

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

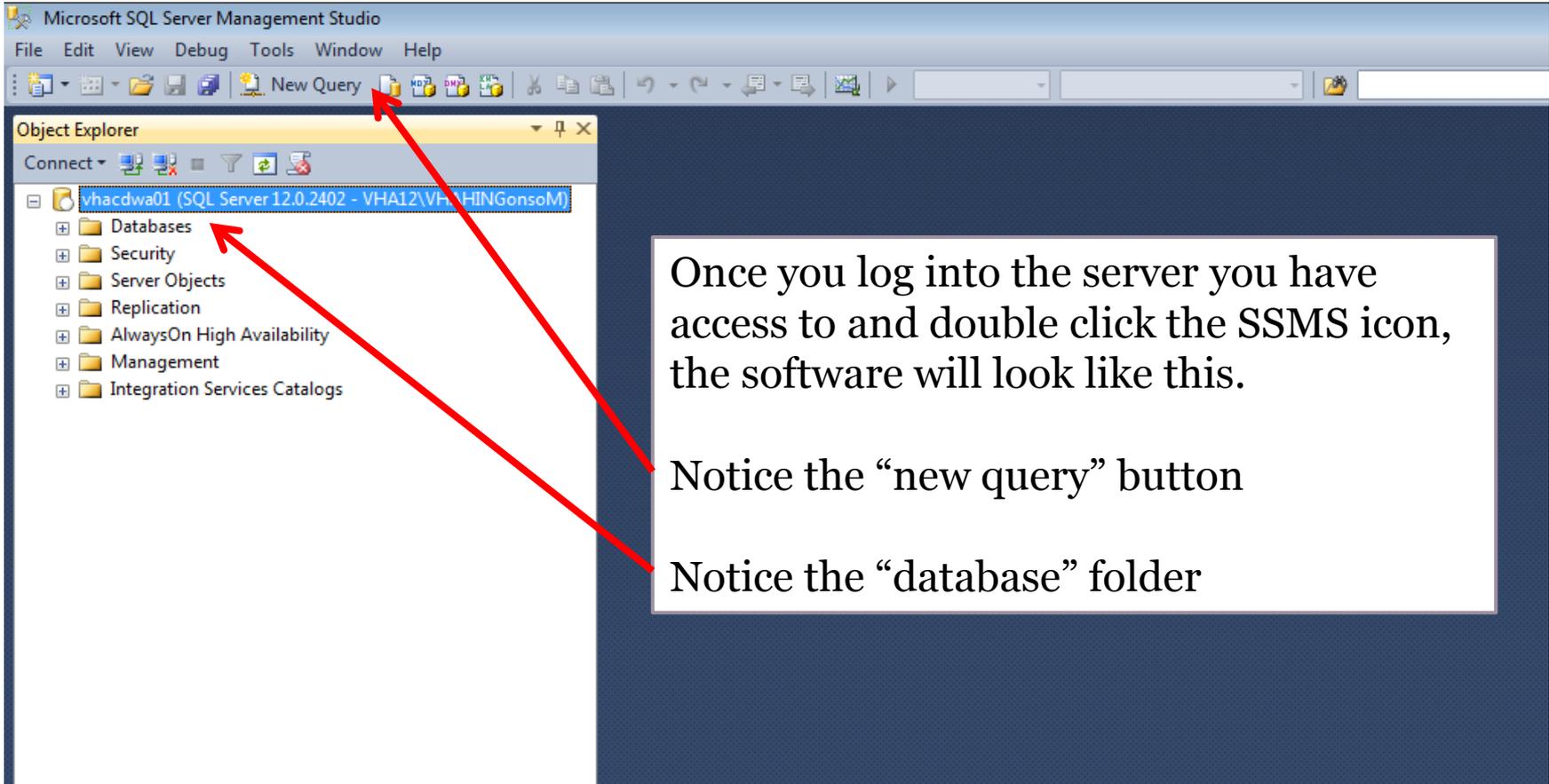
4

5 Very experienced with CDW

In the Production Domains

The Production table called Dim.ICD9

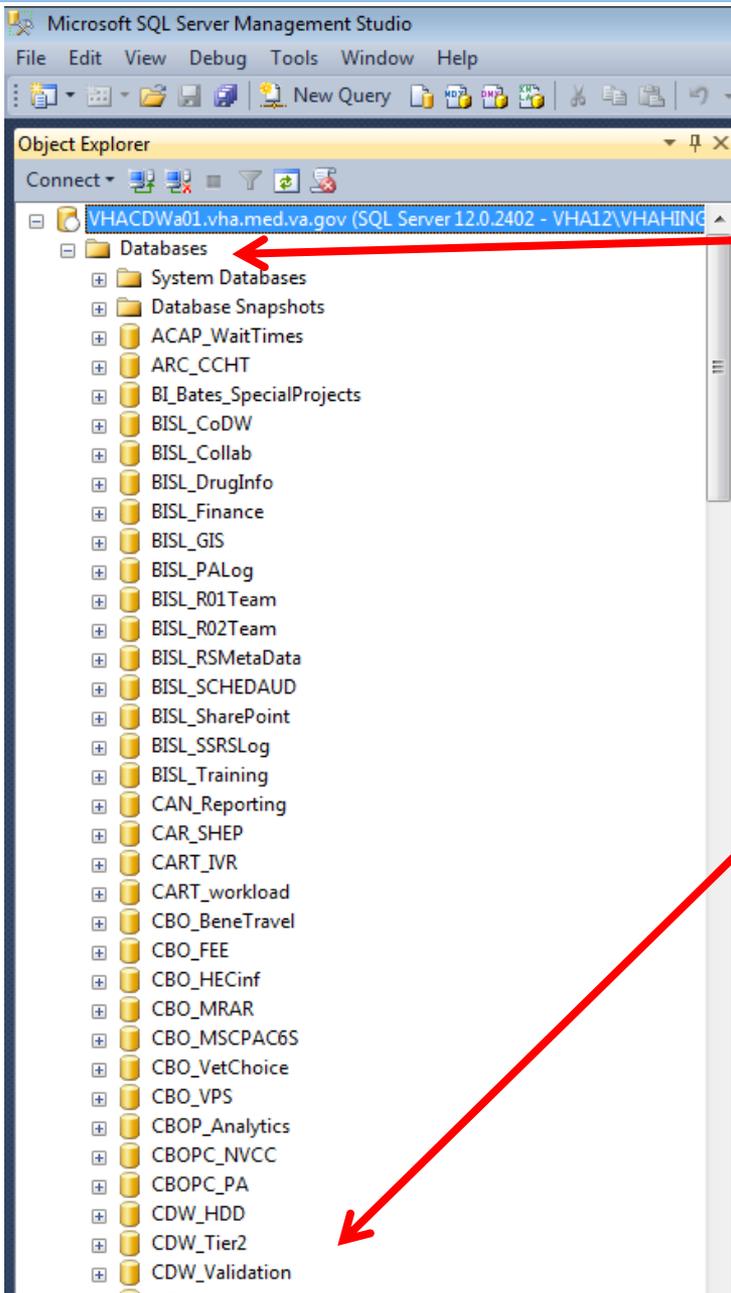
SQL Server Management Studio



Once you log into the server you have access to and double click the SSMS icon, the software will look like this.

Notice the “new query” button

Notice the “database” folder



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

Name
The name of the connection.

100 %
Connected. (1/1)

