

Bike Incidents in California

Our data

```
Rows: 290
Columns: 13
$ county      <chr> "ALAMEDA", "ALAMEDA", "ALAMEDA", "ALAMEDA", "ALAMEDA", "ALPI...
$ highway_type <chr> "Conventional", "Expressway", "Freeway", "One-Way City Stree...
$ total       <dbl> 10, 0, 6, 0, 16, 3, 0, 0, 0, 3, 1, 0, 0, 0, 1, 3, 0, 0, 1, 4...
$ fatality    <dbl> 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ...
$ injury      <dbl> 10, 0, 5, 0, 15, 3, 0, 0, 0, 3, 1, 0, 0, 0, 1, 3, 0, 0, 1, 4...
$ latitude    <dbl> 37.6017, 37.6017, 37.6017, 37.6017, 37.6017, 38.5974, 38.597...
$ longitude   <dbl> -121.7195, -121.7195, -121.7195, -121.7195, -121.7195, -119....
$ state       <chr> "California", "California", "California", "California", "Cal...
$ year        <dbl> 2022, 2022, 2022, 2022, 2022, 2022, 2022, 2022, 2022, 2022, ...
$ pop         <dbl> 1663823, 1663823, 1663823, 1663823, 1663823, 1515, 1515, 151...
$ total_capita <dbl> 6.010255e-06, 0.000000e+00, 3.606153e-06, 0.000000e+00, 9.61...
$ fatality_capita <dbl> 0.000000e+00, 0.000000e+00, 6.010255e-07, 0.000000e+00, 6.01...
$ injury_capita <dbl> 6.010255e-06, 0.000000e+00, 3.005127e-06, 0.000000e+00, 9.01...
```

Creating a Data Frame with County Data

```
#Created a data frame with the county names, latitudes, and longitudes using the help of ChatGPT
```{r}
county_data <- data.frame(
 county = c("ALAMEDA", "ALPINE", "AMADOR", "BUTTE", "CALAVERAS", "COLUSA", "CONTRA COSTA", "DEL NORTE",
 "EL DORADO", "FRESNO", "GLENN", "HUMBOLDT", "IMPERIAL", "INYO", "KERN", "KINGS",
 "LAKE", "LASSEN", "LOS ANGELES", "MADERA", "MARIN", "MARIPOSA", "MENDOCINO", "MERCED",
 "MODOC", "MONO", "MONTEREY", "NAPA", "NEVADA", "ORANGE", "PLACER", "PLUMAS",
 "RIVERSIDE", "SACRAMENTO", "SAN BENITO", "SAN BERNARDINO", "SAN DIEGO", "SAN FRANCISCO",
 "SAN JOAQUIN", "SAN LUIS OBISPO", "SAN MATEO", "SANTA BARBARA", "SANTA CLARA", "SANTA CRUZ",
 "SHASTA", "SIERRA", "SISKIYOU", "SOLANO", "SONOMA", "STANISLAUS", "SUTTER", "TEHAMA",
 "TRINITY", "TULARE", "TUOLUMNE", "VENTURA", "YOLO", "YUBA"),
 latitude = c(37.6017, 38.5974, 38.3489, 39.6254, 38.196, 39.1789, 37.9191, 41.7435,
 38.7787, 36.9859, 39.5989, 40.745, 32.8397, 36.3093, 35.3433, 36.0741,
 39.1012, 40.6739, 34.0522, 36.9859, 38.0834, 37.4849, 39.5501, 37.2083,
 41.5885, 37.938, 36.2168, 38.5025, 39.1347, 33.7175, 38.9045, 40.0036,
 33.9534, 38.5816, 36.6115, 34.9592, 32.7157, 37.7749, 37.9577, 35.3102,
 37.563, 34.4208, 37.3541, 36.9741, 40.7909, 39.5774, 41.6639, 38.3105,
 38.5779, 37.5091, 39.0446, 40.1251, 40.6503, 36.1342, 37.8675, 34.3705, 38.7646, 39.2547),
 longitude = c(-121.7195, -119.8203, -120.7741, -121.537, -120.6805, -122.2342, -121.9283, -123.8974,
 -120.5231, -119.2321, -122.3935, -123.8695, -115.6121, -117.546, -118.7278, -119.8155,
 -122.7533, -120.5579, -118.2437, -120.5824, -122.7633, -119.9663, -123.4384, -120.6977,
 -120.7525, -118.8867, -121.2264, -122.2655, -121.171, -117.8311, -121.1448, -120.8393,
 -117.3962, -121.4944, -121.286, -116.4194, -117.1611, -122.4194, -121.2908, -120.4358,
 -122.3255, -119.6982, -121.9552, -122.0308, -121.8474, -120.5211, -122.545, -121.9018,
 -122.9888, -120.9876, -121.3153, -122.2345, -123.089, -118.8597, -120.2602, -119.1391, -121.9018, -121.3999)
)

#rohan
```
```

Adding per Capita data to our Data Frame

```
```{r}
bikes_capita <- bikes_new %>%
 mutate(total_capita = total/pop, fatality_capita = fatality/pop, injury_capita = injury/pop) #capita counts
print(bikes_capita)
```
```

A tibble: 290 × 13

| | injury
<dbl> | latitude
<dbl> | longitude
<dbl> | state
<chr> | year
<dbl> | pop
<dbl> | total_capita
<dbl> | fatality_capita
<dbl> | injury_capita
<dbl> |
|--|-----------------|-------------------|--------------------|----------------|---------------|--------------|-----------------------|--------------------------|------------------------|
| | 10 | 37.6017 | -121.7195 | California | 2022 | 1663823 | 0.0000060102547 | 0.000000000000000 | 0.0000060102547 |
| | 0 | 37.6017 | -121.7195 | California | 2022 | 1663823 | 0.000000000000000 | 0.000000000000000 | 0.000000000000000 |
| | 5 | 37.6017 | -121.7195 | California | 2022 | 1663823 | 0.0000036061528 | 0.0000006010255 | 0.0000030051273 |
| | 0 | 37.6017 | -121.7195 | California | 2022 | 1663823 | 0.000000000000000 | 0.000000000000000 | 0.000000000000000 |
| | 15 | 37.6017 | -121.7195 | California | 2022 | 1663823 | 0.0000096164075 | 0.0000006010255 | 0.0000090153820 |
| | 3 | 38.5974 | -119.8203 | California | 2022 | 1515 | 0.0019801980198 | 0.000000000000000 | 0.0019801980198 |
| | 0 | 38.5974 | -119.8203 | California | 2022 | 1515 | 0.000000000000000 | 0.000000000000000 | 0.000000000000000 |
| | 0 | 38.5974 | -119.8203 | California | 2022 | 1515 | 0.000000000000000 | 0.000000000000000 | 0.000000000000000 |
| | 0 | 38.5974 | -119.8203 | California | 2022 | 1515 | 0.000000000000000 | 0.000000000000000 | 0.000000000000000 |
| | 3 | 38.5974 | -119.8203 | California | 2022 | 1515 | 0.0019801980198 | 0.000000000000000 | 0.0019801980198 |

