project)	0	0	英語
176	Healthcare Data Visualization Georgia Institute of Technology	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/course/healthdatavisual	0	?	?	?	Free	0	0	英語
177	Introduction to Data Science University of Washington	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/course/datasci	0	0	0	8週間	Free	0	0	英語
178	Introduction to People Analytics Moscow Institute of Physics and Technology	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/learn/people-analytics	0	?	x	7週間	Free	0	0	英語
179	Introduction to Recommender Systems University of Minnesota	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/learn/recommender-systems	0	0	x	8週間	Free	0	0	英語
180	Introduction to Statistics for the Social Sciences University of Zurich	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/course/statistics101	0	0	x	10週間	Free	0	0	英語
181	Linear and Discrete Optimization Ecole Polytechnique Fédérale de Lausanne	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/course/lineartopt	0	0	0	?	Free	0	0	英語

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	イカランタイプ 類型との対応		使用言語
										分析者向け	分析者向け	
182	Machine Learning Stanford University	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/learn/machine-learning/	0	0	0	11週間	Free	0	0	英語
183	Machine Learning University of Washington	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/course/machlearning	0	0	x	10週間	Free	0	0	英語
184	Machine Learning University of Washington	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/specializations/machine-learning	0	0	x	6週間~	\$79 x 5 (courses) + \$79 (Capstone Project)	0	0	英語
185	Marketing Analytics University of Virginia	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/learn/market-analytics	0	0	x	5週間	Free	0	0	英語
186	Methods and Statistics in Social Sciences University of Amsterdam	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/specializations/social-science	0	0	x	8週間~	\$49 x 4 (courses) + \$49 (Capstone Project)	0	0	英語
187	Mining Massive Datasets Stanford University	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/course/mmnds	0	0	0	7週間	Free	0	0	英語
188	Natural Language Processing Columbia University	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/course/nlangp	0	?	x	10週間	Free	0	0	英語
189	Natural Language Processing Stanford University	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/course/nlp	0	?	0	?	Free	0	0	英語
190	Passion Driven Statistics Wesleyan University	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/course/pdstatistics	0	0	x	?	Free	0	0	英語
191	Practical Learning Analytics University of Michigan	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/course/pla	0	0	x	8週間	Free	0	0	英語
192	Probabilistic Graphical Models Stanford University	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/course/pgmm	0	0	0	11週間	Free	0	0	英語
193	Process Mining: Data science in Action Eindhoven University of Technology	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/course/procmin	0	0	x	6週間	Free	0	0	英語
194	Python for Everybody University of Michigan	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/specializations/python	0	0	x	6週間~	\$79 x 4 (courses) + \$79 (Capstone Project)	0	0	英語
195	Questionnaire Design for Social Surveys University of Michigan	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/learn/questionnaire-design	0	0	0	6週間	Free	0	0	英語

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	「オンライン」類型との対応	
										分析者向け	分析者/ビジネス提供者向け
196	Reasoning, Data Analysis and Writing Duke University	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/specializations/reasoning	○	○	×	8週間~	\$49 x 3 (courses) + \$49 (Capstone Project)	○	○
197	Social and Economic Networks: Models and Analysis Stanford University	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/course/networksonline	○	○	×	7週間	Free	○	○
198	Social Network Analysis University of Michigan	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/course/sna	○	?	×	8週間	Free	○	○
199	Statistical Reasoning for Public Health 1: Estimation, Inference, & Interpretation Johns Hopkins University	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/course/statreasoning	○	○	×	8週間	Free	○	○
200	Statistics: Making Sense of Data University of Toronto	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/course/introstats	○	○	×	8週間	Free	○	○
201	Strategic Business Analytics ESSEC Business School	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/specializations/strategic-analytics	○	○	○	4週間~	\$79 x 3 (courses) + \$99 (Capstone Project)	○	○
202	The Caltech-JPL Summer School on Big Data Analytics Caltech	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/course/bigdataschool	○	○	○	2週間	Free	○	○
203	Web Intelligence and Big Data Indian Institute of Technology Delhi	Coursera Inc.	Coursera Inc. Mountain View, CA 問い合わせフォーム有り	https://www.coursera.org/course/bigdata	○	×	○	9週間	Free	○	○
204	Analyzing and Visualizing Data with Excel edX Inc.	edX Inc.	edX Inc. 141 Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/analyzing-visualizing-data-excel-microsoft-dat206x	○	○	×	4週間	Free	○	○
205	Analyzing and Visualizing Data with Power BI edX Inc.	edX Inc.	edX Inc. 141 Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/analyzing-visualizing-data-power-bi-microsoft-dat207x	○	○	×	4週間	Free	○	○
206	Applications of Linear Algebra Part 2 edX Inc.	edX Inc.	edX Inc. 141 Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/applications-linear-algebra-part-2-davidsonx-d003x-2#VHWLk98xE	○	○	×	4週間	Free	○	○
207	Applications of Linear Algebra Part 2 edX Inc.	edX Inc.	edX Inc. 141 Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/applications-linear-algebra-part-2-davidsonx-d003x-2	○	?	×	4週間	Free	○	○

