

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2019年6月27日(27.06.2019)

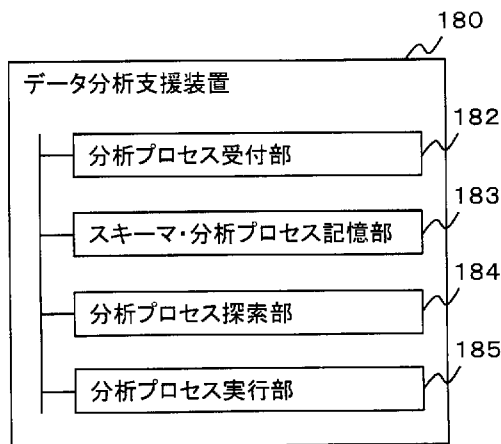


(10) 国際公開番号
WO 2019/123703 A1

- (51) 国際特許分類:
G06F 17/30 (2006.01)
- (21) 国際出願番号: PCT/JP2018/028082
- (22) 国際出願日: 2018年7月26日(26.07.2018)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
62/609,768 2017年12月22日(22.12.2017) US
- (71) 出願人: 日本電気株式会社 (NEC CORPORATION) [JP/JP]; 〒1088001 東京都港区芝五丁目7番1号 Tokyo (JP).
- (72) 発明者: 藤巻 遼平 (FUJIMAKI Ryohei); 〒1088001 東京都港区芝五丁目7番1号 日本電気株式会社内 Tokyo (JP). 楠村 幸貴 (KUSUMURA Yukitaka); 〒1088001 東京都港区芝五丁目7番1号 日本電気株式会社内 Tokyo (JP). 村岡 優輔 (MURAOKA Yusuke); 〒1088001 東京都港区芝五丁目7番1号 日本電気株式会社内 Tokyo (JP).
- (74) 代理人: 岩壁 冬樹, 外(IWAKABE Fuyuki et al.); 〒1040031 東京都中央区京橋二丁目8番7号 読売八重洲ビル6階 サンライズ国際特許事務所 Tokyo (JP).
- (81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS,

(54) Title: DATA ANALYSIS ASSISTANCE DEVICE, DATA ANALYSIS ASSISTANCE METHOD, AND DATA ANALYSIS ASSISTANCE PROGRAM

(54) 発明の名称: データ分析支援装置、データ分析支援方法およびデータ分析支援プログラム



- 180 Data analysis assistance device
- 182 Analysis process reception unit
- 183 Schema/analysis process storage unit
- 184 Analysis process search unit
- 185 Analysis process execution unit

(57) Abstract: An analysis process reception unit 182 receives creation of an analysis process which is a series of processing for data analysis using a column name defined by a schema to be applied to a table. A schema/analysis process storage unit 183 stores information in which the received analysis process is associated with a schema to which the analysis process can be applied. When a selection of a table is received from a user, an analysis process search unit 184 outputs a list of analysis processes which can be applied to the received table, on the basis of information stored in a table/schema storage unit and the information stored in the schema/analysis process storage unit 183. An analysis process execution unit 185 receives the selection of an analysis process from the outputted list, and executes the selected analysis process for the received table.

WO 2019/123703 A1

SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM,
GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類：

- 一 国際調査報告（条約第21条(3)）

(57) 要約：分析プロセス受付部182は、テーブルに適用されるスキーマで定義されているカラム名を用いた、データ分析のための一連の処理である分析プロセスの作成を受け付ける。スキーマ・分析プロセス記憶部183は、受け付けた分析プロセスと、その分析プロセスを適用可能なスキーマとを関連付けた情報を記憶する。分析プロセス探索部184は、テーブルの選択をユーザから受け付けると、テーブル・スキーマ記憶部が記憶する情報、および、スキーマ・分析プロセス記憶部183が記憶する情報に基づいて、受け付けたテーブルに対して適用可能な分析プロセスの一覧を出力する。分析プロセス実行部185は、出力された一覧から分析プロセスの選択を受け付け、受け付けたテーブルに対して選択された分析プロセスを実行する。

明 細 書

発明の名称：

データ分析支援装置、データ分析支援方法およびデータ分析支援プログラム

技術分野

[0001] 本発明は、リレーショナルデータベースを用いたデータの分析を支援するデータ分析支援装置、データ分析支援方法およびデータ分析支援プログラムに関する。

背景技術

[0002] 既存のデータを用いて様々な分析が行われている。特に、データの管理にはリレーショナルデータベース（以下、RDBと記す。）が多く用いられており、RDBを用いた様々なデータ処理方法も提案されている。

[0003] 例えば、特許文献1には、RDBで管理されているデータから、機械学習処理に用いられる特徴量の候補を生成することが記載されている。特許文献1に記載された方法では、特徴量の候補を生成する処理を、Filter条件、map条件およびreduce条件の3つの条件の組合せにより定義することで、特徴量の候補を生成する分析者工数を削減する。

先行技術文献

特許文献

[0004] 特許文献1：国際公開第2017/090475号

発明の概要

発明が解決しようとする課題

[0005] RDBでは、スキーマとテーブルとが一対一に対応し、各テーブルを対象としてデータの分析処理が記述される。言い換えると、同一の構造を有するテーブルが存在する場合、テーブルが異なれば、それぞれのテーブルに含まれるデータに対する分析処理は異なるものとして記述される。

[0006] 検索処理の性能の向上させる観点や、データを分散して管理する観点などから、同じ内容を表す情報が同一のスキーマで定義された複数のテーブルで管理される場合がある。このような環境では、同じ内容を表す情報に対して同じ分析処理を記述しようとしても、テーブルごとに異なる分析処理を記述しなければならないという問題がある。

[0007] 例えば、特許文献1に記載された方法では、分析の対象とするテーブルが異なると、記述する条件の内容や、生成する特徴量生成関数の内容もそれぞれ異なることになる。しかし、同じ内容を含む異なるテーブルに対して、それぞれ異なる分析処理を記述するのは煩雑である。そのため、あるテーブルのデータに対して定義される分析処理を、同様の構造を有する他のテーブルに対しても利用できることが好ましい。

[0008] そこで、本発明は、一のテーブルに対して定義された分析処理を異なるテーブルに対しても実行できるデータ分析支援装置、データ分析支援方法およびデータ分析支援プログラムを提供することを目的とする。

課題を解決するための手段

[0009] 本発明によるデータ分析支援装置は、テーブルに適用されるスキーマで定義されているカラム名を用いた、データ分析のための一連の処理である分析プロセスの作成を受け付ける分析プロセス受付部と、受け付けた分析プロセスと、その分析プロセスを適用可能なスキーマとを関連付けた情報を記憶するスキーマ・分析プロセス記憶部と、テーブルの選択をユーザから受け付けると、テーブルとそのテーブルに適用されるスキーマとを関連付けた情報を記憶するテーブル・スキーマ記憶部が記憶する情報、および、スキーマ・分析プロセス記憶部が記憶する情報に基づいて、受け付けたテーブルに対して適用可能な分析プロセスを特定し、特定された分析プロセスの一覧を出力する分析プロセス探索部と、出力された一覧から分析プロセスの選択を受け付け、受け付けたテーブルに対して選択された分析プロセスを実行する分析プロセス実行部とを備えたことを特徴とする。

[0010] 本発明によるデータ分析支援方法は、テーブルに適用されるスキーマで定

義されているカラム名を用いた、データ分析のための一連の処理である分析プロセスの作成を受け付け、受け付けた分析プロセスと、その分析プロセスを適用可能なスキーマとを関連付けた情報を、スキーマ・分析プロセス記憶部に登録し、テーブルの選択をユーザから受け付けると、テーブルとそのテーブルに適用されるスキーマとを関連付けた情報を記憶するテーブル・スキーマ記憶部が記憶する情報、および、スキーマ・分析プロセス記憶部が記憶する情報に基づいて、受け付けたテーブルに対して適用可能な分析プロセスを特定し、特定された分析プロセスの一覧を出力し、出力された一覧から分析プロセスの選択を受け付け、受け付けたテーブルに対して選択された分析プロセスを実行する

ことを特徴とする。

- [0011] 本発明によるデータ分析支援プログラムは、コンピュータに、テーブルに適用されるスキーマで定義されているカラム名を用いた、データ分析のための一連の処理である分析プロセスの作成を受け付け、受け付けた分析プロセスとその分析プロセスを適用可能なスキーマとを関連付けた情報をスキーマ・分析プロセス記憶部に登録する分析プロセス受付処理、テーブルの選択をユーザから受け付けると、テーブルとそのテーブルに適用されるスキーマとを関連付けた情報を記憶するテーブル・スキーマ記憶部が記憶する情報、および、スキーマ・分析プロセス記憶部が記憶する情報に基づいて、受け付けたテーブルに対して適用可能な分析プロセスを特定し、特定された分析プロセスの一覧を出力する分析プロセス探索処理、および、出力された一覧から分析プロセスの選択を受け付け、受け付けたテーブルに対して選択された分析プロセスを実行する分析プロセス実行処理を実行させることを特徴とする。

発明の効果

- [0012] 本発明によれば、一のテーブルに対して定義された分析処理を異なるテーブルに対しても実行できる。

図面の簡単な説明

- [0013] [図1]本発明によるデータ分析支援装置の第1の実施形態の構成例を示すブロック図である。
- [図2]スキーマ付テーブルからスキーマを抽出する処理の例を示す説明図である。
- [図3]テーブル・スキーマ管理DB30が記憶する情報の例を示す説明図である。
- [図4]分析プロセスを作成する処理の例を示す説明図である。
- [図5]分析プロセスと、その分析プロセスを適用可能なスキーマとを関連付けた情報の例を示す説明図である。
- [図6]分析プロセスを出力する処理の例を示す説明図である。
- [図7]分析プロセスを実行する処理の例を示す説明図である。
- [図8]テーブルを出力する処理の例を示す説明図である。
- [図9]第1の実施形態のデータ分析支援装置を用いて分析プロセスを実行する動作例を示すフローチャートである。
- [図10]第1の実施形態のデータ分析支援装置を用いて分析プロセスを実行する他の動作例を示すフローチャートである。
- [図11]スキーマを管理する動作例を示すフローチャートである。
- [図12]本発明によるデータ分析支援装置の第2の実施形態の構成例を示すブロック図である。
- [図13]列の内容に応じて分析データ型を設定した例を示す説明図である。
- [図14]分析スキーマを抽出する処理の例を示す説明図である。
- [図15]スキーマを管理する動作例を示すフローチャートである。
- [図16]本発明によるデータ分析支援装置の概要を示すブロック図である。

発明を実施するための形態

- [0014] 以下、本発明の実施形態を図面を参照して説明する。なお、以下の説明において、テーブルとは、表形式のデータセット（表型情報）を意味するものとし、スキーマと一体になったテーブル（すなわち、スキーマとテーブルとが関連付けられたもの）のことを、スキーマ付テーブルと記す。また、本発

明においてスキーマとは、テーブルの属性（フィールド、列）を定義した情報であり、属性として、テーブルに含まれる列のカラム名、データ型、制約などが挙げられる。

[0015] 実施形態 1.

