

BMI 6203 Final Project:

Using Geographic and Demographic Health  
Information to Determine Life Insurance Cost

---

Nic Bitter, Patricia Girardi, Zachary Shepelak, Truc Vuong

# Agenda

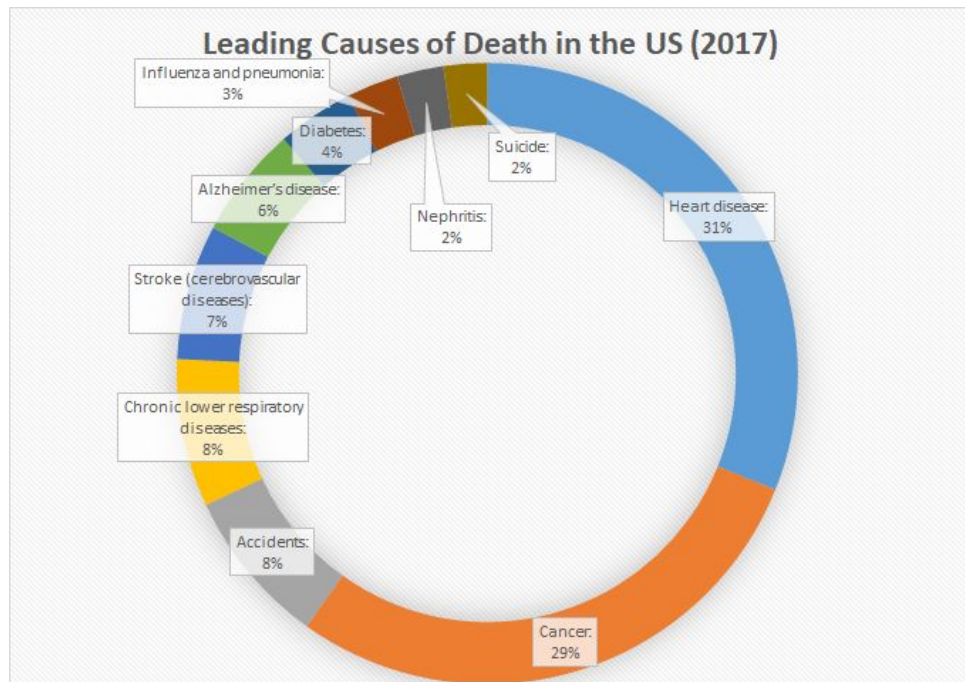
- Introduction
- Business description
- Data requirements
- Entity relationship diagram
- Relational data model
- Physical model
- Database view and queries
- Questions



<https://politicalgraffiti.wordpress.com/2009/07/23/health-insurance-caesar/>

# Introduction

- Heart disease is the leading cause of death in the US.
- There is a robust link between obesity and cancer rates.
  - Obese individuals are 2x as likely to develop cancer of esophagus, liver, and kidney, especially.
- Cancers linked to tobacco use make up 40% of new cancer diagnoses.
- In 2009, Medicare paid \$55 million for bills incurred in the last 2 months of patients' lives.



<https://www.cdc.gov/nchs/fastats/leading-causes-of-death.htm>

# Business Description

- Insurance company wants to gather data on mortality and health risk factors.
- Population is also considered since more rural areas carry higher chance of mortality
- This data is used to determine the cost of the premium.
  - Applicants may provide inaccurate information about their lifestyle to try to reduce their premium's cost.
- We created a database that aggregates data from the CDC on health risk factors and mortality rate at county/state resolution.

## 4 IN 10 CANCER CASES CAN BE PREVENTED...






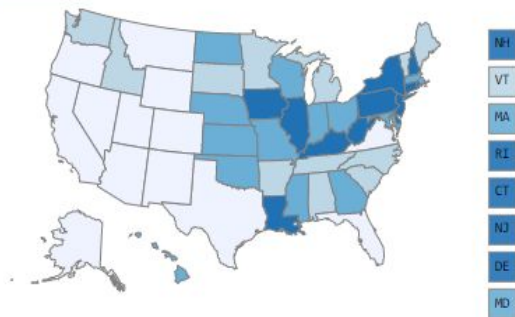
# Business Description

- The basic idea is to combine several data sets into a database that has county-level resolution for various health risk factors and cancer mortality rates.

