



早稲田大学組織経済実証研究所 監修 「Python による人事データ分析入門」

講座概要・シラバス

— 架空人事データを用いた実践的データ分析スキルの習得 —

人的資本情報の開示義務化、女性活躍推進、働き方改革、人財育成、健康増進など様々な経営課題に対する人事の関与への要求は高まる一方です。少ない経営資源をどのような施策に投入していくべきか、施策の効果をどのように評価すべきか、といった問題が議論される中で、人事業務においてもエビデンスに基づく政策決定や、データ分析に基づく KPI(重要業績評価指標)の設定の必要性が広く認識されるようになってきました。

早稲田大学組織経済実証研究所およびその前身である企業内データ計量分析プロジェクトは、2014 年以降、東京大学と早稲田大学で人事データの新しい活用方法を模索する研究会「人事情報活用研究会」を開催してきました。その活動の中で培ってきたノウハウを標準的なカリキュラムにまとめ、より多くの人事の方々に活用していただこうと考え、新たな試みとしてオンライン講座「R による人事データ分析入門」を 2020 年度に開講しました。そして、受講生のご要望を受け、新たに「Python による人事データ分析入門」を 2025 年度に開講する運びとなりました。

「時間・場所」の制約なく、幅広い参加者に、人事データを用いた統計分析の手法を学んでいただきたいと思います。

本講座で扱う汎用プログラミング言語「Python」は、AI 開発やデータ分析など様々な用途で利用されています。特に、データ処理から統計分析、機械学習までを一気通貫で実行できることから、ビジネスのデータ分析の現場で利用されることが増えてきました。また、無料で動作させることができるため、手軽に試すことができるという大きなメリットがあります。一方、「Python」はプログラミング言語であるため、習得するために時間がかかる場合もあります。そこで、予備知識を仮定せず、出来るだけ分かりやすく事例と実践を交えてお伝えします。

本講座では、データサイエンスの中でも因果推論と呼ばれる手法に焦点をあてます。データサイエンスと言うと、機械学習を用いた予測(どの応募者が内定を辞退するか、どの社員が辞める確率が高いかなど)を思い浮かべられるかもしれませんが、組織経済実証研究所では個人に紐づけた予測分析いわゆるプロファイリングを推奨しておりません。個人単位の予測はそもそも誤差やバイアスが大きいですし、私たちは、データの利活用を組織の改善や社員の働く環境の向上のために使われるべきだと考えています。したがって、何らかの統計的傾向を持つ個人を特定する予測分析ではなく、課題の発見や人事施策の効果測定に重きをおいた因果推論にウェイトをおいた講義内容としています。

エビデンスに基づいた社内人材の最適活用の推進に、ぜひともこのeラーニング講座をお役立てください。

本講座の3つの特徴

1. 人事データの分析を目的とする人にとって必要なPythonの機能やコマンドを身につけられるように講義内容が最適化されています。人事データの一つの特徴はパネル構造(個人の情報が日、月、年といった時系列で蓄積されていること)であり、その構造にあった分析方法を提案しています。ご提供する架空人事データを実際に分析しながらハンズオンで人事データ分析を学べるため、すぐに実務に活用できます。学習の中でプログラムしたRのコードを使い回すことで、実際の自社データにも簡単に適用することができます。
2. 効率的な学習を実現するため、全ての講義をオンデマンド動画で提供します。本講座は、非常に使いやすく設計されたLMS(ラーニングマネジメントシステム)を利用しており、パソコンやスマホ、タブレットなどさまざまなデバイスから気軽に受講でき、さらにはLMSのクイズ機能を利用した練習問題・小テストに取り組むことで、講義内容の理解を深めることができます。
3. 学術界・産業界で活躍する人事データ分析の専門家が講師として講義をお届けします。人事データ分析の専門知識を有する研究者や、多くの企業に人事データ活用支援サービスを提供している実務家などを講師陣に迎えており、他に類を見ないほど本格的な人事データ分析を学べる講座となっています。前提知識ゼロからステップ・バイ・ステップで、人事データサイエンティストを養成します。