vhacdwa01 (12.0 RTM) VHA12\VAH\HINSONSO (152) CDWWork 00:00:00 0 rows

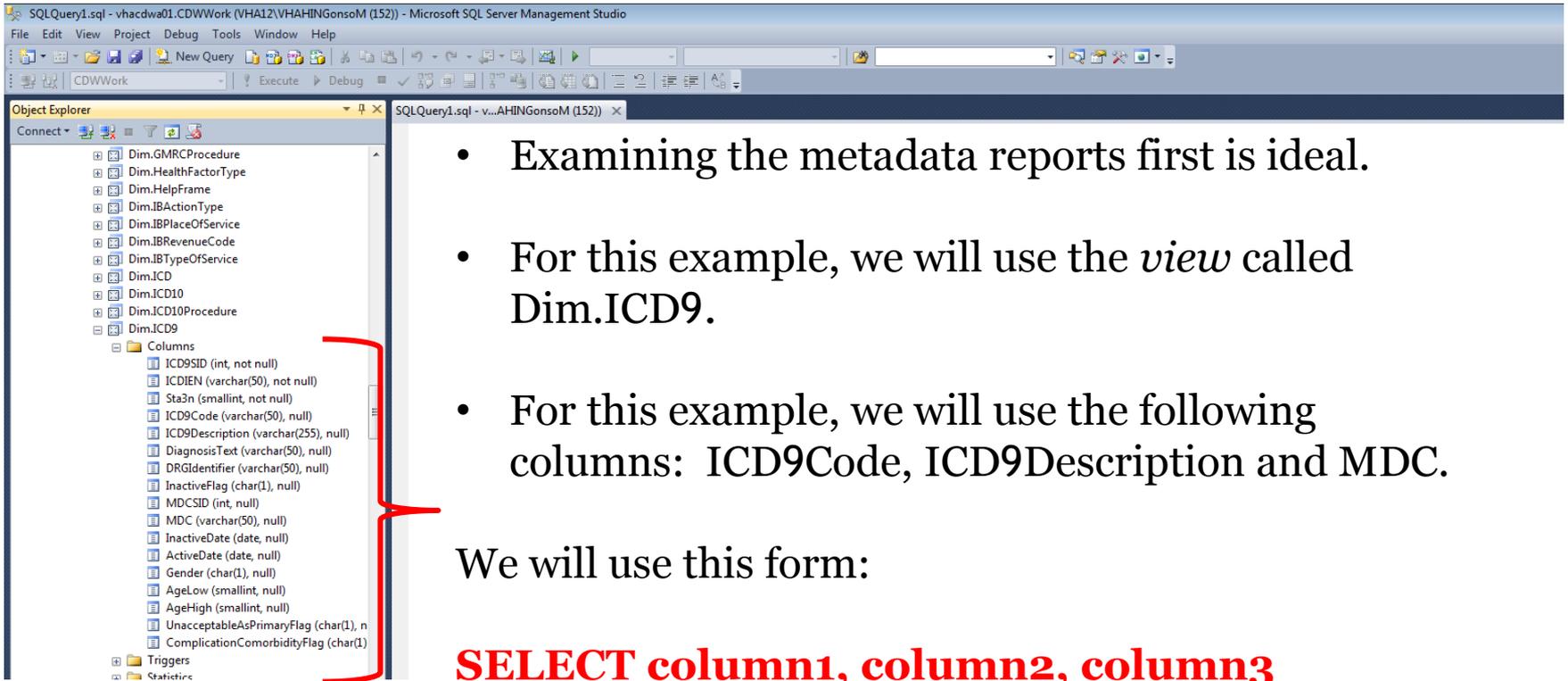
The screenshot shows the Microsoft SQL Server Management Studio interface. On the left, the Object Explorer displays a tree view of databases and views. A context menu is open over the 'Dim.ICD9' view, with 'Select Top 1000 Rows' highlighted. A red '1st' is placed next to this menu item. In the center, the SQL query editor shows a script for the 'SelectTopNRows' command, which selects the top 1000 rows from the 'Dim.ICD9' view. A red '2nd' is placed next to the query. At the bottom, the Results pane shows a table of data with columns: ICD9SID, ICDIEN, Sta3n, ICD9Code, and ICD9C. A red '3rd' is placed next to the results table.

- 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.

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 Management Studio. The 'Columns' folder for the 'Dim.ICD9' view is expanded, showing a list of columns with their data types and nullability. A red bracket highlights this list.

- 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 - vhadwa01.CDWWork (VHA12\VHAHINgonsoM (152))* - Microsoft SQL Server Management Studio

File Edit View Query Project Debug Tools Window Help

CDWWork Execute Debug

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 ;
```


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

SQLQuery1.sql - vhadcwa01.CDWWork (VHA12\VHAHINGonsoM (152))* - Microsoft SQL Server Management Studio

File Edit View Query Project Debug Tools Window Help

CDWWork Debug

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

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