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	「オンライン」類型との対応	
										分析者向け	分析者/ビジネス提供者向け
208	Big Data and Social Physics edX Inc.	edX Inc.	edX Inc. 141 Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/big-data-social-physics-mitx-mas-s69x	○	?	×	1週間	Free	○	○
209	Career Edge: Business and Data Analysis edX Inc.	edX Inc.	edX Inc. 141 Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/career-edge-business-data-analysis-fullbridge-career3x#1	○	?	×	4週間	Free	○	○
210	Communicating Strategically edX Inc.	edX Inc.	edX Inc. 141 Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/communicating-strategically-purdue-pn-15-2	○	?	○	5週間	Free	○	○
211	Data Analysis for Life Sciences 1: Statistics and R edX Inc.	edX Inc.	edX Inc. 141 Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/data-analysis-life-sciences-1-statistics-harvard-ph525-1x	○	○	×	4週間	Free	○	○
212	Data Analysis for Life Sciences 2: Introduction to Linear Models and Matrix Algebra edX Inc.	edX Inc.	edX Inc. 141 Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/data-analysis-life-sciences-2-harvard-ph525-2x#VHWLk98xE	○	○	○	4週間	Free	○	○
213	Data Analysis for Life Sciences 3: Statistical Inference and Modeling for High-throughput Experiments edX Inc.	edX Inc.	edX Inc. 141 Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/data-analysis-life-sciences-3-harvard-ph525-3x	○	○	○	4週間	Free	○	○
214	Data Analysis for Life Sciences 4: High-Dimensional Data Analysis edX Inc.	edX Inc.	edX Inc. 141 Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/data-analysis-life-sciences-4-high-harvard-ph525-4x	○	○	○	4週間	Free	○	○
215	Data Analysis for Life Sciences 5: Introduction to Bioconductor: Annotation and Analysis of Genomes and Genomic Assays edX Inc.	edX Inc.	edX Inc. 141 Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/data-analysis-life-sciences-5-harvard-ph525-5x	○	○	○	4週間	Free	○	○
216	Data Analysis for Life Sciences 6: High-performance Computing for Reproducible Genomics edX Inc.	edX Inc.	edX Inc. 141 Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/data-analysis-life-sciences-6-high-harvard-ph525-6x	○	○	○	4週間	Free	○	○
217	Data Analysis for Life Sciences 7: Case Studies in Functional Genomics edX Inc.	edX Inc.	edX Inc. 141 Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/data-analysis-life-sciences-7-case-harvard-ph525-7x	○	○	○	4週間	Free	○	○

No.	プログラム (講題/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	イカンファア 類型との対応	
										分析者向け	分析者 ビズ提供者向け
218	Data Analysis: Take It to the MAX()	edX Inc.	edx Inc. 問い合わせフォーム有り 141. Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/data-analysis-take-it-max-delftx-ex101x	○	○	×	8週間	Free	○	○
219	Data Science and Machine Learning Essentials	edX Inc.	edx Inc. 問い合わせフォーム有り 141. Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/data-science-machine-learning-essentials-microsoft-dat203x-0	○	○	×	5週間	Free	○	○
220	Data Science Ethics	edX Inc.	edx Inc. 問い合わせフォーム有り 141. Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/data-science-ethics-michiganx-ds101x	○	○	×	5週間	Free	○	○
221	Data Structures and Algorithms Part 2	edX Inc.	edx Inc. 問い合わせフォーム有り 141. Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/s-hu-ju-jie-gou-yu-suan-fa-di-er-bu-fen-pekings-04830050-2x	○	○	○	8週間	Free	○	○
222	Data, Analytics and Learning	edX Inc.	edx Inc. 問い合わせフォーム有り 141. Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/data-analytics-learning-utarlingtonx-link5-10x	○	○	×	9週間	Free	○	○
223	Effective Thinking Through Mathematics	edX Inc.	edx Inc. 問い合わせフォーム有り 141. Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/effctive-thinking-through-mathematics-utaustinx-ut-9-01x	○	?	×	9週間	Free	○	○
224	Enabling Technologies for Data Science and Analytics: The Internet of Things	edX Inc.	edx Inc. 問い合わせフォーム有り 141. Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/enabling-technologies-data-science-columbiadx-ds103x	○	?	×	5週間	Free	○	○
225	Explore Statistics with R	edX Inc.	edx Inc. 問い合わせフォーム有り 141. Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/explore-statistics-r-kix-kieplorx-0#1	○	?	?	?	Free	○	○
226	Foundations of Data Analysis - Part 1: Statistics Using R	edX Inc.	edx Inc. 問い合わせフォーム有り 141. Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/foundations-data-analysis-part-1-utaustinx-ut-7-10x#.VHWtusk98xE	○	○	×	6週間	Free	○	○
227	Foundations of Data Analysis - Part 2: Inferential Statistics	edX Inc.	edx Inc. 問い合わせフォーム有り 141. Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/foundations-data-analysis-part-2-utaustinx-ut-7-20x	○	○	×	6週間	Free	○	○
228	"Heart" Stats: Learning to Love Statistics	edX Inc.	edx Inc. 問い合わせフォーム有り 141. Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/heart-stats-learning-love-statistics-notredamex-soc120x#1	○	?	×	9週間	Free	○	○