人事データ活用推進の3つの指針

この講座は、人事データを用いて自社あるいは顧客の組織や人事制度や生産性の改善に取り組みたいというすべての人事プロフェッショナルを対象としています。しかし、人事データ活用は、分析手法を覚えれば、すぐに実施できるという簡単な話ではありません。以下の3つの取組みを併せて進めていく必要があります。

- (1) 有益なデータを蓄積し、一元管理していくこと、
- (2) 取り組むべき経営課題を明らかにし、優先順位をつけていくこと、
- (3) 分析結果を適切に解釈し、経営にとって重要な含意を整理すること。

こうした総合的な能力は、データ分析を1回きりではなく、継続的に行うことによって高めることができます。本講座の受講がこうした努力の始まりとなることを期待しています。

オンデマンド講義各回の内容

第1章 Python について

プログラミング言語 Python および本講座で使用する Google Colab の特徴とそのユーザーインターフェイスについて、紹介します。

第2章 Python の基本操作

Python の演算、関数、コードを書く際の決まりなど習得します。

第3章 データの構造とデータの読み込み

データ構造の種類や、Python 内でのデータの扱い方法について、基礎知識を学びます。

第4章 データラングリング

ERP など社内システムから分析に必要なデータを抽出した後で、（欠損値や重複の有無の確認、不要なテキストの削除など）データをクリーニングし、さらに目的に合わせて、（変数の整理・作成やフォーマットの変換など）データをラングリングする必要があります。その方法について基本的な操作方法を学びます。

第5章 記述統計量によるデータの要約

どのような分析も、まずは個々のデータ系列の特徴量や分布を把握してから、始めたほうがより効率的な分析や解釈が可能となります。データを要約するための方法を紹介します。

第6章 データビジュアライゼーション

2つの事象の間に因果関係があるのか、施策の効果があったのか、といったことを知りたい時は、まずグラフを描いて視覚的にとらえることが大事です。Python には様々なグラフ作成のコマンドがあり、Excel などと違って、コードをコピーペーストして再利用することで、簡単に数多くのグラフを作成することができます。

第7章 統計的仮説検定

データ系列間の相関を見る場合、それが偶然の産物なのか、背後に相関を生み出すメカニズムがあったのかを確認しながら前に進むことが大事です。そのための検定方法を解説します。

第8章 線形回帰モデル

因果関係や相関関係を明らかにするために、最も一般的に行われる分析が回帰分析です。他の条件を一定とした時に、 X と Y の間にどのような関係があるのか、様々なデータで分析してみましょう。

第9章 二項回帰分析

人事データで扱う情報には、給与や労働時間のように連続的な変数として扱えるものと、昇進、離職のように1（発生）と0（変化なし）で表せるもの（2項変数）や、人事考課のように（S, A, B, C, D など）順序付けられたカテゴリー（離散変数）、あるいは職種のように順序付けのないカテゴリー（離散変数）などがあります。被説明変数がそれぞれのタイプごとに適した回帰分析が存在し、その中でも頻繁に使われる

線形確率モデル、ロジット・モデル、プロビット・モデルの三種類の二項回帰モデルについて解説します。

第 10 章 因果関係の基本と RCT

相関関係ではなく、因果関係があるかどうかを明らかにするための分析アプローチの一部を紹介します。ランダム化比較試験（RCT）はいわゆる実験で、ランダムに介入対象を決めて施策を導入し、導入しなかったグループと比較することで効果を検証します。

第 11 章 奨励デザインと操作変数法

公平性の観点などにより実験ができない場合、施策の実施の代わりに奨励を行うことで、施策の効果を推定することができます。奨励デザインと呼ばれるこのような手法と、回帰分析の枠組みで分析を行う操作変数法を紹介します。