```

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

```

100 %

Results Messages

ICD9Code	ICD9Description	MDC
1... 969.9	POISONING BY UNSPECIFIED PSYCHOTROPIC AG...	INJURY,POISONING,DRUG TOXICITY
1... 970.0	POISONING BY ANALEPTICS	INJURY,POISONING,DRUG TOXICITY
1... 970.1	POISONING BY OPIATE ANTAGONISTS	INJURY,POISONING,DRUG TOXICITY
1... 970.8	POISONING BY OTHER SPECIFIED CENTRAL NERV...	INJURY,POISONING,DRUG TOXICITY
1... 970.81	NULL	INJURY,POISONING,DRUG TOXICITY
1... 970.89	NULL	INJURY,POISONING,DRUG TOXICITY
1... 970.9	POISONING BY UNSPECIFIED CENTRAL NERVOUS...	INJURY,POISONING,DRUG TOXICITY
1... 971.0	POISONING BY PARASYMPATHOMIMETICS (CHOLI...	INJURY,POISONING,DRUG TOXICITY
1... 971.1	POISONING BY PARASYMPATHOLYTICS (ANTICHO...	INJURY,POISONING,DRUG TOXICITY
1... 971.2	POISONING BY SYMPATHOMIMETICS (ADRENERGI...	INJURY,POISONING,DRUG TOXICITY
1... 971.3	POISONING BY SYMPATHOLYTICS (ANTIADRENER...	INJURY,POISONING,DRUG TOXICITY
1... 971.9	POISONING BY UNSPECIFIED DRUG PRIMARILY A...	INJURY,POISONING,DRUG TOXICITY
1... 972.0	POISONING BY CARDIAC RHYTHM REGULATORS	INJURY,POISONING,DRUG TOXICITY
1... 972.1	POISONING BY CARDIOTONIC GLYCOSIDES AND D...	INJURY,POISONING,DRUG TOXICITY
1... 972.2	POISONING BY ANTILPEMIC AND ANTIARTERIOS...	INJURY,POISONING,DRUG TOXICITY
1... 972.3	POISONING BY GANGLION-BLOCKING AGENTS	INJURY,POISONING,DRUG TOXICITY
1... 972.4	POISONING BY CORONARY VASODILATORS	INJURY,POISONING,DRUG TOXICITY
1... 972.5	POISONING BY OTHER VASODILATORS	INJURY,POISONING,DRUG TOXICITY
1... 972.6	POISONING BY OTHER ANTIHYPERTENSIVE AGEN...	INJURY,POISONING,DRUG TOXICITY

