

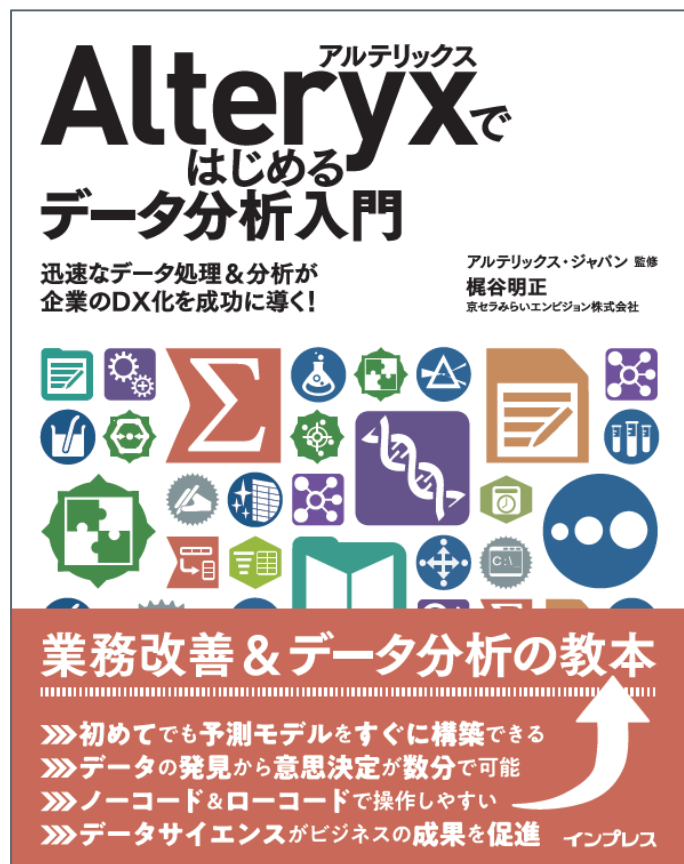
データサイエンス100本ノックをAlteryxで。

Akimasa Kajitani@京セラみらいエンビジョン | Alteryx ACE

2022/11/22

「Alteryxではじめるデータ分析入門」ご紹介

Alteryx初心者向けの書籍となっております、体系的にAlteryxを学習できる入門書となっております。
内容としては、Alteryx Core 認定レベルの内容は網羅しておりますが、実際に企業内で運用する際にどういうことができるのか、ということをご理解頂けることを念頭に置き、マクロ、分析アプリなどの実践的な内容まで含んでおります。



書籍詳細情報

- 著者：梶谷明正 & 京セラみらいエンビジョン株式会社
- 出版社：株式会社インプレス
- 発売日：2022/4/1
- ページ数：260ページ、カラー
- 寸法：B5変形
- 価格：6,600円（税込）
- 購入先：[Amazon](#)
- 対応バージョン：Alteryx 2021.3



データサイエンス100本ノックをAlteryxでやってみる

ウィークリーチャレンジ、難しすぎませんか？

- 初学者にウィークリーチャレンジは思った以上に難しい
 - 初級問題でも触り始めた人だと1時間とか平気でかかる
 - たまにある意味不明の設問・わかりにくい問題文・・・
- インタラクティブレッスンだけだと明らかに学習量不足
 - さわって、ワークフロー作ってなんぼ
- **インタラクティブレッスンとウィークリーチャレンジの間をつなぐものが欲しい**
- Twitter見てたら「データサイエンス100本ノック（構造化データ加工編）」という無料コンテンツが！
 - これは使える！



データサイエンス100本ノックをAlteryxでやってみる

データサイエンス100本ノックとは？

- 一般社団法人データサイエンティスト協会が**無償**でGitHubにて公開している学習コンテンツ
- 実践的な課題に取り組みながらプログラミング、データ分析スキルをデータサイエンス初学者が習得することを目的としている
- 対象はSQL、Python、R
- 関連書籍も発売されています
 - データサイエンス100本ノック構造化データ加工編ガイドブック
- 利用条件は「**大学、企業など組織でのご利用にあたっては、「データサイエンティスト協会スキル定義委員」の「データサイエンス100本ノック（構造化データ加工編）」を利用して**いることを明示いただければ自由に利用してOK」とのこと。

データサイエンス100本ノックをAlteryxでやってみる

Alteryxでも活用しましょう！

- Alteryx向けにウィークリーチャレンジ方式で問題・解答を無償で公開しました
個人ホームページ
<https://analytics-x.tech/archives/294>



