14. リレーショナルデータベースの特徴，データウエアハウス

URL: https://www.kkaneko.jp/de/ds/index.html

金子邦彦

CC BY NC SA

謝辞：この資料では「いらすとや」のイラストを使用しています

アウトライン

番号	項目	説明時間の目安
14-1	リレーショナルデータベースの特徴（総まとめ）	20分
14-2	情報検索とデータベース， Access, MySQLを用いた実演・実習	8分
14-3	索引	5分
14-4	オープンデータ	2分
14-5	日時情報	3分
14-6	データウエアハウス	6分

14-2. 情報検索とデータベース，Access, MySQL を用いた実演・実習

Excel の検索機能の例

- name 属性の値に「**AD**」を**含むもの**

	A	B	C	D	E	F	G	H
1	id	name	job	manager_id	hire_date	salary	commissic	dep_id
2	68319	KAYLING	PRESIDENT		1991/11/18	6000		1001
3	66928	BLAZE	MANAGER	68319	1991/5/1	2750		3001
4	67832	CLARE	MANAGER	68319	1991/6/9	2550		1001
5	65646	JONAS	MANAGER	68319	1991/4/2	2957		2001
6	67858	SCARLET	ANALYST	65646	1997/4/19	3100		2001
7	69062	FRANK	ANALYST	65646	1991/12/3	3100		2001
8	63679	SANDRINE	CLERK	69062	1990/12/18	900		2001
9	64989	ADELYN	SALESMAN	66928	1991/2/20	1700	400	3001
10	65271	WADE	SALESMAN	66928	1991/2/22	1350	600	3001
11	66564	MADDEN	SALESMAN	66928	1991/9/28	1350	1500	3001
12	68454	TUCKER	SALESMAN	66928	1991/9/8	1600	0	3001
13	68736	ADNRES	CLERK	67858	1997/5/23	1200		2001
14	69000	JULIUS	CLERK	66928	1991/12/3	1050		3001
15	69324	MARKER	CLERK	67832	1992/1/23	1400		1001

元データ

検索結果

SQL による検索機能の例

- name 属性の値に「**AD**」を**含むもの**

employee

ID1	id	name	job	manager_id	hire_date	salary	commission
1	68319	KAYLING	PRESIDENT		1991/11/18	6000	
2	66928	BLAZE	MANAGER	68319	1991/05/01	2750	
3	67832	CLARE	MANAGER	68319	1991/06/09	2550	
4	65646	JONAS	MANAGER	68319	1991/04/02	2957	
5	67858	SCARLET	ANALYST	65646	1997/04/19	3100	
6	69062	FRANK	ANALYST	65646	1991/12/03	3100	
7	63679	SANDRINE	CLERK	69062	1990/12/18	900	
8	64989	ADELYN	SALESMAN	66928	1991/02/20	1700	400
9	65271	WADE	SALESMAN	66928	1991/02/22	1350	600
10	66564	MADDEN	SALESMAN	66928	1991/09/28	1350	1500
11	68454	TUCKER	SALESMAN	66928	1991/09/08	1600	0
12	68736	ADNRES	CLERK	67858	1997/05/23	1200	
13	69000	JULIUS	CLERK	66928	1991/12/03	1050	
14	69324	MARKER	CLERK	67832	1992/01/23	1400	

元データ
（employee テーブル）

```
SELECT *
FROM employee
WHERE name like '*AD*';
```

SQL

ID1	id	name	job	manager_id	hire_date	salary	commission	dep_id
8	64989	ADELYN	SALESMAN	66928	1991/02/20	1700	400	3001
9	65271	WADE	SALESMAN	66928	1991/02/22	1350	600	3001
10	66564	MADDEN	SALESMAN	66928	1991/09/28	1350	1500	3001
12	68736	ADNRES	CLERK	67858	1997/05/23	1200		2001