```

SELECT ICD9Code, ICD9Description, MDC
FROM CDWWork.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

SQLQuery1.sql - vhadwa01.CDWWork (VHA12\VHAHINonsoM (152))* - Microsoft SQL Server Management Studio

File Edit View Query Project Debug Tools Window Help

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.ICDProcedure
- 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

```

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

```

100 %

ICD9Code	ICD9Description	MDC	Freq
1... 969.8	POISONING BY OTHER SPECIFIED PSYCHOTROPI...	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 ANTILIPERIC 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

Query executed successfully.

```

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

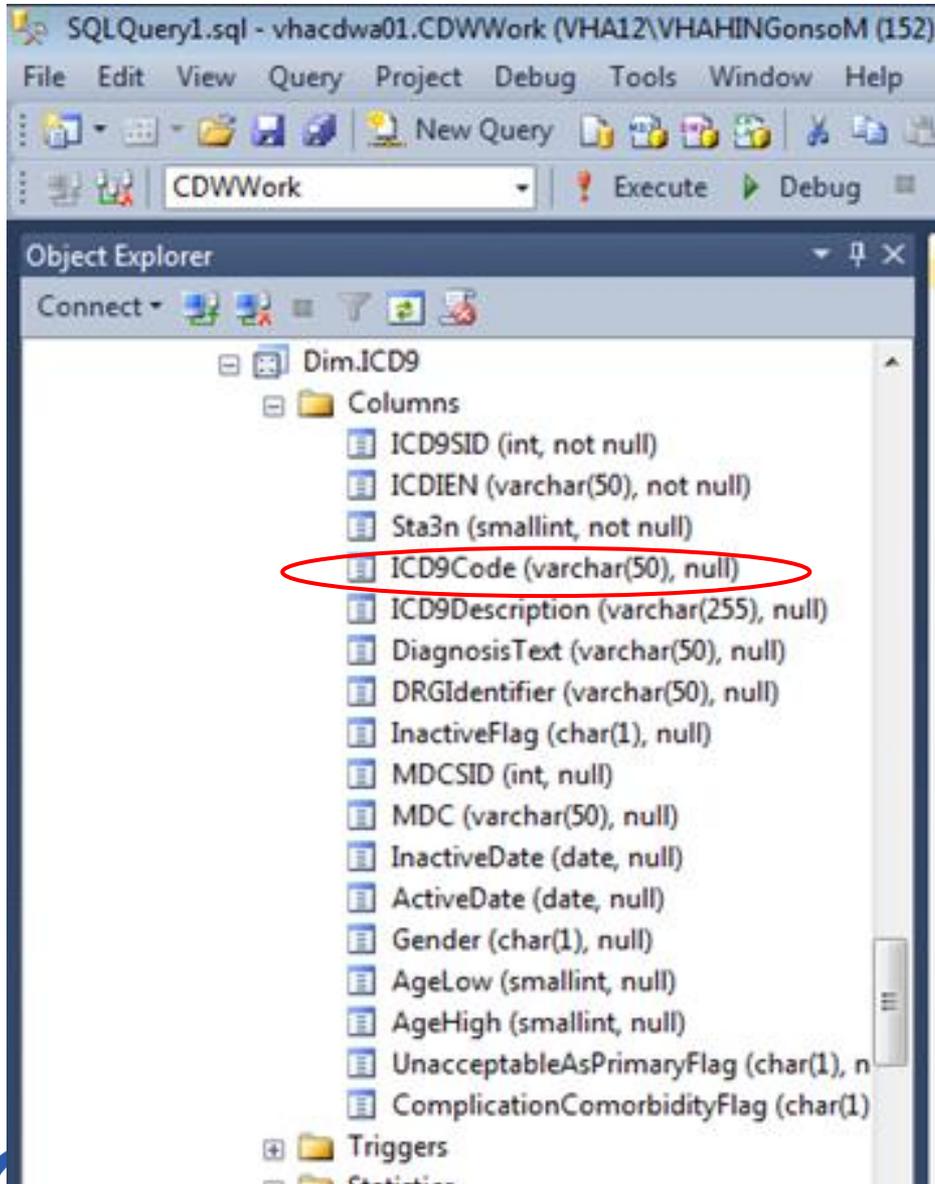
```