図 1 は、本発明によるデータ分析支援装置の第 1 の実施形態の構成例を示すブロック図である。本実施形態のデータ分析支援装置 100 は、スキーマ付テーブル入力部 10 と、スキーマ抽出部 20 と、テーブル・スキーマ管理データベース 30（以下、テーブル・スキーマ管理 DB 30 と記す。）と、分析プロセス受付部 40 と、スキーマ・分析プロセス管理データベース 50（以下、スキーマ・分析プロセス管理 DB 50 と記す。）と、探索部 60 と、分析プロセス実行部 70 とを備えている。

[0016] なお、テーブル・スキーマ管理 DB 30 と、スキーマ・分析プロセス管理 DB 50 とは、具体的には、磁気ディスク装置等に記憶される。

[0017] スキーマ付テーブル入力部 10 は、スキーマ付テーブルを入力する。スキーマ付テーブル入力部 10 は、例えば、RDB が提供するインタフェースを介して、直接 RDB からスキーマ付テーブルを入力してもよい。また、スキーマ付テーブル入力部 10 は、スキーマおよびテーブルの内容が関連付けられたファイルを読み込んでもよい。

[0018] スキーマ抽出部 20 は、スキーマ付テーブルからスキーマを抽出し、抽出されたスキーマと、テーブルとを関連付けてテーブル・スキーマ管理 DB 30 に登録する。図 2 は、スキーマ付テーブルからスキーマを抽出する処理の例を示す説明図である。図 2 に例示するスキーマ付テーブル ST 1 は、2016 年 1 月の顧客リストを表すスキーマ付テーブルであり、スキーマ SC 1 と表型情報であるテーブル TB 1 とを含む。

[0019] スキーマ付テーブル入力部 10 が、図 2 に例示するスキーマ付テーブル ST 1 を入力したとする。このとき、スキーマ抽出部 20 は、スキーマ付テーブル ST 1 から、カラム名、データ型および制約を含むスキーマ SC 1 を抽出する。ただし、スキーマ抽出部 20 が抽出するスキーマの情報は、図 2 に

例示する情報に限定されない。スキーマ抽出部20は、表の属性を表す他の情報を含むスキーマを抽出してもよい。

[0020] なお、テーブル・スキーマ管理DB30に登録する際、スキーマ抽出部20は、カラムの名称およびデータ型が一致するスキーマが登録されていない場合に、抽出されたスキーマを新たなスキーマとしてテーブル・スキーマ管理DB30に登録する。さらに、スキーマ抽出部20は、カラムの名称およびデータ型だけでなく、制約まで一致するスキーマが登録されていない場合に、抽出されたスキーマを新たなスキーマとしてテーブル・スキーマ管理DB30に登録してもよい。

[0021] スキーマ抽出部20は、スキーマを識別する任意の識別子を設定する。図2に示す例では、連番としてスキーマSC1に識別子“001”が設定されている。なお、スキーマ識別子は、図2に例示する数値に限定されない。スキーマ抽出部20は、例えば、ユーザからスキーマ名の指定（例えば、「顧客リスト」など）を受け付け、その指定をスキーマ名として用いてもよい。

[0022] テーブル・スキーマ管理DB30は、スキーマとテーブルとを関連付けて記憶する。テーブル・スキーマ管理DB30は、例えば、スキーマ名とテーブル名とを対応付けて記憶する。

[0023] 図3は、テーブル・スキーマ管理DB30が記憶する情報の例を示す説明図である。図3に示す例では、テーブル・スキーマ管理DB30がテーブル名とスキーマ名とを関連付けて記憶していることを示す。また、図3に示す例では、2016年1月の顧客リストテーブル（顧客リスト2016／1テーブル）のスキーマと、2016年2月の顧客リストテーブル（顧客リスト2016／2テーブル）のスキーマとに、それぞれ同一のスキーマ（スキーマ001）が適用されていることを示す。

[0024] なお、スキーマ付テーブル入力部10、スキーマ抽出部20およびテーブル・スキーマ管理DB30によって、テーブルとスキーマとを分離して管理できることから、スキーマ付テーブル入力部10、スキーマ抽出部20およびテーブル・スキーマ管理DB30を含む装置99を、スキーマ管理装置と

言うことが出来る。なお、本実施形態では、データ分析支援装置100が、スキーマ管理装置を含む場合を例示している。ただし、データ分析支援装置100は、スキーマ管理装置を含んでいなくてもよい。例えば、データ分析装置が外部に存在し、データ分析支援装置100が、外部に存在するデータ分析装置に接続されて各情報を取得するようにしてもよい。

[0025] 分析プロセス受付部40は、スキーマで定義されているカラム名を用いた分析プロセスの作成を受け付ける。分析プロセスとは、テーブルのデータに対して行う一連の処理である。ただし、本実施形態では、テーブルとは切り離れたスキーマをもとに分析プロセスが作成される。分析プロセス受付部40は、予め作成された分析プロセスを受け付けてもよく、分析プロセスを作成するための画面を表示し、ユーザの入力に基づいて作成された分析プロセスを受け付けてもよい。

[0026] 図4は、分析プロセスを作成する処理の例を示す説明図である。例えば、顧客リストの内容に基づいて、各顧客がランクアップするか否か判断する分析（以下、ランクアップ回帰分析）を行うための分析プロセスを作成するとする。また、図4に示す例では、図2に例示するスキーマSC1（スキーマ001）が適用されるテーブルのデータを用いて分析が行われるものとする。

[0027] 例えば、機械学習では、入力データを数値にする必要がある。図2に示す例では、性別のデータ型が `varchar` 型であり、データの内容がMまたはFで表されている。そこで、分析プロセス受付部40は、スキーマ001に含まれる性別のデータを変換する処理P1（例えば、Mを1に、Fを0に変換する処理）を作成してもよい。また、分析プロセス受付部40は、ユーザの属性からランクアップを判別するための回帰式（例えば、`logit`（ランクアップ）=年齢×3+性別+1、など）を用いた判別処理P2を作成してもよい。そして、分析プロセス受付部40は、作成した一連の処理を分析プロセスAP1として受け付ける。

[0028] 分析プロセス受付部40は、作成した分析プロセスをスキーマ・分析プロ

セス管理DB50に登録する。分析プロセス受付部40は、内容が把握できるような名称を分析プロセスに付与して、スキーマ・分析プロセス管理DB50に登録してもよい。例えば、図4に示す例では、分析プロセス受付部40は、「顧客リストに対するランクアップ回帰分析プロセス」のような名称を分析プロセスに付与して、スキーマ・分析プロセス管理DB50に登録してもよい。

[0029] なお、後述する分析プロセス実行部70が処理を実行できる形式であれば、分析プロセスの表現方法は任意である。分析プロセスは、例えば、スクリプトの形式で表現されていてもよい。

[0030] 以上のように、分析プロセス受付部40が、テーブルの定義を含む分析プロセスではなく、スキーマで定義されているカラム名を用いた分析プロセスの作成を受け付ける。そのため、分析対象のテーブルが異なってもスキーマが同一である場合には、同じ内容の分析プロセスを再利用できる。

[0031] スキーマ・分析プロセス管理DB50は、分析プロセスと、その分析プロセスを適用可能なスキーマとを関連付けた情報を記憶する。図5は、分析プロセスと、その分析プロセスを適用可能なスキーマとを関連付けた情報の例を示す説明図である。例えば、図4に例示する分析プロセスは、スキーマ001を用いて定義されており、スキーマ001が適用されるプロセスと言える。そこで、スキーマ・分析プロセス管理DB50は、図5に例示する表の1行目に示すように、図4に例示する分析プロセスと、スキーマ001とを対応付けて記憶する。

[0032] 探索部60は、ユーザからの選択を受け付けて各種情報を探索し、出力する。探索部60は、分析プロセス探索部61と、テーブル探索部62とを含む。

[0033] 分析プロセス探索部61は、テーブルの選択をユーザから受け付ける。分析プロセス探索部61は、テーブル・スキーマ管理DB30が記憶する情報から、受け付けたテーブルに関連付けられているスキーマを抽出する。そして、分析プロセス探索部61は、スキーマ・分析プロセス管理DB50が記

憶する情報から、抽出したスキーマに関連付けられている分析プロセスを特定し、出力する。

[0034] テーブル探索部62は、分析プロセスの選択をユーザから受け付ける。テーブル探索部62は、スキーマ・分析プロセス管理DB50が記憶する情報から、受け付けた分析プロセスに関連付けられているスキーマを抽出する。そして、テーブル探索部62は、テーブル・スキーマ管理DB30が記憶する情報から、抽出したスキーマに関連付けられているテーブルを特定し、出力する。

[0035] 分析プロセス実行部70は、選択されたテーブルに対して分析プロセスを実行する。以下、分析プロセス実行部70が分析プロセスを実行する2つの方法を説明する。

[0036] 探索部60（具体的には、分析プロセス探索部61）は、テーブルの選択をユーザから受け付けた場合に、分析プロセスを出力する。この場合、分析プロセス実行部70は、出力された分析プロセスの一覧から、ユーザの所望する分析プロセスの選択を受け付ける。そして、分析プロセス実行部70は、受け付けたテーブルに対して選択された分析プロセスを実行する。

[0037] 図6は、分析プロセスを出力する処理の例を示す説明図である。探索部60が2016年2月の顧客リストを表す図6に例示するスキーマ付テーブルST2の選択をユーザから受け付けると、分析プロセス探索部61は、図3に例示するテーブル・スキーマ管理DB30が記憶する情報から、受け付けたテーブルに関連付けられているスキーマ001を抽出する。そして、分析プロセス探索部61は、図5に例示するスキーマ・分析プロセス管理DB50が記憶する情報から、抽出したスキーマ001に関連付けられている分析プロセスを特定し、出力する。ここでは、「顧客リストに対するランクアップ回帰分析プロセス」と、「顧客リストに対する性別判別分析プロセス」の2つの分析プロセスが出力される。

[0038] ここで、ユーザが「顧客リストに対するランクアップ回帰分析プロセス」を選択したとする。この場合、分析プロセス実行部70は、受け付けたスキ

ーマ付テーブルS T 2に含まれるテーブルT B 2に対して選択された分析プロセスを実行する。

[0039] 図7は、分析プロセスを実行する処理の例を示す説明図である。ここで、テーブルT B 2に対して、上述する分析プロセスA P 1が適用されるとする。この場合、分析プロセス実行部70は、テーブルT B 2に含まれる性別のデータを変換する処理P 1（Mを1に、Fを0に変換する処理）を行い、回帰式を用いた判別処理P 2を実行する。その結果、図7に例示するランクアップ列の値が算出される。

[0040] なお、図7に示す例では、ランクアップ列の値を算出するため、図6に例示するランクアップ列に値が設定されていない場合を例示した。ただし、分析プロセスに学習処理が定義されている場合、図6に例示する表の列には、実績データとして算出される値が設定されていてもよい。

[0041] 一方、探索部60（具体的には、テーブル探索部62）は、分析プロセスの選択をユーザから受け付けた場合に、テーブルを出力する。この場合、分析プロセス実行部70は、出力されたテーブルの一覧から、ユーザの所望するテーブルの選択を受け付ける。そして、分析プロセス実行部70は、受け付けたテーブルに対して選択された分析プロセスを実行する。

[0042] 図8は、テーブルを出力する処理の例を示す説明図である。探索部60が分析プロセスとして「顧客リストに対するランクアップ回帰分析プロセス」の選択をユーザから受け付けると、テーブル探索部62は、図5に例示するスキーマ・分析プロセス管理DB50が記憶する情報から、受け付けた分析プロセスに関連付けられているスキーマ001を抽出する。そして、テーブル探索部62は、図3に例示するテーブル・スキーマ管理DB30が記憶する情報から、抽出したスキーマ001に関連付けられているテーブルを特定し、出力する。ここでは、2016年1月の顧客リストを含むテーブルと、2016年2月の顧客リストを含むテーブルとが出力される。

[0043] ここで、ユーザが2016年2月の顧客リストを選択したとする。この場合、分析プロセス実行部70は、受け付けたテーブルT B 2に対して選択さ

れた分析プロセスを実行する。分析プロセスを実行する処理は、図7に例示する内容と同様である。

[0044] スキーマ付テーブル入力部10と、スキーマ抽出部20と、分析プロセス受付部40と、探索部60（より具体的には、分析プロセス探索部61と、テーブル探索部62）と、分析プロセス実行部70とは、プログラム（データ分析支援プログラム）に従って動作するコンピュータのプロセッサ（例えば、CPU（Central Processing Unit）、GPU（Graphics Processing Unit）、FPGA（field-programmable gate array））によって実現される。

