

Getting the Information You Need from CDW: SQL Starter Language

by Margaret Gonsoulin, PhD



Thanks to:

- Richard Pham, BISL/CDW
- Colleagues at VIREC

Agenda for Today

- This is follow-up to:
 - [First Time Research User's Guide to CDW](#)
 - [Seeing the Data When You Can't See the Data](#)
- Review key terms and definitions
- Orient new SQL users to CDW's data organization
- Provide new CDW users with a few basic SQL programming skills

Reminder - 2 Domain Types

- Production Domains – contains tables that have been structured by database architects to support their re-joining.
- Raw Domains – contains tables that are direct extracts from the source system (e.g., VistA) that are simply being housed with no editing performed on them.
- These types of domains also vary in terms of their naming traditions.

CDW, VINCI & SQL

- CDW –Corporate Data Warehouse – a relational database that contains significant portions of the VA’s electronic health record.
- VINCI – VA Informatics and Computing Infrastructure – (in part) a remote server that provides software and holds a copy of CDW.
- SQL – Structured Query Language - a programming language used to retrieve information from a relational database.

SSMS -SQL Server Management Studio

- SSMS is computer software that can be used to write and execute SQL code.
 - This talk assumes that you have access to CDW data and SSMS.
 - The examples included in this talk use oversimplified material from the CDW in order to illustrate the logic of SQL.
 - Therefore, this talk is not intended to demonstrate a viable research investigation.

By the end of this talk,

We hope that a new CDW user will:

- Have a basic understanding of SQL
- Be able to read a basic SQL query
- Be able to write a basic SQL query
 - For tables in the production domains
 - For tables in the raw domains

Poll #1: Your CDW Experience

Rate your level of experience with CDW data on a scale of 1 to 5...

1 Not worked with it at all

2

3

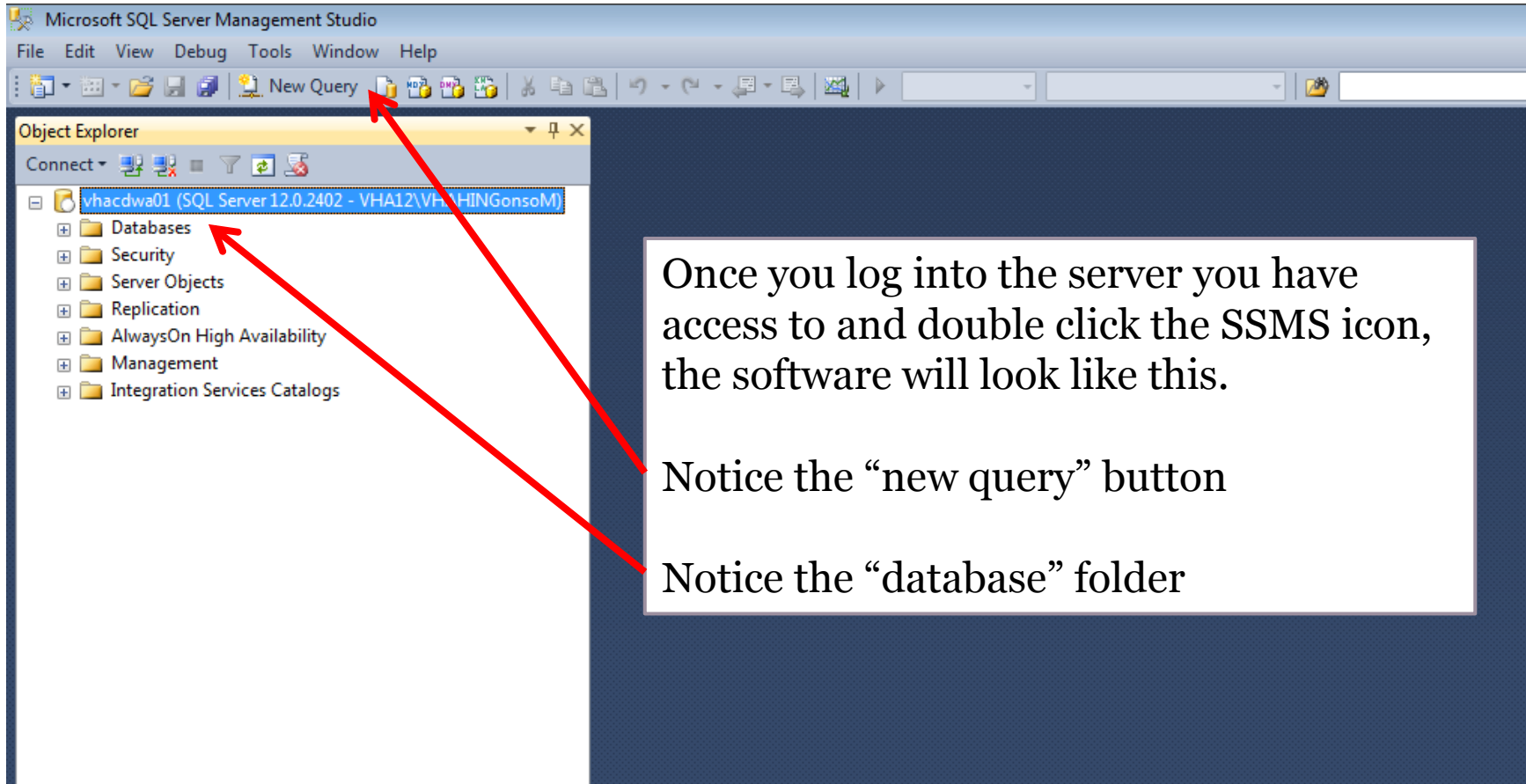
4

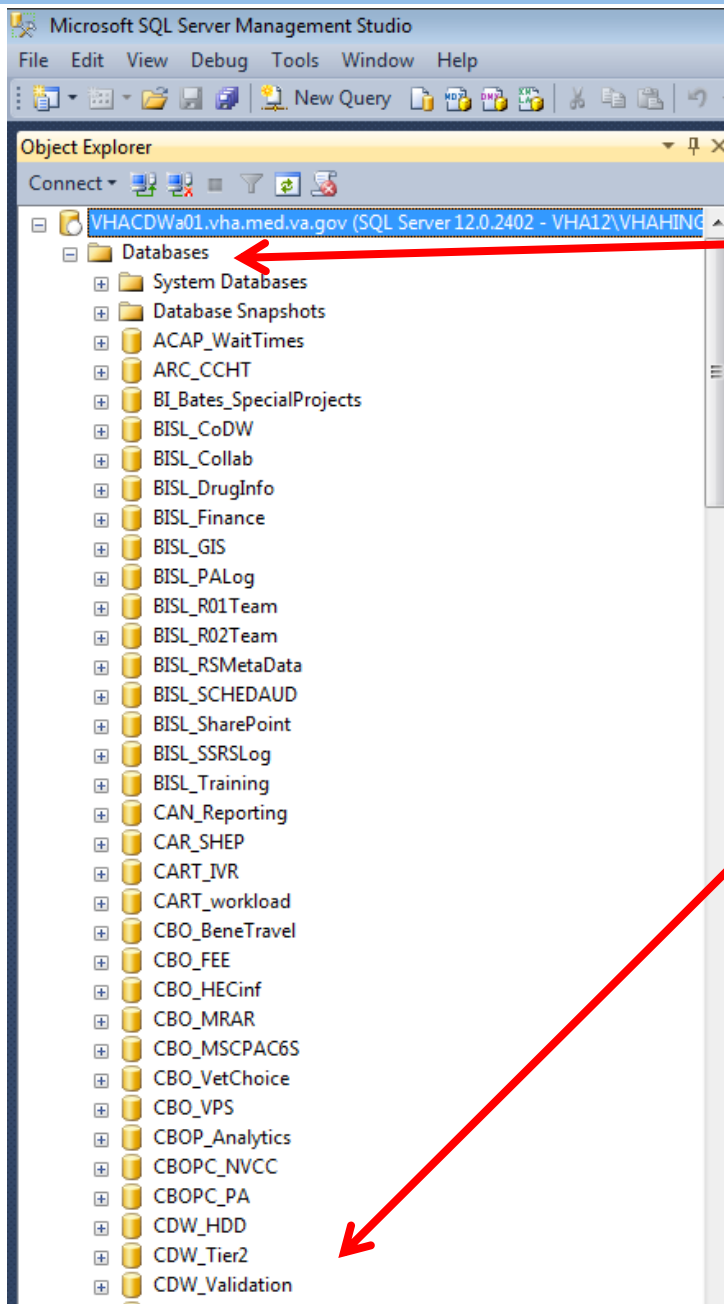
5 Very experienced with CDW

In the Production Domains

The Production table called Dim.ICD9

SQL Server Management Studio





Expand “Databases”

Then, scroll down to find “CDWWork”

Note: this example uses a view of a dimension table in the database folder called CDWWork; all people with basic access to CDW will be able to see this view.

New Query

Once you click “new query”, this window will open in the middle of the screen. You will type your SQL query here.

Use the list of *views* in CDWWork to help guide your effort.

End users of CDW work with the “views” rather than the “tables” in CDW.

Current connection parameters

Aggregate Status	
Connection failure	
Elapsed time	
Finish time	
Name	vhacdwa01
Rows returned	0
Start time	
State	Open

Connection

Connection Details	
Connection name	vhacdwa01 (VHA12\VAH)
Connection elapsed	
Connection finish	
Connection rows	0
Connection start t	
Connection state	Open
Display name	vhacdwa01
Login name	VHA12\VAH
Server name	vhacdwa01
Server version	12.0.2402
Session Tracing ID	SPID
SPID	152

Name
The name of the connection.

Connected. (1/1)

vhacdwa01 (12.0 RTM) | VHA12\VAH | CDWWork | 00:00:00 | 0 rows

The screenshot shows the Microsoft SQL Server Management Studio interface. The title bar indicates the connection to 'CDWWWork' and the database 'VHA12\VHAHNGonsoM (287)'. The Object Explorer on the left shows a tree of database objects, with 'Dim.ICD9' selected. A context menu is open over 'Dim.ICD9', with 'Select Top 1000 Rows' highlighted. A red '1st' is placed next to this menu. The SQL Query window on the right shows a query script for 'SelectTopNRows' command from 'CDWWWork'. A red '2nd' is placed next to the query. The Results window at the bottom shows the output of the query, with a red '3rd' placed next to it. The results table has columns: ICD9SID, ICDIEN, Sta3n, ICD9Code, and ICD9C.

1st Right click the view name for this drop-down menu and then click “select top 1000 rows”

2nd The query seen in the right top screen is automatically written

3rd The results below it automatically appear.

```

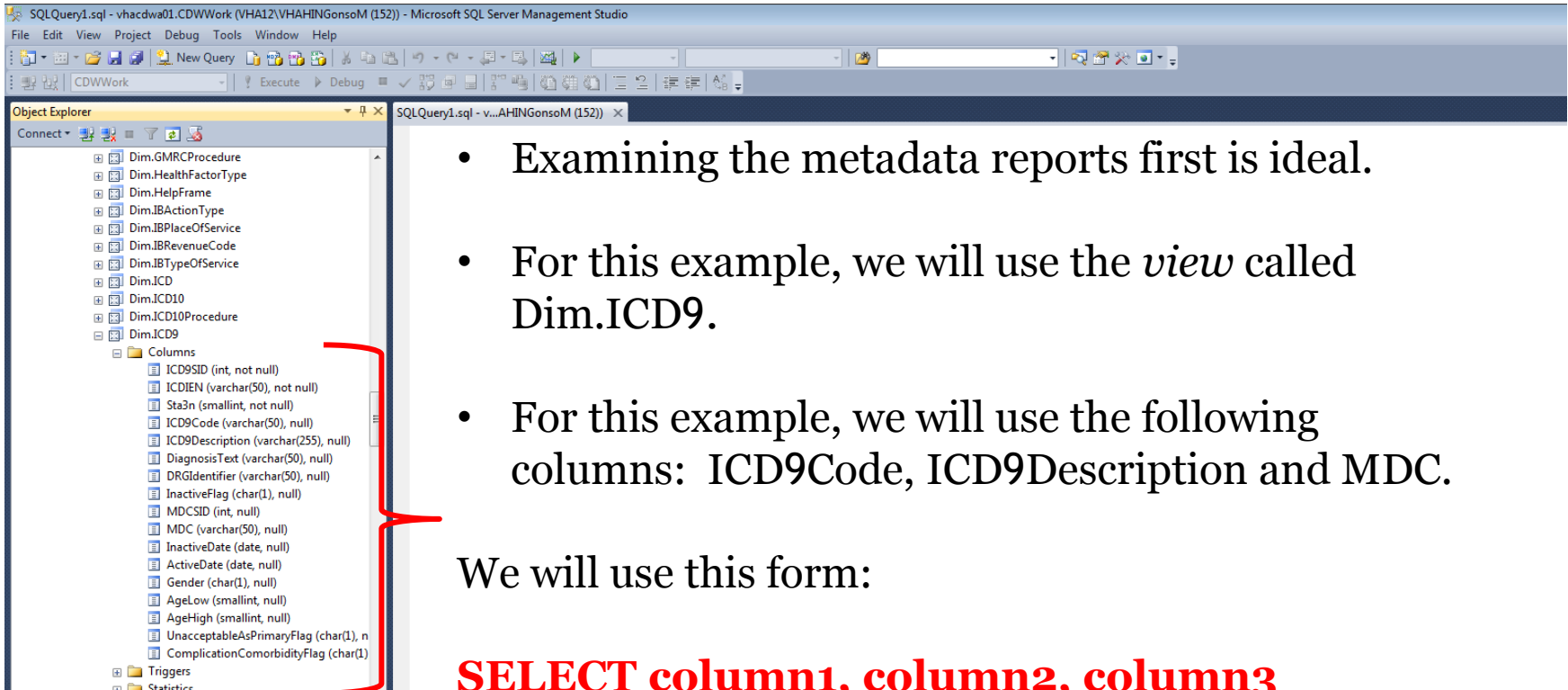
/***** Script for SelectTopNRows command from
SELECT TOP 1000 [ICD9SID]
      , [ICDIEN]
      , [Sta3n]
      , [ICD9Code]
      , [ICD9Description]
      , [DiagnosisText]
      , [DRGIdentifier]
      , [InactiveFlag]
      , [MDCSID]
      , [MDC]
      , [InactiveDate]
      , [ActiveDate]
      , [Gender]
      , [AgeLow]
      , [AgeHigh]
      , [UnacceptableAsPrimaryFlag]
      , [ComplicationComorbidityFlag]
FROM [CDWWWork].[Dim].[ICD9]
  