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.CDWork (VHA12\VHAHingonsoM (660))* - Microsoft SQL Server Management Studio

File Edit View Project Debug Tools Window Help

CDWork Execute Debug

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.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

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

```
SELECT ICD9Code, ICD9Description, MDC
FROM CDWork.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 CDWork.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

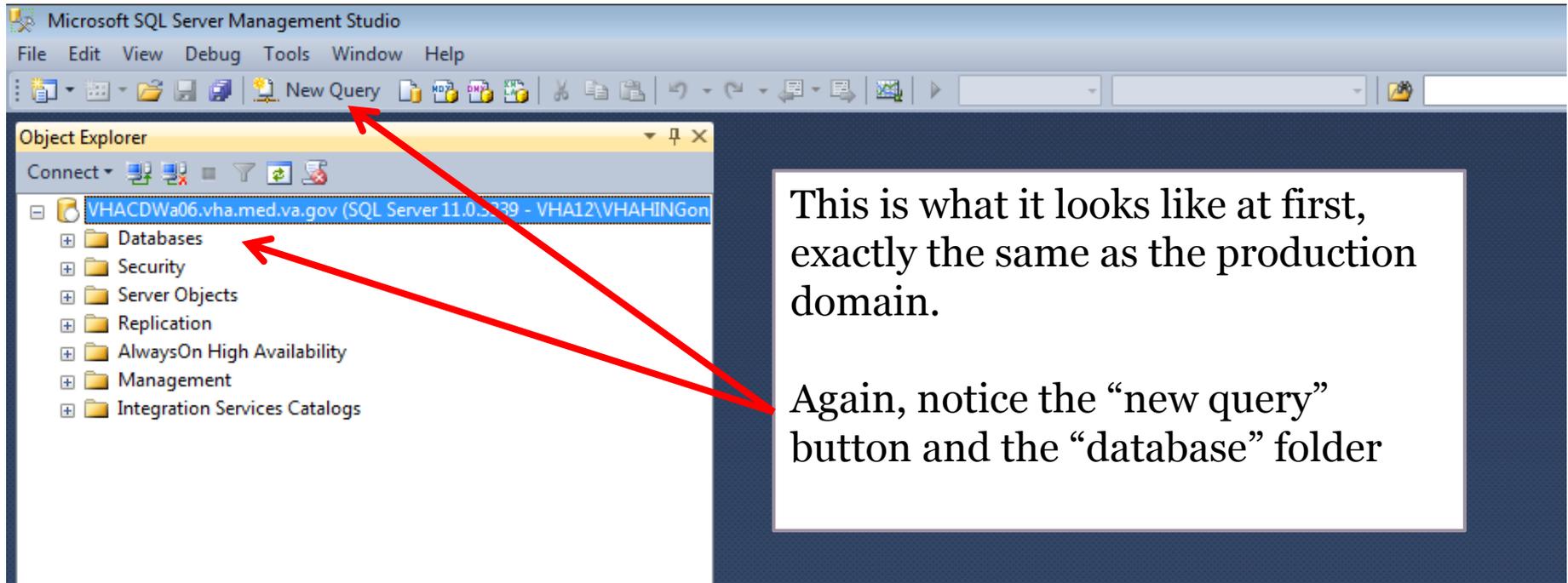
Which of the following best describes your role in the VA?

- 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

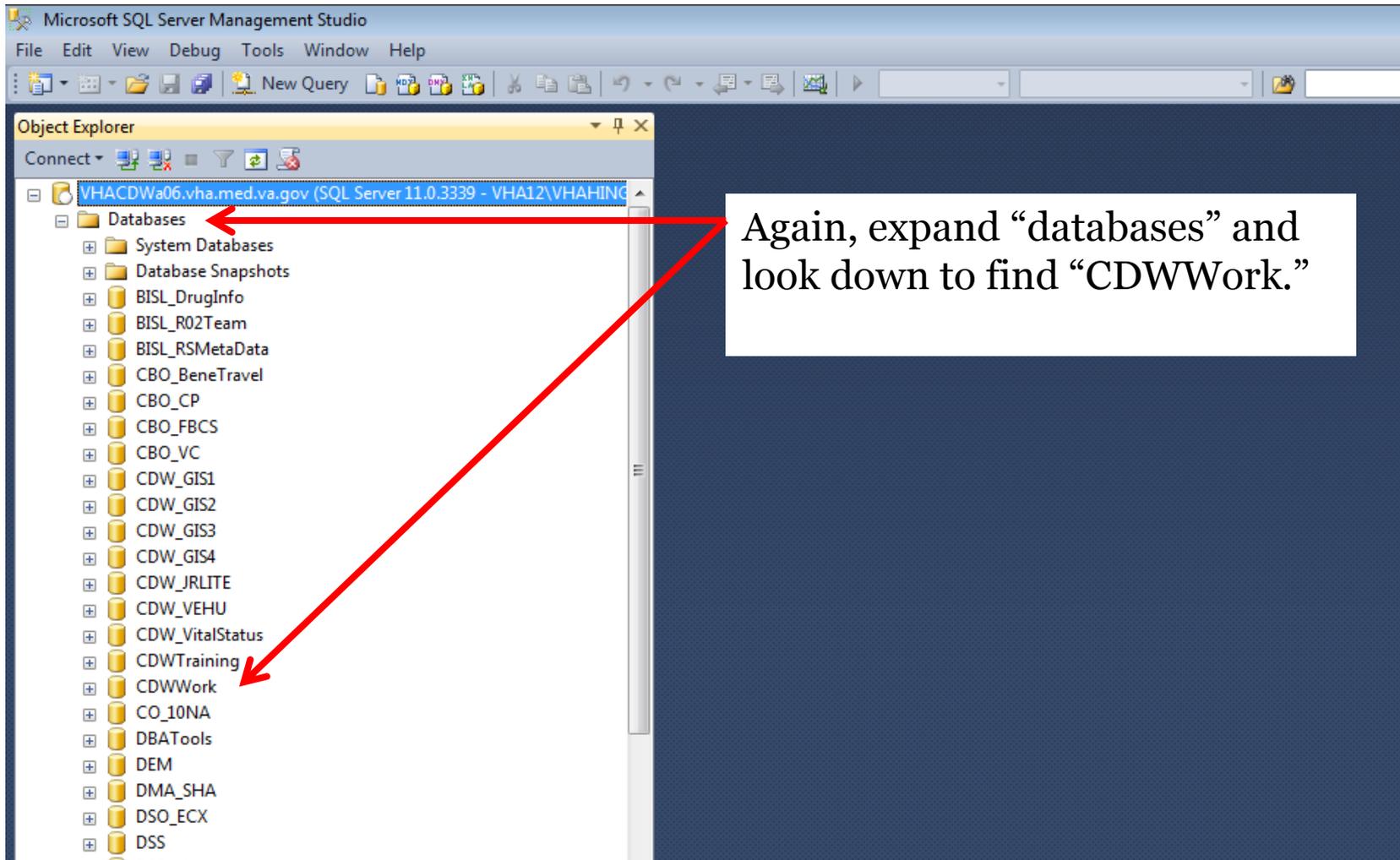


This is what it looks like at first, exactly the same as the production domain.

Again, notice the “new query” button and the “database” folder

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