[0045] 上記プログラムは、例えば、記憶部（図示せず）に記憶され、プロセッサは、そのプログラムを読み込み、プログラムに従って、スキーマ付テーブル入力部10、スキーマ抽出部20、分析プロセス受付部40、探索部60（より具体的には、分析プロセス探索部61と、テーブル探索部62）および分析プロセス実行部70として動作してもよい。また、データ分析支援装置の機能がSaaS（Software as a Service）形式で提供されてもよい。

[0046] スキーマ付テーブル入力部10と、スキーマ抽出部20と、分析プロセス受付部40と、探索部60（より具体的には、分析プロセス探索部61と、テーブル探索部62）と、分析プロセス実行部70とは、それぞれが専用のハードウェアで実現されていてもよい。また、各装置の各構成要素の一部又は全部は、汎用または専用の回路（circuitry）、プロセッサ等やこれらの組合せによって実現されもよい。これらは、単一のチップによって構成されてもよいし、バスを介して接続される複数のチップによって構成されてもよい。各装置の各構成要素の一部又は全部は、上述した回路等とプログラムとの組合せによって実現されてもよい。

[0047] また、データ分析支援装置の各構成要素の一部又は全部が複数の情報処理装置や回路等により実現される場合には、複数の情報処理装置や回路等は、集中配置されてもよいし、分散配置されてもよい。例えば、情報処理装置や回路等は、クライアントサーバシステム、クラウドコンピューティングシス

テム等、各々が通信ネットワークを介して接続される形態として実現されてもよい。

[0048] 次に、本実施形態のデータ分析支援装置の動作を説明する。図9は、本実施形態のデータ分析支援装置を用いて分析プロセスを実行する動作例を示すフローチャートである。

[0049] 分析プロセス受付部40は、スキーマで定義されているカラム名を用いた分析プロセスの作成を受け付け（ステップS11）、スキーマ・分析プロセス管理DB50に、分析プロセスとスキーマとを関連付けた情報を登録する（ステップS12）。

[0050] 分析プロセス探索部61は、テーブルの選択をユーザから受け付けると（ステップS13）、テーブル・スキーマ管理DB30が記憶する情報およびスキーマ・分析プロセス管理DB50が記憶する情報に基づいて、受け付けたテーブルに対して適用可能な分析プロセスを特定する（ステップS14）。そして、分析プロセス探索部61は、特定された分析プロセスの一覧を出力する（ステップS15）。

[0051] 分析プロセス実行部70は、ユーザより、出力された分析プロセスの一覧から分析プロセスの選択を受け付ける（ステップ16）。そして、分析プロセス実行部70は、受け付けたテーブルに対して選択された分析プロセスを実行する（ステップS17）。

[0052] 図10は、本実施形態のデータ分析支援装置を用いて分析プロセスを実行する他の動作例を示すフローチャートである。図10に例示するフローチャートは、図9に例示するフローチャートと比較して探索部60および分析プロセス実行部70の処理が異なる。分析プロセスとスキーマとを関連付けた情報を登録するステップS11からステップS12の処理は、図9に例示する処理と同様である。

[0053] テーブル探索部62は、分析プロセスの選択をユーザから受け付けると（ステップS21）、テーブル・スキーマ管理DB30が記憶する情報およびスキーマ・分析プロセス管理DB50が記憶する情報に基づいて、受け付け

た分析プロセスで用いるテーブルを特定する（ステップS 2 2）。そして、テーブル探索部 6 2 は、特定されたテーブルの一覧を出力する（ステップS 2 3）。

[0054] 分析プロセス実行部 7 0 は、ユーザより、出力されたテーブルの一覧からテーブルの選択を受け付ける（ステップS 2 4）。そして、分析プロセス実行部 7 0 は、受け付けたテーブルに対して選択された分析プロセスを実行する（ステップS 2 5）。

[0055] 図 1 1 は、スキーマを管理する動作例を示すフローチャートである。スキーマ付テーブル入力部 1 0 が、スキーマとテーブルとが関連付けられたスキーマ付テーブルを入力すると（ステップS 3 1）、スキーマ抽出部 2 0 は、スキーマ付テーブルから、スキーマを抽出する（ステップS 3 2）。そして、スキーマ抽出部 2 0 は、抽出されたスキーマと、テーブルとを関連付けてテーブル・スキーマ管理DB 3 0 に登録する（ステップS 3 3）。その際、スキーマ抽出部 2 0 は、カラムの名称およびデータ型が一致するスキーマがテーブル・スキーマ管理DB 3 0 に登録されていない場合に、抽出されたスキーマを新たなスキーマとして登録する。

[0056] 以上のように、本実施形態では、分析プロセス受付部 4 0 が分析プロセスの作成を受け付け、受け付けた分析プロセスと、その分析プロセスを適用可能なスキーマとを関連付けた情報を、スキーマ・分析プロセス管理DB 5 0 に登録する。その後、テーブルの選択をユーザから受け付けると、分析プロセス探索部 6 1 は、テーブル・スキーマ管理DB 3 0 が記憶する情報及びスキーマ・分析プロセス管理DB 5 0 が記憶する情報に基づいて、受け付けたテーブルに対して適用可能な分析プロセスを特定し、特定された分析プロセスの一覧を出力する。そして、分析プロセス実行部 7 0 は、出力された分析プロセスの一覧から分析プロセスの選択を受け付け、受け付けたテーブルに対して選択された分析プロセスを実行する。よって、一のテーブルに対して定義された分析処理を異なるテーブルに対しても実行できる。

[0057] また、本実施形態では、分析プロセス受付部 4 0 が分析プロセスの作成を

受け付け、受け付けた分析プロセスと、その分析プロセスを適用可能なスキーマとを関連付けた情報を、スキーマ・分析プロセス管理DB50に登録する。その後、分析プロセスの選択をユーザから受け付けると、テーブル探索部62は、テーブル・スキーマ管理DB30が記憶する情報、および、スキーマ・分析プロセス管理DB50が記憶する情報に基づいて、受け付けた分析プロセスで用いるテーブルを特定し、特定されたテーブルの一覧を出力する。そして、分析プロセス実行部70は、出力されたテーブルの一覧からテーブルの選択を受け付け、受け付けたテーブルに対して選択された分析プロセスを実行する。よって、上述する方法と同様、一のテーブルに対して定義された分析処理を異なるテーブルに対しても実行できる。

[0058] また、本実施形態では、スキーマ付テーブル入力部10がスキーマ付テーブルを入力し、スキーマ抽出部20がスキーマ付テーブルから、スキーマを抽出し、抽出されたスキーマと、テーブルとを関連付けてテーブル・スキーマ管理DB30に登録する。その際、スキーマ抽出部20が、カラムの名称およびデータ型が一致するスキーマがテーブル・スキーマ管理DB30に登録されていない場合に、抽出されたスキーマを新たなスキーマとして登録する。よって、一般的なRDBで利用されるスキーマ付テーブルを、スキーマとテーブルとに分離して管理できる。その結果、スキーマに対して分析プロセスを定義することで、一のテーブルに対して定義された分析処理を異なるテーブルに対しても実行できる。

[0059] 実施形態2.

次に、本発明によるデータ分析支援装置の第2の実施形態を説明する。第1の実施形態では、スキーマ抽出部20が、カラムの名称およびデータ型が一致するスキーマが登録されていないときに、抽出されたスキーマをテーブル・スキーマ管理DB30に登録する場合について説明した。

[0060] 一方、RDBのバージョンの違いや、テーブルの設計変更などにより、同一の内容を示す列であっても、異なるデータ型が定義されているテーブルも存在する。また、同じ数値型や文字列型であっても、RDBのメモリ管理等

の観点から複数種類のデータ型が定義されていることもある。

[0061] しかし、データ分析の観点では、同一の内容を示す列は、同じデータ型として扱えることが好ましく、RDBが想定する種類のデータ型までは必要としない場合も少なくない。そこで、本実施形態では、データ型を抽象化したデータ型である分析データ型を用いて、分析プロセスを管理する方法を説明する。

[0062] 本実施形態において、分析データ型とは、分析処理のために便宜上定義される抽象化されたデータ型であり、実際にRDBで用いられるデータ型とは別に設けられる。具体的には、分析データ型には、同値判定が可能なデータ型を表すカテゴリ変数、連続値のデータ型を表す数値変数、および、順序関係を有し時間軸上の一点を表す情報を抽出可能なデータ型を表す時間変数が含まれる。

[0063] 具体的には、数値変数は、回帰分析等で用いられる実数値などの連続値を表すデータ型であり、例えば、四則演算などの演算を適用可能なデータ型である。ただし、分析データ型に含まれる内容は、上記内容に限定されない。例えば、経度および緯度で表現される地理的な一地点を示すデータ型を、分析データ型に含めてもよい。

[0064] 図12は、本発明によるデータ分析支援装置の第2の実施形態の構成例を示すブロック図である。本実施形態のデータ分析支援装置200は、スキーマ付テーブル入力部10と、分析スキーマ抽出部21と、テーブル・分析スキーマ管理データベース31（以下、テーブル・分析スキーマ管理DB31と記す。）と、分析プロセス受付部40と、分析スキーマ・分析プロセス管理データベース51（以下、分析スキーマ・分析プロセス管理DB51と記す。）と、探索部60と、分析プロセス実行部70とを備えている。

[0065] なお、テーブル・分析スキーマ管理DB31と、分析スキーマ・分析プロセス管理DB51とは、具体的には、磁気ディスク装置等に記憶される。

[0066] スキーマ付テーブル入力部10は、第1の実施形態と同様に、スキーマ付テーブルを入力する。

- [0067] 分析スキーマ抽出部 2 1 は、第 1 の実施形態におけるスキーマ抽出部 2 0 と同様に、スキーマ付テーブルからスキーマを抽出する。さらに、分析スキーマ抽出部 2 1 は、抽出したスキーマに含まれるデータ型を分析データ型に変換する。そして、分析スキーマ抽出部 2 1 は、データ型を変換したスキーマと、テーブルとを関連付けてテーブル・分析スキーマ管理 DB 3 1 に登録する。以下の説明では、分析データ型にデータ型を変換したスキーマのことを、分析スキーマと記すこともある。
- [0068] 具体的には、分析スキーマ抽出部 2 1 は、抽出したスキーマに含まれるデータ型を、列の内容（具体的には、カラム名、データ型など）に応じて予め定めた分析データ型に変換してもよい。また、分析スキーマ抽出部 2 1 は、抽出したスキーマに含まれるデータ型に対する分析データ型への変換指示をユーザから受け付けてもよい。このように、分析スキーマ抽出部 2 1 は、スキーマに含まれるカラムのデータ型を分析データ型へ変換することから、データ型変換部とすることができる。
- [0069] 図 1 3 は、列の内容に応じて分析データ型を設定した例を示す説明図である。図 1 3 に例示するように、分析目的に応じた分析データ型を予め設定しておいてもよい。分析スキーマ抽出部 2 1 は、カラムに対して予め分析データ型への変換ルールが設定されている場合、その設定に基づいてデータ型を分析データ型へ変換してもよい。
- [0070] また、分析スキーマ抽出部 2 1 は、上述する処理を組み合わせてもよい。例えば、データ型やカラム名に応じた分析データ型への変換ルールを予め設定して記憶部（図示せず）に記憶させておく。まず、分析スキーマ抽出部 2 1 は、この変換ルールに従い、抽出したスキーマに含まれるデータ型を分析データ型に一括で変換する。次に、分析スキーマ抽出部 2 1 は、変換後の分析データ型をカラム名とともに出力し、個別に分析データ型の変更を受け付ける。なお、分析スキーマ抽出部 2 1 は、全ての分析データ型への変更を個別に受け付けてもよい。具体的には、分析スキーマ抽出部 2 1 は、スキーマのカラムごとに分析データ型への変換指示を受け付け、抽出したスキーマに