## Rate of New Obesity-associated Cancers by State

All Obesity-associated Cancers, Male and Female, United States, 2012-2016  
Rate per 100,000 people

 Map  Table  Export



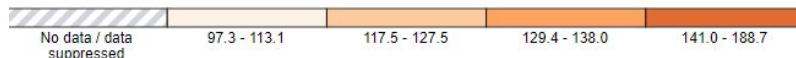
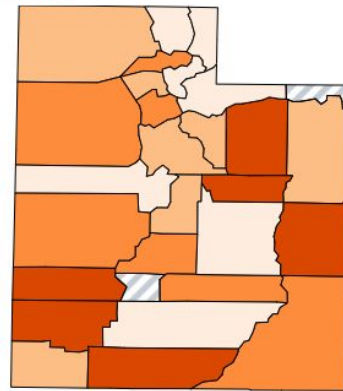
Rate per 100,000 people



## Rate of Cancer Deaths in Utah

All Types of Cancer, All Ages, All Races/Ethnicities, Male and Female, 2012-2016  
Rate per 100,000 people

 Map  Table  Export



<https://gis.cdc.gov/Cancer/USCS/DataViz.html>

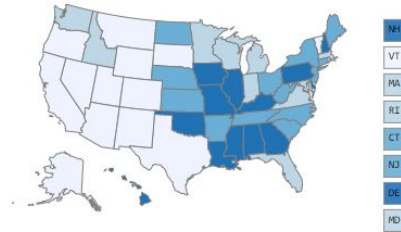
# Data Requirements

- Geographic/demographic data is used to approximate risk information for each group.
  - At county level: tobacco use, heart disease, and cancer.
  - At state level: obesity and heart disease.
- The database is designed so that it is easy to add to it as more risk information becomes available.
- Risk information is then analyzed and processed into a “policy” risk factor, which informs the percent added to the premium.
- The monthly premium is calculated as a percent added to premium, (\$67 = base premium), and then weighted inversely proportional to the population.

## Rate of New Alcohol-associated Cancers by State

All Alcohol-associated Cancers, Male and Female, United States, 2016  
Rate per 100,000 people

 Map  Table  Export



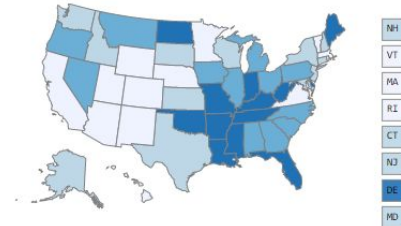
Rate per 100,000 people



## Rate of New HPV-associated Cancers by State

All HPV-associated Cancers, Male and Female, United States, 2016  
Rate per 100,000 people

