# Do vaccinated and unvaccinated respondents in underserved Chicago community areas differ in their trust for doctors, state/local officials, family/friends, and religious leaders?

Camargo, Lucky

2022-12-04

## Research Question

Do vaccinated and unvaccinated respondents in underserved Chicago community areas differ in their trust for doctors, state/local officials, family/friends, and religious leaders?

## Background

Vaccine hesitancy and antivaccination sentiments are labels ascribed to people in the US that refuse the COVID-19 vaccine. The lack of vaccine understanding and mistrust are oftentimes associated with people's refusal to take the COVID-19 vaccine. By analyzing data between vaccinated and unvaccinated respondents I am looking to identify the differences in trust between vaccinated and unvaccinated respndents. Survey information is taken from a Chicago based sample of 1362 survey respondents. In doing so, I will look to provide policy recommendations to help address vaccine hesitancy/antivaccination.

## Data Description:

Electronic and paper surveys were collected between July 2021 and April 2022 within Chicagoland. The survey respondents were 18 and over and were asked questions regarding the respondents vaccination status, demographics, experience getting vaccinated, motivators and barriers to vaccinations and trusted messengers. In total 1362 surveys were collected. The survey was funded by the Rockefeller Foundation and was conducted in five (5) U.S cities, with the Chicago Community Trust acting as the anchor organization in Chicago. Mathematica acted as the data collector for all 5 cities including Chicago, and Sinai Urban Health Institute acted as the data and evaluation lead for Chicago.

Independent Variables-narrowing it down to 2-3 - Vaccine status-dichotomous - Gender-dichotomous - Income-ordinal - Race/ethnicity-dummy variable - Age-ordinal - Education levels-ordinal - health insurance status

Dependent Variables Sources of trusted information-ordinal Likert Scale within ordinal level of measurement (1=A great deal; 2=A fair amount; 3=Not much; 4=Not at All)

- Religious leaders
- State and local government
- Your family and friends

- Your own doctor/healthcare provider

-Is the COVID-19 vaccine safe? Y/N

Data analysis consists of uploading data in excel format to R programming and performing a series of recoding, renaming of variables and creating a variety of datset summaries to generate the graphs/plots that indicate response percentages and demographic information. Diverging graphs and tables have been generated to visualize the data manipulation. The report includes the step by step data manipulation code used to obtain the outcomes. In some cases the data code structure was duplicated with the substitution of corresponding variables to render different visual outcomes.

Set working directory and load packages

```
setwd("C:/Users/17737/Desktop/R Files_434/project/DATA")

library(scales)
library(tidyverse)
library(readxl)
library(ggplot2)
library(kableExtra)
library(ggrepel)
```

Upload file

```
covidvax=read_excel("C:/Users/17737/Desktop/R Files_434/project/DATA/PulseData_2.xlsx")
```

Renaming the file to pulse_surv as a precaution and making conversions to the variable class to explore and do preliminary data manipulation

```
pulse_surv=covidvax


  pulse_surv$black=as.numeric(pulse_surv$black)
  pulse_surv$white=as.numeric(pulse_surv$white)
  pulse_surv$asian=as.numeric(pulse_surv$asian)
  pulse_surv$hispanic=as.numeric(pulse_surv$hispanic)
  pulse_surv$vax_stat=as.numeric(pulse_surv$vax_stat)
  pulse_surv$female=as.numeric(pulse_surv$female)
  pulse_surv$male=as.numeric(pulse_surv$male)
```

Recoding original female variable in dataset to a factor and simultaneously creating a new variable called gender

```
pulse_surv=
  pulse_surv%>%
  mutate(gender=recode(female,
                       "0" = "Male",
                       "1" = "Female"))


pulse_surv$gender=as.factor(pulse_surv$gender)
```

Recoding vax_stat variable into new variable vax_stat1 with categories/levels and converting it into a factor

```
pulse_surv=
  pulse_surv%>%
  mutate(vax_stat1=recode(vax_stat,
                          "0"= "no_vax",
                          "1"= "vax"))


pulse_surv$vax_stat1=as.factor(pulse_surv$vax_stat1)
```

Recoding education variable called educ variable into new variable educ1 with categories/levels and converting it into a factor

```
pulse_surv=
  pulse_surv%>%
  mutate(educ1=recode(educ,
                      "1"="8th or less",
                      "2"="some hs",
                      "3"="hs_GED",
                      "4"="trade_voc",
                      "5"="some col_2yrdeg",
                      "6"="bach",
                      "7"="master_highr",
                      "M"="NA"))


pulse_surv$educ1=as.factor(pulse_surv$educ1)
```

