科目名：IoTとデータ活用Ⅱ
受講者数：7名
期間：R7年10月～R8年1月　16コマ
概要：IoT基礎知識、データ分析スキルを学ぶ。特にⅡでは、IoT収集データを用いたデータ分析演習を行っていく。
▼教材

IoTデータ活用

ここでは、生成AIを活用してのデータ分析を学習します。
前期で基礎的なデータ分析手法をExcelを活用して学んできましたが、ここでは同様の分析を生成AIを用いてどのくらい容易に分析が可能になるかを体感してみます。

2

前期の内容を少し振り返ってみましょう

これまでデータ分析でやったこと





3

本科目の進め方

データサイエンスとは？

データサイエンスとは、大量のデータから有益な情報を抽出し、分析、解釈するための学問分野。
統計学や情報工学、アルゴリズムなどの手法を組み合わせて、データから問題解決に必要な知見を引き出します。データサイエンスでは以下のようなプロセスで作業を進めていきます

1. **データの収集**
2. **データの前処理**
3. **モデルの作成・分析**
4. **モデルの評価**

データサイエンスは、交通システム、金融市場や販売の予測と様々な分野で活用されており、農業分野でも活用が始まっています。

データサイエンス各論（活用事例） - 筑波大学 (jv-campus.org)

4

本科目の進め方

データサイエンスとは？

- データを用いて、新たな価値を引き出す手法
- データサイエンスとAIとは密接な関係

データサイエンス活用の3つのポイント

- データサイエンスを使って解決しようとしている「問題」は何か？
- どのような「データ」を収集しているか？
- 現場にデータサイエンスをどのように「導入」しているか？

5

業界の理解

業界の理解

「農産物出荷量（価格）の予測」というテーマを選定しているので、予測を行う上で必要となる基本的な業界知識を理解する

- 農産物市場の基礎（日本の農業概要、卸売市場、価格形成メカニズムなど）
- 農産物の栽培時期と地域（分析対象となる農産物の栽培特性と主要地域）
- 農業と気象の関係（気象の基礎知識および農産物と気象条件の関係）
- 農業におけるデータサイエンス事例

6

予測対象の決定

Step 1　予測対象の候補選定

- どの農産物の出荷量（価格）を予測するのか
  例えば、
  比較的長期間収穫できる農産物と短期間で収穫が終わるものを比較してみる
  降水量の影響をあまり受けない施設栽培と影響を受ける露地栽培の農産物を比較してみる
- どこの卸売市場での出荷量（価格）を予測するのか
  例えば、
  収穫と消費が比較的狭いエリア内で完結する地方卸売市場と多くの生産地から集荷して消費する中央卸売市場の２つを比較してみる
- いつの時期の出荷量（価格）を予測するのか
  日本列島は南北に長く、山地が多く同じ農産物でも時期や作型によって栽培時期が変わります。予測する農産物のどの時期の収穫量（価格）を予測するか決める必要があります。

7

予測対象の決定

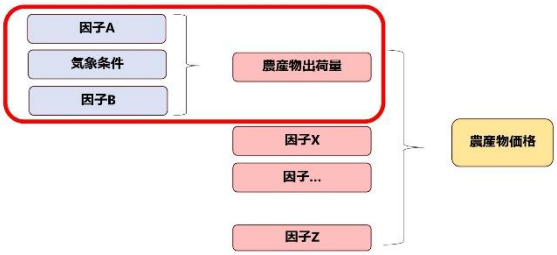
Step2　収集データの検討（因子分析-1）

- 農産物の出荷量に影響を与える可能性のある因子を抽出する
- 抽出したデータの信頼性、収集可能性を確認する
- 抽出された因子から今回の価格予測モデルに適用するデータを決定する

8

予測対象の決定

目標は「農産物価格の予測」であるが、かなり複雑に因子が絡み合っていそう（要検証）
基本的には価格は需要と供給のバランスによって決定されるので、価格決定因子の中で最も影響力のある「農産物の出荷量」を予測し、その後、農産物価格を予測する



9

科目名：IoTとデータ活用Ⅱ
受講者数：7名
期間：R7年10月～R8年1月　16コマ
概要：IoT基礎知識、データ分析スキルを学ぶ。特にⅡでは、IoT収集データを用いたデータ分析演習を行っていく。
▼教材

予測対象の決定

Step３　予測対象テーマの最終決定

いつからいつまでのどのようなデータを用いて分析し、どの卸売市場のどの農産物のいつの出荷量を予測するのかを決定する

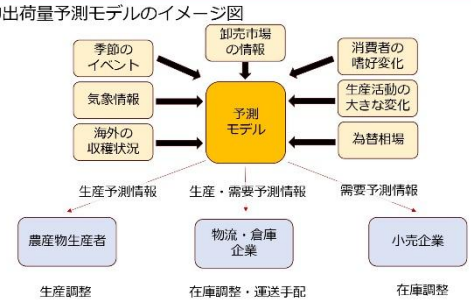
- 対象とする卸売市場
- 対象とする農産物と生産地
- 利活用データの収集期間
- 出荷量の予測時期