 Map  Table  Export

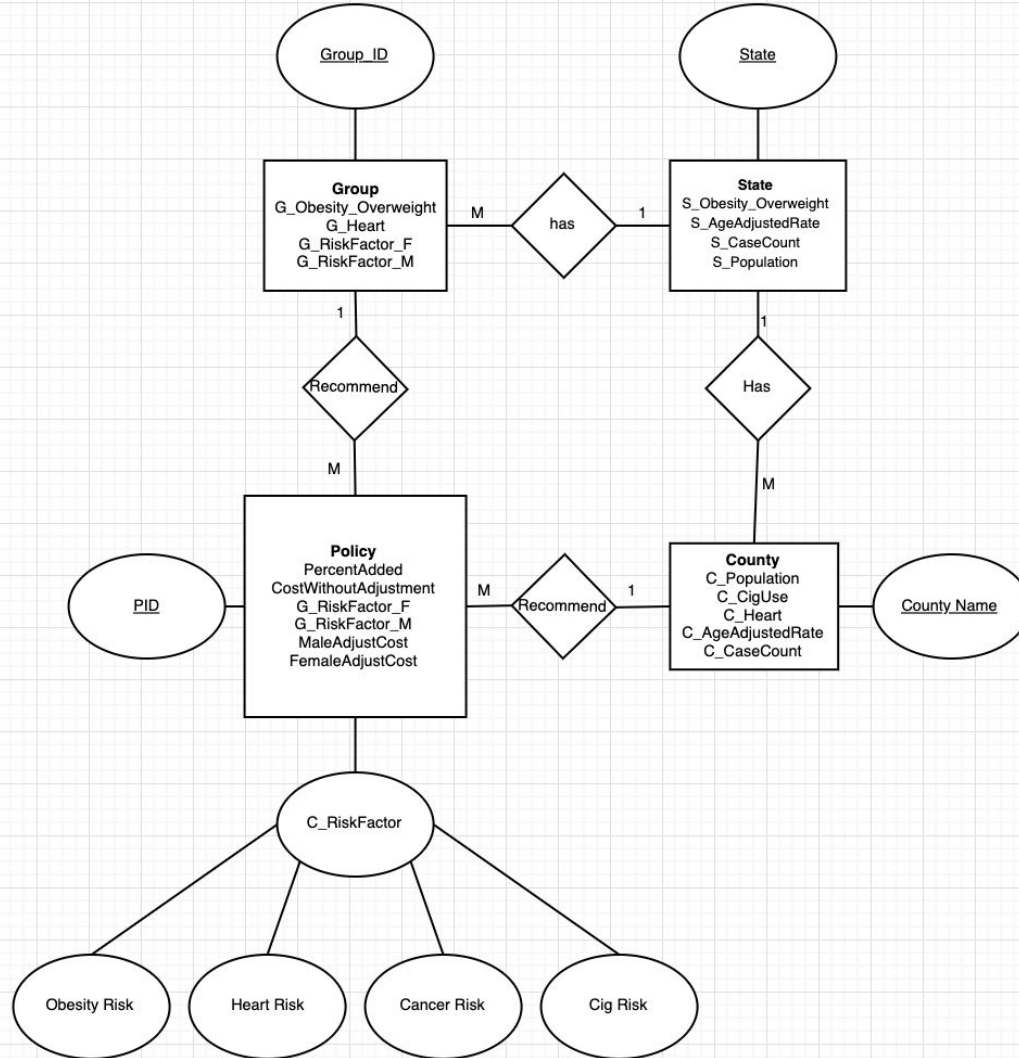


Rate per 100,000 people

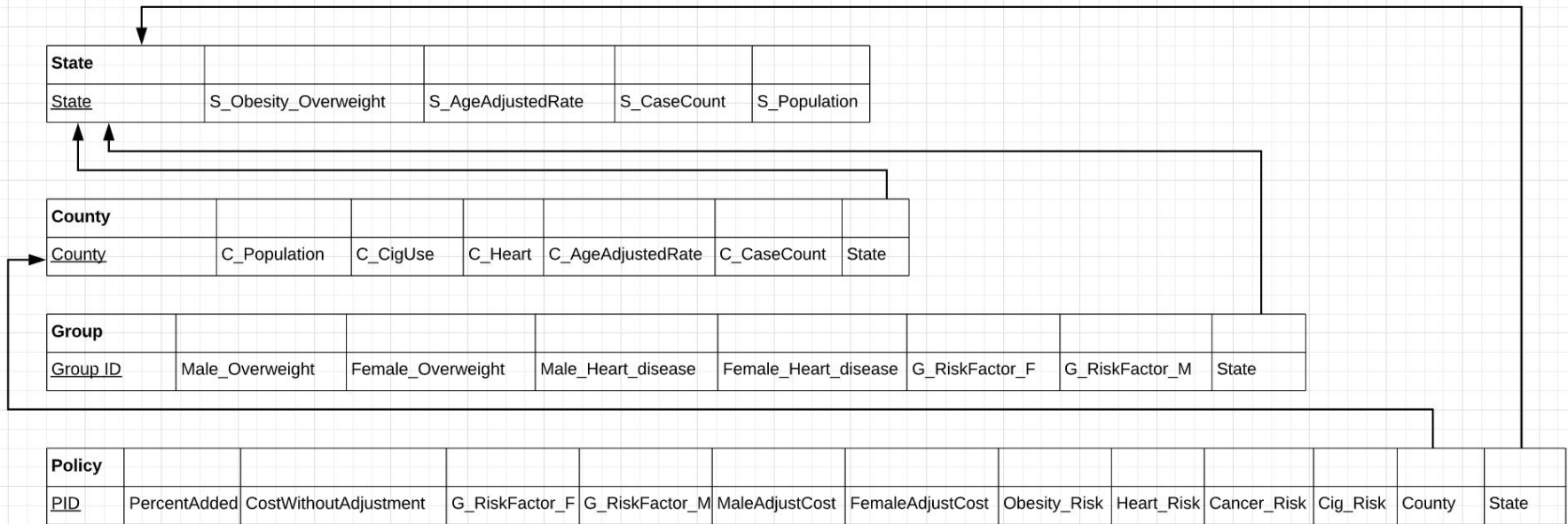


<https://gis.cdc.gov/Cancer/USCS/DataViz.html>

# Entity Relationship Diagram



# Relational Data Model





# Data Import using Wizard

Table Data Import

Select Destination

Select destination table and additional options.

☐ Use existing table:

☒ Create new table:  .

☐ Drop table if exists

Table Data Import

Configure Import Settings

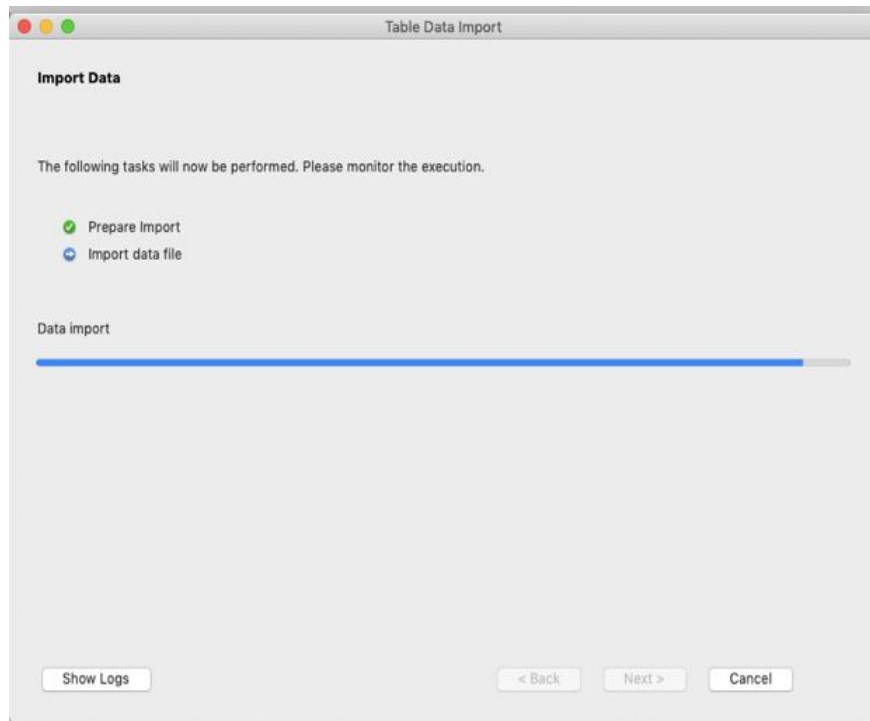
Detected file format: csv

Encoding:

<input checked="" type="checkbox"/> Source Column	Field Type
<input checked="" type="checkbox"/> State	<input type="text" value="text"/>
<input checked="" type="checkbox"/> AgeAdjustedRate	<input type="text" value="double"/>
<input checked="" type="checkbox"/> CaseCount	<input type="text" value="int"/>
<input checked="" type="checkbox"/> Population	<input type="text" value="int"/>
<input checked="" type="checkbox"/> Obesity Rate	<input type="text" value="double"/>

State	AgeAdjuste...	CaseCount	Population	Obesity Rate
New Mexi...	374.3	45580	10419654	32.3
Arizona	387	152239	33577674	29.5
Colorado	401.4	113506	26761747	23
Nevada	392.6	62045	14192998	29.5

# Data Import using Wizard (continued)



# Key Selection

CHEMAS

Filter objects

bmiu0259362

Tables

County Table

Group Table

Policy Table

State Table

Views

Stored Procedures

Functions

newinga

Name: State Table

Schema: bmiu0259362

Column	Datatype	PK	NN	UQ	BIN	UN	ZF	AI	G	Default / Expression
State	VARCHAR(35)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
AgeAdjusted...	DOUBLE	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
CaseCount	INT(11)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
Population	INT(11)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
Obesity Rate	DOUBLE	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
<click to edit>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