Recoding income variable called inc variable into new variable inc1 with categories/levels and converting it into a factor

```
pulse_surv=
  pulse_surv%>%
  mutate(inc1=recode(inc,
                      "1"="$9,999-below",
                      "2"="$10K-$19,999",
                      "3"="$20K-$29,999",
                      "4"="$30K-$39,999",
                      "5"="$40K-$59,999",
                      "6"="$60K-$79,999",
                      "7"="$80K-$99,999",
                      "8"="$100K-$149,999",
                      "9"="$150K-$199,999",
                      "10"="$200K-above",
                      "R"= "NA",
                      "M"="NA"))


pulse_surv$inc1=as.factor(pulse_surv$inc1)
```

There are dummy variables for each race/ethnicity variable.Combining the dummy variables for race-there are four. Doing so by finding the race/ethinicy column numbers and then using the which command and then applying the loop function shown below

```
which( colnames(pulse_surv)=="black" )
```

```
## [1] 49
```

```
which( colnames(pulse_surv)=="asian" )
```

```
## [1] 50
```

```
which( colnames(pulse_surv)=="white" )
```

```
## [1] 51
```

```
which( colnames(pulse_surv)=="hispanic" )
```

```
## [1] 52
```

```
pulse_surv$re = names(pulse_surv[49:52])[apply(pulse_surv[49:52], 1, match, x = 1)]

pulse_surv$re=as.factor(pulse_surv$re)
```

Explore dataset: histogram of age frequencies

```
hist(pulse_surv$age)
```

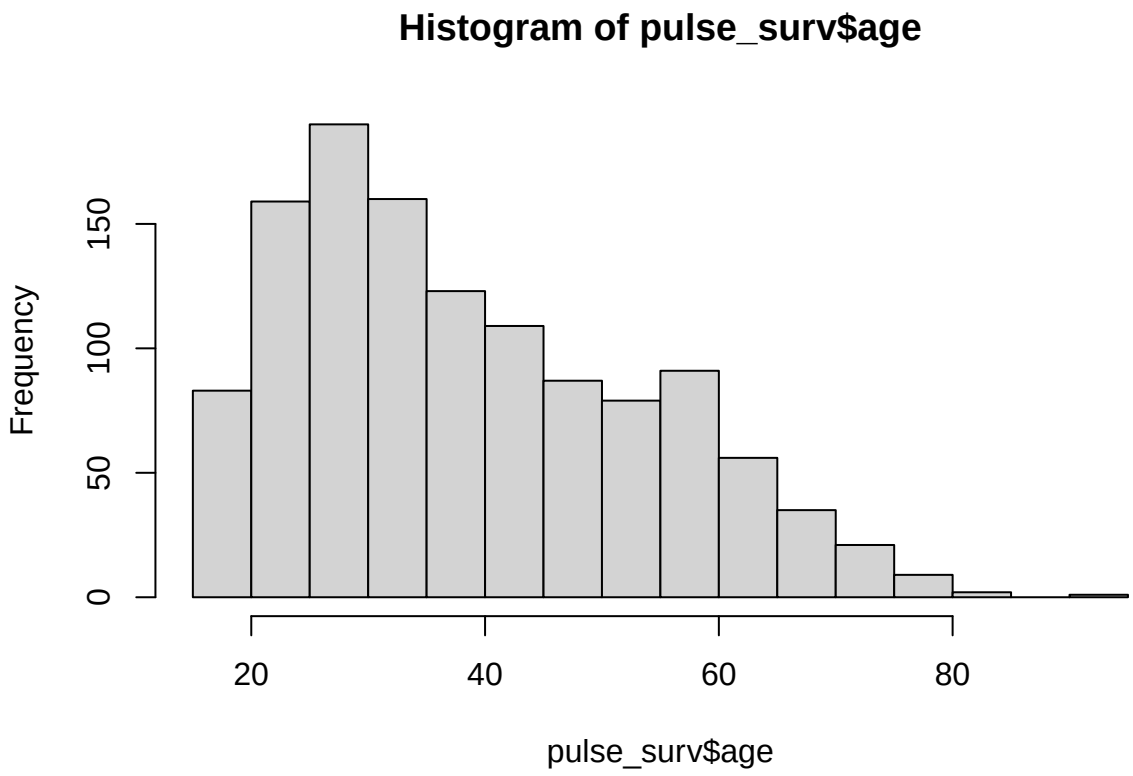


Table 1: Pulse Survey Respondent count by Vaccination Status

| Vaccination status | Black | Hispanic | White | Asian |
|---|---|---|---|---|
| no_vax | 114 | 57 | 33 | 6 |
| vax | 447 | 482 | 197 | 33 |

Recoding and renaming variable that asks respondents to rate the statement: The COVID-19 vaccine is safe scale from 1 to 5, where 1 means strongly disagree and 5 means strongly agree.

