

データ分析における ChatGPTの実践的活用と AIのブラックボックス性への対応

中部大学 生命健康科学部 生命医科学科
AI数理データサイエンスセンター
新谷正嶺

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

1

本日の構成

前半(メイン)

2023年10月2日に行った高校生向けのAI授業

後半(+ α)

研究者向けのAI活用の紹介

※前半も後半も、「今までAIを活用したことのない方」を対象として想定しています。

大学全体のAIリテラシー向上を目指して、AI数理の実践を通して一人でも活用できる人を増やすことが目的です。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

2

それでは早速本日のメインである
前半

「2023年10月2日に行った高校生向けのAI授業」
の発表を始めます。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

3

授業の目標

現在AIの活用は一般にも広く知れ渡り注目を集めています。
しかし、実際にAIの特性を理解し使いこなしている人はどれほどいる
でしょうか？

本講義では、AIの中でも今話題のChatGPTを例として挙げ、実際に皆さんに解析を行って頂き、AIの持つメリットやデメリットを実感してもらいます。

その後、特にAIの持つデメリットであるブラックボックス性の危険性について考えます。そこでこのデメリットを排除する為、私が提唱したAI2L(AI to Learn : AIを自身の学習に使う)という指針を紹介します。これにより、AIの正しい使用方法の理解を深めることを目的とします。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

4

授業の概要項目

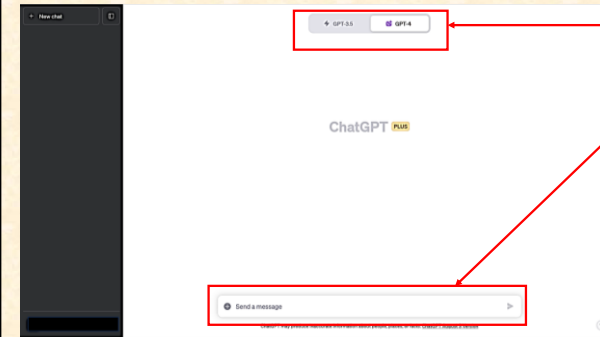
- ① ChatGPTの使い方
- ② ChatGPTの使用例
- ③ データ解析におけるChatGPTの実践
～例題にて練習～
- ④ データ解析の練習
～データ解析の基本であるExcelも練習～
- ⑤ ChatGPTにてデータ解析の練習
～Excelの練習にて使用した問題を使用～
- ⑥ ChatGPT使用時の注意点
- ⑦ AIを使いこなすための新たな指針(AI2L)
- ⑧ まとめ

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

5

① ChatGPTの使い方

ChatGPTの基本的な使い方を説明します。



まず、「GPT-4」を選択します。

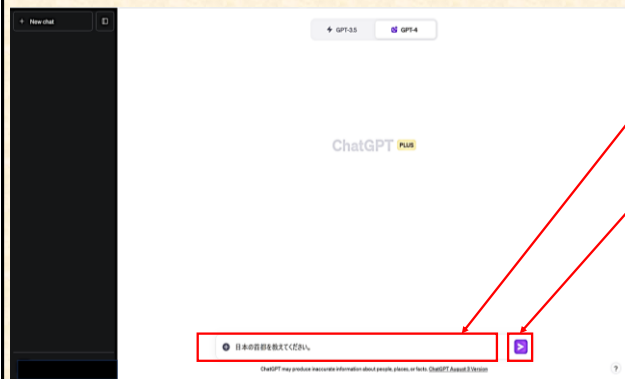
次に、この欄に質問したいことや、命令したいことを入力します。

早速次のスライドで簡単な練習をしてみましょう。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

6

① ChatGPTの使い方



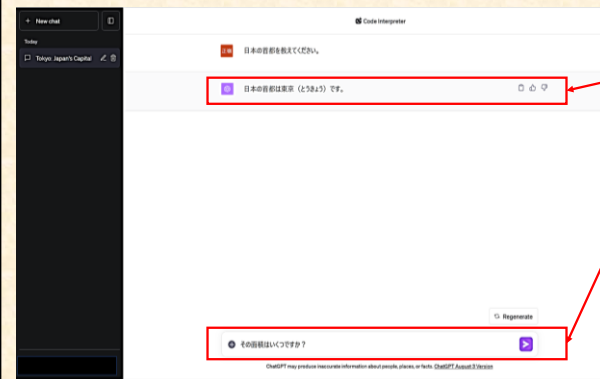
質問内容を入力します。今回は例として「日本の首都」について質問してみます。

質問内容を入力したら、こちらの実行ボタンを押します。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

7

① ChatGPTの使い方



質問に対する回答がこちらに表示されます。

ChatGPTは、前の会話の流れを踏まえて質問を続けることができるので、続けて質問してみます。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

8

① ChatGPTの使い方

追加の質問と、それに対する回答が得られます。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved 9

② ChatGPTの使用例

先ほどはChatGPTの基本的な使用方法について学びました。

では、社会では具体的にどのような使用例が挙げられるでしょうか？

有名な使用例は、「日本語→英語」の翻訳とその解説です。

次ページにて説明します。

質問

回答

「私は日本人です。ピアノを弾くのが好きです。あなたはどこの国から来ましたか？」という文章を英語に翻訳し、その翻訳の文法的な解説をした上で、再度日本語に逆翻訳するという使用例があります。

これまでネット上の情報は英語が主流で、日本語の情報の数十倍といわれており、情報収集において、言語による障壁が大きな問題でした。

しかし、ChatGPTの翻訳、解説機能により、英語の学習教材としてのみならず、アクセスしたい情報や英語の情報を簡単に翻訳し、文法的なチェックを行うことがより簡単になりました。詳細を次ページに示します。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved 11

ChatGPTの回答詳細

了解しました。まず、提供された日本語の文章を英語に翻訳します。

私は日本人です。
ピアノを弾くのが好きです。
あなたはどこの国から来ましたか？

翻訳: **日本語から英語への翻訳**

I am Japanese.
I like playing the piano.
Where are you from?

文法的な説明:
"I am" は主語と動詞を組み合わせたもので、自分の国籍や職業などの自己紹介に使います。"Japanese" は日本人を意味する形容詞です。
"I like" は好きであることを表す表現です。"playing the piano" は動名詞の形を取っており、ピアノを弾く行為を指します。
"Where are you from?" は相手の出身地や国籍を尋ねる一般的な表現です。

次に、上記の英語を再度日本語に逆翻訳します。

再度日本語に逆翻訳した結果は以下の通りです。

文法的な解説

私は日本人です。
ピアノを弾くのが好きです。
あなたはどこの国から来ましたか？

再度英語から日本語へ逆翻訳し、元の文章との違いを検証

この結果を見ると、元の日本語の文章と同じ内容になっています。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved 12

③データ解析におけるChatGPTの実践 ～例題にて練習～

前項ではChatGPTの実用例について紹介しました。

現在、翻訳・校正の分野でChatGPTは広く活用されています。

ただし、本講義では、ChatGPTの「解析」における活用方法を学習します。会社や学校など各事業体で、様々なデータを取り扱う際にもChatGPTは有効に使用することができます。自身の仕事や研究で長年培った知見と併せて、ChatGPTを活用することで、さらに深い解析を可能にします。受講生の皆さんの持つ経験や専門性に、ChatGPTという「AI」を上手く活用することにより、「鬼に金棒」の「金棒」の部分を提供する授業を目指します。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

13

④データ解析の練習 ～データ解析の基本であるExcelも練習～

今後、皆さんが大学に進学したり、会社に就職した際に、データ解析の能力はとて役に立ちます。

では、具体的にどのような内容で練習するのが良いでしょうか？
本講義では、皆さんにも馴染みのあるコンビニを例として使用します。

【設定】

あなたは、あるコンビニの店長です。
より売上を上げるためにはどのような商品を積極的に仕入れれば良いか？何をすればいいのか？について悩んでいます。
まずは、現状を理解するための解析を行ってみましょう。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

14

④データ解析の練習 ～データ解析の基本であるExcelも練習～

コンビニには様々な商品が売られています。
また、同じ「酒」「たばこ」「雑誌」でもメーカーが多く種類も様々です。

今回はそのように複雑なデータから現状を理解するための解析として何が出来るのかを学んでいきます。

次ページにてあなたのコンビニの売上データを提示しますので、一度じっくりと見てどのようなことが分かりそうか考えてみてください。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

15

～例題にて練習～ コンビニの売上データ

購入日時	購入額 (円)	性別	年齢層	天気	購入商品
2023-07-01 08:00:00	1176	男性	50代以上	晴	雑誌、タバコ
2023-07-01 08:30:00	2	64歳 男性	50代以上	晴	雑誌、スタンプ
2023-07-01 10:00:00	2	2000 男性	50代以上	晴	雑誌、傘、タバコ
2023-07-01 16:30:00	4	102 男性	20代	晴	アイスクリーム
2023-07-01 16:30:00	5	800 男性	20代	晴	雑誌、お菓子
2023-07-01 17:00:00	6	1011 女性	20代	晴	雑誌、インスタント食品
2023-07-01 17:00:00	7	1350 女性	40代	晴	アイスクリーム、お菓子
2023-07-01 18:00:00	8	126 男性	20代	晴	お菓子
2023-07-01 18:30:00	9	130 男性	50代以上	晴	タバコ
2023-07-01 08:30:00	10	1913 男性	10代	晴	雑誌、おにぎり、雑誌、アイスクリーム
2023-07-01 14:30:00	11	1606 男性	20代	晴	タバコ、アイスクリーム、おにぎり、スタンプ、雑誌
2023-07-01 22:00:00	12	509 女性	50代以上	晴	アイスクリーム、おにぎり
2023-07-01 08:15:00	13	953 男性	20代	晴	タバコ
2023-07-01 08:30:00	14	1180 男性	20代	晴	雑誌、インスタント食品、お菓子
2023-07-01 13:30:00	15	648 男性	50代以上	晴	雑誌
2023-07-01 10:23:00	16	780 女性	20代	晴	雑誌
2023-07-01 08:30:00	17	1058 男性	40代	晴	雑誌、アイスクリーム、お菓子、おにぎり
2023-07-01 08:30:00	18	925 女性	40代	晴	インスタント食品、お菓子、スタンプ、お菓子
2023-07-01 08:30:00	19	1256 男性	40代	晴	アイスクリーム、お菓子
2023-07-01 13:34:00	20	644 男性	50代以上	晴	雑誌
2023-07-01 18:00:00	21	732 女性	40代	晴	アイスクリーム、お菓子
2023-07-01 18:00:00	22	764 男性	10代	晴	スタンプ、タバコ
2023-07-01 18:00:00	23	1606 男性	50代以上	晴	雑誌、おにぎり
2023-07-01 22:00:00	24	1110 女性	20代	晴	雑誌、タバコ
2023-07-01 14:30:00	25	1014 男性	50代以上	晴	雑誌
2023-07-01 10:00:00	26	425 女性	40代	晴	雑誌、インスタント食品、スタンプ
2023-07-01 20:18:00	27	1640 男性	20代	晴	雑誌、アイスクリーム
2023-07-01 22:30:00	28	700 女性	40代	晴	雑誌
2023-07-01 09:30:00	29	803 女性	40代	晴	雑誌、アイスクリーム
2023-07-01 17:00:00	30	130 男性	50代	晴	スタンプ、アイスクリーム、タバコ

2023年7月1日～31日のコンビニの売り上げデータを使用します。(Excel形式)

- ・購入日時
- ・顧客ID
- ・購入額
- ・性別
- ・年齢層
- ・天気
- ・購入商品

上記のデータが入力されています。

まずは、データの確認を行いましょ。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

16

コンビニの売上データ:Excel内容の確認

購入日時	顧客ID	購入額 (¥)	性別	年齢層	天気	購入商品
2023-07-01 04:54:00	1	1378	女性	50代以上	曇り	雑誌2、タバコ2
2023-07-01 04:36:00	2	944	男性	50代以上	曇り	雑誌2、スナック5
2023-07-01 10:46:00	3	2080	男性	10代	曇り	雑誌3、弁当2、タバコ4
2023-07-01 16:19:00	4	102	男性	20代	曇り	アイスクリーム4
2023-07-01 16:13:00	5	806	女性	20代	曇り	タバコ5、お茶4
2023-07-01 17:40:00	6	1011	女性	30代	曇り	雑誌4、インスタント食品5
2023-07-01 15:16:00	7	1385	女性	40代	曇り	アイスクリーム1、弁当5、酒1
2023-07-01 10:02:00	8	126	男性	20代	曇り	お茶5
2023-07-01 05:21:00	9	159	男性	50代以上	曇り	スナック5
2023-07-01 08:26:00	10	1973	男性	10代	曇り	酒5、おにぎり5、雑誌1、アイスクリーム5
2023-07-01 14:39:00	11	1496	男性	10代	曇り	タバコ1、アイスクリーム3、スナック4、酒4

本講義では一般的にデータ解析でよく使用されるExcel形式のデータを用います。

先ほどのデータの内容を、一部を抜き出して確認してみましょう。

どのぐらいの年齢層の人(男性or女性)がいつ、どんな天気の時に行ったのかを知ることができます。

※データ情報の補足事項:

- ・「雑誌1」や「雑誌2」は雑誌を買った個数ではなく、雑誌の種類を示しています。
- ・今回のデータでは練習用として簡単にするために、同じ商品を複数買うことはなく、一つの商品を一個だけ購入した場合のデータになっています。よって、複数の別の商品を同時に購入することはあっても、各商品の購入個数は1個です。つまり、「雑誌2、タバコ2」の場合、「雑誌2」を1個、「タバコ2」を1個買ったことになります。