- Alteryx Communityでも5問程度にまとめてディスカッションにてウィークリーチャレンジ方式で公開しています
 - 1~4問目 : <https://bit.ly/3SxqVl3>
 -
 -
 -

ウィークリーチャレンジ方式の出題

コミュニティ > ディスカッション > ラベルのついたディスカッション: データサイエンス100本ノック

🗨️ ラベルのついたディスカッション: データサイエンス100本ノック

ラベルデータサイエンス100本ノックを持つトピックを表示 購読 (このラベル)。すべてのトピックを表示する

	データサイエンス100本ノック 構造化データ加工編をAlteryxでやってみる (87問目~88問目: ...   投稿者:  AkimasaKajitani 日曜日・最終投稿 昨日、投稿者: Yoshiro_Fujimori	2 件の返信	70 件の閲覧回数
	データサイエンス100本ノック 構造化データ加工編をAlteryxでやってみる (84問目~86問目: ...   投稿者:  AkimasaKajitani 2週間前・最終投稿 土曜日、投稿者:  AkimasaKajitani	3 件の返信	123 件の閲覧回数
	データサイエンス100本ノック 構造化データ加工編をAlteryxでやってみる (79問目~83問目: ...   投稿者:  AkimasaKajitani 2週間前・最終投稿 2週間前、投稿者:  AkimasaKajitani	2 件の返信	166 件の閲覧回数
	データサイエンス100本ノック 構造化データ加工編をAlteryxでやってみる (75問目~78問目: ...   投稿者:  AkimasaKajitani 3週間前・最終投稿 2週間前、投稿者: Yoshiro_Fujimori	4 件の返信	187 件の閲覧回数
	データサイエンス100本ノック 構造化データ加工編をAlteryxでやってみる (68問目~74問目: ...   投稿者:  AkimasaKajitani、投稿日時: 10-16-2022 08:36 AM・最終投稿 4週間前、投稿者:  AkimasaKajitani	3 件の返信	205 件の閲覧回数
	データサイエンス100本ノック 構造化データ加工編をAlteryxでやってみる (63問目~67問目: ...   投稿者:  AkimasaKajitani、投稿日時: 10-09-2022 11:40 AM・最終投稿日時: 10-15-2022 10:33 AM、投稿者:  AkimasaKajitani	3 件の返信	223 件の閲覧回数

ラベル

ADAPT Program	2
Admin UI	1
Apps	1
basic data profile	1
CSV	1
dcmh8twy229t4@d6.gmob...	1
Designer	1
Error	1
Excel	1
Gallery	1
GIS	1
Help	1
importer	1



ウィークリーチャレンジ方式の出題



68 問目

P-068: 商品データ (product.csv) の各商品について、消費税率10%の税込み金額を求めよ。1円未満の端数は切り捨てとし、結果を10件表示せよ。ただし、単価 (unit_price) には欠損が生じていることに注意せよ。