```

	ICD9SID	ICDIEN	Sta3n	ICD9Code	ICD9C
	-1	-1	-1	*Missing*	*Missi
	0	0	0	*Unknown at this time*	*Unkr
3	800000000	236	459	151.9	MALIK
4	800000001	237	459	152.1	MALIK
5	800000002	238	459	152.2	MALIK
6	800000003	239	459	152.3	MALIK
7	800000004	24	459	110.6	DEEF
8	800000005	240	459	152.8	MALIK
9	800000006	241	459	152.9	MALIK
10	800000007	242	459	153.1	MALIK
11	800000008	243	459	153.2	MALIK
12	800000009	244	459	153.3	MALIK
13	800000010	245	459	153.4	MALIK
14	800000011	246	459	153.5	MALIK
15	800000012	247	459	153.6	MALIK
16	800000013	248	459	153.7	MALIK
17	800000014	249	459	153.8	MALIK

SELECT and FROM

- SELECT allows a programmer to list the columns (variables) that they would like to see in the results of their query
 - Each column name should be followed by a comma except the last one in the list
- FROM selects that appropriate *view* from which the columns will be collected
 - The name of the view should be written in this format “database.schema.view”

SELECT & FROM



The screenshot shows the Object Explorer in Microsoft SQL Server Enterprise Manager. The 'Columns' section under the 'Dim.ICD9' table is highlighted with a red bracket. The columns listed are:

- ICD9SID (int, not null)
- ICDIEN (varchar(50), not null)
- Sta3n (smallint, not null)
- ICD9Code (varchar(50), null)
- ICD9Description (varchar(255), null)
- DiagnosisText (varchar(50), null)
- DRGIdentifier (varchar(50), null)
- InactiveFlag (char(1), null)
- MDCSID (int, null)
- MDC (varchar(50), null)
- InactiveDate (date, null)
- ActiveDate (date, null)
- Gender (char(1), null)
- AgeLow (smallint, null)
- AgeHigh (smallint, null)
- UnacceptableAsPrimaryFlag (char(1), null)
- ComplicationComorbidityFlag (char(1), null)

- Examining the metadata reports first is ideal.
- For this example, we will use the *view* called Dim.ICD9.
- For this example, we will use the following columns: ICD9Code, ICD9Description and MDC.

We will use this form:

**SELECT column1, column2, column3
FROM Database.Schema.View;**

SQLQuery1.sql - vhacdwa01.CDWWork (VHA12\VHAHINGoSoM (152))* - Microsoft SQL Server Management Studio

File Edit View Query Project Debug Tools Window Help

CDWWork

Object Explorer

Connect

Dim.ICD9

Columns

- ICD9SID (int, not null)
- ICDIEN (varchar(50), not null)
- Sta3n (smallint, not null)
- ICD9Code (varchar(50), null)
- ICD9Description (varchar(255), null)
- DiagnosisText (varchar(50), null)
- DRGIdentifier (varchar(50), null)
- InactiveFlag (char(1), null)
- MDCSID (int, null)
- MDC (varchar(50), null)
- InactiveDate (date, null)
- ActiveDate (date, null)
- Gender (char(1), null)
- AgeLow (smallint, null)
- AgeHigh (smallint, null)
- UnacceptableAsPrimaryFlag (char(1), null)
- ComplicationComorbidityFlag (char(1), null)

Triggers

Statistics

Dim.ICD9Procedure

Dim.ICDDescription

Dim.ICDPProcedure

Dim.IdentificationFormat

Dim.ImmunizationName

Dim.Institution

Dim.InterventionRecommendation

Dim.InterventionType

Dim.IRTDeficiency

Dim.IRTDeficiencyCategory

Dim.IRTStatus

Dim.IVAdditiveIngredient

Dim.IVAdditiveIngredientElectrolytes

Dim.IVAdditiveIngredientSynonym

Dim.IVRoom

Dim.IVSolutionIngredient

Dim.IVSolutionIngredientElectrolytes

Dim.IVSolutionIngredientSynonym

Dim.LabChemTest

Dim.LabChemTestPanelList

Dim.LabChemTestUrgency

Dim.LabCode

Dim.LabCodeSubtype

Dim.LabSection

Dim.LocalDrug

Dim.Location

SQLQuery1.sql - v...AHINGonSoM (152))*

```
SELECT ICD9Code , ICD9Description , MDC
FROM CDWWork.Dim.ICD9
```

100 %

Results Messages

	ICD9Code	ICD9Description	MDC
1	"Missing"	"Missing"	"Missing"
2	"Unknown at this time"	"Unknown at this time"	"Unknown at this time"
3	151.9	MALIGNANT NEOPLASM OF STOMACH, UNSPECIFIED	DIGESTIVE SYSTEM
4	152.1	MALIGNANT NEOPLASM OF JEJUNUM	DIGESTIVE SYSTEM
5	152.2	MALIGNANT NEOPLASM OF ILEUM	DIGESTIVE SYSTEM
6	152.3	MALIGNANT NEOPLASM OF MECKEL'S DIVERTICULUM	DIGESTIVE SYSTEM
7	110.6	DEEP SEATED DERMATOPHYTOSIS	SKIN.BREAST.SUBCUTANEOUS T
8	152.8	MALIGNANT NEOPLASM OF OTHER SPECIFIED SITES OF S...	DIGESTIVE SYSTEM
9	152.9	MALIGNANT NEOPLASM OF SMALL INTESTINE, UNSPECIFIED	DIGESTIVE SYSTEM
10	153.1	MALIGNANT NEOPLASM OF TRANSVERSE COLON	DIGESTIVE SYSTEM
11	153.2	MALIGNANT NEOPLASM OF DESCENDING COLON	DIGESTIVE SYSTEM
12	153.3	MALIGNANT NEOPLASM OF SIGMOID COLON	DIGESTIVE SYSTEM
13	153.4	MALIGNANT NEOPLASM OF CECUM	DIGESTIVE SYSTEM
14	153.5	MALIGNANT NEOPLASM OF APPENDIX VERMIFORMIS	DIGESTIVE SYSTEM
15	153.6	MALIGNANT NEOPLASM OF ASCENDING COLON	DIGESTIVE SYSTEM
16	153.7	MALIGNANT NEOPLASM OF SPLENIC FLEXURE	DIGESTIVE SYSTEM
17	153.8	MALIGNANT NEOPLASM OF OTHER SPECIFIED SITES OF LA...	DIGESTIVE SYSTEM
18	110.8	DERMATOPHYTOSIS OF OTHER SPECIFIED SITES	SKIN.BREAST.SUBCUTANEOUS T
19	153.9	MALIGNANT NEOPLASM OF COLON, UNSPECIFIED	DIGESTIVE SYSTEM

Query executed successfully.

**SELECT ICD9Code, ICD9Description, MDC
FROM CDWWork.Dim.ICD9;**



2,025,871 results
or rows appear
here in a new
window.

We can ask for the
results in order.

ORDER BY

- The ORDER BY command allows the programmer to see the results in alpha-numeric order based on the column that is chosen.
- ORDER BY must be the last command in a query.

```
SELECT column1, column2, column3  
FROM Database.Schema.View  
ORDER BY column1 ;
```

ORDER BY

The screenshot shows the Microsoft SQL Server Management Studio interface. The title bar indicates the file is 'SQLQuery1.sql' and the server is 'vha12\VAHAGonsoM (152)'. The menu bar includes File, Edit, View, Query, Project, Debug, Tools, Window, and Help. The toolbar contains icons for New Query, Execute, and other standard database operations. The Object Explorer on the left shows a tree view of the database schema, including tables like Dim.ICD9, Dim.ICD9Procedure, and Dim.LabChemTest. The main query window displays the following SQL code:

```
SELECT ICD9Code, ICD9Description, MDC
FROM CDWork.Dim.ICD9
ORDER BY ICD9Code ;
```

