U.S. Population data analysis

Problem description

On the Source files, we find different information about the population in various counties in the USA and how it changed over the years including size age, income, race, age.

Data source: https://www.census.gov/data/tables/time-series/demo/popest/2020s-counties-detail.html

The situation of analysis:

The year is 2021.
You have your own business, a consulting firm specialized in location intelligence marketing. In other words, you advise new and existing business franchises where they should open new locations based on various informations about the locations.
With the information in hand, answere the questions to the following business situations:

Client 1: Your client #1 has several Senior Living Communities in the USA and is looking for the next 3 locations (Counties) to open.
The client's clientele is typically a mid-income senior.
Which locations would you recommend and why? Justify your choice.

Client 2: Your client #2 is a pre-K school and is looking to open a 2 new locations somewhere in the USA.
The school is focused on mid and low income areas.
The client is also specialized in bilingual education (english and spanish).
Find the 2 new locations. Justify your choice.

Client 3: Your client #3 has many stores specialized in food and entertainment for teenagers.
Some of the main features of their locations are:
- They sell "World Sweets", meaning, they sell cakes, candies and other sweet food from different places of the world
- They are an "upscale" type of business so their public is very wealthy
- In addition to food, they also feature modern music (sometimes live), video game tournments

Which 2 places would you recommend them to install their next location? Justify your choice.

Client 4: The client #4 is a beauty store focused on products for womans. Their typical consumer has between 17 and 55 years and is a mid class or above.
Which 5 places would you recommend them to install their stores? Justify your choice.

Client 5: Your client #5 is Disney.
Recently Universal Studios announced that they will open a new park in Frisco (Denton county).
Universal has other similar parks with different attactions based in Florida and in California.
This new park in Frisco will be focused on the young kids and attractions' theme like Trolls, Shrek, Madagascar among others.
Based on the data that you have, should Disney also open a park nearby? Justify your answer.

Assumptions:

Income range:

I'm considering the below income ranges to classify the county as low/ mid or high-income group, based on the median and quartile values from the historical data:

Minimum = $0

Median = $41831

Maximum = $230141

Quartiles (0.5 and 0.75) = $41831.0 and $48718.0

Approximating the values:

Per capita income value	**Classification**
<40000	Low income
40000-55000	Middle income
>55000	High income

Client 1

client #1 has several Senior Living Communities in the USA and is looking for the next 3 locations (Counties) to open.
The client's clientele is typically a mid-income senior.
Which locations would you recommend and why? Justify your choice.

Steps:

Step 1: Extracting income-based population data to identify the middle-income counties. I'm considering the recent 3 years of data for the analysis and only income entries that show the per capita personal income for each county.

Step 2: Grouping by county name, aggregating the income data values and finding the median per capita for each county.

Step 3: Extracting population-by age data and grouping by county, to find the senior citizen population for each county.

Step 4: Identifying the top 20 counties with maximum number of population 65+.