検索結果

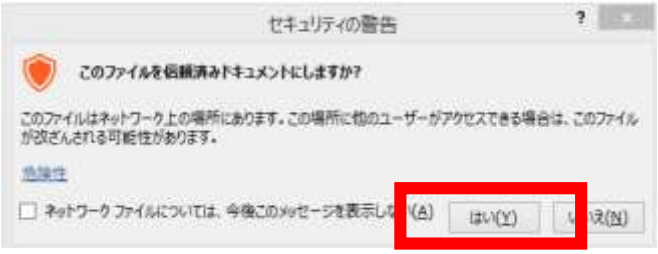
実演・実習用のデータベースファイル

- **実演・実習用の Access データベースファイル**

※ セレッソの利用者は，セレッソからダウンロード可能

- 「**コンテンツの有効化**」のメッセージが出たときは、確認のうえ、次にすすむ

- つぎのような表示が出たときは、確認のうえ、「**はい**」



Access を用いた実演・実習

① データベースファイルは、**大学のセレッソ**よりダウンロードできる

② 次の SQL を実行

```
SELECT *
FROM employee
WHERE name like '*AD*';
```

③ 結果を確認

ID1	id	name	job	manager_id	hire_date	salary	commission	dep_id
8	64989	ADELYN	SALESMAN	66928	1991/02/20	1700	400	3001
9	65271	WADE	SALESMAN	66928	1991/02/22	1350	600	3001
10	66564	MADDEN	SALESMAN	66928	1991/09/28	1350	1500	3001
12	68736	ADNRES	CLERK	67858	1997/05/23	1200		2001

Paiza.IO を用いた実演・実習

① URL は https://paiza.io/ja

② 「**コード作成を試してみる**」をクリック

③ 「**MySQL**」を選ぶ

次ページへ続く

- 次の SQL を実行、結果を確認

```sql
CREATE TABLE employee (
    id integer,
    name text,
    job text,
    manager_id integer,
    hire_date datetime,
    salary integer,
    commission integer,
    dep_id integer
);
insert into employee values(68319,'KAYLING','PRESIDENT',NULL,'1991/11/18',6000,NULL,1001);
insert into employee values(66928,'BLAZE','MANAGER',68319,'1991/5/1',2750,NULL,3001);
insert into employee values(67832,'CLARE','MANAGER',68319,'1991/6/9',2550,NULL,1001);
insert into employee values(65646,'JONAS','MANAGER',68319,'1991/4/2',2957,NULL,2001);
insert into employee values(67858,'SCARLET','ANALYST',65646,'1997/4/19',3100,NULL,2001);
insert into employee values(69062,'FRANK','ANALYST',65646,'1991/12/3',3100,NULL,2001);
insert into employee values(63679,'SANDRINE','CLERK',69062,'1990/12/18',900,NULL,2001);
insert into employee values(64989,'ADELYN','SALESMAN',66928,'1991/2/20',1700,400,3001);
insert into employee values(65271,'WADE','SALESMAN',66928,'1991/2/22',1350,600,3001);
insert into employee values(66564,'MADDEN','SALESMAN',66928,'1991/9/28',1350,1500,3001);
insert into employee values(68454,'TUCKER','SALESMAN',66928,'1991/9/8',1600,0,3001);
insert into employee values(68736,'ADNRES','CLERK',67858,'1997/5/23',1200,NULL,2001);
insert into employee values(69000,'JULIUS','CLERK',66928,'1991/12/3',1050,NULL,3001);
insert into employee values(69324,'MARKER','CLERK',67832,'1992/1/23',1400,NULL,1001);
SELECT * FROM employee WHERE name like '%AD%';
```

出力 入力 コメント 0

id	name	job	manager_id	hire_date	salary	commission	dep_id
64989	ADELYN	SALESMAN	66928	1991-02-20 00:00:00	1700	400	3001
65271	WADE	SALESMAN	66928	1991-02-22 00:00:00	1350	600	3001
66564	MADDEN	SALESMAN	66928	1991-09-28 00:00:00	1350	1500	3001
68736	ADNRES	CLERK	67858	1997-05-23 00:00:00	1200	NULL	2001