No.	プログラム (講題/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	イカンファア 類型との対応	
										分析者向け	分析者 ビズ提供者向け
229	Implementing Predictive Analytics with Hadoop in Azure HDInsight	edX Inc.	edx Inc. 問い合わせフォーム有り 141. Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/implementing-predictive-analytics-hadoop-microsoft-dat202-3x	○	○	○	5週間	Free	○	○
230	Implementing Real-Time Analytics with Hadoop in Azure HDInsight	edX Inc.	edx Inc. 問い合わせフォーム有り 141. Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/implementing-real-time-analytics-hadoop-microsoft-dat202-2x	○	○	○	3週間	Free	○	○
231	Introduction to Applied Biostatistics: Statistics for Medical Research	edX Inc.	edx Inc. 問い合わせフォーム有り 141. Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/introduction-applied-biostatistics-osakaux-med101x	○	○	×	6週間	Free	○	○
232	Introduction to Big Data with Apache Spark	edX Inc.	edx Inc. 問い合わせフォーム有り 141. Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/introduction-big-data-apache-spark-uc-berkeleyx-cs100-1x#1	○	○	○	5週間	Free	○	○
233	Introduction to Computational Thinking and Data Science	edX Inc.	edx Inc. 問い合わせフォーム有り 141. Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/introduction-computational-thinking-data-mitx-6-00-2x-2#1	○	○	○	10週間	Free	○	○
234	Introduction to Probability - The Science of Uncertainty	The edX Inc.	edx Inc. 問い合わせフォーム有り 141. Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/introduction-probability-science-mitx-6-041x-0	○	?	×	16週間	Free	○	○
235	Introduction to R Programming	edX Inc.	edx Inc. 問い合わせフォーム有り 141. Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/introduction-r-programming-microsoft-dat204x-0	○	○	×	4週間	Free	○	○
236	Introduction to Statistics: Descriptive Statistics	edX Inc.	edx Inc. 問い合わせフォーム有り 141. Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/introduction-statistics-descriptive-uc-berkeleyx-stat2-1x	○	?	×	5週間	Free	○	○
237	Introduction to Statistics: Inference	edX Inc.	edx Inc. 問い合わせフォーム有り 141. Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/introduction-statistics-inference-uc-berkeleyx-stat2-3x#.VHWZLck98xE	○	?	×	5週間	Free	○	○
238	Introduction to Statistics: Probability	edX Inc.	edx Inc. 問い合わせフォーム有り 141. Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/introduction-statistics-probability-uc-berkeleyx-stat2-2x	○	?	×	5週間	Free	○	○
239	Knowledge Management and Big Data in Business	edX Inc.	edx Inc. 問い合わせフォーム有り 141. Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/knowledge-management-big-data-business-hkpolyux-ise101x-0	○	?	×	6週間	Free	○	○

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	「カンファ」類型との対応	
										分析者向け	分析者/ビジネス提供者向け
240	Machine Learning for Data Science and Analytics	edX Inc.	edX Inc. 問い合わせ先 141 Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/machine-learning-data-science-analytics-columbiacx-ds102x	○	?	x	5週間	Free	○	○
241	Mathematical Modeling in Biology and Life Sciences	edX Inc.	edX Inc. 問い合わせ先 141 Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/mathematical-modeling-biology-life-pekingx-01139732x	○	○	x	10週間	Free	○	○
242	Open Government	edX Inc.	edX Inc. 問い合わせ先 141 Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/open-government-delftx-og101x	○	○	x	5週間	Free	○	○
243	Predictive Analytics	edX Inc.	edX Inc. 問い合わせ先 141 Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/predictive-analytics-iimbx-qm901x	○	○	○	7週間	Free	○	○
244	Querying with Transact-SQL	edX Inc.	edX Inc. 問い合わせ先 141 Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/querying-transact-sql-microsoft-dat201x-0	○	○	x	5週間	Free	○	○
245	Scalable Machine Learning	edX Inc.	edX Inc. 問い合わせ先 141 Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/scalable-machine-learning-uc-berkeleyx-cs190-1x#.VHWYX8k98xE	○	○	○	5週間	Free	○	○
246	Statistical Thinking for Data Science and Analytics	edX Inc.	edX Inc. 問い合わせ先 141 Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/statistical-thinking-data-science-columbiacx-ds101x	○	○	x	5週間	Free	○	○
247	Statistics for Business - I	edX Inc.	edX Inc. 問い合わせ先 141 Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/statistics-business-i-iimbx-qm101-1x	○	?	x	5週間	Free	○	○
248	Statistics for Business - II	edX Inc.	edX Inc. 問い合わせ先 141 Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/statistics-business-ii-iimbx-qm101-2x	○	○	x	7週間	Free	○	○
249	Tackling the Challenges of Big Data	edX Inc.	edX Inc. 問い合わせ先 141 Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/tackling-challenges-big-data-ying-dui-da-mit-professional-education-6-bdxszh	○	○	x	6週間	Free	○	○
250	Text Mining and Analytics	edX Inc.	edX Inc. 問い合わせ先 141 Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/text-mining-analytics-delftx-bt1x	○	○	x	6週間	Free	○	○