含まれるデータ型を受け付けた分析データ型に個別に変換してもよい。

[0071] 図14は、分析スキーマを抽出する処理の例を示す説明図である。図14に例示する2つのスキーマ付テーブルST3, ST4は、いずれも顧客リストを含むテーブルであるが、スキーマの内容（具体的には、データ型）が異なる。例えば、2016年の顧客リストテーブルST3の顧客IDは、数値で表されていることから、RDB上ではデータ型longで管理されている。一方、例えば、2001年の顧客リストテーブルST4の顧客IDも、数値で表されているが、バージョン等の違いにより、RDB上ではデータ型intで管理されている。

[0072] 一方、顧客IDは、数値計算の対象とされるよりも、同値（非同値）判定の対象とされることが多いと考えられる。そこで、図13に例示するように、分析スキーマ抽出部21は、顧客IDをカテゴリ値として分析できるように、分析データ型への変換を行う。

[0073] まず、分析スキーマ抽出部21は、スキーマ付テーブルST3, ST4から、それぞれスキーマSC2, SC3を抽出する。そして、分析スキーマ抽出部21は、図13に例示する変換ルールに基づいて、各列のデータ型を分析データ型へ変換したスキーマSC4を作成する。

[0074] テーブル・分析スキーマ管理DB31は、分析スキーマとテーブルとを関連付けて記憶する。テーブル・分析スキーマ管理DB31は、例えば、分析スキーマ名とテーブル名とを対応付けて記憶する。テーブル・分析スキーマ管理DB31が分析スキーマ名とテーブル名とを対応付けて記憶する態様は、第1の実施形態におけるテーブル・スキーマ管理DB30と同様である。

[0075] 分析プロセス受付部40は、第1の実施形態と同様、分析スキーマで定義されているカラム名を用いた分析プロセスの作成を受け付ける。そして、分析プロセス受付部40は、作成した分析プロセスを分析スキーマ・分析プロセス管理DB51に登録する。

[0076] 分析スキーマ・分析プロセス管理DB51は、分析プロセスと、その分析プロセスを適用可能な分析スキーマとを関連付けた情報を記憶する。分析ス

キーマ・分析プロセス管理DB51が分析プロセスと分析スキーマとを対応付けて記憶する態様は、第1の実施形態におけるスキーマ・分析プロセス管理DB50と同様である。

[0077] 探索部60は、第1の実施形態と同様、分析プロセス探索部61と、テーブル探索部62とを含む。分析プロセス探索部61は、テーブルの選択をユーザから受け付ける。分析プロセス探索部61は、テーブル・分析スキーマ管理DB31が記憶する情報から、受け付けたテーブルに関連付けられている分析スキーマを抽出する。そして、分析プロセス探索部61は、分析スキーマ・分析プロセス管理DB51が記憶する情報から、抽出した分析スキーマに関連付けられている分析プロセスを特定し、出力する。

[0078] このとき、分析プロセス実行部70は、出力された分析プロセスの一覧から、ユーザの所望する分析プロセスの選択を受け付ける。そして、分析プロセス実行部70は、受け付けたテーブルに対して選択された分析プロセスを実行する。

[0079] また、テーブル探索部62は、分析プロセスの選択をユーザから受け付ける。テーブル探索部62は、分析スキーマ・分析プロセス管理DB51が記憶する情報から、受け付けた分析プロセスに関連付けられている分析スキーマを抽出する。そして、テーブル探索部62は、テーブル・分析スキーマ管理DB31が記憶する情報から、抽出した分析スキーマに関連付けられているテーブルを特定し、出力する。

[0080] このとき、分析プロセス実行部70は、出力されたテーブルの一覧から、ユーザの所望するテーブルの選択を受け付ける。そして、分析プロセス実行部70は、受け付けたテーブルに対して選択された分析プロセスを実行する。

[0081] このように、探索部60（より具体的には、分析プロセス探索部61と、テーブル探索部62）および分析プロセス実行部70の動作は、スキーマが分析スキーマに変更された以外は、第1の実施形態と同様である。

[0082] なお、スキーマ付テーブル入力部10と、分析スキーマ抽出部21と、分

析プロセス受付部40と、探索部60（より具体的には、分析プロセス探索部61と、テーブル探索部62）と、分析プロセス実行部70とは、プログラム（データ分析支援プログラム）に従って動作するコンピュータのプロセッサによって実現される。また、第1の実施形態と同様に、スキーマ付テーブル入力部10、分析スキーマ抽出部21およびテーブル・分析スキーマ管理DB31を含む装置199を、スキーマ管理装置とすることが出来る。なお、第1の実施形態と同様、本実施形態のデータ分析支援装置200が、スキーマ管理装置を含んでいなくてもよい。例えば、データ分析装置が外部に存在し、データ分析支援装置200が、外部に存在するデータ分析装置に接続されて各情報を取得するようにしてもよい。

[0083] 次に、本実施形態のデータ分析支援装置の動作を説明する。図15は、スキーマを管理する動作例を示すフローチャートである。なお、スキーマを抽出するまでの処理は、図11に例示するステップS31からステップS32までの処理と同様である。

[0084] スキーマを抽出後、分析スキーマ抽出部21は、スキーマに含まれるカラムのデータ型を分析データ型へ変換する（ステップS41）。そして、分析スキーマ抽出部21は、分析スキーマとテーブルとを関連付けてテーブル・分析スキーマ管理DB31に登録する（ステップS42）。

[0085] 以上のように、本実施形態では、分析スキーマ抽出部21が、スキーマに含まれるカラムのデータ型を分析データ型へ変換し、分析データ型で定義されるスキーマとテーブルとを関連付けた情報をテーブル・分析スキーマ管理DB31に登録する。また、分析プロセス受付部40は、分析スキーマ・分析プロセス管理DB51に、分析プロセスと、分析データ型で定義されるスキーマとを関連付けた情報を登録する。よって、第1の実施形態の効果に加え、データ型が異なるスキーマが定義されたテーブルに対しても、同じ分析プロセスを用いて同じ処理を実行することが可能になる。

[0086] 例えば、数値情報を含むカラムのデータに対して、繰り返し処理を行う状況を考える。繰り返し処理の一例として、「数値型の全てのカラムの対数を

新しいカラムとして追加する」、「数値型の全てのカラムの一か月の平均値を新しいカラムとして追加する」などが挙げられる。

[0087] 例えば、需給、引出額および預入額は、一般に数値情報で表される。一方、RDBでは、需給がint型、引出額がlong型、預入額がlong型で定義されているとする。この場合、引出額と預入額のデータ型は同一であるが、需給とデータ型が異なる。そのため、一般的に、それぞれのカラムのデータを考慮して個別に処理を記載する必要がある。

[0088] 一方、本実施形態では、数値情報を列に含むテーブルのスキーマのデータ型を分析データ型に変換する。このような変換を行うことで、分析に則したデータ型に応じた繰り返し処理を簡単に記述することが可能になる。したがって、定義されたデータ型が異なるカラムに対しても、同様の分析プロセスを実行することが可能になる。

[0089] また、逆に、ATM (Automated Teller Machine) のID、引出額および預入額のデータ型がいずれもlong型に定義されているとする。一方、ATMのIDは、演算の対象とされる情報でない場合が一般的である。この場合、分析の観点では数値情報の意味が異なるため、やはり一般的には個別に処理を記述する必要がある。

[0090] 一方、本実施形態では、列の意味を考慮してスキーマのデータ型を分析データ型に変換する。このような変換を行うことで、定義されたデータ型が同じカラムに対しても、その意味に応じて分析プロセスを区別することが可能になる。

[0091] 次に、本発明の概要を説明する。図16は、本発明によるデータ分析支援装置の概要を示すブロック図である。本発明によるデータ分析支援装置180（例えば、データ分析支援装置100）は、テーブルに適用されるスキーマで定義されているカラム名を用いた、データ分析のための一連の処理である分析プロセスの作成を受け付ける分析プロセス受付部182（例えば、分析プロセス受付部40）と、受け付けた分析プロセスと、その分析プロセスを適用可能なスキーマとを関連付けた情報を記憶するスキーマ・分析プロセ

ス記憶部 183（例えば、スキーマ・分析プロセス管理DB50）と、テーブルの選択をユーザから受け付けると、テーブルとそのテーブルに適用されるスキーマとを関連付けた情報を記憶するテーブル・スキーマ記憶部（例えば、テーブル・スキーマ管理DB30）が記憶する情報、および、スキーマ・分析プロセス記憶部 183 が記憶する情報に基づいて、受け付けたテーブルに対して適用可能な分析プロセスを特定し、特定された分析プロセスの一覧を出力する分析プロセス探索部 184（例えば、分析プロセス探索部 61）と、出力された一覧から分析プロセスの選択を受け付け、受け付けたテーブルに対して選択された分析プロセスを実行する分析プロセス実行部 185（例えば、分析プロセス実行部 70）とを備えている。

[0092] そのような構成により、一のテーブルに対して定義された分析処理を異なるテーブルに対しても実行できる。

[0093] また、データ分析支援装置 180（例えば、データ分析支援装置 200）は、スキーマに含まれるカラムのデータ型を、分析処理に用いられるデータ型として定義された分析データ型へ変換するデータ型変換部を備えていてもよい。ここで、分析データ型は、少なくとも同値判定が可能なデータ型を表すカテゴリ変数、および、数値変数を含む。そして、データ型変換部は、テーブル・スキーマ記憶部（例えば、テーブル・分析スキーマ管理DB31）に、分析データ型で定義されるスキーマとテーブルとを関連付けた情報を登録し、分析プロセス受付部 182 は、スキーマ・分析プロセス記憶部 183（例えば、分析スキーマ・分析プロセス管理DB51）に、分析プロセスと、分析データ型で定義されるスキーマとを関連付けた情報を登録してもよい。

[0094] そのような構成によれば、データ型が異なるスキーマが定義されたテーブルに対しても、同じ分析プロセスを用いて同じ処理を実行することが可能になる。

[0095] このとき、データ型変換部は、データ型またはカラム名に応じた分析データ型への変換ルールに応じて、抽出したスキーマに含まれるデータ型を分析

データ型に一括で変換してもよい。

[0096] また、データ型変換部は、スキーマのカラムごとに分析データ型への変換指示を受け付け、抽出したスキーマに含まれるデータ型を受け付けた分析データ型に個別に変換してもよい。

[0097] また、分析データ型は、カテゴリ変数、数値変数、および、順序関係を有する時間軸上の一点を示すデータ型を表す時間変数を含んでいてもよい。

[0098] 以上、実施形態及び実施例を参照して本願発明を説明したが、本願発明は上記実施形態および実施例に限定されるものではない。本願発明の構成や詳細には、本願発明の範囲内で当業者が理解し得る様々な変更をすることができる。

[0099] この出願は、2017年12月22日に提出された米国仮出願第62/609,768号を基礎とする優先権を主張し、その開示の全てをここに取り込む。

符号の説明

- [0100] 10 スキーマ付テーブル入力部
20 スキーマ抽出部
21 分析スキーマ抽出部
30 テーブル・スキーマ管理DB
31 テーブル・分析スキーマ管理DB
40 分析プロセス受付部
50 スキーマ・分析プロセス管理DB
51 分析スキーマ・分析プロセス管理DB
60 探索部
61 分析プロセス探索部
62 テーブル探索部
70 分析プロセス実行部
99 スキーマ管理装置
100, 200 データ分析支援装置