14-3. 索引（インデックス）

索引（インデックス）とは

- **各種データ処理の高速化のためのもの**
- 本体となるデータに対して作成される

見出し語	場所
アセンブラ	35
データベース	110
マシンラーニング	169

索引（インデックス）

データ本体

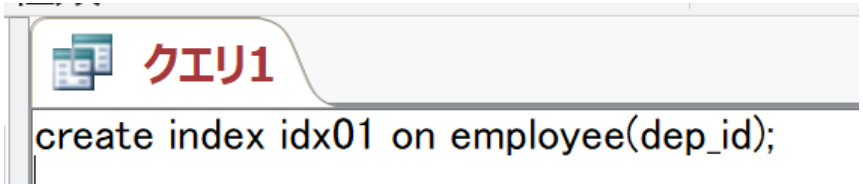
索引（インデックス）の作成

- SQL を用いて作成可能
- 作成時に、**テーブル名**と、**属性名**を**指定**

MySQL

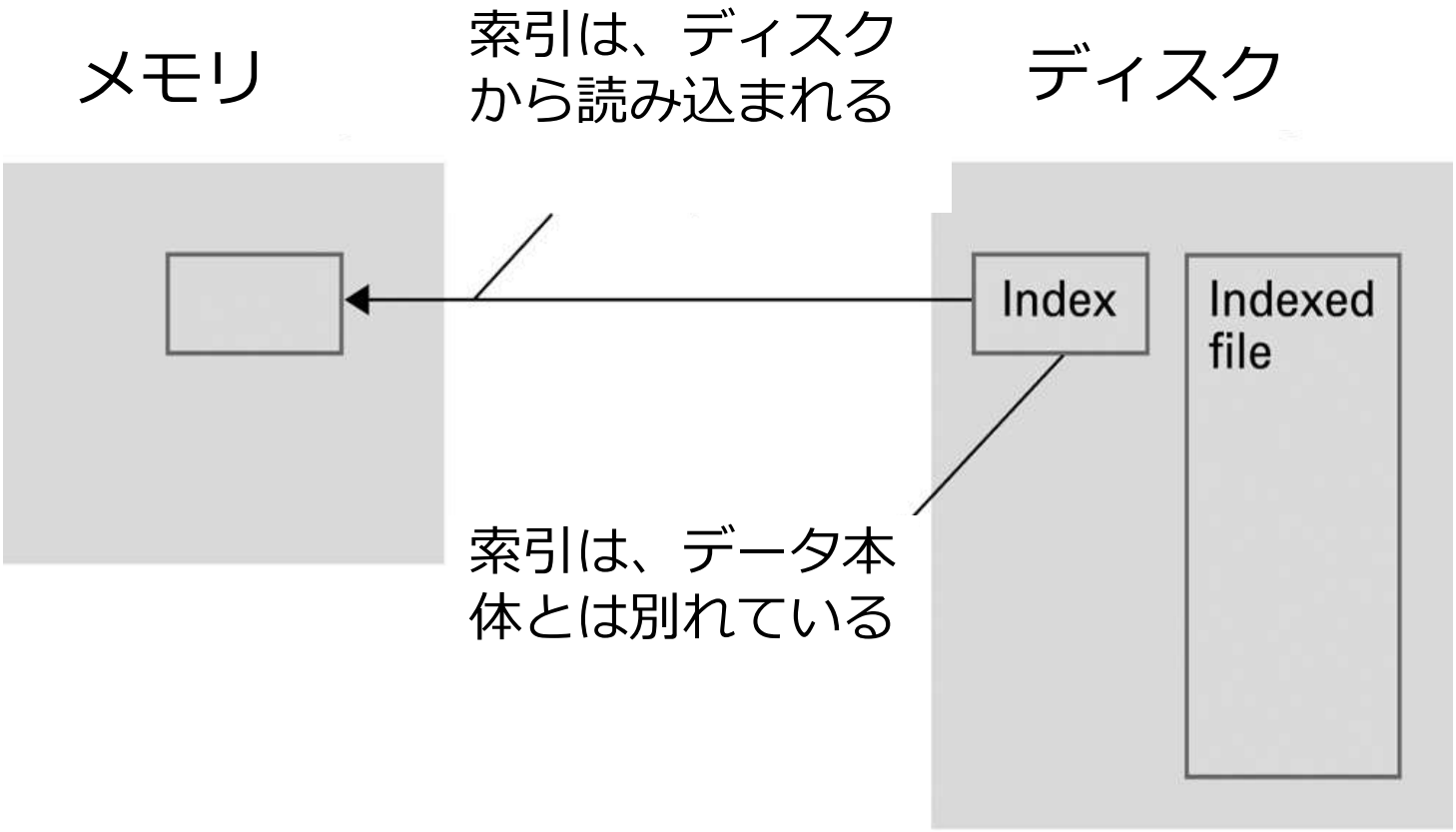
```
create index idx01 on employee(dep_id);
```

