

Texas Home Buyers Guide



Brainstorming Ideas



Who is the Audience?

Texas Homeowners



Audience Interest?

Housing price
COVID rate?

Is the place safe?



How to achieve it?

Assign a score to each category.

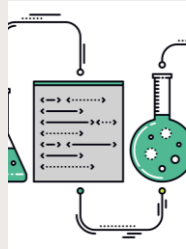
Find livability grade by county using the score

Main Questions ?

- **Housing Price:** Where are housing prices rising the most in the past 6 years? Where have prices fallen? AKA where do you want to stay away from in your home investment?
- **COVID transmission:** Does the place have higher COVID transmission rate?
- **Public safety:** Is the place safe to live in? What is the crime rate?



Data Source Search



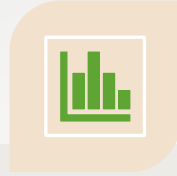
- **Housing Price:** Zillow Housing Data
<https://www.zillow.com/research/data/>
- **COVID transmission:** Texas Health and Human service
COVID-19 data <https://www.dshs.texas.gov/covid-19-coronavirus-disease-2019/texas-covid-19-data>
- **Public safety:** Texas Open Data Portal
https://data.texas.gov/browse?Dataset-Category_Category-Tile=Public+Safety
- **Population Count:** API Call to <https://data.cdc.gov/>



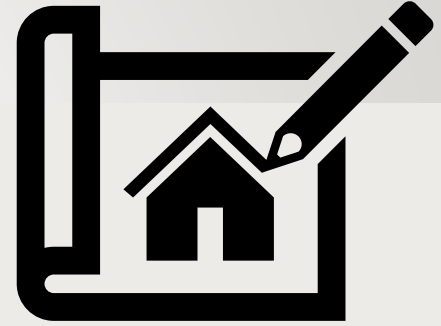
REVIEW THE
DATASHEET



CLEAN THE
DATASHEET



PERFORM
ANALYSIS



| Analysis

Housing Analysis – Approach

Datasheet observation	Cleaning the Datasheet	Identifying Limitations
<ul style="list-style-type: none">•Original csv file had every US county•Format was average home value by month•Date range was 2010-2023	<ul style="list-style-type: none">•Filter Data by State (254 Texas Counties remained)•Retain only columns from April 2017- March 2023 (6 Years)•Remove null values (211 Counties Remaining)	<ul style="list-style-type: none">•43 Counties missing data and were excluded•Data was average home value by county, did not include the range of values within a given county