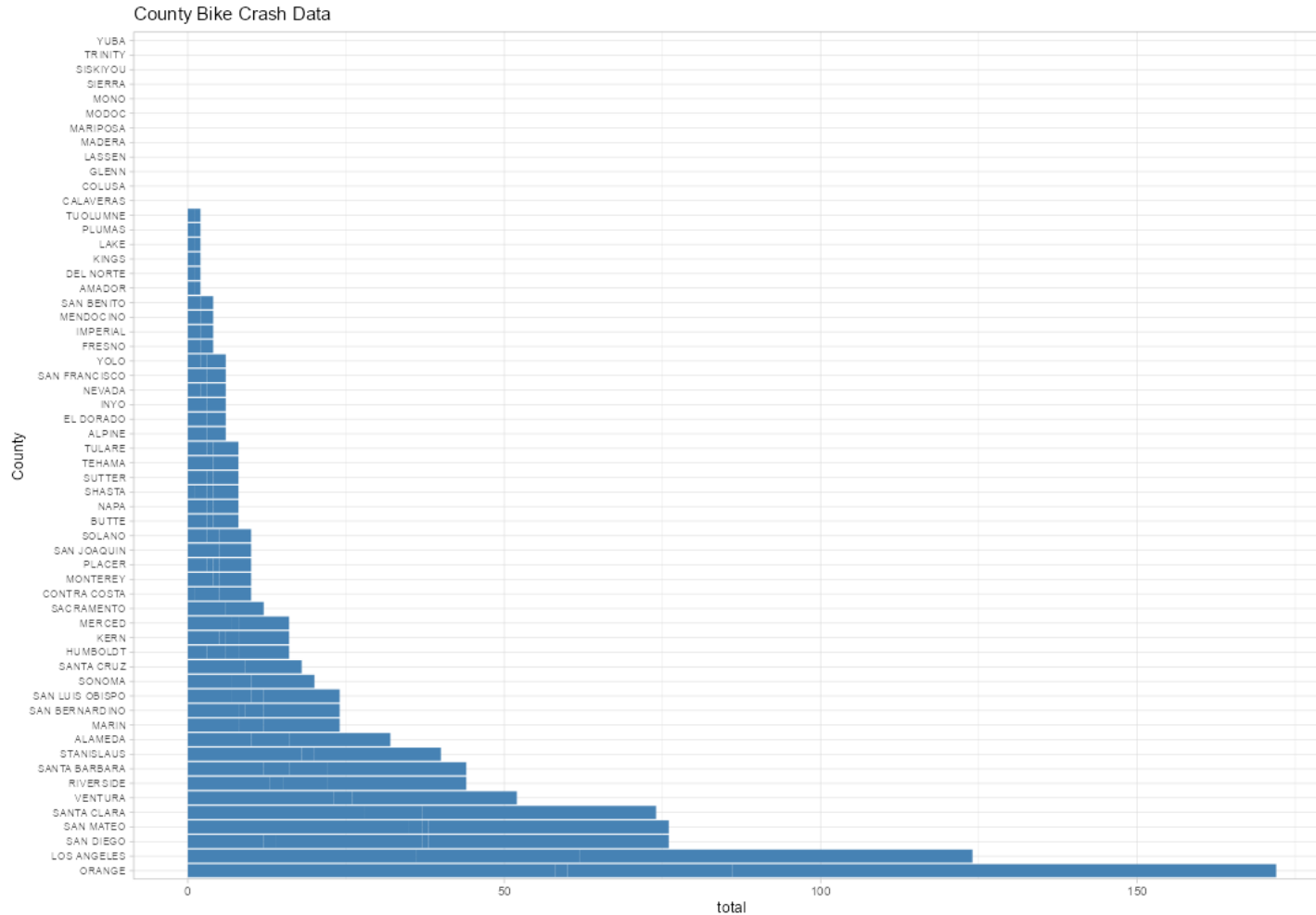
1-10 of 290 rows | 5-13 of 13 columns

Previous 1 2 3 4 5 6 ... 29 Next

Which county has the most bike crashes?

