



# Geographically Discrete Data: Mining the Internet to Find Authoritative Health Data

---

Helena VonVille

Library Director

The University of Texas School of Public Health at  
Houston



## Goal of CHARTing Web site

---

- Provide a comprehensive collection of links
  - Geographically discrete data for Texas
  - Illness/disease incidence/prevalence
  - Describe determinants of health
- Select from authoritative data providers
- Site stability important
- Actively update site



# Determinants of Health

---

- Income and Social Status
- Social Support Networks
- Education
- Employment/Working Conditions
- Social Environments
- Physical Environments
- Personal Health Practices and Coping Skills
- Healthy Child Development
- Biology and Genetic Endowment
- Health Services
- Gender
- Culture



## Geographically Discrete Data: Definition

---

- Also known as “localized data”
- Describes a locale at the county, city, or neighborhood level
- Dependent upon the needs of person using the data
  - Researchers in urban areas may need more discrete data than those in rural areas



# Community Health Assessment Resource for Texas (CHART)

---

- Funded by NN/LM SCR
- Geocode Texas vital statistics and Census data
- GIS-based site would be publicly available
- Library role: dissemination and education

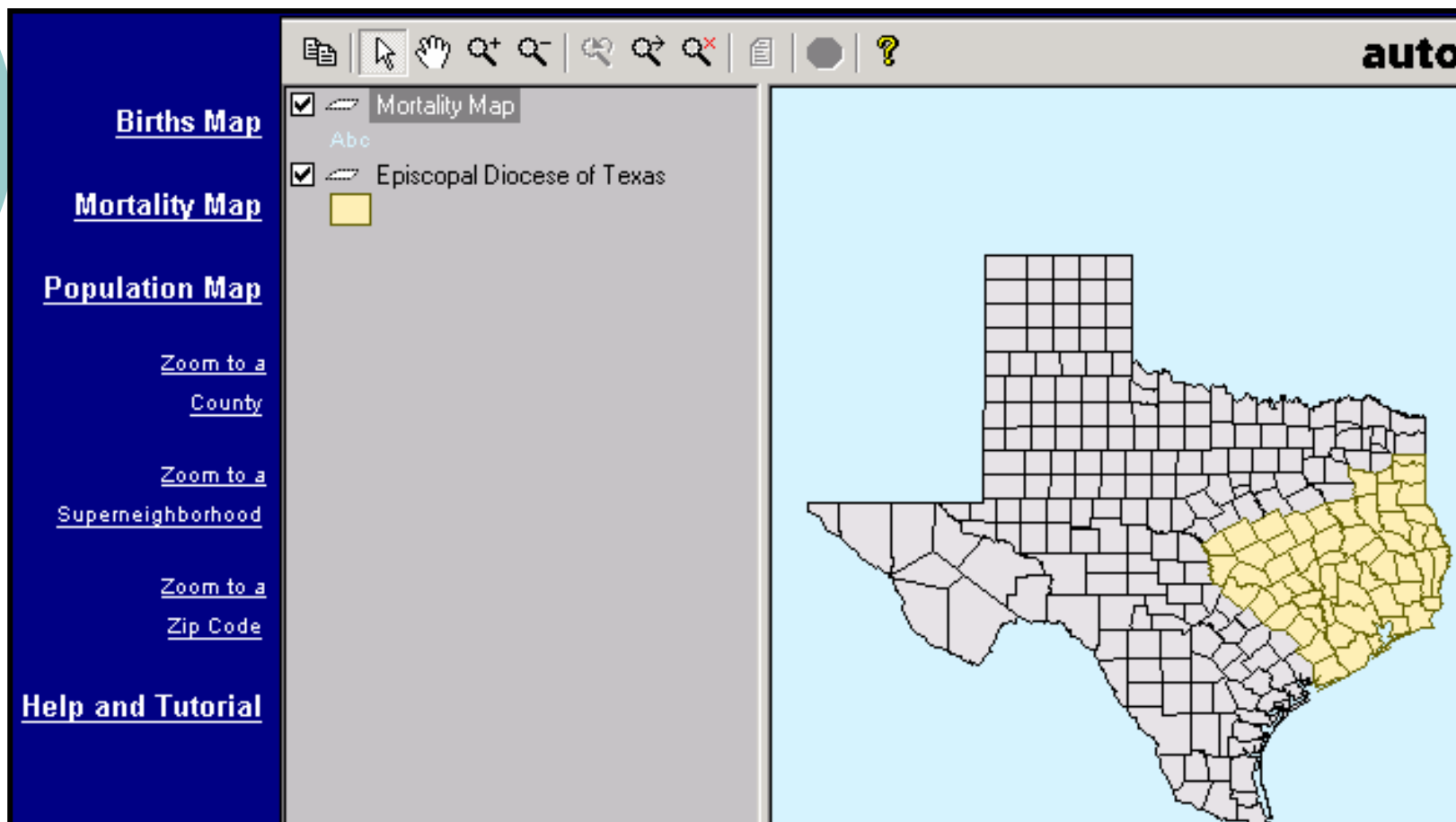


# Problems with Geocoding Data

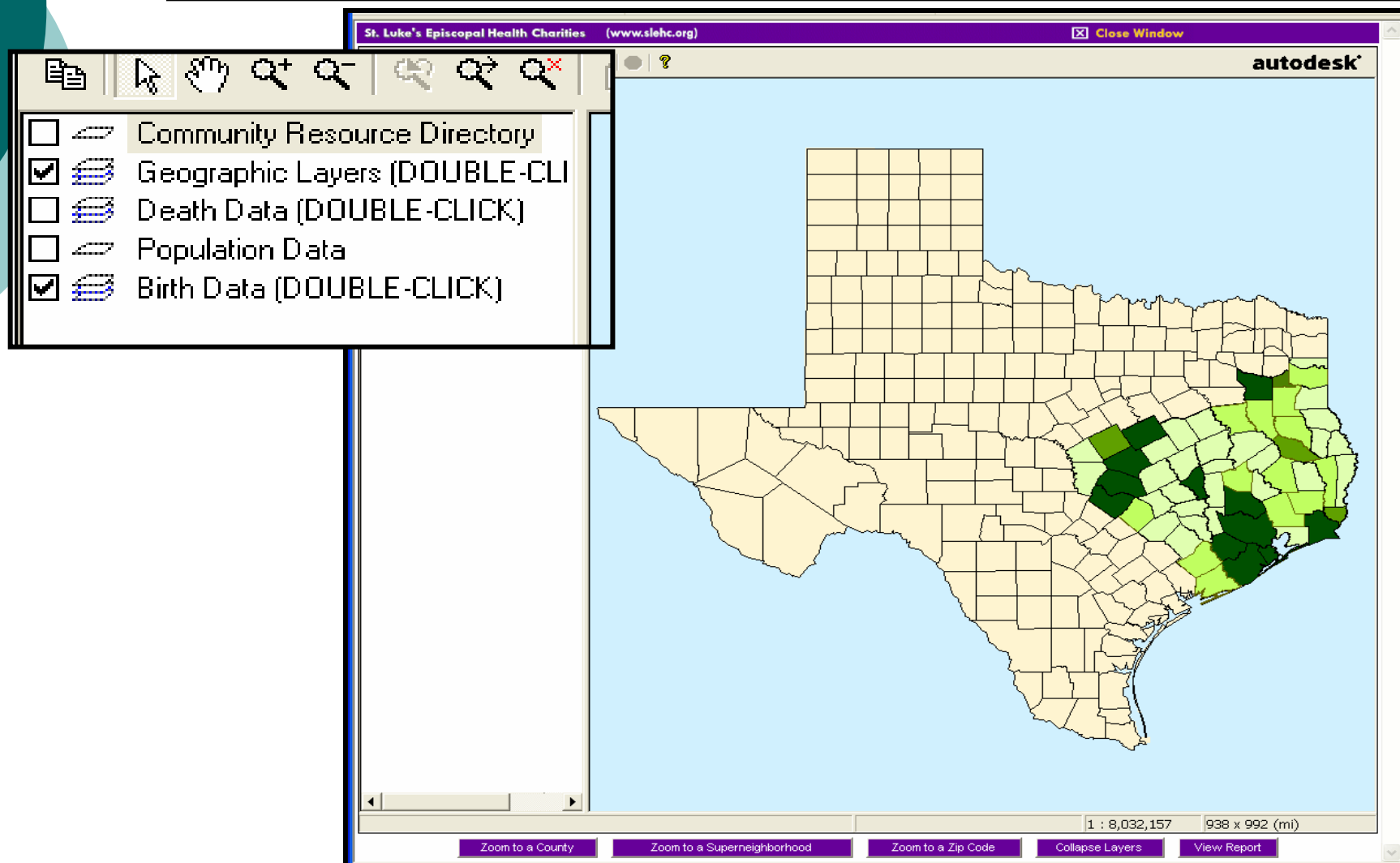
---

- Incorrect address on birth/death certificates
- Incorrect county on birth/death certificates
- Street names misspelled
  - MacGregor vs. McGregor
- Numbers transposed or read incorrectly during input process
  - 4 that looks like a 9
- Post office boxes used in rural areas