もし、このデータを「解析して分かったこと」を報告するように依頼されたら、どのような項目に注目して進めていけばいいのでしょうか？

コンビニの売上データ:Excelの解析

購入日時	顧客ID	購入額 (¥)	性別	年齢層	天気	購入商品
2023-07-01 04:54:00	1	1378	女性	50代以上	曇り	雑誌2、タバコ2
2023-07-01 04:36:00	2	944	男性	50代以上	曇り	雑誌2、スナック5
2023-07-01 10:46:00	3	2080	男性	10代	曇り	雑誌3、弁当2、タバコ4
2023-07-01 16:19:00	4	102	男性	20代	曇り	アイスクリーム4
2023-07-01 16:13:00	5	806	女性	20代	曇り	タバコ5、お茶4
2023-07-01 17:40:00	6	1011	女性	30代	曇り	雑誌4、インスタント食品5
2023-07-01 15:16:00	7	1385	女性	40代	曇り	アイスクリーム1、弁当5、酒1
2023-07-01 10:02:00	8	126	男性	20代	曇り	お茶5
2023-07-01 05:21:00	9	159	男性	50代以上	曇り	スナック5
2023-07-01 08:26:00	10	1973	男性	10代	曇り	酒5、おにぎり5、雑誌1、アイスクリーム5
2023-07-01 14:39:00	11	1496	男性	10代	曇り	タバコ1、アイスクリーム3、スナック4、酒4

本講義では、単にChatGPTの使用方法をレクチャーするだけでなく、データ解析の考え方についても触れていきます。まずは、渡されたデータを理解し、どのようなことが分かるかを想像して解析してみましょう。

・お題: このデータを「解析して分かったこと」を報告する。

どのような項目に注目して進めていけばいいのでしょうか？ 着目すべき点について、一度自分自身で考えてみてください。

コンビニの売上データ:Excelの解析

購入日時	顧客ID	購入額 (¥)	性別	年齢層	天気	購入商品
2023-07-01 04:54:00	1	1378	女性	50代以上	曇り	雑誌2、タバコ2
2023-07-01 04:36:00	2	944	男性	50代以上	曇り	雑誌2、スナック5
2023-07-01 10:46:00	3	2080	男性	10代	曇り	雑誌3、弁当2、タバコ4
2023-07-01 16:19:00	4	102	男性	20代	曇り	アイスクリーム4
2023-07-01 16:13:00	5	806	女性	20代	曇り	タバコ5、お茶4
2023-07-01 17:40:00	6	1011	女性	30代	曇り	雑誌4、インスタント食品5
2023-07-01 15:16:00	7	1385	女性	40代	曇り	アイスクリーム1、弁当5、酒1
2023-07-01 10:02:00	8	126	男性	20代	曇り	お茶5
2023-07-01 05:21:00	9	159	男性	50代以上	曇り	スナック5
2023-07-01 08:26:00	10	1973	男性	10代	曇り	酒5、おにぎり5、雑誌1、アイスクリーム5
2023-07-01 14:39:00	11	1496	男性	10代	曇り	タバコ1、アイスクリーム3、スナック4、酒4

このデータを眺めて、あなたはどうなことを調べてみたいと考えたでしょうか？

「解析」は自由な発想で切り口を見つけることが大切です。

・お題: このデータを「解析して分かったこと」を報告する。

あなたはどのようなことを調べたいと考えましたか？

以下に、解析の切り口の例を挙げます。

例)

- ①一番売れている商品をランキング形式で調べる。
- ②性別や年代別に売上金額の違いを調べる。
- ③天気によってよく売れる商品がないか調べる。

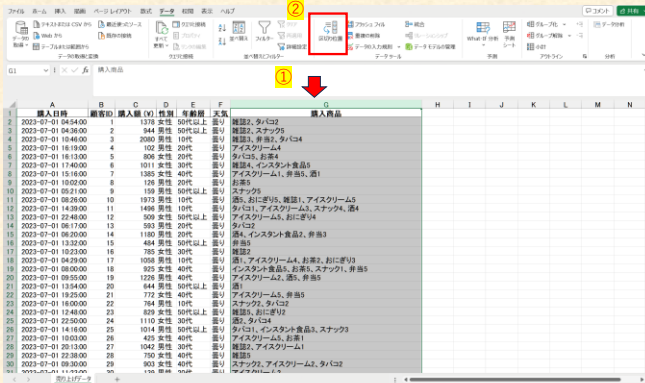
今回は簡単な練習として例①の「一番売れている商品をランキング形式で調べる」に取り組んでみましょう。

コンビニの売上データ:Excelの解析 売れ筋のランキング調査

それでは、Excelデータに戻って、「一番売れている商品をランキング形式で調べる」ために、従来の解析方法を具体的に学びましょう。

購入商品の欄を見ると、1人当たりの購入商品が1つだけとは限らず、複数の場合もあることが分かります。今回のデータでは、一人当たり1~4個までの商品を購入していました。各商品をカウントするためには1つのセル(枠)の中のデータを複数ではなく、一つずつに分ける必要があります。次ページにて購入商品の区分けを行います。

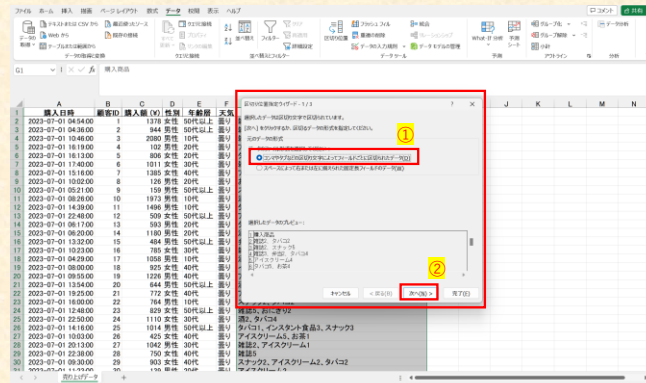
コンビニの売上データ:Excelの解析 売れ筋のランキング調査



まず、①のように、「購入商品」の列を選択します。

そして、②の「区切り位置」のボタンを押します。

コンビニの売上データ:Excelの解析 売れ筋のランキング調査

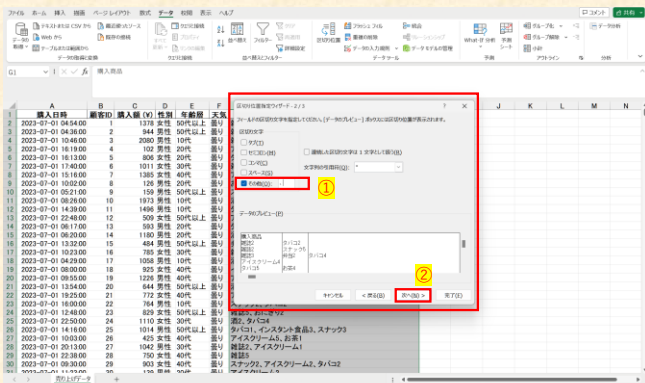


「区切り位置」のボタンを押すとこのような画面が出てきます。

これは、セル(枠)の中に複数のデータが含まれるときに、それを別々のデータに分けるのに便利な機能です。

①の「コンマやタブなどの区切り文字によってフィールドごとに区切られたデータ」を選択し、②の「次へ」のボタンを押します。

コンビニの売上データ:Excelの解析 売れ筋のランキング調査

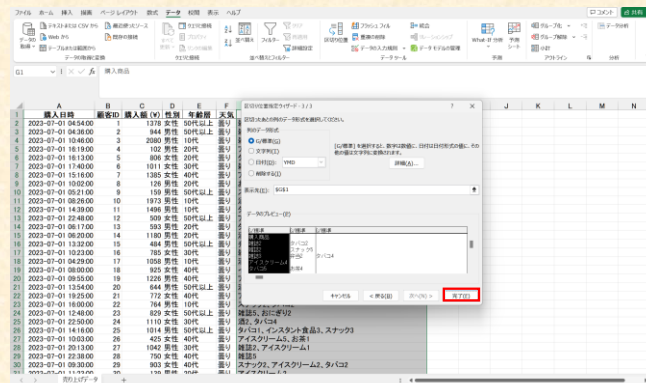


次の画面が出てきます。

今回は、「雑誌2、タバコ2」のように複数の商品が「」で区切られていたデータなので区切り文字を「その他」の欄で①の「」に指定します。

②の「次へ」のボタンを押します。

コンビニの売上データ:Excelの解析 売れ筋のランキング調査



次の画面が出てきます。

この画面では何も設定せずにそのまま「次へ」のボタンを押します。

コンビニの売上データ:Excelの解析 売れ筋のランキング調査

購入日時	販売印	購入人数 (V)	性別	年齢層	売上	購入商品
2023-07-01 04:54:00	1	1378	男性	50代以上	買戻	タバコ
2023-07-01 04:30:00	2	844	男性	50代以上	買戻	スタンプ
2023-07-01 04:30:00	3	2000	男性	10代	買戻	スタンプ
2023-07-01 16:19:00	4	102	男性	20代	買戻	アイスクリーム
2023-07-01 16:13:00	5	800	女性	20代	買戻	タバコ
2023-07-01 17:40:00	6	1011	女性	30代	買戻	お茶
2023-07-01 15:18:00	7	1385	男性	40代	買戻	アイスクリーム
2023-07-01 10:20:00	8	126	男性	20代	買戻	お茶
2023-07-01 09:21:00	9	159	男性	50代以上	買戻	スタンプ
2023-07-01 08:28:00	10	1973	男性	10代	買戻	アイスクリーム
2023-07-01 14:39:00	11	1486	男性	10代	買戻	アイスクリーム
2023-07-01 22:48:00	12	500	男性	50代以上	買戻	アイスクリーム
2023-07-01 08:17:00	13	583	男性	20代	買戻	タバコ
2023-07-01 09:20:00	14	1100	男性	20代	買戻	アイスクリーム
2023-07-01 13:32:00	15	484	男性	50代以上	買戻	お茶
2023-07-01 10:23:00	16	785	男性	30代	買戻	アイスクリーム
2023-07-01 04:29:00	17	1058	男性	10代	買戻	アイスクリーム
2023-07-01 08:00:00	18	925	男性	40代	買戻	アイスクリーム
2023-07-01 09:55:00	19	1228	男性	40代	買戻	アイスクリーム
2023-07-01 13:54:00	20	844	男性	50代以上	買戻	お茶
2023-07-01 19:25:00	21	772	女性	40代	買戻	アイスクリーム
2023-07-01 16:00:00	22	784	男性	10代	買戻	スタンプ
2023-07-01 12:48:00	23	829	男性	50代以上	買戻	アイスクリーム
2023-07-01 22:50:00	24	1110	女性	30代	買戻	アイスクリーム
2023-07-01 14:18:00	25	1014	男性	30代	買戻	アイスクリーム
2023-07-01 10:53:00	26	425	女性	40代	買戻	アイスクリーム
2023-07-01 20:13:00	27	1042	男性	30代	買戻	アイスクリーム
2023-07-01 22:38:00	28	750	女性	40代	買戻	アイスクリーム
2023-07-01 09:20:00	29	803	男性	40代	買戻	アイスクリーム
2023-07-01 11:47:00	30	150	男性	40代	買戻	アイスクリーム

完了すると、このような画面になり、1つのセル(枠)に複数記載されていた商品が、「」によって分解され、各商品が別々のセル(枠)に表示されます。

この状態にすれば各セル(枠)に入力されている商品は1つだけになるのでカウントするのが簡単になります。

ただ、この状態だと見づらいので各セル(枠)の幅を綺麗にそろえましょう。

コンビニの売上データ:Excelの解析 売れ筋のランキング調査

購入日時	販売印	購入人数 (V)	性別	年齢層	売上	購入商品
2023-07-01 04:54:00	1	1378	男性	50代以上	買戻	タバコ
2023-07-01 04:30:00	2	844	男性	50代以上	買戻	スタンプ
2023-07-01 04:30:00	3	2000	男性	10代	買戻	スタンプ
2023-07-01 16:19:00	4	102	男性	20代	買戻	アイスクリーム
2023-07-01 16:13:00	5	800	女性	20代	買戻	タバコ
2023-07-01 17:40:00	6	1011	女性	30代	買戻	お茶
2023-07-01 15:18:00	7	1385	男性	40代	買戻	アイスクリーム
2023-07-01 10:20:00	8	126	男性	20代	買戻	お茶
2023-07-01 09:21:00	9	159	男性	50代以上	買戻	スタンプ
2023-07-01 08:28:00	10	1973	男性	10代	買戻	アイスクリーム
2023-07-01 14:39:00	11	1486	男性	10代	買戻	アイスクリーム
2023-07-01 22:48:00	12	500	男性	50代以上	買戻	アイスクリーム
2023-07-01 08:17:00	13	583	男性	20代	買戻	タバコ
2023-07-01 09:20:00	14	1100	男性	20代	買戻	アイスクリーム
2023-07-01 13:32:00	15	484	男性	50代以上	買戻	お茶
2023-07-01 10:23:00	16	785	男性	30代	買戻	アイスクリーム
2023-07-01 04:29:00	17	1058	男性	10代	買戻	アイスクリーム
2023-07-01 08:00:00	18	925	男性	40代	買戻	アイスクリーム
2023-07-01 09:55:00	19	1228	男性	40代	買戻	アイスクリーム
2023-07-01 13:54:00	20	844	男性	50代以上	買戻	お茶
2023-07-01 19:25:00	21	772	女性	40代	買戻	アイスクリーム
2023-07-01 16:00:00	22	784	男性	10代	買戻	スタンプ
2023-07-01 12:48:00	23	829	男性	50代以上	買戻	アイスクリーム
2023-07-01 22:50:00	24	1110	女性	30代	買戻	アイスクリーム
2023-07-01 14:18:00	25	1014	男性	30代	買戻	アイスクリーム
2023-07-01 10:53:00	26	425	女性	40代	買戻	アイスクリーム
2023-07-01 20:13:00	27	1042	男性	30代	買戻	アイスクリーム
2023-07-01 22:38:00	28	750	女性	40代	買戻	アイスクリーム
2023-07-01 09:20:00	29	803	男性	40代	買戻	アイスクリーム
2023-07-01 11:47:00	30	150	男性	40代	買戻	アイスクリーム