Access



テーブル employee の属性名 dep_id を用いた検索が高速に。

索引（インデックス）の仕組み



14-4. オープンデータ

オープンデータ

- 政府等が保有するデータがデジタルで公開

データカタログサイト

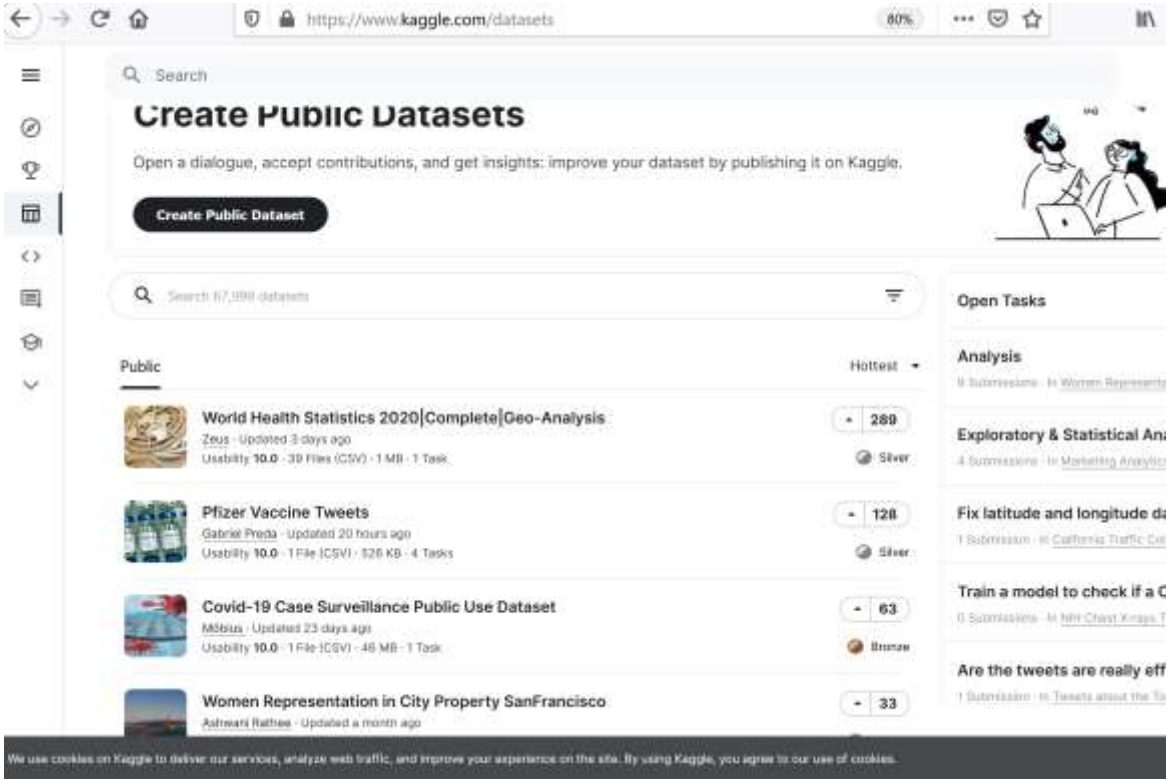
https://www.data.go.jp/

専門家の間でも、データの共有は当然

- Kaggle のデータセットのページ

https://www.kaggle.com/datasets

機械学習、データサイエンス、人工知能研究等のため



14-5. 日時情報

現在日時

- MySQL での現在日時の取得: **now()**

```
select now();
```

出力 入力 コメント 0

```
now()
2021-01-13 04:10:54
```

Paiza.IO での実行結果

- Access での現在日時の取得: **now()**

```
select now();
```

Expr1000
2021/01/13 13:13:27

Access での実行結果

- 多くの**リレーショナルデータベース管理システム**（Access や MySQL など）では、**now** を用いて、**現在日時の取得**ができる。

- 「日時」は、SQL では **datetime**

```
CREATE TABLE employee (
    id integer,
    name text,
    job text,
    manager_id integer,
    hire_date datetime,
    salary integer,
    commission integer,
    dep_id integer
);
```

日時

14-6. データウエアハウス

データウエアハウスとは

- データの**問い合わせ（クエリ）**、特に、**集計・集約**に向いたシステム
- レコード（行）の**挿入**は行える
- レコード（行）の**更新、削除**は原則**行わない**

氏名	予約内容	予約日	キャンセル日
XX	おせちA	2019/12/1	
YY	おせちB	2019/12/4	
ZZ	おせちB	2019/12/5	2019/12/6
ZZ	おせちA	2019/12/6	

予約テーブル

商品	価格	価格改定日
おせちA	12000	2019/11/10
おせちA	10000	2019/11/20