第 12 章 傾向スコアマッチング

実験ができなくても、社内の特定の部署や社員を対象に何らかの施策が行われた場合、実験に近い状況を作り出して、効果検証を行うことが出来ます。例えば、施策対象職場に非常に似た職場を探して来て比較するとか、施策対象の社員に似た属性の社員を集めてきて比較するといったアプローチです。そのための手法を解説します。

第 13 章 パネルデータ分析

個人について継続的に記録されるパネルデータを用いた分析手法を紹介します。固定効果モデルは、回帰分析の係数にバイアスを引き起こす、観測できない個人の能力の影響を取り除くことを可能にします。

第 14 章 主成分分析による次元縮約

多くの変数をもつ(多次元)データについて、情報をできるだけ損なわずに少ない主成分に要約し、次元を縮約する主成分分析を紹介します。多次元データを主成分に縮約することで、視覚化や解釈が容易になり、分析の変数として活用できる利点があります。

第 15 章 クラスター分析によるグルーピング

与えられたデータから、類似しているデータを同じカテゴリーにまとめる手法であるクラスター分析を紹介します。クラスター分析によって、例えばアンケートデータを用いて回答をいくつかの分類に分け、回答者の傾向を分析するなどの分析を行うことができます。

第 16 章 決定木によるハイパフォーマー予測

条件の組み合わせで回帰や分類を行う、決定木分析と呼ばれる分析手法について解説します。特に、実践的にどのような条件の組み合わせでハイパフォーマーの割合が高くなるかを分析するハイパフォーマー予測を実施し、結果を解釈する方法を学びます。

質問対応

受講者と講師陣のオンライン交流スペースを設けています。講義内容に関する質問などをチャットで随時受け付けいたします。

受講特典

本講座の受講特典として、人事領域で注目を集める専門家をお招きする講演会シリーズ「ピープルアナリティクスアカデミー」に無料でご招待いたします。

受講期間

受講期間は 10 ヶ月間です。例えば、4 月 1 日に受講開始案内を受け取られた方は、翌年 1 月末日終了となります。

講座期間中にすべての受講動画を視聴し、各章の確認テストに合格された方には、修了証としてオープンバッジ(※注 1)を発行します。

(注 1) オープンバッジ <https://www.openbadge.or.jp/about-ob>

受講料金

2025 年度の受講料金は以下の通りです。

【受講費：一般】

法人：100,000 円／名(税別)

個人：100,000 円／名(税別)

※本協会の正会員として会費を納めておられる企業・団体等の法人の場合、またはその法人にお務めの個人の方は下記割引価格が適用されます。

【受講費：会員割引】

法人：50,000 円／名(税別)

個人：50,000 円／名(税別)

契約主体		法人	個人
一般	1 名	10 万円	10 万円
協会の会員企業所属	1 名	5 万円	5 万円

(注) 税別

受講お申し込みと手続きの流れ

下記のフォームよりお申し込みください。

<https://req.qubo.jp/pahrtech/form/IMr8IH64>

- ・受講申込フォームにご入力いただいてから数日以内に請求書をお送りいたします。
- ・毎月 20 日までに手続き完了（フォームからのお申込み・受講料お振込み）された方へ翌月 1 日に受講開始のご案内メールをお送りします。
- ・20 日にお振込み完了され、翌営業日に入金になるものは、20 日に手続き完了されたものとして扱います。
- ・21 日以降のお申込みは、自動的に翌々月 1 日開始分となります。

監修：早稲田大学組織経済実証研究所

早稲田大学組織経済実証研究所は、企業内データ、HR サービス事業者データを用いた組織経済学、人事経済学、行動経済学分野における実証研究をテーマに活動を行う研究組織です。所長である大湾秀雄(早稲田大学政治経済学術院教授)を中心に、人事・組織の課題をテーマに数多くの企業と産学連携の共同研究を推し進めています。

お問い合わせ

ピープルアナリティクス & HR テクノロジー協会事務局 e ラーニング講座担当
e-learning-pa@peopleanalytics.or.jp