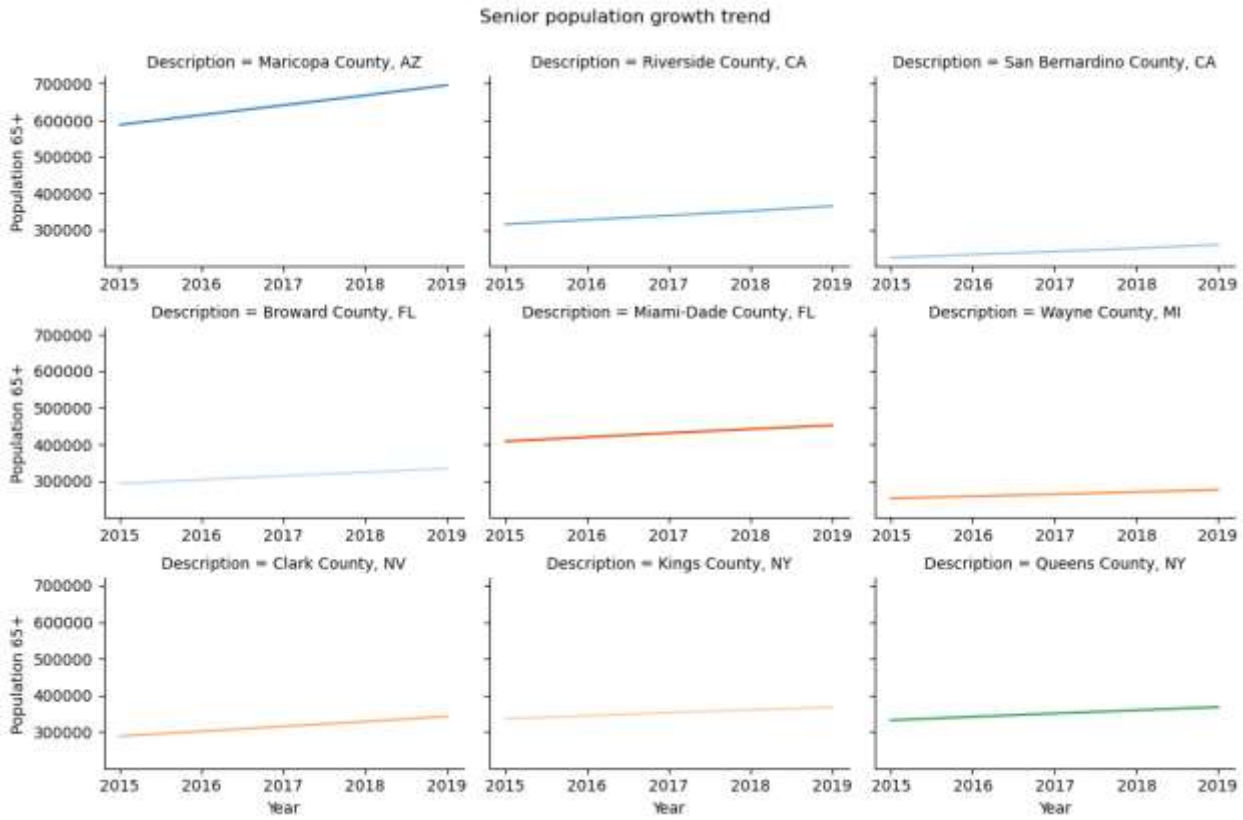
Step 5: Joining this data set with income data values and filtering the middle-income group which is the range between $40000 and $50000 (derived based on median and max values of per capita income).

Results of analysis:

The table shows the top few counties that have maximum senior population, with mid-income.

Description	Population 65+	Median Per-capita income
San Bernardino County, CA	260520	40150
Wayne County, MI	276530	43088
Broward County, FL	334596	50538
Clark County, NV	342659	47759
Riverside County, CA	365942	40587
Kings County, NY	367884	53346
Queens County, NY	368333	50279
Miami-Dade County, FL	452607	53148
Maricopa County, AZ	696331	48113

This graph shows the trend of senior population in these counties. The high count of population and steep increase in Maricopa County, AZ is a good indicator. The next highest value with an increasing trend is Miami-Dade county, FL. For the third place- Kings County, NY and Queens County, NY have very similar values and I'm choosing Queens County with slightly higher senior population and a more appropriate candidate for a middle-income county.



As a result, I would suggest the 3 counties (with the high average senior population in mid-income group):

Maricopa County, AZ

Miami-Dade County, FL

Queens County, NY

Client 2

Your client #2 is a pre-K school and is looking to open 2 new locations somewhere in the USA.
The school is focused on mid and low-income areas.
The client is also specialized in bilingual education (English and Spanish).
Find the 2 new locations. Justify your choice.

Steps for analysis:

Step 1: Identify the top 50 counties with maximum population in ages 0-4

Step 2: Group the per-capita income by county and find the median of the same

Step 3: Finding out the amount of people who will benefit from bilingual education in each county. Summing up the races “white only”, “black only” and “Hispanic” population for this category.