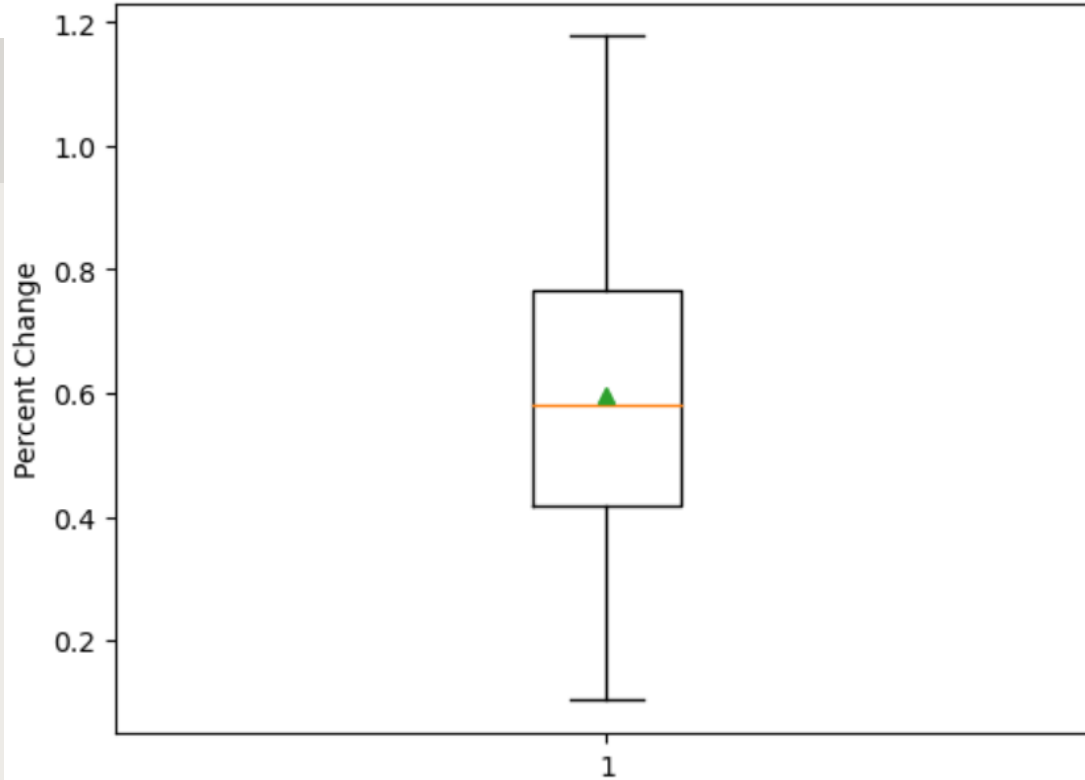
Housing Data Analysis

SUMMARY STATISTICS

	Ave House Value (Mar '23)	Total Change
count	211.000000	211.000000
mean	210191.215213	0.595651
std	116365.047633	0.231988
min	51395.705853	0.103070
25%	132042.741768	0.417610
50%	183460.563318	0.579587
75%	264249.683744	0.764944
max	655275.199060	1.177120

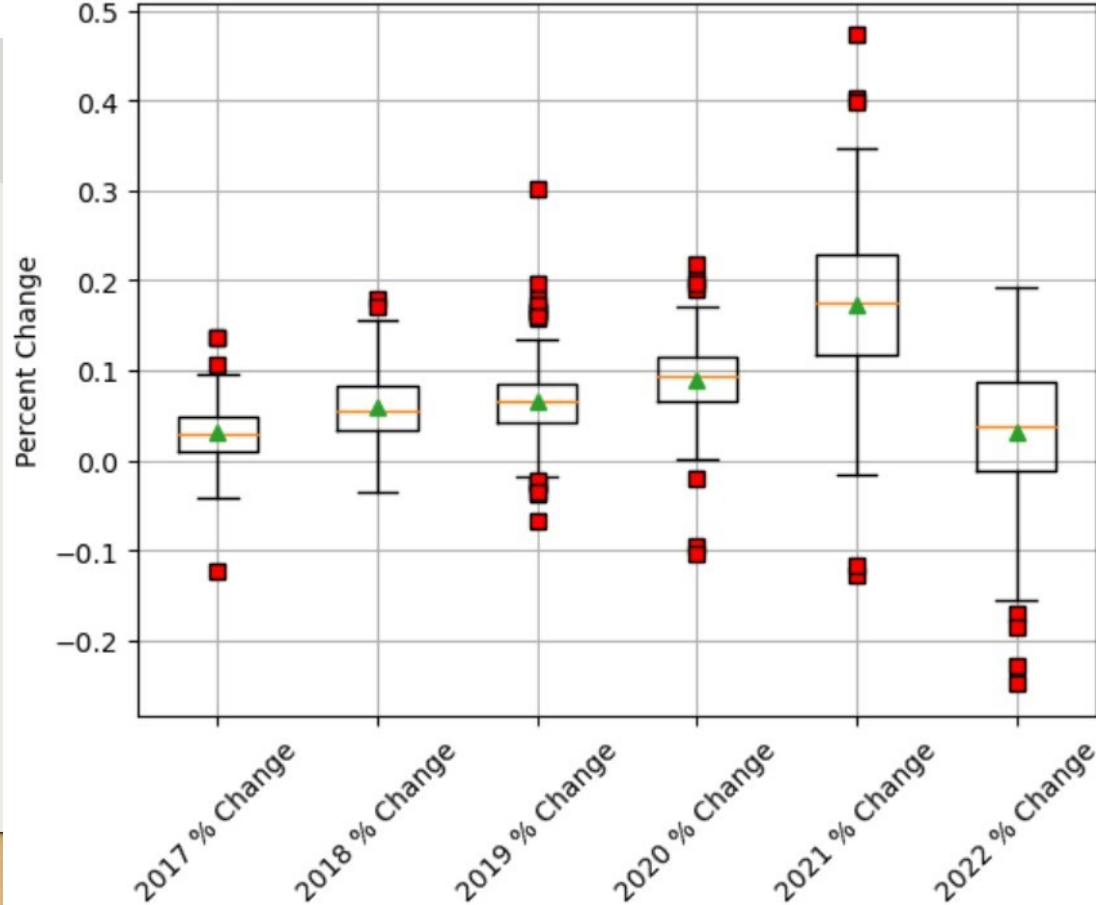
Total Percent Change (by County)

April 2017-March 2023



Texas Housing

Housing Value % Change (by Texas County)



Texas Housing

So... What's a good investment?