New recoded variable vaxsafe_r

```
pulse_surv=
  pulse_surv%>%
  mutate(vaxsafe_r=recode(q7_a_vaxsafe,
                          "1"="strongly disagree",
                          "2"="somewhat disagree",
                          "3"="neutral",
                          "4"="somewhat agree",
                          "5"="strongly agree",
                          "M"="NA"))
```

Creating dataset that provides vaccination frequencies by race/ethnicity and vaccination status

```
re1=
  pulse_surv%>%
select(black,white,hispanic,asian,vax_stat1)%>%
  group_by(vax_stat1)%>%
summarise(blk=sum(black),
hispanic =sum(hispanic),
white=sum(white),
asian=sum(asian))


re1%>%
  kable(col.names=c("Vaccination status", "Black","Hispanic","White","Asian"),
                    align="clc",
                    caption="Pulse Survey Respondent count by Vaccination Status",
                    digits=c(0,2,2),
                    format.arg=list(big.mark=","))%>%


kable_classic(font_size=20,
              full_width = F,
              html_font= "Cambria")
```

```
re2=
pulse_surv%>%
  select(vax_stat1, re)%>%
  filter(!is.na(re))%>%
```

Table 2: Pulse Survey Respondent count

| Race/Ethnicity | Vaccination Status | Count | Percent |
|---|---|---|---|
| asian | no_vax | 2 | 6% |
| asian | vax | 32 | 94% |
| black | no_vax | 114 | 20% |
| black | vax | 447 | 80% |
| hispanic | no_vax | 52 | 10% |
| hispanic | vax | 454 | 90% |
| white | no_vax | 30 | 14% |
| white | vax | 191 | 86% |

```
  group_by(re, vax_stat1)%>%
  count(name="n_answers")%>%
  group_by(re)%>%
  mutate(percent_answers=n_answers/sum(n_answers))%>%
  ungroup()%>%
  mutate(percent_answers_label = percent(percent_answers, accuracy=1))%>%
  select(re, vax_stat1, n_answers, percent_answers_label)


re2%>%
  kable(col.names=c("Race/Ethnicity", "Vaccination Status","Count","Percent"),
                    align="clc",
                    caption="Pulse Survey Respondent count",
                    digits=c(0,2,2),
                    format.arg=list(big.mark=","))%>%


kable_classic(font_size=20,
              full_width = F,
              html_font= "Cambria")
```

#Create a data set that consists of unvaccinated responses by race with their responses to vax safe_r qs

```
unvaxed_re_vaxsafe=
  pulse_surv%>%
  select(vax_stat1, re, vaxsafe_r)%>%
  filter(vax_stat1=="no_vax")%>%
  filter(vaxsafe_r!= "NA")%>%
  filter(!is.na(re))%>%
  group_by(re, vaxsafe_r)%>%
  count(name="n_answers")%>%
  group_by(re)%>%
```

```
  mutate(percent_answers=n_answers/sum(n_answers))%>%
  ungroup()%>%
  mutate(percent_answers_label = percent(percent_answers, accuracy=1))



head(unvaxed_re_vaxsafe, 5)
```

```
## # A tibble: 5 x 5
##   re    vaxsafe_r         n_answers percent_answers percent_answers_label
##   <fct> <chr>                 <int>           <dbl> <chr>
## 1 asian neutral                   1          0.5    50%
## 2 asian somewhat disagree         1          0.5    50%
## 3 black neutral                  37          0.336  34%
## 4 black somewhat agree           10          0.0909 9%
## 5 black somewhat disagree        18          0.164  16%
```