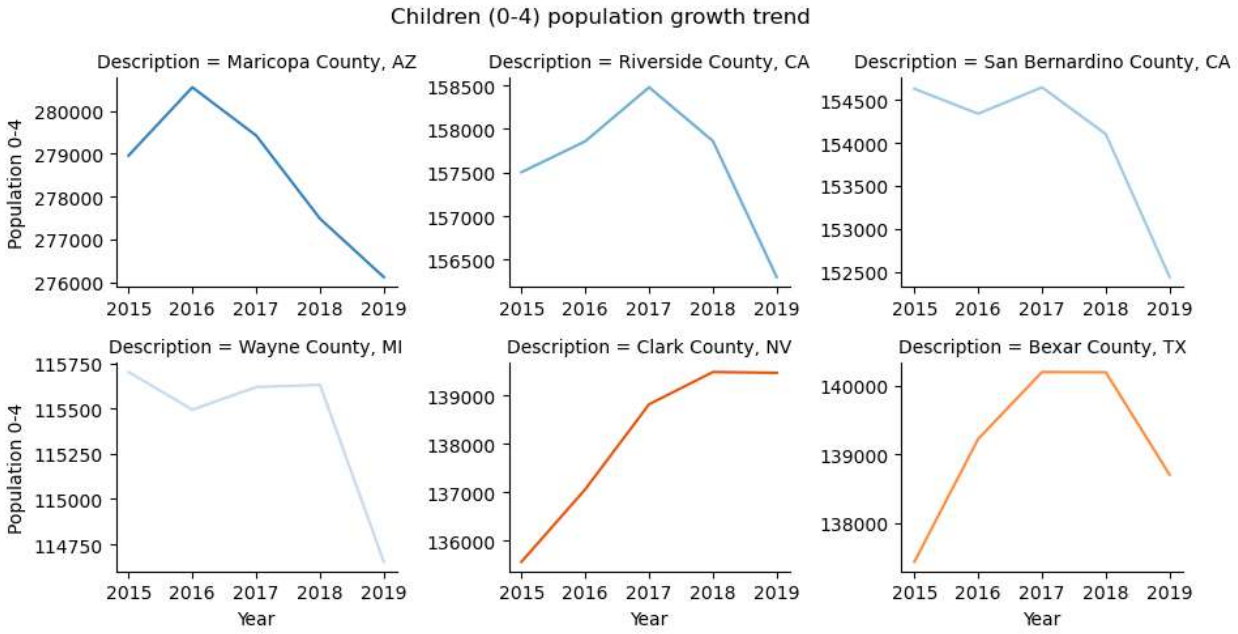
Step 4: Merging the data from above 3 steps and filtering the per capita income range <50000 USD (representing low- and middle-income groups) and total population benefiting from bilingual education to be >163000 (this value is derived from average value of bilingual population of the top 50 counties from step 1.

Based on the above steps, we get the following data:

Description	Population 0-4	Median Per-capita income	Bilingual_Total
Maricopa County, AZ	279426	48113	5409070
Riverside County, CA	158477	40587	3381875
San Bernardino County, CA	154650	40150	3060238
Bexar County, TX	140198	46531	3076308
Clark County, NV	139489	47759	2589233
Wayne County, MI	115630	43088	1749404

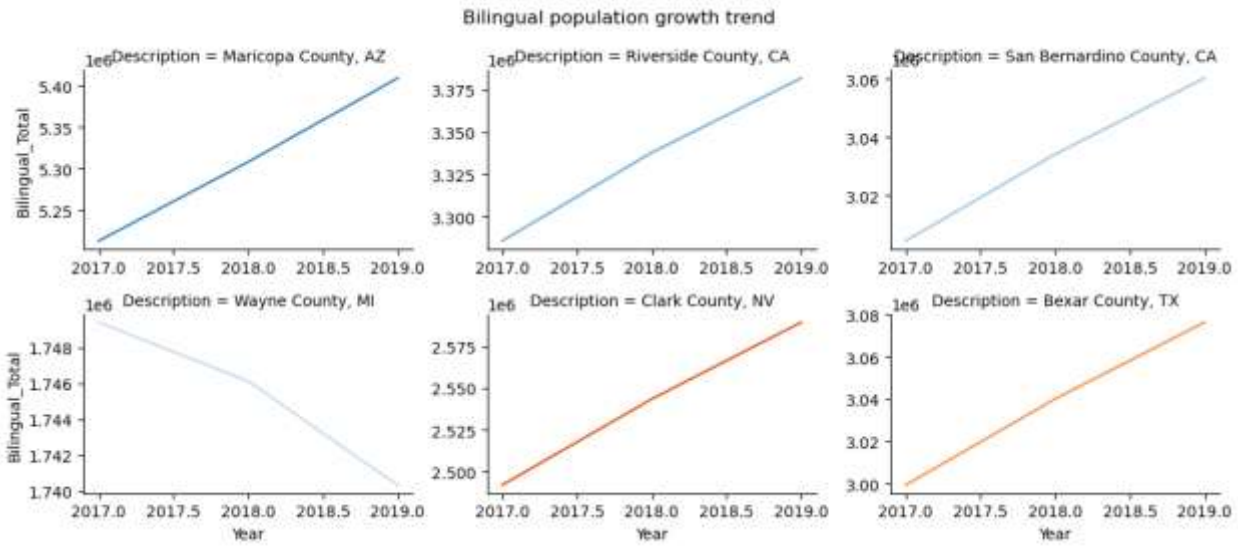
Supporting graphs:

Graph showing trend of population (ages 0-4):



This graph shows that in all counties except Clark County, NV the populations tend to reduce by the year. Though the population seems to reduce (by a small amount), Maricopa County, AZ has a very high current count of both child population and bilingual population. So, this is a good candidate. Secondly, Clark County, NV has a positive increase in child population, which is also a good candidate for pre-k location.