County Name	Percent Change	Housing Score
Blanco	1.177	5
Llano	1.137	5
Bosque	1.113	5
Hunt	1.083	5
Burnet	1.079	5

Highest increases in value from 2017 - 2023

County Name	Percent Change	Housing Score
Terry	0.103	1
Ochiltree	0.113	1
Hemphill	0.115	1
Floyd	0.150	1
Collingsworth	0.166	1

Lowest increases in value from 2017 - 2023

Safety Analysis – Approach



Datasheet observation

- 4 years data (2017 – 2020)
- 254 Unique Counties records.

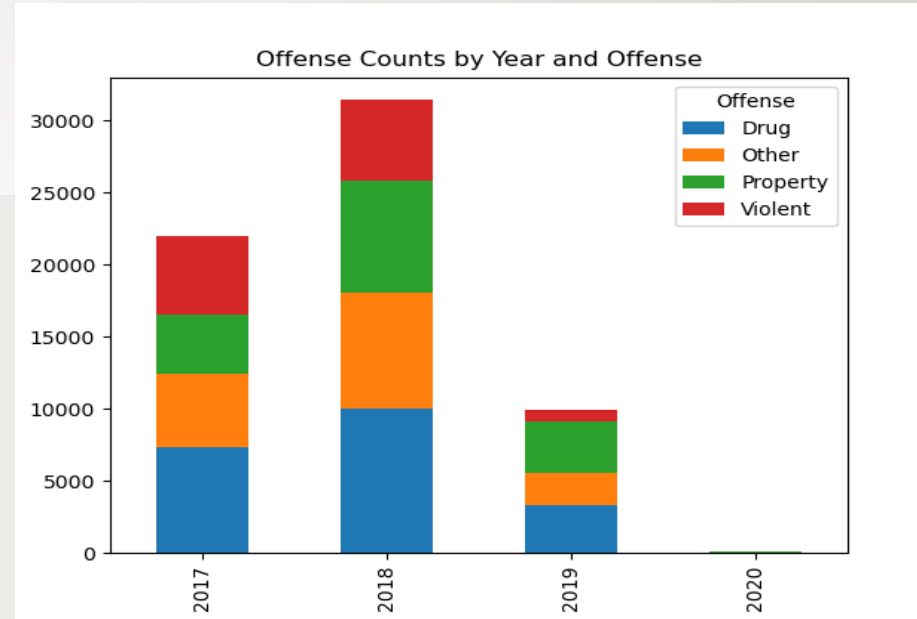
Cleaning the Datasheet

- Used only Offense date.
- Retain only data from 2017-2020.
- Remove null values