Adjusting the percentages with the negative feedback as negatives to position to one side of the plot

```
unvaxed_re_vaxsafe_diverging=
  unvaxed_re_vaxsafe%>%
  mutate(percent_answers=if_else(vaxsafe_r %in% c("strongly agree", "somewhat agree", 'neutral'),
                                 percent_answers, -percent_answers))%>%
  mutate(percent_answers_label=percent(percent_answers, accuracy=1))
```

Converting the negative label to absolute values so that negative symbols are not revealed on the plot

```
unvaxed_re_vaxsafe_diverging_good_labels=
  unvaxed_re_vaxsafe_diverging%>%
  mutate(percent_answers_label =abs(percent_answers))%>% #converts the negative values to absolute
  mutate(percent_answers_label= percent(percent_answers_label, accuracy =1))
```

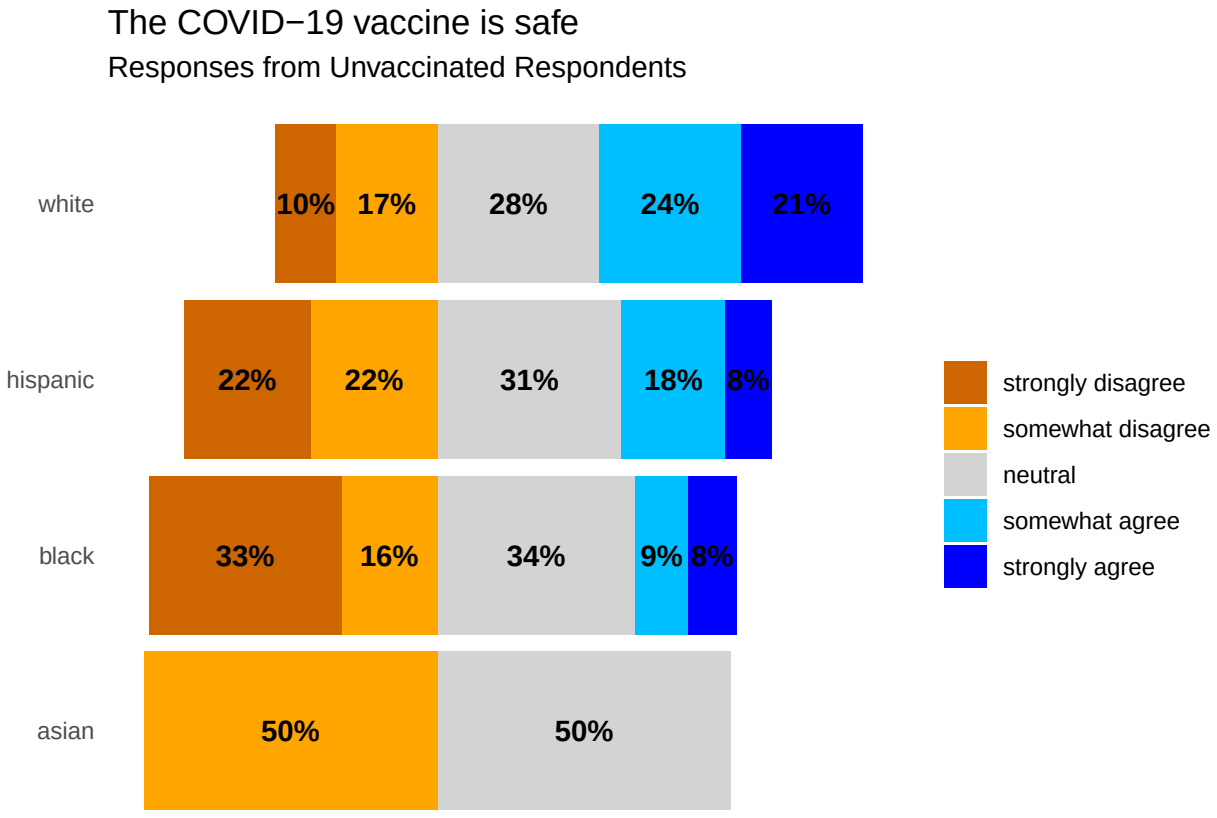
Reordering the bars and creating new dataset

```
unvaxed_re_vaxsafe_diverging_right_order=
  unvaxed_re_vaxsafe_diverging_good_labels%>%
  mutate(vaxsafe_r=fct_relevel(vaxsafe_r,
                               "strongly disagree", "somewhat disagree",
                              "strongly agree" ,"somewhat agree", "neutral"))
```