  - 80% to 90% of urban data generally matched on first sweep
  - 50 to 65% of rural data can be initially matched
- Texas is really big

# Community Health Information System (CHIS) (2003)



# Community Health Information System (CHIS) (2004)







# Original CHARTing Menu

---

- Glossary of terms
- County overviews
- Natality, mortality, and morbidity data
- Sociodemographic data
- Environmental & occupational exposure data

# County Overviews

---

[Texas Health Facts](#) (formerly County Fact Sheets; Texas Department of Health; 1990s) Good 2 page overview of the health of a **county** that includes: Demography, Social Indicators, Medicaid, Natality, Communicable Diseases, Mortality, Health Professional Facilities, and Texas Department of Health Program Summaries. Scroll down for each county.

[The Landscape Project](#) (Texas Institute for Health Policy Research; **county** level data and **Metropolitan Statistical Area** level data)

The Landscape Project creates reports based on 14 indicators: Communicable Diseases, Environmental Health, Government Finance, the Health Care Sector, Health Insurance Information, Infant and Maternal Health, Mortality Statistics, Needs Based Program Projections, Population Distribution, Social Security, and Socioeconomic Characteristics.

[Community Profiles](#) (Texas Economic Development; **city** and **county** level data)

The Community Profiles give very brief overviews of each county and include demographic, educational enrollment, farming statistics, public assistance expenditures and poverty data.



# Changes to CHARTing Menu

---

- Replaced Morbidity, Mortality, Natality with:
  - Infectious Diseases
  - Maternal Health, Childbirth, & Children