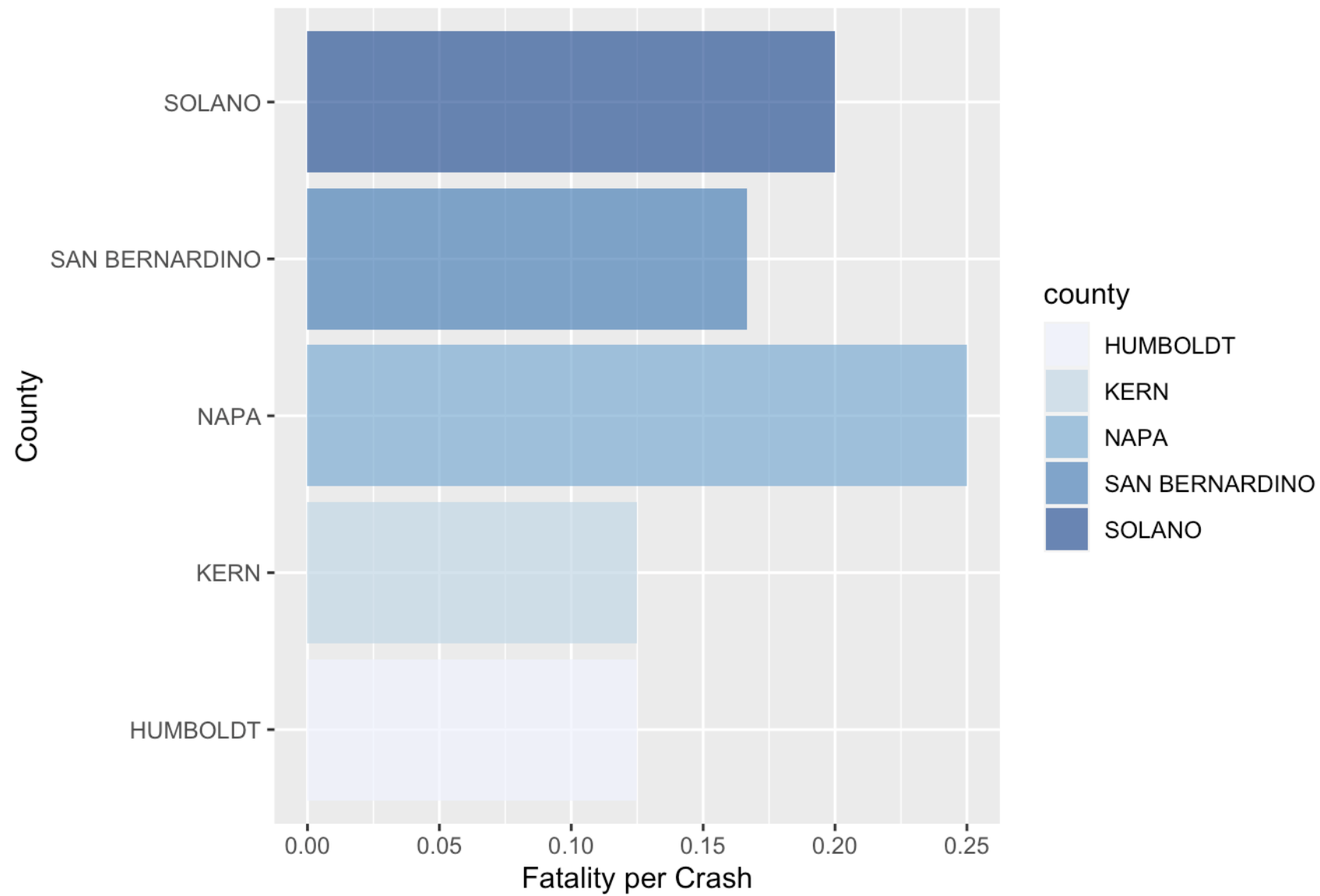
Select Y Variable

total



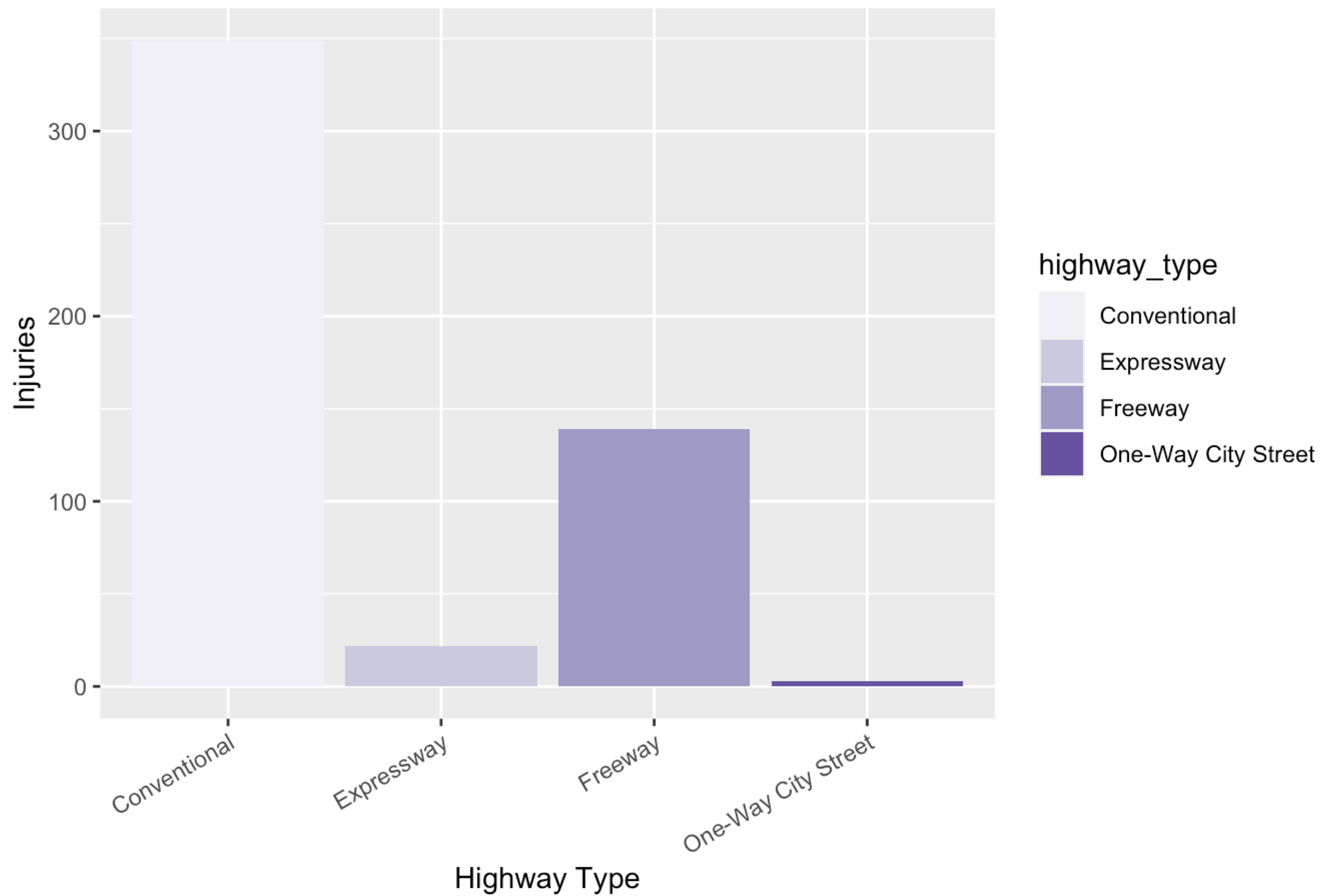
Which County Has the Highest
Fatality:Crash Ratio?