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	「カンファ」類型との対応	
										分析者向け	分析者/ビジネス提供者向け
251	The Analytics Edge	edX Inc.	edX Inc. 問い合わせ先 141 Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/analytics-edge-mibx-15-071x-0#.VHWYX8k98xE	○	○	x	12週間	Free	○	○
252	Wiretaps to Big Data: Privacy and Surveillance in the Age of Interconnection	edX Inc.	edX Inc. 問い合わせ先 141 Portland St., 9th floor, Cambridge, MA 02139	https://www.edx.org/course/wiretaps-big-data-privacy-surveillance-cornellx-engr1280x	○	?	x	10週間	Free	○	○
253	Making Sense of Data	Google	Forum有り。 https://datasense.withgoogle.com/forum	https://datasense.withgoogle.com/course	○	○	x	-	Free	○	○
254	Tackling the Challenges of Big Data	Massachusetts Institute of Technology	MIT Professional Education PHONE : +1 617-324-7693 E-mail : onlinex@mit.edu	http://web.mit.edu/professiona/onlinex/programs/courses/tackling_the_challenges_of_big_data.html	○	○	x	約4週間	Free	○	○
255	A/B Testing: Online Experiment Design and Analysis	Udacity, Inc.	Udacity, Inc. 2465 Latham Street, 3rd Floor Mountain View, CA 94040	https://www.udacity.com/course/ab-testing--ud257	○	○	x	4週間	Free	○	○
256	Data Analysis with R: Visually Analyze and Summarize Data Sets	Udacity, Inc.	Udacity, Inc. 2465 Latham Street, 3rd Floor Mountain View, CA 94040	https://www.udacity.com/course/data-analysis-with-r--ud651	○	○	x	約8週間	Free	○	○
257	Data Analyst Nanodegree: Discover Insights from Data	Udacity, Inc.	Udacity, Inc. 2465 Latham Street, 3rd Floor Mountain View, CA 94040	https://www.udacity.com/course/data-analyst-nanodegree--nd002	○	○	○	?	\$200/month after 1 week free trial	○	○
258	Data Visualization and D3.js: Communicating with Data	Udacity, Inc.	Udacity, Inc. 2465 Latham Street, 3rd Floor Mountain View, CA 94040	https://www.udacity.com/course/data-visualization-and-d3js--ud507	○	○	○	約7週間	Free	○	○
259	Data Wrangling with MongoDB: Data Manipulation and Retrieval	Udacity, Inc.	Udacity, Inc. 2465 Latham Street, 3rd Floor Mountain View, CA 94040	https://www.udacity.com/course/data-wrangling-with-mongodb--ud032	○	○	x	約8週間	Free	○	○
260	Intro to Data Analysis: Data Analysis Using NumPy and Pandas	Udacity, Inc.	Udacity, Inc. 2465 Latham Street, 3rd Floor Mountain View, CA 94040	https://www.udacity.com/course/intro-to-data-analysis--ud170	○	○	○	約6週間	Free	○	○
261	Intro to Data Science: Learn What It Takes to Become a Data Scientist	Udacity, Inc.	Udacity, Inc. 2465 Latham Street, 3rd Floor Mountain View, CA 94040	https://www.udacity.com/course/intro-to-data-science--ud359	○	○	○	約8週間	Free	○	○
262	Intro to Descriptive Statistics: Mathematics for Understanding Data	Udacity, Inc.	Udacity, Inc. 2465 Latham Street, 3rd Floor Mountain View, CA 94040	https://www.udacity.com/course/intro-to-descriptive-statistics--ud827	○	○	x	約8週間	Free	○	○
263	Intro to Hadoop and MapReduce: How to Process Big Data	Udacity, Inc.	Udacity, Inc. 2465 Latham Street, 3rd Floor Mountain View, CA 94040	https://www.udacity.com/course/intro-to-hadoop-and-mapreduce--ud617	○	○	○	約4週間	\$199/month after 14-day trial	○	○

No.	プログラム (講座/教材) 名	提供元	問い合わせ先	URL	online	実習等	前提条件	期間	料金	「カンパ」類型との対応	
										分析者向け	意思決定者向け
264	Intro to Inferential Statistics: Making Predictions From Data	Udacity, Inc.	Udacity, Inc. 2465 Latham Street, 3rd Floor Mountain View, CA 94040	https://www.udacity.com/course/intro-to-inferential-statistics--ud201	○	○	○	約8週間	Free	○	○
265	Intro to Machine Learning: Pattern Recognition for Fun and Profit	Udacity, Inc.	Udacity, Inc. 2465 Latham Street, 3rd Floor Mountain View, CA 94040	https://www.udacity.com/course/intro-to-machine-learning--ud120	○	○	○	約10週間	Free	○	○
266	Intro to Statistics: Making Decisions Based on Data	Udacity, Inc.	Udacity, Inc. 2465 Latham Street, 3rd Floor Mountain View, CA 94040	https://www.udacity.com/course/intro-to-statistics--st101	○	○	x	約8週間	Free	○	○
267	Machine Learning Engineer Nanodegree: Making Predictive Models	Udacity, Inc.	Udacity, Inc. 2465 Latham Street, 3rd Floor Mountain View, CA 94040	https://www.udacity.com/course/machine-learning-engineer-nanodegree--nd009	○	○	○	10-12 months	\$200/month after 1 week free trial	○	○
268	Machine Learning: Reinforcement Learning	Udacity, Inc.	Udacity, Inc. 2465 Latham Street, 3rd Floor Mountain View, CA 94040	https://www.udacity.com/course/machine-learning-reinforcement-learning--ud820	○	○	○	約4週間	Free	○	○
269	Machine Learning: Supervised Learning	Udacity, Inc.	Udacity, Inc. 2465 Latham Street, 3rd Floor Mountain View, CA 94040	https://www.udacity.com/course/machine-learning-supervised-learning--ud675	○	○	○	約8週間	Free	○	○
270	Machine Learning: Unsupervised Learning	Udacity, Inc.	Udacity, Inc. 2465 Latham Street, 3rd Floor Mountain View, CA 94040	https://www.udacity.com/course/machine-learning-unsupervised-learning--ud741	○	○	○	約4週間	Free	○	○
271	Real-Time Analytics with Apache Storm: The "Hadoop of Real-Time"	Udacity, Inc.	Udacity, Inc. 2465 Latham Street, 3rd Floor Mountain View, CA 94040	https://www.udacity.com/course/real-time-analytics-with-apache-storm--ud381	○	○	○	約2週間	Free	○	○

