Which is the best county to expand clothing stores to in California?

1) Finding out how many clothing stores are present in each county

```
select count(location_name) as number_of_stores,
SUBSTRING(poi_cbg, 3, 3) as county_sub , c.county
from `finalprojectmod.Finalmod.visits` a
inner join `finalprojectmod.Finalmod.brands` b
on a.safegraph_brand_ids = b.safegraph_brand_id
inner join `finalprojectmod.Finalmod.cbg_fips` c
on SUBSTRING(poi_cbg, 3, 3) = c.county_fips
where poi_cbg like '06%'
and b.top_category = 'Clothing Stores'
and c.state = 'CA'
group by county_sub , c.county
order by number_of_stores desc
;
```

Result : 37 counties have clothing stores

Saved this table as “Numberofstores_per_county_cali”

2) Listing all counties in Cali and joining the previous result to identify which counties currently have no stores

```
select a.county, a.county_fips, b.number_of_stores
from `finalprojectmod.Finalmod.cbg_fips` a
left join `Final_MOD.Numberofstores_per_county_cali` b
on a.county_fips = b.county_sub
where state = 'CA'
order by b.number_of_stores
;
```

Result: 21 counties have null values

Saved this table as “Numberofstores_per_county_Cali_all”

3) Finding population of high income group (>125K) by counties in California

```
SELECT
SUM(`inc_125-150` + `inc_150-200` + `inc_gte200`) AS
total_population_high_income,
SUBSTRING(cbg, 2, 3) AS cbg_county_substring, b.county, b.state
FROM
`finalprojectmod.Finalmod.cbg_demographics` a
inner join `finalprojectmod.Finalmod.cbg_fips` b
```

```
ON SUBSTRING(a.cbg, 2, 3) = b.county_fips
WHERE
cbg LIKE '6%'
and b.state = 'CA'
AND LENGTH(cbg) = 11
GROUP BY
cbg_county_substring, b.county, b.state
ORDER BY
total_population_high_income desc
;
```

Saved this table as “Highincomepop_by_county_cali”

4) Joining the above two tables and finding out which counties that have no stores (null) have the largest group of high income population

```
select a.* , b.total_population_high_income from
`Final_MOD.Numberofstores_per_county_Cali_all` a
inner join `Final_MOD.Highincomepop_by_county_cali` b
on a.county_fips = b.cbg_county_substring
order by number_of_stores , total_population_high_income desc
;
```

Conclusion:
El Derado county has no clothing stores and the largest population of high income group, making it a lucrative option to open clothing stores

Analyzimg consumer behavior and peak visit times in areas with existing clothing stores
Motive: So the client can design targeted sales and incentives during those periods, with the aim of boosting clothing sector growth.

1) Finding out the average daily visit counts, average raw visitor counts in each county in Cali

```
SELECT
SUBSTRING(poi_cbg, 3, 3) AS cbg_county_substring,c.county,
ROUND(AVG(CAST(JSON_EXTRACT_SCALAR(popularity_by_day, '$.Monday') AS INT64)))
AS avg_Monday,
ROUND(AVG(CAST(JSON_EXTRACT_SCALAR(popularity_by_day, '$.Tuesday') AS INT64)))
AS avg_Tuesday,
ROUND(AVG(CAST(JSON_EXTRACT_SCALAR(popularity_by_day, '$.Wednesday') AS
INT64))) AS avg_Wednesday,
ROUND(AVG(CAST(JSON_EXTRACT_SCALAR(popularity_by_day, '$.Thursday') AS INT64)))
AS avg_Thursday,
ROUND(AVG(CAST(JSON_EXTRACT_SCALAR(popularity_by_day, '$.Friday') AS INT64)))
AS avg_Friday,
ROUND(AVG(CAST(JSON_EXTRACT_SCALAR(popularity_by_day, '$.Saturday') AS INT64)))
AS avg_Saturday,
ROUND(AVG(CAST(JSON_EXTRACT_SCALAR(popularity_by_day, '$.Sunday') AS INT64)))
AS avg_Sunday,
round (avg(raw_visitor_counts)) as raw_avg_visitor_counts
FROM
`finalprojectmod.Finalmod.visits` a
INNER JOIN `finalprojectmod.Finalmod.brands` b
ON a.safegraph_brand_ids = b.safegraph_brand_id
inner join `finalprojectmod.Finalmod.cbg_fips` c
ON SUBSTRING(a.poi_cbg, 3, 3) = c.county_fips
WHERE
poi_cbg LIKE '06%'
AND b.top_category = "Clothing Stores"
and c.state = 'CA'
GROUP BY
cbg_county_substring, c.county
order by
Cbg_county_substring
;
```

