



DATABASE SYSTEMS

Online banking data analysis



BUDAPESTI MŰSZAKI
ÉS GAZDASÁGTUDOMÁNYI EGYETEM
Építőmérnöki Kar - építőmérnöki képzés 1782 óta

Fotogrammetria és Térinformatika Tanszék

Bence Molnár

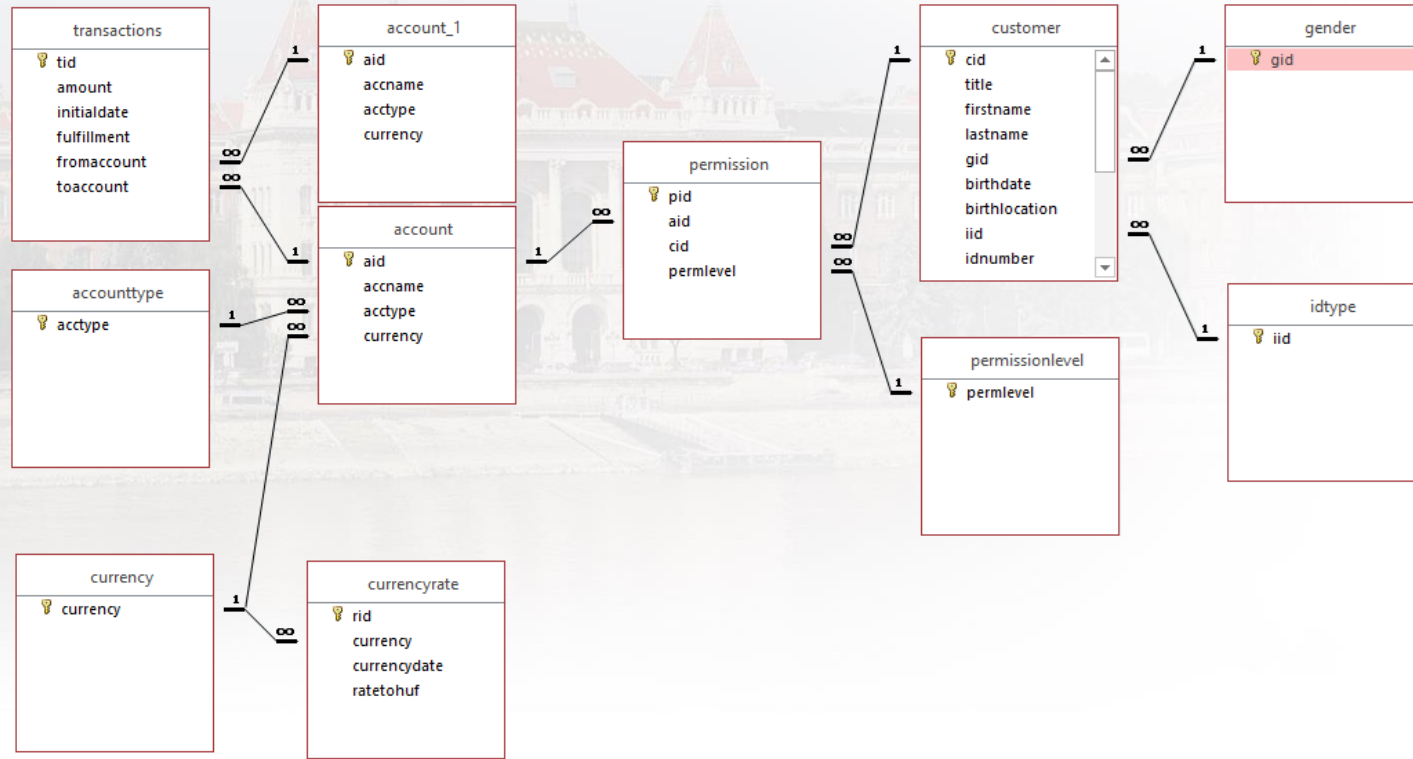
2022.02.14.

DATABASE OF A BANK

We have a database of a bank, which includes:

- Customer list
- Account list
- Transaction list
- Currency rate historical data

RELATIONAL SCHEMA DIAGRAM



QUESTIONS TO ANSWER

- List all customer in an ordered list (based on full name)
- Return the number of all customer
- Give the total cache flow (sum of all transactions)
- What was the highest rate of HUF-EUR currency exchange?
- Which customer has access to most accounts?
- Which account has the most customer access for it?

BASICS OF MS ACCESS INTERFACE

Download and open the sample dataset.