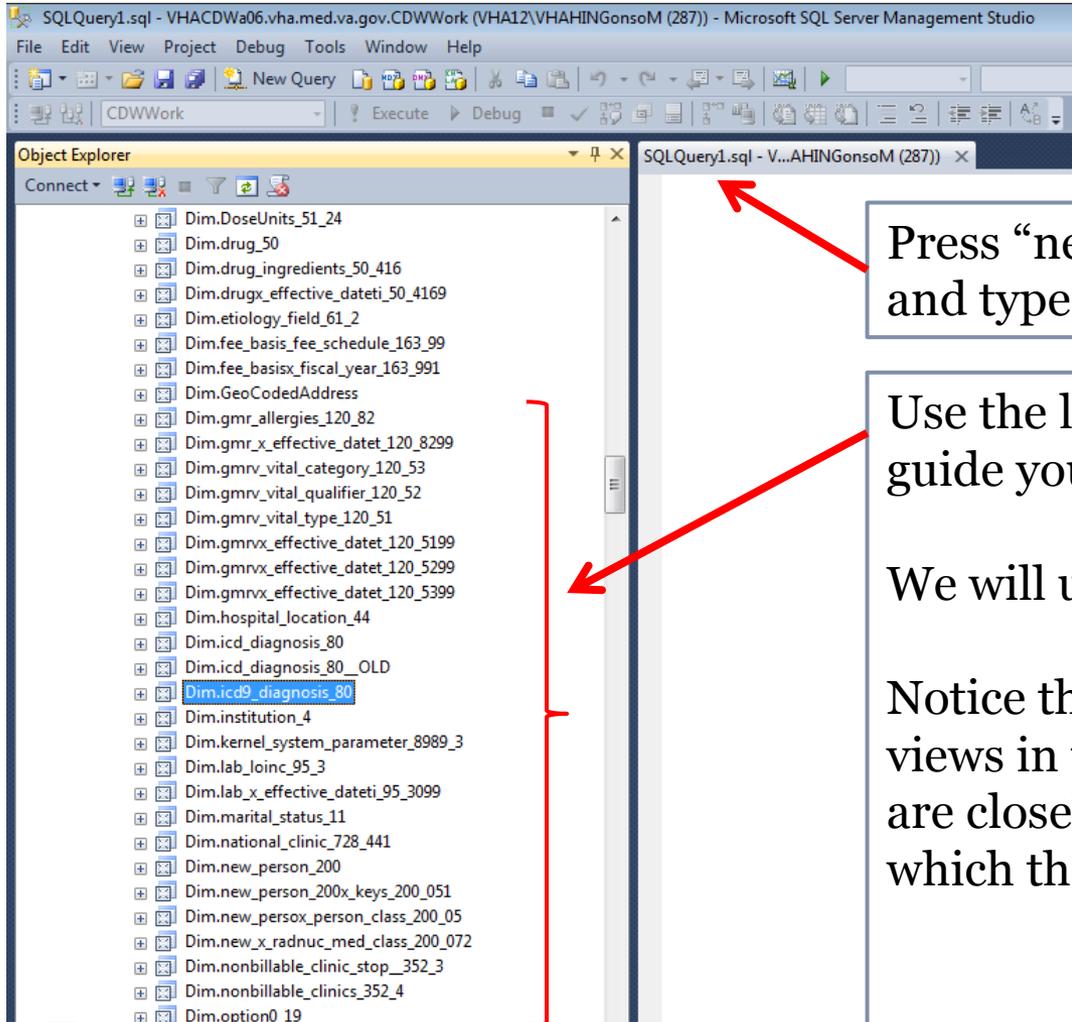
The screenshot shows the Microsoft SQL Server Management Studio interface. The Object Explorer pane is expanded to show the 'Databases' folder under the server instance 'VHACDWa06.vha.med.va.gov (SQL Server 11.0.3339 - VHA12\VHAHIN...)'. The list of databases includes:

- System Databases
- Database Snapshots
- BISL_DrugInfo
- BISL_R02Team
- BISL_RSMetaData
- CBO_BeneTravel
- CBO_CP
- CBO_FBCS
- CBO_VC
- CDW_GIS1
- CDW_GIS2
- CDW_GIS3
- CDW_GIS4
- CDW_JRLITE
- CDW_VEHU
- CDW_VitalStatus
- CDWTraining
- CDWork
- CO_10NA
- DBATools
- DEM
- DMA_SHA
- DSO_ECX
- DSS

A red arrow points to the 'Databases' folder, and another red arrow points to the 'CDWork' database.

Again, expand “databases” and look down to find “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.

SQLQuery3.sql - VHACDwa06.vha.med.va.gov.CDWWork (VHA12)VHAHINGoSoM (72) - Microsoft SQL Server Management Studio

File Edit View Project Debug Tools Window Help

