Toxic Release Inventory Basic Data File Format Documentation

Prepared by: The Environmental Protection Agency
Office of Environmental Information
Office of Information Analysis and Access
Toxics Release Inventory Program Division
Information and Outreach Branch

Date: July 14, 2009

Table of Contents

TRI Basic Data File Format - Introduction

The *TRI Basic Data File* are a convenient and easy-to-use file format that contains the data elements most frequently requested by data users. Over the years, the Toxics Release Inventory (TRI) Program has found that data users are most interested in the quantity data reported by TRI facilities. These include the quantities of toxic chemicals released on-site, transferred off-site and summary data concerning releases, recycling, energy recovery and treatment as it appears in the Pollution Prevention portion of TRI's form R.

The *TRI Basic Data Files* were created in response to user requests for state and national data sets. Frequently, users would contact the TRI program division and request *all* the TRI data. However, upon further correspondence, it was discovered they only wanted the quantity data elements (i.e. the quantity of On-site Air Releases, etc.). But, they still wanted these data elements for all the years. Users noted that they couldn't get this information from the TRI on-line query tools. They said using the tools for their purpose proved cumbersome, time consuming and in many cases just didn't work as the tools would "time out" in many cases because of the long running queries.

So, the TRI program division developed the *TRI Basic Data Files* to meet this need. This document defines and describes the data file format of these files. It lists the record layout and specifically defines each field in that layout. Finally, it gives users instructions on how to load these data files into two common desktop products, Microsoft (MS) Excel and Access.

Note that, this file format does not contain data from non-quantitative areas of the Form R or A (the two forms TRI uses to collect data from facilities) such as "Chemical Activities and Uses", "Treatment Methods and Efficiency", "Recycling Processes" and "Energy Recover Processes". To see the data from these and other non-quantitative sections of the Form R and A, refer to the "TRI Basic Plus Data File" at the TRI Early Data Sharing Website.

File Format Information

The general categories that users are interested in and that are represented in this file are as follows:

- Facility Name, Address, Latitude & Longitude Coordinates and SIC or NAICS codes
- Chemical identification and classification information
- On-site Release quantities
- Publicly Owned Treatment Works (POTW) transfer quantities
- Off-site Transfer quantities for release/disposal and further waste management
- Summary Pollution Prevention quantities (Section 8 of the Form R)

The file contains approximately 100 data fields. The data is simple ASCII text that is delimited by TABs. This format loads easily into many desktop products such as spreadsheets and database software. See Appendix B and C for instructions on how to load the data files into MS Excel and MS Access.

The record layout for the file appears in the next section. There are seven columns in the layout format. The first column (identified by the column heading '#') is a sequential field number identifier. The second column, "Field" is the name of the data field as it will appear in the data file. Many of the field names begin with a section reference,

such as "5.1 - Fugitive Air". The "5 .1" represents the section of the Form R where the data came from. Many users find the data fields easier to use when they are prefaced with the section number.

The third and fourth columns, "Maximum Length" and "Data Type", specify the maximum length and the data type of the field. The "Maximum Length" column also indicates the format of numeric data. Comma notation is used for numbers that may contain decimals.   For example, a "Maximum Length" value of "22,7" indicates that the number can be 22 digits long with 7 digits to the right of the decimal point. There are two possible values for the "Data Type". They are 'C' for Character/Text data and 'N' for numeric data.

The fifth and sixth columns under the "Form R Reference" heading indicate the "Part" and "Section" of the Form R or A where the data originates from. Refer to the "Toxic Chemical Release Reporting Forms and Instructions" documentation at http://www.epa.gov/tri/report/index.htm for more information on the Form R and A and a supplemental information for of each data field in this file.

The "Definition" column gives a description of each data element and provides notes about its origin and use. There are several data fields that represent totals in the data file. The "Definition" column tells which data fields are added together to obtain the totals.

This file format can be loaded easily into several common desk top products. Two products that many users have are Microsoft Excel and Access. Appendices B and C provide instructions for loading these files into those two products.

File Content Types

On average, over the course of the TRI program, approximately 23,000 facilities report to TRI each year. They report releases and other waste management data for 85,000 to 105,000 toxic chemicals annually. Each record in the *TRI Basic Data Files* will represent one of these chemicals as reported by one TRI facility per year.

There are three specific types of the *TRI Basic Data Files* based on content.

The first type, known as the "State Data File", will only contain data for one state, district or US territory per one reporting year. For instance, one "State File" might be all the data for the state of Alabama for reporting year 2008.

The second type, known as the "National Data File", will contain all the data for the Nation as a whole. This includes data for all 50 states and the six US districts and territories (i.e. American Samoa, District of Columbia, Guam, Northern Mariana Islands, Puerto Rico and the Virgin Islands).

The third type, known as the "Federal Facility Data File", will contain data for all government owned and operated federal sites.

The data for each file type, whether is it is the State, National and Federal Facility file will be zipped into a self-extracted zip file with a file extension of .exe   Users will be able to download these self-extracting zipped files to their computers and unzip it (just by clicking on the file name) automatically without the need for separate WinZip software.

File Naming Conventions and Version Numbers

Each file type will have a slightly different naming convention. However, all file names will follow this standard pattern:

File Name = "TRI" + Reporting Year + File Content + Version

Where the text string "TRI" identifies the file as TRI data. The "Reporting Year" indicates which reporting the file represents. The "File Content" will indicate which of the three file types (State, National or 56 States) the file is. It will be either…

- A State Abbreviation   … for a State File
- The word "US"   … for a National File
- The words "States"   … for the 56 States File

The "Version" indicates which year of data the file was created with. TRI facilities can revise their data submissions at any time. So, historic data sometimes changes. The version number specifies the point in time or reporting year from which the data was created from. For instance, the 2003 National Data file that has a version number of "v08" (i.e. TRI_2003_US_v08.txt) is data that reporting for the 2003 data cycle but extracted as it existed as of reporting year 2003.

Here are some examples and explanations of each file type:

State File:

> TRI_2008_AL_v08.txt

The file name above, "TRI_2008_AL_v08.txt" is a "State File" type. It contains the facility and chemical identification as well as quantity data for all chemicals reported by facilities in Alabama (AL) for reporting year 2008. The version number is "v08" signifying that the file was created with Reporting Year 2008 data.

National File:

> TRI_2000_US_v08.txt

The file name above, "TRI_2000_US_v08.txt" is a "National File" type. It contains the facility and chemical identification as well as quantity data for all chemicals reported by all facilities in reporting year 2008.   The version number is "v08" signifying that the file was created with Reporting Year 2008 data. In other words, the file contains the 2000 data as is existed during the reporting year 2008 cycle.

Federal Facility File:

> TRI_2005_FED_v08.txt

The file name above, "TRI_2005_FED_v08.txt" is a "Federal Facility File" type.   It contains the facility and chemical identification as well as quantity data for all chemicals reported by All Federal Facilities in reporting year 2005. The version number is "v08" signifying that the file was created with Reporting Year 2008 data. In other words, the file contains the 2005 data as is existed during the reporting year 2008 cycle.

File Usage – Limiting Factors

Appendices B and C describe the steps for loading the *TRI Basic Data Files* into Microsoft Excel and Access respectively.   Note that for versions of Microsoft Excel prior to and including Version 2003 (found in Microsoft Office 2003) there is a limitation of 65,535 rows that can be loaded into a worksheet.   Because of this limitation, users will not be able to load the "National Files" containing 85,000 or more records into these versions of Excel. However, the limit on the 2008 version of Excel is 1,048,576. So, users of that version of Excel will be able to load the National Files as well as the other two file types without a problem.

Microsoft Access, a desktop database tool, does not have a limit on the number of records that can be loaded into a table. Check the limitations of any other tool to be sure that the TRI Basic Data Files can be loaded in full.

Zeroes in the Data

The *TRI Basic Data Files* were created to be loaded into different tools such as spreadsheets, databases and statistical applications. Some of those tools require that numeric data be populated with a number (and not a blank) in order for the functionality of the tool to work correctly. For instance, to total a column in many spreadsheet applications, all rows in that column must contain a number and not be blank.

In light of this, zeroes have been input into the TRI Basic Data Files in places where numeric data was blank. There are three instances where this normally occurs. First, facilities can file either a Form R or a Form A when reporting to TRI. The Form A is a short form that is used when chemical releases and other waste management quantities are below certain thresholds. It only requires that facility enter its identifying information and the identity of the chemical it's releasing or otherwise managing as waste. It does not require the facility to enter any actual quantity data. For chemicals submitted as Form As, all the quantity data is represented as zeroes. See the Toxic Chemical Release Reporting Forms and Instructions at http://www.epa.gov/tri/report/index.htm for more information on Form As and reporting thresholds.

The second case where zeroes have been substituted for blanks occurs when facilities report "NA" or "Not Applicable" for a quantity on the Form R. Reporting "NA" means that the release or waste management quantity is not possible for that facility. For example, if a facility is not located near a water body, it will not have the ability to release any of the chemical to water. Therefore, in section 5.3 of the form R where facilities are asked to enter their on-site water releases, the facility would enter "NA" because the release is not possible. See the Toxic Chemical Release Reporting Forms and Instructions at http://www.epa.gov/tri/report/index.htm for more information on the use of NA in TRI reporting.

The third case where zeroes have been substituted for blanks occurs when facilities do not respond to quantity questions on the form R and leave them blank. This occurs most often with the submission of paper forms. Data submitted via the TRI automated reporting tools (i.e. TRI-ME desktop and TRI-ME web) do not allow for blanks in quantity data reporting. They require the submitter to enter a number or indicate "NA".