Plot that displays the perception of UNVACCINATED by race/ethnicity The COVID-19 vaccine is safe

```
unvaxed_re_vaxsafe_diverging_right_order%>%
  ggplot(aes(x = re,
             y = percent_answers,
             fill = vaxsafe_r)) +
  geom_col() +
  geom_text(aes(label = percent_answers_label),
            position = position_stack(vjust = 0.5),
            color = "black",
            fontface = "bold") +
```

```
coord_flip() +
scale_x_discrete() +

scale_fill_manual(breaks=c("strongly disagree", "somewhat disagree", "neutral", "somewhat agree", "st
                  values=c(
                    "strongly disagree"="darkorange3",
                    "somewhat disagree"="orange",
                    "neutral"= "light grey",
                    "somewhat agree"="deepskyblue",
                    "strongly agree"="blue"
                  ))+


labs(title = "The COVID-19 vaccine is safe",
     subtitle = "Responses from Unvaccinated Respondents",
     x = NULL,
     fill = NULL) +
theme_minimal() +
theme(axis.text.x = element_blank(),
      axis.title.x = element_blank(),
      panel.grid = element_blank(),
      legend.position = "right")
```



## Create a data set that consists of vaccinated responses by race with their responses to vax safe_r qs Vaccinated by race -ask is vaccine is safe

Create a data set that is vaccinated by race and responses to vax safe_r qs vaxstat1 variable for vaccination status Vaxsafe_r variable for statement asking opinion about the safety of vaccine re variable of race and ethnicty group

```
vaxed_re_vaxsafe=
  pulse_surv%>%
  select(vax_stat1, re, vaxsafe_r)%>%
  filter(vax_stat1=="vax")%>%
  filter(vaxsafe_r!= "NA")%>%
  filter(!is.na(re))%>%
  group_by(re, vaxsafe_r)%>%
  count(name="n_answers")%>%
  group_by(re)%>%
  mutate(percent_answers=n_answers/sum(n_answers))%>%
  ungroup()%>%
  mutate(percent_answers_label = percent(percent_answers, accuracy=1))
```

Adjusting the percentages with the negative feedback as negatives to position to one side

```
vaxed_re_vaxsafe_diverging=
  vaxed_re_vaxsafe%>%
  mutate(percent_answers=if_else(vaxsafe_r %in% c("strongly agree", "somewhat agree", 'neutral'),
                                 percent_answers, -percent_answers))%>%
  mutate(percent_answers_label=percent(percent_answers, accuracy=1))
```

Converting the negative labels to absolute values-create a new dataset

```
vaxed_re_vaxsafe_diverging_good_labels=
  vaxed_re_vaxsafe_diverging%>%
  mutate(percent_answers_label =abs(percent_answers))%>% #converts the negative values to absolute
  mutate(percent_answers_label= percent(percent_answers_label, accuracy =1))
```