今回のデータは最大でも1人当たり4個種類までしか購入していないので左図のようにG~J列までの4列を選択します。

①として、カーソルをJ列の横に合わせてダブルクリックをします。

コンビニの売上データ:Excelの解析 売れ筋のランキング調査

購入日時	販売印	購入人数 (V)	性別	年齢層	売上	購入商品
2023-07-01 04:54:00	1	1378	男性	50代以上	買戻	タバコ
2023-07-01 04:30:00	2	844	男性	50代以上	買戻	スタンプ
2023-07-01 04:30:00	3	2000	男性	10代	買戻	スタンプ
2023-07-01 16:19:00	4	102	男性	20代	買戻	アイスクリーム
2023-07-01 16:13:00	5	800	女性	20代	買戻	タバコ
2023-07-01 17:40:00	6	1011	女性	30代	買戻	お茶
2023-07-01 15:18:00	7	1385	男性	40代	買戻	アイスクリーム
2023-07-01 10:20:00	8	126	男性	20代	買戻	お茶
2023-07-01 09:21:00	9	159	男性	50代以上	買戻	スタンプ
2023-07-01 08:28:00	10	1973	男性	10代	買戻	アイスクリーム
2023-07-01 14:39:00	11	1486	男性	10代	買戻	アイスクリーム
2023-07-01 22:48:00	12	500	男性	50代以上	買戻	アイスクリーム
2023-07-01 08:17:00	13	583	男性	20代	買戻	タバコ
2023-07-01 09:20:00	14	1100	男性	20代	買戻	アイスクリーム
2023-07-01 13:32:00	15	484	男性	50代以上	買戻	お茶
2023-07-01 10:23:00	16	785	男性	30代	買戻	アイスクリーム
2023-07-01 04:29:00	17	1058	男性	10代	買戻	アイスクリーム
2023-07-01 08:00:00	18	925	男性	40代	買戻	アイスクリーム
2023-07-01 09:55:00	19	1228	男性	40代	買戻	アイスクリーム
2023-07-01 13:54:00	20	844	男性	50代以上	買戻	お茶
2023-07-01 19:25:00	21	772	女性	40代	買戻	アイスクリーム
2023-07-01 16:00:00	22	784	男性	10代	買戻	スタンプ
2023-07-01 12:48:00	23	829	男性	50代以上	買戻	アイスクリーム
2023-07-01 22:50:00	24	1110	女性	30代	買戻	アイスクリーム
2023-07-01 14:18:00	25	1014	男性	30代	買戻	アイスクリーム
2023-07-01 10:53:00	26	425	女性	40代	買戻	アイスクリーム
2023-07-01 20:13:00	27	1042	男性	30代	買戻	アイスクリーム
2023-07-01 22:38:00	28	750	女性	40代	買戻	アイスクリーム
2023-07-01 09:20:00	29	803	男性	40代	買戻	アイスクリーム
2023-07-01 11:47:00	30	150	男性	40代	買戻	アイスクリーム

すると、左図のように、各セル(枠)の幅が綺麗に整います。

今回のデータは1人当たり1~4つの商品を購入しているデータです。

購入した商品は左のセル(枠)から順に表示されています。また、商品の後ろに表記されている数字は個数ではなく、商品の種類です。各商品1~5種類あります。例えば「雑誌2」の場合は雑誌を2冊買ったという意味ではなく、雑誌の「2」という種類を1冊買ったことを意味します。

コンビニの売上データ:Excelの解析 売れ筋のランキング調査

購入日時	販売印	購入人数 (V)	性別	年齢層	売上	購入商品1	購入商品2	購入商品3	購入商品4
2023-07-01 04:54:00	1	1378	男性	50代以上	買戻	タバコ			
2023-07-01 04:30:00	2	844	男性	50代以上	買戻	スタンプ			
2023-07-01 04:30:00	3	2000	男性	10代	買戻	スタンプ			
2023-07-01 16:19:00	4	102	男性	20代	買戻	アイスクリーム			
2023-07-01 16:13:00	5	800	女性	20代	買戻	タバコ			
2023-07-01 17:40:00	6	1011	女性	30代	買戻	お茶			
2023-07-01 15:18:00	7	1385	男性	40代	買戻	アイスクリーム			
2023-07-01 10:20:00	8	126	男性	20代	買戻	お茶			
2023-07-01 09:21:00	9	159	男性	50代以上	買戻	スタンプ			
2023-07-01 08:28:00	10	1973	男性	10代	買戻	アイスクリーム			
2023-07-01 14:39:00	11	1486	男性	10代	買戻	アイスクリーム			
2023-07-01 22:48:00	12	500	男性	50代以上	買戻	アイスクリーム			
2023-07-01 08:17:00	13	583	男性	20代	買戻	タバコ			
2023-07-01 09:20:00	14	1100	男性	20代	買戻	アイスクリーム			
2023-07-01 13:32:00	15	484	男性	50代以上	買戻	お茶			
2023-07-01 10:23:00	16	785	男性	30代	買戻	アイスクリーム			
2023-07-01 04:29:00	17	1058	男性	10代	買戻	アイスクリーム			
2023-07-01 08:00:00	18	925	男性	40代	買戻	アイスクリーム			
2023-07-01 09:55:00	19	1228	男性	40代	買戻	アイスクリーム			
2023-07-01 13:54:00	20	844	男性	50代以上	買戻	お茶			
2023-07-01 19:25:00	21	772	女性	40代	買戻	アイスクリーム			
2023-07-01 16:00:00	22	784	男性	10代	買戻	スタンプ			
2023-07-01 12:48:00	23	829	男性	50代以上	買戻	アイスクリーム			
2023-07-01 22:50:00	24	1110	女性	30代	買戻	アイスクリーム			
2023-07-01 14:18:00	25	1014	男性	30代	買戻	アイスクリーム			
2023-07-01 10:53:00	26	425	女性	40代	買戻	アイスクリーム			
2023-07-01 20:13:00	27	1042	男性	30代	買戻	アイスクリーム			
2023-07-01 22:38:00	28	750	女性	40代	買戻	アイスクリーム			
2023-07-01 09:20:00	29	803	男性	40代	買戻	アイスクリーム			
2023-07-01 11:47:00	30	150	男性	40代	買戻	アイスクリーム			

左図のように、各セル(枠)の幅が綺麗に整います。

早速、購入商品1~4の商品をカウントして一番売れている商品をランキング形式で解析してみましょう。

次ページから各列の商品をカウントする方法を提示します。今回は「ピボットテーブル」という方法でカウントを行います。「ピボットテーブル」は便利な方法なので是非覚えてください。

コンビニの売上データ:Excelの解析 売れ筋のランキング調査

購入日時	販売ID	購入年齢 (X)	性別	年齢層	天候	購入商品1	購入商品2	購入商品3	購入商品4
2023-07-01 04:36:00	1	1370	男性	50代以上	晴	アイスクリ-ム1	アイスクリ-ム2		
2023-07-01 04:36:00	2	844	男性	50代以上	曇り	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4
2023-07-01 10:46:00	3	2003	男性	10代	晴	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4
2023-07-01 16:18:00	4	102	男性	20代	曇り	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4
2023-07-01 16:18:00	5	808	男性	20代	曇り	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4
2023-07-01 17:40:00	6	1011	女性	30代	曇り	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4
2023-07-01 18:10:00	7	1385	女性	40代	曇り	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4
2023-07-01 18:10:00	8	138	男性	20代	曇り	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4
2023-07-01 08:21:00	9	109	男性	50代以上	曇り	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4
2023-07-01 08:21:00	10	1972	男性	10代	曇り	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4
2023-07-01 14:20:00	11	1486	男性	10代	曇り	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4
2023-07-01 22:40:00	12	509	男性	50代以上	曇り	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4
2023-07-01 08:17:00	13	583	男性	20代	曇り	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4
2023-07-01 08:20:00	14	1182	男性	20代	曇り	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4
2023-07-01 13:22:00	15	484	男性	50代以上	曇り	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4
2023-07-01 13:22:00	16	785	男性	30代	曇り	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4
2023-07-01 04:20:00	17	1058	男性	10代	曇り	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4
2023-07-01 08:00:00	18	925	男性	40代	曇り	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4
2023-07-01 08:58:00	19	1276	男性	40代	曇り	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4
2023-07-01 13:54:00	20	844	男性	50代以上	曇り	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4
2023-07-01 18:25:00	21	772	女性	40代	曇り	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4
2023-07-01 18:50:00	22	764	男性	10代	曇り	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4
2023-07-01 12:40:00	23	629	男性	50代以上	曇り	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4
2023-07-01 22:30:00	24	1110	男性	30代	曇り	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4
2023-07-01 16:18:00	25	1014	男性	50代以上	曇り	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4
2023-07-01 16:18:00	26	475	男性	40代	曇り	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4
2023-07-01 20:10:00	27	1542	男性	30代	曇り	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4
2023-07-01 22:30:00	28	750	男性	40代	曇り	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4
2023-07-01 08:20:00	29	802	男性	40代	曇り	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4
2023-07-01 08:20:00	30	844	男性	50代以上	曇り	アイスクリ-ム1	アイスクリ-ム2	アイスクリ-ム3	アイスクリ-ム4

購入商品をカウントするために「ピボットテーブル」を使用します。

まず左図のようにデータの記載されている範囲を全て選択してください。

そして、「挿入」メニューの中の「ピボットテーブル」を選択します。

コンビニの売上データ:Excelの解析 売れ筋のランキング調査

「ピボットテーブル」を選択すると、左図のような画面が出てきます。

そのまま、「OK」のボタンを押します。

コンビニの売上データ:Excelの解析 売れ筋のランキング調査

そうすると、左図のように、新しいシートにピボットテーブルの画面が出現します。

そして、右側にピボットテーブルを操作するためのメニュー画面が表示されます。

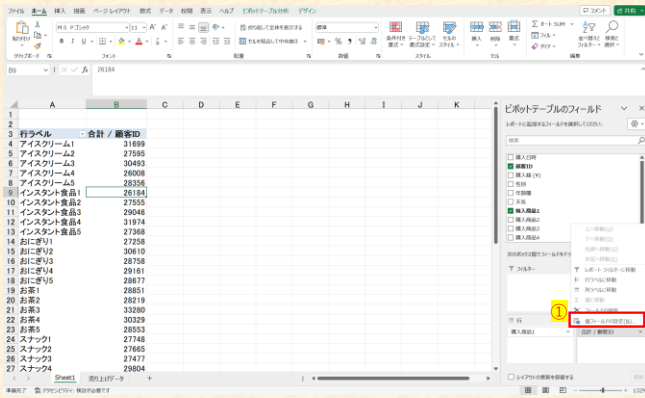
今回は、「購入商品1」～「購入商品4」の列に記載されている商品をカウントしたいので、各列に各商品がいくつあったかをそれぞれカウントしていきます。

コンビニの売上データ:Excelの解析 売れ筋のランキング調査

まず、購入商品1の商品の数をカウントするために、①「行」の部分に「購入商品1」をドラッグします。②次に、「値」の部分に「顧客ID」をドラッグします。

そうすると、左図のように各商品に対する、顧客IDの数量の合計が出てきます。しかし、ここで注意していただきたいのは、ここに出てきた合計の数字は「顧客IDのナンバー」をそのまま合計してしまっている点です。例えば顧客IDが「10」の人と「132」の人が購入したのであれば、購入数は2人の「2」にならなければなりません。このままでは、「10+132」の「142」と表記されてしまいます。

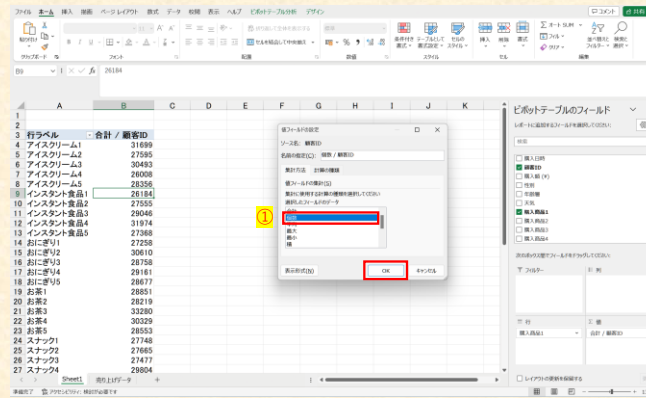
コンビニの売上データ:Excelの解析 売れ筋のランキング調査



そこで顧客IDのナンバーの合計ではなく、購入した顧客の人数の合計を計算するために、「顧客ID」の値の設定を変更します。

①「合計/顧客ID」をクリックし、「値フィールドの設定」のボタンを押します。

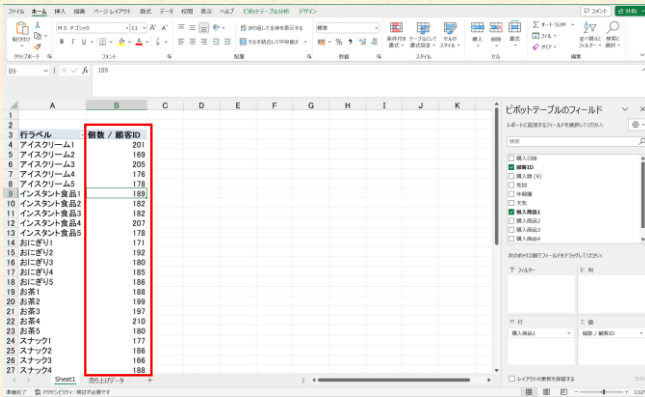
コンビニの売上データ:Excelの解析 売れ筋のランキング調査



そこで顧客IDのナンバーの合計ではなく、購入した顧客の人数の合計を計算するために、「顧客ID」の値の設定を変更します。

①設定を「合計」→「個数」に変更して「OK」を押します。

コンビニの売上データ:Excelの解析 売れ筋のランキング調査

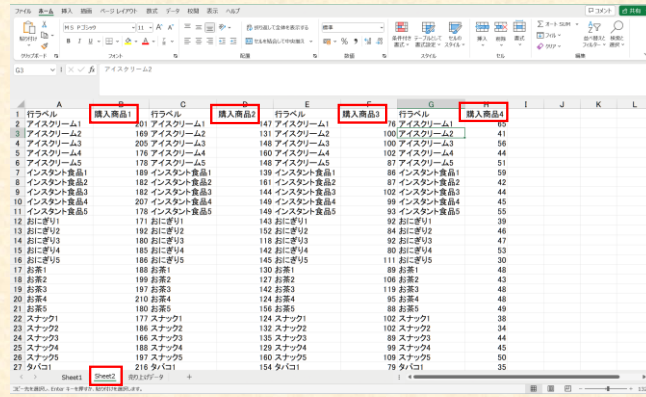


値の設定を変更したことにより、各商品を購入した人数の合計が表示されるようになりました。

今回のデータでは、購入客は違う商品を複数買うことはあっても、同じ商品を複数買うことはないため、購入人数をそのまま購入個数としてカウントして構いません。

得られた結果を別のシートにコピーして、残りの購入商品2,3,4についても同様の作業を繰り返します。

コンビニの売上データ:Excelの解析 売れ筋のランキング調査



左図のように別のシートを作成し、ピボットテーブルから得られたデータをコピーしました。また、「個数/顧客ID」という表記では分かりにくいので、各セル(特に「購入商品1~4」と記入しました。