  - Chronic Diseases
  - Cancer
  - Cardiovascular Disease
  - Suicide, Accidents, & Injuries
  - Other Causes



# Texas Department of Health

---

- Birth Defects
- Behavioral Risk Factor Surveillance
- Breast & Cervical Cancer Control
- Cancer Registry
- Center for Health Statistics
- Epidemiology
- HIV/STD, AIDS
- Infectious Disease Epidemiology and Surveillance



# Morbidity Data from TDH

## Public Health Region 8 Statistical Summaries

- [Reported AIDS & HIV Infection Rates - 1999](#)
- [Reported Selected Gastrointestinal Disease Rates - 1999](#)
- [Reported Hepatitis Rates - 1999](#)
- [Reported Other Selected Disease Rates - 1999](#)
- [Reported Sexually Transmitted Disease Rates - 1999](#)
- [Reported Vaccine-Preventable Disease Rates - 1999](#)

# Morbidity Data from TDH

REPORTED OTHER SELECTED DISEASE RATES (CASES PER 100,000 POPULATION) PUBLIC HEALTH REGION 8 - 1999							
COUNTY	1999 POP.	ASEPTIC MENINGITIS		CHICKENPOX		ENCEPHALITIS	
		CASES	RATES	CASES	RATES	CASES	RATES
ATASCOSA	36,915	1	2.7	1	2.7	0	0.0
BANDERA	13,915	0	0.0	3	21.6	0	0.0
BEXAR	1,360,411	147	10.8	2	0.1	1	0.7
CALHOUN	20,099	0	0.0	0	0.0	0	0.0
COMAL	73,767	1	1.4	31	42.0	0	0.0
DEWITT	20,314	0	0.0	11	54.1	0	0.0
DIMMIT	11,251	0	0.0	41	364.4	0	0.0
EDWARDS	2,497	0	0.0	0	0.0	0	0.0



## Morbidity Data from TDH

---

[1999 Annual Epidemiology Report](#)  
(Texas Department of Health)

[1998 Annual Epidemiology Report](#)  
(Texas Department of Health)

[1997 Annual Epidemiology Report](#)  
(Texas Department of Health)

Go to **Regional Statistical Summaries** for reports by county for chicken pox, hepatitis, STD's, measles, mumps, etc. Go to **Reported Cases of Selected Diseases** for reports by public health regions.