注) 各コンテストについて、本事業が推薦、あるいは支持するものではない。

No.	名称	主催団体	URL	概要	参加資格	参加費	開催頻度	問い合わせ先
1	Deep Analytics: Competition	オプト データサイエンスラボ	https://datasciencelab.jp/compelist	コンペティションを通じて、データ分析ニーズとデータサイエンティストのマッチングを目指している。	なし	なし	随時	OPT DataScienceLab 〒102-0081 東京都千代田区四番町6 東急東横ビル Email: 問い合わせ用のメールフォームを利用
2	データビジネス創造コンテスト	慶應義塾大学SFC研究所 データビジネス創造ラボ	https://www.sfc.keio.ac.jp/~data-business-creation/	学生がデータを広く深くDIGして新たな知の抽出や価値の創出を競う場を提供するとともに、才能の原石をあきらめずに磨き出す機会を提供することを目的とする。	高校生、大学生、大学院生 (社会人経験者は除く)	なし	年2回	慶応義塾大学SFC研究所 データビジネス創造ラボ事務局 Email: 問い合わせ用のメールフォームを利用
3	スポーツデータ解析コンペティション	日本統計学会 統計分科会	http://estat.sci.kagoshima-u.ac.jp/sports/	野球やサッカーなど数多くのスポーツの場において、欧米諸国を中心として様々なデータ収集や高度な統計分析が行われ、科学的な見地から戦術解析や選手評価がされている。また、これらの国々では学術誌やコンgresなど学術的な場においてもこうした研究発表が盛んに行われており、統計学の専門家が積極的に関与している。日本においては、2009年に日本統計学会スポーツ統計分科会を立ち上げるなどの活動が行われてきたが、スポーツの場における統計学の活用、および統計学専門家によるスポーツに関する研究はまだ十分であるとはいえない。そこで、実データに基づく応用研究の促進と、研究成果を現場に還元するきっかけを与えること、また研究者の裾野を広げることとを目的としてコンペティションが実施されている。	①日本統計学会会員または同スポーツ統計分科会員で、大学で統計学・統計科学を教える教員もしくは専攻する博士課程 (博士後期課程) の大学院生を代表者とする個人またはチーム。 ②修士課程 (博士前期課程) の学生や学部学生が参加する場合は必ず教員を代表者とすること。 ③チーム構成員には民間企業等に勤める者を含まないこと。 ④中学校・高校生のチームも参加可能 (代表は教員)	なし	年1回	開催年によって異なる。
4	INSIGHT SIGNAL	野村総合研究所	https://www.is.nri.co.jp/contest/index.html	野村総合研究所が調査を行った消費者マーケティングデータを提供し、データ分析による新たなビジネスの法則、マーケティング指標等を導き、その内容を競う。	なし	なし	年1回	株式会社野村総合研究所 消費サービス・ヘルスクアコンサルティング部 マーケティング分析コンテスト事務局 〒105-0005 東京都千代田区丸の内1-6-5 丸の内北口ビル Email: mac2015@nri.co.jp TEL: 03-5533-2647
5	学生データサイエンスリーグ	株式会社リクルートキャリア DATA リア	https://careerful.com/programDetails/?pi=K003&isc=cna0005h	リクルートグループが2014年から開催している。実データをもちに、統計学や機械学習などを活用し、予測モデルを構築するデータ分析が競われる。予測精度にカ点が置かれている。	学生であること。	なし	年1回	問い合わせ用のメールフォームを利用する。

付録 D : データサイエンス関連コンテストのリスト (平成 27 年 12 月 10 日現在)

No.	名称	主催団体	URL	概要	参加資格	参加費	開催頻度	問い合わせ先
6	DATA LEAGUE～社会人 株式会社リクルートキャリアから学ぶ！データ活用実リマ 講講座～データ分析アイ デアソン		https://careeful.com/ programDetails? pi=T003#aboutDetail	その街の魅力を身につけ出し、プロモーション施策を考 える実践型の「データ分析アイデアソン」。横浜 F・ マリノスの活動データ×横浜市のデータ×総務省のオー プンデータを素材に、各データの関係性から街の魅力 を見つけ、プロモーション施策を導き出す。 また、本プログラムでの発表内容は「YOKOHAMA YOUTH Ups! 2015」にも同時ノミネートされる。	学生であること。	なし	年1回	問い合わせ用のメールアドレスを利 用する。
7	データ解析コンベンティ ション	経営科学系研究会連合 協議会	http://jasmac- j.jimdo.com/	1994年より「共通の軍データを示し、参加者が分析を 競う」ことを目的とした「データ解析コンベンティショ ン」。POSデータのような取引データや、生活者の意 識を調査したアンケートデータなどが提供される。コ ンベンティションの成果は、学会での研究発表や、学術 論文発表として発表されている。	なし	なし	年1回	データ解析コンベンティション事務局 〒113-0033 東京都文京区本郷5- 27-8 赤門樋口ビル201 (株) 産業科学研究開発センター内 データ解析コンベンティション事務局 Email: 問い合わせ用のメール フォームを利用 TEL: (03) 5840-6467 FAX: (03) 5840-6415
8	データサイエンス・アド ベンチャー杯	独立行政法人科学技術振 興機構	http://www.sascom.j p/AAC/overview.html #Outline	科学技術イノベーションの創出に向けた我が国の科学 技術基本計画の中核的実施機関である独立行政法人科 学技術振興機構が主催。分析データとして、JST科学技 術データを使用することが義務づけられている。必要 に応じて、その他のデータを使用しても良い。	高校生以上で、学校や企業等の組織に所属する人。学 生は、必ず指導教員と一緒に応募しなければならな い。	なし	年1回	独立行政法人科学技術振興機構 アドベンチャー杯事務局 〒160-0022 東京都新宿区新宿6- 27-56 新宿スクエア5F TEL:03-4530-6157 Email: info-aac@sascom.jp kunal.jain@analyticavidhya.com
9	The Ultimate Battle: Ground of Data Scientists	Analytics Vidhya	http://datahack.analyticsvidhya.com/	?	?	?	?	
10	Data Hackathon 3.x	Analytics Vidhya	http://datahack.analyticsvidhya.com/contest /all	none	none	free	at any time	kunal.jain@analyticavidhya.com
11	Hansa Cequity Hiring Hack	Analytics Vidhya	http://datahack.analyticsvidhya.com/contest /all	Hansa Cequity is looking to hire the best among all the participants. And they are not looking to hire 1 or 2 scientists, they are looking to hire up to 10.	none	free	at any time	kunal.jain@analyticavidhya.com
12	LearnUp - Learn Data Science From Scratch	Analytics Vidhya	http://datahack.analyticsvidhya.com/contest /all	This meetup is organized for beginners keen to learn data science / analytics. During the session, we'll teach them useful basic tricks in a hands on manner and then help them with their career related queries.	none	free	at any time	kunal.jain@analyticavidhya.com