Fatalities per Crash by County



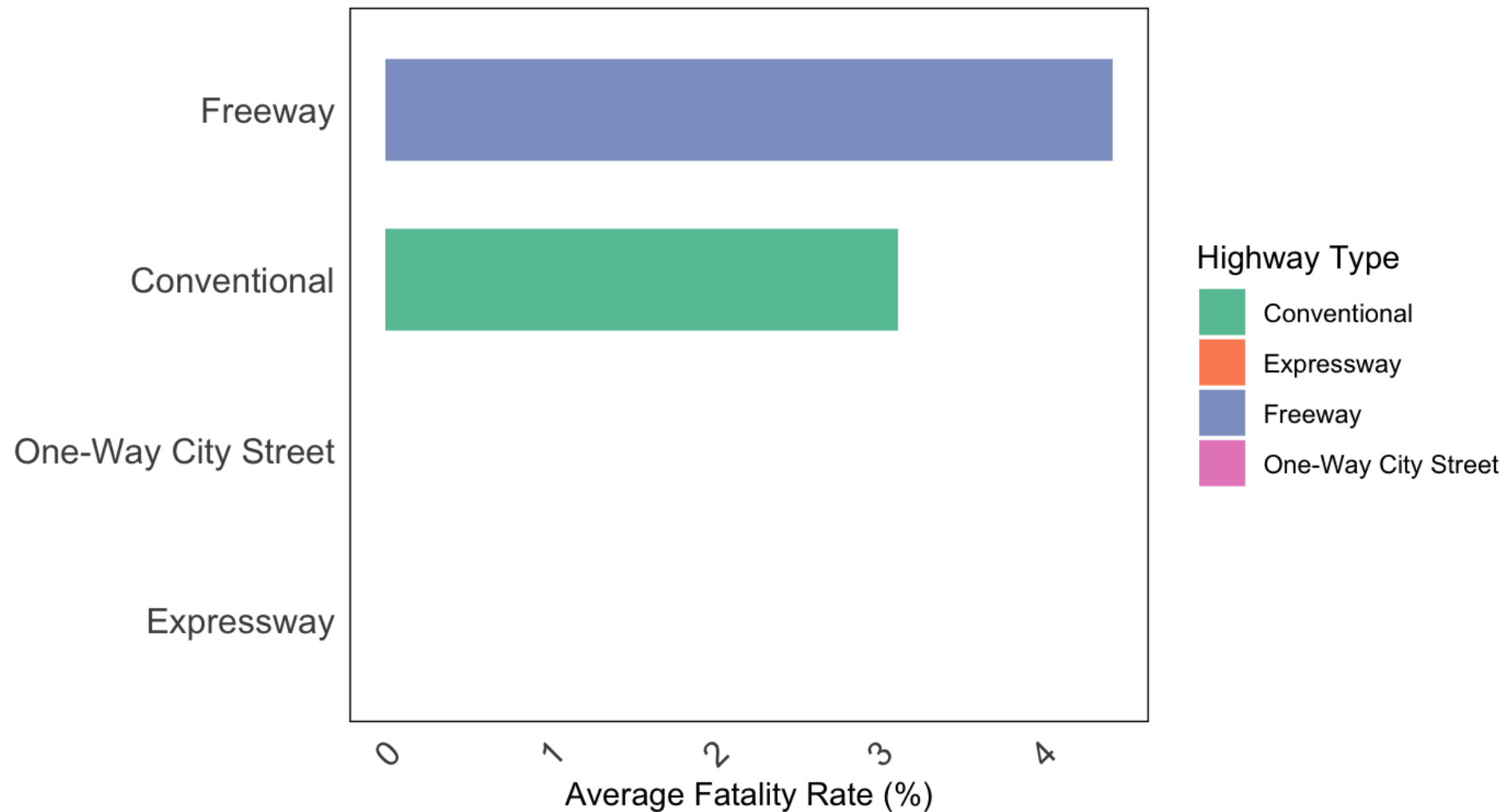
Which Type of Road Leads to the
Most injuries?

Injuries by Highway Type



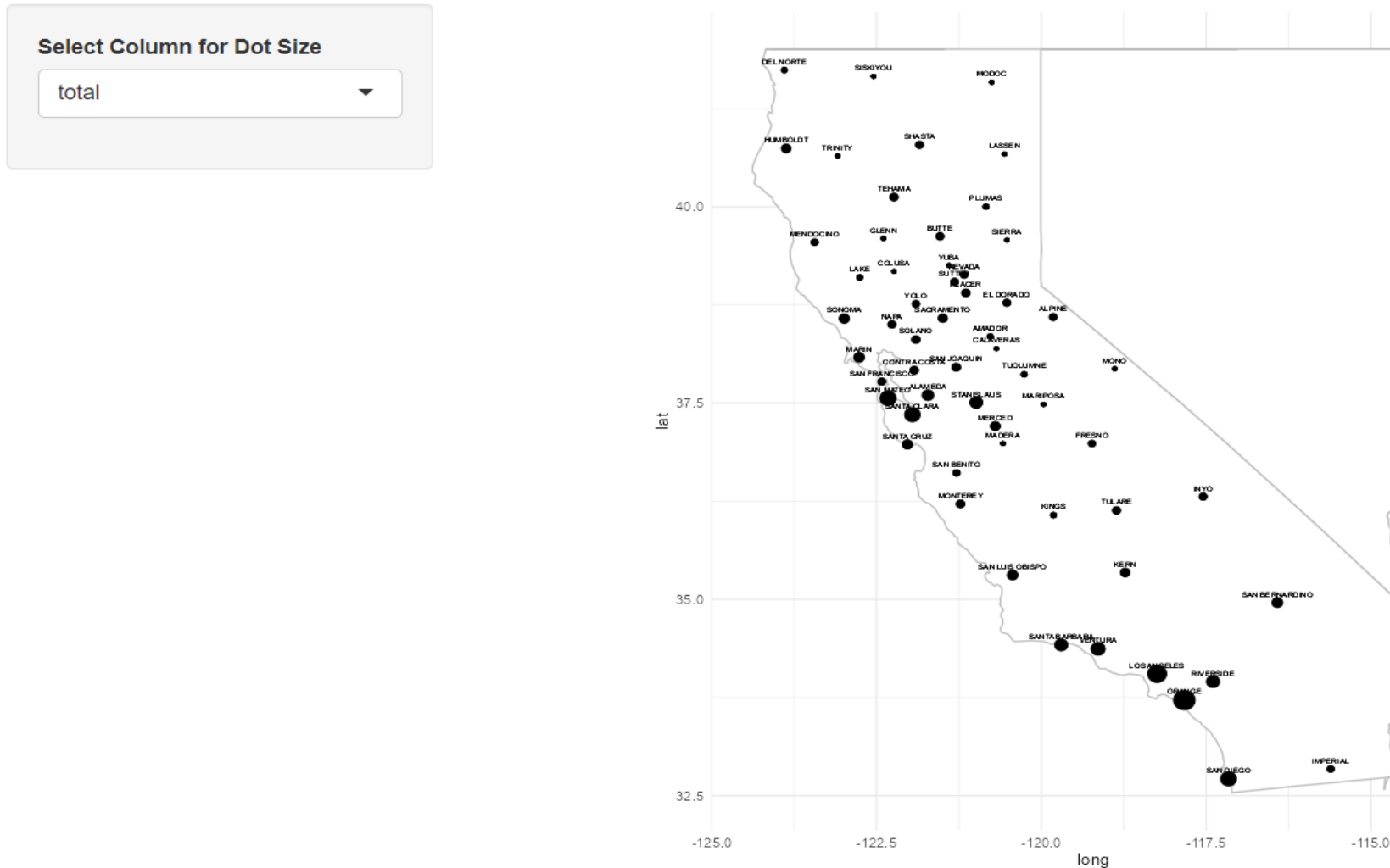
What street type is the most fatal?