請求の範囲

- [請求項1] テーブルに適用されるスキーマで定義されているカラム名を用いた、データ分析のための一連の処理である分析プロセスの作成を受け付ける分析プロセス受付部と、
- 受け付けた分析プロセスと、当該分析プロセスを適用可能なスキーマとを関連付けた情報を記憶するスキーマ・分析プロセス記憶部と、
- テーブルの選択をユーザから受け付けると、テーブルと当該テーブルに適用されるスキーマとを関連付けた情報を記憶するテーブル・スキーマ記憶部が記憶する当該情報、および、前記スキーマ・分析プロセス記憶部が記憶する情報に基づいて、受け付けたテーブルに対して適用可能な前記分析プロセスを特定し、特定された分析プロセスの一覧を出力する分析プロセス探索部と、
- 出力された一覧から前記分析プロセスの選択を受け付け、受け付けたテーブルに対して選択された分析プロセスを実行する分析プロセス実行部とを備えた
- ことを特徴とするデータ分析支援装置。
- [請求項2] スキーマに含まれるカラムのデータ型を、分析処理に用いられるデータ型として定義された分析データ型へ変換するデータ型変換部を備え、
- 前記分析データ型は、少なくとも同値判定が可能なデータ型を表すカテゴリ変数、および、数値変数を含み、
- 前記データ型変換部は、テーブル・スキーマ記憶部に、分析データ型で定義されるスキーマとテーブルとを関連付けた情報を登録し、
- 分析プロセス受付部は、スキーマ・分析プロセス記憶部に、分析プロセスと、分析データ型で定義されるスキーマとを関連付けた情報を登録する
- 請求項1記載のデータ分析支援装置。
- [請求項3] データ型変換部は、データ型またはカラム名に応じた分析データ型

への変換ルールに応じて、抽出したスキーマに含まれるデータ型を分析データ型に一括で変換する

請求項2記載のデータ分析支援装置。

[請求項4] データ型変換部は、スキーマのカラムごとに分析データ型への変換指示を受け付け、抽出したスキーマに含まれるデータ型を受け付けた分析データ型に個別に変換する

請求項2または請求項3記載のデータ分析支援装置。

[請求項5] 分析データ型は、カテゴリ変数、数値変数、および、順序関係を有する時間軸上の一点を示すデータ型を表す時間変数を含む

請求項2から請求項4のうちのいずれか1項に記載のデータ分析支援装置。

[請求項6] テーブルに適用されるスキーマで定義されているカラム名を用いた、データ分析のための一連の処理である分析プロセスの作成を受け付け、

受け付けた分析プロセスと、当該分析プロセスを適用可能なスキーマとを関連付けた情報を、スキーマ・分析プロセス記憶部に登録し、

テーブルの選択をユーザから受け付けると、テーブルと当該テーブルに適用されるスキーマとを関連付けた情報を記憶するテーブル・スキーマ記憶部が記憶する当該情報、および、前記スキーマ・分析プロセス記憶部が記憶する情報に基づいて、受け付けたテーブルに対して適用可能な前記分析プロセスを特定し、

特定された分析プロセスの一覧を出力し、

出力された一覧から前記分析プロセスの選択を受け付け、

受け付けたテーブルに対して選択された分析プロセスを実行する

ことを特徴とするデータ分析支援方法。

[請求項7] スキーマに含まれるカラムのデータ型を、分析処理に用いられるデータ型として定義された分析データ型へ変換し、

前記分析データ型は、少なくとも同値判定が可能なデータ型を表す

カテゴリ変数、および、数値変数を含み、

テーブル・スキーマ記憶部に、分析データ型で定義されるスキーマとテーブルとを関連付けた情報を登録し、

スキーマ・分析プロセス記憶部に、分析プロセスと、分析データ型で定義されるスキーマとを関連付けた情報を登録する

請求項6記載のデータ分析支援方法。

[請求項8]

コンピュータに、

テーブルに適用されるスキーマで定義されているカラム名を用いた、データ分析のための一連の処理である分析プロセスの作成を受け付け、受け付けた分析プロセスと当該分析プロセスを適用可能なスキーマとを関連付けた情報をスキーマ・分析プロセス記憶部に登録する分析プロセス受付処理、

テーブルの選択をユーザから受け付けると、テーブルと当該テーブルに適用されるスキーマとを関連付けた情報を記憶するテーブル・スキーマ記憶部が記憶する当該情報、および、前記スキーマ・分析プロセス記憶部が記憶する情報に基づいて、受け付けたテーブルに対して適用可能な前記分析プロセスを特定し、特定された分析プロセスの一覧を出力する分析プロセス探索処理、および、

出力された一覧から前記分析プロセスの選択を受け付け、受け付けたテーブルに対して選択された分析プロセスを実行する分析プロセス実行処理

を実行させるためのデータ分析支援プログラム。

[請求項9]

コンピュータに、

スキーマに含まれるカラムのデータ型を、分析処理に用いられるデータ型として定義された分析データ型へ変換するデータ型変換処理を実行させ、

前記分析データ型は、少なくとも同値判定が可能なデータ型を表すカテゴリ変数、および、数値変数を含み、

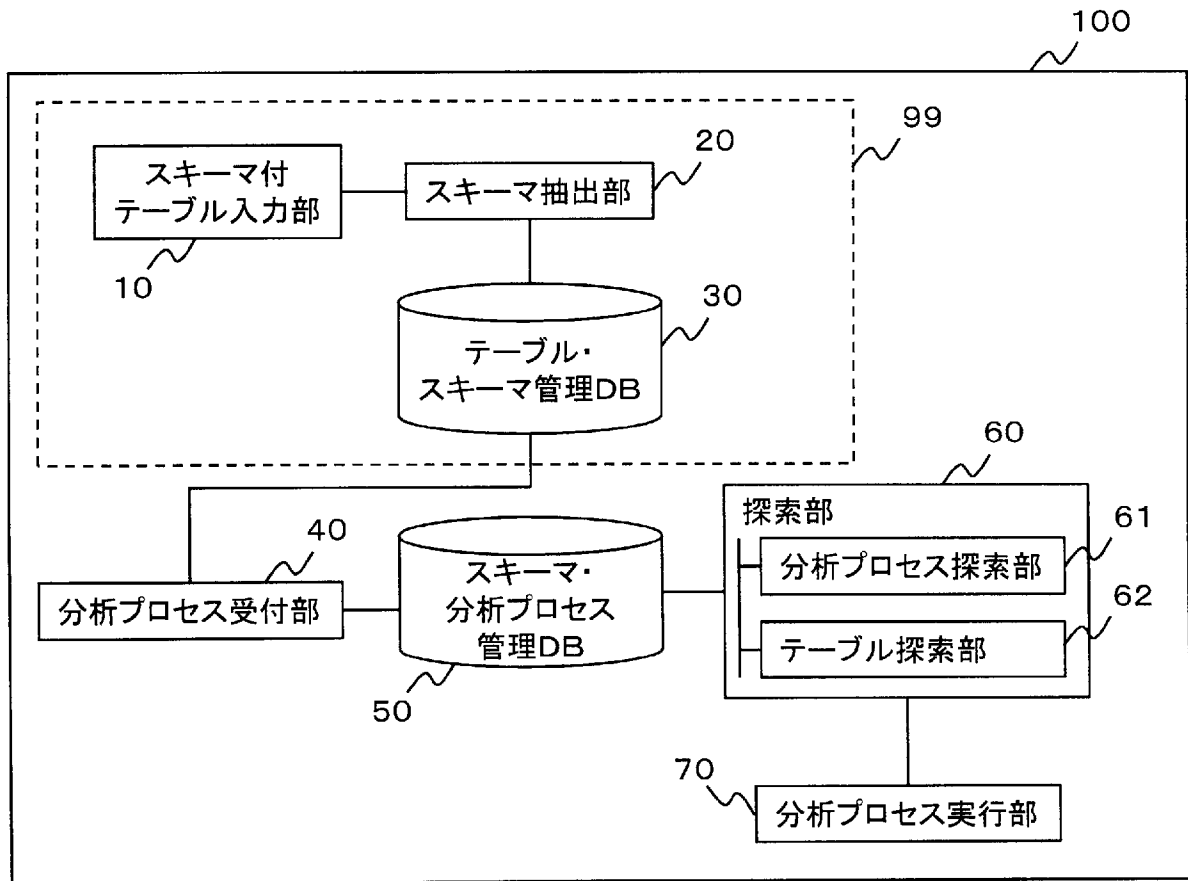
前記データ型変換処理で、テーブル・スキーマ記憶部に、分析データ型で定義されるスキーマとテーブルとを関連付けた情報を登録させ

、

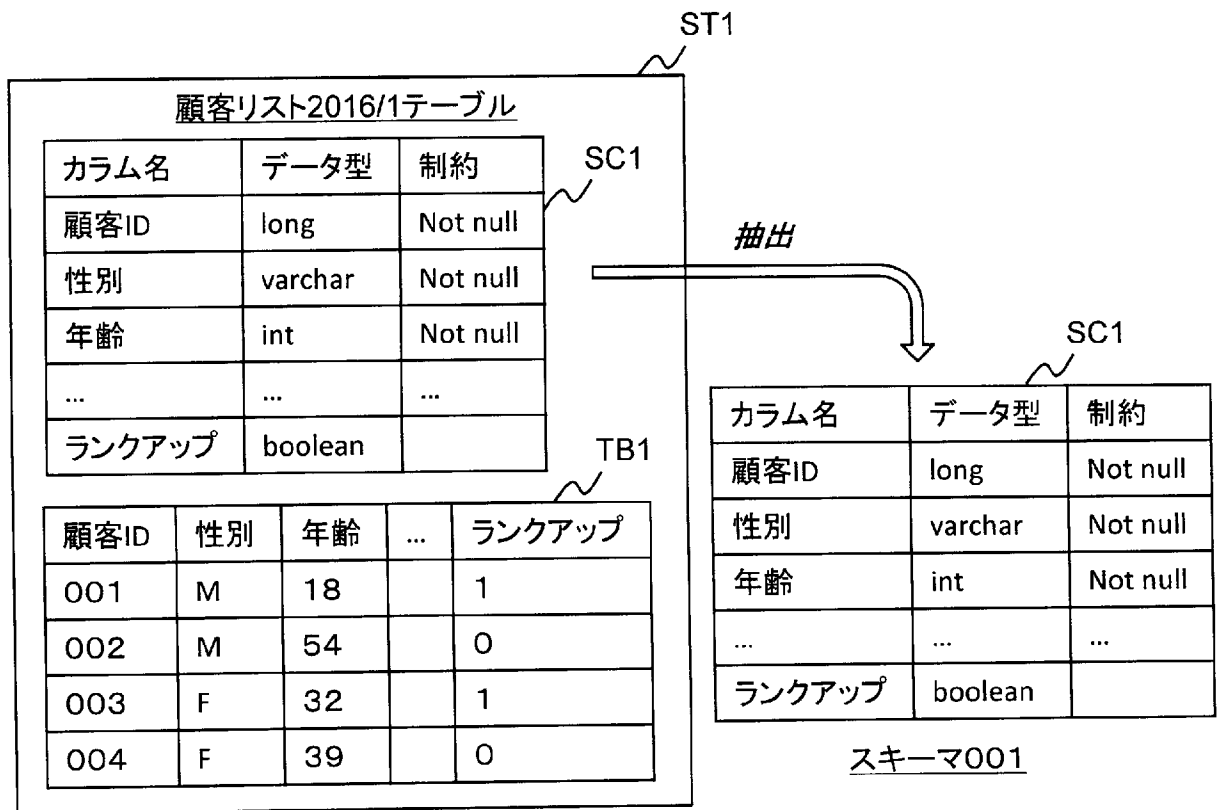
分析プロセス受付処理で、スキーマ・分析プロセス記憶部に、分析プロセスと、分析データ型で定義されるスキーマとを関連付けた情報を登録させる

請求項8記載のデータ分析支援プログラム。

[図1]