Other On-line Tools for Accessing TRI Data

Besides the data in this file and the other data sources referenced above, users can also access and query data on-line via two TRI dissemination tools. TRI-Explorer is an easy to on-line application that queries TRI data. It generate reports and data files. TRI-Explorer allows a user to build a query, generate results and then drill down into those results for more information. TRI-Explorer focuses on Release, Off-site Transfers and Waste Quantity (Section 8 of the Form R) Summary Reports. For each report, users can focus on the areas of Chemicals, Facilities, Federal Facilities, Trends, Geography and Industries. TRI-Explorer can be found at http://www.epa.gov/triexplorer/chemical.htm. See the introduction to the tool at http://www.epa.gov/triexplorer/introduction.htm.

The "TRI" section of Envirofacts provides all the data contained in TRI-Explorer and more. Envirofacts works similar to TRI-Explorer. It allows the user to build queries by filling text boxes and choosing from drop down lists. It allows the user to generate reports and down loadable data files. Users can generate reports in the areas of Facilities Summaries, Form R facsimiles, EZ Query reports that are easy to use and focus on specific areas of reporting, the Form R download which allows users to chose any data elements off the form R and create files and reports and the very flexible Customized Query. The TRI section of Envirofacts can be found at http://www.epa.gov/enviro/html/tris/tris_query.html. See the overview of the tool at http://www.epa.gov/enviro/html/tris/index.html .

Record Layout

Field Documentation for the "Format II Data" File

#	Field	Maximum Length	Data Type	Form R Reference Part	Section	Definition
1	Year	4	C	I	1	The Reporting Year - Year the chemical was released or waste managed
2	TRI Facility ID	15	C	I	4.1	The TRI Facility Identification Number assigned by EPA/TRI
3	Facility Name	62	C	I	4.1	Facility Name
4	Street Address	62	C	I	4.1	Street Address where facility is located
5	City	28	C	I	4.1	City Name where facility is located
6	County	25	C	I	4.1	County Name where facility is located
7	ST	2	C	I	4.1	State Abbreviation where the facility is located
8	ZIP	9	C	I	4.1	ZIP code where facility is located. Either 5 or 9 characters. No hyphens.
9	Latitude	9,6	N	None	None	Facility Latitude represented as decimal data
10	Longitude	10,6	N	None	None	Facility Longitude represented as decimal data
11	Primary SIC	4	C	I	4.5	Primary Standard Industrial Code (SIC) Code that represents the Facility's primary Business activity. SIC codes were discontinued in 2006 when NAICS codes became the new business activity indicator for TRI.
12	SIC 2	4	C	I	4.5	Supplemental SIC code representing other business activities of the facility.
13	SIC 3	4	C	I	4.5	Supplemental SIC code representing other business activities of the facility.
14	SIC 4	4	C	I	4.5	Supplemental SIC code representing other business activities of the facility.
15	SIC 5	4	C	I	4.5	Supplemental SIC code representing other business activities of the facility.
16	SIC 6	4	C	I	4.5	Supplemental SIC code representing other business activities of the facility.
17	Primary NAICS	6	C	I	4.5	Primary North American Industry Code System (NAIC) code that represents the Facility's primary Business activity. NAICS codes were first reported in 2006. For years prior to 2006, NAICS codes have been assigned.
18	NAICS 2	6	C	I	4.5	Supplemental NAICS code representing other business activities of the facility
19	NAICS 3	6	C	I	4.5	Supplemental NAICS code representing other business activities of the facility
20	NAICS 4	6	C	I	4.5	Supplemental NAICS code representing other business activities of the facility

Field Documentation for the "Format II Data" File

#	Field	Maximum Length	Data Type	Form R Reference Part	Form R Reference Section	Definition
21	NAICS 5	6	C	I	4.5	Supplemental NAICS code representing other business activities of the facility
22	NAICS 6	6	C	I	4.5	Supplemental NAICS code representing other business activities of the facility
23	Doc_Ctrl_Num	13	C	None	None	The Document Control Number is a unique ID that is assigned to each form.
24	Chemical	70	C	II	1.2	Name of Chemical
25	CAS # / Compound ID	9	C	II	1.1	The Chemical Abstract Service Number of the chemical or chemical compound category
26	Clean Air Act Chemical	3	C	None	None	Indication if the chemical is a Clean Air Act Chemical (Yes or No)
27	Classification	6	C	None	None	Classification of the chemical. Values are as follows: TRI - Standard TRI Chemical PBT - Persistent Bioaccumulatvie Toxic Dioxin - Dioxin or Dioxin-like Compound
28	Metal	3	C	None	None	Indication if the chemical is a metal (Yes or No)
29	Metal Category	1	C	None	None	Category of Metal. Values are either 1, 2, 3, or 4 for metals. See Appendix A for definitions and lists of Chemicals that belong to each category
30	Carcinogen	3	C	None	None	Indication if the chemical is a carcinogen (Yes or No)
31	Form Type	1	C	None		The form the data was submitted on. Values are: A – Form A R – Form R
32	Unit of Measure	6	C	None	None	The units of measure the chemical is displayed in (Grams or Pounds)
33	5.1 - Fugitive Air	22,7	N	II	5.1	On-site Fugitive Air Releases
34	5.2 - Stack Air	22,7	N	II	5.2	On-site Stack Air Releases
35	5.3 - Water	22,7	N	II	5.3	On-site Water Releases
36	5.4.1 - Underground Class I	22,7	N	II	5.4.1	On-site Underground Injection Releases to Class I Wells
37	5.4.2 - Underground Class II-V	22,7	N	II	5.4.2	On-site Underground Injection Releases to Class II-V Wells
38	5.5.1A - RCRA C Landfills	22,7	N	II	5.5.1A	On-site RCRA C Landfills Releases
39	5.5.1B - Other Landfills	22,7	N	II	5.5.1B	On-site Other Landfills Releases
40	5.5.2 - Land Treatment	22,7	N	II	5.5.2	On-site Land Treatment Releases

Field Documentation for the "Format II Data" File

#	Field	Maximum Length	Data Type	Form R Reference Part	Section	Definition
41	5.5.3 - Surface Impoundment	22,7	N	II	5.5.3	On-site Surface Impoundment. In reporting years prior to 2003, there was no distinction between RCRA and Other Surface impoundments. For those years, this was the only total quantity of on-site surface impoundment reported. This field will contain mostly zeroes for years 2003 and after. To obtain total Surface Impoundment for any year, add fields 41, 42 and 43 together.
42	5.5.3A - RCRA Surface Impoundment	22,7	N	II	5.5.3A	On-site RCRA Surface Impoundment Releases. This sub category of surface impoundment was created in 2003. Prior to 2003, all on-site surface impoundment was reported in variable #41, "Surface Impoundment". This field will contain mostly zeroes for years prior to 2003. To obtain total Surface Impoundment for any year, add fields 41, 42 and 43 together.
43	5.5.3B - Other Surface Impoundment	22,7	N	II	5.5.3B	On-site NON-RCRA/Other Surface Impoundment Releases. This sub category of surface impoundment was created in 2003. Prior to 2003, all on-site surface impoundment was reported in variable #41, "Surface Impoundment". This field will contain mostly zeroes for years prior to 2003. To obtain total Surface Impoundment for any year, add fields 41, 42 and 43 together.
44	5.5.4 - Other Disposal	22,7	N	II	5.5.4	On-site Other Disposal Releases
45	On-site Release Total	22,7	N	II		Total Releases On-site for a chemical at a facility. This is a summation of all releases in section 5 (fields 33 through 44).
46	6.1 - POTW - Metals and Metal Compounds	22,7	N	II	6.1	If a chemical is a metal (or metal compound) the amount of transfer to a POTW will appear under this data element. Metals and Metal Compounds can not be treated by most POTWs. Therefore, when a metal is transferred to a POTW, TRI considers it an Off-site Release. It will be added to the Off-site Release total.
47	6.1 - POTW - Non-Metals	22,7	N	II	6.1	If a chemical is not a metal, its POTW release amount is listed here. This amount will not be added toward the off-site release total.
48	6.1 - POTW - Total Transfers	22,7	N	II	6.1	This amount is the total transfers to a POTW of a chemical. This amount will match either 6.1 POTW - Metals and Metal Compounds or 6.1 POTW - Non-Metals
49	6.2 - M10	22,7	N	II	6.2	Off-site Storage
50	6.2 - M41	22,7	N	II	6.2	Off-site Solidification/Stabilization for Metals and Metal Compounds Only

Field Documentation for the "Format II Data" File

#	Field	Maximum Length	Data Type	Form R Reference Part	Section	Definition
51	6.2 - M62	22,7	N	II	6.2	Off-site Wastewater Treatment (Excluding POTWs) for Metals and Metal Compounds Only
52	6.2 - M71	22,7	N	II	6.2	Off-site Underground Injection. There was no distinction between off-site underground injections into class I wells and class II-V wells prior to 2003. For those years, this was the only quantity for off-site underground injection. In 2003, two sub categories were added to distinguish between underground injections to class I wells and class II-V wells. This field will contain mostly zeroes for years 2003 and after. To obtain total off-site underground injection for any year, add fields 52, 53 and 54 together.
53	6.2 - M81	22,7	N	II	6.2	Off-site Underground Injection to Class I Wells. This sub category of off-site underground injection was created in 2003. Prior to 2003, all off-site underground injection quantities were reported in variable #52, "M71". This field will contain mostly zeroes for years prior to 2003. To obtain total off-site underground injection for any year, add fields 52, 53 and 54 together.
54	6.2 - M82	22,7	N	II	6.2	Off-site Underground Injection to Class II-V Wells. This sub category of off-site underground injection was created in 2003. Prior to 2003, all off-site underground injection quantities were reported in variable #52, "M71". This field will contain mostly zeroes for years prior to 2003. To obtain total off-site underground injection for any year, add fields 52, 53 and 54 together.
55	6.2 - M72	22,7	N	II	6.2	Off-site Landfill/Disposal Surface Impoundment. Prior to 2002, all Landfill and Surface Impoundment releases were reported as "M72". In 2002, "M72" was split into three subcategories representing Surface Impoundment (M63), Releases to Other Landfills (M64) and Releases to RCRA Subtitle landfills (M65). Subsequently, "M63" was subdivided into Subtitle C Surface Impoundment (M66) and Other Surface Impoundment (M67) in 2003. This field will contain mostly zeroes for years 2002 and after. To obtain total off-site Landfill/Disposal Surface Impoundment for any year, add fields 55 though 60 together.