基本的に2023年度までのデータは存在するので、2022年度までのデータを収集し2023年度のデータを予測し、実値と比較して検証する。
（時間的に可能なら、プレゼン発表時には2023年度までのデータを利用した予測モデルを生成し、2024年度を予測し実値と比較検証する）

10

予測モデルのイメージ

農産物出荷量予測モデルのイメージ図

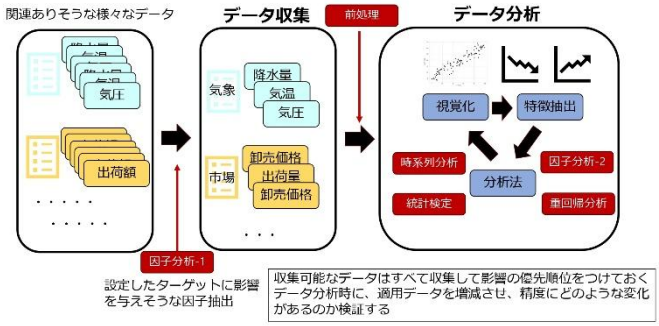


廃棄ロスや販売機会ロス、運送ロスの軽減に寄与する

11

データ分析の進め方

データ分析の進め方イメージ



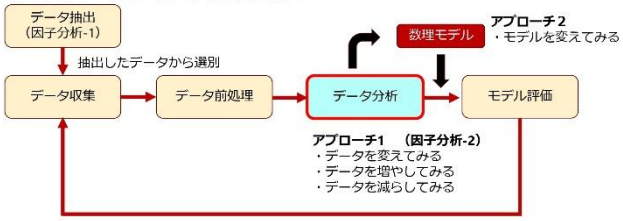
収集可能なデータはすべて収集して影響の優先順位をつけておくデータ分析時に、適用データを増減させ、精度にどのような変化があるのか検証する

12

データ分析の進め方

データ分析の進め方

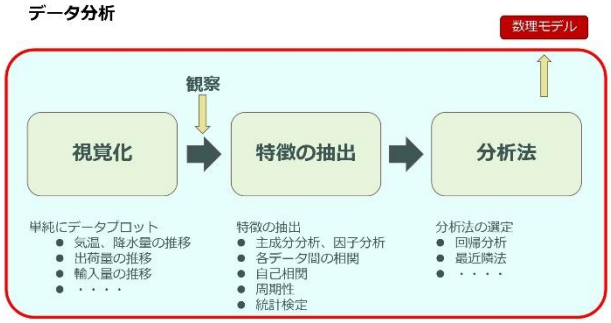
データ収集以降のフェーズでは、データの収集、数理モデルの決定、データの分析、モデル評価を行い、精度を上げていくために下記の手順を繰り返します



13

データ分析の進め方　データ分析

データ分析



14

データ分析の思考プロセス

①　分析目的の明確化

現状の課題を明確にすることにより、「なぜデータ分析を行うのか？」といった方針を決めます。
仮に課題が明確でない場合でも、「分析によって何をしたいのか」を明らかにしておく必要があります。
あくまで、データ分析という作業は「課題の解決」のための手段に過ぎないことを理解しておきましょう

なぜお米（野菜）の価格が上がっているのだろう？　分析して何が原因なのか確かめてみたい

②　仮説の洗い出し

目的の計画化では「何をしたいのか」をもとにしていましたが、このステップでは「何を知りたいのか」を突き詰めていきます。最終的には、「真偽はともかくとして、最も確からしい答え」、つまり仮説を洗い出し、その検証のためにデータ分析を行います。仮説を立ててデータ分析を行うことで、より精度の高い結果を期待することができます。

収穫量が減ったから？　お米（野菜）を食べる人が増えたから？

15

データ分析の思考プロセス

③　データの収集

「分析目的」や「仮説」を考慮しつつ、必要なデータ収集を行っていきます。収集対象やデータ収集期間や時期、データ形式なども考えておく必要があります。

「20年分のコメ（野菜）の収穫量および需要の推移」、「収穫量変化の要因（例えば気象変化、農業従事者の変化）」、「コメ需要の変化（外国人観光客の増加による需要の増加や日本人の健康志向による和食回帰）」

④　データ前処理

ビッグデータの中身を確認してみると、形式が異なるものや重複したもの、外れ値などが含まれている可能性があります。そのままデータ分析を進めてしまうと、集計や分析に時間を要したり、誤った結果を導くことにもなりかねません。そのため、データ分析に入る前に、ある程度データの前処理を行うことが必要となります（データクレンジング）

16

データ分析の思考プロセス

⑤　データ分析および分析結果の実ビジネスへの反映

得られた分析結果について、「なぜその結果になったのか」を検証し、真の課題を突き止めて「その対策をどのように講じていくのか」といったことを考え、実際に行動を起こしていってこそ、真にデータ分析という作業が実を結ぶことになります。

「作付面積を増やす」「農業従事者数を増やす」「大規模化、スマート化により生産性を上げる」「温暖化に耐えうる品種改良を行う」「輸入米を増やす」等

17