The Results pane at the bottom shows the output of the query. It displays a table with three columns: ICD9Code, ICD9Description, and MDC. The data shows multiple rows for the same ICD9Code (971.0, 971.1, 971.11, 971.12, 971.13, 971.14, 971.15, 971.16, 971.17, 971.18, 971.19, 971.2, 971.21, 971.22, 971.23, 971.24, 971.25, 971.26, 971.27, 971.28, 971.29, 971.3, 971.31, 971.32, 971.33, 971.34, 971.35, 971.36, 971.37, 971.38, 971.39, 971.4, 971.41, 971.42, 971.43, 971.44, 971.45, 971.46, 971.47, 971.48, 971.49, 971.5, 971.51, 971.52, 971.53, 971.54, 971.55, 971.56, 971.57, 971.58, 971.59, 971.6, 971.61, 971.62, 971.63, 971.64, 971.65, 971.66, 971.67, 971.68, 971.69, 971.7, 971.71, 971.72, 971.73, 971.74, 971.75, 971.76, 971.77, 971.78, 971.79, 971.8, 971.81, 971.82, 971.83, 971.84, 971.85, 971.86, 971.87, 971.88, 971.89, 971.9, 971.91, 971.92, 971.93, 971.94, 971.95, 971.96, 971.97, 971.98, 971.99, 972.0, 972.01, 972.02, 972.03, 972.04, 972.05, 972.06, 972.07, 972.08, 972.09, 972.1, 972.11, 972.12, 972.13, 972.14, 972.15, 972.16, 972.17, 972.18, 972.19, 972.2, 972.21, 972.22, 972.23, 972.24, 972.25, 972.26, 972.27, 972.28, 972.29, 972.3, 972.31, 972.32, 972.33, 972.34, 972.35, 972.36, 972.37, 972.38, 972.39, 972.4, 972.41, 972.42, 972.43, 972.44, 972.45, 972.46, 972.47, 972.48, 972.49, 972.5, 972.51, 972.52, 972.53, 972.54, 972.55, 972.56, 972.57, 972.58, 972.59, 972.6, 972.61, 972.62, 972.63, 972.64, 972.65, 972.66, 972.67, 972.68, 972.69, 972.7, 972.71, 972.72, 972.73, 972.74, 972.75, 972.76, 972.77, 972.78, 972.79, 972.8, 972.81, 972.82, 972.83, 972.84, 972.85, 972.86, 972.87, 972.88, 972.89, 972.9, 972.91, 972.92, 972.93, 972.94, 972.95, 972.96, 972.97, 972.98, 972.99, 973.0, 973.01, 973.02, 973.03, 973.04, 973.05, 973.06, 973.07, 973.08, 973.09, 973.1, 973.11, 973.12, 973.13, 973.14, 973.15, 973.16, 973.17, 973.18, 973.19, 973.2, 973.21, 973.22, 973.23, 973.24, 973.25, 973.26, 973.27, 973.28, 973.29, 973.3, 973.31, 973.32, 973.33, 973.34, 973.35, 973.36, 973.37, 973.38, 973.39, 973.4, 973.41, 973.42, 973.43, 973.44, 973.45, 973.46, 973.47, 973.48, 973.49, 973.5, 973.51, 973.52, 973.53, 973.54, 973.55, 973.56, 973.57, 973.58, 973.59, 973.6, 973.61, 973.62, 973.63, 973.64, 973.65, 973.66, 973.67, 973.68, 973.69, 973.7, 973.71, 973.72, 973.73, 973.74, 973.75, 973.76, 973.77, 973.78, 973.79, 973.8, 973.81, 973.82, 973.83, 973.84, 973.85, 973.86, 973.87, 973.88, 973.89, 973.9, 973.91, 973.92, 973.93, 973.94, 973.95, 973.96, 973.97, 973.98, 973.99, 974.0, 974.01, 974.02, 974.03, 974.04, 974.05, 974.06, 974.07, 974.08, 974.09, 974.1, 974.11, 974.12, 974.13, 974.14, 974.15, 974.16, 974.17, 974.18, 974.19, 974.2, 974.21, 974.22, 974.23, 974.24, 974.25, 974.26, 974.27, 974.28, 974.29, 974.3, 974.31, 974.32, 974.33, 974.34, 974.35, 974.36, 974.37, 974.38, 974.39, 974.4, 974.41, 974.42, 974.43, 974.44, 974.45, 974.46, 974.47, 974.48, 974.49, 974.5, 974.51, 974.52, 974.53, 974.54, 974.55, 974.56, 974.57, 974.58, 974.59, 974.6, 974.61, 974.62, 974.63, 974.64, 974.65, 974.66, 974.67, 974.68, 974.69, 974.7, 974.71, 974.72, 974.73, 974.74, 974.75, 974.76, 974.77, 974.78, 974.79, 974.8, 974.81, 974.82, 974.83, 974.84, 974.85, 974.86, 974.87, 974.88, 974.89, 974.9, 974.91, 974.92, 974.93, 974.94, 974.95, 974.96, 974.97, 974.98, 974.99, 975.0, 975.01, 975.02, 975.03, 975.04, 975.05, 975.06, 975.07, 975.08, 975.09, 975.1, 975.11, 975.12, 975.13, 975.14, 975.15, 975.16, 975.17, 975.18, 975.19, 975.2, 975.21, 975.22, 975.23, 975.24, 975.25, 975.26, 975.27, 975.28, 975.29, 975.3, 975.31, 975.32, 975.33, 975.34, 975.35, 975.36, 975.37, 975.38, 975.39, 975.4, 975.41, 975.42, 975.43, 975.44, 975.45, 975.46, 975.47, 975.48, 975.49, 975.5, 975.51, 975.52, 975.53, 975.54, 975.55, 975.56, 975.57, 975.58, 975.59, 975.6, 975.61, 975.62, 975.63, 975.64, 975.65, 975.66, 975.67, 975.68, 975.69, 975.7, 975.71, 975.72, 975.73, 975.74, 975.75, 975.76, 975.77, 975.78, 975.79, 975.8, 975.81, 975.82, 975.83, 975.84, 975.85, 975.86, 975.87, 975.88, 975.89, 975.9, 975.91, 975.92, 975.93, 975.94, 975.95, 975.96, 975.97, 975.98, 975.99, 976.0, 976.01, 976.02, 976.03, 976.04, 976.05, 976.06, 976.07, 976.08, 976.09, 976.1, 976.11, 976.12, 976.13, 976.14, 976.15, 976.16, 976.17, 976.18, 976.19, 976.2, 976.21, 976.22, 976.23, 976.24, 976.25, 976.26, 976.27, 976.28, 976.29, 976.3, 976.31, 976.32, 976.33, 976.34, 976.35, 976.36, 976.37, 976.38, 976.39, 976.4, 976.41, 976.42, 976.43, 976.44, 976.45, 976.46, 976.47, 976.48, 976.49, 976.5, 976.51, 976.52, 976.53, 976.54, 976.55, 976.56, 976.57, 976.58, 976.59, 976.6, 976.61, 976.62, 976.63, 976.64, 976.65, 976.66, 976.67, 976.68, 976.69, 976.7, 976.71, 976.72, 976.73, 976.74, 976.75, 976.76, 976.77, 976.78, 976.79, 976.8, 976.81, 976.82, 976.83, 976.84, 976.85, 976.86, 976.87, 976.88, 976.89, 976.9, 976.91, 976.92, 976.93, 976.94, 976.95, 976.96, 976.97, 976.98, 976.99, 977.0, 977.01, 977.02, 977.03, 977.04, 977.05, 977.06, 977.07, 977.08, 977.09, 977.1, 977.11, 977.12, 977.13, 977.14, 977.15, 977.16, 977.17, 977.18, 977.19, 977.2, 977.21, 977.22, 977.23, 977.24, 977.25, 977.26, 977.27, 977.28, 977.29, 977.3, 977.31, 977.32, 977.33, 977.34, 977.35, 977.36, 977.37, 977.38, 977.39, 977.4, 977.41, 977.42, 977.43, 977.44, 977.45, 977.46, 977.47, 977.48, 977.49, 977.5, 977.51, 977.52, 977.53, 977.54, 977.55, 977.56, 977.57, 977.58, 977.59, 977.6, 977.61, 977.62, 977.63, 977.64, 977.65, 977.66, 977.67, 977.68, 977.69, 977.7, 977.71, 977.72, 977.73, 977.74, 977.75, 977.76, 977.77, 977.78, 977.79, 977.8, 977.81, 977.82, 977.83, 977.84, 977.85, 977.86, 977.87, 977.88, 977.89, 977.9, 977.91, 977.92, 977.93, 977.94, 977.95, 977.96, 977.97, 977.98, 977.99, 978.0, 978.01, 978.02, 978.03, 978.04, 978.05, 978.06, 978.07, 978.08, 978.09, 978.1, 978.11, 978.12, 978.13, 978.14, 978.15, 978.16, 978.17, 978.18, 978.19, 978.2, 978.21, 978.22, 978.23, 978.24, 978.25, 978.26, 978.27, 978.28, 978.29, 978.3, 978.31, 978.32, 978.33, 978.34, 978.35, 978.36, 978.37, 978.38, 978.39, 978.4, 978.41, 978.42, 978.43, 978.44, 978.45, 978.46, 978.47, 978.48, 978.49, 978.5, 978.51, 978.52, 978.53, 978.54, 978.55, 978.56, 978.57, 978.58, 978.59, 978.6, 978.61, 978.62, 978.63, 978.64, 978.65, 978.66, 978.67, 978.68, 978.69, 978.7, 978.71, 978.72, 978.73, 978.74, 978.75, 978.76, 978.77, 978.78, 978.79, 978.8, 978.81, 978.82, 978.83, 978.84, 978.85, 978.86, 978.87, 978.88, 978.89, 978.9, 978.91, 978.92, 978.93, 978.94, 978.95, 978.96, 978.97, 978.98, 978.99, 979.0, 979.01, 979.02, 979.03, 979.04, 979.05, 979.06, 979.07, 979.08, 979.09, 979.1, 979.11, 979.12, 979.13, 979.14, 979.15, 979.16, 979.17, 979.18, 979.19, 979.2, 979.21, 979.22, 979.23, 979.24, 979.25, 979.26, 979.27, 979.28, 979.29, 979.3, 979.31, 979.32, 979.33, 979.34, 979.35, 979.36, 979.37, 979.38, 979.39, 979.4, 979.41, 979.42, 979.43, 979.44, 979.45, 979.46, 979.47, 979.48, 979.49, 979.5, 979.51, 979.52, 979.53, 979.54, 979.55, 979.56, 979.57, 979.58, 979.59, 979.6, 979.61, 979.62, 979.63, 979.64, 979.65, 979.66, 979.67, 979.68, 979.69, 979.7, 979.71, 979.72, 979.73, 979.74, 979.75, 979.76, 979.77, 979.78, 979.79, 979.8, 979.81, 979.82, 979.83, 979.84, 979.85, 979.86, 979.87, 979.88, 979.89, 979.9, 979.91, 979.92, 979.93, 979.94, 979.95, 979.96, 979.97, 979.98, 979.99, 980.0, 980.01, 980.02, 980.03, 980.04, 980.05, 980.06, 980.07, 980.08, 980.09, 980.1, 980.11, 980.12, 980.13, 980.14, 980.15, 980.16, 980.17, 980.18, 980.19, 980.2, 980.21, 980.22, 980.23, 980.24, 980.25, 980.26, 980.27, 980.28, 980.29, 980.3, 980.31, 980.32, 980.33, 980.34, 980.35, 980.36, 980.37, 980.38, 980.39, 980.4, 980.41, 980.42, 980.43, 980.44, 980.45, 980.46, 980.47, 980.48, 980.49, 980.5, 980.51, 980.52, 980.53, 980.54, 980.55, 980.56, 980.57, 980.58, 980.59, 980.6, 980.61, 980.62, 980.63, 980.64, 980.65, 980.66, 980.67, 980.68, 980.69, 980.7, 980.71, 980.72, 980.73, 980.74, 980.75, 980.76, 980.77, 980.78, 980.79, 980.8, 980.81, 980.82, 980.83, 980.84, 980.85, 980.86, 980.87, 980.88, 980.89, 980.9, 980.91, 980.92, 980.93, 980.94, 980.95, 980.96, 980.97, 980.98, 980.99, 981.0, 981.01, 981.02, 981.03, 981.04, 981.05, 981.06, 981.07, 981.08, 981.09, 981.1, 981.11, 981.12, 981.13, 981.14, 981.15, 981.16, 981.17, 981.18, 981.19, 981.2, 981.21, 981.22, 981.23, 981.24, 981.25, 981.26, 981.27, 981.28, 981.29, 981.3, 981.31, 981.32, 981.33, 981.34, 981.35, 981.36, 981.37, 981.38, 981.39, 981.4, 981.41, 981.42, 981.43, 981.44, 981.45, 981.46, 981.47, 981.48, 981.49, 981.5, 981.51, 981.52, 981.53, 981.54, 981.55, 981.56, 981.57, 981.58, 981.59, 981.6, 981.61, 981.62, 981.63, 981.64, 981.65, 981.66, 981.67, 981.68, 981.69, 981.7, 981.71, 981.72, 981.73, 981.74, 981.75, 981.76, 981.77, 981.78, 981.79, 981.8, 981.81, 981.82, 981.83, 981.84, 981.85, 981.86, 981.87, 981.88, 981.89, 981.9, 981.91, 981.92, 981.93, 981.94, 981.95, 981.96, 981.97, 981.98, 981.99, 982.0, 982.01, 982.02, 982.03, 982.04, 982.05, 982.06, 982.07, 982.08, 982.09, 982.1, 982.11, 982.12, 982.13, 982.14, 982.15, 982.16, 982.17, 982.18, 982.19, 982.2, 982.21, 982.22, 982.23, 982.24, 982.25, 982.26, 982.27, 982.28, 982.29, 982.3, 982.31, 982.32, 982.33, 982.34, 982.35, 982.36, 982.37, 982.38, 982.39, 982.4, 982.41, 982.42, 982.43, 982.44, 982.45, 982.46, 982.47, 982.48, 982.49, 982.5, 982.51, 982.52, 982.53, 982.54, 982.55, 982.56, 982.57, 982.58, 982.59, 982.6, 982.61, 982.62, 982.63, 982.64, 982.65, 982.66, 982.67, 982.68, 982.69, 982.7, 982.71, 982.72, 982.73, 982.74, 982.75, 982.76, 982.77, 982.78, 982.79, 982.8, 982.81, 982.82, 982.83, 982.84, 982.85, 982.86, 982.87, 982.88, 982.89, 982.9, 982.91, 982.92, 982.93, 982.94, 982.95, 982.96, 982.97, 982.98, 982.99, 983.0, 983.01, 983.02, 983.03, 983.04, 983.05, 983.06, 983.07, 983.08, 983.09, 983.1, 983.11, 983.12, 983.13, 983.14, 983.15, 983.16, 983.17, 983.18, 983.19, 983.2, 983.21, 983.22, 983.23, 983.24, 983.25, 983.26, 983.27, 983.28, 983.29, 983.3, 983.31, 983.32, 983.33, 983.34, 983.35, 983.36, 983.37, 983.38, 983.39, 983.4, 983.41, 983.42, 983.43, 983.44, 983.45, 983.46, 983.47, 983.48, 983.49, 983.5, 983.51, 983.52, 983.53, 983.54, 983.55, 983.56, 983.57, 983.58, 983.59, 983.6, 983.61, 983.62, 983.63, 983.64, 983.65, 983.66, 983.67, 983.68, 983.69, 983.7, 983.71, 983.72, 983.73, 983.74, 983.75, 983.76, 983.77, 983.78, 983.79, 983.8, 983.81, 983.82, 983.83, 983.84, 983.85, 983.86, 983.87, 983.88, 983.89, 983.9, 983.91, 983.92, 983.93, 983.94, 983.95, 983.96, 983.97, 983.98, 983.99, 984.0, 984.01, 984.02, 984.03, 984.04, 984.05, 984.06, 984.07, 984.08, 984.09, 984.1, 984.11, 984.12, 984.13, 984.14, 984.15, 984.16, 984.17, 984.18, 984.19, 984.2, 984.21, 984.22, 984.23, 984.24, 984.25, 984.26, 984.27, 984.28, 984.29, 984.3, 984.31, 984.32, 984.33, 984.34, 984.35, 984.36, 984.37, 984.38, 984.39, 984.4, 984.41, 984.42, 984.43, 984.44, 984.45, 984.46, 984.47, 984.48, 984.49, 984.5, 984.51, 984.52, 984.53, 984.54, 984.55, 984.56, 984.57, 984.58, 984.59, 984.6, 984.61, 984.62, 984.63, 984.64, 984.65, 984.66, 984.67, 984.68, 984.69, 984.7, 984.71, 984.72, 984.73, 984.74, 984.75, 984.76, 984.77, 984.78, 984.79, 984.8, 984.81, 984.82, 984.83, 984.84, 984.85, 984.86, 984.87, 984.88, 984.89, 984.9, 984.91, 984.92, 984.93, 984.94, 984.95, 984.96, 984.97, 984.98, 984.99, 985.0, 985.01, 985.02, 985.03, 985.04, 985.05, 985.06, 985.07, 985.08, 985.09, 985.1, 985.11, 985.12, 985.13, 985.14, 985.15, 985.16, 985.17, 985.18, 985.19, 985.2, 985.21, 985.22, 985.23, 985.24, 985.25, 985.26, 985.27, 985.28, 985.29, 985.3, 985.31, 985.32, 985.33, 985.34, 985.35, 985.36, 985.37, 985.38, 985.39, 985.4, 985.41, 985.42, 985.43, 985.44, 985.45, 985.46, 985.47, 985.48, 985.49, 985.5, 985.51, 985.52, 985.53, 985.54, 985.55, 985.56, 985.57, 985.58, 985.59, 985.6, 985.61, 985.62, 985.63, 985.64, 985.65, 985.66, 985.67, 985.68, 985.69, 985.7, 985.71, 985.72, 985.73, 985.74, 985.75, 985.76, 985.77, 985.78, 985.79, 985.8, 985.81, 985.82, 985.83, 985.84, 985.85, 985.86, 985.87, 985.88, 985.89, 985.9, 985.91, 985.92, 985.93, 985.94, 985.95, 985.96, 985.97, 985.98, 985.99, 986.0, 986.01, 986.02, 986.03, 986.04, 986.05, 986.06, 986.07, 986.08, 986.09, 986.1, 986.11, 986.12, 986.13, 986.14, 986.15, 986.16, 986.17, 986.18, 986.19, 986.2, 986.21, 986.22, 986.23, 986.24, 986.25, 986.26, 986.27, 986.28, 986.29, 986.3, 986.31, 986.32, 986.33