Get familiar with data structure (Database tools->Relationship)

On the left hand side we can find the existing tables, queries, forms and reports.

We have multiple options to get information from a table/query. These options are available on the left side of Home tab of ribbon menu

- Datasheet View (view and edit data stored in table)
- Design View (Table structure or query editing)
- SQL View (Only available for queries, enables a better control on query definition)

CREATING NEW OBJECTS

New tables and queries can be created on the Create tab of ribbon menu

- By a wizard with limited setting options
- Interactive interface with more functionalities

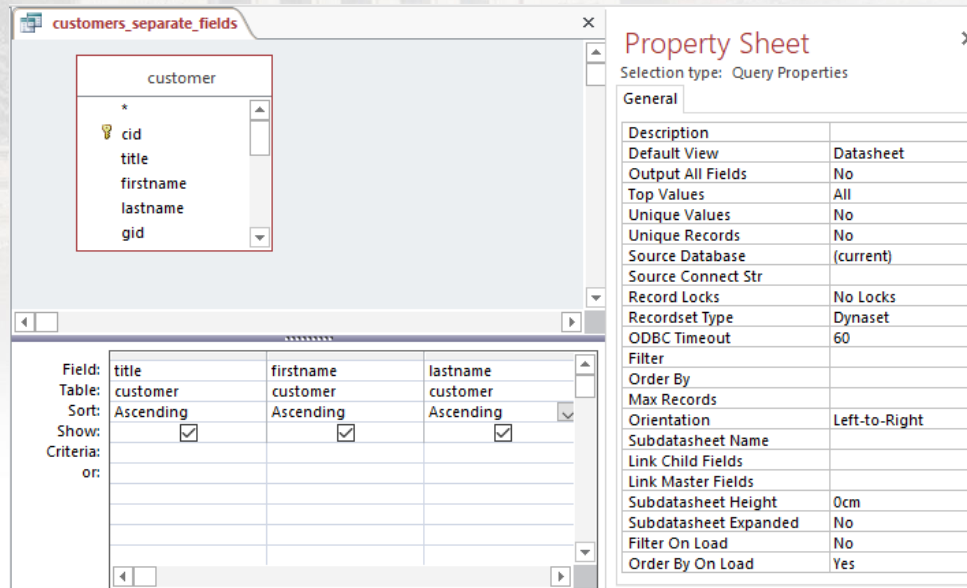
CREATING A QUERY

- Go to schema diagram and get an overview of the entire structure
- Select the required tables
- Which columns do we need?
- Do we need to filter rows?
- Do we need any aggregation?
- Do we need to group records?
- Do we need to sort the results?
- Do we need only to display a portion of the result set (paging)?



LIST OF CUSTOMERS

First just get names: title, firstname and lastname in separate columns. If sorting, which fields to consider and which order?



LIST OF CUSTOMERS

Get names, but joined into a single field.

Expr1: [customer]![title] & " " & [customer]![firstname] & " " & [customer]![lastname]

Derived field name

Table name

Column name

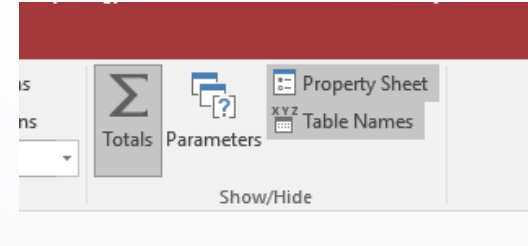
Concatenate operator

The screenshot displays the Microsoft Access interface. The main window shows a query named 'customers_singlefield' with a field list containing 'cid', 'title', 'firstname', 'lastname', and 'gid'. Below the field list, the 'Expr1' field is defined as '[customer]![title] & " " & [customer]![firstname] & " " & [customer]![lastname]'. The 'Property Sheet' is open on the right, showing the 'General' tab with various query properties.