# CDC Wonder: Natality

## Natality Data Request

[Data Use Restrictions](#)

set has counts of births occurring within the United States to U.S. residents and non-residents. Counts can be requested by state, county, child's gender and weight, maternal race, maternal age, maternal education, gestation period, prenatal care, and birth plurality. The data is derived from birth certificates. [More information.](#)

If you have trouble displaying the list of states below, click [here](#) for the screen without java applets.

### 1. Select location and time span:

[Select state, and then select county:](#)  
(Mother's State and County of Residence)

The United States

Year(s) (1995 - 2001): 1995 Through: 2001

### 2. Select maternal and child characteristics:

Gender of Child: Any

Birth Weight: 500 gms or less through: Not stated

Plurality of Birth: Any

Maternal Age: Less than 15 years through: 50-54 years

Maternal Race: Any Races

Hispanic Origin of Mother: Any Hispanic and Non-Hispanic

Maternal Education: 0 - 8 years through: Not stated

Prenatal Care Began: No Prenatal care Through: Unknown or not stated

Gestation Period: Under 20 weeks through: Not stated





# CHARTing Link to Federal Reports

---

- [Tuberculosis Surveillance](#) (Centers for Disease Control & Prevention; data available for **MSAs** and **selected cities**)

Scroll down to Tables 38 through 43 for TB morbidity by MSA and selected cities.



# Landscape Project (Texas Institute for Health Policy Research)

---

- Utilizes data from CDC, TDH, FBI, and other agencies
  - Works in collaboration with an advisory committee from state agencies, health care professionals and community representatives
  - Provides reports for counties or MSAs
  - Reports include 120 indicators in 14 categories
- Communicable Diseases
  - Crime
  - Environmental Health
  - Government Finance
  - The Health Care Sector
  - Health Insurance
  - Household Information
  - Infant and Maternal Health
  - Mortality Statistics
  - Needs Based Programs
  - Population Projections
  - Population Distribution
  - Social Security
  - Socioeconomic Characteristics



# Missing Sources for County-Level Data

---

- Non-reportable infectious diseases
  - Can find emerging disease data (SARS, WNV)
    - Extremely small incidence
- Chronic conditions
  - Asthma, diabetes, obesity, and cardiovascular disease
- Mental illness
- Current birth defects data only recently made available



## Texas Hospital Inpatient Discharges

---

- Multiple report variables may be selected to create tables
- Utilizes DRGs, not the more specific ICD-9 or ICD-10
- Public Health Region level of data

# Texas Hospital Inpatient Discharges

Hospital Discharge Statistics for Public Health Region 6																		
Year: 2000																		
Age of Patient: 0 to 17																		
Diagnosis: (4) Diseases and disorders of the respiratory system																		
Status	Payer																	
	Self-Pay		Worker's Compensation		Medicare		Medicaid		CHAMPUS and other Federal Programs		Commercial/Blue Cross		Other		Unknown		Total	
	Number	Total Length of Stay	Number	Total Length of Stay	Number	Total Length of Stay	Number	Total Length of Stay	Number	Total Length of Stay	Number	Total Length of Stay	Number	Total Length of Stay	Number	Total Length of Stay	Number	Total Length of Stay
Routine Discharge	1,011	3,079	3	8.0	14	34	4,346	18,860	33	81	3,704	13,586	166	914	55	133	9,332	36,695
Discharge to Another Facility	20	48	0	0.0	0	0.0	95	1,048	1	1.0	67	646	6	86	2	5.0	191	1,834
Discharge to Home Health Service	0	0.0	0	0.0	0	0.0	37	346	0	0.0	16	164	0	0.0	0	0.0	53	510
Left Against Medical Advice	3	9.0	0	0.0	2	2.0	16	29	0	0.0	4	11.0	0	0.0	1	3.0	26	54
Expired	0	0.0	0	0.0	0	0.0	25	423	0	0.0	14	207	5	19.0	0	0.0	44	649
Other	0	0.0	0	0.0	0	0.0	1	3.0	0	0.0	4	9.0	0	0.0	1	3.0	6	15.0
Unknown	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
<b>Total</b>	<b>1,034</b>	<b>3,136</b>	<b>3</b>	<b>8.0</b>	<b>16</b>	<b>36</b>	<b>4,520</b>	<b>20,709</b>	<b>34</b>	<b>82</b>	<b>3,809</b>	<b>14,623</b>	<b>177</b>	<b>1,019</b>	<b>59</b>	<b>144</b>	<b>9,652</b>	<b>39,757</b>