Field Documentation for the "Format II Data" File

#	Field	Maximum Length	Data Type	Form R Reference Part	Section	Definition
56	6.2 - M63	22,7	N	II	6.2	Off-site Surface Impoundment. This sub category of Off-site Landfill/Disposal and Surface Impoundment was established in 2002. Prior to 2002, all off-site landfill/disposal surface impoundment quantities were reported in variable #55, "M72". This field will contain mostly zeroes for years prior to 2002. In 2003, this field was subdivided into Subtitle C Surface Impoundment (M66) and Other Surface Impoundment (M67). To obtain total off-site Surface Impoundment for any year, add field Ns 56, 57 and 58 together.
57	6.2 - M66	22,7	N	II	6.2	Off-site Subtitle C Surface Impoundment. This sub category of Off-site Surface Impoundment was established in 2003. Prior to 2003, all off-site surface impoundment quantities were reported in variable #56, "M63". This field will contain mostly zeroes for years prior to 2003. To obtain total off-site Surface Impoundment for any year, add field numbers 56, 57 and 58 together.
58	6.2 - M67	22,7	N	II	6.2	Off-site Other Surface Impoundment. This sub category of Off-site Surface Impoundment was established in 2003. Prior to 2003, all off-site surface impoundment quantities were reported in variable #56, "M63. This field will contain mostly zeroes for years prior to 2003. To obtain total off-site Surface Impoundment for any year, add field Ns 56, 57 and 58 together.
59	6.2 - M64	22,7	N	II	6.2	Off-site Other Landfills. This sub category of Off-site Landfill/Disposal and Surface Impoundment was established in 2002. Prior to 2002, all off-site landfill/disposal surface impoundment quantities were reported in variable #55, "M72". This field will contain mostly zeroes for years prior to 2002.
60	6.2 - M65	22,7	N	II	6.2	Off-site RCRA Subtitle C Landfill. This sub category of Off-site Landfill/Disposal and Surface Impoundment was established in 2002. Prior to 2002, all off-site landfill/disposal surface impoundment quantities were reported in variable #55, "M72". This field will contain mostly zeroes for years prior to 2002.
61	6.2 - M73	22,7	N	II	6.2	Off-site Land Treatment
62	6.2 - M79	22,7	N	II	6.2	Off-site Other Land Disposal
63	6.2 - M90	22,7	N	II	6.2	Off-site Other Off-site Management
64	6.2 - M94	22,7	N	II	6.2	Off-site Transfer to Waste Broker – Disposal
65	6.2 - M99	22,7	N	II	6.2	Off-site Unknown

Field Documentation for the "Format II Data" File

#	Field	Maximum Length	Data Type	Form R Reference Part	Form R Reference Section	Definition
66	Off-Site Release Total	22,7	N	None	None	The Off-site Release Total equals the sum of M10 + M41 + M62 + M71 + M72 + M63 + M64 + M65 + M73 + M79 + M90 + M94 + M99 + M40 (if the chemical is a category 1,3 or 4metal) + M61 (if the chemical is a category 1,3 or 4 metal) + (6.1 POTW - Metals and Metal Compounds Only)
67	6.2 - M20	22,7	N	II	6.2	Off-site Solvents/Organics Recovery
68	6.2 - M24	22,7	N	II	6.2	Off-site Metals Recovery
69	6.2 - M26	22,7	N	II	6.2	Off-site Other Reuse or Recovery
70	6.2 - M28	22,7	N	II	6.2	Off-site Acid Regeneration
71	6.2 - M93	22,7	N	II	6.2	Off-site Transfer to Waste Broker - Recycling
72	Off-Site Recycled Total	22,7	N	None	None	The sum of M20 + M24 + M26 + M28 + M93
73	6.2 - M56	22,7	N	II	6.2	Off-site Energy Recovery
74	6.2 - M92	22,7	N	II	6.2	Off-site Transfer to Waste Broker - Energy Recovery
75	Off-Site Recovery Total	22,7	N	None	None	The sum of M56 + M92
76	6.2 - M40	22,7	N	II	6.2	Off-site Solidification/Stabilization
77	6.2 - M50	22,7	N	II	6.2	Off-site Incineration/Thermal Treatment
78	6.2 - M54	22,7	N	II	6.2	Off-site Incineration/Insignificant fuel value
79	6.2 - M61	22,7	N	II	6.2	Off-site Waster Treatment (Excluding POTW)
80	6.2 - M69	22,7	N	II	6.2	Off-site Other Waste Treatment
81	6.2 - M95	22,7	N	II	6.2	Off-site Transfer to Waste Broker - Waste Treatment
82	Off-Site Treated Total	22,7	N	None	None	The sum of M40 + M50 + M54 + M61 + M69 + M95
83	Total Releases	22,7	N	None	None	The total on and off-site releases from sections 5 and 6 of the Form R. Equals On-site Release Total (field #45) + Off-site Release Total (field #66).
84	8.1 - Releases	22,7	N	II	8.1	Amount of Total On- and Off-site Releases as reported in Section 8, Source Reduction and Recycling Activities / Pollution Prevention. Reported from RY 1987 through 2002.
85	8.1a - On-site Contained Releases	22,7	N	II	8.1a	Beginning in RY 2003, the total releases in Section 8 of the form R were broken up into 4 sub categories. For this data element, facilities reported Total ON-SITE disposal to Class I Underground Injection Wells, RCRA Subtitle C landfills and other landfills.
86	8.1b - On-site Other Releases	22,7	N	II	8.1b	Beginning in RY 2003, the total releases in Section 8 of the form R were broken up into 4 sub categories. For this data element, facilities reported their other ON-SITE disposal or releases not covered in 8.1a.

Field Documentation for the "Format II Data" File

#	Field	Maximum Length	Data Type	Part	Section	Definition
				Form R Reference		
				Part	Section	
87	8.1c - Off-site Contained Releases	22,7	N	II	8.1c	Beginning in RY 2003, the total releases in Section 8 of the form R were broken up into 4 sub categories. For this data element, facilities reported Total OFF-SITE disposal to Class I Underground Injection Wells, RCRA Subtitle C landfills and other landfills.
88	8.1d - Off-site Other Releases	22,7	N	II	8.1d	Beginning in RY 2003, the total releases in Section 8 of the form R were broken up into 4 sub categories. For this data element, facilities reported their other OFF-SITE disposal or releases not covered in 8.1c.
89	8.2 - Energy Recovery On-site	22,7	N	II	8.2	Amount of Energy Recovery On-site
90	8.3 - Energy Recovery Off-site	22,7	N	II	8.3	Amount of Energy Recovery Off-site
91	8.4 - Recycling On-Site	22,7	N	II	8.4	Amount of Recycling On-site
92	8.5 - Recycling Off-Site	22,7	N	II	8.5	Amount of Recycling Off-site
93	8.6 - Treatment On-site	22,7	N	II	8.6	Amount of Treatment On-site
94	8.7 - Treatment Off-site	22,7	N	II	8.7	Amount of Treatment Off-site
95	Production Waste (8.1 thru 8.7)	22,7	N	II	8.1-8.7	The Total Production Waste Quantity. This is the summation of the quantities in Section 8.1 through 8.7 or field numbers 83 through 94.
96	8.8 - One-time Release	22,7	N	II	8.8	Quantity released to the environment as a result of remedial actions, catastrophic events, or one-time events not associated with production processing.
97	8.9 - Production Ratio	9,2	N	II	8.9	Production Ratio or Activity index.
98	Parent CO Name	60	C	I	5.1	Name of Parent Company.
99	Parent CO DB NUM	9	C	I	5.2	Parent Company's Dun & Bradstreet Number.