Average Fatality Rate Among Types of Streets



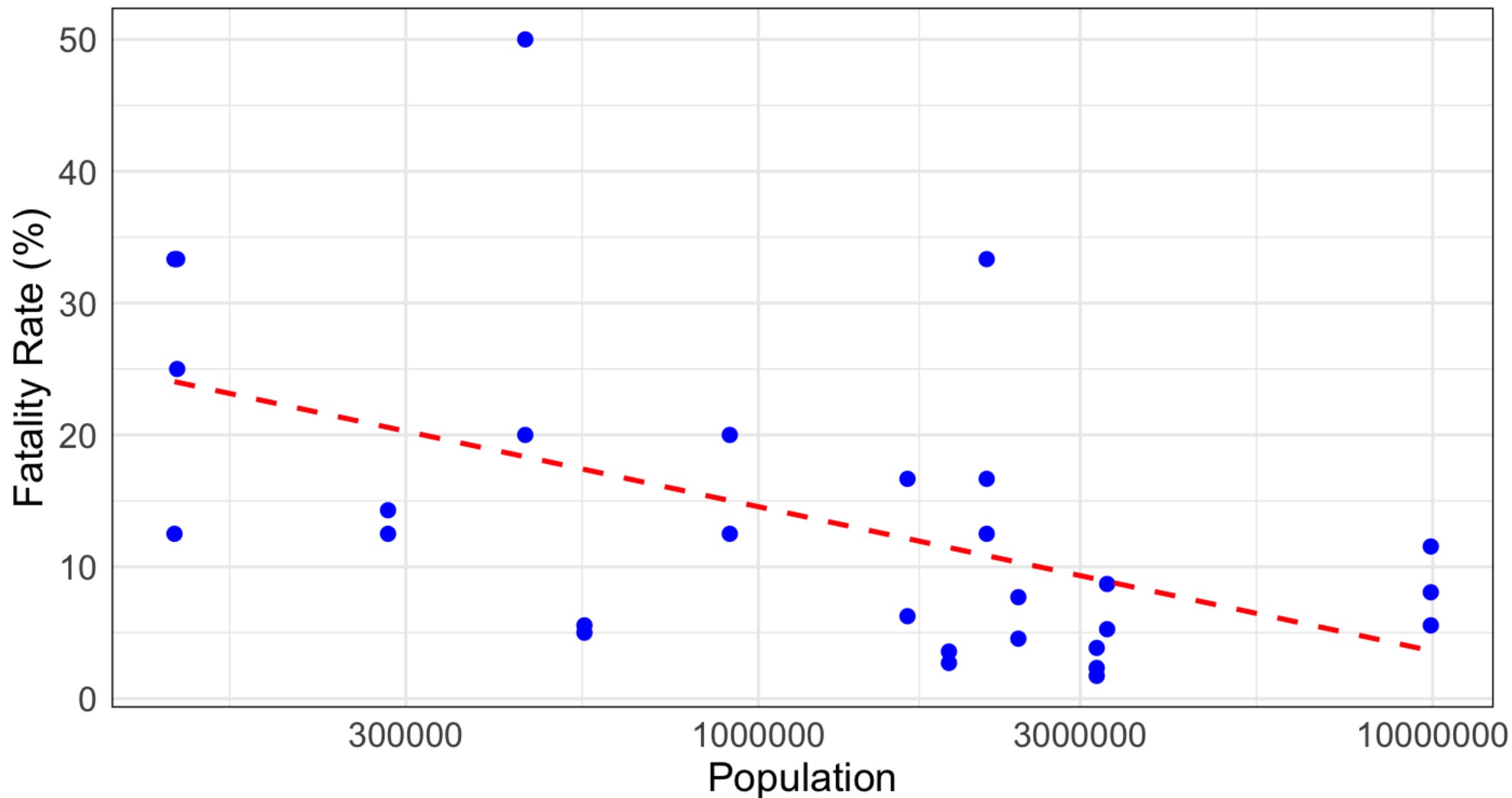
Map of total reported bike crashes in California

Bike Crash Data in California by County



Are you more likely to get into a fatal accident
in a county with a larger population?