No.	名称	主催団体	URL	概要	参加資格	参加費	開催頻度	問い合わせ先
13	The Data Science Bowl	Booz Allen Hamilton, Inc. and Kaggle, Inc.	http://www.datasien.com/ http://www.kaggle.com/	We are pioneering a new language to understand our incredibly beautiful and complex world. A language that is forward-looking rather than retrospective, different from the words of historians and famed novelists. It is data science; and through it, we have the power to use insights from our past to build an unprecedented future. We need your help building that future. The Data Science Bowl offers tremendous potential to modernize the way we understand and address a major challenge.	none	free	once a year	http://www.boozallen.com/cont act-us
14	Data Science Competitions	CloudANALYTIX, Inc.	https://www.crowdan alytics.com/	The CrowdANALYTIX Community: where data experts collaborate & compete to solve business challenges for global enterprises.	none	free	at any time	CrowdANALYTIX, Inc. US Office: 440 N Wolfe Rd, Suite 170 Sunnyvale, CA 94085, USA India Office: 2621, 2nd Floor, 27th Main Rd, Sector 1, HSR Layout, Bangalore 560102 India Email: connect@crowdanalytix.com
15	Data Science Game	Data Science Game	http://www.datasien.com/	An international challenge around data-driven issues, divided into two events: 1) An online non- eliminatory phase to practice with the software and understand our expectations. 2) The real deal.	1st or 2nd year MSc student, or even 1st year PhD candidate (if you come from abroad)	free	once a year	http://www.datasiencegame.co m/#register/cngp
16	Find a Hackathon	Devpost, Inc.	http://devpost.com/h ackathons	none	none	free	at any time	Devpost, Inc. 433 West 14th Street, Suite 3F, NYC 10014 Email: support@devpost.com
17	Competitions	DrivenData, Inc.	http://www.drivendat a.org/competitions/	Open-innovation competition where freelance developers and data scientists submit statistical models. Using our competition platform and evaluation engine, the models are ranked based on how well they predict data withheld from the competitors.	none	free	at any time	DrivenData Inc. PO Box 390664 Cambridge, MA 02139 Email: info@drivendata.org

No.	名称	主催団体	URL	概要	参加資格	参加費	開催頻度	問い合わせ先
	Music Data Hackathon	Data Science London	http://datascienceglobb.org/music-data-hackathon/	Insight performs extensive market research about their artists by interviewing thousands of people around the world. This research has produced EMI One Million Interview Dataset; one of the largest music preference datasets in the world today, that connects data about people—who they are, where they live, how they engage with music in their daily lives— with their opinions about EMI's artists. This Data Science London hackathon will focus on one key subset of this data: understanding what it is about people and artists that predicts how much people are going to like a particular track. We have taken a sample of the data from the United Kingdom that provides a granular mixture of profile, word-association, and rating data. The goal of this weekend hackathon is to design an algorithm that combines users' (a) demographics, (b) artist and track ratings, (c) answers to questions about their preferences for music, and (d) words that they use to describe EMI artists in order to predict how much they like tracks they have just heard. There is also a Data Visualization track for data artists to show off their work.	none	free	once a year	Richard O'Brien EMI Group, London. Tel: +44 (0)20 7795 7447 Email: obrienr@emigroup.com Ralph Risk, Marketing Director, EMEA Lightspeed Research Ltd Tel: +44 (0)207.896.1950 Email: risk@lightspeedresearch.com Jenny Sneyd Brands2Life/EMC Tel: +44 (0)20 7592 1200 Email: jennysneyd@brands2life.com Paige Schoknecht Cutline Communications (for Kaggle) Tel: +1 (415) 348-2708 Email: pschoknecht@cutline.com

18

Data Science Competition	Government Statistical Service (GSS), Government Operational Research Service (GORS), Government Economic Service (GES), and Government Social Research Service (GSR)	https://gss.civilservice.gov.uk/statistics/open-data/win-data-science-learning/	The GSS, GORS, GES and GSR professions have teamed up to create an exciting competition designed to promote data science and innovative analytical techniques in government. The prize also provides a fantastic opportunity to develop skills and knowledge in the field but best part is that it is so easy to enter! The competition is a platform to showcase the great work already done across government by analysts of all backgrounds.	Anyone from the Civil Service can enter to win the prize.	free	once a year	GSS Professional Support Team Office for National Statistics, Cardiff Road, Newport, NP10 8XG. Email: gssnet@ons.gsi.gov.uk
--------------------------	---	---	---	---	------	-------------	--