※税込み金額はtax_priceというフィールド名とすること



69 問目

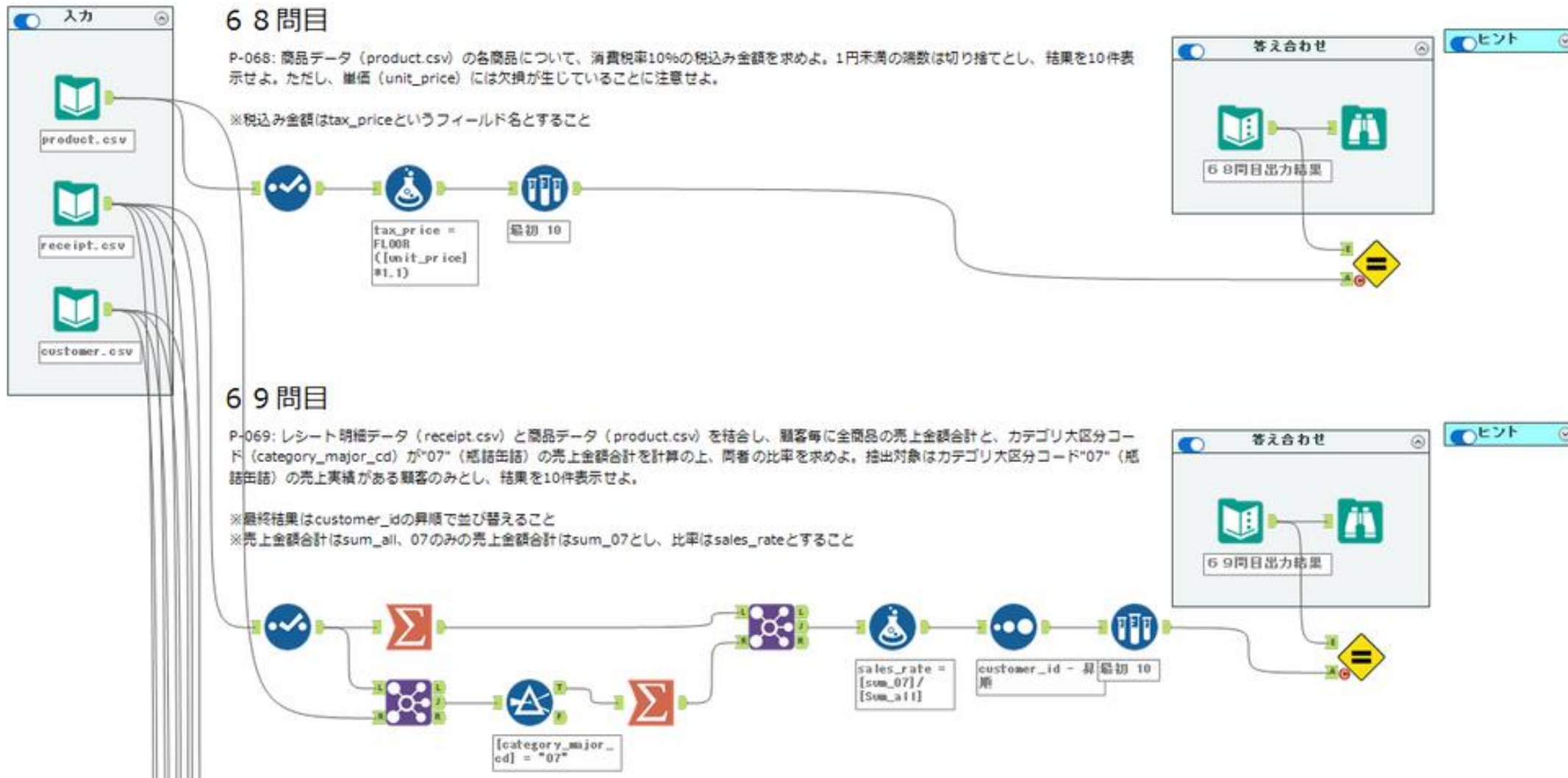
P-069: レシート明細データ (receipt.csv) と商品データ (product.csv) を結合し、顧客毎に全商品の売上金額合計と、カテゴリ大区分コード (category_major_cd) が"07" (瓶詰缶詰) の売上金額合計を計算の上、両者の比率を求めよ。抽出対象はカテゴリ大区分コード"07" (瓶詰缶詰) の売上実績がある顧客のみとし、結果を10件表示せよ。

※最終結果はcustomer_idの昇順で並び替えること

※売上金額合計はsum_all、07のみの売上金額合計はsum_07とし、比率はsales_rateとすること



ウィークリーチャレンジ方式で解答



データサイエンス100本ノックをAlteryxでやってみる

どのようなことが学べるのか？

- 基本的なデータ加工、統計処理等

1. 列に対する操作
2. 行に対する操作
3. あいまい条件
4. ソート
5. 集計
6. 副問合せ
7. 結合
8. 縦横変換
9. データ変換
10. 数値変換
11. 四則演算
12. 日付型の計算
13. サンプリング
14. 外れ値・異常値
15. 欠損値
16. 乗算エラー対応
17. 座標データ
18. 名寄せ
19. データ分割
20. 不均衡データ
21. 正規化・非正規化
22. ファイル入出力

※データサイエンス100本ノックのマニュアルより

どのツールが学べるのか？

- 8カテゴリ37ツールが学べます

準備	17
結合	8
パース	3
データ調査	2
空間	2
入出力	2
変換	2
開発者	1
総計	37

インピュテーション オートフィールド
オーバーサンプルフィールド クロスタブ
サンプリング サンプル作成 セレクト ソート
タイル データクレンジング データ出力
データ入力 テキスト入力 フィールドサマリー
フィールド付加 フィルター フォーミュラ
ポイント作成 ユニーク ユニオン
ランダム%サンプリング レコードID
レコードカウント レコード選択 閲覧
基本データプロファイル 距離 結合
検索置換 集計 正規表現 転置
動的リネーム 日時 複数フィールドフォーミュラ
複数結合 複数行フォーミュラ 列分割

