

Python・R・SQLを用いた

データサイエンス 実践セミナー

eラーニング付

■ 必要な事前知識

- Python、R、SQL の基本文法を理解していること
※プログラミング入門セミナーを受講済みであることが望ましい

■ このような課題をお持ちの方におすすめ

- データサイエンスで使うプログラミングをどのように学んで良いかわからない
- グラフを描いたりすることはできるが、どのように読み解いて良いかわからない
- データサイエンスで使うライブラリが多様なので、なかなか覚えられない
- データ分析プロジェクトの一連の流れを想像できない

■ 概要

データサイエンスプロジェクトでは、Python、R、SQLの3つのツールがよく用いられます。機械学習や統計学といったデータサイエンスのスキルを活用してデータ活用するとは、どういうことなのでしょう？その実態は課題設定から地道なデータ収集、分析という試行錯誤の連続です。

本セミナーでは、実践に即した形でデータハンドリング、データ可視化、統計解析のライブラリや関数を使えるようになることを目指します。また、課題設定からデータハンドリング、分析、意思決定までの流れを講師と一緒にウォークスルーを行います。データ分析プロジェクトのリアルを体感することで実務での勘所を養います。

■ 学習ゴール

- Pythonのpandasやmatplotlibといったライブラリを用いた、データハンドリングや可視化の方法を理解する
- Rのデータを読み込んでから仮説検定をするまでの流れを理解する
- SQLでサブクエリを用いた複雑な集計をできるようになる
- データ分析プロジェクトの課題設定からデータ収集、分析、結果を出すまでの流れを理解する
- アルゴリズム・手法の選択、精度指標の考え方を数理ではなく、ビジネスの文脈で理解する



■ プログラム

※事前予習 eラーニングを視聴していただきます。

1 日目 9:45 ~ 16:45

- Pythonを使ったデータハンドリング
- Pythonを使ったデータ可視化
- **ショートプロジェクト** データを可視化して読み解こう
- Rを使ったデータハンドリング
- Rを使った仮説検定と回帰分析
- **ショートプロジェクト** Rを使ったデータ処理
- SQLを使ったデータハンドリング (JOIN、サブクエリ、CASE式)
- **ショートプロジェクト** アプリのログデータを集計してみよう

2 日目 9:45 ~ 16:45

- **ウォークスルー①** UberとLyftの価格を予測する
 - ・ Pythonを使ったデータハンドリング・データ可視化
 - ・ 機械学習のおさらい
 - ・ scikit-learnを使った機械学習のモデリング
 - ・ 精度評価
 - ・ データサイエンスの世界からビジネスの世界への意思決定へ
- **ウォークスルー②** メーカーのリポート最適化プロジェクト
 - ・ Rでデータハンドリング・データ可視化
 - ・ 回帰分析のおさらい
 - ・ データサイエンスの世界からビジネスの世界への意思決定へ

※内容は、変更される場合があります。あらかじめご了承ください。

■ 開催日時

2025年3月11日(火)~12日(水)

【時間】 2日間とも 9:45 ~ 16:45

■ 形式 Zoomを使ったオンラインセミナー

■ 講師 (敬称略)

株式会社データミックス 講師陣

2017年設立。データサイエンス領域における教育事業やソリューション事業の企画開発、データ分析実務スキル検定の企画・運営などをおこなう。

■ 参加料 (税込)

● 法人会員…121,000円/1名

● 会員外…144,100円/1名