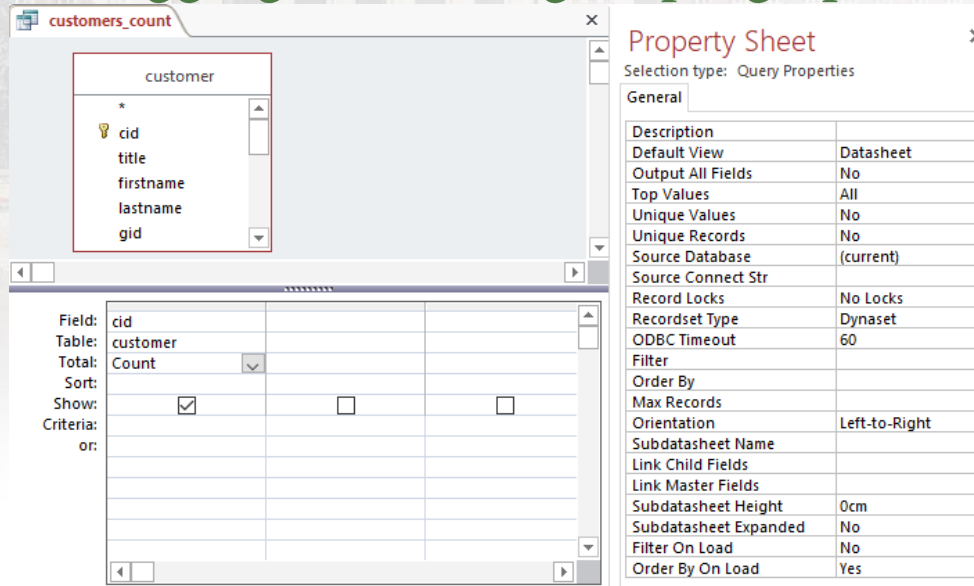
Field:	Expr1: [customer]![title]	firstname	lastname
Table:	customer	customer	customer
Sort:		Ascending	Ascending
Show:	☒	☐	☐
Criteria:			
or:			

Property Sheet	
Selection type: Query Properties	
General	
Description	Datasheet
Default View	Datasheet
Output All Fields	No
Top Values	All
Unique Values	No
Unique Records	No
Source Database	(current)
Source Connect Str	
Record Locks	No Locks
Recordset Type	Dynaset
ODBC Timeout	60
Filter	
Order By	
Max Records	
Orientation	Left-to-Right
Subdatasheet Name	
Link Child Fields	
Link Master Fields	
Subdatasheet Height	0cm
Subdatasheet Expanded	No
Filter On Load	No
Order By On Load	Yes

NUMBER OF CUSTOMERS



At ribbon menu, you can turn on and off the totals. If turned on, you can find aggregation and grouping options below.



TOTAL CASH FLOW

Apply the SUM function

The screenshot displays the Microsoft Access interface. On the left, a query window titled 'transaction_total' shows a list of fields from the 'transaction' table: 'tid' (marked with a key icon), 'amount', 'initialdate', 'fulfillment', and 'fromaccount'. Below this, a design grid is visible with the following settings:

Field:	amount		
Table:	transaction		
Total:	Sum		
Sort:			
Show:	☒	☐	☐
Criteria:			
or:			

On the right, the 'Property Sheet' for the query is open, showing 'Query Properties'. The 'General' tab is active, displaying various settings:

General	
Description	
Default View	Datasheet
Output All Fields	No
Top Values	All
Unique Values	No
Unique Records	No
Source Database	(current)
Source Connect Str	
Record Locks	No Locks
Recordset Type	Dynaset
ODBC Timeout	60
Filter	
Order By	
Max Records	
Orientation	Left-to-Right
Subdatasheet Name	
Link Child Fields	
Link Master Fields	
Subdatasheet Height	0cm
Subdatasheet Expanded	No
Filter On Load	No
Order By On Load	Yes

HUF-EUR CURRENCY RATE

First we only display the maximum value

The screenshot displays the Microsoft Access interface. On the left, a query window titled 'currencyrate_max_eur' shows a table named 'currencyrate' with fields: '*', 'rid' (marked as a primary key), 'currency', 'currencydate', and 'ratetohuf'. Below the table, the 'Criteria' section is set to 'EUR'. The 'Show' section has checkboxes for 'currency', 'ratetohuf', and 'currencydate', all of which are checked. The 'Total' section is set to 'Group By' for 'currency' and 'Max' for 'ratetohuf'. The 'Sort' section is empty. The 'Filter' section is empty. The 'Recordset Type' is set to 'Dynaset'. The 'ODBC Timeout' is set to '60'. The 'Orientation' is set to 'Left-to-Right'. The 'Subdatasheet Name' is empty. The 'Link Child Fields' and 'Link Master Fields' are empty. The 'Subdatasheet Height' is set to '0cm'. The 'Subdatasheet Expanded' is set to 'No'. The 'Filter On Load' is set to 'No'. The 'Order By On Load' is set to 'Yes'.