Appendix A: Chemical Classifications

Category 1 Metals
ANTIMONY
ANTIMONY COMPOUNDS
ARSENIC
ARSENIC COMPOUNDS
BERYLLIUM
BERYLLIUM COMPOUNDS
CADMIUM
CADMIUM COMPOUNDS
CHROMIUM
CHROMIUM COMPOUNDS (EXCEPT CHROMITE ORE MINED IN THE TRANSVAAL REGION)
COBALT
COBALT COMPOUNDS
COPPER
COPPER COMPOUNDS
LEAD
LEAD COMPOUNDS
MANGANESE
MANGANESE COMPOUNDS
MERCURY
MERCURY COMPOUNDS
NICKEL
NICKEL COMPOUNDS
SELENIUM
SELENIUM COMPOUNDS
SILVER
SILVER COMPOUNDS
THALLIUM
THALLIUM COMPOUNDS
VANADIUM COMPOUNDS
ZINC COMPOUNDS

Category 2 Metals
ALUMINUM OXIDE (FIBROUS FORMS)
ALUMINUM PHOSPHIDE
ASBESTOS (FRIABLE)
BIS(TRIBUTYLTIN) OXIDE
BORON TRICHLORIDE
BORON TRIFLUORIDE
C.I. DIRECT BLUE 218
C.I. DIRECT BROWN 95
FENBUTATIN OXIDE
FERBAM
IRON PENTACARBONYL
LITHIUM CARBONATE
MANEB
METIRAM
MOLYBDENUM TRIOXIDE
OSMIUM TETROXIDE
POTASSIUM BROMATE
SODIUM NITRITE
THORIUM DIOXIDE
TITANIUM TETRACHLORIDE
TRIBUTYLTIN FLUORIDE
TRIBUTYLTIN METHACRYLATE
TRIPHENYLTIN CHLORIDE
TRIPHENYLTIN HYDROXIDE
ZINEB

Category 3 Metals
BARIUM
BARIUM COMPOUNDS

Appendix B: Loading the "TRI Basic Data Files" into Microsoft (MS) Excel

These files contain ASCII text data with TAB delimiters. This makes them easy to load into most desktop spreadsheets, databases and processing environments.

To load these files into MS Excel do the following:

1. Save the *TRI Basic Data File* to your computer.

    For this example, the file name is **TRI_2008_AL_v08.txt**
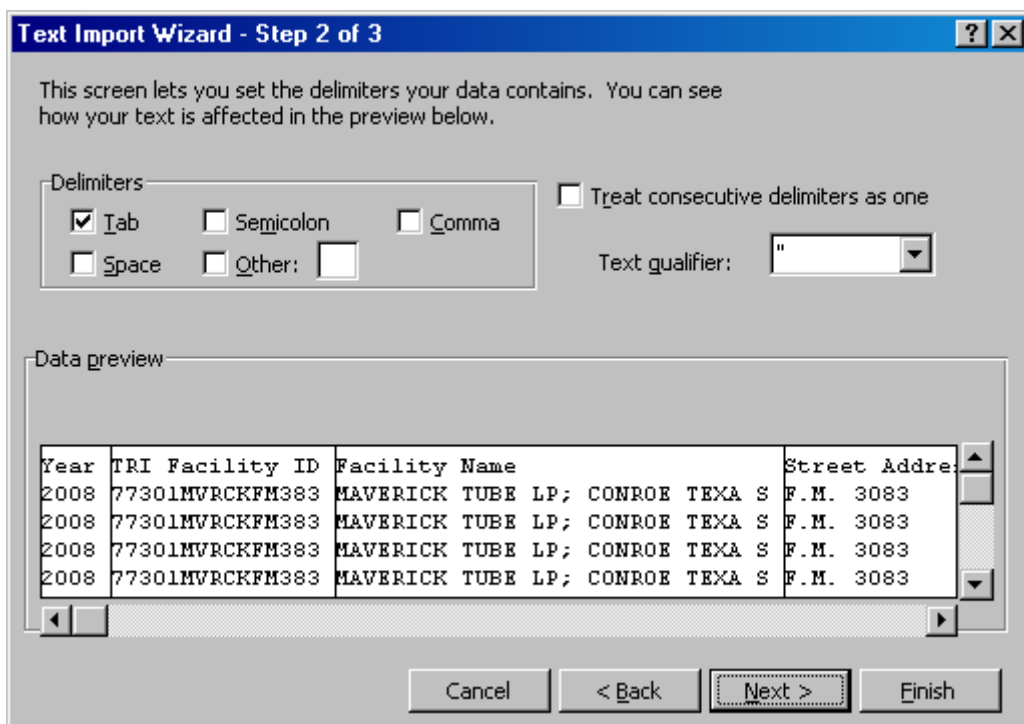   The file will be saved to the **C:\My_Dir directory**

2. Open Excel

3. From the menu bar click on "File" then "Open".

4. A pop-up box entitled "Open" will appear. At the bottom of the pop-up box, for the field "Files of Type" select "All Files (*.*)

5. Navigate to the directory where you saved the file.

6. Highlight (click on) the *TRI Basic Data File* you want to open and then click the "Open" button. See the example below where I navigate to the C:\My_Dir directory and select the TRI_2008_AL_v08.txt file to open
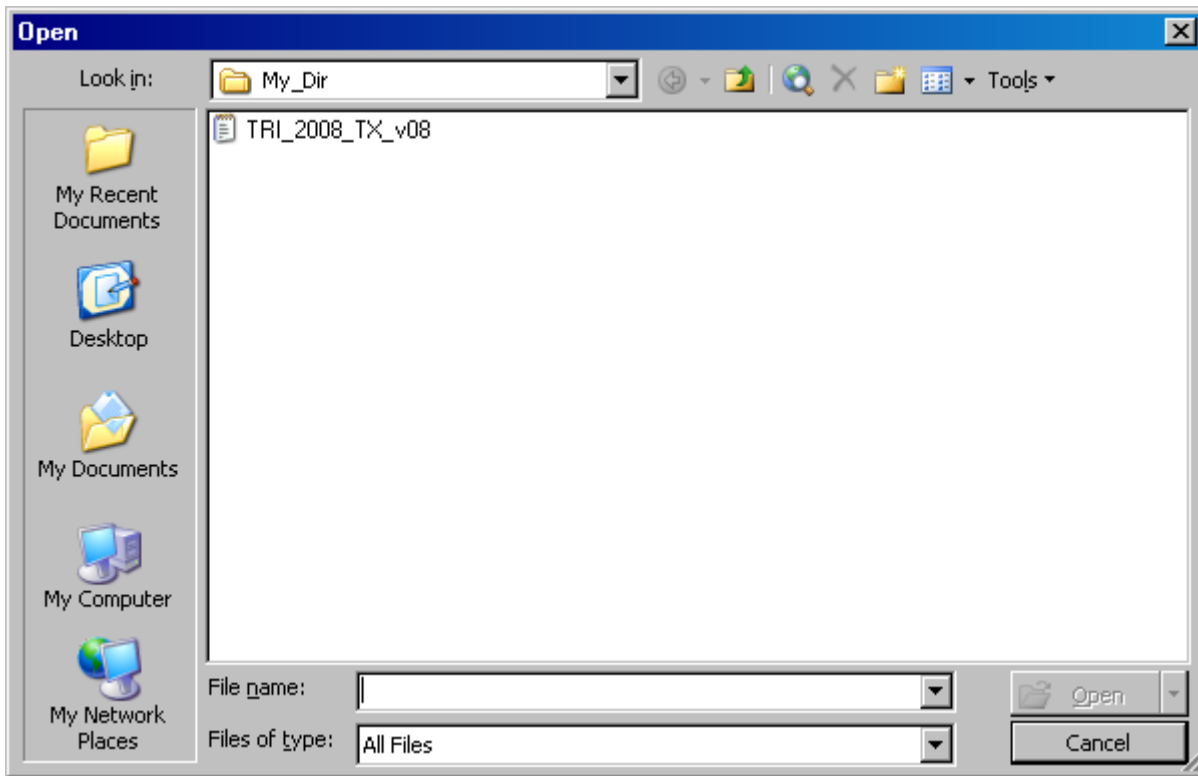
7. A new pop-up window appears entitled "Text Import Wizard – Step 1 of 3". It asks you to "Choose the file type that best describes your data". This is the first of three pop-up wizard windows that will appear. Choose, "Delimited" and click "Next". See example below.
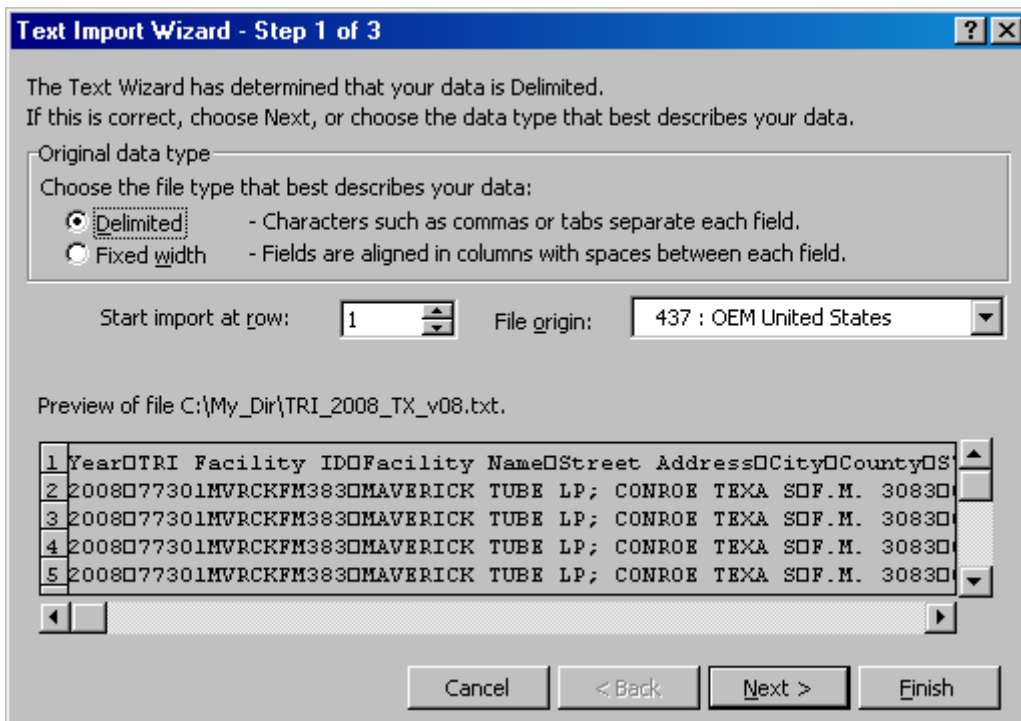
8. The next pop-up window asks you to specify the Delimiter. TAB is selected by default. The file is TAB delimited so leave the delimiter set as "TAB" Also, choose "None" as the Text qualifier. Then click "Next". See example below.