GROUP BY

- The GROUP BY command collapses identical values into a single entry rather than allowing repetitive rows to exist in the output.
- Remembering that ORDER BY must be the last command in a query, GROUP BY should precede it.
- Every column that is not an aggregated* measure must be included in the GROUP BY clause.

```
SELECT column1, column2, column3  
FROM Database.Schema.Table  
GROUP BY column1, column2 , column3  
ORDER BY column1 ;
```

GROUP BY

The screenshot shows the Microsoft SQL Server Management Studio interface. The Object Explorer on the left displays the database structure, including tables like Dim.ICD9, Dim.ICD9Procedure, and Dim.ICD9Description. The query window in the center contains the following SQL query:

```
SELECT ICD9Code, ICD9Description, MDC
FROM CDWork.Dim.ICD9
GROUP BY ICD9Code, ICD9Description, MDC
ORDER BY ICD9Code ;
```

The Results pane at the bottom shows the output of the query. It displays a table with three columns: ICD9Code, ICD9Description, and MDC. The data is grouped by ICD9Code, ICD9Description, and MDC. A red star is placed next to the first row of the results, which has an ICD9Code of 971.0.

ICD9Code	ICD9Description	MDC
969.9	POISONING BY UNSPECIFIED PSYCHOTROPIC AG...	INJURY,POISONING,DRUG TOXICITY
970.0	POISONING BY ANALEPTICS	INJURY,POISONING,DRUG TOXICITY
970.1	POISONING BY OPIATE ANTAGONISTS	INJURY,POISONING,DRUG TOXICITY
970.8	POISONING BY OTHER SPECIFIED CENTRAL NERV...	INJURY,POISONING,DRUG TOXICITY
970.81	NULL	INJURY,POISONING,DRUG TOXICITY
970.89	NULL	INJURY,POISONING,DRUG TOXICITY
970.9	POISONING BY UNSPECIFIED CENTRAL NERVOUS...	INJURY,POISONING,DRUG TOXICITY
971.0	POISONING BY PARASYMPATHOMIMETICS (CHOLI...	INJURY,POISONING,DRUG TOXICITY
971.1	POISONING BY PARASYMPATHOLYTICS (ANTICHO...	INJURY,POISONING,DRUG TOXICITY
971.2	POISONING BY SYMPATHOMIMETICS (ADRENERGI...	INJURY,POISONING,DRUG TOXICITY
971.3	POISONING BY SYMPATHOLYTICS (ANTIADRENER...	INJURY,POISONING,DRUG TOXICITY
971.9	POISONING BY UNSPECIFIED DRUG PRIMARILY A...	INJURY,POISONING,DRUG TOXICITY
972.0	POISONING BY CARDIAC RHYTHM REGULATORS	INJURY,POISONING,DRUG TOXICITY
972.1	POISONING BY CARDIOTONIC GLYCOSIDES AND D...	INJURY,POISONING,DRUG TOXICITY
972.2	POISONING BY ANTILPEMIC AND ANTIARTERIOS...	INJURY,POISONING,DRUG TOXICITY
972.3	POISONING BY GANGLION-BLOCKING AGENTS	INJURY,POISONING,DRUG TOXICITY
972.4	POISONING BY CORONARY VASODILATORS	INJURY,POISONING,DRUG TOXICITY
972.5	POISONING BY OTHER VASODILATORS	INJURY,POISONING,DRUG TOXICITY
972.6	POISONING BY OTHER ANTIHYPERTENSIVE AGEN...	INJURY,POISONING,DRUG TOXICITY

SELECT ICD9Code, ICD9Description, MDC
FROM CDWork.Dim.ICD9
GROUP BY ICD9Code, ICD9Description, MDC
ORDER BY ICD9Code ;

Now there are 16,645 rows.
We can see the ICD code of
971.0 only has a single row.

We can ask to see how many
times each value repeats.

COUNT function

- The COUNT function can be added to the SELECT phrase in the query to create a new column.
- Because the new column is a sum (an aggregate measure), the programmer is obligated to include the GROUP BY clause in the query for all non-aggregate columns.
- The AS allows for the new column to be given a name by the programmer.

```
SELECT column1, column2, column3, COUNT (column1) AS Freq  
FROM Database.Schema.View  
GROUP BY column1, column2 , column3  
ORDER BY column1 ;
```

COUNT

The screenshot shows the Microsoft SQL Server Management Studio interface. The title bar indicates the connection is to 'SQLQuery1.sql - vhaadwa01.CDWWork (VHA12\VHAHNGonsoM (152))*'. The Object Explorer on the left shows the 'Dim.ICD9' database with various tables and columns. The query window on the right contains the following SQL query:

```
SELECT ICD9Code , ICD9Description , MDC , COUNT (ICD9Code) AS Freq
FROM CDWWork.Dim.ICD9
GROUP BY ICD9Code , ICD9Description , MDC
ORDER BY ICD9Code ;
```

The Results pane at the bottom displays the output of the query as a table with four columns: ICD9Code, ICD9Description, MDC, and Freq. The table shows multiple rows of data, with the frequency (Freq) column consistently showing the value 130 for each row. A red star is placed next to the row where ICD9Code is 971.0.

ICD9Code	ICD9Description	MDC	Freq
1... 969.8	POISONING BY OTHER SPECIFIED PSYCHOTROPIC...	INJURY,POISONING,DRUG TOXICITY	130
1... 969.9	POISONING BY UNSPECIFIED PSYCHOTROPIC AG...	INJURY,POISONING,DRUG TOXICITY	130
1... 970.0	POISONING BY ANALEPTICS	INJURY,POISONING,DRUG TOXICITY	130
1... 970.1	POISONING BY OPIATE ANTAGONISTS	INJURY,POISONING,DRUG TOXICITY	130
1... 970.8	POISONING BY OTHER SPECIFIED CENTRAL NERV...	INJURY,POISONING,DRUG TOXICITY	130
1... 970.81	NULL	INJURY,POISONING,DRUG TOXICITY	130
1... 970.89	NULL	INJURY,POISONING,DRUG TOXICITY	130
1... 970.9	POISONING BY UNSPECIFIED CENTRAL NERVOUS...	INJURY,POISONING,DRUG TOXICITY	130
1... 971.0	POISONING BY PARASYMPATHOMIMETICS (CHOLI...	INJURY,POISONING,DRUG TOXICITY	130
1... 971.1	POISONING BY PARASYMPATHOLYTICS (ANTICHO...	INJURY,POISONING,DRUG TOXICITY	130
1... 971.2	POISONING BY SYMPATHOMIMETICS (ADRENERGI...	INJURY,POISONING,DRUG TOXICITY	130
1... 971.3	POISONING BY SYMPATHOLYTICS (ANTIADRENER...	INJURY,POISONING,DRUG TOXICITY	130
1... 971.9	POISONING BY UNSPECIFIED DRUG PRIMARILY A...	INJURY,POISONING,DRUG TOXICITY	130
1... 972.0	POISONING BY CARDIAC RHYTHM REGULATORS	INJURY,POISONING,DRUG TOXICITY	130
1... 972.1	POISONING BY CARDIOTONIC GLYCOSIDES AND D...	INJURY,POISONING,DRUG TOXICITY	130
1... 972.2	POISONING BY ANTIPEMIC AND ANTIARTERIOS...	INJURY,POISONING,DRUG TOXICITY	130
1... 972.3	POISONING BY GANGLION-BLOCKING AGENTS	INJURY,POISONING,DRUG TOXICITY	130
1... 972.4	POISONING BY CORONARY VASODILATORS	INJURY,POISONING,DRUG TOXICITY	130
1... 972.5	POISONING BY OTHER VASODILATORS	INJURY,POISONING,DRUG TOXICITY	130

The status bar at the bottom indicates 'Query executed successfully.'