ここで、「購入商品1~4」の商品を確認すると、全ての購入商品の欄に、「アイスクリーム1~5」のように、各商品1~5が抜けもれなくあるデータであることが分かります。今回のデータでは、偶然にも購入商品1~4に全ての商品が最低でも1回は購入されたため、例えば「スナック3」は無しという状況が発生しなかったようです。

コンビニの売上データ:Excelの解析 売れ筋のランキング調査

購入商品1	購入回数	購入商品1	購入回数	購入商品1	購入回数	購入商品1	購入回数
スタック4	188	スタック4	60	スタック4	45	スタック4	38
スタック24	187	スタック24	60	スタック24	60	スタック24	60
タバコ1	218	タバコ1	35	タバコ1	35	タバコ1	35
タバコ2	184	タバコ2	43	タバコ2	43	タバコ2	43
タバコ3	182	タバコ3	47	タバコ3	47	タバコ3	47
タバコ4	148	タバコ4	63	タバコ4	63	タバコ4	63
タバコ5	185	タバコ5	34	タバコ5	34	タバコ5	34
お菓1	145	お菓1	46	お菓1	46	お菓1	46
お菓2	171	お菓2	46	お菓2	46	お菓2	46
お菓3	192	お菓3	34	お菓3	34	お菓3	34
お菓4	203	お菓4	42	お菓4	42	お菓4	42
お菓5	184	お菓5	34	お菓5	34	お菓5	34
お菓6	180	お菓6	38	お菓6	38	お菓6	38
お菓7	148	お菓7	37	お菓7	37	お菓7	37
お菓8	184	お菓8	39	お菓8	39	お菓8	39
お菓9	212	お菓9	40	お菓9	40	お菓9	40
お菓10	171	お菓10	38	お菓10	38	お菓10	38
お菓11	186	お菓11	49	お菓11	49	お菓11	49
お菓12	184	お菓12	39	お菓12	39	お菓12	39
お菓13	182	お菓13	39	お菓13	39	お菓13	39
お菓14	189	お菓14	34	お菓14	34	お菓14	34
お菓15	182	お菓15	41	お菓15	41	お菓15	41
お菓16	8472	お菓16	6378	お菓16	6378	お菓16	6378
お菓17	8472	お菓17	6378	お菓17	6378	お菓17	6378

しかし、最後の方を見ると、「購入商品1」には「空白」の欄がありません。これは、購入客が最低でも1個の商品を購入しているため「購入商品1」に空白欄が生じなかったためです。

よって、データの並びをそろえるために、「購入商品1」にも「空白」の欄を作成し、「0」という数字を入力します。

コンビニの売上データ:Excelの解析 売れ筋のランキング調査

購入商品1	購入回数	購入商品1	購入回数	購入商品1	購入回数	購入商品1	購入回数
スタック4	188	スタック4	60	スタック4	45	スタック4	38
スタック24	187	スタック24	60	スタック24	60	スタック24	60
タバコ1	218	タバコ1	35	タバコ1	35	タバコ1	35
タバコ2	184	タバコ2	43	タバコ2	43	タバコ2	43
タバコ3	182	タバコ3	47	タバコ3	47	タバコ3	47
タバコ4	148	タバコ4	63	タバコ4	63	タバコ4	63
タバコ5	185	タバコ5	34	タバコ5	34	タバコ5	34
お菓1	145	お菓1	46	お菓1	46	お菓1	46
お菓2	171	お菓2	46	お菓2	46	お菓2	46
お菓3	192	お菓3	34	お菓3	34	お菓3	34
お菓4	203	お菓4	42	お菓4	42	お菓4	42
お菓5	184	お菓5	34	お菓5	34	お菓5	34
お菓6	180	お菓6	38	お菓6	38	お菓6	38
お菓7	148	お菓7	37	お菓7	37	お菓7	37
お菓8	184	お菓8	39	お菓8	39	お菓8	39
お菓9	212	お菓9	40	お菓9	40	お菓9	40
お菓10	171	お菓10	38	お菓10	38	お菓10	38
お菓11	186	お菓11	49	お菓11	49	お菓11	49
お菓12	184	お菓12	39	お菓12	39	お菓12	39
お菓13	182	お菓13	39	お菓13	39	お菓13	39
お菓14	189	お菓14	34	お菓14	34	お菓14	34
お菓15	182	お菓15	41	お菓15	41	お菓15	41
お菓16	8472	お菓16	6378	お菓16	6378	お菓16	6378
お菓17	8472	お菓17	6378	お菓17	6378	お菓17	6378

左図のように(空白)欄を作成し、「0」という数字を入力しました。

これにより、「購入商品1〜4」の各項目が完全に一致しました。

データ処理において、同じデータをずれなく表記することは思わぬ間違いを防ぐ大切な作業なのでよくチェックすることを癖づけてください。

コンビニの売上データ:Excelの解析 売れ筋のランキング調査

行ラベル	購入商品1	行ラベル	購入商品1	行ラベル	購入商品1	行ラベル	購入商品1	計算用行ラベル	購入回数合計
1 行ラベル		1 行ラベル		1 行ラベル		1 行ラベル			
2 アイスクリーム1	201	アイスクリーム1	147	アイスクリーム1	76	アイスクリーム1	1	アイスクリーム1+	425
3 アイスクリーム2	183	アイスクリーム2	131	アイスクリーム2	100	アイスクリーム2	2	アイスクリーム2+	414
4 アイスクリーム3	205	アイスクリーム3	148	アイスクリーム3	100	アイスクリーム3	58	アイスクリーム3+	451
5 アイスクリーム4	176	アイスクリーム4	160	アイスクリーム4	102	アイスクリーム4	44	アイスクリーム4+	478
6 アイスクリーム5	178	アイスクリーム5	148	アイスクリーム5	87	アイスクリーム5	51	アイスクリーム5+	462
7 インスタント食品1	182	インスタント食品1	139	インスタント食品1	88	インスタント食品1	59	インスタント食品1+	448
8 インスタント食品2	182	インスタント食品2	161	インスタント食品2	87	インスタント食品2	42	インスタント食品2+	472
9 インスタント食品3	182	インスタント食品3	144	インスタント食品3	102	インスタント食品3	44	インスタント食品3+	470
10 インスタント食品4	201	インスタント食品4	149	インスタント食品4	89	インスタント食品4	45	インスタント食品4+	485
11 インスタント食品5	178	インスタント食品5	149	インスタント食品5	93	インスタント食品5	55	インスタント食品5+	465
12 おにぎり1	171	おにぎり1	143	おにぎり1	92	おにぎり1	39	おにぎり1+	425
13 おにぎり2	182	おにぎり2	152	おにぎり2	84	おにぎり2	46	おにぎり2+	464
14 おにぎり3	180	おにぎり3	118	おにぎり3	92	おにぎり3	47	おにぎり3+	467
15 おにぎり4	185	おにぎり4	142	おにぎり4	80	おにぎり4	53	おにぎり4+	460
16 おにぎり5	188	おにぎり5	145	おにぎり5	111	おにぎり5	30	おにぎり5+	474
17 お茶1	188	お茶1	130	お茶1	89	お茶1	30	お茶1+	437
18 お茶2	199	お茶2	127	お茶2	106	お茶2	43	お茶2+	485
19 お茶3	197	お茶3	142	お茶3	119	お茶3	48	お茶3+	475
20 お茶4	210	お茶4	124	お茶4	85	お茶4	46	お茶4+	465
21 お茶5	180	お茶5	156	お茶5	88	お茶5	49	お茶5+	473
22 スナック1	177	スナック1	124	スナック1	102	スナック1	38	スナック1+	421
23 スナック2	185	スナック2	132	スナック2	102	スナック2	44	スナック2+	463
24 スナック3	168	スナック3	135	スナック3	88	スナック3	44	スナック3+	435
25 スナック4	188	スナック4	129	スナック4	89	スナック4	46	スナック4+	452
26 スナック5	181	スナック5	160	スナック5	109	スナック5	50	スナック5+	490
27 タバコ1	218	タバコ1	154	タバコ1	79	タバコ1	35	タバコ1+	486

それではよいよ各商品の購入回数を合計して一番売れている商品を調査しましょう。

左図の赤枠のように「計算用行ラベル」を追加し、「購入回数合計」の列を作ります。

そして、購入商品1〜4の数を合計する数式を入力します。

コンビニの売上データ:Excelの解析 売れ筋のランキング調査

行ラベル	購入商品1	行ラベル	購入商品1	行ラベル	購入商品1	行ラベル	購入商品1	計算用行ラベル	購入回数合計
1 行ラベル		1 行ラベル		1 行ラベル		1 行ラベル			
2 アイスクリーム1	201	アイスクリーム1	147	アイスクリーム1	76	アイスクリーム1	1	アイスクリーム1+	425
3 アイスクリーム2	183	アイスクリーム2	131	アイスクリーム2	100	アイスクリーム2	2	アイスクリーム2+	414
4 アイスクリーム3	205	アイスクリーム3	148	アイスクリーム3	100	アイスクリーム3	58	アイスクリーム3+	451
5 アイスクリーム4	176	アイスクリーム4	160	アイスクリーム4	102	アイスクリーム4	44	アイスクリーム4+	478
6 アイスクリーム5	178	アイスクリーム5	148	アイスクリーム5	87	アイスクリーム5	51	アイスクリーム5+	462
7 インスタント食品1	182	インスタント食品1	139	インスタント食品1	88	インスタント食品1	59	インスタント食品1+	448
8 インスタント食品2	182	インスタント食品2	161	インスタント食品2	87	インスタント食品2	42	インスタント食品2+	472
9 インスタント食品3	182	インスタント食品3	144	インスタント食品3	102	インスタント食品3	44	インスタント食品3+	470
10 インスタント食品4	201	インスタント食品4	149	インスタント食品4	89	インスタント食品4	45	インスタント食品4+	485
11 インスタント食品5	178	インスタント食品5	149	インスタント食品5	93	インスタント食品5	55	インスタント食品5+	465
12 おにぎり1	171	おにぎり1	143	おにぎり1	92	おにぎり1	39	おにぎり1+	425
13 おにぎり2	182	おにぎり2	152	おにぎり2	84	おにぎり2	46	おにぎり2+	464
14 おにぎり3	180	おにぎり3	118	おにぎり3	92	おにぎり3	47	おにぎり3+	467
15 おにぎり4	185	おにぎり4	142	おにぎり4	80	おにぎり4	53	おにぎり4+	460
16 おにぎり5	188	おにぎり5	145	おにぎり5	111	おにぎり5	30	おにぎり5+	474
17 お茶1	188	お茶1	130	お茶1	89	お茶1	30	お茶1+	437
18 お茶2	199	お茶2	127	お茶2	106	お茶2	43	お茶2+	485
19 お茶3	197	お茶3	142	お茶3	119	お茶3	48	お茶3+	475
20 お茶4	210	お茶4	124	お茶4	85	お茶4	46	お茶4+	465
21 お茶5	180	お茶5	156	お茶5	88	お茶5	49	お茶5+	473
22 スナック1	177	スナック1	124	スナック1	102	スナック1	38	スナック1+	421
23 スナック2	185	スナック2	132	スナック2	102	スナック2	44	スナック2+	463
24 スナック3	168	スナック3	135	スナック3	88	スナック3	44	スナック3+	435
25 スナック4	188	スナック4	129	スナック4	89	スナック4	46	スナック4+	452
26 スナック5	181	スナック5	160	スナック5	109	スナック5	50	スナック5+	490
27 タバコ1	218	タバコ1	154	タバコ1	79	タバコ1	35	タバコ1+	486

数式により、左図のように「489」と合計の数字が出力されます。

①全ての商品について同様の計算を行うため「489」の計算結果を選択し、セル(枠)の右下でカーソルが+になる部分に合わせてダブルクリックを行います。

コンビニの売上データ:Excelの解析 売れ筋のランキング調査

行ラベル	購入商品	行ラベル	購入商品	行ラベル	購入商品	行ラベル	購入商品	計算用行ラベル	購入数合計
1	201 アイスクリーム-1	147 アイスクリーム-1	76 アイスクリーム-1	65	アイスクリーム-1	481			481
2	189 アイスクリーム-2	131 アイスクリーム-2	100 アイスクリーム-2	41	アイスクリーム-2	441			441
3	205 アイスクリーム-3	140 アイスクリーム-3	100 アイスクリーム-3	56	アイスクリーム-3	506			506
4	176 アイスクリーム-4	160 アイスクリーム-4	100 アイスクリーム-4	44	アイスクリーム-4	482			482
5	178 アイスクリーム-5	146 アイスクリーム-5	97 アイスクリーム-5	51	アイスクリーム-5	484			484
6	188 インスタント食品1	139 インスタント食品1	86 インスタント食品1	59	インスタント食品1	473			473
7	187 インスタント食品2	141 インスタント食品2	87 インスタント食品2	42	インスタント食品2	472			472
8	182 インスタント食品3	144 インスタント食品3	102 インスタント食品3	44	インスタント食品3	472			472
9	207 インスタント食品4	149 インスタント食品4	99 インスタント食品4	45	インスタント食品4	500			500
10	178 インスタント食品5	149 インスタント食品5	93 インスタント食品5	55	インスタント食品5	475			475
11	171 おにぎり1	143 おにぎり1	84 おにぎり1	46	おにぎり1	474			474
12	180 おにぎり2	152 おにぎり2	84 おにぎり2	46	おにぎり2	474			474
13	180 おにぎり3	118 おにぎり3	82 おにぎり3	47	おにぎり3	437			437
14	180 おにぎり4	142 おにぎり4	82 おにぎり4	53	おにぎり4	465			465
15	188 おにぎり5	145 おにぎり5	111 おにぎり5	40	おにぎり5	472			472
16	189 おにぎり6	127 おにぎり6	96 おにぎり6	43	おにぎり6	475			475
17	197 おにぎり7	142 おにぎり7	119 おにぎり7	46	おにぎり7	473			473
18	210 おにぎり8	124 おにぎり8	94 おにぎり8	48	おにぎり8	477			477
19	189 おにぎり9	142 おにぎり9	95 おにぎり9	46	おにぎり9	475			475
20	210 おにぎり10	124 おにぎり10	94 おにぎり10	48	おにぎり10	477			477
21	180 スナック1	132 スナック1	86 スナック1	43	スナック1	473			473
22	177 スナック2	124 スナック2	102 スナック2	38	スナック2	441			441
23	180 スナック3	132 スナック3	102 スナック3	47	スナック3	474			474
24	186 スナック4	135 スナック4	85 スナック4	44	スナック4	434			434
25	180 スナック5	129 スナック5	99 スナック5	44	スナック5	461			461
26	180 スナック6	132 スナック6	100 スナック6	44	スナック6	474			474
27	216 9/13	184 9/13	79 9/13	43	9/13	481			481
28	187 9/12	189 9/12	89 9/12	42	9/12	481			481
29	187 9/13	121 9/13	74 9/13	47	9/13	426			426
30	219 9/12	149 9/12	83 9/12	57	9/12	508			508
31	185 9/14	149 9/14	84 9/14	54	9/14	485			485