# Texas Hospital Discharge Database

**2. Enter ICD-9 Code**

250 [ICD-9 Code LookUp](#)

**Example:** Enter "250.7" for All ICD-9 Codes that begin with "250.7".

**Example:** Enter "001.0" for "CHOLERA DUE TO VIBRIO CHOLERAЕ", the specific ICD-9 Code Description for "001.0".

**3. Select Patient Gender**

Male  
 Female  
 Both Genders  
 Unknown  
 All

**4. Select Patient Age**

35 - 39 years  
 40 - 44 years  
 45 - 49 years  
 50 - 54 years  
 55 - 59 years  
 60 - 64 years  
 65 - 69 years  
 70 - 74 years  
 75 - 79 years  
 80 - 84 years

**5. Select Patient Race**

American Indian/Eskimo/Aleut  
 Asian or Pacific Islander  
 Black  
 White  
 Other  
 Any

**6. Select Patient Eth**

Hispanic Origin  
 Not of Hispanic Or  
 Any

**7. Select State/County/Zip Code of Patient Residence**

Instructions on Selecting All Counties/Zip Codes

AR	Fort Bend	77053
LA	Franklin	77053
NM	Freestone	77406
OK	Frio	77417
TX	Gaines	77441
	Galveston	77444
	Garza	77451
	Gillespie	77459
	Glasscock	77461
	Goliad	77464

**8. Select County/Hospital of Discharge in Tex**

Instructions on Selecting All Counties/Hospitals

TX	Hansford	BayCoast Medical Cen
	Hardeman	Bayou City Medical Ce
	Hardin	Bayou City Medical Ce
	Harris	Bayshore Medical Cen
	Harrison	Beacon Health Westch
	Haskell	Bellaire Medical Cente
	Hays	Bellaire Medical Cente
	Hemphill	Ben Taub General Hos
	Henderson	Casa A Special Hospit
	Hidalgo	Charter Behavioral Hea

○ Report variables include:

- Quarter/Year of discharge
- ICD-9 code
- Gender of patient
- Race/ethnicity
- Patient's state, county, zip code
- Hospital state, county, zip

# Texas Hospital Discharge Database

- If 3-digit ICD-9 used, long list of 4- and 5-digit ICD-9's returned
- Final result at bottom of page

Diabetes Research Center  
Texas Tech University Health Sciences Center at El Paso  
4800 Alberta Ave.  
El Paso, TX 79905

Home Types of Diabetes News Areas of Excellence Resources Data and Statistics TTUHSC - El Paso

Bookmark Us - Contact Us - Mission Statement - Privacy Policy - Site Map - Site Requirements - Suggestions