Saved this table as “Foot_traffic_by_county_cali”

2) Finding Population of each county in California

```
select sum(pop_total) as population_per_county,
SUBSTRING(cbg, 2, 3) AS cbg_county_substring, b.county
from
`finalprojectmod.Finalmod.cbg_demographics` a
inner join `finalprojectmod.Finalmod.cbg_fips` b
on SUBSTRING(a.cbg, 2, 3) = b.county_fips
where cbg like '6%'
and length (cbg) = 11
and b.state ='CA'
group by cbg_county_substring, county
order by county
;
```

Saved this table as “Population_per_county_cali”

3) Joining these two tables to find per capita results based on number of stores (using the Numberofstores_per_county_cali table created in 1.1)

```
SELECT
a.*,
b.population_per_county,
(c.number_of_stores / b.population_per_county) as number_of_stores_per_capita,
c.number_of_stores
FROM
`Final_MOD.Foot_traffic_by_county_cali` a
INNER JOIN
`Final_MOD.Population_per_county_cali` b
ON
a.cbg_county_substring = b.cbg_county_substring
inner join
`Final_MOD.Numberofstores_per_county_cali`c
on a.cbg_county_substring = c.county_sub
ORDER BY
CAST(c.number_of_stores / b.population_per_county AS FLOAT64) DESC
;
```

Result: Napa County has the highest number of stores per capita and have the daily visits highest at Saturday, suggesting Saturday is a good day to offer deals and discounts to attract customers to maximize profits

4) Joining these two tables to find per capita results based on raw visitor counts in a month (using the Numberofstores_per_county_cali table created in 1.1)

```
SELECT
a.*,
b.population_per_county,
(a.raw_avg_visitor_counts/ b.population_per_county) as
avg_raw_vistors_per_capita,
c.number_of_stores
FROM
`Final_MOD.Foot_traffic_by_county_cali` a
INNER JOIN
`Final_MOD.Population_per_county_cali` b
ON
a.cbg_county_substring = b.cbg_county_substring
inner join
`Final_MOD.Numberofstores_per_county_cali`c
on a.cbg_county_substring = c.county_sub
ORDER BY
CAST(a.raw_avg_visitor_counts / b.population_per_county AS FLOAT64) DESC
;
```

Result: Tehama County has the highest number of raw visitors per capita per month and have the daily visits highest at Friday, suggesting Friday is a good day to offer deals and discounts to attract customers to maximize profits

Analyzing potential for expansion of clothing stores based on gender and age.

1) Finding number of Children and Infant Clothing stores in each county

```
select count(location_name) as number_of_stores,
SUBSTRING(poi_cbg, 3, 3) as county_sub
from `finalprojectmod.Finalmod.visits` a
inner join `finalprojectmod.Finalmod.brands` b
on a.safegraph_brand_ids = b.safegraph_brand_id
where poi_cbg like '06%'
and b.top_category = 'Clothing Stores'
and b.sub_category = ("Children's and Infants' Clothing Stores")
group by county_sub
order by county_sub
;
```

Saved this table as “child_infant_stores_by_county”

2) Finding population of infants and children in each county

```
SELECT
SUM(`pop_m_lt5`+`pop_f_lt5`+`pop_m_5-9`+
`pop_m_10-14`+`pop_f_5-9`+`pop_f_10-14`) AS total_sum,
SUBSTRING(cbg, 2, 3) AS cbg_county_substring, b.county, b.state
FROM
`finalprojectmod.Finalmod.cbg_demographics` a
inner join `finalprojectmod.Finalmod.cbg_fips` b
ON SUBSTRING(a.cbg, 2, 3) = b.county_fips
WHERE
cbg LIKE '6%'
and b.state = 'CA'
AND LENGTH(cbg) = 11
GROUP BY
cbg_county_substring, b.county, b.state
ORDER BY
Total_sum
;
```