左図のように全ての行に購入商品1~4の購入数の合計が出力されました。

最後に、この購入数合計の数字を大きい数字から順にランキング形式に並べ替えると、目的である「一番売れている商品をランキング形式で調べ」という目的に対する答えの完成です。

あと最後の一事です。頑張りましょう。

左図のようにたくさんある数字を大きい順から順番に並べ替える方法を次ページにて説明します。

コンビニの売上データ:Excelの解析 売れ筋のランキング調査

ランキング形式への並べ替えには少し注意が必要です。

①まず、得られた合計の部分のコピーして別の場所に貼り付け直します。なぜ、貼り付け直すかという、購入数合計の数字は、表記上は「489」などの数字ですが、そのセル(枠)の中には足算の計算式が入力されているので、そのままランキング処理をすると間違った順番になってしまいます。

そのため別の場所に貼り付け直します。②その際にセル(枠)の中の計算式を除去し、数字のみの情報にするために「貼り付けのオプション」を左図のように「値」を選択して貼り付けます。

コンビニの売上データ:Excelの解析 売れ筋のランキング調査

貼り付けが終わると左図のように見た目の上では全く同じ内容がコピーされました。

ただし、右のデータには数式が含まれておらず純粋な数字のデータなのでランキング形式に並べ替えることが出来ます。

次ページにて並べ替えの方法を説明します。

①まず、「計算用行ラベル」と「購入数合計」のセル(枠)を選択します。

コンビニの売上データ:Excelの解析 売れ筋のランキング調査

「計算用行ラベル」と「購入数合計」のセル(枠)を選択した状態のまま、ホームメニューの中の「並べ替えとフィルター」を押し、「フィルター」を押します。

コンビニの売上データ:Excelの解析 売れ筋のランキング調査

「フィルター」を設定すると左図のようにセル(枠)の右側に「▼」の印が表示されます。

①今回は購入数の合計を多い順に表示させたいので、「購入数合計」の「▼」をクリックします。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved 45

コンビニの売上データ:Excelの解析 売れ筋のランキング調査

「購入数合計」の「▼」をクリックすると、このような画面になります。

①今回は、数字の大きい順にランキングをしたいので、「降順」のボタンを押し、順番を並べ替えます。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved 46

コンビニの売上データ:Excelの解析 売れ筋のランキング調査

「降順」の処理を終えると左図のように数字の大きい順に並べ替えられます。

「総計」と「空白」は商品ではないので無視して、それより下の部分を見ると「弁当1」518個、「スナック5」516個、「アイスクリーム3」509個・・・の順に売れていることがわかります。

これにより、どの商品がどの順番でいくつ売れていることがわかりました。

目的とする一番売れている商品をランキング形式で調べ、調べることで出来ました。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved 47

コンビニの売上データ:Excelの解析 売れ筋のランキング調査

売上ランキング	購入商品	個数
1位	弁当1	518
2位	スナック5	516
3位	アイスクリーム3	509
4位	タバコ4	508
5位	お茶3	506
6位	インスタント食品4	500
7位	イン酒	496
8位	弁当3	494
9位	酒5	491
10位	アイスクリーム1	489

左図がこのコンビニの2023年7月の1か月間で売れた商品のランキングです。

実際にExcelで解析操作をしてみてもどう感じたでしょうか？

ただ単に、売上のランキングを作成するだけでも、かなりの労力を割くことを実感できたのではないのでしょうか。

Excelでの解析方法は様々なため、ランキングを作成する方法は今回の方法に限りませんが、なるべく多くのExcel機能を紹介できる方法でランキングを作成しました。

では、本題であるChatGPTを使用するとどれくらい簡単に出来るでしょうか？また、その結果は本当に正しいのでしょうか？

次ページより早速見てみましょう。


Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved 48

⑤ ChatGPTにてデータ解析の練習 ～Excelの練習にて使用した問題を使用～

先ほどまでの講義で、「あるコンビニの1か月間の売上データから一番売れた商品をランキング形式で表示する」方法をExcelを使用して学びました。


「Excel」を現在よく使用される従来の解析方法とすると、「ChatGPTによる解析」は新しい方法です。

先ほどのお題をChatGPTで解析した場合どれくらい簡単に結果までたどり着けるでしょうか？



従来のExcelによる解析

➔

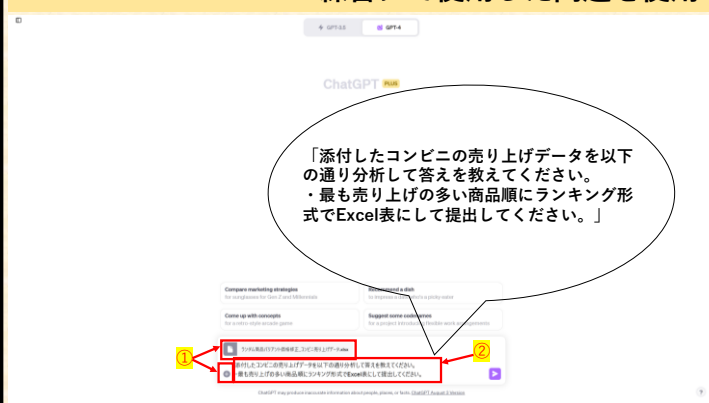


ChatGPTによる解析

どのくらい便利になるでしょうか？

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved 49

⑤ ChatGPTにてデータ解析の練習 ～Excelの練習にて使用した問題を使用～



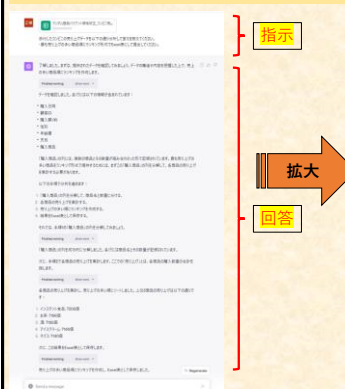
Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved 50

それでは、いよいよChatGPTでデータ解析をしてみましょう。

- ①まず、「+」のボタンを押し、解析するデータを添付します。今回はExcel形式のデータなので、そのデータを添付します。
- ②次に命令文に目的の解析を行うように文章を入力します。

※見ずらいので次ページより拡大した画面を表示します。

⑤ ChatGPTにてデータ解析の練習 ～Excelの練習にて使用した問題を使用～



Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved 51

指示

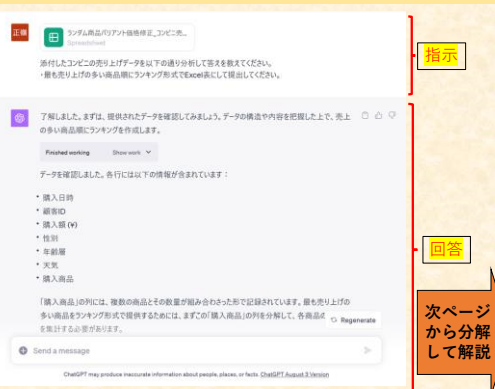
➔

拡大

回答

➔

次ページから分解して解説



Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved 51

⑤ ChatGPTにてデータ解析の練習 ～ChatGPTの回答～

了解しました。まずは、提供されたデータを確認してみましょう。データの構造や内容を把握した上で、売上の多い商品順にランキングを作成します。データを確認しました。各行には以下の情報が含まれています：

- 購入日時
- 顧客ID
- 購入額 (¥)
- 性別
- 年齢層
- 天気
- 購入商品

①

ChatGPTの回答が長いので、少しずつ説明します。

まず①の部分でこちらが与えたデータの内容をChatGPTが確認していることが分かります。

ChatGPTはデータ分析において、その手順を説明しながら解析するのでこちらが確認するのに役立ちます。

※見ずらいので拡大した画面で表示しています。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved 52

⑤ ChatGPTにてデータ解析の練習 ～ChatGPTの回答～

「購入商品」の列には、複数の商品とその数量が組み合わさった形で記録されています。最も売上げの多い商品をランキング形式で提供するためには、まずこの「購入商品」の列を分解して、各商品の売上げを集計する必要があります。

先ほどのExcel操作でいうと、Excelのこのセルの部分に記入された複数の商品の一つ一つのセルに分解した操作と同じようなことをしようとしています。

②の部分で、先ほどExcelで行った作業のように、1つのセルに複数表示されている商品を、一つずつカウントできるように、1つのセルにある複数の商品を分割しようとしていることが分かります。

しかし、「複数の商品とその数量」という表現から「おにぎり2」の「2」を数量として認識していることが推察されます。

※見づらいので拡大した画面で表示しています。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

53

⑤ ChatGPTにてデータ解析の練習 ～ChatGPTの回答～

以下の手順で分析を進めます：

- ① 「購入商品」の列を分解して、商品名と数量に分ける。
- ② 各商品の売上げを集計する。
- ③ 売上げの多い順にランキングを作成する。
- ④ 結果をExcel表として保存する。

それでは、手順1の「購入商品」の列を分解してみましょう。「購入商品」の列を成功地に分解しました。各行には商品名とその数量が記録されています。

手順(1)

③の部分で、分析の手順を(1)～(4)に分けて説明をしています。手順(1)の作業を実行し商品名と数量に分けることに成功しています。

※しかし、この表現から、商品名の後ろの数字、例えば「おにぎり2」の「2」をおにぎりの種類ではなく、数量として認識していることが分かります。ChatGPTに指示を与えるときはしっかりと前提条件を伝える必要があります。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

54

⑤ ChatGPTにてデータ解析の練習 ～ChatGPTの回答～

次に、手順2で各商品の売上げを集計します。ここでの「売上げ」とは、各商品の購入数量の合計を指します。

手順(2)

手順(2)の作業を実行しています。ここで、「売上げ」の定義を確認しています。与えたデータ

の中には「購入額(¥)」もあるため、解釈によっては「売上金額」のランキングを作成される可能性もありました。今回はこちらの意図の通り購入数をカウントしていますが、ここでも指示不足がありました。

ChatGPTに指示した命令文を振り返ってみると・・・

「添付したコンビニの売上げデータを以下の通り分析して答えを教えてください。
・最も売上げの多い商品順にランキング形式でExcel表にして提出してください。」

この表現は解釈によっては売上げの「個数」ではなく、「金額」にもなりうる。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

55

⑤ ChatGPTにてデータ解析の練習 ～ChatGPTの回答～

各商品の売上げを集計し、売上げの多い順にソートしました。上位5商品の売上げは以下の通りです：

インスタント食品: 7208個
お茶: 7196個
酒: 7185個
アイスクリーム: 7146個
タバコ: 7140個

次に、この結果をExcel表として保存します。売上げの多い商品順にランキングを作成し、Excel表として保存しました。このExcelファイルをクリックしてダウンロードしてください。何か他に質問や要求があれば、お気軽にお知らせください。

手順(3)

手順(3)の作業を実行しています。結果の一部を表記しています。手順(4)にて、こちらの指示通りExcel形式でランキングの結果を出力しています。

手順(4)

以上が、最初の命令文によって得られた結果です。しかし、先ほどから指摘しているように、今回のランキングは数量に誤りがあります。ChatGPTに再度指示を出して修正した結果を出力させましょう。

こちらの指示不足により「おにぎり2」という表記を「おにぎり」を1個買った訳ではなく、「おにぎり」を2個買ったものとしてカウントしています。他の「インスタント食品」「お茶」なども同様に間違えています。この部分を修正して再度正しい答えを出してもらいましょう。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

56

⑤ ChatGPTにてデータ解析の練習 ～ChatGPTの回答～

正解 結果の数量に誤りがあります。今回のデータの「おにぎり2」の「2」は個数ではなくおにぎりの種類を意味しています。つまり、「雑穀2」を1個、「タバコ2」を1個買ったこととなります。そのような分類で再度計算しなおして結果をExcelファイルで提出してください。

申し訳ございません、誤解しました。指定された方法で再計算を行い、結果をExcelファイルに保存します。

まず、商品名とその種類の数字を分けて、商品ごとのカウントを行います。

Finished working Show work

各商品の購入回数を集計し、売り上げの多い順にソートしました。上位5商品の購入回数は以下の通りです：

1. 弁当: 2413回
2. インスタント食品: 2392回
3. お茶: 2386回
4. アイスクリーム: 2385回
5. タバコ: 2375回

この結果をExcel表として保存します。

Send a message

指示 それでは、正しい結果を得られるように再度指示を出しましょう。

最初の方でお伝えしたようにChatGPTは会話を続けた場合、前の文脈を踏まえて回答する性質を備えています。

①よって、再度「Send a message」の欄に

結果の数量に誤りがあります。今回のデータの「おにぎり2」の「2」は個数ではなくおにぎりの種類を意味しています。つまり、「雑穀2」の場合、「雑穀2」を1個、「タバコ2」を1個買ったこととなります。そのような分類で再度計算しなおして結果をExcelファイルで提出してください。