Text Import Wizard - Step 2 of 3

This screen lets you set the delimiters your data contains. You can see how your text is affected in the preview below.

Delimiters
☑ Tab   ☐ Semicolon   ☐ Comma
☐ Space   ☐ Other: [ ]

☐ Treat consecutive delimiters as one

Text qualifier: [" ▼]

Data preview

Year	TRI Facility ID	Facility Name	Street Addre
2008	77301MVRCKFM383	MAVERICK TUBE LP; CONROE TEXA S	F.M. 3083
2008	77301MVRCKFM383	MAVERICK TUBE LP; CONROE TEXA S	F.M. 3083
2008	77301MVRCKFM383	MAVERICK TUBE LP; CONROE TEXA S	F.M. 3083
2008	77301MVRCKFM383	MAVERICK TUBE LP; CONROE TEXA S	F.M. 3083

[Cancel]   [< Back]   [Next >]   [Finish]

9.   Another pop-up window (the third and final window of the wizard) appears asking you to set the data format for each column. You can leave most of the columns alone. However, you should change the "ZIP", the "Doc_Ctrl_Num" and the "CAS# / Compound Id" columns to text. To do that, scroll horizontally across until you see the "ZIP" column. It's the 8th column across. Highlight the "ZIP" column by clicking on it. It should turn black with white text. Above in the "Column Data Format" selection area click on the "Text" radio button. That's all that's needed. Now scroll across and find the "Doc_Ctrl_Num" and "CAS#/Compound Id" columns. Follow the same procedure as stated above to convert them to "Text" columns.   When you done, click the "Finish" button at the bottom of the pop-up box.   Note, if any of the other columns don't load as you think they should, try re-opening the file and change any problematic columns to text. See below for an example.

Text Import Wizard - Step 3 of 3

This screen lets you select each column and set the Data Format.

'General' converts numeric values to numbers, date values to dates, and all remaining values to text.

Column data format
- ⦿ General
- ○ Text
- ○ Date: [MDY ▼]
- ○ Do not import column (skip)

[Advanced...]

Data preview

Gener	General	General	General
Year	TRI Facility ID	Facility Name	Street Addre
2008	77301MVRCKFM383	MAVERICK TUBE LP; CONROE TEXA S	F.M. 3083
2008	77301MVRCKFM383	MAVERICK TUBE LP; CONROE TEXA S	F.M. 3083
2008	77301MVRCKFM383	MAVERICK TUBE LP; CONROE TEXA S	F.M. 3083
2008	77301MVRCKFM383	MAVERICK TUBE LP; CONROE TEXA S	F.M. 3083

[Cancel] [< Back] [Next >] [Finish]

10. The file should open up in Excel with everything formatted correctly in each column.   Technically, the file is still an ASCII Text file with TAB delimiters. You should save the file as an Excel (.xls) file so that you can re-open it without going through this whole procedure again. To do that, go up to the menu bar and click on "File" and "Save As". Navigate to the directory where you want to save the file. At the bottom of the pop-up window, choose "Microsoft Excel Workbook (*.xls)" from the drop down list as the "Save as Type". Click the "Save" button. See the example below:
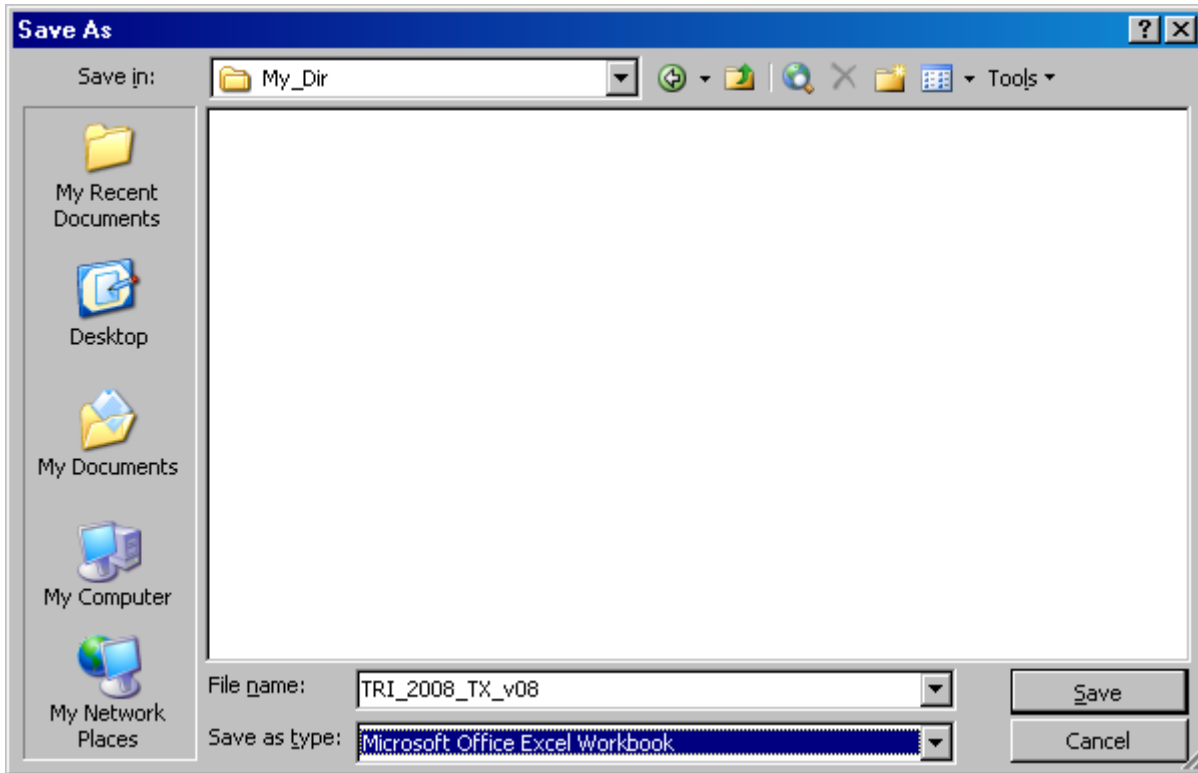
11. The loading of the *TRI Basic Data File* into Microsoft Excel is complete.

Appendix C: Loading the "TRI Basic Data Files" into Microsoft (MS) Access

1. Save the *TRI Basic Data File* to your computer.

   For this example, the file name is **TRI_2008_AL_v08.txt**
   The file will be saved to the **C:\My_Dir directory**

2. Open MS Access

3. An Access database must first exist before you can load the *TRI Basic Data File* into it. So, you either have to create a new database or open an existing database. For instructions on creating a new database go to Sep 4 If you've already created a database and want to load the *TRI Basic Data File* into it, skip to Step 5.