Graph showing trend of bilingual population:



Supporting the choices made above, this graph shows that bilingual population is increasing in both Maricopa County, AZ and Clark County, NV (which also has a better per-capita income).

Result:

For this client, I would suggest

Maricopa County, AZ

Clark County, NV

Client 3

Your client #3 has many stores specialized in food and entertainment for teenagers. Some of the main features of their locations are:- They sell "World Sweets", meaning, they sell cakes, candies and other sweet food from different places of the world - They are an "upscale" type of business so their public is very wealthy - In addition to food, they also feature modern music (sometimes live), video game tournaments.Which 2 places would you recommend them to install their next location? Justify your choice.

Steps followed:
Step 1: Identify the top 50 counties with maximum population in ages 5-17

Step 2: Group the per-capita income by county and find the median of the same

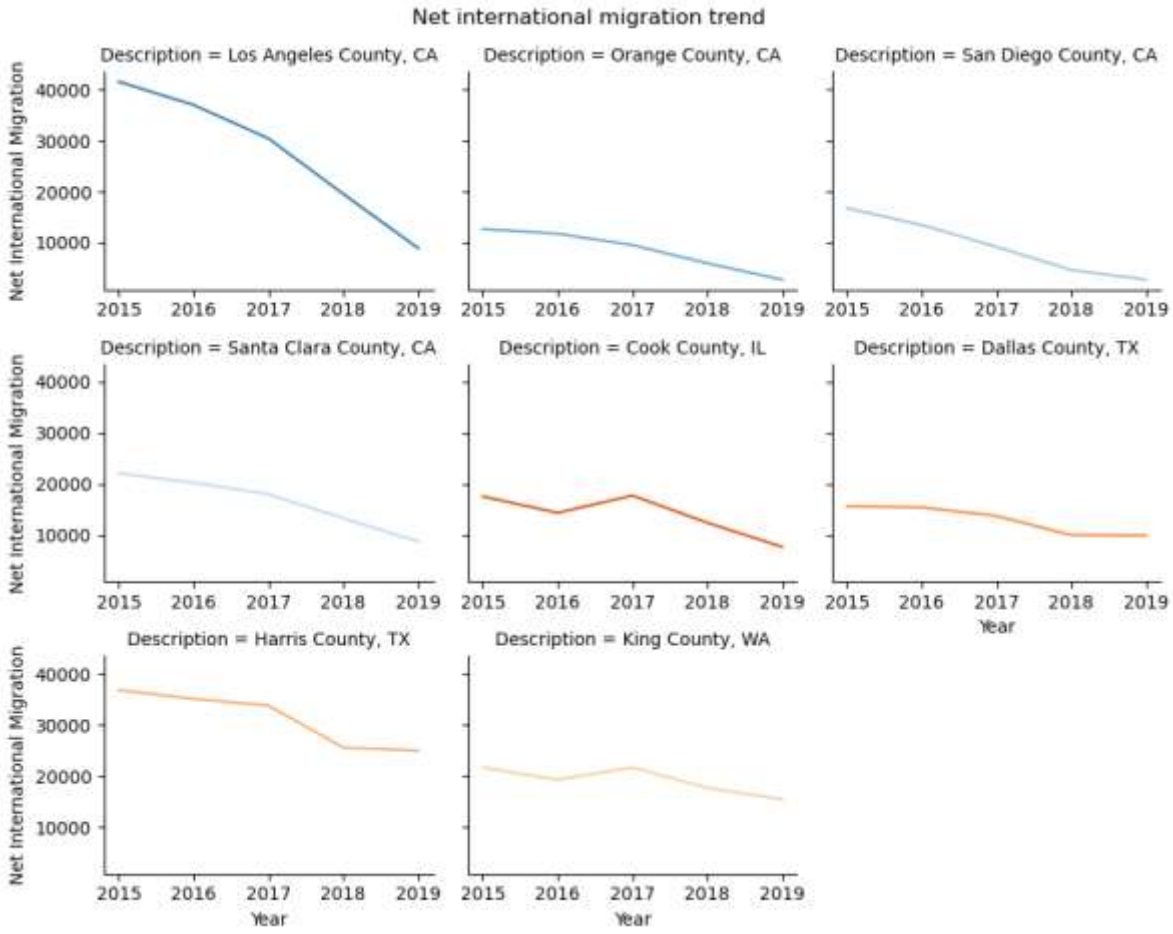
Step 3: Finding out the international population count in each county.