と入力し、訂正を促しました。ChatGPTはどのように回答するでしょうか？次ページに示します

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved 57

⑤ ChatGPTにてデータ解析の練習 ～ChatGPTの回答～

申し訳ございません、誤解しました。指定された方法で再計算を行い、結果をExcelファイルに保存します。

まず、商品名とその種類の数字を分けて、商品ごとのカウントを行います。各商品の購入回数を集計し、売り上げの多い順にソートしました。上位5商品の購入回数は以下の通りです：

弁当: 2413回
インスタント食品: 2392回
お茶: 2386回
アイスクリーム: 2385回
タバコ: 2375回

この結果をExcel表として保存します。申し訳ございません、指定された方法に従って再計算しました。

商品売り上げランキング(修正版)をダウンロード

このExcelファイルをクリックしてダウンロードしてください。何か他に質問や要求があれば、お気軽にお知らせください。

①指示内容を理解し、再度計算しなおしています。

②しかし、再計算の結果である上位5商品の内容を見てみると「弁当」「インスタント食品」...となっており、「弁当1」や「インスタント食品2」などの表現が無いことから、こちらの意図を誤解し続けていることが分かります。

もう一度指示をして、正しい回答を得られるように挑戦してみましょう。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved 58

⑤ ChatGPTにてデータ解析の練習 ～ChatGPTの回答～

このExcelファイルをダウンロードしてください。何か他に質問や要求があれば、お気軽にお知らせください。

正解 結果が間違っています。弁当やインスタント食品にはそれぞれ違う種類の商品として「弁当1」や「弁当2」などの種類があります。それぞれの商品、つまり「弁当1」や「弁当2」を違う商品としてそれぞれカウントして再度ランキング形式でエクセル形式で提出してください。

申し訳ございません、再度正確に計算いたします。

商品名とその種類（例：「弁当1」「弁当2」）を分けて、それぞれの商品のカウントを行います。そして、カウントの多い順にランキングを作成します。

申し訳ございません、前回の説明が不十分でした。正確に各商品（「弁当1」「弁当2」などの種類を区別して）の購入回数を集計しました。

上位5商品の購入回数は以下の通りです：

1. 弁当: 2413回
2. インスタント食品: 2392回

Send a message

①再度「Send a message」の欄に

結果が間違っています。弁当やインスタント食品にはそれぞれ違う種類の商品として「弁当1」や「弁当2」などの種類があります。それぞれの商品、つまり「弁当1」や「弁当2」を違う商品としてそれぞれカウントして再度ランキング形式でエクセル形式で提出してください。

と入力し、訂正を促しました。ChatGPTはどのように回答するでしょうか？次ページに示します

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved 59

⑤ ChatGPTにてデータ解析の練習 ～ChatGPTの回答～

申し訳ございません、再度正確に計算いたします。商品名とその種類（例：「弁当1」「弁当2」）を分けて、それぞれの商品のカウントを行います。そして、カウントの多い順にランキングを作成します。

申し訳ございません、前回の説明が不十分でした。正確に各商品（「弁当1」「弁当2」などの種類を区別して）の購入回数を集計しました。

上位5商品の購入回数は以下の通りです：

弁当: 2413回
インスタント食品: 2392回
お茶: 2386回
アイスクリーム: 2385回
タバコ: 2375回

各商品の売り上げランキングをダウンロード

このExcelファイルをクリックしてダウンロードしてください。何か他に質問や要求があれば、お気軽にお知らせください。

①指示内容を理解し、再度計算しなおしています。

②しかし、再計算の結果である上位5商品の内容を見てみると未だに、「弁当1」や「インスタント食品2」などの表現が無いことから、こちらの意図を誤解し続けていることが分かります。

もう一度指示をして、正しい回答を得られるように挑戦してみましょう。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved 60

⑤ ChatGPTにてデータ解析の練習 ～ChatGPTの回答～

指示

①再度「Send a message」の欄に提出されたExcelファイルは未だに間違っています。「弁当1」や「弁当2」を別々の商品として認識し、それぞれについてカウントしてランキングを計算し、エクセル形式で提出してください。

回答

と入力し、訂正を促しました。ChatGPTはどのように回答するでしょうか？次ページに示します

①指示内容を理解し、再度計算しなおしています。

②再計算が上手くいかずエラーの原因を探し再計算しています。

③結果を見ると「弁当1」や「スナック5」と表記されており、どうやら正しい結果のようです。

④結果をExcel形式で出力しているのでダウンロードして確認しましょう。

⑤ ChatGPTにてデータ解析の練習 ～ChatGPTの回答～

申し訳ございません、再度、指定された方法で計算を行います。「弁当1」、「弁当2」など、商品名とその種類を合わせた名前ごとにカウントを行い、ランキングを作成します。再度計算しましたが、結果が以前と同じになってしまっています。お手数をおかして申し訳ございません。

エラーの原因を特定し、修正いたします。正確に「弁当1」や「弁当2」などの商品名とその種類を区別してカウントを行い、ランキングを作成します。申し訳ございませんが、正確に計算しました。

上位5商品の購入回数は以下の通りです：
 弁当1: 518回
 スナック5: 516回
 アイスクリーム3: 509回
 タバコ4: 508回
 お茶3: 506回

正確な商品売り上げランキングをダウンロード
このExcelファイルをクリックしてダウンロードしてください。何か他に質問や要求があれば、お気軽にお知らせください。

①指示内容を理解し、再度計算しなおしています。

②再計算が上手くいかずエラーの原因を探し再計算しています。

③結果を見ると「弁当1」や「スナック5」と表記されており、どうやら正しい結果のようです。

④結果をExcel形式で出力しているのでダウンロードして確認しましょう。

⑤ ChatGPTにてデータ解析の練習 ～ChatGPTの回答～

上位5商品の購入回数は以下の通りです：
 1. 弁当1: 518回
 2. スナック5: 516回
 3. アイスクリーム3: 509回
 4. タバコ4: 508回
 5. お茶3: 506回

正確な商品売り上げランキングをダウンロード

このExcelファイルをクリックしてダウンロードしてください。何か他に質問や要求があれば、お気軽にお知らせください。

左図の「正確な商品売り上げランキングをダウンロード」をクリックすると、指示した通りExcel形式のファイルがダウンロードされます。

自分で計算した結果とChatGPTが計算した結果を比べて、ChatGPTが正しい答えを導き出しているか確認しましょう。

⑤ ChatGPTにてデータ解析の練習 ～ChatGPTの回答～

Excelで自分で計算した結果

ChatGPTが出した結果

Excelで自分で計算した結果

ChatGPTが出した結果

左図に「自分で計算した結果」と「ChatGPTが計算した結果」を比べました。少し見づらいので二つの結果を横に並べました。

どうでしょうか？同じ個数の売上の場合、表示の順番に違いはあるものの、結果は最後までピッタリ同じ結果になりました。

ChatGPTを使用すると最初の指示を含めてたった4回の指示で欲しい結果を簡単に得ることが出来ました。時間もかなり短縮して答えを得られました。これがChatGPTによる解析援助の例として挙げられると思います。

⑥ChatGPT使用時の注意点

	A	B	C	D	E
	自分で計算した結果	購入数合計	ChatGPTの結果	購入回数	
1	菓子	12988			
2	菓子	518	菓子	518	
3	アイス	518	アイス	518	
4	アイス	509	アイス	509	
5	アイス	509	アイス	509	
6	アイス	509	アイス	509	
7	アイス	509	アイス	509	
8	アイス	509	アイス	509	
9	アイス	509	アイス	509	
10	アイス	509	アイス	509	
11	アイス	499	アイス	499	
12	アイス	499	アイス	499	
13	アイス	499	アイス	499	
14	アイス	489	アイス	489	
15	アイス	489	アイス	489	
16	アイス	489	アイス	489	
17	アイス	489	アイス	489	
18	アイス	489	アイス	489	
19	アイス	489	アイス	489	
20	アイス	479	アイス	479	
21	アイス	479	アイス	479	
22	アイス	479	アイス	479	
23	アイス	479	アイス	479	
24	アイス	479	アイス	479	
25	アイス	479	アイス	479	
26	アイス	479	アイス	479	
27	アイス	479	アイス	479	
28	アイス	479	アイス	479	
29	アイス	479	アイス	479	
30	アイス	479	アイス	479	
31	アイス	479	アイス	479	
32	アイス	479	アイス	479	
33	アイス	479	アイス	479	
34	アイス	479	アイス	479	
35	アイス	479	アイス	479	
36	アイス	479	アイス	479	
37	アイス	479	アイス	479	
38	アイス	479	アイス	479	
39	アイス	479	アイス	479	
40	アイス	479	アイス	479	
41	アイス	479	アイス	479	
42	アイス	479	アイス	479	
43	アイス	479	アイス	479	
44	アイス	479	アイス	479	
45	アイス	479	アイス	479	
46	アイス	479	アイス	479	
47	アイス	479	アイス	479	
48	アイス	479	アイス	479	
49	アイス	479	アイス	479	
50	アイス	479	アイス	479	
51	アイス	479	アイス	479	
52	アイス	479	アイス	479	
53	アイス	479	アイス	479	
54	アイス	479	アイス	479	
55	アイス	479	アイス	479	
56	アイス	479	アイス	479	
57	アイス	479	アイス	479	
58	アイス	479	アイス	479	
59	アイス	479	アイス	479	
60	アイス	479	アイス	479	
61	アイス	479	アイス	479	
62	アイス	479	アイス	479	
63	アイス	479	アイス	479	
64	アイス	479	アイス	479	
65	アイス	479	アイス	479	
66	アイス	479	アイス	479	
67	アイス	479	アイス	479	
68	アイス	479	アイス	479	
69	アイス	479	アイス	479	
70	アイス	479	アイス	479	
71	アイス	479	アイス	479	
72	アイス	479	アイス	479	
73	アイス	479	アイス	479	
74	アイス	479	アイス	479	
75	アイス	479	アイス	479	
76	アイス	479	アイス	479	
77	アイス	479	アイス	479	
78	アイス	479	アイス	479	
79	アイス	479	アイス	479	
80	アイス	479	アイス	479	
81	アイス	479	アイス	479	
82	アイス	479	アイス	479	
83	アイス	479	アイス	479	
84	アイス	479	アイス	479	
85	アイス	479	アイス	479	
86	アイス	479	アイス	479	
87	アイス	479	アイス	479	
88	アイス	479	アイス	479	
89	アイス	479	アイス	479	
90	アイス	479	アイス	479	
91	アイス	479	アイス	479	
92	アイス	479	アイス	479	
93	アイス	479	アイス	479	
94	アイス	479	アイス	479	
95	アイス	479	アイス	479	
96	アイス	479	アイス	479	
97	アイス	479	アイス	479	
98	アイス	479	アイス	479	
99	アイス	479	アイス	479	
100	アイス	479	アイス	479	

あなたはコンビニの店長として、7月の1か月間の売上においてどの商品が売れているのかをランキング形式で知ることが出来ました。しかし、今回の講義を通して、「おや？結局ChatGPTの答えを、自分のExcel計算による結果と比較して正しいかどうかを確認したということは、結局自分で計算する手間が発生するわけだからChatGPTによって楽になるわけではないのでは？少なくとも自分でやれないと意味がないのでは？」と思ったあなたは、鋭い感性をお持ちです。

結論から申し上げますと、その通りなのです。実はこれこそが今回の講義で一番お伝えしたいメッセージの一つです。それは、ChatGPTで何でもできるわけではなく、そのチェック能力、つまりデータ分析力の基礎知識を持ち合わせていないと、ChatGPTを上手く活用できないのです。よって、ChatGPTを使用した結果が間違っている場合に気が付かない、という問題が生じてしまうのです。

⇒これがAIのブラックボックス性の良い例です。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

65

⑥ChatGPT使用時の注意点

【AIの持つブラックボックス性について】

実際に今回のChatGPTとの対話でも何度も間違った答えを得ました。今回は幸いにして、専門的な知識が無くともChatGPTの回答が以下のように

インスタント食品: 7208個
お茶: 7196個
酒: 7185個
アイスクリーム: 7146個
タバコ: 7140個

という結果から「おや？なぜ『おにぎり1』『おにぎり2』などの表記出ないのだろうか？」という疑問を持って、得られた結果を疑うことが出来ました。もし、これが「おにぎり1」7200個「おにぎり2」7100個のような答えだったらどうでしょうか？

そのまま回答を信用して大量の仕入れをしてしまうことにつながっていたのではないのでしょうか？実際の売り上げは500個程度にも関わらず10倍以上の仕入れをしてしまったら売れ残りが大量に生じ、大変な赤字になってしまいます。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

66

⑥ChatGPT使用時の注意点

【AIの持つブラックボックス性について】

二回目の修正によっても

弁当: 2413回
インスタント食品: 2392回
お茶: 2386回
アイスクリーム: 2385回
タバコ: 2375回

という間違った結果を得られました。しかし、これも商品ラベルから間違えを推測するしかありませんでした。

ChatGPTを人間との会話のように、その答えを盲目的に信用してしまうと間違えに気づけません。AI全般に言えることですが、このようにAIがどのような計算過程を経てその答えを出したかを理解せずを使用してしまおうとAIの中で何を計算しているのか分からず、誤った答えを見抜けません。つまり、AIのブラックボックス性を排除できません。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

67

⑥ChatGPT使用時の注意点

では、結局ChatGPTで今回の回答を得るためには自分でExcelなどを使って計算するしかないのでしょうか？そんなことはありません。

では、どのようにチェックするのでしょうか？

幸いにもChatGPTには、「Python」のプログラムコードと実行結果を見られる機能がついています。よって、各回答に対してどのように計算したのかを確認することが出来ますし、プログラムコードが分からなくても計算の根拠を計算方法を直接ChatGPTに聞けば回答してくれる機能があります。

⇒具体的にChatGPTの回答がなぜ7200個や2400個など出てきたのかをチェックを通して考えてみましょう。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