おせちB	5000	2019/11/10

価格テーブル

データウエアハウスの特色

- **過去データがシステム内に残り続ける**
- レコード（行）の**更新、削除**を原因とする**異状**は考えなくてよい　（更新、削除を行わないから）
- **正規形、正規化**を考えなくてよい

氏名	予約内容	予約日	キャンセル日
XX	おせちA	2019/12/1	
YY	おせちB	2019/12/4	
ZZ	おせちB	2019/12/5	2019/12/6
ZZ	おせちA	2019/12/6	

予約テーブル

商品	価格	価格改定日
おせちA	12000	2019/11/10
おせちA	10000	2019/11/20
おせちB	5000	2019/11/10

価格テーブル

データウエアハウスと、普通のデータベースの比較

- 普通のデータベース：**最新情報の保持**

氏名	予約内容
XX	おせちA
YY	おせちB
ZZ	おせちA

予約テーブル

商品	価格
おせちA	10000
おせちB	5000

価格テーブル

- データウエアハウス：**過去データをすべて残す**

氏名	予約内容	予約日	キャンセル日
XX	おせちA	2019/12/1	
YY	おせちB	2019/12/4	
ZZ	おせちB	2019/12/5	2019/12/6
ZZ	おせちA	2019/12/6	

予約テーブル

商品	価格	価格改定日
おせちA	12000	2019/11/10
おせちA	10000	2019/11/20
おせちB	5000	2019/11/10

価格テーブル

データウエアハウスの活用例 ①

◆大量の商品が出品され、購入されるオークションサイト

・不正行動の監視

・出品者、購入者が「どうすればより満足するか」の分析

eBay Web ページ
http://www.ebay.co.jp/

出品、購入の状況を丸ごと記録

出典: Teradata 社 Web ページ
http://jpn.teradata.jp/library/nyumon/ins_1904.html

データウエアハウスの活用例 ②

◆「空席で飛ばすくらいなら、安値でも売りたい」と思っている航空会社

・満席になりそうか、空席はいくつになりそうかを的確に予測

・空席をぴったり埋めるために、ダイナミックに値下げ／値上げ

予約、キャンセルの状況を丸ごと記録

出典: Teradata 社 Web ページ
http://jpn.teradata.jp/library/nyumon/ins_1904.html