

Project 1

Topic: COVID Data vs Age & Race

Brought to you by:



Introduction & Problem

- We are looking at the relationship between certain COVID data metrics and demographic metrics
- COVID
 - Vaccination Status
 - Confirmed Case Totals
 - Death Count Totals
- Demographic
 - Age
 - Race

Our Data Source

LA Times Archive of California Coronavirus Data

<https://github.com/datadesk/california-coronavirus-data>

- California Public Dept. of Health - Covid-19 Cases by Age Group
(cdph-age.csv)
- California Public Dept. of Health - Covid-19 Cases by Race / Ethnicity
(cdph-race-ethnicity.csv)
- California Public Dept. of Health - Covid-19 Vaccinations by Age Group
(cdph-vaccination-state-by-age.csv)
- California Public Dept. of Health - Covid-19 Vaccinations by Race / Ethnicity
(cdph-vaccination-state-by-race-ethnicity.csv)

Research Question 1

Which **age groups** have the highest **vaccination numbers** in
California?

RQ1 Findings

What age groups have the highest vaccination numbers in California?

Bar Chart

- Shows that 18-49 is the largest age group
- 50-64 and 65+ were the same

Limitations

- Age groups
- Comparisons w/ US Census Data

RQ 1 EDA

What age groups have the highest vaccination numbers in California?

- Replaced certain values
 - Vacc_age =
vacc_age.replace(to_replace="17-Dec",value = "12-17")
 - Vacc_age =
vacc_age.replace(to_replace="11-May",value = "5-11")

Research Question 2

Which **races** have the highest **vaccination numbers** in
California?

RQ 2 Findings

What races have the highest vaccination numbers in California?

Pie Chart

- Whites and Latinos had the highest percentage of fully vaccinated
- Formatting

Limitations

- Unknown?
- Comparisons w/ US Census Data
 - California's population makeup is 39% Latino, 36% White, 15% Asian and 6% Black
 - Latinos 28.6% vaccinated
 - Asians 18% vaccinated

Research Question 3

Which races have the highest **positive test rates** and **death counts**?

RQ3 EDA

Which races have the highest positive
test rates and death counts?

- All Visuals (4)
 - Replaced/combined cdph-other into other
 - Coded to ignore any rows with specific age groups in data set
 - Only calculated the case total and death count total columns in rows with “all” as age group
- Bar Charts
 - Grouped by race
 - Descending order of confirmed cases and death totals
 - Capitalized race names for bar labels
- Other Formatting
 - Titles
 - Labels
 - Colors
 - Font Size
 - Figure Size
 - Rotation

RQ3 Findings

Which races have the highest positive
test rates and death counts?

- Latinos had the highest number of confirmed cases
 - 55% of total cases
- Blacks had the lowest number of cases
 - 4.6% of total cases
- Latinos also had the highest number of deaths
 - 46.5% of total deaths
- Blacks had the lowest number of deaths
 - 2.7% of total deaths

Research Question 4

Which **ages** have the highest **positive test rates** and **death counts**?

RQ 4 Findings

Which ages have the highest positive
test rates and death counts?

Race

Latinos and White Populations have the highest deaths in the state of California.

Latinos have the highest conversions from reported or confirmed cases to actual deaths at 59%.

Whites make up 33%, the second highest conversions from reported/confirmed cases to actual deaths.

Age

Ages 50-79 have low variance when it comes to deaths and average around 500,000 deaths.

Those above age 80 account for 71% of all deaths and have the highest death counts at 1,989,221.

Limitations & Challenges

1. When combining datasets, we have to take into account whether they're pulling from the original raw dataset with access to the same amount of data.
2. Our data only covers California, therefore, any assumptions or recommendations we make may not apply to other states or countries.
3. Due to time constraints, we did not use any datasets with current population counts, therefore we cannot make any per capita calculations.