68

⑥ChatGPT使用時の注意点

それでは、手順1の「購入商品」の列を分解してみましょう。

Finished working Show work

「購入商品」の列を成功的に分解しました。各行には商品名とその数量が記録されています。

次に、手順2で各商品の売り上げを集計します。ここでの「売り上げ」とは、各商品の購入数量の合計を指します。

Finished working Show work

各商品の売り上げを集計し、売り上げの多い順にソートしました。上位5商品の売り上げは以下の通りです：

1. インスタント食品: 7208個
2. お茶: 7196個

ChatGPTには左図のように実行したプログラムをPythonコードで表示する機能がついているため、どのような過程で計算したのかを確認することが出来ます。

この部分をクリックして約7200個の回答を出力した際の計算内容をチェックしてみましょう。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

69

⑥ChatGPT使用時の注意点

```
python
# 購入商品の列を", "で分割して各商品とその数量をリストに格納する
product_lists = data[ '購入商品' ].str.split(", ").expand(True).stack().reset_index()

# 各商品の名称と数量を分ける
product_name = product_lists.str.extract( "(?P<name>).*")
product_qty = product_lists.str.extract( "(?P<qty>).*").astype(float)

# 商品名と数量のテーブルを作成する
product_df = pd.DataFrame({
    '商品名': product_name(),
    '数量': product_qty()
})

product_df.head()
```

RESULT	商品名	数量
0	おにぎり	2.0
0	タバコ	2.0
1	おにぎり	2.0
1	アイスクリーム	9.0
2	おにぎり	9.0

Pythonコードは左図のような内容になっていました。Pythonの細かい説明は割愛しますが、左図を見ると、商品名とその後の数字をその商品の数量としてカウントしていることが分かります。

つまり「おにぎり5」という表記があった際は、こちらは「おにぎり5」を1個買った、とカウントしてほしいわけですが、ChatGPTは「おにぎり」を5個買ったという理解をしてカウントしている訳です。

では、2回目に間違えて約2400個の回答をしたのはなぜでしょうか？

⇒次ページにて確認していきましょう。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved 70

⑥ChatGPT使用時の注意点

まず、商品名とその種類の数字を分けて、商品ごとのカウントを行います。

```
python
# 各商品の購入回数のカウントする
product_count = product_df[ '商品名' ].value_counts().reset_index()
product_count.columns = [ '商品名', '購入回数' ]

# 購入回数の多い順にソートする
product_count_sorted = product_count.sort_values(by='購入回数', ascending=False)
product_count_sorted.head()
```

RESULT	商品名	購入回数
0	お茶	2413
1	インスタント食品	2392
2	お茶	2286
3	アイスクリーム	2385
4	タバコ	2375

各商品の購入回数を集計し、売り上げの多い順にソートしました。上位5商品の購入回数は以下の通りです：

1. お茶: 2413回

Pythonコードは左図のような内容になっていました。ここでは「おにぎり1」や「おにぎり2」の違いを無視して「おにぎり」という括りの数え上げをしています。

よって、「おにぎり1」は「おにぎり」の後ろの「1」を無視して「おにぎり」を1個、「おにぎり2」は「おにぎり」の後ろの「2」を無視して「おにぎり」を1個とカウントして、最終的に「おにぎり1~5」の数をそれぞれ1個として「おにぎり」という概念で合計して算出したことが分かります。

ですから、約7200→約2400と1/3程度に減ったわけですね。

今回のデータは各商品1~5まで同程度に出現するデータなので、1~5の平均値である3で割った数になった訳です。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

71

⑥ChatGPT使用時の注意点

【AIのブラックボックス性の排除】

このようにChatGPTには、「Python」のプログラムコードと実行結果を見られる機能がついています。よって、各回答に対してどのように計算したのかを確認することが出来ますし、プログラムコードが分からなくても計算の根拠や計算方法を直接ChatGPTに聞けば回答してくれる機能があります。

ChatGPTに頼めばなんだか分からないけど回答を出してくれるからといって、その解析のプロセスを確認せずにブラックボックスを通して出した結果を使用してはいけません。

私は、このAIの持つブラックボックス性を排除して使用方法をAI2L(AI to Learn: AIを自身の学習に使う)と名付けて、AIを「学習活用」として積極的に用いることを提唱しました。

今回の例でも分かるように、ExcelやPythonコードが分からなくてもAIを通して学ぶことも可能ですし、学習目的にAIを使用することは後述する様々な危険を阻止できるので有益だと考えています。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

72

⑥ChatGPT使用時の注意点

【AI活用の危険なデメリット：情報の取り扱い】

まず、注意点として「情報の取り扱い」が挙げられます。これは、あらゆるパソコン操作に共通する感性ですが、先ほどのように、仮にあなたがコンビニの店長だとして、実際のコンビニの売上データをそのままChatGPTに計算してもらおう、という方法は控えた方が良いでしょう。

なぜなら、企業として社外秘である「売上データ」をChatGPTに提供してしまい、最終的に情報流出を起こす可能性が有るからです(情報リテラシー)。

実際、今回の講義に使用した元データも、ChatGPTにコンビニの売上のダミーデータ(仮想データ)を作成するように指示して作成したもののなのです。

実際のコンビニの売上データは使用していません。自分自身で大量のそれらしいデータを作成するのは手間なのでChatGPTにダミーデータの作成を依頼し、作成しました。しかし、データをよく見た方はお気づきかもしれませんが、一部現実的でない部分が含まれています。例えば、「酒2」や「タバコ3」などを10代の人買ったデータがいくつか含まれています。これは違法行為ですから、コンビニでこのようなことが多々生じることは考えにくく、ダミーデータの出来栄としては不適切とは言えますが、今回は敢えて残しました。

ChatGPTを上手く 사용하기 위해서는、明確な指示が必要となる良い例ですね。

Copyright © 2023 Seime A. Shintani Chubu Univ. All Rights Reserved

73

⑥ChatGPT使用時の注意点

【指示内容の明瞭化：プログラミングの習得】

今回コンビニの売上ランキングを作成する際に、ChatGPTとの間で多くの言い方による解釈の齟齬が生じ、回答が間違っていました。ChatGPTに指示を与えるときはなるべく前提条件と、やって欲しい内容を明確に伝える必要があります。

ただし、どこまで正確に伝えれば言語上の齟齬が起きずに処理できるかはやってみないと分からないことでもあるので、できればプログラミングの習得をすることが望ましいでしょう。

ChatGPTをデータ分析で使用する際に最低でもExcelの知識やPythonなどのプログラム知識を必要とします。

特にExcelは教育や会社などでも一般的に用いられるソフトなので高校生の皆さんには基礎体力としてぜひ身に付けて欲しい分析力だと考えます。

もし、Excelも触ったことがないという方であれば、まずはExcelに触れて様々な機能があることを習得してください。

Copyright © 2023 Seime A. Shintani Chubu Univ. All Rights Reserved

74

⑥ChatGPT使用時の注意点

【AI活用の危険なデメリット：指示内容からChatGPT利用者の興味関心が分かってしまう】

先ほど、ChatGPTを使用する側が実際のデータを使わないようにして情報流出を防いでくださいとお伝えしました。しかし、利用者がそもそもChatGPTに特定の質問を多くすることによって、利用者の興味関心がChatGPT側に把握されてしまうことにも注意が必要です。

例えば質問者がある特定の分野の質問を繰り返すと、ChatGPT側に質問者の興味関心を把握される恐れがあります。

皆さんもネット検索などで、自分の興味関心に沿った広告が出てくることを経験したことはないでしょうか？

ChatGPTもネットにつながっているので注意が必要です。

次ページにChatGPTの使用時の注意点をまとめます。

Copyright © 2023 Seime A. Shintani Chubu Univ. All Rights Reserved

75

⑥ChatGPT使用時の注意点(まとめ)

- ・ ChatGPTが出した結果に対して自分自身で責任をもって内容をチェックすること。そのためのデータ分析の力を養うためにExcelやPythonなどのソフトやプログラミングを勉強すること。
- ・ ChatGPTに限らず情報リテラシーとして、個人情報を安易に入力しないこと。
- ・ ChatGPTに対して会話特有の解釈の齟齬が生まれないように具体的指示を出すこと。
- ・ あまりにも偏った質問ばかりをして自分自身の考えている内容を明らかにしすぎないこと。
- ・ 学習支援ツールとしてのAI活用としてAI 2L(AI to Learn)のような上手い使い方をして自身の学習した内容の理解を深めるためのツールとして使用してみる。
- ・ ChatGPTの問題点ばかりを気にして使用しないのではなく、積極的に使用して自分のデータ解析の能力を向上させること。まずはダミーデータ作成などで練習すること。

Copyright © 2023 Seime A. Shintani Chubu Univ. All Rights Reserved

76

⑦AIを使いこなすための新たな指針(AI2L)

ChatGPTを含むAIの使用例として上手い使用例としてはまず「教育、学習用」として使用するのが安全だと考えられます。

今回のようなダミーデータの作成や自身が学習したい内容、例えば生物や物理の概念について、高校生でも分かる言葉で100字以内に簡潔に説明してもらい、自分で教科書や参考書でチェックして学習支援ツールとして使用するの、安全かつ有益な使用方法だと考えられます。

質問者の興味関心が例えChatGPT側に伝わったとしても、教育や学習目的であればさほど個人情報を特定されたり、自分自身の考えを詮索される恐れがないためです。私はこれを学習支援ツールとしてのAI活用としてAI2L(AI to Learn)と名付けました。要は学習支援ツールとしてのChatGPTの活用です。

また、これは少し高度な話ですが、研究用でも繰り返しの作業や数式の法則探索には相性が良く有益な利用方法だと考えられます。

以上の注意点を踏まえて新しいChatGPTという新技術を上手く利用していただくことを期待しています。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

77

⑧まとめ

以上で本講義を終わります。
皆さんはどのような感想を持ちましたか？

データ解析における簡便化をお伝えしたうえで、問題点もお伝えしましたが、実際のデータを使用した場合でも工夫次第でChatGPTを活用することは可能です。

たとえば、ある学校の先生が生徒の成績を点数の高い順に並べ替えたり、平均値や標準偏差を計算した際に、一覧表さえあればChatGPTで計算自体は出来ませんが、生徒の氏名が入力されているため、情報流出につながります。しかし、生徒の名前を番号に置き換えて、得られた結果を最終的に生徒の名前に自分で突き合わせれば個人情報の流出は防げるでしょう。

このように、少し手間かけることで、ChatGPTにやらせること、自分自身で抑えるべき情報を抑えて、便利な部分だけを、チェック機能を備えたうえで使用することが出来れば、理想的な使い方が出来るでしょう。(AI2L)

いきなりChatGPTに触らなくても、今回のExcelの練習問題を実際に自分で解いてみるのも十分勉強になると思いますので、是非活用してみてください。きっと、大学や職場で役立つと思います。皆さんの学びに貢献できれば幸いです。
以上です。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

78

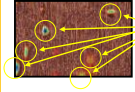
研究におけるAI2L活用の実施例

チタン板表面



熟練者じゃないと研磨が十分かどうかの判別が難しい。

研磨不足部位を指摘



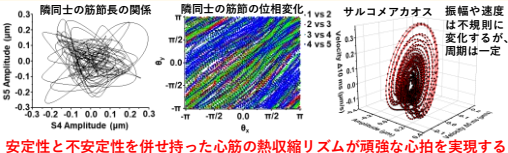
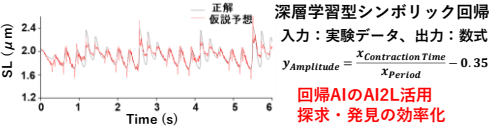
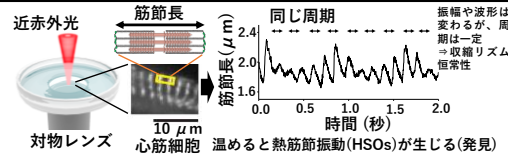
研磨十分と不十分を分類できる分類AIを作成。その判断根拠を可視化。
分類AIのAI2L活用
学習の効率化

チタン板表面



上記の結果から特徴を学べば、学習者が研磨不足部位を指摘可能になる。

言語化が難しかった専門的識別能力の習得・学習を速やかに進めるようにする！



以上が前半の授業内容です。

それでは早速本日の + α である
後半

「研究者向けのAI活用の紹介」の発表を始めます。研究者の自身の研究にも簡単に取り入れられ、かつ、研究時間の短縮につながると考えられる例を紹介します。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

80

後半(+ α)の目標

AI数理の専門家以外の研究者はどれほどAIを活用しているのでしょうか？

実はChatGPTの登場により、プログラミング知識の習得に対するハードルは格段に下がりました。

本講演では実際に研究で得られたデータの解析を紹介し、ご自身の研究活動にAI数理を活用するメリットを実感していただくことを目的としています。

特に、AIの持つブラックボックス性や研究データの情報保護の観点から、AIのデメリットを排除する方法として私が提唱したAI2L(AI to Learn: AIを自身の学習に使う)という指針を紹介します。ご自身の研究に役立てていただければ幸いです。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

81

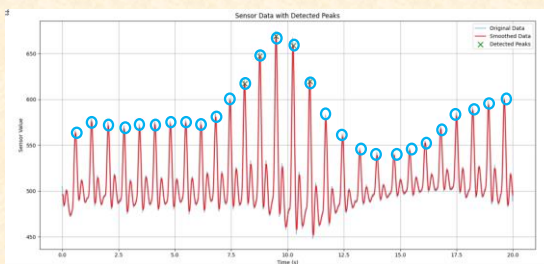
後半の概要項目

- ①実際のデータの提示、お題の設定
- ②ChatGPTによるダミーデータの作成
- ③ダミーデータをChatGPTに解析させ、得たいプログラムコードを習得
- ④得られたプログラムコードの検証
- ⑤実際のデータをChatGPTを**通さず**に解析
- ⑥まとめ：AIを使いこなすための新たな指針(AI2L)

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

82

①実際のデータの提示、お題の設定



左図は脈波のデータです。

研究者は、この脈波のデータから、大きい波のピークの数と、その間隔を測りたいと考えています。

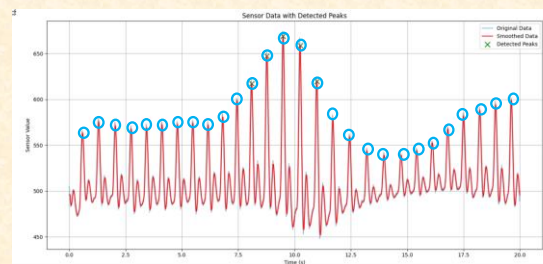
このデータがテキスト形式で出力された場合、そのピーク位置を検出して、脈波の値とその時間軸の値を算出して表にするには少し手間がかかります。