Step 4: Merging the data from above 3 steps and filtering the per capita income range >550000 USD (representing high-income groups) and total immigrant (international) population as a relatively high value for profits.

Results of the analysis:

Description	Population 5-17	Median Per-capita income	Net International Migration
San Diego County, CA	513057.2	61147	2756
Orange County, CA	516692.4	68917	2765
Cook County, IL	822917.8	63436	7659
Santa Clara County, CA	311609.8	110344	8703
Los Angeles County, CA	1603287.4	62300	8911
Dallas County, TX	492369	60780	9889
King County, WA	319768.4	91161	15419
Harris County, TX	893225.8	58235	24967

Graph of international migration trend:



Although there is a steep decline in the net international migration in Los Angeles County, CA, the overwhelmingly high number of population in age group 5-17 makes it a good candidate. But if the client would like more of a higher income county to operate, Santa Clara would be a good fit. My second recommendation would be Dallas County TX, because of the relatively flatter line with no steep decrease and a good international population count. Also, it is a different geographical area so that the client will have a wider audience and more accessible from many states.

My suggestions for this client:

California (Los Angeles County or Santa Clara County)
Dallas County TX.

Client 4

The client #4 is a beauty store focused on products for woman's. Their typical consumer has between 17 and 55 years and is a mid-class or above.
Which 5 places would you recommend them to install their stores? Justify your choice

Steps involved:
Step 1: Find the % of women population in each county (based on historical data)

Step 2: Find the % of population in the ages 18-54 for each county
Step 3: Combine with income-based data to identify middle class or above counties.
Step 4: Merge the data and find the optimal count based on percentage population values.

Results of analysis:

Description	Female Population Percentage	18-54 Population Percentage	Median Per-capita income
Chittenden County, VT	51.04115431	54.38583332	58179
Middlesex County, MA	51.07520944	52.16553593	79970
Wake County, NC	51.41366567	53.28520218	56298
Fulton County, GA	51.6363929	55.19188481	81809
Alexandria Independent City, VA	51.73336176	59.61114833	84923
Suffolk County, MA	51.7540554	61.12288468	76782
Davidson County, TN	51.84586441	55.45599183	63101
Mecklenburg County, NC	52.03430588	54.3131685	58750
Orange County, NC	52.3223508	54.8239388	63062
District of Columbia, DC	52.70699428	59.88701332	79221
New York County, NY	52.73770169	58.08314129	183568

I have chosen the counties based on the female population % >50 and relatively higher population in 18-54 age group, also with a competitive per-capita income for making more profits. The counties that I would suggest are:

New York County, NY
District of Columbia, DC
Suffolk county, MA
Alexandria Independent city, VA
Fulton County, GA

Client 5

Your client #5 is Disney.
Recently Universal Studios announced that they will open a new park in Frisco (Denton county).
Universal has other similar parks with different attactions based in Florida and in California.
This new park in Frisco will be focused on the young kids and attractions' theme like
Trolls, Shrek, Madagascar among others.
Based on the data that you have, should Disney also open a park nearby? Justify your answer.

Since Universal Studios and Disney are world famous theme parks and resorts, I'm considering the entire state population rather than only the county in which the parks are actually present.

Disney (our client) already has theme parks and resorts in California and Florida, as do Universal studios. Below are the results of comparisons drawn between the states of California, Florida and Texas:

Mean population stats (2015 – 2019):

Description	Total Population	Population 0-4	Population 5-17	Population 18-24	Population Under 18
California	39283494	2454293.4	6569004.8	3789582.6	9023298.2
Florida	20901637.2	1130882.8	3052130.8	1750303.8	4183013.6
Texas	28260857.2	2002879.2	5335614.2	2783493.8	7338493.4

Most recent population stats (2019):