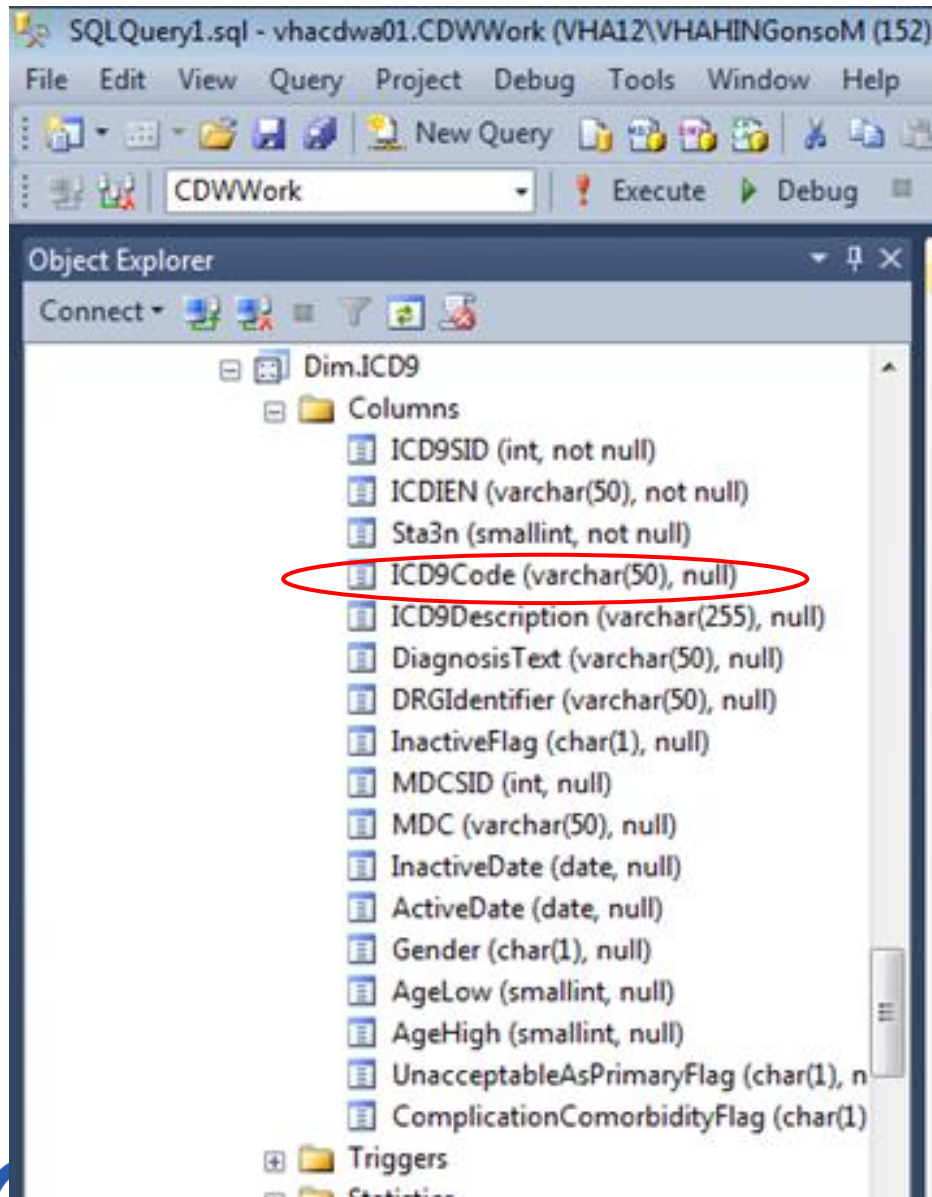
```
SELECT ICD9Code, ICD9Description, MDC, COUNT (ICD9Code) AS Freq
FROM CDWWork.Dim.ICD9
GROUP BY ICD9Code, ICD9Description, MDC
ORDER BY ICD9Code ;
```



Now we have the equivalent of a frequency table. The ICD code of 971.0 appears 130 times in the view Dim.ICD9, once for each VistA system.

WHERE statements

- The command WHERE allows a user to limit their search to include only a chosen subset of the data.
- Let's select the ICD9 code (309.81), indicating post-traumatic stress disorder.
- Although 309.81 looks like a number, you will need to know whether ICD9Code is a character or a numeric entry in CDW before you can write a WHERE statement.



VARCHAR = a string of variable length

(50) = the maximum number of characters

Other examples seen here:

Char = string of fixed length

Int = integer

Smallint = integer

Date = a date

LIKE vs. =

- LIKE is typically used for string variables
- = is typically used for numeric variables
- So, because ICD9Code is a string, we will use LIKE in the WHERE phrase
- Always use single quotes around your characters
- Numeric values would not require quotes

```
SELECT column1, column2, column3  
FROM Database.Schema.View  
WHERE column1 LIKE '-----' ; --note the value will go in the quotes
```

SQLQuery1.sql - VHACDWa01.vha.med.va.gov.CDWWork (VHA12\VHAHingonsoM (660))* - Microsoft SQL Server Management Studio

File Edit View Project Debug Tools Window Help

CDWWork

Object Explorer

Connect

- Dim.IBRevenueCode
- Dim.IBTypeOfService
- Dim.ICD
- Dim.ICD10
- Dim.ICD10Procedure
- Dim.ICD9
 - Columns
 - ICD9SID (int, not null)
 - ICDIEN (varchar(50), not null)
 - Sta3n (smallint, not null)
 - ICD9Code (varchar(50), null)
 - ICD9Description (varchar(255), null)
 - DiagnosisText (varchar(50), null)
 - DRGIdentifier (varchar(50), null)
 - InactiveFlag (char(1), null)
 - MDCSID (int, null)
 - MDC (varchar(50), null)
 - InactiveDate (date, null)
 - ActiveDate (date, null)
 - Gender (char(1), null)
 - AgeLow (smallint, null)
 - AgeHigh (smallint, null)
 - UnacceptableAsPrimaryFlag (char(1), null)
 - ComplicationComorbidityFlag (char(1), null)
 - Triggers
 - Statistics
- Dim.ICD9Procedure
- Dim.ICDDescription
- Dim.ICDPProcedure
- Dim.IdentificationFormat
- Dim.ImmunizationName
- Dim.Institution
- Dim.InterventionRecommendation
- Dim.InterventionType
- Dim.JRTDeficiency
- Dim.JRTDeficiencyCategory
- Dim.JRTStatus
- Dim.IVAdditiveIngredient
- Dim.IVAdditiveIngredientElectrolytes
- Dim.IVAdditiveIngredientSynonym
- Dim.IVRoom
- Dim.IVSolutionIngredient
- Dim.IVSolutionIngredientElectrolytes
- Dim.IVSolutionIngredientSynonym
- Dim.LabChemTest
- Dim.LabChemTestPanelList
- Dim.LabChemTestUrgency

SQLQuery1.sql - V...AHINGonsoM (660))*

```
SELECT ICD9Code, ICD9Description, MDC
FROM CDWWork.Dim.ICD9
WHERE ICD9Code LIKE '309.81' ;
```

100 %

Results Messages

	ICD9Code	ICD9Description	MDC
1	309.81	POSTTRAUMATIC STRESS DISORDER	MENTAL DISEASES & DISORDERS
2	309.81	POSTTRAUMATIC STRESS DISORDER	MENTAL DISEASES & DISORDERS
3	309.81	POSTTRAUMATIC STRESS DISORDER	MENTAL DISEASES & DISORDERS
4	309.81	POSTTRAUMATIC STRESS DISORDER	MENTAL DISEASES & DISORDERS
5	309.81	POSTTRAUMATIC STRESS DISORDER	MENTAL DISEASES & DISORDERS
6	309.81	POSTTRAUMATIC STRESS DISORDER	MENTAL DISEASES & DISORDERS
7	309.81	POSTTRAUMATIC STRESS DISORDER	MENTAL DISEASES & DISORDERS
8	309.81	POSTTRAUMATIC STRESS DISORDER	MENTAL DISEASES & DISORDERS
9	309.81	POSTTRAUMATIC STRESS DISORDER	MENTAL DISEASES & DISORDERS
10	309.81	POSTTRAUMATIC STRESS DISORDER	MENTAL DISEASES & DISORDERS
11	309.81	POSTTRAUMATIC STRESS DISORDER	MENTAL DISEASES & DISORDERS
12	309.81	POSTTRAUMATIC STRESS DISORDER	MENTAL DISEASES & DISORDERS
13	309.81	POSTTRAUMATIC STRESS DISORDER	MENTAL DISEASES & DISORDERS
14	309.81	POSTTRAUMATIC STRESS DISORDER	MENTAL DISEASES & DISORDERS
15	309.81	POSTTRAUMATIC STRESS DISORDER	MENTAL DISEASES & DISORDERS
16	309.81	POSTTRAUMATIC STRESS DISORDER	MENTAL DISEASES & DISORDERS
17	309.81	POSTTRAUMATIC STRESS DISORDER	MENTAL DISEASES & DISORDERS
18	309.81	POSTTRAUMATIC STRESS DISORDER	MENTAL DISEASES & DISORDERS
19	309.81	POSTTRAUMATIC STRESS DISORDER	MENTAL DISEASES & DISORDERS

Query executed successfully.

SELECT ICD9Code, ICD9Description, MDC
FROM CDWWork.Dim.ICD9
WHERE ICD9Code LIKE '309.81' ;

Note: In an actual research investigation, one might find it necessary to address varied entries with a wildcard such as '309%81' or similar logic.

Poll #2: About You

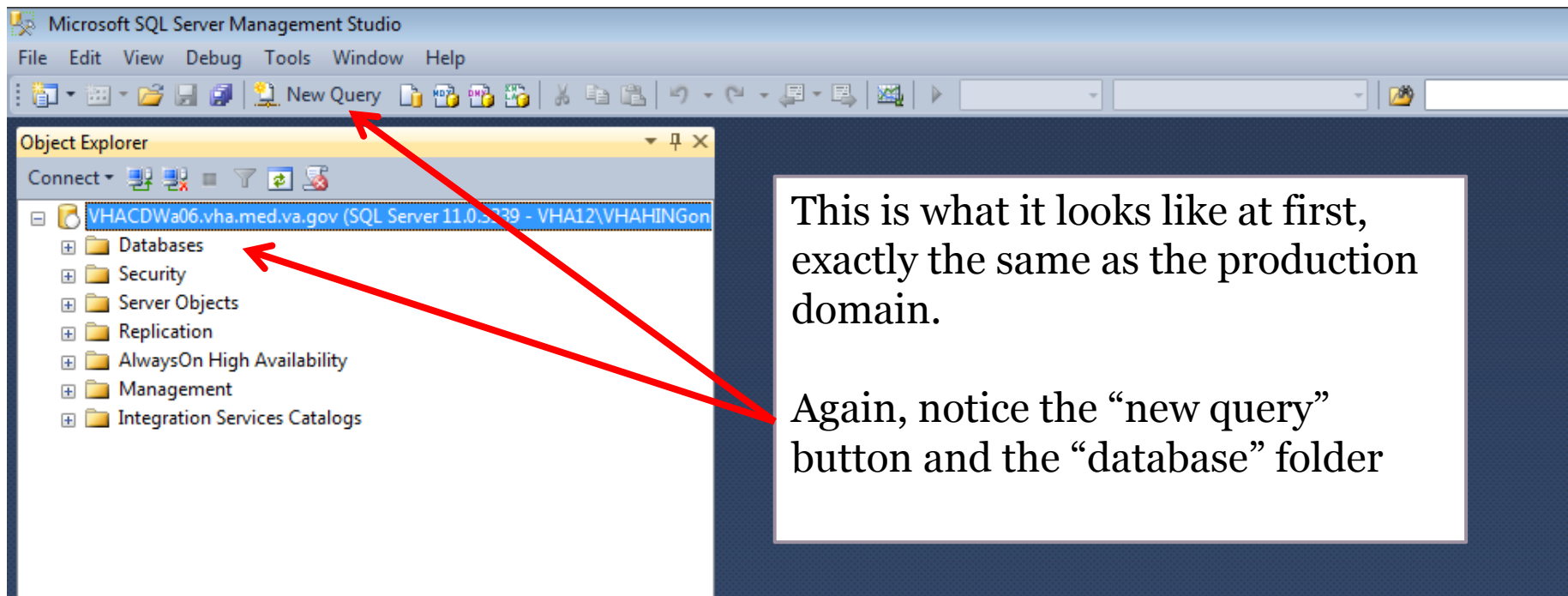
Which of the following best describes your role in the VA?