4. To create a new database, choose File | New from the file menu. Then choose "Blank database". A pop-up box will appear that will allow you to name the new database, navigate to a directory and save it there. Name the new database by typing a name into the "File Name" text box. In the example below, I named the database "My_Database.mdb". Then navigate to a directory and save the file by clicking the "Create" button. In the example below, I save the database to C:\My_Dir. See the example below. Once you've created the new database skip to Step 6.
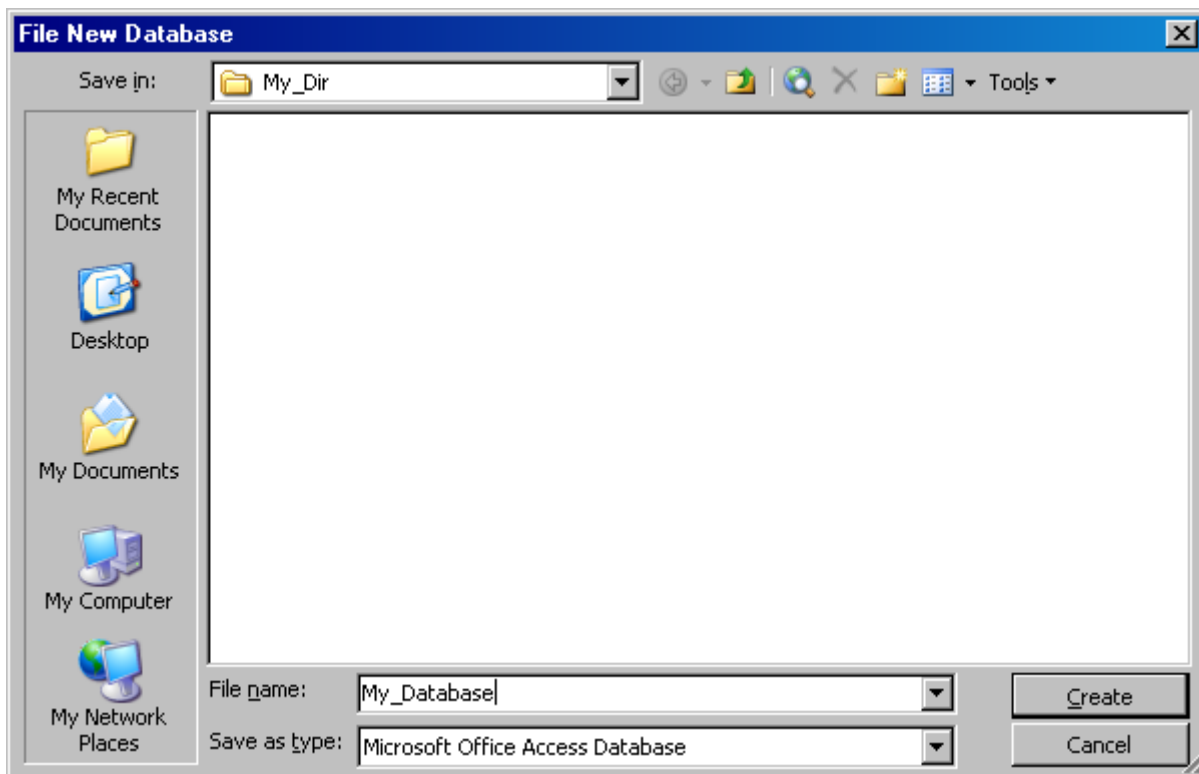
5. To open an existing database, choose File | Open from the file menu. A pop-up box will appear that will allow you to search for the database you want to open. Navigate to the correct directory and then click on the file name of the database you want to open. In the example below, I navigate to C:\My_Dir and click on the My_Database.mdb database. Finally, click the "Open" button to open the database. See the example below.
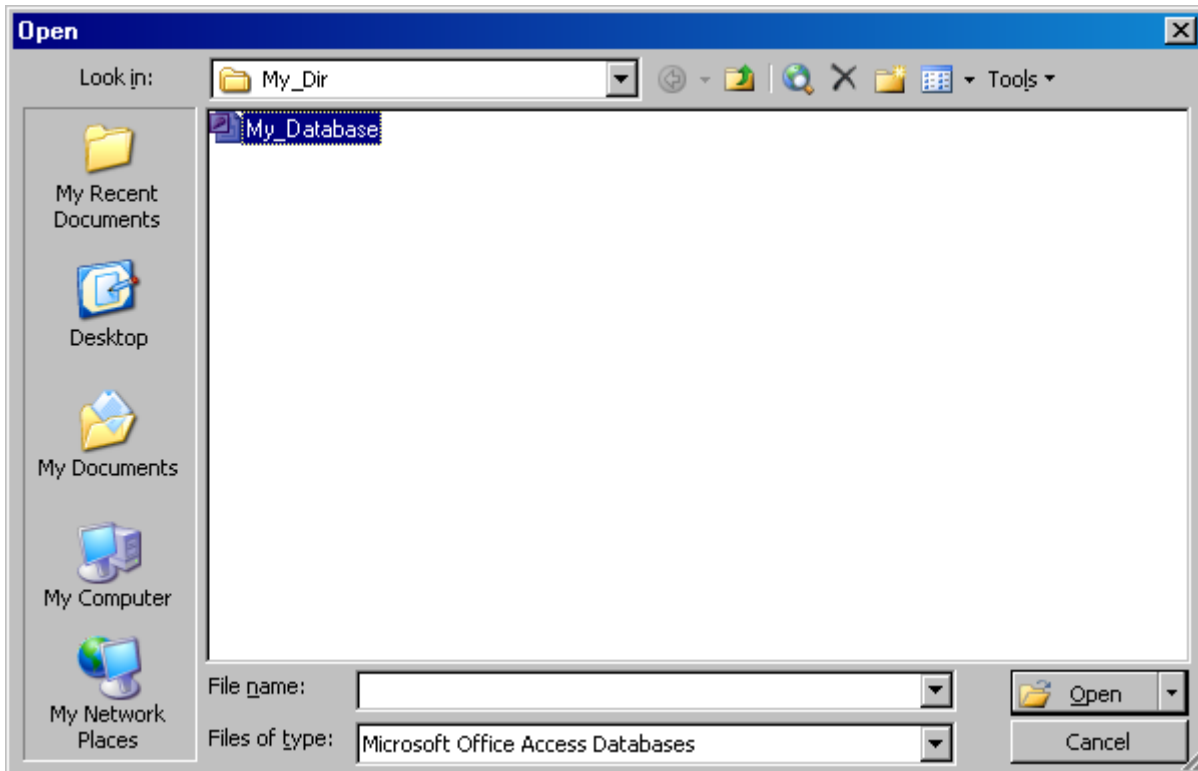

6. Now that you have you've created a new database or you have opened an existing database, you can actually load the *TRI Basic Data File*. To do this, click on "File" then "Get External Data" then "Import".