お題：
○で示した心拍毎の大きいピークを検出して、その値と時間の値を表にまとめる。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

83

②ChatGPTによるダミーデータの作成



しかし、この実測データをそのままChatGPTに添付して解析してしまうと、研究者の命ともいえる大切なデータの情報管理が十分に出来ません(情報リテラシー)。

もしかすると、自分の意図しない組織間でデータを活用されてしまう恐れがあります。

そこで、絶対にネット上にあげたくないデータを解析したい場合は、自身のデータを直接ChatGPTに解析させるのではなく、まず、ダミーデータを作成してから解析することをお勧めします。

お題：
○で示した心拍毎の大きいピークを検出して、その値と時間の値を表にまとめる。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

84

② ChatGPTによるダミーデータの作成



Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

85

早速ダミーデータを作成してみましょう。

左図はChatGPTにダミーデータ作成を依頼した場面です。

「周期的に振動するセンサー値のダミーデータのCSVファイルを作成してください。1列目は時間（単位は「秒」）、2列目はセンサーの値（0~1000）としてください。センサーの値の変動周期は2秒周期、20秒間のデータとしてください。」

と入力し、ダミーデータを作成しました。

次ページにダミーデータ作成時のPythonコードを表示します。

② ChatGPTによるダミーデータの作成

```
python
import csv
import math
import os

# File path
file_path = "/mnt/data/sensor_data.csv"

# Generate dummy data
time_values = [i * 0.01 for i in range(2001)]
sensor_values = [500 + 500 * math.sin(2 * math.pi * t) for t in time_values]

# Write to CSV
with open(file_path, 'w', newline='') as csvfile:
    writer = csv.writer(csvfile)
    writer.writerow(["Time (s)", "Sensor Value"])
    for t, value in zip(time_values, sensor_values):
        writer.writerow(t, value)

file_path
```

左図のようなフローでダミーデータを作成したことが確認できます。

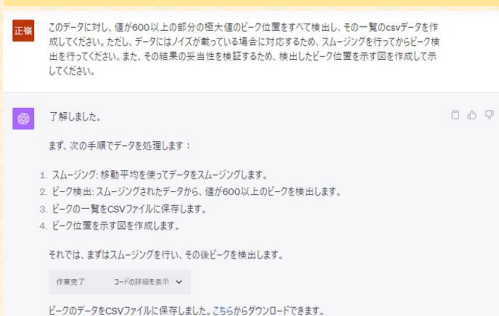
本講演では、細かいPythonコードの説明は割愛しますが、ChatGPTの場合はPythonコードを一から勉強しなくても左図のようにプログラムコードを表示してくれるので、プログラミングの学習に対するハードルが格段に下がったことが実感できると思います。

また、Pythonは様々な作業を指示するコアとなるプログラミングコードをたくさんネットで検索できるのでプログラミングの勉強も頑張れば自力で出来るでしょう。もしくは、得られたプログラムコードの意味をChatGPTに解説してもらうという方法で調べてもいいでしょう。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved


86

③ ダミーデータをChatGPTに解析させ、得たいプログラムコードを習得



Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

87

それでは、いよいよ得られたダミーデータをChatGPTに解析させ、目的である「で示した心拍毎の大きいピークを検出して、その値と時間の値を表にまとめる。」ための実行コードを得てみましょう。

※スムージングは波形データに生じるノイズの除去を意味しています。

今回は、ピーク検出の妥当性を確認するために図の作成も依頼しました。次ページにてChatGPTの回答を確認しましょう。

④ 得られたプログラムコードの検証

```
python
import numpy as np
import pandas as pd
from scipy.signal import find_peaks

# Load the data
df = pd.read_csv(file_path)

# Apply a moving average for smoothing
window_size = 5 # Define the window size for moving average
df["smoothed_value"] = df["sensor_value"].rolling(window_size).mean()

# Detect peaks
peaks, _ = find_peaks(df["smoothed_value"].values, height=600)

# Get peak positions and values
peak_times = df.iloc[peaks][["Time (s)", "sensor_value"].values]
peak_values = df.iloc[peaks][["smoothed_value"].values]

# Save to CSV
peak_data_path = "/mnt/data/peaks_data.csv"
df_peaks = pd.DataFrame({
    "Time (s)": peak_times,
    "Peak Value": peak_values
})
df_peaks.to_csv(peak_data_path, index=False)

peak_data_path
```

左図がChatGPTによって得られたピーク検出のプログラムコードです。

細かい説明は割愛しますが、脈波の値が600以上の値に対してピーク検出を正しく行いピークを正しく検出できていることが分かります。

ここは、少しPythonの知識が必要とされる部分ですので、まだ習得されていない方はPythonの本で勉強されるか、是非、私の紀要論文の「生命医学系分野の学生のためのプログラミング・人工知能教育カリキュラム-Python言語の基本から、生成の人工知能アルゴリズムまで-」を見てみてください。

次に、得られた「ピーク検出」が上手にできているかどうかを念のために確認するために作成させた図を表示します。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

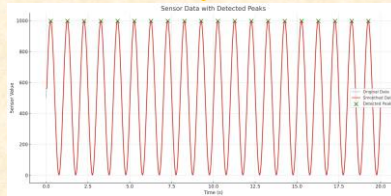
88

④ 得られたプログラムコードの検証

```
python
import matplotlib.pyplot as plt

plt.figure(figsize=(8, 7))
plt.plot(df["time (s)", df["Sensor Value"], label="Original Data", alpha=0.5)
plt.plot(df["time (s)", df["Smoothed Value"], label="Smoothed Data", color="green")
plt.scatter(peak_times, peak_values, color="green", marker="x", s=100, label="Detected Peaks")
plt.xlabel("Time (s)")
plt.ylabel("Sensor Value")
plt.legend()
plt.title("Sensor Data with Detected Peaks")
plt.grid(True)
plt.tight_layout()
plt.show()
```

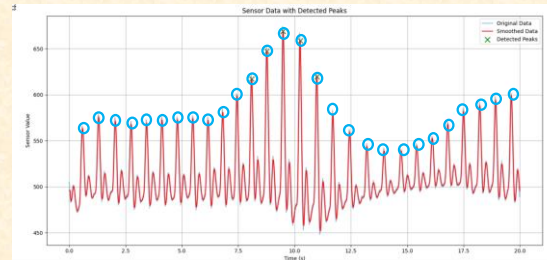
左図がChatGPTによって得られた「与えたダミーデータに対するピーク検出を図で示す」為のPythonコードです。以下が、得られた図です。きちんとピーク検出できていることが確認できます。



Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

89

⑤ 実際のデータをChatGPTを通さずに解析



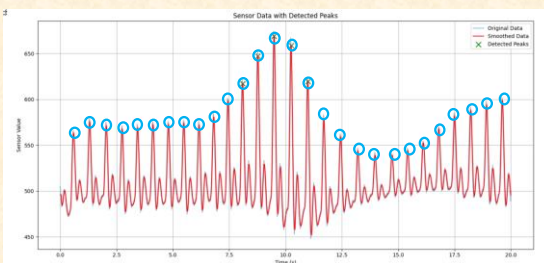
それでは、いよいよ自身が実際に計測したデータに対して、ChatGPTによって得られたピーク検出のコードを利用して、ピーク検出を行きましょう。

この方法のポイントは、ChatGPTに実際の研究データを提供せずに、「ダミーデータを作成する」というプロセスを一つ追加するだけで、やって欲しいプログラミングコードを取得し、それを自分のデータに適用させてオフラインの状態での解析が出来る点です。自身の研究データに関する情報管理の観点から安心して使用できます。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

90

⑤ 実際のデータをChatGPTを通さずに解析



実際に左図の実測データに対して上手くピーク検出(○)できるでしょうか。

早速やってみましょう。

次ページに得られたコードを実行する方法を示します。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

91

⑤ 実際のデータをChatGPTを通さずに解析

```
In [1]: import numpy as np
import pandas as pd
from scipy.signal import find_peaks

# Load the data
df = pd.read_csv(file_path)
df = pd.read_csv("sensor_data.csv")

# Apply a moving average for smoothing
window_size = 5 # Define the window size for moving average
df["Smoothed Value"] = df["Sensor Value"].rolling(window=window_size).mean().fillna(method="ffill")

# Detect peaks
peaks, _ = find_peaks(df["Smoothed Value"].values, height=600)

# Extract peak positions and values
peak_times = df.iloc[peaks]["Time (s)"].values
peak_values = df.iloc[peaks]["Smoothed Value"].values

# Save to CSV
peak_data_path = "peaks_data.csv"
df_peaks = pd.DataFrame([
    {"Peak Time (s)": peak_times,
     "Peak Value": peak_values}
])
df_peaks.to_csv(peak_data_path, index=False)

peak_data_path

Out[1]: 'peaks_data.csv'
```

左図は、ChatGPTから得たPythonコードをPython実行環境に貼り付け、実際の自身のデータを解析する画面です。

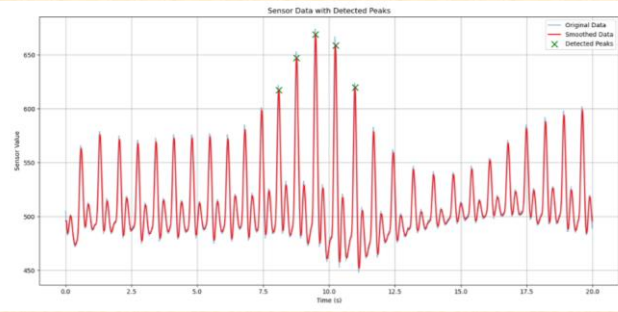
細かい説明は割愛しますが、脈波の値が600以上のピークを検出するコードになっているため、大きい波の方のピークを上手く検出できそうです。

結果を見てみましょう。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

92

⑤ 実際のデータをChatGPTを通さずに解析



左図が結果です。
Pythonのコードが
600以上のピークを検
出するコードになっ
ているため、全ての
ピークを検出するこ
とが出来ませんでした。

再度、閾値を大きい
脈波のピークに合う
ように再度Python
コードを修正して試
してみましょう。

上手くいくでしょ
うか？

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

93

⑤ 実際のデータをChatGPTを通さずに解析

```
In [3]: import numpy as np
import pandas as pd
from scipy.signal import find_peaks

# Load the data
df = pd.read_csv(file_path)
df = pd.read_csv("sensor_data.csv")

# Apply a moving average for smoothing
window_size = 5 # Define the window size for moving average
df["Smoothed_Value"] = df["Sensor Value"].rolling(window=window_size).mean().fillna(method="bfill")

# Detect peaks
peaks, _ = find_peaks(df["Smoothed_Value"].values, height=600)

# Extract peak positions and values
peak_times = df.iloc[peaks]["Time (s)"].values
peak_values = df.iloc[peaks]["Smoothed_Value"].values

# Save to CSV
peak_data_path = "peaks_data.csv"
df_peaks = pd.DataFrame({
    "Peak Time (s)": peak_times,
    "Peak Value": peak_values
})
df_peaks.to_csv(peak_data_path, index=False)

peak_data_path

Out[3]: 'peaks_data.csv'
```

もう一度Pythonコードに戻っ
て閾値を600以上から530以上
に設定しなおしました。

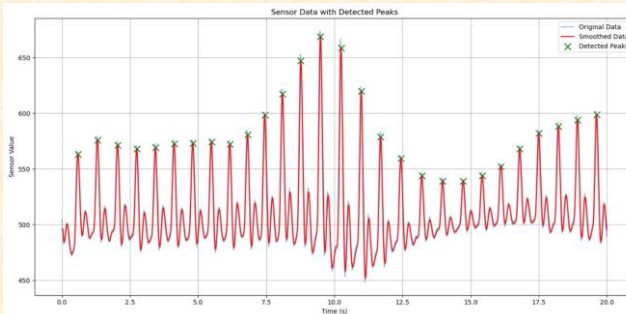
今度は上手くいくでしょうか？

次ページにて結果を見てみま
しょう。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

94

⑤ 実際のデータをChatGPTを通さずに解析



左図が結果です。
今度は上手くいきま
した。
お題である「心拍毎
の大きいピークを検
出」する、という部
分まで完了しました。

最後に、その値と時
間の値を表にまとめ
た結果を次ページに
示します。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

95

⑤ 実際のデータをChatGPTを通さずに解析

#	A	B
1	Peak Time (s)	Peak Value
2	0.6	563.2
3	1.31	576
4	2.05	571.6
5	2.75	566
6	3.44	569.4
7	4.12	572.8
8	4.81	573
9	5.49	574.2
10	6.17	572.2
11	6.82	580.8
12	7.45	586.6
13	8.1	637.2
14	8.77	647.2
15	9.49	669
16	10.25	658.8
17	10.99	619.8
18	11.7	578.8
19	12.45	559.4
20	13.21	544
21	13.97	539
22	14.73	539.2
23	15.43	544.2
24	16.11	552.4
25	16.8	566
26	17.5	582

左図が目的とするピークに関する脈波の値とその時間の値の表です。

今回、研究者は脈波の値とその時間の間隔について調べたかったので無事に結果を得ることが出来ました。

Copyright © 2023 Seine A. Shintani Chubu Univ. All Rights Reserved

96

⑥まとめ：AIを使いこなすための新たな指針 (AI2L)

今回の講演では情報リテラシーに関してかなり厳密に情報管理する方法を紹介しました。

しかし、ダミーデータを作成するというワンステップをはさむことによって、自分の貴重なデータが第三者に利用されない安心感や、研究協力者の個人情報保護の観点からは重要な指針でした。

AI2Lでは、このように、自身の研究においてもピーク検出の為に膨大な時間やプログラムコードを作成する手間を相当省けたことが分かったかと思います。

また、今まで既存の計測機器に備え付けの解析ソフトにライセンスが切れる度に何十万円や何百万円も払っていたケースでも自分自身で解析が出来てしまうという利点があります。研究者にとって最も重要な「時間」「労力」「費用」を上手く効率化できる簡単な例として紹介しました。

大学のAI数理の利用者の数が増えることにより、大学全体の研究力向上と、皆様の研究に貢献できれば幸いです。
以上です。