Saved this table as “ Child_infant_pop_per_county_cali”

3) Using the above two tables to find children and infant stores per capita

```
select a.*, b.number_of_stores,
(number_of_stores / total_sum) as stores_per_capita_infant
from `Final_MOD.Child_infant_pop_per_county_cali` a
inner join `Final_MOD.child_infant_stores_by_county`b
on a.cbg_county_substring = b.county_sub
order by stores_per_capita_infant
;
```

Result: Kern County has the lowest number of stores for infants and children per capita, indicating a potential for growth

4) Finding number of Women Clothing stores in each county

```
select count(location_name) as number_of_stores,
SUBSTRING(poi_cbg, 3, 3) as county_sub
from `finalprojectmod.Finalmod.visits` a
inner join `finalprojectmod.Finalmod.brands` b
on a.safegraph_brand_ids = b.safegraph_brand_id
where poi_cbg like '06%'
and b.top_category = 'Clothing Stores'
and b.sub_category = ("Women's Clothing Stores")
group by county_sub
order by county_sub
;
```

Saved this table as “women_stores_by_county”

5) Finding population of Women in each county

```
SELECT
SUM(pop_f_total) AS total_sum,
SUBSTRING(cbg, 2, 3) AS cbg_county_substring, b.county, b.state
FROM
`finalprojectmod.Finalmod.cbg_demographics` a
inner join `finalprojectmod.Finalmod.cbg_fips` b
ON SUBSTRING(a.cbg, 2, 3) = b.county_fips
WHERE
cbg LIKE '6%'
and b.state = 'CA'
AND LENGTH(cbg) = 11
GROUP BY
cbg_county_substring, b.county, b.state
```

```
ORDER BY
Total_sum
;
```

Saved this table as " Women_pop_per_county_cali"

6) Using the above two tables to find women's stores per capita

```
select a.*, b.number_of_stores,
(number_of_stores / total_sum) as stores_per_capita_women
 from `Final_MOD.Women_pop_per_county_cali` a
inner join `Final_MOD.women_stores_by_county` b
on a.cbg_county_substring = b.county_sub
order by stores_per_capita_women
;
```

Result: San Mateo County has the lowest number of stores for women per capita, indicating a potential for growth

7) Finding number of Men Clothing stores in each county

```
select count(location_name) as number_of_stores,
SUBSTRING(poi_cbg, 3, 3) as county_sub
from `finalprojectmod.Finalmod.visits` a
inner join `finalprojectmod.Finalmod.brands` b
on a.safegraph_brand_ids = b.safegraph_brand_id
where poi_cbg like '06%'
and b.top_category = 'Clothing Stores'
and b.sub_category = ("Men's Clothing Stores")
group by county_sub
order by county_sub
;
```

Saved this table as "men_stores_by_county"

8) Finding population of Men in each county

```
SELECT
SUM(pop_m_total) AS total_sum,
SUBSTRING(cbg, 2, 3) AS cbg_county_substring, b.county, b.state
FROM
`finalprojectmod.Finalmod.cbg_demographics` a
inner join `finalprojectmod.Finalmod.cbg_fips` b
ON SUBSTRING(a.cbg, 2, 3) = b.county_fips
WHERE
```

```
cbg LIKE '6%'
and b.state = 'CA'
AND LENGTH(cbg) = 11
GROUP BY
cbg_county_substring, b.county, b.state
ORDER BY
Total_sum;
```

Saved this table as “men_pop_per_county_cali”

9) Using the above two tables to find men’s stores per capita

```
select a.*, b.number_of_stores,
(number_of_stores / total_sum) as stores_per_capita_men
 from `Final_MOD.men_pop_per_county_cali` a
inner join `Final_MOD.men_stores_by_county` b
on a.cbg_county_substring = b.county_sub
order by stores_per_capita_men;
```

Result: Santa Barbara County has the lowest number of stores for men per capita, indicating a potential for growth