- ☐ Research Investigator / PI
- ☐ Data Manager / Analyst
- ☐ Project Coordinator
- ☐ Operations
- ☐ Other

In the Raw Domains

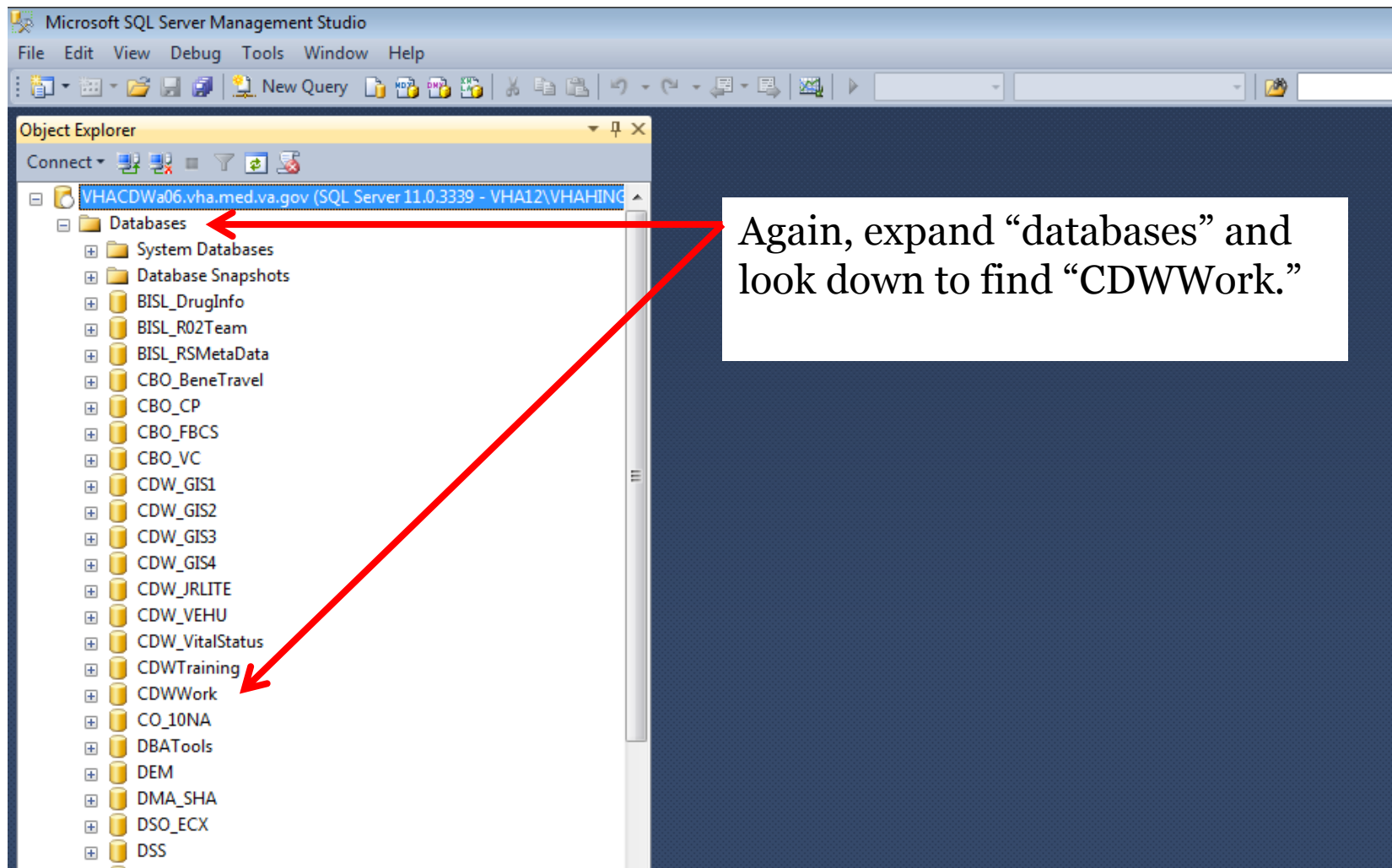
We will repeat the process with Dim.ICD9_Diagnosis_80

SQL Server Management Studio

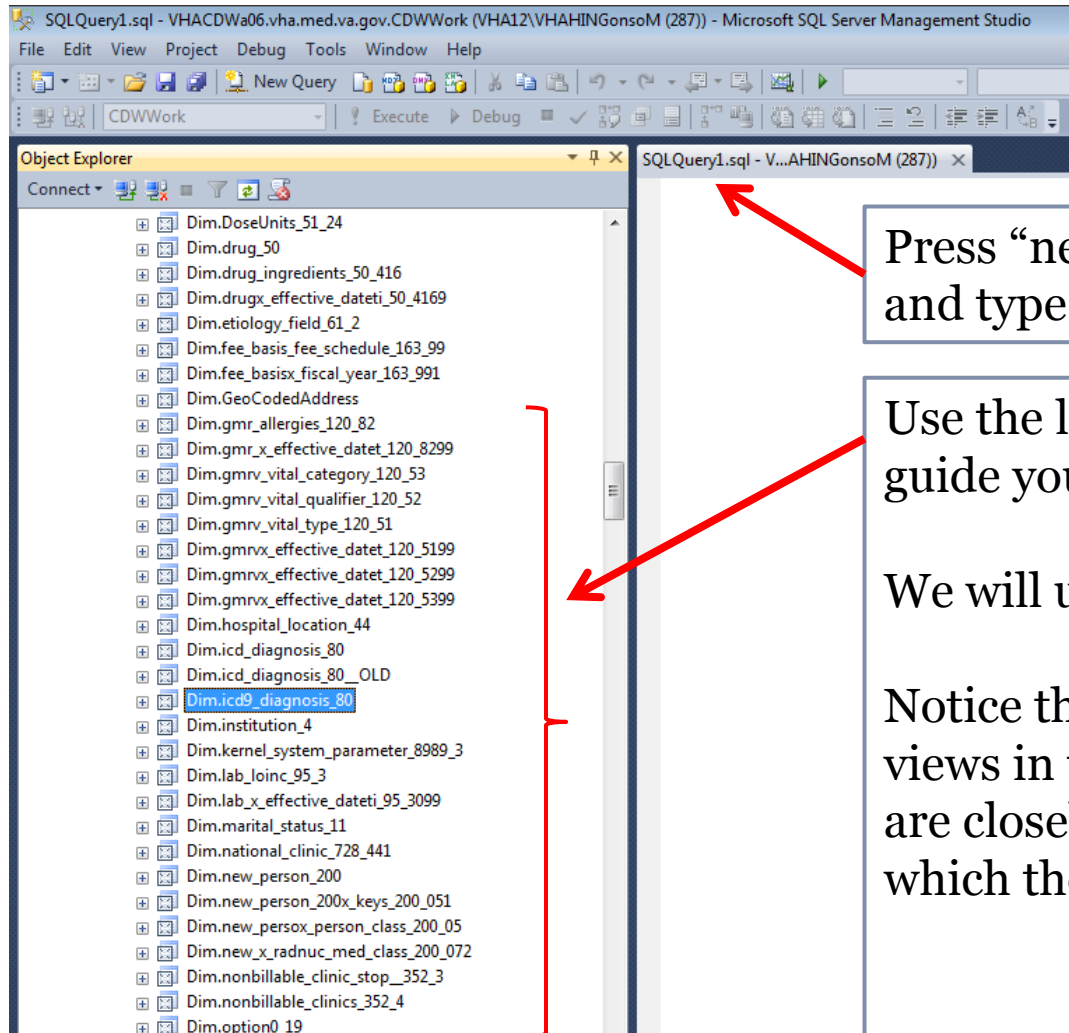


The only difference from the previous example is that I am now working off of a different server that holds the raw domains.

Expand Databases and Scroll to CDWork



New Query



Press “new query” to open this window, and type your SQL query here.

Use the list of *views* in CDWWork to help guide your effort.

We will use Dim.icd9_diagnosis_80

Notice the difference in the naming styles; views in the raw domains have names that are closely related to the VistA’s files from which they originate.

The screenshot shows the Microsoft SQL Server Management Studio interface. The Object Explorer on the left lists various database objects. A context menu is open over the 'Dim.icd9_diagnosis_80' view, with the 'Select Top 1000 Rows' option highlighted. A red '1st' is placed next to this menu. The main query window shows a SQL query that has been automatically generated based on the selected view and the 'Select Top 1000 Rows' command. A red '2nd' is placed next to the query text. The query results are displayed in a table at the bottom, with a red '3rd' placed next to the table header. The table contains 18 rows of data, including columns for station_no, row_id, code_number, identifier, and diagnosis.

1st

2nd

3rd

```

/***** Script for SelectTopNRows command from SSMS *****/
SELECT TOP 1000 [station_no]
, [row_id]
, [code_number]
, [identifier]
, [diagnosis]
, [major_diagnostic_category]
, [mdc13]
, [mdc24]
, [mdc25]
, [icd_expanded]
, [sex]
, [description]
, [age_low]
, [age_high]
, [activation_date]
, [drga]
, [drgb]
, [drgc]
, [drgd]
, [drge]
, [drgf]
, [complicationcomorbidity]
, [inactive_flag]
, [unacceptable_as_principal_dx]
, [inactive_date]
FROM [CDWWork].[dbo].[Dim.icd9_diagnosis_80]

```

station_no	row_id	code_number	identifier	diagnosis
116	1.000000	100.81	CQ	LEPTOSPIRAL MENINGIT
116	2.000000	100.89	CQ	LEPTOSPIRAL INFECT NI
116	3.000000	100.9	NULL	LEPTOSPIROSIS NOS
116	4.000000	102.1	NULL	MULTIPLE PAPILLOMAT
116	5.000000	102.2	NULL	EARLY SKIN YAWS NEC
116	6.000000	102.3	NULL	HYPERKERATOSIS OF Y
116	7.000000	102.4	NULL	GUMMATA AND ULCERS
116	8.000000	102.5	Y	GANGOSA
116	9.000000	102.6	NULL	YAWS OF BONE JOINT
116	10.000000	102.7	NULL	YAWS MANIFESTATIONS
116	11.000000	102.8	NULL	LATENT YAWS
116	12.000000	102.9	NULL	YAWS NOS
116	13.000000	103.1	NULL	PINTA INTERMED LESIO
116	14.000000	103.2	NULL	PINTA LATE LESIONS
116	15.000000	103.3	NULL	PINTA MIXED LESIONS
116	16.000000	103.9	NULL	PINTA NOS
116	17.000000	104.8	NULL	SPIROCHETAL INFECT N
116	18.000000	104.9	NULL	SPIROCHETAL INFECT N

Again,

1st Right click the view name for this drop-down menu and click “select top 1000 rows”

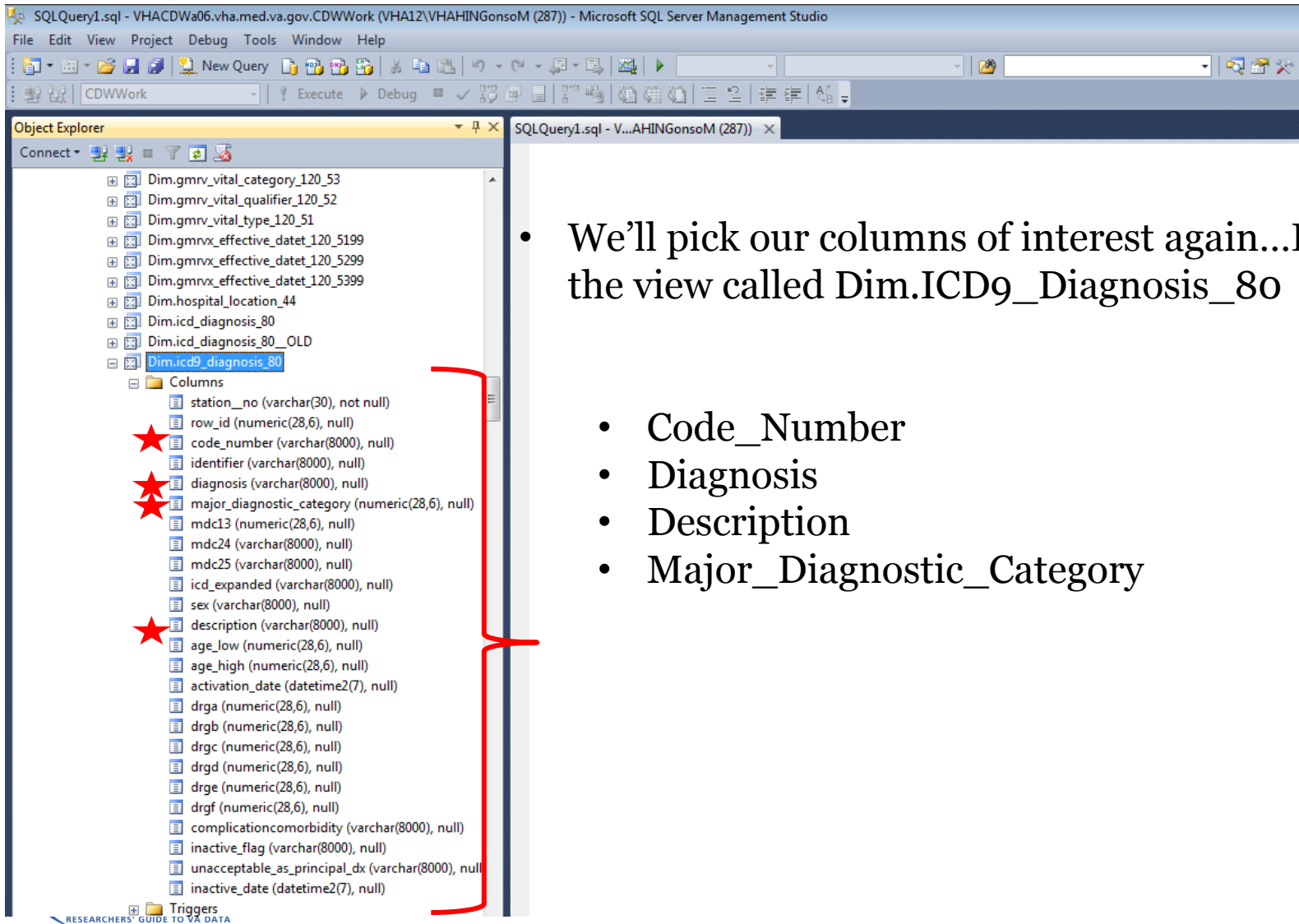
2nd, the query in the circle is automatically written