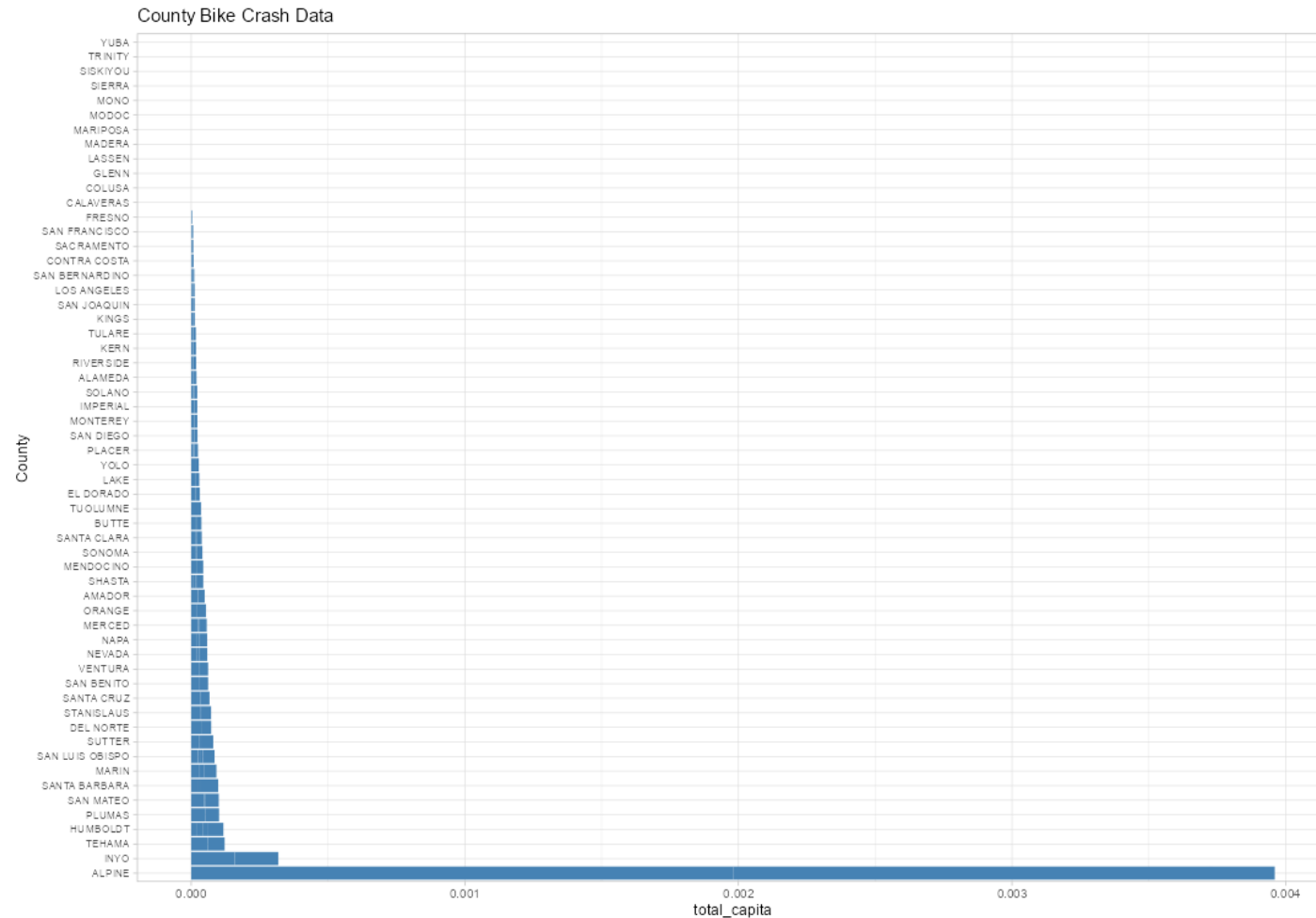
Correlation between Population Size and Fatality Rate



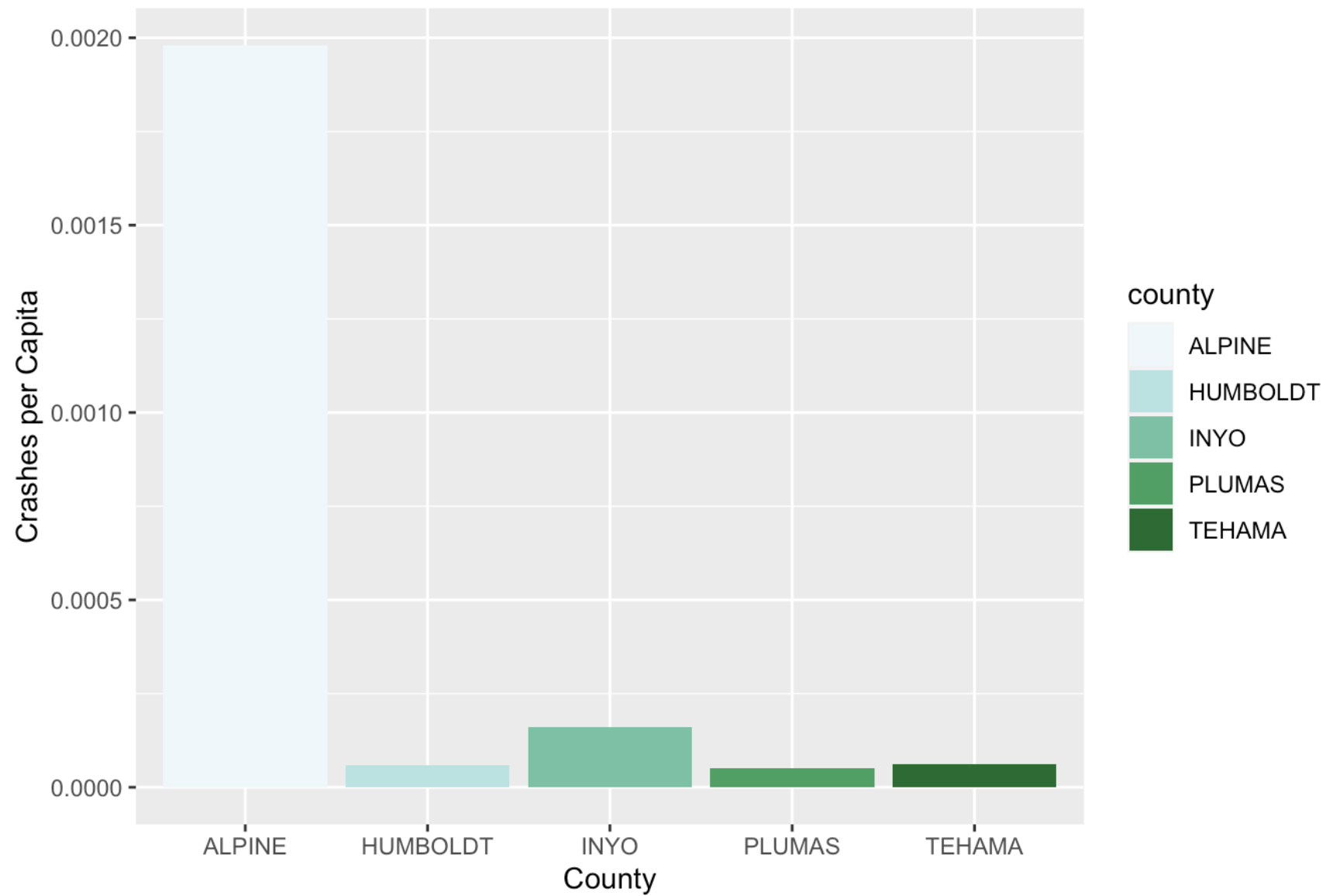
Which county has highest bike crash per CAP?