Column details 'State'

Column Name: State

Charset/Collation: Default Charset

Comments:

Datatype: VARCHAR(35)

Default

Storage: ☐ VIRTUAL ☐ STORED

☒ Primary Key

☒ Not NULL

☐ Unique

☐ Binary

☐ Unsigned

☐ ZeroFill

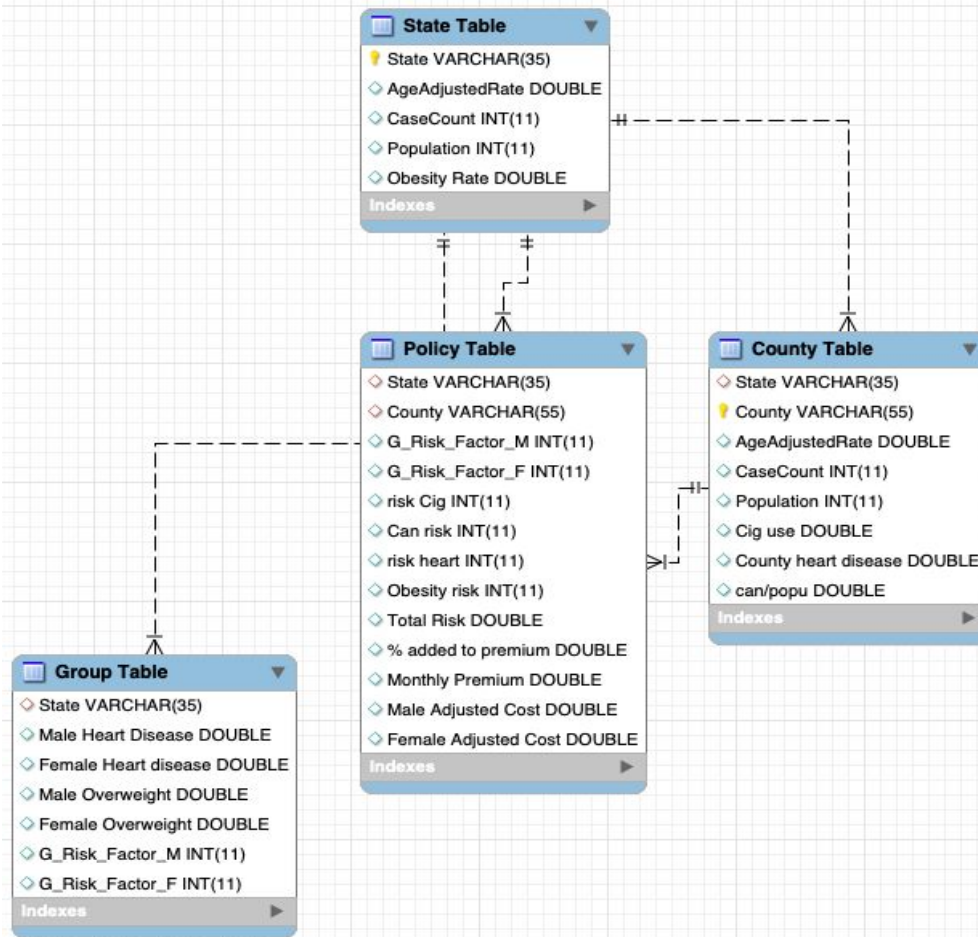
☐ Auto Increment

☐ Generated

# Foreign Keys

-  ALTER TABLE `Group Table` ADD(  
    Constraint Group\_fk1  
    FOREIGN KEY (State) REFERENCES `State Table` (State)  
);
-  ALTER TABLE `County Table` ADD(  
    Constraint County\_fk1  
    FOREIGN KEY (State) REFERENCES `State Table` (State)  
);
-  ALTER TABLE `Policy Table` ADD(  
    Constraint Policy\_fk1  
    FOREIGN KEY (State) REFERENCES `State Table` (State)  
);
-  ALTER TABLE `Policy Table` ADD(  
    Constraint Policy\_fk2  
    FOREIGN KEY (County) REFERENCES `County Table` (County)  
);