データサイエンス100本ノックをAlteryxでやってみる

どのようなことが学べるのか？（詳細）

- 項目の削除、項目名の変更
- ソート
- サンプルング
- 抽出（一致、前方一致、後方一致、値比較、桁数一致、TopN、ランダム、層化抽出、外れ値）
- 順位付け
- レコード数カウント
- ユニーク化
- 四則演算
- レコード間計算
- 集計（合計、最頻値、平均、中央値、分散、標準偏差、パーセンタイル）
- 日付（最新、最古、形式変換、年月日の個別抽出、経過日数・月数・年数・秒数計算、曜日）
- 結合（内部結合、完全外部結合、レコードをずらして結合）
- データ補完（Nullを0に、平均値、中央値）
- 組み合わせ取得
- クロス集計
- 横持ち、縦持ち変換
- 階級
- 二値化、コード値
- 四分位
- ダミー変数化
- 標準化、正規化
- 常用対数化、自然対数化
- 消費税計算
- 小数点処理（切り捨て、切り上げ、四捨五入）
- IQR
- 欠損数確認
- 欠損データ削除
- 割合計算
- 距離計算
- 名寄せ
- 予測モデル作成のためのデータ分割
- アンダーサンプリング
- 第三正規化、非正規化
- ファイル入力、出力（CSV、TSV、文字コード）

データサイエンス100本ノックをAlteryxでやってみる

データサイエンス100本ノックと他コンテンツの棲み分けとは？

- 難易度：ウィークリーチャレンジ >>> データサイエンス100本ノック > インタラクティブレッスン
- それぞれの違いとは？
 - ウィークリーチャレンジ：具体的なユースケース中心。難易度は初級から上級まで幅広い。ただし、わかりにくくて時間がかかる問題等もあり。
 - データサイエンス100本ノック：網羅的、実践的、ややくどいところもあるが、データ処理から統計まで一通りカバーしている。
 - インタラクティブレッスン：網羅されているが基礎的、ツールの使い方がメイン。
- どう進めるのが良いのか？
 - インタラクティブレッスン→データサイエンス100本ノック→ウィークリーチャレンジ