7. The "Import" window will open. Navigate to the directory where you saved the *TRI Basic Data File*. Remember to go to the bottom of the "Import" window and select "Text Files (*.txt; *.csv; *.tab; *.asc)" for the "Files of Type". The *TRI Basic Data File* you saved should now appear in the window. Click on it. Then, click the "Import" button at the bottom of the screen. See below for an example.
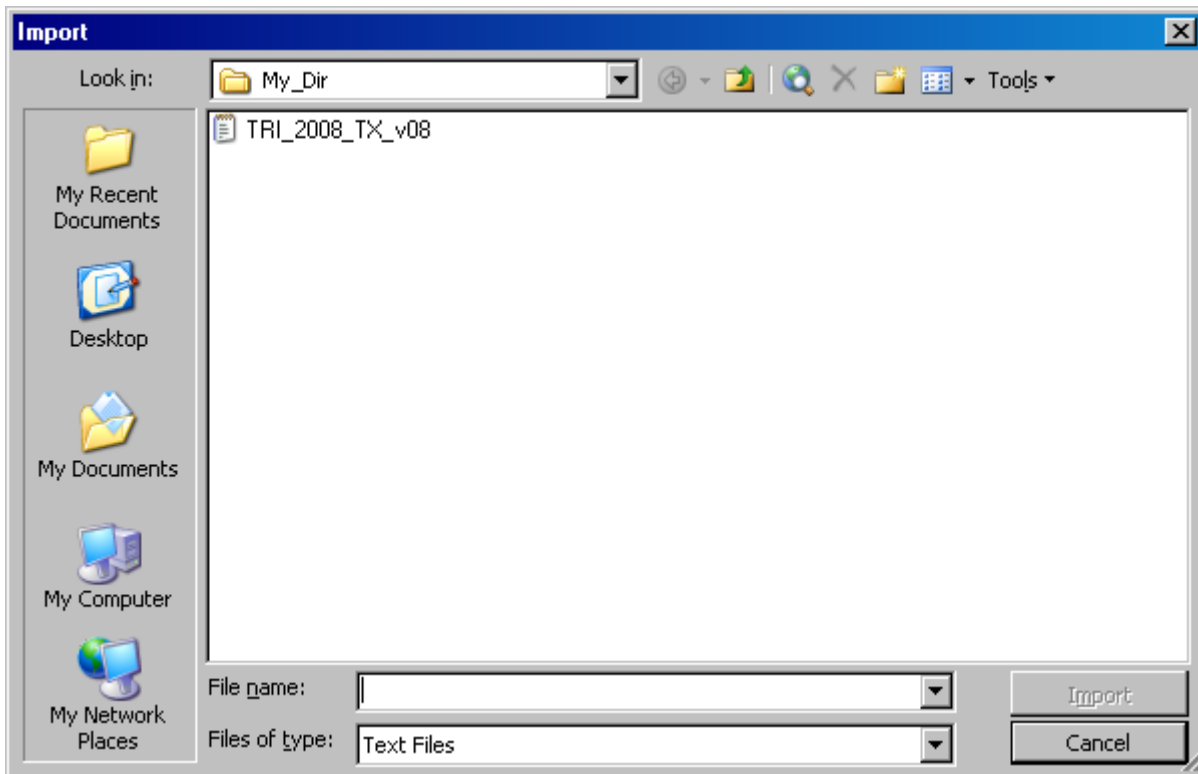
8. An "Import C Wizard" should appear. It initially asks you to choose the format that best describes that data file you want to import. Choose "Delimited". Then click the "Next" button at the bottom of the screen.
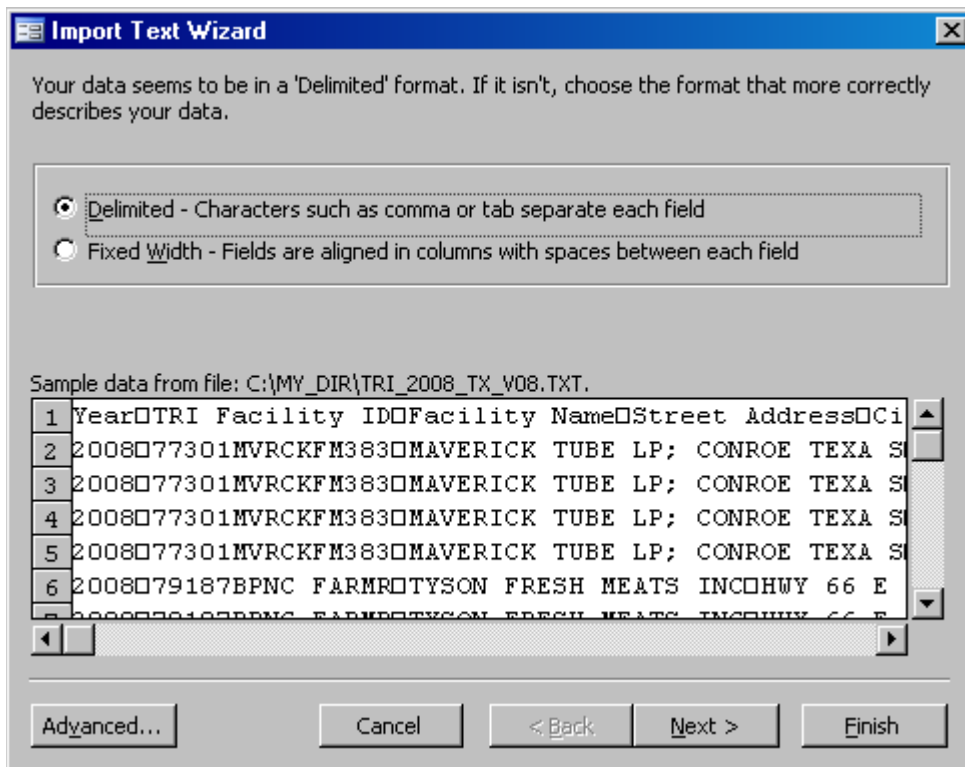
9. The next window asks you to choose the delimiter that separates your fields. Choose, "TAB". Also click the check box next to "First Row Contains Fields Names". Choose "None" as the text qualifier. It should be the default. Finally, click the "Next" button at the bottom of the screen. See the example below.
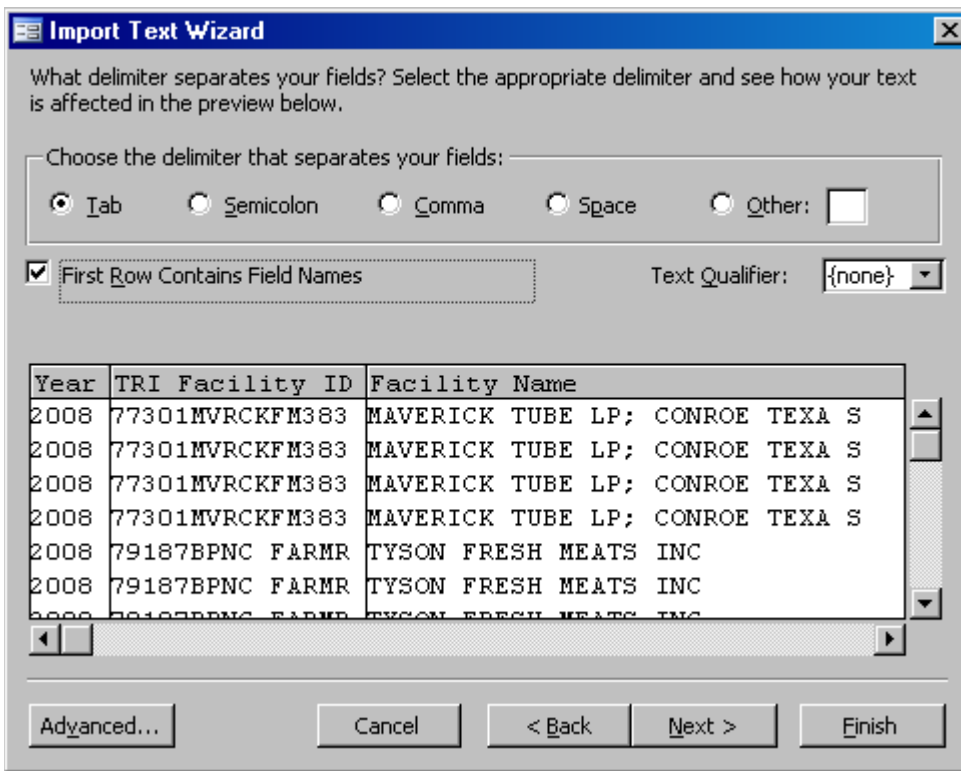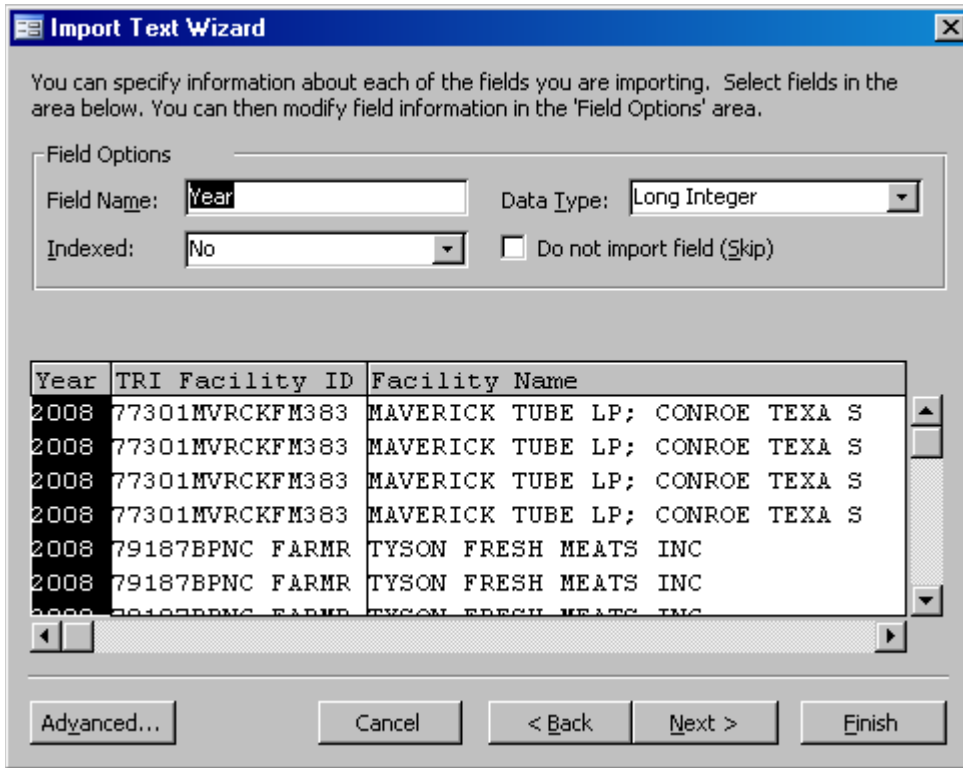