**Texas Hospital Discharge Database Query Results**

**ICD-9 Code(s) Selected:**  
250.3 : DIABETES WITH OTHER COMA  
250.43: DIABETES MELLITUS WITH RENAL MANIFESTATIONS, TYPE I, UNCONTROLLED  
250.91: DIABETES MELLITUS WITH UNSPECIFIED COMPLICATION, TYPE I NOT STATED AS UNCONTROLLED  
250.11: DIABETES MELLITUS WITH KETOACIDOSIS, TYPE I NOT STATED AS UNCONTROLLED  
250.01: DIABETES MELLITUS WITHOUT MENTION OF COMPLICATION, TYPE I, NOT STATED AS UNCONTROLLED  
250.13: DIABETES MELLITUS WITH KETOACIDOSIS, TYPE I, UNCONTROLLED  
250.2 : DIABETES WITH HYPEROSMOLARITY  
250.52: DIABETES MELLITUS WITH OPHTHALMIC MANIFESTATIONS, TYPE II OR UNSPECIFIED TYPE, UNCONTROLLED  
250.53: DIABETES MELLITUS WITH OPHTHALMIC MANIFESTATIONS, TYPE I, UNCONTROLLED  
250.6 : DIABETES WITH NEUROLOGICAL MANIFESTATIONS  
250.00: DIABETES MELLITUS WITHOUT MENTION OF COMPLICATION, TYPE II OR UNSPECIFIED TYPE, NOT STATED AS UNCONTROLLED  
250.22: DIABETES MELLITUS WITH HYPEROSMOLARITY, TYPE II OR UNSPECIFIED TYPE, UNCONTROLLED  
250.60: DIABETES MELLITUS WITH NEUROLOGICAL MANIFESTATIONS, TYPE II OR UNSPECIFIED TYPE, NOT STATED AS UNCONTROLLED  
250.61: DIABETES MELLITUS WITH NEUROLOGICAL MANIFESTATIONS, TYPE I NOT STATED AS UNCONTROLLED  
250.73: DIABETES MELLITUS WITH PERIPHERAL CIRCULATORY DISORDERS, TYPE I, UNCONTROLLED  
250.82: DIABETES MELLITUS WITH OTHER SPECIFIED MANIFESTATIONS, TYPE I, UNCONTROLLED

**Discharge Quarter(s) Selected:**  
1999: No discharge dates for 1999 were selected.  
2000: No discharge dates for 2000 were selected.  
2001: Jan. 2001 - Mar. 2001  
Apr. 2001 - June 2001  
July 2001 - Sept. 2001  
Oct. 2001 - Dec. 2001

**Patient Gender Selected:** All Genders  
**Patient Age Selected:** 65 - 69 years  
**Patient Race Selected:** Any Race  
**Patient Ethnicity Selected:** Any Ethnicity

**Patient Residence Location Selected:**  
**State:** Texas  
**County:** Fort Bend  
**Zip Code:** All Zip Codes in Fort Bend were selected.

**Discharge Hospital Selected:**  
**State:** Texas  
**County:** Harris  
**Hospital:** All Hospitals in Harris were selected.

**Number of Encounters:**  
60

[Perform another query](#)

Website developed and maintained by TTUHSC at El Paso Department of Information Technology. Copyright © 2003.  
Texas Hospital Inpatient Discharge Public Use Data File, 1st quarter, 1999 - 4th quarter, 2001.  
Texas Health Care Information Council, Austin, Texas, 2003.

**Discharge Hospital Selected:**

**State:**

Texas

**Number of Encounters:**

60

[Perform another query](#)

# Developing Context: What's Wrong with This Data?

As you view these links to various sources of health data, it is important to keep in mind that data does not tell the whole story. Data and data gathering are not perfect-- if you see an anomaly or large deviation in the data, find out why. Don't assume.

## 1. Do you know what data is really gathered?

Births for Harris County								
Marital Status								
	Married		Not Married		Unknown		All Mothers	
Year	Number	Percent	Number	Percent	Number	Percent	Number	Percent
1992	45,851	78.2	12,801	21.8	0	0.0	58,652	100.0
1993	<b>45,873</b>	<b>78.8</b>	<b>12,332</b>	<b>21.2</b>	0	0.0	58,205	100.0
1994	<b>39,130</b>	<b>68.6</b>	<b>17,691</b>	<b>31.0</b>	205	0.4	57,026	100.0
1995	38,208	67.0	18,783	32.9	66	0.1	57,057	100.0

*(Texas Health Data Births to Texas Residents)*



# 2004 Changes to CHARTing

---

## ○ Partnership Links

- National Network of Libraries of Medicine South Central Region
- Community Health Assessment Resource for Texas
- Texas State Strategic Health Partnership [Data context](#)
- Texas Public Health Training Center



South Central Region



## ○ New pages added (“Other Sources of Data”)

- Access to Health Care
- Criminal Justice Data
- Regional Web Sites

# CHART Today

The screenshot displays the CHART web application interface. At the top left is the logo for The University of Texas School of Public Health at Houston. The main header reads "Center for Health Policy Studies" and "Community Health Informatics". Below this is the title "The Community Health Assessment Resource for Texas (CHART)". A sidebar on the left contains "Important Links" and "CHART Links" dropdown menus, and a link for "First Time Users Click Here". The main content area features a map of Texas with county boundaries. A legend on the left of the map lists "Click Here to See Texas County Reports" with checkboxes for years 2001 through 1996. The year 2001 is checked. A dark blue square highlights a county on the map. The Autodesk logo is visible in the top right of the map area. A "View Report" dialog box is open in the foreground, listing available reports for 2001, including birth and mortality reports by various demographic factors. The dialog box has "OK", "Cancel", and "Help" buttons. At the bottom of the map area, a status bar shows "reports' selected 1 : 3,519,888 311 x 257 (mi)".