Description	Total Population	Population 0-4	Population 5-17	Population 18-24	Population Under 18
California	39512223	2383716	6510925	3678035	8894641
Florida	21477737	1139742	3090187	1742768	4229929
Texas	28995881	1990891	5408919	2813300	7399810

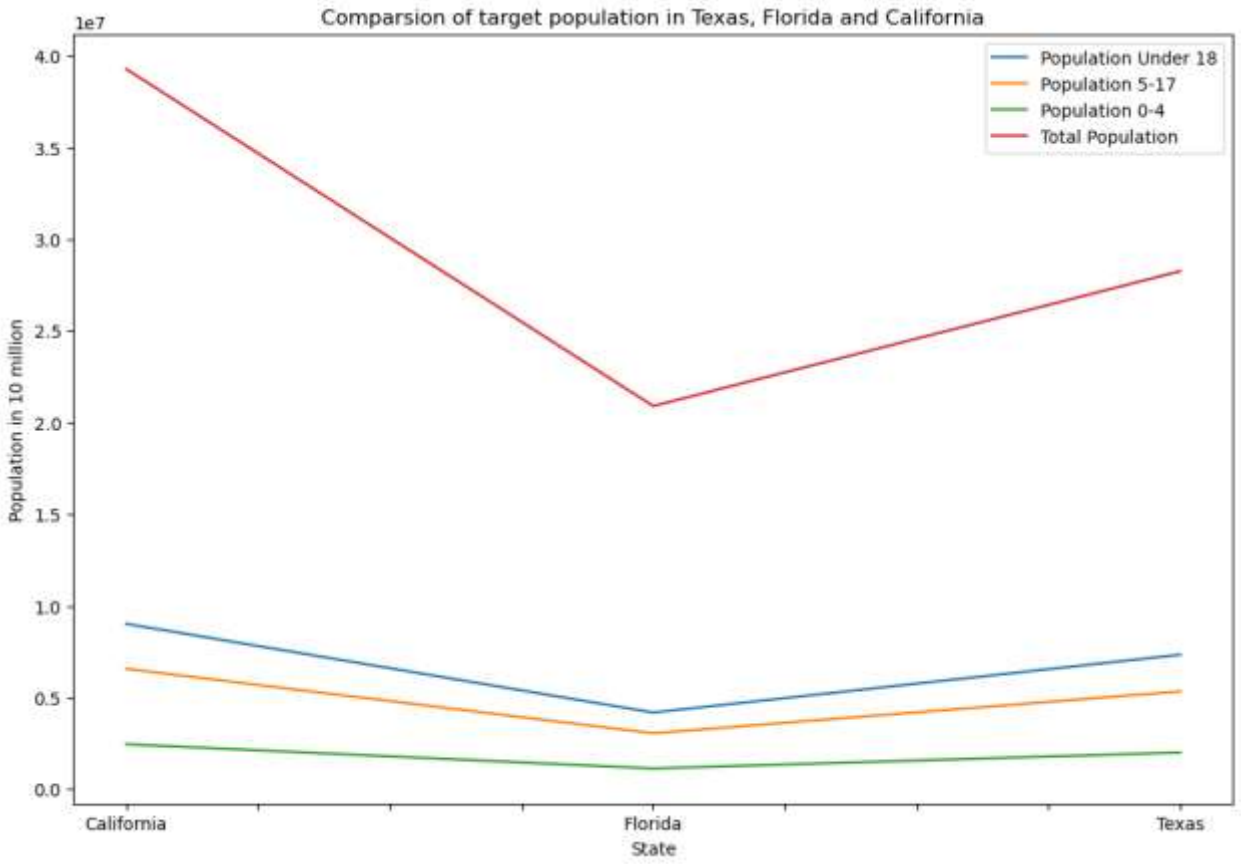
Total Population estimates for 2020:

Description	Year	Population	Count or estimate	State or County Release	Total population: Actual 2019 - Estimate 2020
California	2020	39368078	Estimate	County	-144145
Florida	2020	21733312	Estimate	County	255575
Texas	2020	29360759	Estimate	County	364878

From the above stats, we see that the population of Texas (child and teens) is nearly comparable to the California location. Also, based on the population estimate for 2020, we see that the total population of Texas is estimated to increase by 364878 which is higher than California and Florida.

Overall, based on population data, Texas is a good candidate location for the new Disneyland.

Graph showing the population distribution:



Provided the Florida location of Disney theme park is thriving with profits, Texas location would definitely be a good choice for Disney.

Answer:
Yes, Disney should open a park nearby in Texas.