Identifying Limitations

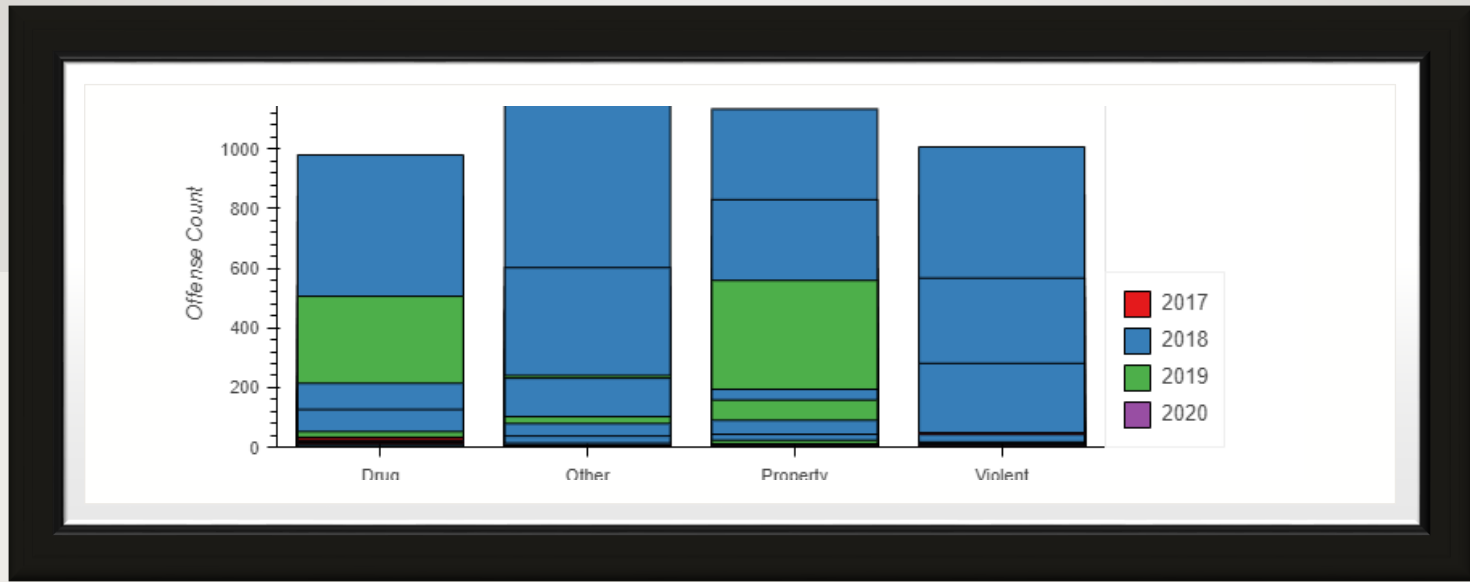
- Only 3 years have complete data. 2020 data is not complete or might not be published for public records.

Safety Analysis – Visualization



Offense count By Year

Year	Drug	Other	Property	Violence
2017	7351	5079	4109	5432
2018	10053	7970	7777	5617
2019	3360	2157	3570	795
2020	26	12	34	-



Offense Type over years

Top 5 Least Rated Counties based on Safety index

County	Avg Offense Count	Offense Percentage	Safety Rank
Anderson	43.0	0.202	0.736
Hill	68.0	0.424	0.818
Smith	268.0	1.690	0.968
Hood	64.0	0.301	0.779
Hopkins	76.0	0.358	0.791

Top 5 Least Rated Counties based on Safety index

County	Avg Offense Count	Offense Percentage	Safety Rank
Anderson	43.0	0.202	0.736
Hill	68.0	0.424	0.818
Smith	268.0	1.690	0.968
Hood	64.0	0.301	0.779
Hopkins	76.0	0.358	0.791

COVID Analysis – Initial prework



Datasheet observation

- 4 years data(2020 to 2023).
- 254 Unique Counties records.
- Each row is cumulative addition
- Informational contents
- Total and Unknowns date records

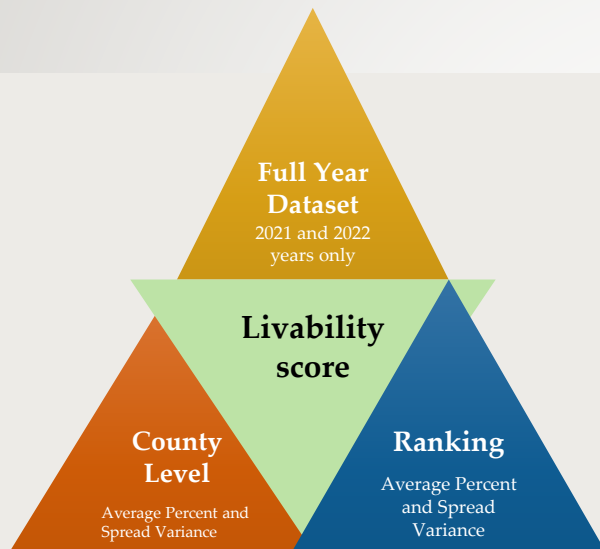
Cleaning the Datasheet

- Remove all informational text
- Retain only month wise data
- Remove null values

Identifying Limitations

- Only the Cumulative Confirmed Cases over Time by County dataset will be used for this analysis.
- Only 2 years have full year dataset (year 2021 and 2022)