**View Report**

Available reports:

- 2001 County Birth Report- Births by Birthweight
- 2001 County Birth Report- Births by Mother's Age
- 2001 County Birth Report- Births by Mother's Edu.
- 2001 County Birth Report- Births by Mother's Race/Eth.
- 2001 County Birth Report- Births by Prenatal Care
- 2001 Texas Mortality Report- Child Deaths by Ethnicity
- 2001 Texas Mortality Report- Deaths by Age
- 2001 Texas Mortality Report- Deaths by Cause
- 2001 Texas Mortality Report- Deaths by Ethnicity
- 2001 Texas Mortality Report- Infant Deaths by Ethnicity

reports' selected 1 : 3,519,888 311 x 257 (mi)

# CHART Today

## 2001 Birth Report

### Births by Trimester of Prenatal Care and Ethnicity

#### 2001 Number of Births by Trimester of Prenatal Care and Ethnicity

COUNTY	1st Trimester Prenatal Care	2nd Trimester Prenatal Care	3rd Trimester Prenatal Care	No Prenatal Care	TOTAL
Milam	304	39	8	8	359
<b>TOTAL</b>	<b>304</b> 84.7%	<b>39</b> 10.9%	<b>8</b> 2.2%	<b>8</b> 2.2%	<b>359</b> 100.0%

#### 2001 Number of Births by Trimester of Prenatal Care and Ethnicity

##### White

COUNTY	1st Trimester Prenatal Care	2nd Trimester Prenatal Care	3rd Trimester Prenatal Care	No Prenatal Care	TOTAL
Milam	179	20	0	5	204
<b>TOTAL</b>	<b>179</b>	<b>20</b>	<b>0</b>	<b>5</b>	<b>204</b>

# Possible New Layout

Injury, or  children   Exposure	<a href="#">Morbidity (Incidence)</a>	<a href="#">Mortality (Deaths)</a>
	<h2>Morbidity (Incidence)</h2> <p>For an overview of data covering incidence of STD's, hepatitis, and chicken pox, see <a href="#">County Overviews</a>.</p>	
	<a href="#">Texas Hospital Discharge Database</a> Texas Tech University Health Science Center at El Paso <b>zip code</b> (for patients) and <b>hospital level data</b> 1999 - 2001	<b>Select:</b> <b>Includes:</b> All ICD-9 codes <b>Other Information:</b> Interactive database. To find your ICD-9 code, use the CDC's mortality worktable <a href="#">1998 GMWKI</a> (pdf document, see pages 1 - 110). Report variables include: gender, age (including neonatal and infant patients), patient race, ethnicity, county and/or zip code of patient residence, and county and/or hospital of patient discharge. For other discharge information, see the <a href="#">Texas Hospital Inpatient Discharges</a> .
	<a href="#">Selected Vaccine-Preventable Disease Statistics by County</a> Texas Department of Health <b>County level data</b> 1999 to 2002	<b>Select:</b> Your public health region for county reports. <b>Includes:</b> Chickenpox; Hepatitis B; Haemophilus influenzae type b (Hib); measles; mumps; pertussis (whooping cough); rubella; and tetanus. Additional reports are available from the <a href="#">Immunization Division</a> . <b>Other Information:</b>
<a href="#">Diseases from the Zoonosis Control Division</a>	<b>Select:</b> "Regional Statistical Summaries of Public Health Regions" for county reports	



# Conceptual Framework

---

- Hillemeier, M.M., Lynch, J., Harper, S., and Casper, M. 2003.  
“Measuring Contextual Characteristics for Community Health.” *HSR: Health Services Research* 38 (6, Pt. II): 1645-1717)



---

To visit the CHARTing site or to contact me about the CHARTing challenge:

<http://www.sph.uth.tmc.edu/library/chartinghealthinfo.htm>

**Thank you!**