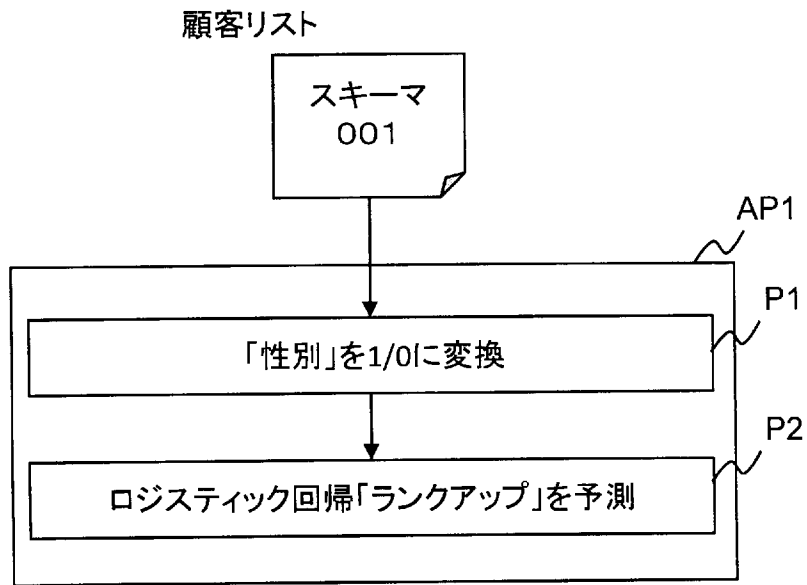
[図2]



[図3]

テーブル	スキーマ
顧客リスト 2016/1テーブル	001
顧客リスト 2016/2テーブル	001

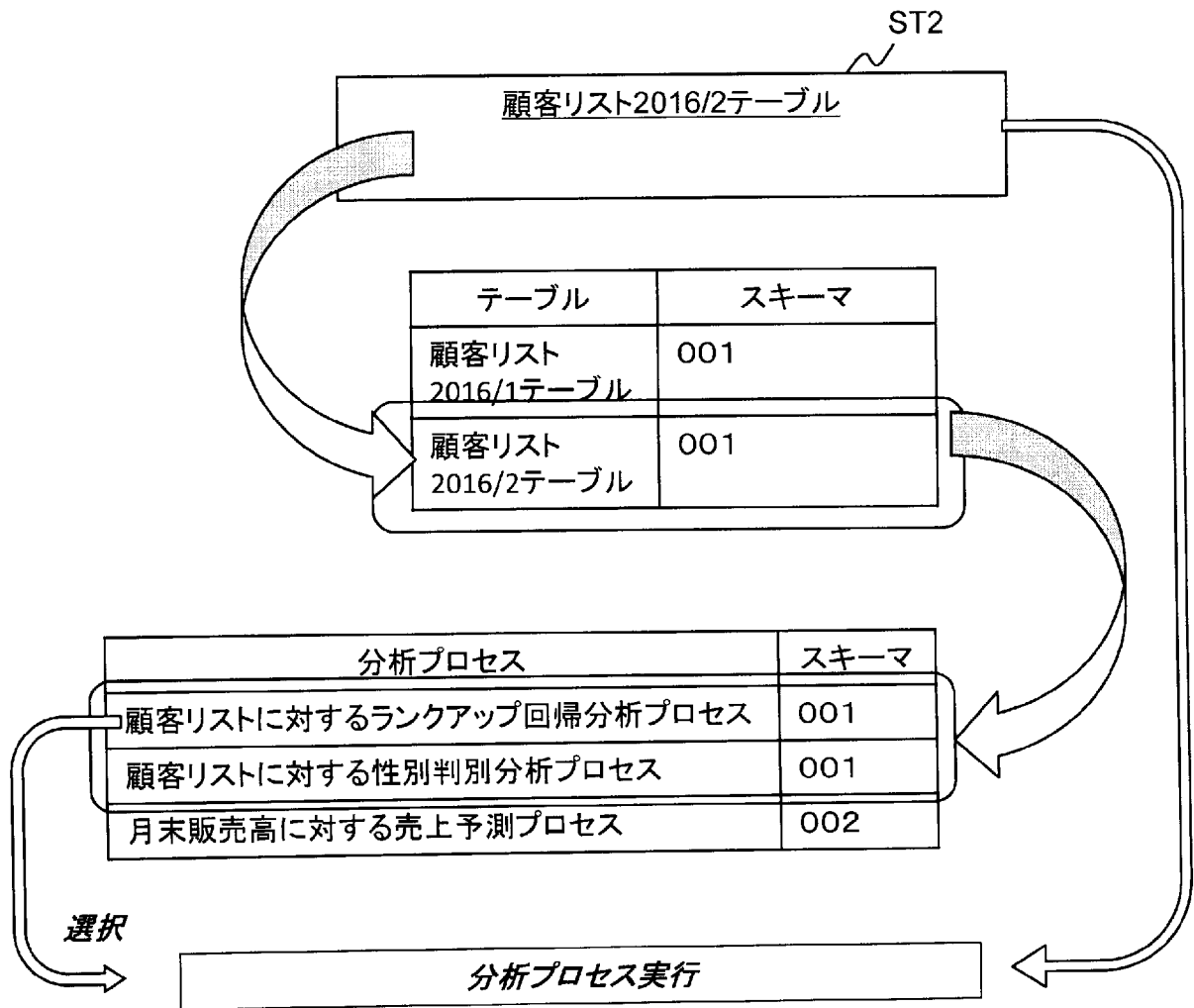
[図4]



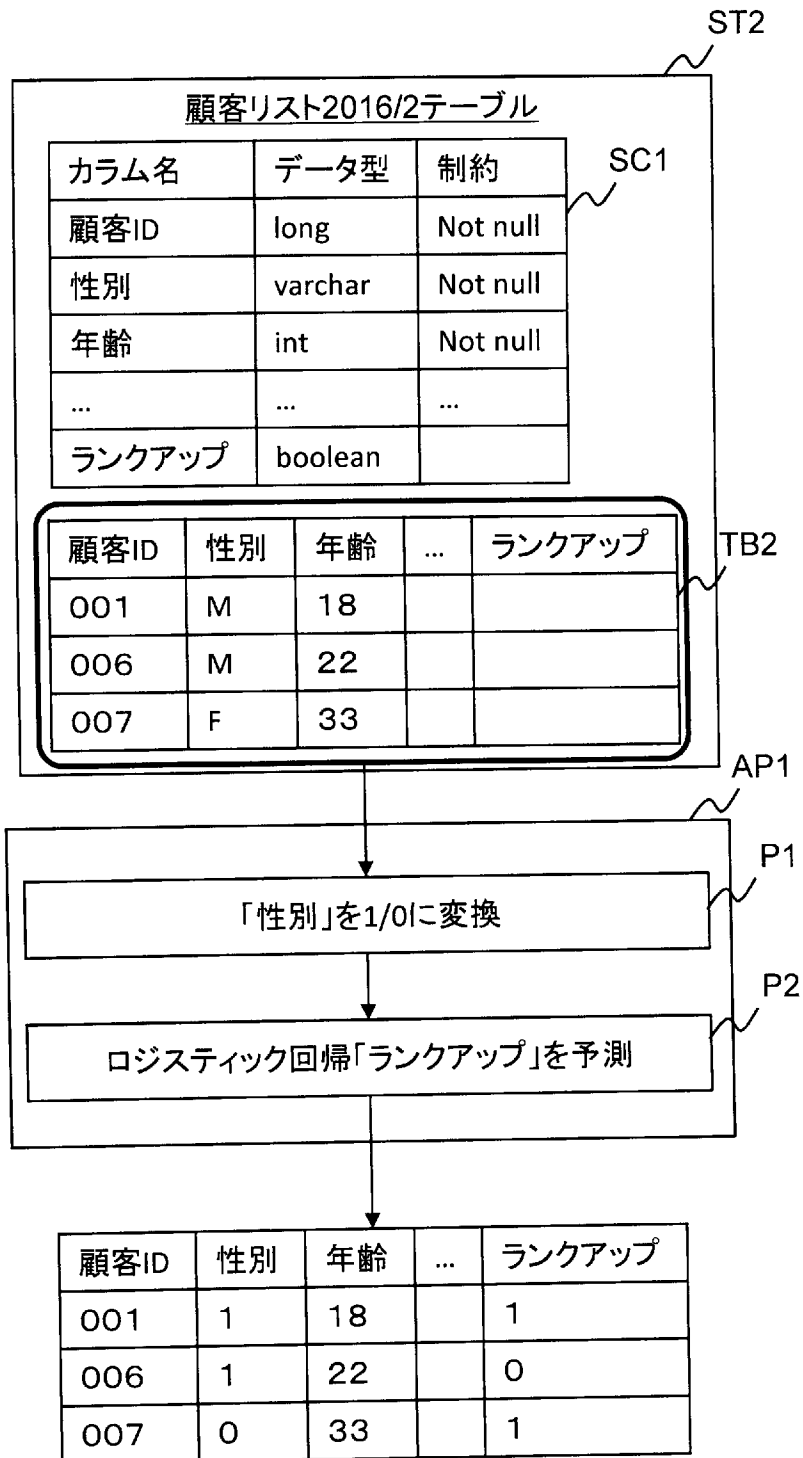
[図5]

分析プロセス	スキーマ
顧客リストに対するランクアップ回帰分析プロセス	001
顧客リストに対する性別判別分析プロセス	001
月末販売高に対する売上予測プロセス	002

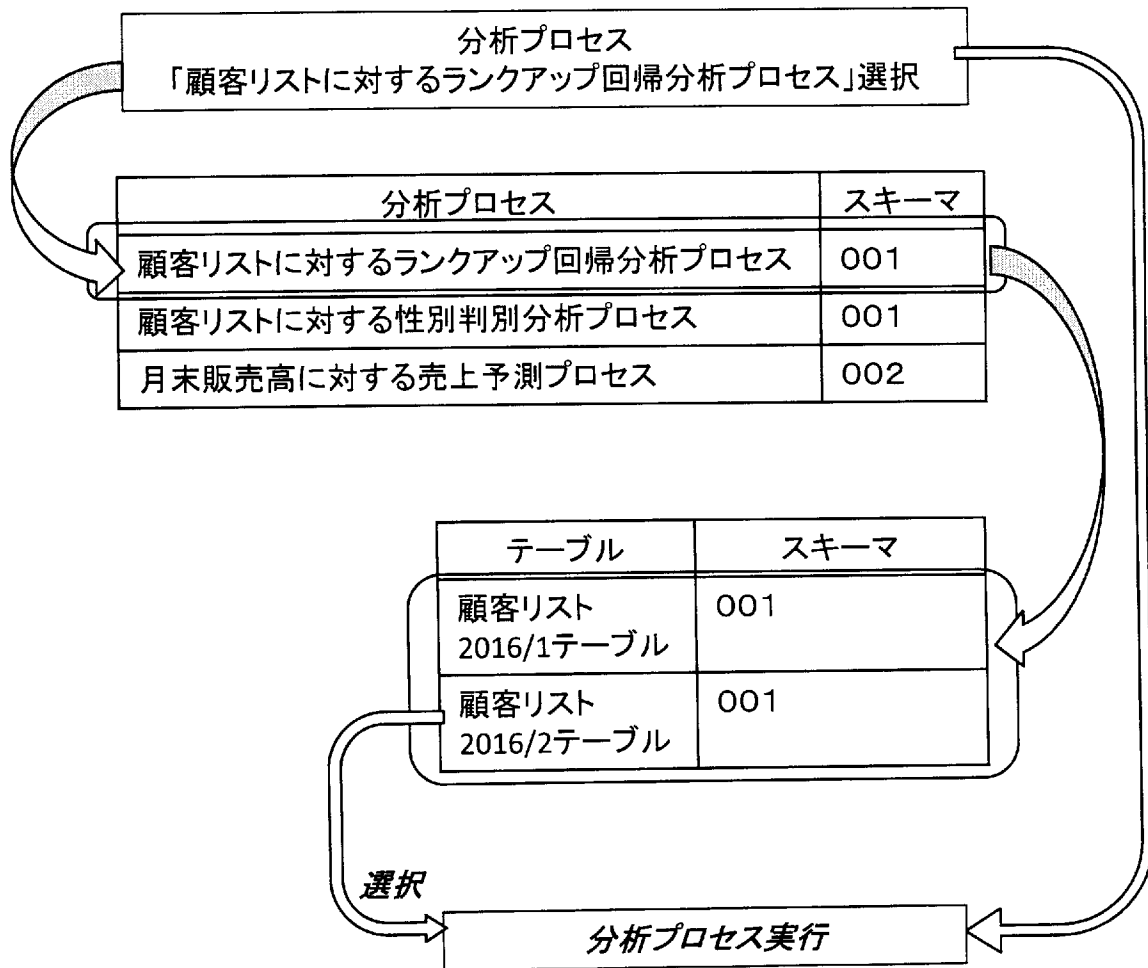
[図6]