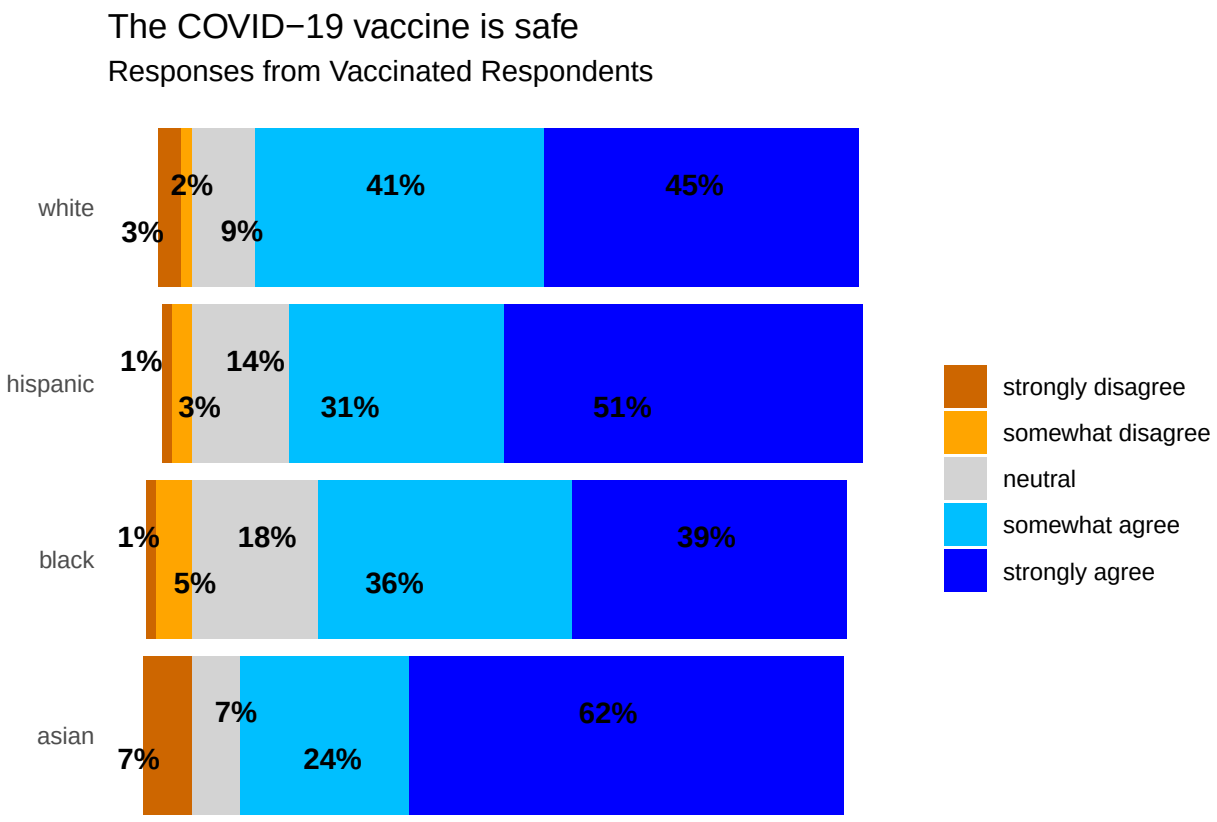
Now releveling the vaxsafe_r responses to display coherently on plot -create new dataset

```
vaxed_re_vaxsafe_diverging_right_order=
  vaxed_re_vaxsafe_diverging_good_labels%>%
  mutate(vaxsafe_r=fct_relevel(vaxsafe_r,
                               "strongly disagree", "somewhat disagree",
                               "strongly agree" ,"somewhat agree", "neutral"))
```

Plot that displays the perception of VACCINATED by race/ethnicity The COVID-19 vaccine is safe

```
vaxed_re_vaxsafe_diverging_right_order%>%
  ggplot(aes(x = re,
             y = percent_answers,
             fill = vaxsafe_r)) +
  geom_col() +
```

```
 geom_text_repel(aes(label = percent_answers_label),
           position = position_stack(vjust = 0.4),
           color = "black",
           fontface = "bold") +
coord_flip() +
scale_x_discrete() +

scale_fill_manual(breaks=c("strongly disagree", "somewhat disagree", "neutral", "somewhat agree", "st
                   values=c(
                     "strongly disagree"="darkorange3",
                     "somewhat disagree"="orange",
                     "neutral"= "light grey",
                     "somewhat agree"="deepskyblue",
                     "strongly agree"="blue"
                   ))+
labs(title = "The COVID-19 vaccine is safe",
     subtitle = "Responses from Vaccinated Respondents",
     x = NULL,
     fill = NULL) +
theme_minimal() +
theme(axis.text.x = element_blank(),
      axis.title.x = element_blank(),
      panel.grid = element_blank(),
      legend.position = "right")
```




This is dataset summary created to prepare for plot that will consist of unvaccinated survey respondents by

health insurance status and race

```
Unvaxed_by_ins_by_re=
  pulse_surv%>%
  select(vax_stat1, re, health_ins)%>%
  filter(vax_stat1=="no_vax")%>%
  filter(!is.na(health_ins))%>%
  filter(!is.na(re))%>%
  filter(health_ins!="M")%>%
  group_by(re, health_ins)%>%
  count(name="n_answers")%>%
  group_by(re)%>%
  mutate(percent_answers=n_answers/sum(n_answers))%>%
  ungroup()%>%
  mutate(percent_answers_label = percent(percent_answers, accuracy=1))



head(Unvaxed_by_ins_by_re, 5)
```

```
## # A tibble: 5 x 5
##   re       health_ins n_answers percent_answers percent_answers_label
##   <fct>    <chr>          <int>           <dbl> <chr>
## 1 asian    0                  1           0.5   50%
## 2 asian    1                  1           0.5   50%
## 3 black    0                 25           0.223 22%
## 4 black    1                 87           0.777 78%
## 5 hispanic 0                 26           0.5   50%
```