CDWWork

Object Explorer

Connect

Dim.DoseUnits_51_24
Dim.drug_50
Dim.drug_ingredients_50_416
Dim.drugx_effective_dateti_50_4169
Dim.etiology_field_61_2
Dim.fee_basis_fee_schedule_163_99
Dim.fee_basis_fiscal_year_163_991
Dim.GeoCodedAddress
Dim.gmr_allergies_120_82
Dim.gmr_x_effective_datet_120_8299
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
Dim.institution_4
Dim.kernel_system
Dim.lab_loinc_95_3
Dim.lab_x_effective
Dim.marital_status_
Dim.national_clinic
Dim.new_person_20
Dim.new_person_20
Dim.new_perso_x_pe
Dim.new_x_radnuc
Dim.nonbillable_clin
Dim.nonbillable_clin
Dim.option0_19
Dim.option0_19x_m
Dim.patient_2
Dim.patientx_race_i
Dim.patix_ethnicity
Dim.provider_class
Dim.ptf_45
Dim.race_10
Dim.religion_13
Dim.security_key_19_1
Dim.servicesection_49
Dim.signsymptoms_120_83
Dim.signx_effective_datet_120_8399
Dim.site_codes
Dim.state_5

2nd

1st

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].[Dim].[icd9_diagnosis_80]