Property Sheet

Selection type: Query Properties

General

Description	
Default View	Datasheet
Output All Fields	No
Top Values	All
Unique Values	No
Unique Records	No
Source Database	(current)
Source Connect Str	
Record Locks	No Locks
Recordset Type	Dynaset
ODBC Timeout	60
Filter	
Order By	
Max Records	
Orientation	Left-to-Right
Subdatasheet Name	
Link Child Fields	
Link Master Fields	
Subdatasheet Height	0cm
Subdatasheet Expanded	No
Filter On Load	No
Order By On Load	Yes

HUF-EUR CURRENCY RATE

If we need to display corresponding date as well, that makes things a little bit complicated. We will learn it later: there is a need for sub-queries, joins or views.

HUF-* CURRENCY RATE

By grouping based on currency and removing filtering on HUF we can obtain all currencies maximum rate. This is where we are convinced, that databases gives a new perspective to automate analysis in an easy way.

CUSTOMERS ACCESS TO MOST ACCOUNTS

Here we need multiple tables: permission and customer

Customer is required to resolve names based on ids. Account table is not required as permission includes account numbers.

Note: we set TOP 1, but multiple records are returned because of equality.

The screenshot displays the Microsoft Access interface. The main window shows a query named 'account_count' with two tables, 'customer' and 'permission', connected by a one-to-many relationship. The 'customer' table has fields: cid (primary key), title, firstname, lastname, and gid. The 'permission' table has fields: pid (primary key), aid, cid, and permlevel. The query design grid shows the following configuration:

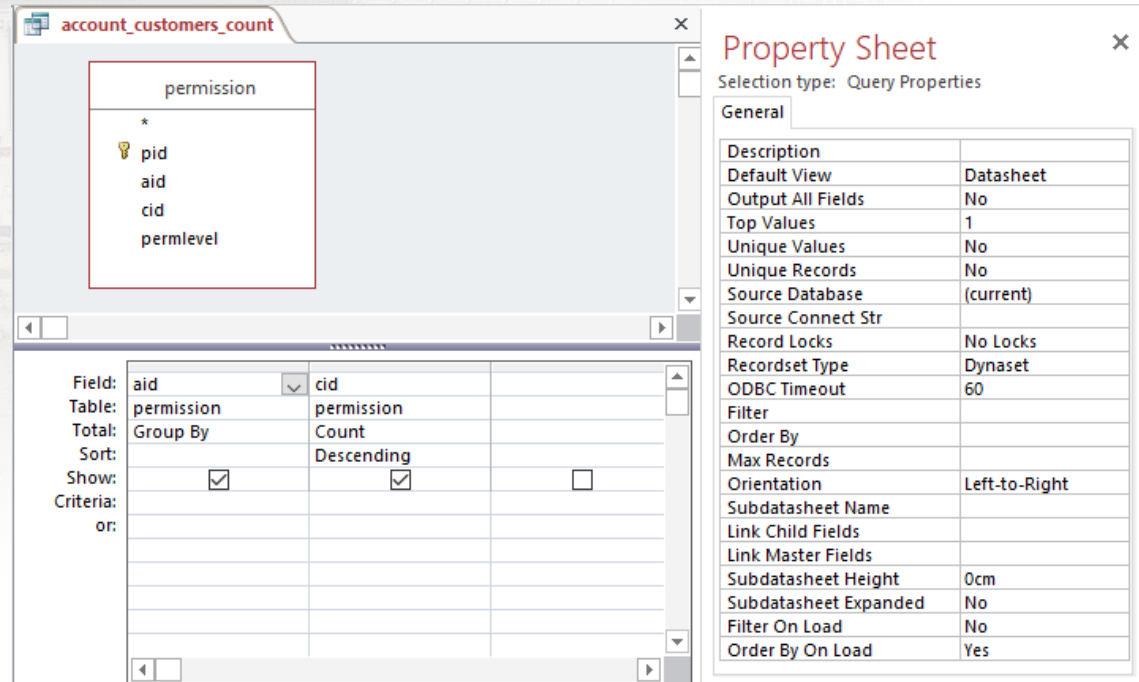
Field:	aid	title	firstname	lastname
Table:	permission	customer	customer	customer
Total:	Count	Group By	Group By	Group By
Sort:	Descending			
Show:	☒	☒	☒	☒
Criteria:				
or:				

The Property Sheet for the query is open, showing the 'General' tab. The 'Output All Fields' property is set to 'No', and the 'Top Values' property is set to '1'. The 'Description' property is set to 'Datasheet'.