# Physical Model



# Database View

```
1 • SELECT * FROM bmiu0259362.`Policy Table`;
```

100% 1:1

Result Grid Filter Rows: Search Export: Fetch rows:

State	County	G_Risk_Factor_M	G_Risk_Factor_F	risk Cig	Can risk	risk heart	Obesity risk	Total Risk	% added to premium	Monthly Premium	Male Adjusted Cost
Alabama	Barbour County, AL	2	3	2	2	3	3	2.5	50.11	100.57	113.98
Alabama	Bibb County, AL	2	3	3	2	3	3	2.75	58.47	106.17	120.33
Alabama	Blount County, AL	2	3	3	2	3	3	2.75	58.39	106.12	120.27
Alabama	Bullock County, AL	2	3	2	2	2	3	2.25	41.95	95.11	107.79
Alabama	Butler County, AL	2	3	2	2	3	3	2.5	50.15	100.6	114.01
Alabama	Calhoun County, AL	2	3	2	2	3	3	2.5	50.03	100.52	113.92
Alabama	Chambers County, AL	2	3	3	3	2	3	2.75	58.42	106.14	120.29
Alabama	Cherokee County, AL	2	3	2	2	3	3	2.5	50.12	100.58	113.99

# Example Queries (Averages)

/\*Avg policy risk factor of each state\*/

```
SELECT 'State', ROUND(AVG('Total Risk'),2) as 'Average Risk'
FROM 'Policy Table'
group by State;
```

/\*Average premium cost for males in United States\*/

```
SELECT ROUND(AVG('Male Adjusted Cost'),2) FROM 'Policy Table'
```

100% 1:96

Result Grid

Filter Rows: Search

Export: 

State	Average R...
Hawaii	1.12
Utah	1.19
Colorado	1.22
District of Columbia	1.25
Massachusetts	1.29
California	1.32
Connecticut	1.36
Idaho	1.41

100% 62:91

Result Grid

Filter Rows: Search

Export: 

ROUND(AVG('Male Adjusted Cost'),2)

104.69

# Example Queries- Highest monthly premiums

Loving county, TX has very low population of about 100 people. Despite lower risk profile it will have higher premiums.

29

30 /\*Top 5 highest policy risk factor for females counties and groups\*/

31 • SELECT 'Female Adjusted Cost', 'Total Risk', 'Policy Table'.County

32 FROM 'Policy Table'

33 LEFT JOIN 'County Table' on 'County Table'.County= 'Policy Table'.County

34 GROUP BY 'Female Adjusted Cost'

35 DESC LIMIT 5;

36

100% 16:35

Result Grid Filter Rows: Search Export: Fetch rows:

Female Adjusted Cost	Total Risk	County	
▶ 139.39	2.333333333	Loving County, TX	
134.46	3	Calhoun County, AR	
134.36	3	Hancock County, TN	
134.35	3	Woodruff County, AR	
134.32	3	Monroe County, AR	







# Lowest Premiums and associated counties

Slight differences between premiums due to differences in population.

```
36
37  /*Top 5 lowest policy risk factor for females counties and groups*/
38  • SELECT `Female Adjusted Cost`, `Total Risk`, `Policy Table`.County
39  FROM `Policy Table`
40  LEFT JOIN `County Table` on `County Table`.County= `Policy Table`.County
41  GROUP BY `Female Adjusted Cost`
42  ASC LIMIT 5;
43
44
45
46
```

100% 15:42

**Result Grid**   Filter Rows:  Export:  Fetch rows: 

	Female Adjusted Cost	Total Risk	County
▶	71.47	1	Contra Costa County, CA
	71.48	1	Imperial County, CA
	71.49	1	Missoula County, MT
	71.5	1	San Benito County, CA
	71.51	1	Eagle County, CO

# States with highest/lowest obesity rates

```
43
44 /*5 Fattest States*/
45 • SELECT State, `Obesity Rate`
46 FROM `State Table`
47 GROUP BY `Obesity Rate`
48 DESC LIMIT 5;
49
50
51
52
```