Select Y Variable

total_capita



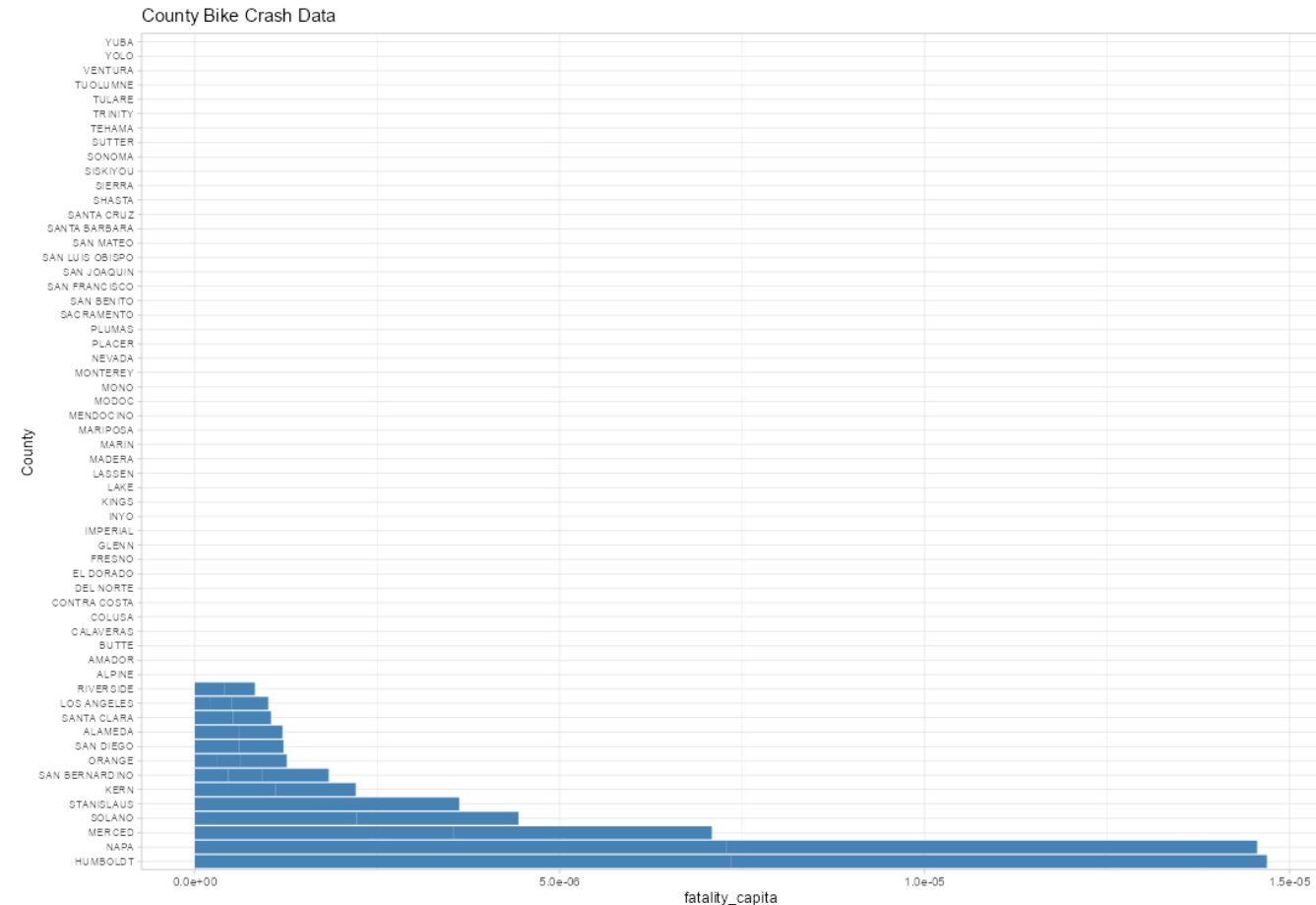
Crashes per Capita by County



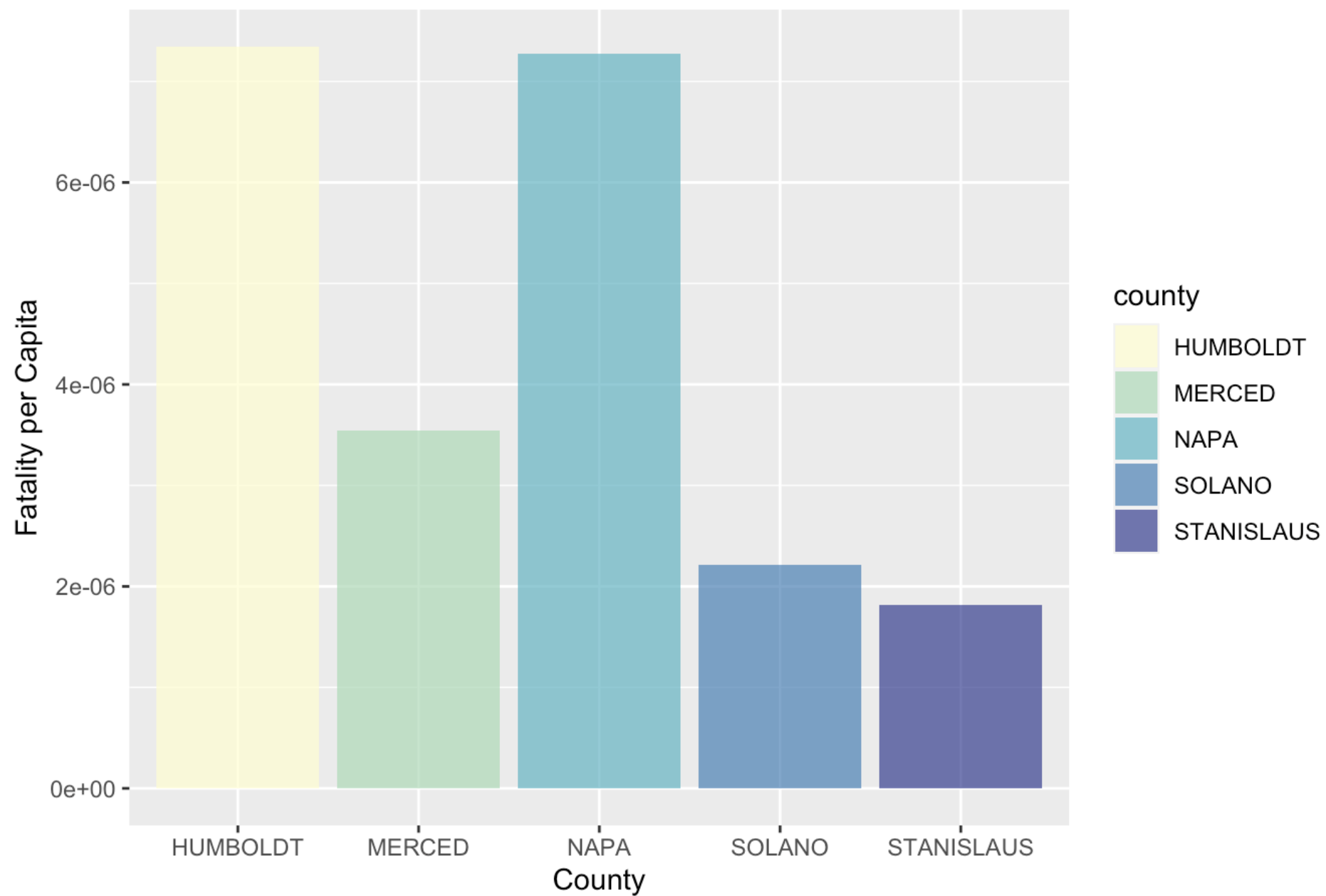
Which County Has the Highest Fatalities per Capita?

Select Y Variable

fatality_capita

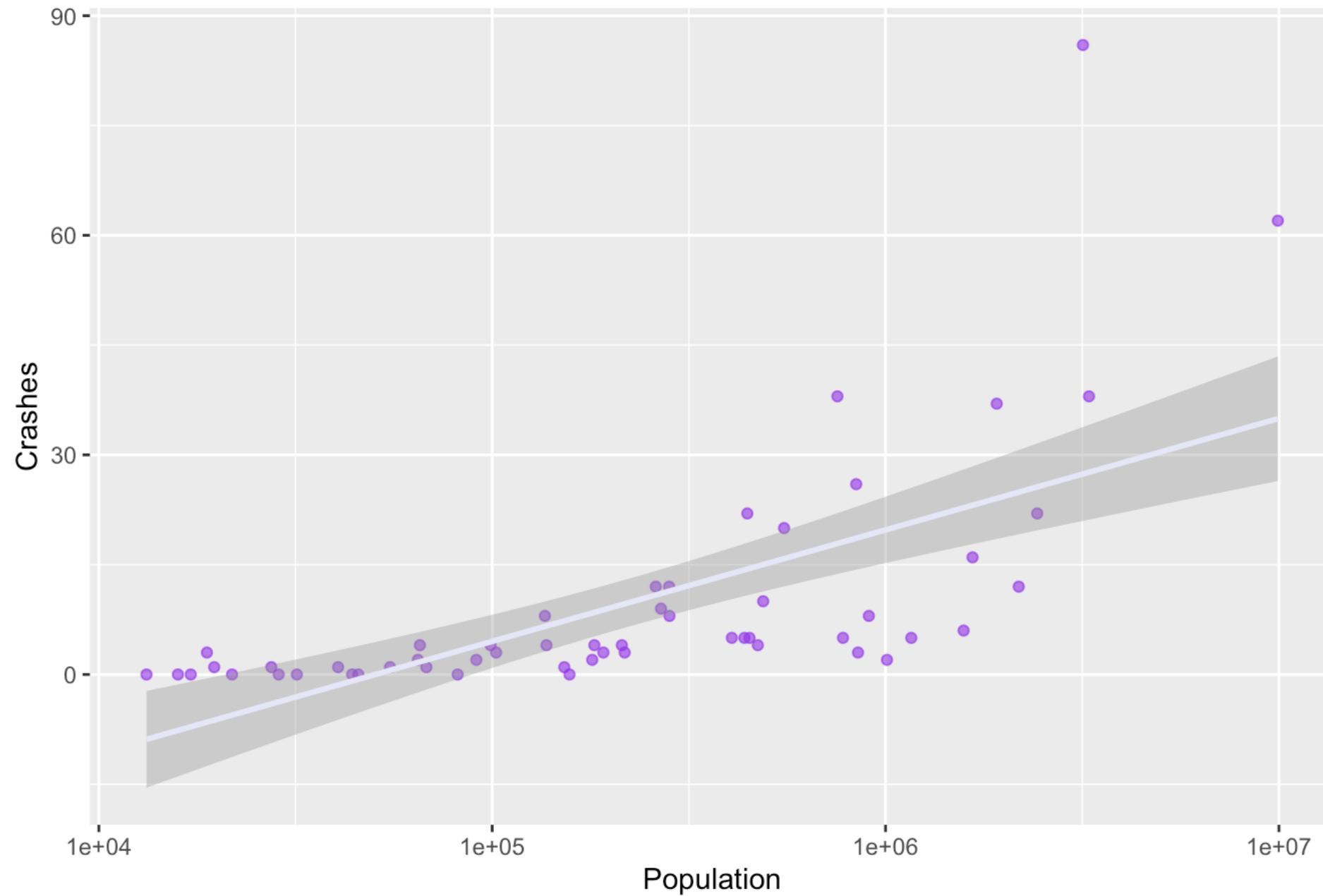


Fatality per Capita by County



Is There a Correlation Between
Population and Crashes?

Crashes by Population



Yolo County and Side Note about Data Science

- Yolo county, had only 3 reported injuries in 2022...
- Do we think only 3 people were hit on bikes in 2022?