19

Innocentive Challenges	InnoCentive, Inc.	https://www.innocenti.com/ar/challenge/browse	InnoCentive crowdsources innovation solutions from the world's smartest people who compete to provide ideas and solutions to important business, social, policy, scientific, and technical challenges.	none	free	at any time	InnoCentive, Inc. 265 Winter Street, 2nd Floor Waltham, MA 02451 USA Phone: +1-978-482-3300 Email: 問い合わせ用のメール フォームを利用
------------------------	-------------------	---	--	------	------	-------------	---

20

No.	名称	主催団体	URL	概要	参加資格	参加費	開催頻度	問い合わせ先
	kaggle competitions	Kaggle, Inc.	https://www.kaggle.com/	Kaggle is the world's largest community of data scientists. They compete with each other to solve complex data science problems, and the top competitors are invited to work on the most interesting and sensitive business problems from some of the world's biggest companies through Masters competitions. Kaggle provides cutting-edge data science results to companies of all sizes. We have a proven track-record of solving real-world problems across a diverse array of industries including life sciences, financial services, energy, information technology, and retail.	none	free	at any time	Kaggle, Inc. 188 King St., Unit 502, San Francisco, CA 94107 Email: support@kaggle.com

21

Academic Competitions	Kaggle, Inc.	https://www.kaggle.com/	Kaggle hosts free projects for hundreds of universities around the globe. Engage students with an opportunity to apply machine learning to real problems.	none	free	at any time	Kaggle, Inc. 188 King St., Unit 502, San Francisco, CA 94107 Email: support@kaggle.com
-----------------------	--------------	---	---	------	------	-------------	--

22

KDD Cup	KDD Cup	http://www.kdd.org/	KDD Cup is the annual Data Mining and Knowledge Discovery competition organized by ACM Special Interest Group on Knowledge Discovery and Data Mining, the leading professional organization of data miners.	none	free	once a year	ACM Special Interest Group on Knowledge Discovery and Data Mining Email: 問い合わせ用のメール フォームを利用
---------	---------	---	---	------	------	-------------	---

23

Data Science Competition	R&D Lab	http://indatalabs.com/blog/datascience_blog/	Indata Labs, and the Research Institute for Applied Problems of Mathematics and Informatics of Belarusian State University	none	free	once a year	Indata Labs Email: lab@indatalabs.com
--------------------------	---------	---	--	------	------	-------------	--

24

Data Science Competition	Strategenics Pty Ltd.	http://strategenics.com.au/who-we-are/news-and-events/2015-data-science-competition/	ModelOff is the world's largest professional competition for Financial Modeling, Financial Analysis, Investment Analysis and Microsoft Excel. The global online competition attracts thousands of industry-leading professionals and students from more than 110 countries working across Finance, Banking, Accounting, Insurance, Actuarial, Engineering, Consulting and many Quantitative disciplines.	?	free	once a year	Strategenics Pty Ltd. GPO BOX 5066, Brisbane, 4000, Queensland Email: contact@strategenics.com.au
--------------------------	-----------------------	---	--	---	------	-------------	--

25

MODELLOFF - Financial Analysis & Modeling	PSCL - Professional Services Champions League	http://www.modeloff.com/	ModelOff is the world's largest professional competition for Financial Modeling, Financial Analysis, Investment Analysis and Microsoft Excel. The global online competition attracts thousands of industry-leading professionals and students from more than 110 countries working across Finance, Banking, Accounting, Insurance, Actuarial, Engineering, Consulting and many Quantitative disciplines.	none	free	once a year	PSCL - Professional Services Champions League Email: info@psciglobal.com
---	---	---	--	------	------	-------------	---

26

TEXATA - The World's Big Data Analytics Showdown	PSCL - Professional Services Champions League	http://www.texata.com/	TEXATA is a fun, innovative and challenging annual competition for students and professionals to develop and test their big data analytics skills against friends, colleagues and top data scientists around the world. Competitors undertake two Online Qualification Rounds before the World's Top 12 progress to a Live Finals Event held in Austin, Texas.	none	free	once a year	PSCL - Professional Services Champions League Email: info@psciglobal.com
--	---	---	--	------	------	-------------	---