Alteryxの基礎トレーニングとしてぜひご利用いただければと思います。

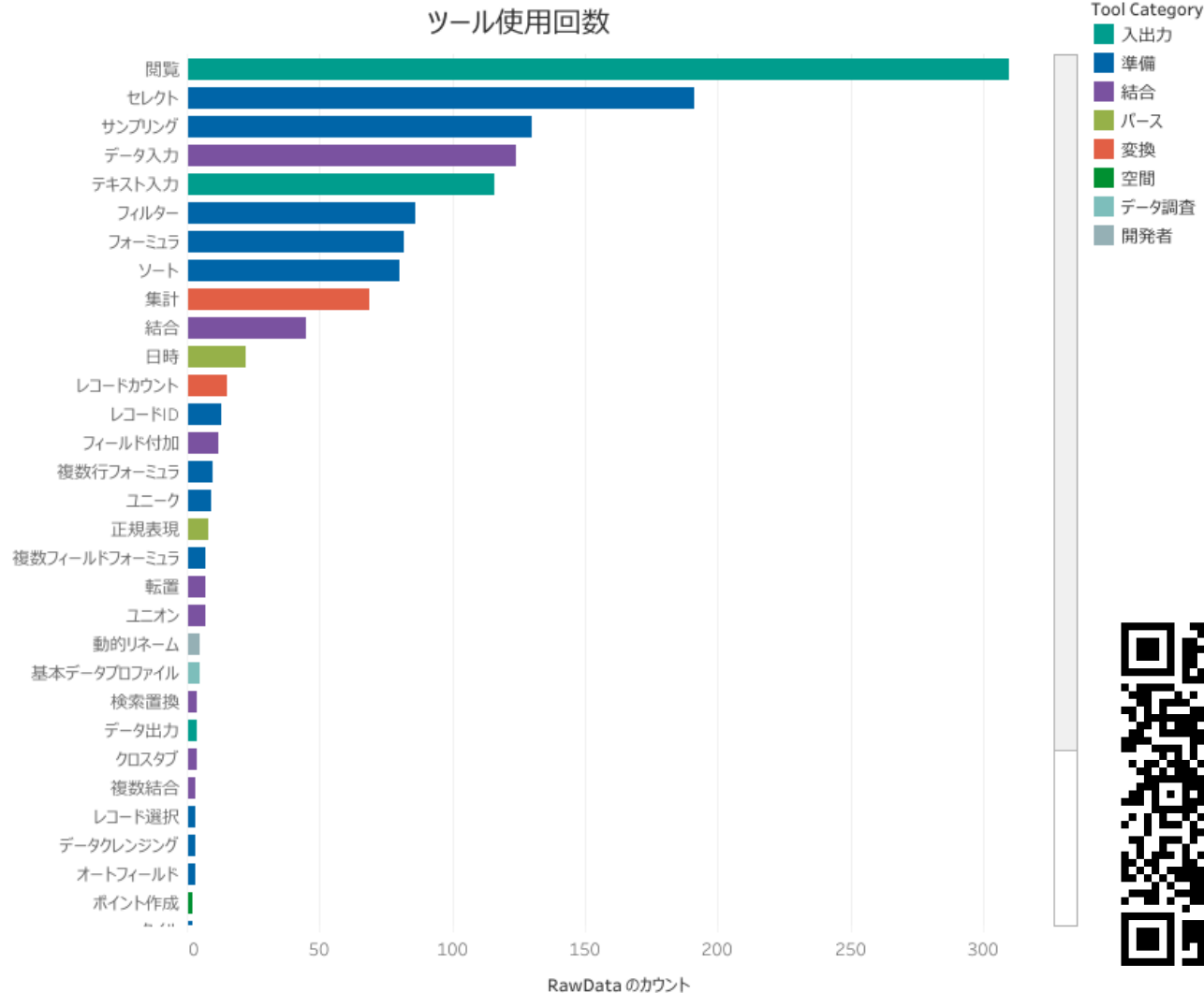
データサイエンス100本ノックをAlteryxでやってみる

データサイエンス100本ノックをAlteryxでやってみた結果。

問題文

001	002	003	004	005	006	007	008	009	010
011	012	013	014	015	016	017	018	019	020
021	022	023	024	025	026	027	028	029	030
031	032	033	034	035	036	037	038	039	040
041	042	043	044	045	046	047	048	049	050
051	052	053	054	055	056	057	058	059	060
061	062	063	064	065	066	067	068	069	070
071	072	073	074	075	076	077	078	079	080
081	082	083	084	085	086	087	088	089	090
091	092	093	094	095	096	097	098	099	100

難易度
 ★
 ★★
 ★★★



おまけ：アイデアにぜひ投票ください

英語版コミュニティには、新規アイデアを募集する掲示板があります。ここに日本ユーザー向けの機能についてアップしているのでぜひご投票ください。

Date tool preset at Japanese

- <https://community.alteryx.com/t5/Alteryx-Designer-Ideas/Date-tool-preset-at-Japanese/idc-p/1002419#>
- 日時ツールにはなぜか日本でよく使われている「YYYY/MM/DD」形式がありません。これをプリセットに追加して欲しい。



Allow 2 byte character field name at Shape File

- <https://community.alteryx.com/t5/Alteryx-Designer-Ideas/Allow-2-byte-character-field-name-at-Shape-File/idi-p/617109>
- Shapeファイルという空間データのデファクトスタンダードの形式があるのですが、多くのGISツールではフィールド名に日本語が使えるのにDesignerでは使えません。日本語が使いたい！



おまけ：新認定MicroCredential

今年に入り、新しくMicro-Credentialシリーズの認定が始まっています。このマイクロクレデンシャルシリーズは、特定の分野での能力を証明するために設計された短縮版の認定です。



Alteryx Foundation Micro-Credential (Designer基礎)

- データ分析と基本的な Designerの基本的な理解



Alteryx Machine Learning Fundamentals Micro-Credential (機械学習基礎)

- 機械学習の基礎に関する知識
 - 機械学習の概念
 - データの準備と分析
 - モデルの評価と解釈

コミュニティ > 学ぶ > アカデミー > [認定](#)
にて無料で受験可能です (いずれも1時間で40問の試験です)。



alteryx

ありがとうございました

Akimasa Kajitani@KYOCERA Mirai Envision | Alteryx ACE
akimasa-kajitani@kcme.jp

alteryx.com