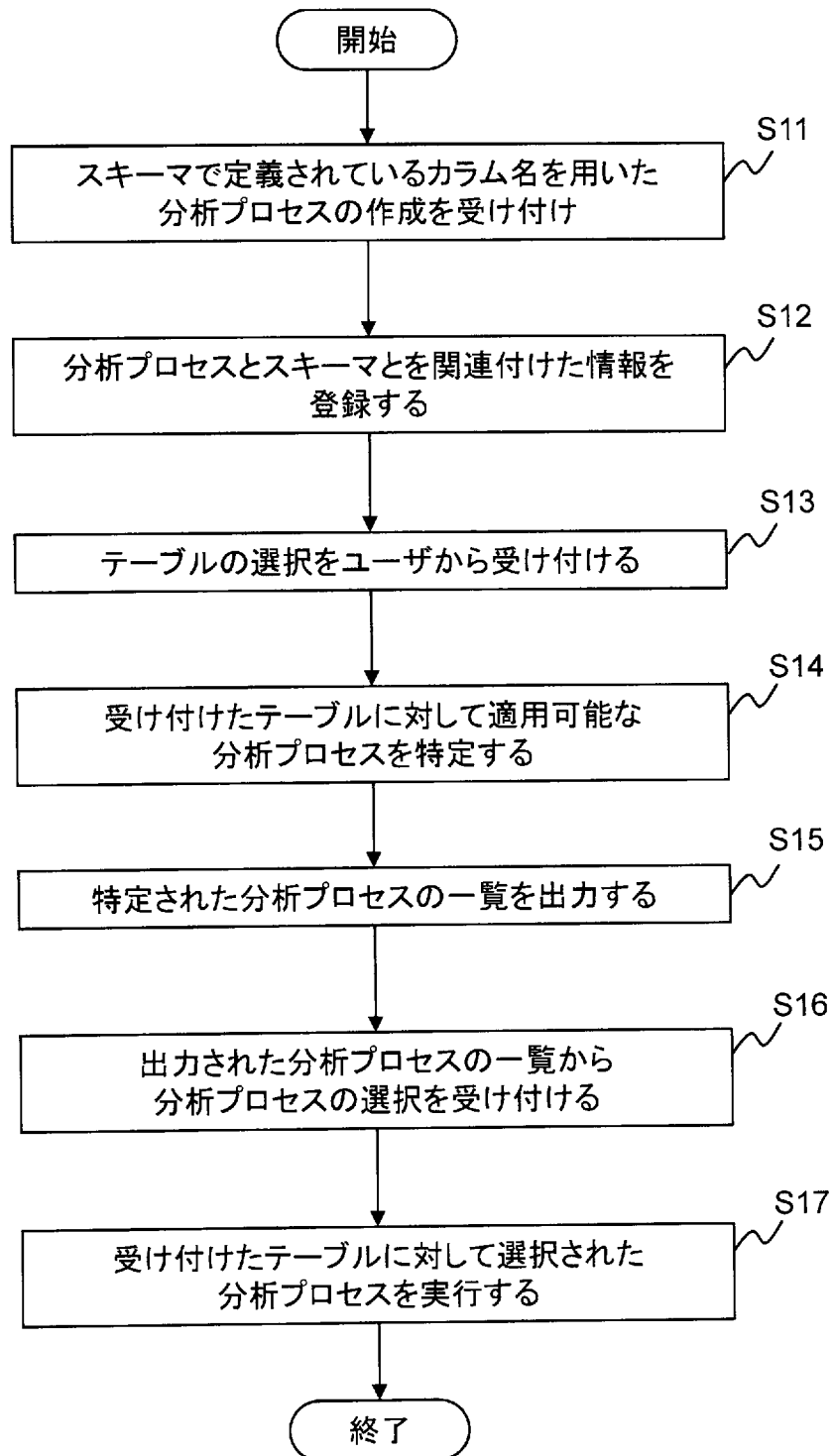
[図7]



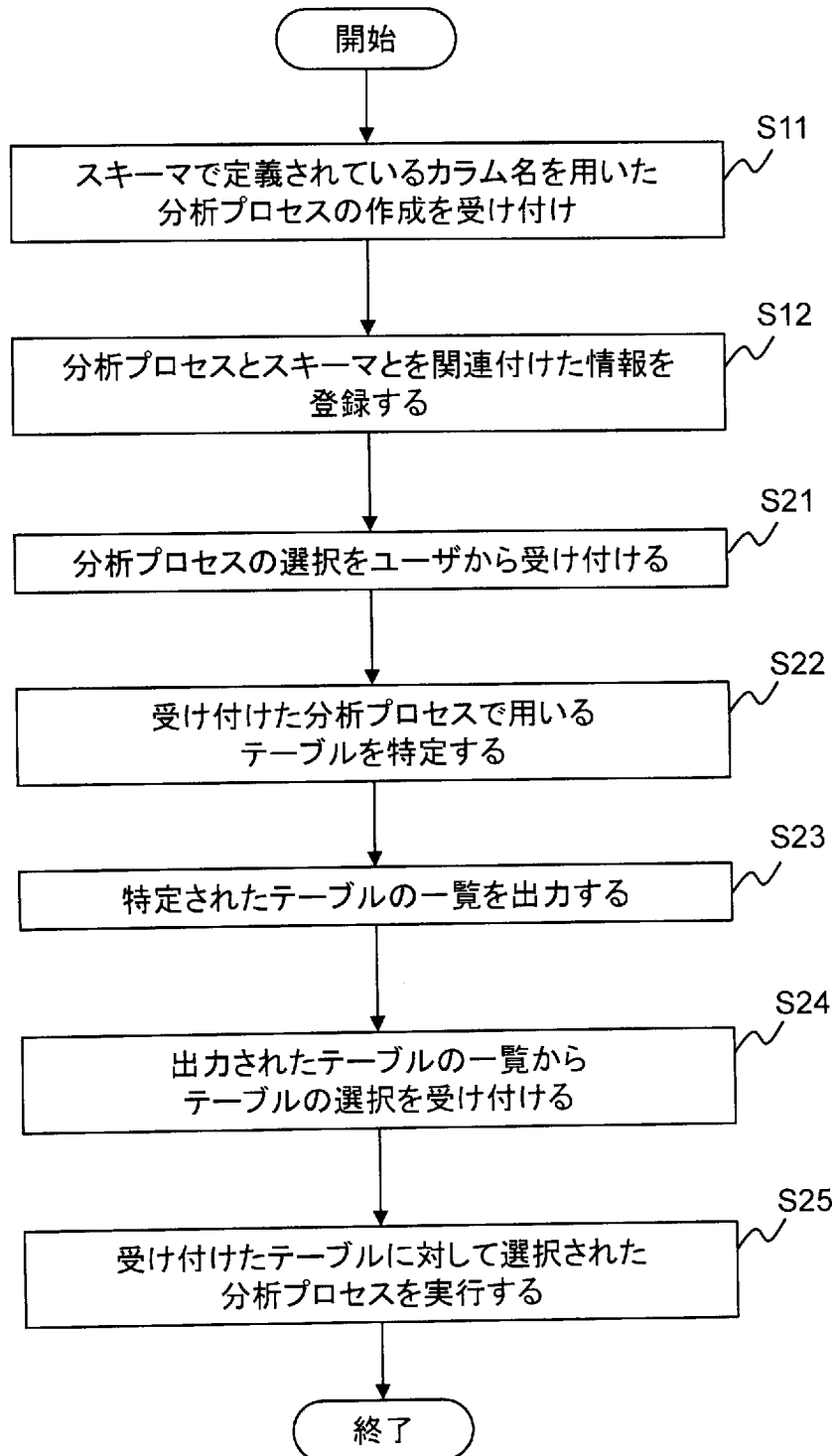
[図8]



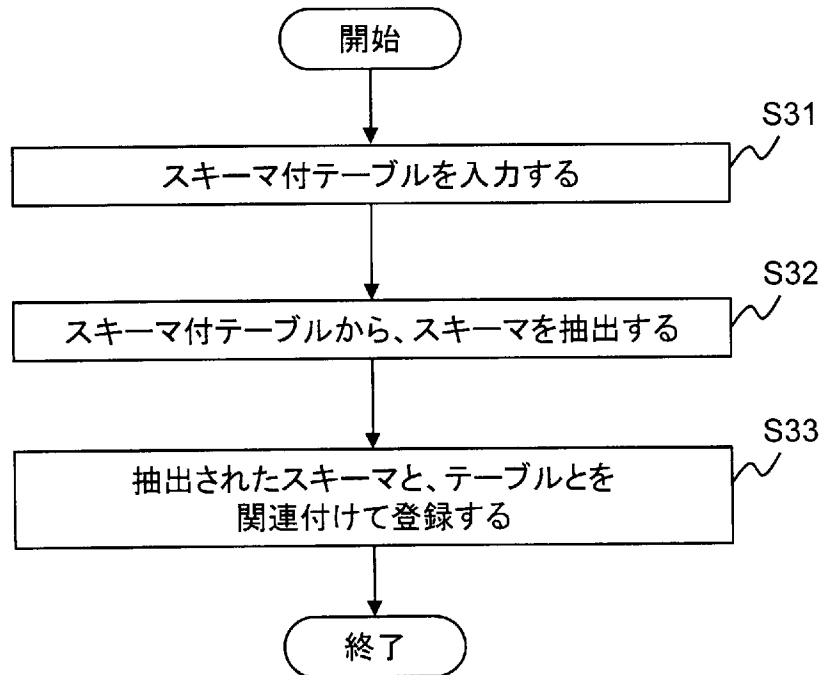
[図9]