27

No.	名称	主催団体	URL	概要	参加資格	参加費	開催頻度	問い合わせ先
	LOYALTYGAMES - Customer Loyalty & Gamification	PSCL - Professional Services Champions League	http://www.theloyaltygames.com/	Loyalty, Rewards and Gamification are among the fastest growing and critically important disciplines within Professional Services and Enterprise worldwide. The omnipotence power of social, digital marketing and mobile platforms makes these skills more important. LoyaltyGames, the Loyalty and Gamification World Championships, is a challenging case study event that promotes analytical skills for professionals and students working in Loyalty Marketing, Rewards Programs, Management Consulting, Gamification, Games-Based Learning and other innovative industries.	none	free	once a year	PSCL - Professional Services Champions League Email: info@psciglobal.com
28	HIQORA - High IQ World Championships	PSCL - Professional Services Champions League	http://www.hiqora.com/	The High IQ World Championships are an educational and intellectually-stimulating competition for the world's brightest and most gifted minds. Participants of all ages and countries engage in a wide-range of innovative activities and questions to test intelligence in quantitative puzzles, games-based learning and real-world IQ aptitude tests. HIQORA helps foster intellectual education and continuous learning across areas of ability and interest for the high-achieving intellectual community.	none	free	once a year	PSCL - Professional Services Champions League Email: info@psciglobal.com
30	ACTURIAN - Actuarial World Championships	PSCL - Professional Services Champions League	http://www.acturian.com/	The Actuarial World Championships (Acturian™) is an educational and challenging competition for actuarial professionals and university students in Life Insurance, General Insurance, Pensions, Risk, Investments, Mathematics and Quantitative Disciplines.	none	free	once a year	PSCL - Professional Services Champions League Email: info@psciglobal.com
31	PENTIOS - The World's Finest Analytical Minds	PSCL - Professional Services Champions League	http://www.pentios.com/	PENTIOS World Championship is an intellectual business competition for professionals and students that is completed over 24 hours with multi-disciplinary business skills. It involves the completion of five continuous and sequential skills disciplines across Professional Services and Business Consulting industries. While several variations of a pentios competition can exist, the PENTIOS World Championship involves five non-stop challenges within 24 hours including: Financial Modelling, Actuarial Sciences, Big Data Analytics, Customer Loyalty Analytics and High IQ Intelligence.	none	free	once a year	PSCL - Professional Services Champions League Email: info@psciglobal.com

No.	名称	主催団体	URL	概要	参加資格	参加費	開催頻度	問い合わせ先
	topcoder	topcoder	https://www.topcoder.com/	Topcoder gathers the world's experts in design, development and data science to work on interesting and challenging problems. Members are provided with opportunities to demonstrate their expertise, improve their skills, and win cash, while helping real world organizations solve real world problems. Topcoder has been hosting technical competitions since 2001 and is seen as one of the top online technical communities in the world.	none	free	at any time	Topcoder, Inc. 760 Market Street San Francisco, CA 94102 Email: 問い合わせ用のメールアドレス フォームを利用
32	TunedIT Challenges	TunedIT Solutions	http://tunedit.org/challenges	TunedIT Challenges is the first web platform for hosting data mining competitions, launched in 2009 on top of the Research framework for the evaluation of data-driven algorithms. Tunedit Challenges provide extremely flexible and easy way to launch data mining contests of any type, in every application domain.	none	free	at any time	TunedIT Solutions ul. Żwirki i Wigury 93 lok. 3049 02-089 Warszawa Poland E-mail: tunedit@tunedit.org Phone: +48 22 662 3196
34	ALGORITHMS COMPETITION	Two Roads Tech Solutions	http://tworoads.co.in/#aboutus	The Algorithms Competition is to write a very fast and correct implementation of a tertiary random forest.	none	free	at any time	Two Roads Tech Solutions Suite 354, Evoma, # 14 Bhattarahalli, K R Puram, Bangalore 560049. India Email: info@tworoads.co.in Phone: +91 80 3061 3896
35	ONLINE LEARNING CHALLENGE	Two Roads Tech Solutions	http://tworoads.co.in/#aboutus	The Online Learning Challenge is about finding the optimal way for mixing a number of signals to form a strategy that performs best over unseen data. Being a purely machine learning/statistical problem, this does not require any experience in finance.	none	free	at any time	Two Roads Tech Solutions Suite 354, Evoma, # 14 Bhattarahalli, K R Puram, Bangalore 560049. India Email: info@tworoads.co.in Phone: +91 80 3061 3896
36	QUANT CHALLENGE	Two Roads Tech Solutions	http://tworoads.co.in/#aboutus	The Quant Challenge is to find a portfolio investing strategy form data that performs admirably in out of sample data.	none	free	at any time	Two Roads Tech Solutions Suite 354, Evoma, # 14 Bhattarahalli, K R Puram, Bangalore 560049. India Email: info@tworoads.co.in Phone: +91 80 3061 3896

No.	名称	主催団体	URL	概要	参加資格	参加費	開催頻度	問い合わせ先
	DATA SCIENCE BASED TRADING COMPETITION	Two Roads Tech Solutions	http://tworoads.co.in/#aboutus	<p>The Data Science based trading competition is a trading competition with an emphasis on the machine learning aspect i.e. how best to trade a security, perhaps by predicting a price in the future. The profit/loss (PnL) of the trading strategy depends upon how well we can predict future price of the security to be traded and the execution logic which specifies how to trade. Since this competition primarily tests machine learning skills, we have coded up the execution logic for you. Although it is possible to work without any knowledge of the execution logic, we suggest you go through the script <code>simulate_signal.py</code> to better understand how predicted prices translates to PnL. To help you predict prices, in addition to the prices of the security to be traded we have provided you with a set of features.</p>	none	free	at any time	Two Roads Tech Solutions Suite 354, Evoma, #14 Bhattarahalli, K R Puram, Bangalore 560049, India Email: info@tworoads.co.in Phone: +91 80 3061 3896



文部科学省委託事業

データサイエンティスト育成ネットワークの形成

文部科学省委託事業
「データサイエンティスト育成ネットワークの形成」
平成 27 年度事業報告書

発行日／平成 28 年 3 月

情報・システム研究機構

統計数理研究所 データサイエンティスト育成ネットワークの形成
〒190-8562 東京都立川市緑町 10-3 Tel : 050-5533-8500(代表)

e-mail: info@datascientist.ism.ac.jp