10. The next window asks you where you'd like to store your data. You can choose "In a New Table" or "In an Existing Table. For this example, I'm going to store the data into a new table. If you load multiple *TRI Basic Data Files* you many want to initially load the data into a new table, and then subsequently load other data files into the same (existing) table. Click "Next" to continue. See the example below.
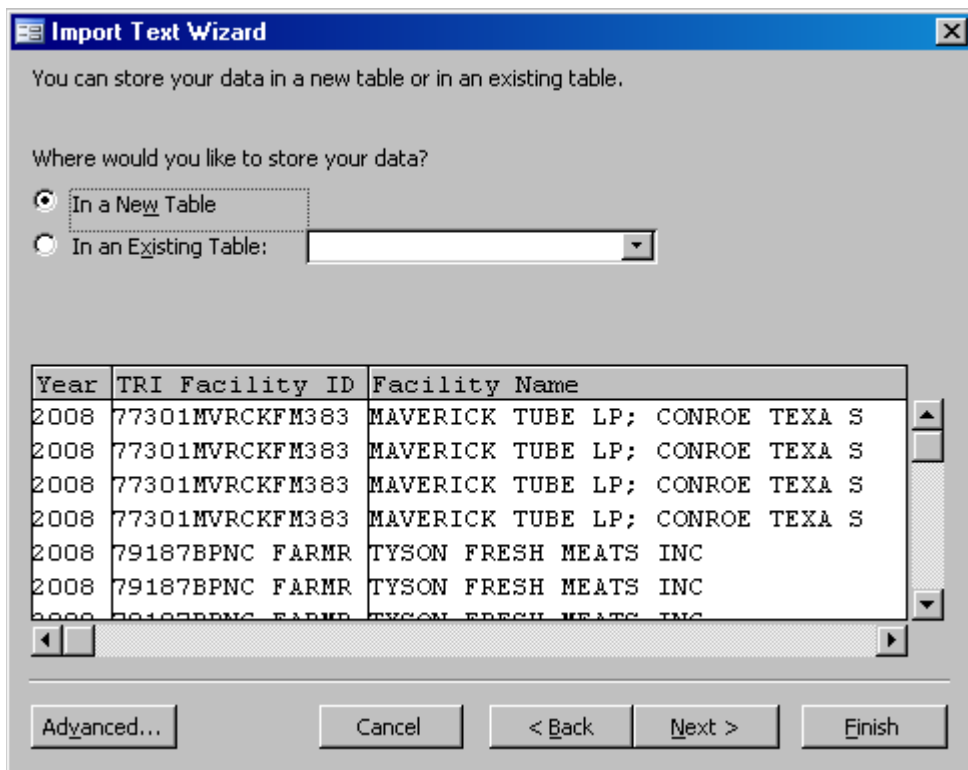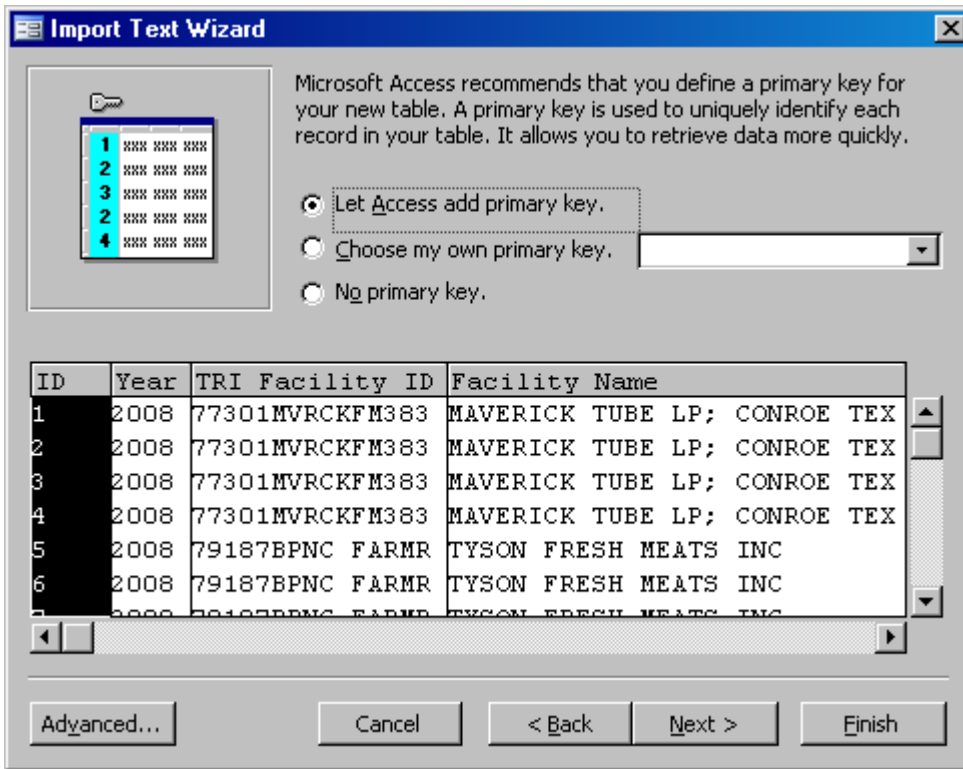
11. The final window asks you to specify information about each data field you are importing (loading) into MS Access. You must change the data type of a few fields to "Text" get them to load correctly. Some may already be shown as "Text" in Access. If they aren't, you'll have to change them. The fields are as follows:
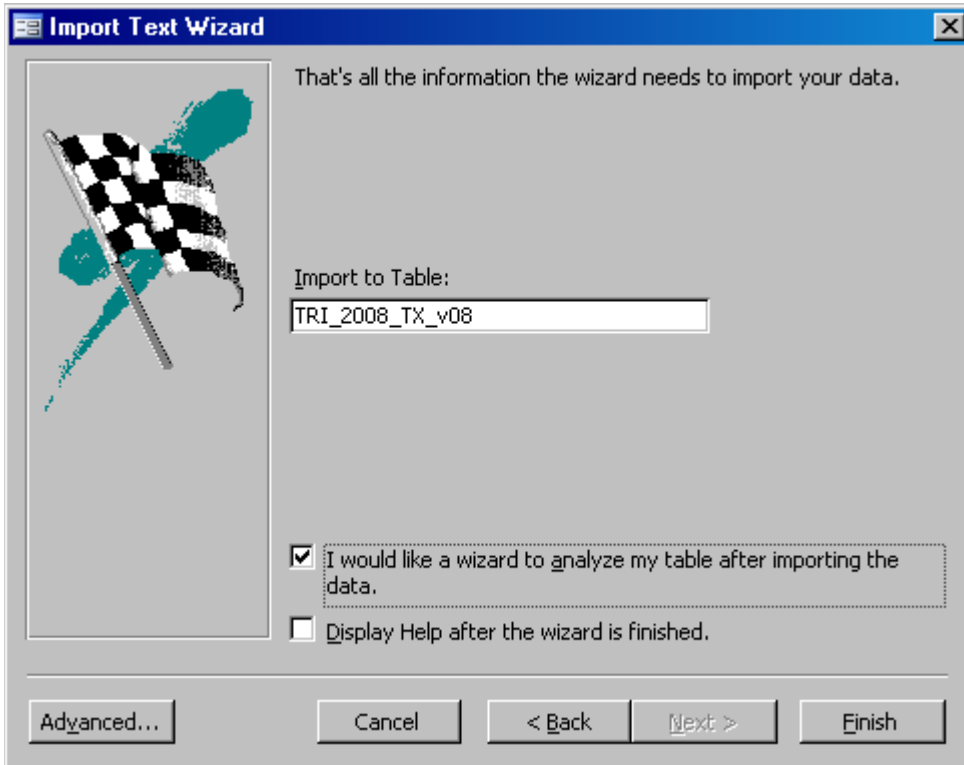
    a. ZIP
    b. All the SIC Codes {Primary SIC, SIC2, SIC3, SIC4, SIC5, SIC6}
    c. All the NAICS Codes {Primary NAICS, NAICS1, NAICS2, … NAICS6}
    d. Doc_Ctrl_Num
    e. CAS# / Compound Id

To change the data type for each field, simply use the scroll bar at the bottom of the window to locate each field (i.e. ZIP) . When it comes into view on the window, click on the column. It will turn to black with white text in it. In the top right corner of the pop-up, click on "Text" from the selections in the "Data Type" drop down. Make this adjustment for each of the fields listed above. If "Text" is already chosen for any of these columns, simply leave it alone and move on to the next column you have to change. When you've addressed each of the fields, click the "Finish" button on the bottom of the window. See below for an example. Another way to change the data types of the fields is to click the "Advanced" button and follow the instructions.

12. Microsoft Access recommends that you define a primary key for your table

Import Text Wizard

That's all the information the wizard needs to import your data.

Import to Table:

TRI_2008_TX_v08

☑ I would like a wizard to analyze my table after importing the data.

☐ Display Help after the wizard is finished.

[Advanced...]   [Cancel]   [< Back]   [Next >]   [Finish]

13. The data will load into new table that will have the same name as the input file. In this example, the new table name is TRI_2008_AL_v08.   See below for an example.
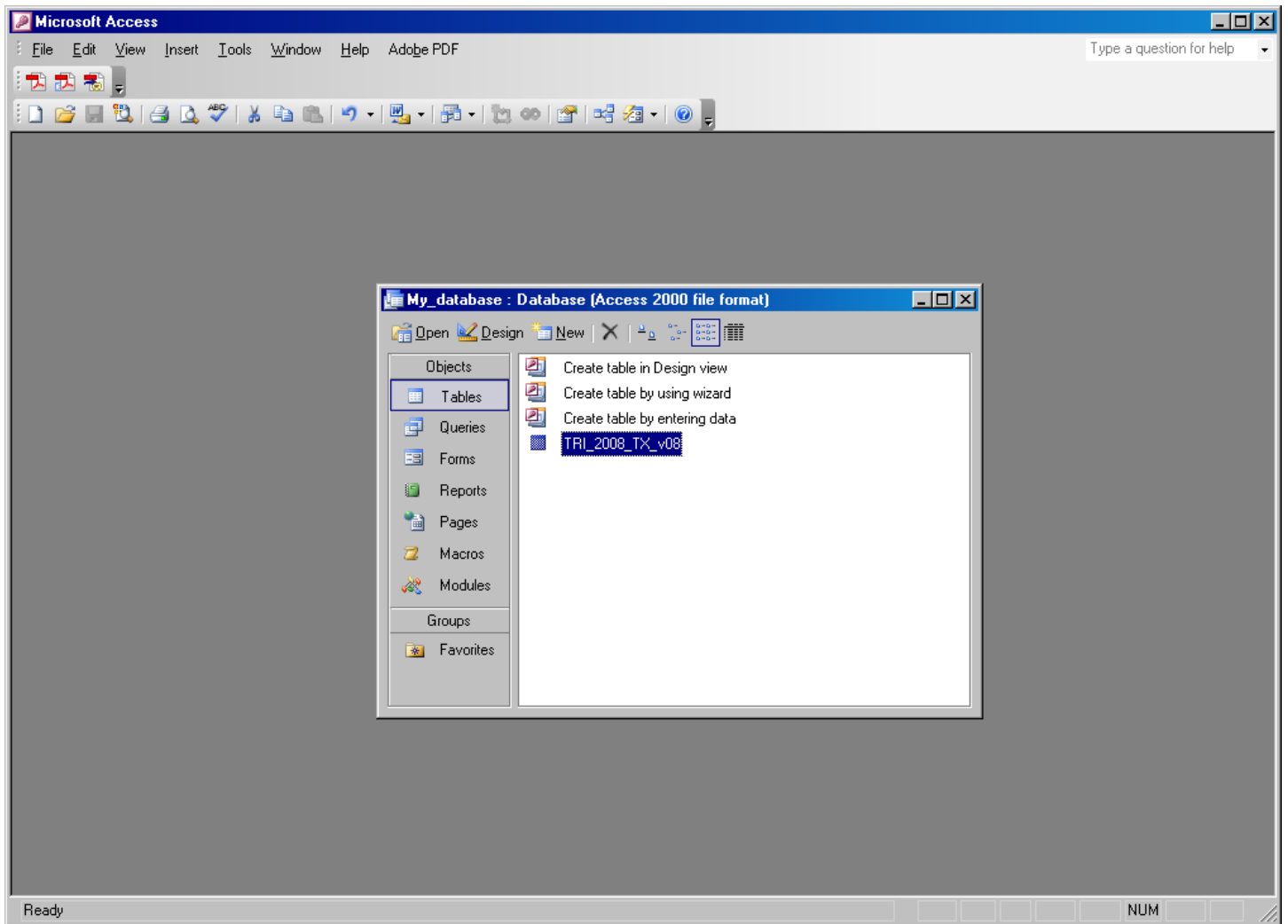
14. The loading of the *TRI Basic Data File* into Microsoft Access is complete.