```

Results

station_no	row_id	code_number	identifier	diagnosis	
1	116	1.000000	100.81	CQ	LEPTOSPIRAL MENINGIT
2	116	2.000000	100.89	CQ	LEPTOSPIRAL INFECT NI
3	116	3.000000	100.9	NULL	LEPTOSPIROSIS NOS
4	116	4.000000	102.1	NULL	MULTIPLE PAPILLOMAT/
5	116	5.000000	102.2	NULL	EARLY SKIN YAWS NEC
6	116	6.000000	102.3	NULL	HYPERKERATOSIS OF Y
7	116	7.000000	102.4	NULL	GUMMATA AND ULCERS
8	116	8.000000	102.5	Y	GANGOSA
9	116	9.000000	102.6	NULL	YAWS OF BONE JOINT
10	116	10.000000	102.7	NULL	YAWS MANIFESTATIONS
11	116	11.000000	102.8	NULL	LATENT YAWS
12	116	12.000000	102.9	NULL	YAWS NOS
13	116	13.000000	103.1	NULL	PINTA INTERMED LESIO
14	116	14.000000	103.2	NULL	PINTA LATE LESIONS
15	116	15.000000	103.3	NULL	PINTA MIXED LESIONS
16	116	16.000000	103.9	NULL	PINTA NOS
17	116	17.000000	104.8	NULL	SPIROCHETAL INFECT N
18	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

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

File Edit View 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.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

RESEARCHERS' GUIDE TO VA DATA

- 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

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)
 - 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... *

```
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 PAPILOMATA	MULTIPLE PAPILOMATA 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

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

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 ;
```


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 - VHACDW06.vha.med.va.gov.CDWWork (VHA12\VHA\HNGonsoM (287))* - Microsoft SQL Server Management Studio

File Edit View Query Project Debug Tools Window Help

CDWWork

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_dateti_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_Catego	
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 ;
```

COUNT

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

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)
 - 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_keys_200_051

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

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

Results

Code_Number	Diagnosis	Description	Major_Diagnostic_Category	Freq	
7...	309.4	ADJ REACT-EMOTION/CON...	ADJUSTMENT REACTION WITH MIX...	19.000000	1
7...	309.4	ADJ DIS-EMOTION/CONDUCT	ADJUSTMENT DISORDER WITH MI...	19.000000	3
7...	309.81	NULL	1	NULL	131
7...	309.81	POSTTRAUMATIC STRESS ...	POSTTRAUMATIC STRESS DISORD...	19.000000	3
7...	309.81	PROLONG POSTTRAUM ST...	PROLONGED POSTTRAUMATIC ST...	19.000000	1
7...	309.82	NULL	1	NULL	131
7...	309.82	ADJUST REACT-PHYS SYMPT	ADJUSTMENT REACTION WITH PH...	19.000000	4
7...	309.83	ADJUST REACT-WITHDRAW...	ADJUSTMENT REACTION WITH WI...	19.000000	4
7...	309.83	NULL	1	NULL	131
7...	309.89	NULL	1	NULL	131
7...	309.89	ADJUSTMENT REACTION NEC	OTHER SPECIFIED ADJUSTMENT R...	19.000000	4
7...	309.9	NULL	1	NULL	131
7...	309.9	ADJUSTMENT REACTION N...	UNSPECIFIED ADJUSTMENT REACT...	19.000000	4
7...	310.0	NULL	1	NULL	131
7...	310.0	FRONTAL LOBE SYNDROME	FRONTAL LOBE SYNDROME	19.000000	4
7...	310.1	ORGANIC PERSONALITY SY...	ORGANIC PERSONALITY SYNDROME	19.000000	1
7...	310.1	NULL	1	NULL	131
7...	310.1	PERSONALITY CHG OTH DIS	PERSONALITY CHANGE DUE TO CO...	19.000000	3
7...	310.2	POSTCONCUSSION SYNDR...	POSTCONCUSSION SYNDROME	1.000000	4

Query executed successfully.

```
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 decorative graphic consisting of a solid blue horizontal bar that spans the width of the slide. Below this bar, on the right side, there are several horizontal lines of varying lengths and colors, including light blue and white, creating a layered, stepped effect.