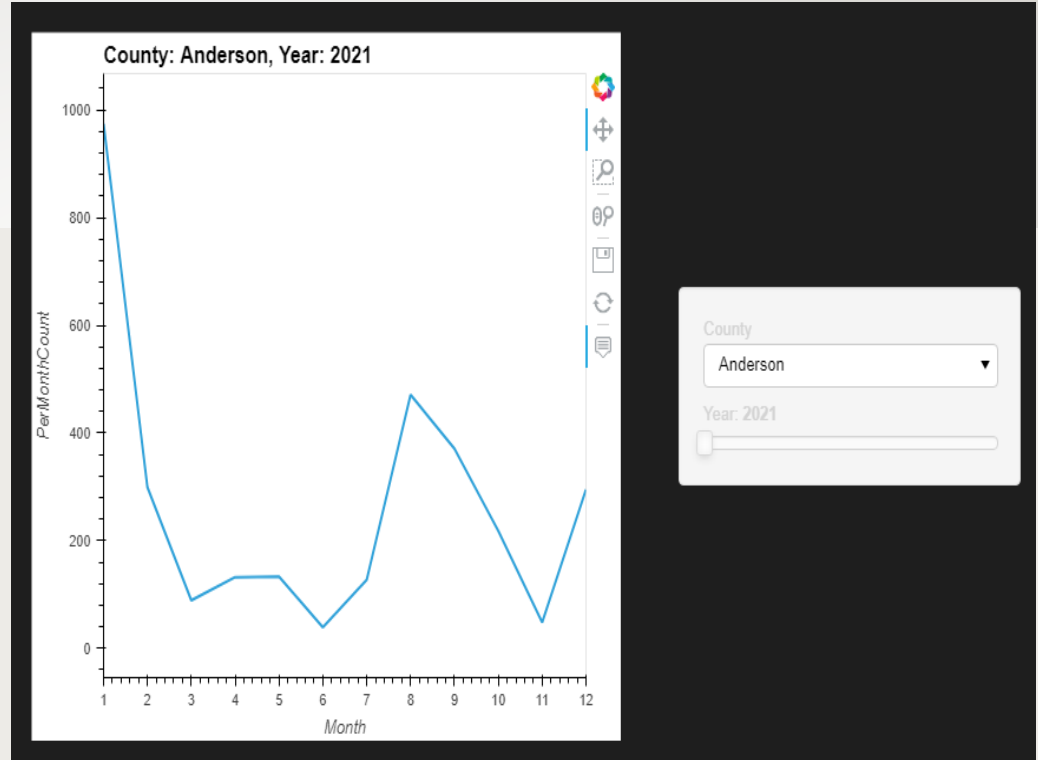
COVID Analysis – Approach





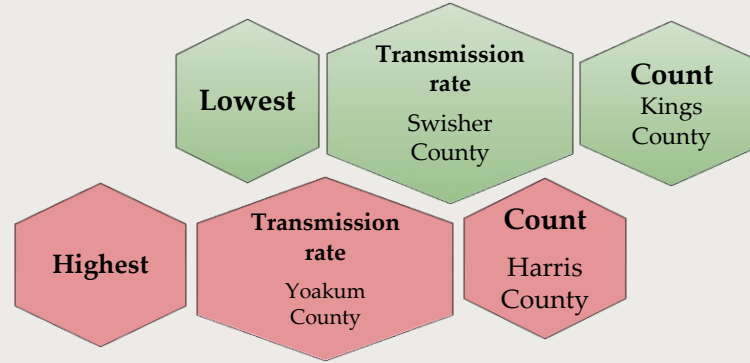
COVID Data Analysis

TIMESERIES VISUALIZATION



COVID Data Analysis

INTERESTING
FINDINGS




Top 5 Best Counties based on livability score



COUNTY	Avg Count	Count Rank	Spread Rate	Spread Rank	Livability Score
King	0.15	0.00	-44.44	0.05	5.0
Shackelford	0.86	0.09	-61.86	0.01	5.0
Foard	0.74	0.07	-54.17	0.02	5.0
Stonewall	0.16	0.01	-59.09	0.01	5.0
Runnels	3.29	0.26	-44.25	0.05	4.5

COVID Analysis – Livability Score

Top 5 Least preferred Counties based on livability score



COUNTY	Avg Count	Count Rank	Spread Rate	Spread Rank	Livability Score
Grimes	29.52	0.72	15.63	0.73	1.0
McLennan	197.98	0.91	30.59	0.83	1.0
Williamson	594.69	0.96	22.19	0.79	1.0
Cameron	318.43	0.94	52.39	0.91	1.0
Maverick	92.99	0.88	32.72	0.83	1.0