3rd the results automatically appear in the bottom window

SELECT and FROM

- SELECT your columns (a.k.a, variables)
 - `SELECT column1, column2, column3`
- FROM your *view* of choice
 - `From database.schema.view_name`
- Don't forget to end with a semi-colon

SELECT & FROM



The screenshot shows the Microsoft SQL Server Management Studio interface. The Object Explorer on the left displays the database structure, with the view `Dim.icd9_diagnosis_80` selected. The Columns list for this view is expanded, showing various columns. A red bracket highlights the following columns, which are also listed in the adjacent text:

- `code_number` (varchar(8000), null)
- `diagnosis` (varchar(8000), null)
- `major_diagnostic_category` (numeric(28,6), null)
- `description` (varchar(8000), null)

The other columns listed in the view are:

- `station_no` (varchar(30), not null)
- `row_id` (numeric(28,6), null)
- `identifier` (varchar(8000), null)
- `mdc13` (numeric(28,6), null)
- `mdc24` (varchar(8000), null)
- `mdc25` (varchar(8000), null)
- `icd_expanded` (varchar(8000), null)
- `sex` (varchar(8000), null)
- `age_low` (numeric(28,6), null)
- `age_high` (numeric(28,6), null)
- `activation_date` (datetime2(7), null)
- `drga` (numeric(28,6), null)
- `drgb` (numeric(28,6), null)
- `drgc` (numeric(28,6), null)
- `drgd` (numeric(28,6), null)
- `drge` (numeric(28,6), null)
- `drgf` (numeric(28,6), null)
- `complicationcomorbidity` (varchar(8000), null)
- `inactive_flag` (varchar(8000), null)
- `unacceptable_as_principal_dx` (varchar(8000), null)
- `inactive_date` (datetime2(7), null)

- We'll pick our columns of interest again...From the view called `Dim.ICD9_Diagnosis_80`

- `Code_Number`
- `Diagnosis`
- `Description`
- `Major_Diagnostic_Category`

SQLQuery1.sql - VHACDWA06.vha.med.va.gov.CDWWork (VHA12\VHAHingonSOM (287)) Executing... - Microsoft SQL Server Management Studio

File Edit View Query Project Debug Tools Window Help

Object Explorer

Connect

- Dim.gmr_vital_category_120_53
- Dim.gmr_vital_qualifier_120_52
- Dim.gmr_vital_type_120_51
- Dim.gmr_v_effective_datet_120_5199
- Dim.gmr_v_effective_datet_120_5299
- Dim.gmr_v_effective_datet_120_5399
- Dim.hospital_location_44
- Dim.icd_diagnosis_80
- Dim.icd_diagnosis_80_OLD
- Dim.icd9_diagnosis_80
- Columns
 - station_no (varchar(30), not null)
 - row_id (numeric(28,6), null)
 - code_number (varchar(8000), null)
 - identifier (varchar(8000), null)
 - diagnosis (varchar(8000), null)
 - major_diagnostic_category (numeric(28,6), null)
 - mdc13 (numeric(28,6), null)
 - mdc24 (varchar(8000), null)
 - mdc25 (varchar(8000), null)
 - icd_expanded (varchar(8000), null)
 - sex (varchar(8000), null)
 - description (varchar(8000), null)
 - age_low (numeric(28,6), null)
 - age_high (numeric(28,6), null)
 - activation_date (datetime2(7), null)
 - drga (numeric(28,6), null)
 - drgb (numeric(28,6), null)
 - drgc (numeric(28,6), null)
 - drgd (numeric(28,6), null)
 - drge (numeric(28,6), null)
 - drgf (numeric(28,6), null)
 - complicationcomorbidity (varchar(8000), null)
 - inactive_flag (varchar(8000), null)
 - unacceptable_as_principal_dx (varchar(8000), null)
 - inactive_date (datetime2(7), null)
- Triggers
- Indexes
- Statistics
- Dim.institution_4
- Dim.kernel_system_parameter_8989_3
- Dim.lab_loinc_95_3
- Dim.lab_x_effective_dateti_95_3099
- Dim.marital_status_11
- Dim.national_clinic_728_441
- Dim.new_person_200
- Dim.new_person_200x_keyv_200_051

SQLQuery1.sql - VHA... (287)) Executing... X

```
SELECT Code_Number, Diagnosis, Description, Major_Diagnostic_Category
FROM CDWWork.Dim.ICD9_Diagnosis_80 ;
```

100 %

Results Messages

	Code_Number	Diagnosis	Description	Major_Diagnostic_Category
1	100.81	LEPTOSPIRAL MENINGITIS	LEPTOSPIRAL MENINGITIS (ASEPTIC)	1.000000
2	100.89	LEPTOSPIRAL INFECT NEC	OTHER SPECIFIED LEPTOSPIRAL INFECTIONS	1.000000
3	100.9	LEPTOSPIROSIS NOS	LEPTOSPIROSIS, UNSPECIFIED	18.000000
4	102.1	MULTIPLE PAPILLOMATA	MULTIPLE PAPILLOMATA DUE TO YAWS AND WET CRAB YAWS	9.000000
5	102.2	EARLY SKIN YAWS NEC	OTHER EARLY SKIN LESIONS OF YAWS	9.000000
6	102.3	HYPERKERATOSIS OF YAWS	HYPERKERATOSIS DUE TO YAWS	9.000000
7	102.4	GUMMATA AND ULCERS, YAWS	GUMMATA AND ULCERS DUE TO YAWS	9.000000
8	102.5	GANGOSA	GANGOSA	3.000000
9	102.6	YAWS OF BONE JOINT	BONE AND JOINT LESIONS DUE TO YAWS	9.000000
10	102.7	YAWS MANIFESTATIONS NEC	OTHER MANIFESTATIONS OF YAWS	18.000000
11	102.8	LATENT YAWS	LATENT YAWS	18.000000
12	102.9	YAWS NOS	YAWS, UNSPECIFIED	18.000000
13	103.1	PINTA INTERMED LESIONS	INTERMEDIATE LESIONS OF PINTA	9.000000
14	103.2	PINTA LATE LESIONS	LATE LESIONS OF PINTA	18.000000
15	103.3	PINTA MIXED LESIONS	MIXED LESIONS OF PINTA	9.000000
16	103.9	PINTA NOS	PINTA, UNSPECIFIED	18.000000
17	104.8	SPIROCHETAL INFECT NEC	OTHER SPECIFIED SPIROCHETAL INFECTIONS	18.000000
18	104.9	SPIROCHETAL INFECT NOS	SPIROCHETAL INFECTION, UNSPECIFIED	18.000000
19	110.1	DERMATOPHYTOSIS OF NAIL	DERMATOPHYTOSIS OF NAIL	9.000000

11,243,873 results or rows appear here in a new window.

ORDER BY

- The ORDER BY command will allow us to request the results in order by our chosen column(s).
- Remember that ORDER BY must be the last command in any query.

```
SELECT column1, column2, column3  
FROM Database.Schema.View  
ORDER BY column1 ;
```

ORDER BY

SQLQuery1.sql - VHACDwa06.vha.med.va.gov.CDWWork (VHA12\VHAHINonsoM (287)) Executing... - Microsoft SQL Server Management Studio

File Edit View Query Project Debug Tools Window Help

Object Explorer

Connect

- Dim.gmrv_vital_category_120_53
- Dim.gmrv_vital_qualifier_120_52
- Dim.gmrv_vital_type_120_51
- Dim.gmrvx_effective_datet_120_5199
- Dim.gmrvx_effective_datet_120_5299
- Dim.gmrvx_effective_datet_120_5399
- Dim.hospital_location_44
- Dim.icd_diagnosis_80
- Dim.icd_diagnosis_80_OLD
- Dim.icd9_diagnosis_80
- Columns
 - station_no (varchar(30), not null)
 - row_id (numeric(28,6), null)
 - code_number (varchar(8000), null)
 - identifier (varchar(8000), null)
 - diagnosis (varchar(8000), null)
 - major_diagnostic_category (numeric(28,6), null)
 - mdc13 (numeric(28,6), null)
 - mdc24 (varchar(8000), null)
 - mdc25 (varchar(8000), null)
 - icd_expanded (varchar(8000), null)
 - sex (varchar(8000), null)
 - description (varchar(8000), null)
 - age_low (numeric(28,6), null)
 - age_high (numeric(28,6), null)
 - activation_date (datetime2(7), null)
 - drga (numeric(28,6), null)
 - drgb (numeric(28,6), null)
 - drgc (numeric(28,6), null)
 - drgd (numeric(28,6), null)
 - drge (numeric(28,6), null)
 - drgf (numeric(28,6), null)
 - complicationcomorbidity (varchar(8000), null)
 - inactive_flag (varchar(8000), null)
 - unacceptable_as_principal_dx (varchar(8000), null)
 - inactive_date (datetime2(7), null)
- Triggers
- Indexes
- Statistics
- Dim.institution_4
- Dim.kernel_system_parameter_8989_3
- Dim.lab_loinc_95_3
- Dim.lab_x_effective_datet_95_3099
- Dim.marital_status_11
- Dim.national_clinic_728_441
- Dim.new_person_200
- Dim.new_person_200x_keys_200_051

SQLQuery1.sql - VHA... (287)) Executing... *

```

SELECT Code_Number, Diagnosis, Description, Major_Diagnostic_Category
FROM CDWWork.Dim.ICD9_Diagnosis_80
ORDER BY Code_Number ;

```

Results

	Code_Number	Diagnosis	Description	Major_Diagnostic_Category
7...	482.30	NULL	1	NULL
7...	482.30	NULL	1	NULL
7...	482.30	NULL	1	NULL
7...	482.30	NULL	1	NULL
7...	482.30	NULL	1	NULL
7...	482.30	NULL	1	NULL
7...	482.30	NULL	1	NULL
7...	482.30	NULL	1	NULL
7...	482.30	NULL	1	NULL
7...	482.30	NULL	1	NULL
7...	482.30	NULL	1	NULL
7...	482.30	NULL	1	NULL
7...	482.30	NULL	1	NULL
7...	482.30	NULL	1	NULL
7...	482.30	NULL	1	NULL
7...	482.30	NULL	1	NULL
7...	482.30	NULL	1	NULL
7...	482.30	PNEUM. UNSPEC. STREPTOCOCCUS	PNEUMONIA DUE TO STREPTOCOCCUS UNSPECIFIED	4.000000
7...	482.30	NULL	1	NULL
7...	482.30	NULL	1	NULL

SELECT Code_Number
, Diagnosis
, Description
, Major_Diagnostic_Category
FROM CDWWork.Dim.ICD9_Diagnosis_80
ORDER BY Code_Number ;



We see the same repetition in the raw domain.