Description	Value
Default View	Datasheet
Output All Fields	No
Top Values	1
Unique Values	No
Unique Records	No
Source Database	(current)
Source Connect Str	
Record Locks	No Locks
Recordset Type	Dynaset
ODBC Timeout	60
Filter	
Order By	
Max Records	
Orientation	Left-to-Right
Subdatasheet Name	
Link Child Fields	
Link Master Fields	
Subdatasheet Height	0cm
Subdatasheet Expanded	No
Filter On Load	No
Order By On Load	Yes

THE ACCOUNT WITH TOP NUMBER OF ELIGIBLE CUSTOMERS (WITH PERMISSION)

We need only a single table



The screenshot displays the Microsoft Access interface. The main window shows a query named 'account_customers_count' in Design View. The query is based on the 'permission' table, which has fields: pid (primary key), aid, cid, and permlevel. The query is set to show the top 1 values, sorted by aid in descending order. The Property Sheet on the right shows the query's properties, including the default view (Datasheet), output all fields (No), top values (1), unique values (No), unique records (No), source database (current), source connect string, record locks (No Locks), recordset type (Dynaset), ODBC timeout (60), filter, order by, max records, orientation (Left-to-Right), subdatasheet name, link child fields, link master fields, subdatasheet height (0cm), subdatasheet expanded (No), filter on load (No), and order by on load (Yes).

account_customers_count

permission

- pid
- aid
- cid
- permlevel

Field: aid cid
Table: permission permission
Total: Group By Count
Sort: Descending
Show: ☒ ☒ ☐
Criteria:
or:

Property Sheet

Selection type: Query Properties

General

Description	
Default View	Datasheet
Output All Fields	No
Top Values	1
Unique Values	No
Unique Records	No
Source Database	(current)
Source Connect Str	
Record Locks	No Locks
Recordset Type	Dynaset
ODBC Timeout	60
Filter	
Order By	
Max Records	
Orientation	Left-to-Right
Subdatasheet Name	
Link Child Fields	
Link Master Fields	
Subdatasheet Height	0cm
Subdatasheet Expanded	No
Filter On Load	No
Order By On Load	Yes

THE ACCOUNT WITH TOP NUMBER OF ELIGIBLE CUSTOMERS (WITH PERMISSION)

Considering permission levels

The screenshot displays the Microsoft Access interface. The main window shows a query named 'account_customers_count_permlevel'. The query design grid is visible, showing fields from the 'permission' table: 'pid' (with a key icon), 'aid', 'cid', and 'permlevel'. The 'pid' field is marked with an asterisk, indicating it is the primary key. The 'Where' clause is highlighted with a red circle, containing the text 'owner' Or 'partner'.

The Property Sheet on the right shows the 'General' tab. The 'Selection type' is set to 'Query Properties'. The 'Description' is 'Datasheet'. The 'Default View' is 'No'. The 'Output All Fields' is 'No'. The 'Top Values' is '1'. The 'Unique Values' is 'No'. The 'Unique Records' is 'No'. The 'Source Database' is '(current)'. The 'Source Connect Str' is 'No Locks'. The 'Recordset Type' is 'Dynaset'. The 'ODBC Timeout' is '60'. The 'Filter' is 'None'. The 'Order By' is 'None'. The 'Max Records' is 'None'. The 'Orientation' is 'Left-to-Right'. The 'Subdatasheet Name' is 'None'. The 'Link Child Fields' is 'None'. The 'Link Master Fields' is 'None'. The 'Subdatasheet Height' is '0cm'. The 'Subdatasheet Expanded' is 'No'. The 'Filter On Load' is 'No'. The 'Order By On Load' is 'Yes'.

Field:	aid	cid	permlevel
Table:	permission	permission	permission
Total:	Group By	Count	Where
Sort:		Descending	
Show:	☒	☒	☐
Criteria:			"owner" Or "partner"
or:			

Description	
Default View	Datasheet
Output All Fields	No
Top Values	1
Unique Values	No
Unique Records	No
Source Database	(current)
Source Connect Str	No Locks
Recordset Type	Dynaset
ODBC Timeout	60
Filter	None
Order By	None
Max Records	None
Orientation	Left-to-Right
Subdatasheet Name	None
Link Child Fields	None
Link Master Fields	None
Subdatasheet Height	0cm
Subdatasheet Expanded	No
Filter On Load	No
Order By On Load	Yes

SUMMARY

- Available dataset and structure
- Creating queries
 - Required columns
 - Column concatenation
 - Filtering rows
 - Aggregating rows
 - Grouping rows
 - Query multiple tables



Questions?



BUDAPESTI MŰSZAKI
ÉS GAZDASÁGTUDOMÁNYI EGYETEM

Építőmérnöki Kar - építőmérnöki képzés 1782 óta

Fotogrammetria és Térinformatika Tanszék