100% 1:49

Result Grid Filter Rows: Search Edit: Export/Import: Fet

State	Obesity Rate
Mississippi	39.5
Arkansas	37.1
Louisiana	36.8
Kentucky	36.6
Alabama	36.2
NULL	NULL

```
49
50
51 /*5 leanest States*/
52 • SELECT State, `Obesity Rate`
53 FROM `State Table`
54 GROUP BY `Obesity Rate`
55 ASC LIMIT 5;
56
57
```

100% 13:55

Result Grid Filter Rows: Search Edit: Export/Import: Fet

State	Obesity Rate
Colorado	23
District of Columbia	24.7
Hawaii	24.9
Massachusetts	25.7
California	25.8
NULL	NULL

State Table 23

# Example Queries (High/Low)

```
68  
69 /* Counties with lowest Cancer rates*/  
70 • SELECT `can/popu`, `CaseCount`, `County`, `Population`  
71 FROM `County Table`  
72 GROUP BY `can/popu`  
73 ASC LIMIT 5;
```

100% 13:73

Result Grid Filter Rows: Search Edit: Export/Import:

can/popu	CaseCount	County	Population
0	0	Arthur County, NE	2307
0.001118412	16	Ziebach County, SD	14306
0.001863754	107	Chattahoochee County, GA	57411
0.002008276	381	Madison County, ID	189715
0.002399934	351	Summit County, CO	146254
NULL	NULL	NULL	NULL

County Table 29

```
74  
75 /* Counties with Highest Ciggarrete use*/  
76 • SELECT `Cig use`, `County`  
77 FROM `County Table`  
78 GROUP BY `Cig use`  
79 DESC LIMIT 5;
```

100% 14:79

Result Grid Filter Rows: Search Edit: Export/Import: Fetch rows:

Cig use	County
40.18235294	Southeast Fairbanks County, AK
39.24117647	Buffalo County, SD
38.87647059	Nome County, AK
36.77647059	Clay County, KY
36.75882353	McCreary County, KY
NULL	NULL

# Example Queries (Utah)

```
82  /*For UTAH list counties with Highest preimium*/
83  •  SELECT `Cig use`,`County heart disease`, `can/popu`,`County Table`.`County`,`County Table`.State,`Male Adjusted Cost`
84  FROM `County Table`
85  LEFT JOIN `Policy Table` on `County Table`.County= `Policy Table`.County
86  WHERE `County Table`.State='Utah'
87  GROUP BY `Male Adjusted Cost`
88  DESC LIMIT 5;
89
90
91
```

100% 1:90

Result Grid

Filter Rows:

Export:

Fetch rows:

	Cig use	County heart disease	can/popu	County	State	Male Adjusted Cost	
▶	20.92941176	424.6	0.006751283	Plute County, UT	Utah	90.77	
	21.15882353	379.9	0.007035901	Daggett County, UT	Utah	85.3	
	18.75882353	465.7	0.005748409	Kane County, UT	Utah	83.67	
	18.37058824	478.8	0.004002965	Wayne County, UT	Utah	78.21	
	18.74117647	416.9	0.003199025	Juab County, UT	Utah	77.63	

# References

(Including data sources)

<https://gis.cdc.gov/Cancer/USCS/DataViz.html>

<https://www.stateofobesity.org/adult-obesity/>

<https://www.stateofobesity.org/physical-inactivity/>

<https://www.cdc.gov/statesystem/cigaretteuseadult.html>

<https://www.countyhealthrankings.org/app/alabama/2019/measure/factors/11/data>

<https://www.cancer.gov/about-cancer/causes-prevention/risk/obesity/obesity-fact-sheet>

<https://www.cdc.gov/media/releases/2016/p1110-vital-signs-cancer-tobacco.html>

<https://nccd.cdc.gov/DHDSPAtlas/Reports.aspx>

<https://www.cdc.gov/heartdisease/facts.htm>

<https://scienceblog.cancerresearchuk.org/2014/12/26/600000-preventable-cancers-the-size-of-the-healthy-living-prize/>

<https://www.cdc.gov/nchs/fastats/leading-causes-of-death.htm>

<https://pophealthmetrics.biomedcentral.com/articles/10.1186/1478-7954-12-5#MOESM4>

- Electronic supplemental material 3

# Questions?