COVID Analysis – Livability Score

Time to
Merge



Housing Score

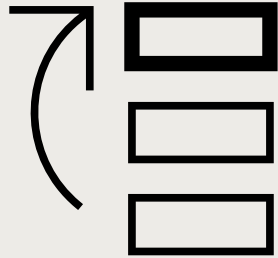


COVID Score



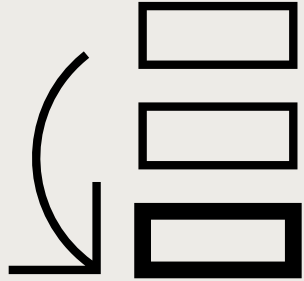
Safety Score

Top 5 best county by Livability Score



COUNTY	Population	LIVABILITY GRADE	COVID SCORE	HOUSING SCORE	SAFETY SCORE
Mason	4274	4.5	4.5	5	4
Stephens	9366	4.17	4.5	5	3
Delta	5331	4.17	3.5	5	4
Presidio	6704	4.17	2.5	5	5
San Augustine	8237	4	4	4	4

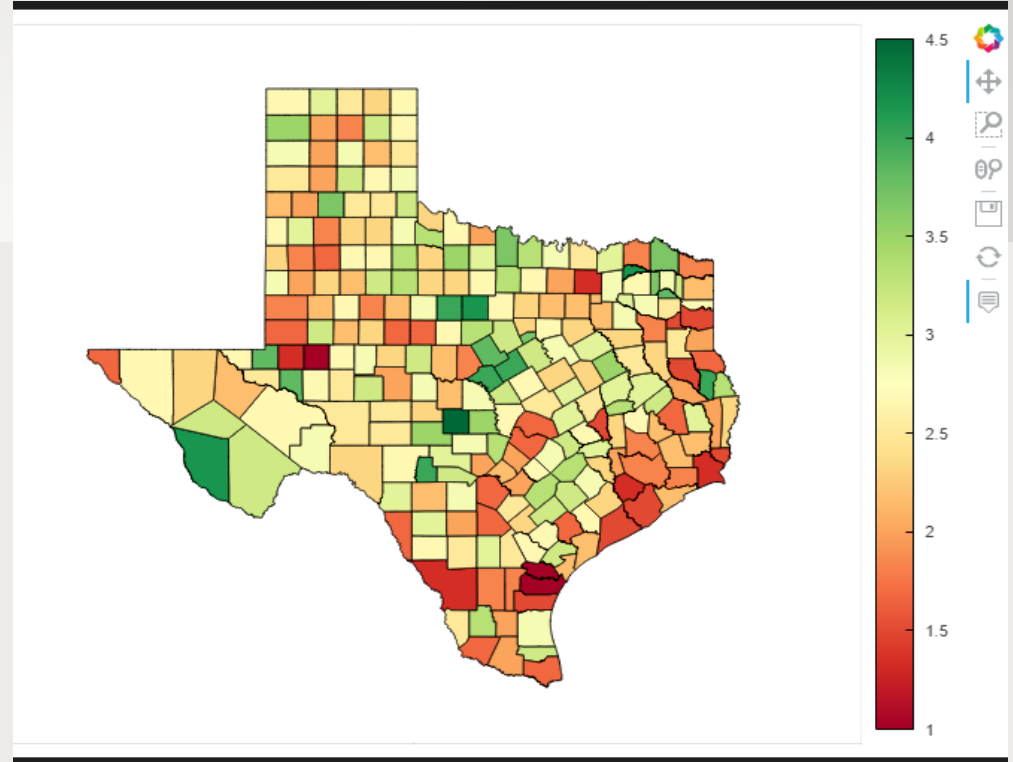
Top 5 least preferred county by Livability Score



COUNTY	Population	LIVABILITY GRADE	COVID SCORE	HOUSING SCORE	SAFETY SCORE
Midland	176832	1	1	1	1
San Patricio	66730	1	1	1	1
Nueces	362294	1	1	1	1
Webb	276652	1.33	1	2	1
Fort Bend	811688	1.33	1	2	1

Texas Map Visualization

Counties by their
livability grade





Questions ?



Thank you!!!