GROUP BY

- Again, we will collapse the repeated entries using the GROUP BY command.
- Remember:
 - ORDER BY must be the last command
 - Every column that is not an aggregate measure must be included in the GROUP BY clause

```
SELECT column1, column2, column3  
FROM Database.Schema.View  
GROUP BY column1, column2 , column3  
ORDER BY column1 ;
```

GROUP BY

SQLQuery1.sql - VHACDWa06.vha.med.va.gov.CDWWork (VHA12\VHAHINGonsoM (287))* - Microsoft SQL Server Management Studio

File Edit View Query Project Debug Tools Window Help

Connect * CDWWork

Object Explorer

- Dim.gmrvt_vital_category_120_53
- Dim.gmrvt_vital_qualifier_120_52
- Dim.gmrvt_vital_type_120_51
- Dim.gmrvt_effective_datet_120_5199
- Dim.gmrvt_effective_datet_120_5299
- Dim.gmrvt_effective_datet_120_5399
- Dim.hospital_location_44
- Dim.icd_diagnosis_80
- Dim.icd_diagnosis_80_OLD
- Dim.icd9_diagnosis_80
- Columns
 - station_no (varchar(30), not null)
 - row_id (numeric(28,6), null)
 - code_number (varchar(8000), null)
 - identifier (varchar(8000), null)
 - diagnosis (varchar(8000), null)
 - major_diagnostic_category (numeric(28,6), null)
 - mdc13 (numeric(28,6), null)
 - mdc24 (varchar(8000), null)
 - mdc25 (varchar(8000), null)
 - icd_expanded (varchar(8000), null)
 - sex (varchar(8000), null)
 - description (varchar(8000), null)
 - age_low (numeric(28,6), null)
 - age_high (numeric(28,6), null)
 - activation_date (datetime2(7), null)
 - drga (numeric(28,6), null)
 - drgb (numeric(28,6), null)
 - drgc (numeric(28,6), null)
 - drgd (numeric(28,6), null)
 - drge (numeric(28,6), null)
 - drgf (numeric(28,6), null)
 - complicationcomorbidity (varchar(8000), null)
 - inactive_flag (varchar(8000), null)
 - unacceptable_as_principal_dx (varchar(8000), null)
 - inactive_date (datetime2(7), null)
- Triggers
- Indexes
- Statistics
- Dim.institution_4
- Dim.kernel_system_parameter_8989_3
- Dim.lab_loinc_95_3
- Dim.lab_x_effective_datet_95_3099
- Dim.marital_status_11
- Dim.national_clinic_728_441
- Dim.new_person_200
- Dim.new_person_200x_keys_200_051

SQLQuery1.sql - V...AHINGonsoM (287))*

```
SELECT Code_Number, Diagnosis, Description, Major_Diagnostic_Category
FROM CDWWork.Dim.ICD9_Diagnosis_80
GROUP BY Code_Number, Diagnosis, Description, Major_Diagnostic_Category
ORDER BY Code_Number ;
```

100 %

Results Messages

	Code_Number	Diagnosis	Description	Major_Diagnostic_Category
1	001.0	CHOLERA D/T VIB CHOLERA	CHOLERA DUE TO VIBRIO CHOLERA	6.000000
2	001.0	NULL	1	NULL
3	001.1	NULL	1	NULL
4	001.1	CHOLERA D/T VIB EL TOR	CHOLERA DUE TO VIBRIO CHOLERA EL TOR	6.000000
5	001.9	NULL	1	NULL
6	001.9	CHOLERA NOS	CHOLERA, UNSPECIFIED	6.000000
7	002.0	TYPHOID FEVER	TYPHOID FEVER	18.000000
8	002.0	NULL	1	NULL
9	002.1	PARATYPHOID FEVER A	PARATYPHOID FEVER A	18.000000
10	002.1	NULL	1	NULL
11	002.2	PARATYPHOID FEVER B	PARATYPHOID FEVER B	18.000000
12	002.2	NULL	1	NULL
13	002.3	NULL	1	NULL
14	002.3	PARATYPHOID FEVER C	PARATYPHOID FEVER C	18.000000
15	002.9	NULL	1	NULL
16	002.9	PARATYPHOID FEVER NOS	PARATYPHOID FEVER, UNSPECIFIED	18.000000
17	003.0	SALMONELLA ENTERITIS	SALMONELLA GASTROENTERITIS	6.000000
18	003.0	NULL	1	NULL
19	003.1	SALMONELLA SEPTICEMIA	SALMONELLA SEPTICEMIA	18.000000

Query executed successfully.

SELECT Code_Number , Diagnosis
 , Description , Major_Diagnostic_Code
 FROM CDWWork.Dim.ICD9_Diagnosis_80
 GROUP BY Code_Number , Diagnosis
 , Description , Major_Diagnostic_Code
 ORDER BY Code_Number ;

Now there are 101,438 rows.

COUNT function

- Again, we can use the COUNT function to create a frequency table.
- Using AS we can assign the name “freq” to the count column.
- Remember to use the GROUP BY clause for all non-aggregate columns.

```
SELECT column1, column2, column3, COUNT (column1) AS Freq  
FROM Database.Schema.View  
GROUP BY column1, column2 , column3  
ORDER BY column1 ;
```



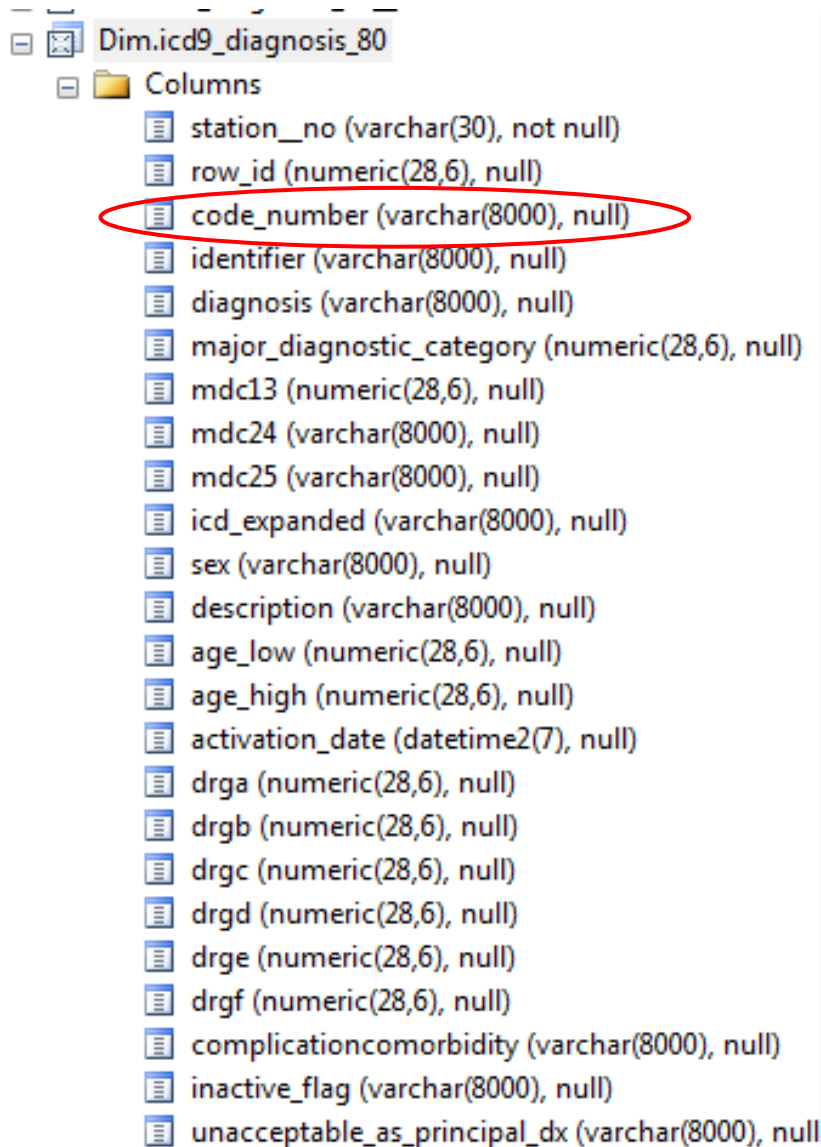
```
SELECT Code_Number , Diagnosis
       , Description , Major_Diagnostic_Code
       , COUNT (Code_Number) AS Freq
FROM CDWWork.Dim.ICD9_Diagnosis_80
GROUP BY Code_Number , Diagnosis
       , Description , Major_Diagnostic_Code
ORDER BY Code_Number ;
```



Now we have the equivalent of a frequency table.

WHERE statements

- We'll use the WHERE again to choose a subset of the data.
- We'll select the Code_Number (309.81) for PTSD again.
- Again, we'll check to see whether Code_Number is a character or a numeric before writing our WHERE statement.



Dim.icd9_diagnosis_80	
Columns	
station_no	(varchar(30), not null)
row_id	(numeric(28,6), null)
code_number	(varchar(8000), null)
identifier	(varchar(8000), null)
diagnosis	(varchar(8000), null)
major_diagnostic_category	(numeric(28,6), null)
mdc13	(numeric(28,6), null)
mdc24	(varchar(8000), null)
mdc25	(varchar(8000), null)
icd_expanded	(varchar(8000), null)
sex	(varchar(8000), null)
description	(varchar(8000), null)
age_low	(numeric(28,6), null)
age_high	(numeric(28,6), null)
activation_date	(datetime2(7), null)
drga	(numeric(28,6), null)
drgb	(numeric(28,6), null)
drgc	(numeric(28,6), null)
drgd	(numeric(28,6), null)
drge	(numeric(28,6), null)
drgf	(numeric(28,6), null)
complicationcomorbidity	(varchar(8000), null)
inactive_flag	(varchar(8000), null)
unacceptable_as_principal_dx	(varchar(8000), null)

VARCHAR = a string of variable length

(8000) = the maximum number of characters

Therefore, we will tend use the LIKE function rather than an equal sign in our WHERE clause.

Pulling it all together in one query

- SELECT, COUNT, AS, WHERE, LIKE, GROUP BY and ORDER BY
- Remember the rules!

```
SELECT column1, column2, column3, COUNT (columnX) AS NewName
FROM Database.Schema.View
WHERE column1 LIKE '-----'  --note the value will go in the quotes
GROUP BY column1, column2...  -- include all non-aggregate columns
ORDER BY column1...           -- name column(s) to order by
;                               -- don't forget to end with a ;
```

SQLQuery1.sql - VHACDWa06.vha.med.va.gov.CDWWork (VHA12\VHAHINGoSoM (287))* - Microsoft SQL Server Management Studio

File Edit View Query Project Debug Tools Window Help

CDWWork

Execute Debug

Object Explorer

Connect

- Dim.gmr_vital_category_120_53
- Dim.gmr_vital_qualifier_120_52
- Dim.gmr_vital_type_120_51
- Dim.gmr_vx_effective_datet_120_5199
- Dim.gmr_vx_effective_datet_120_5299
- Dim.gmr_vx_effective_datet_120_5399
- Dim.hospital_location_44
- Dim.icd_diagnosis_80
- Dim.icd_diagnosis_80_OLD
- Dim.icd9_diagnosis_80
 - Columns
 - station_no (varchar(30), not null)
 - row_id (numeric(28,6), null)
 - code_number (varchar(8000), null)
 - identifier (varchar(8000), null)
 - diagnosis (varchar(8000), null)
 - major_diagnostic_category (numeric(28,6), null)
 - mdc13 (numeric(28,6), null)
 - mdc24 (varchar(8000), null)
 - mdc25 (varchar(8000), null)
 - icd_expanded (varchar(8000), null)
 - sex (varchar(8000), null)
 - description (varchar(8000), null)
 - age_low (numeric(28,6), null)
 - age_high (numeric(28,6), null)
 - activation_date (datetime2(7), null)
 - drga (numeric(28,6), null)
 - drgb (numeric(28,6), null)
 - drgc (numeric(28,6), null)
 - drgd (numeric(28,6), null)
 - drge (numeric(28,6), null)
 - drgf (numeric(28,6), null)
 - complicationcomorbidity (varchar(8000), null)
 - inactive_flag (varchar(8000), null)
 - unacceptable_as_principal_dx (varchar(8000), null)

SQLQuery1.sql - V...AHINGoSoM (287))*

```

SELECT Code_Number, Diagnosis, Description, Major_Diagnostic_Category, COUNT (Code_Number) AS Freq
FROM CDWWork.Dim.ICD9_Diagnosis_80
WHERE Code_Number LIKE '309.81'
GROUP BY Code_Number, Diagnosis, Description, Major_Diagnostic_Category
ORDER BY Code_Number ;

```

```

SELECT Code_Number, Diagnosis
      , Description, Major_Diagnostic_Code
      , COUNT (Code_Number) AS Freq
FROM CDWWork.Dim.ICD9_Diagnosis_80
WHERE Code_Number LIKE '309.81'
GROUP BY Code_Number, Diagnosis
      , Description, Major_Diagnostic_Code
ORDER BY Code_Number ;

```

100 %

Results Messages

	Code_Number	Diagnosis	Description	Major_Diagnostic_Category	Freq
1	309.81	NULL	1	NULL	131
2	309.81	POSTTRAUMATIC STRESS DIS	POSTTRAUMATIC STRESS DISORDER	19.000000	3
3	309.81	PROLONG POSTTRAUM STRESS	PROLONGED POSTTRAUMATIC STRESS DISORDER	19.000000	1

Summary/Conclusion

- These basic commands (SELECT, FROM, WHERE, GROUP BY and ORDER BY) form the core of most SQL queries.
- Using this basic structure, you should be able to explore the content of any view in CDW.
- The goal of the next cyberseminar will be to show you how to combine columns from multiple views; in other words, it will step the new user through a JOIN in SQL.

Contact Information

Margaret Gonsoulin, PhD

VIReC@va.gov

708-202-2413

Questions?

A series of horizontal lines in various shades of blue and white, extending from the left edge of the slide and ending on the right side, creating a modern, layered effect.