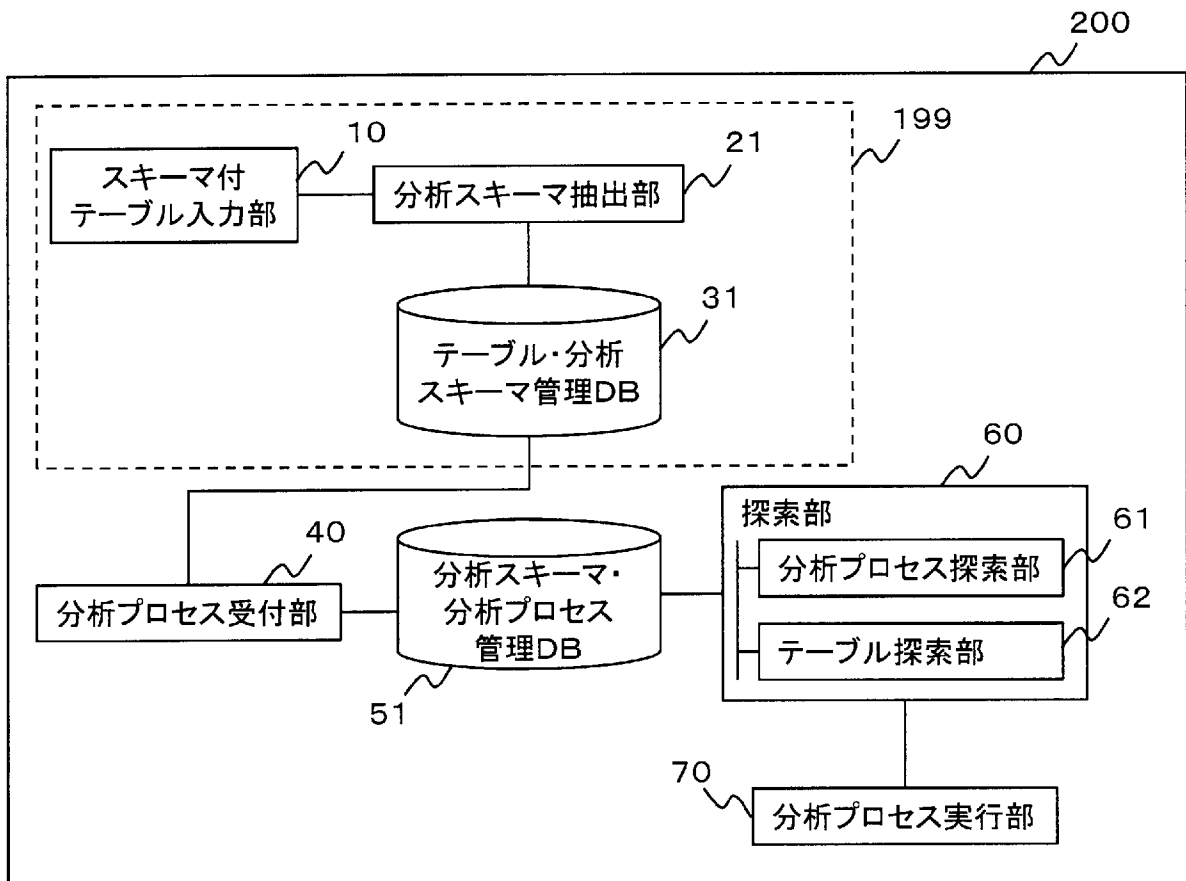
[図10]



[図11]



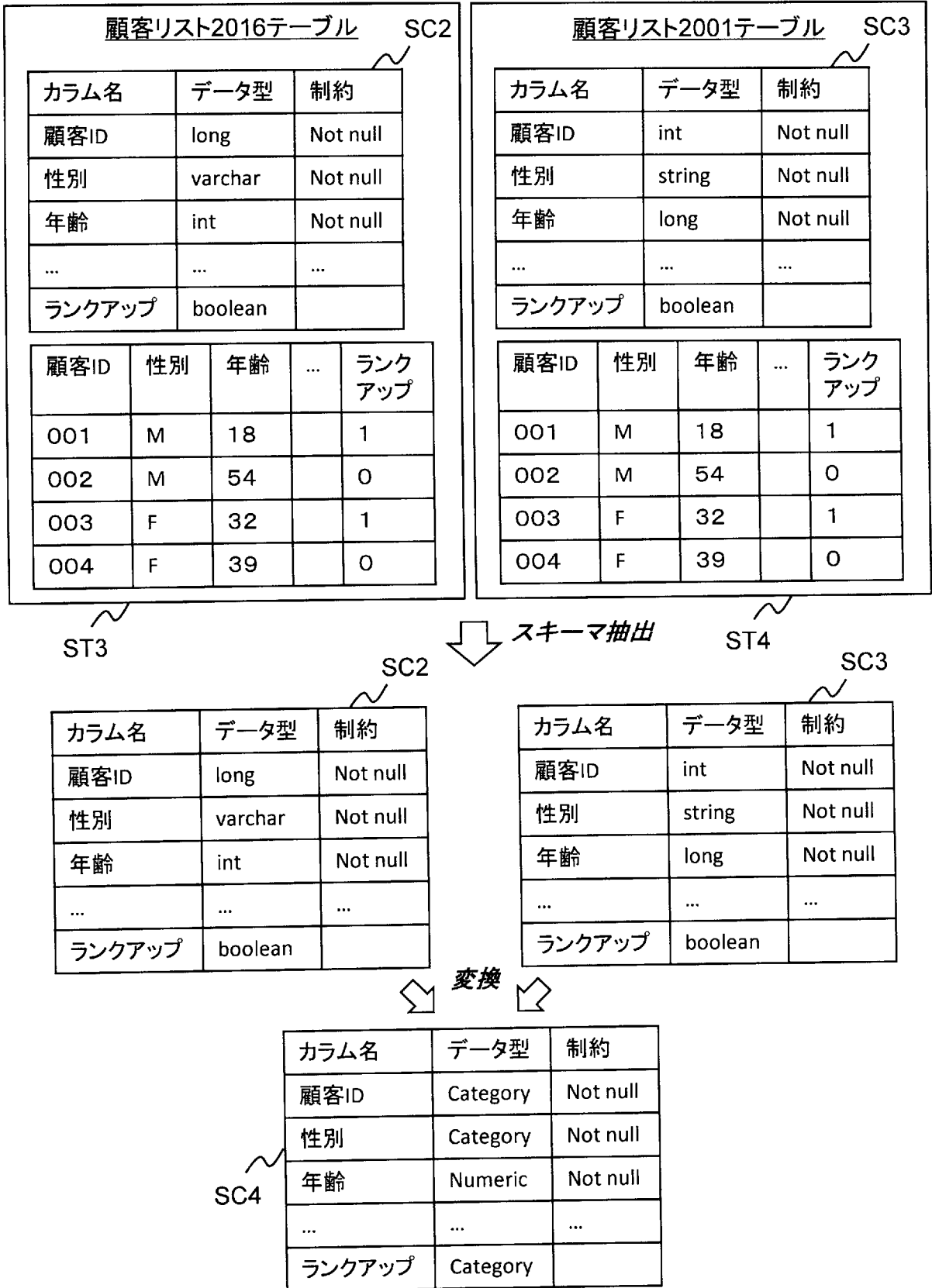
[図12]



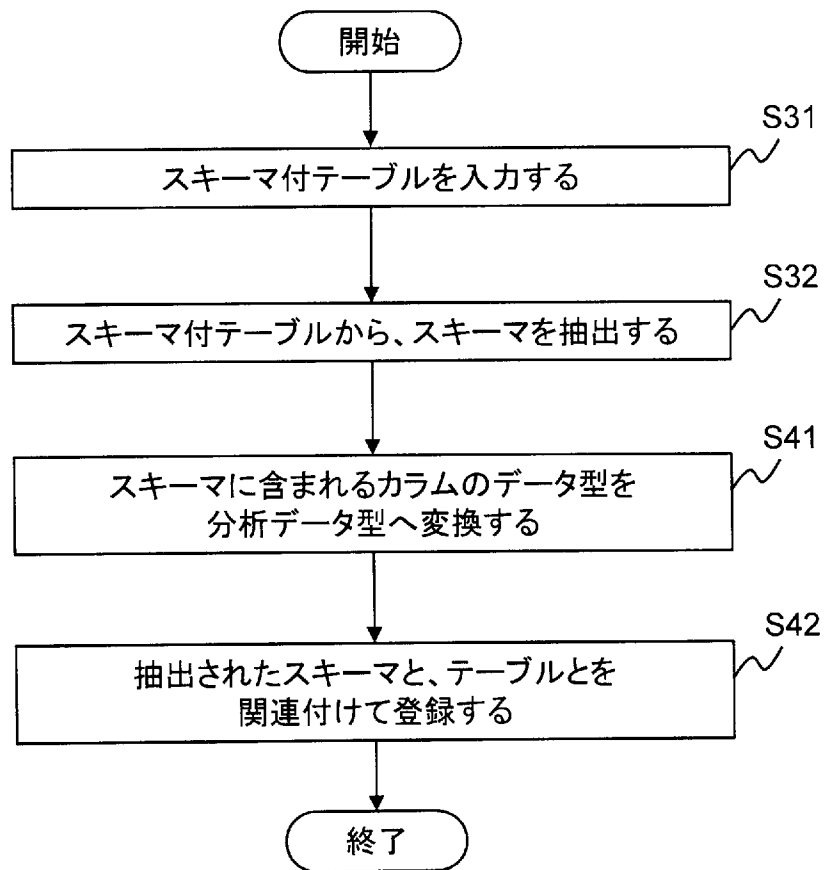
[図13]

カラム名	分析データ型
需給	Numeric(数値)
引出額	Numeric(数値)
預入額	Numeric(数値)
ATM ID	Category(カテゴリ値)
日付	Date Time(日付)
日時	Date Time(日付)

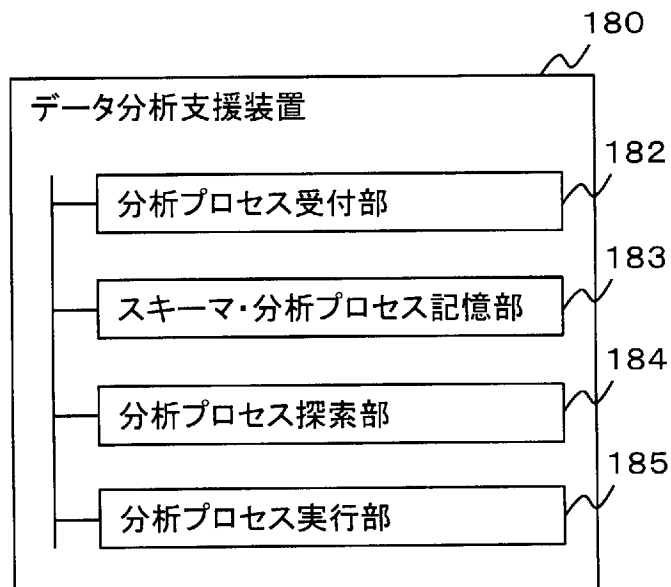
[図14]



[図15]



[図16]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2018/028082

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl. G06F17/30 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl. G06F17/30

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan	1922-1996
Published unexamined utility model applications of Japan	1971-2018
Registered utility model specifications of Japan	1996-2018
Published registered utility model applications of Japan	1994-2018

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2002/0147599 A1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 10 October 2002, entire text, all drawings (Family: none)	1-9
A	US 2005/0102303 A1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 12 May 2005, entire text, all drawings (Family: none)	1-9
A	JP 2011-257812 A (FUJITSU LTD.) 22 December 2011, entire text, all drawings & US 2011/0302187 A1	1-9

Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents:	
“A” document defining the general state of the art which is not considered to be of particular relevance	“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
“E” earlier application or patent but published on or after the international filing date	“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
“O” document referring to an oral disclosure, use, exhibition or other means	“&” document member of the same patent family
“P” document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 04.10.2018	Date of mailing of the international search report 16.10.2018
---	--

Name and mailing address of the ISA/ Japan Patent Office 3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan	Authorized officer Telephone No.
--	---

A. 発明の属する分野の分類（国際特許分類（IPC））

Int.Cl. G06F17/30(2006.01)i

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（IPC））

Int.Cl. G06F17/30

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922-1996年
日本国公開実用新案公報	1971-2018年
日本国実用新案登録公報	1996-2018年
日本国登録実用新案公報	1994-2018年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	US 2002/0147599 A1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 2002.10.10, 全文, 全図 (ファミリーなし)	1-9
A	US 2005/0102303 A1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 2005.05.12, 全文, 全図 (ファミリーなし)	1-9
A	JP 2011-257812 A (富士通株式会社) 2011.12.22, 全文, 全図 & US 2011/0302187 A1	1-9

☐ C欄の続きにも文献が列挙されている。

☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

- 「A」特に関連のある文献ではなく、一般的技術水準を示すもの
- 「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
- 「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
- 「O」口頭による開示、使用、展示等に言及する文献
- 「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

- 「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
- 「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
- 「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
- 「&」同一パテントファミリー文献

国際調査を完了した日

04.10.2018

国際調査報告の発送日

16.10.2018

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)
郵便番号 100-8915
東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

齊藤 貴孝

電話番号 03-3581-1101 内線 3599

5M

4774