Recoding the variable health ins and created new dataset name

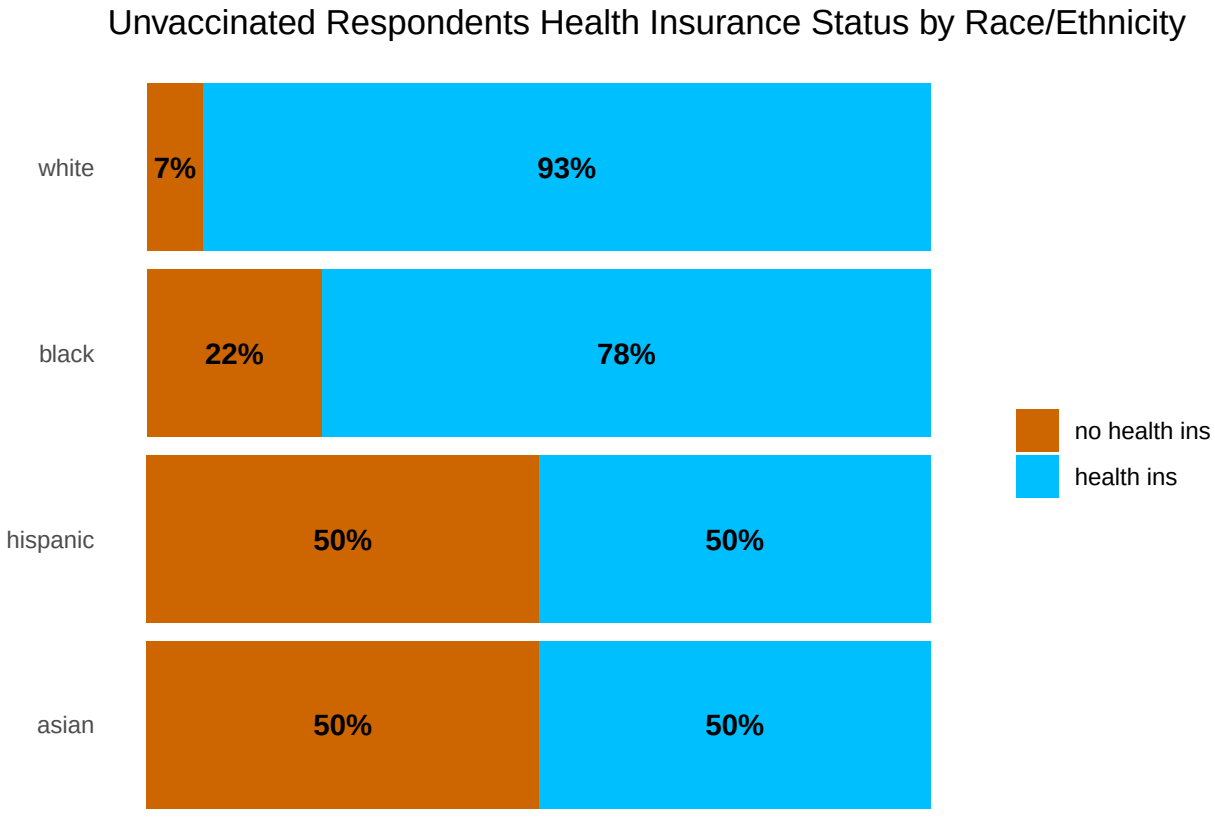
```
Unvaxed_by_ins_by_re2=
  Unvaxed_by_ins_by_re%>%
  mutate(health_ins=recode(health_ins,
                           "1"="health ins",
                           "0"="no health ins"))%>%
  mutate(re=fct_relevel(re,
                        "asian",
                        "hispanic",
                        "black",
                        "white"))
```

Exploratory plot showing the percentage of unvaccinated survey respondents with with health insurance by race/ethnicity:

Insert comments and observations here

```
Unvaxed_by_ins_by_re2%>%
  ggplot(aes(x = re,
             y = percent_answers,
             fill = health_ins)) +
  geom_col() +
  geom_text(aes(label = percent_answers_label),
```

```
            position = position_stack(vjust = 0.5),
            color = "black",
            fontface = "bold") +
  coord_flip() +
  scale_x_discrete() +
  scale_fill_manual(breaks =c("no health ins", "health ins"),#changes the order of the legend text list
                    values = c(
                      "no health ins"= "darkorange3",
                      "health ins"= "deepskyblue"
                    )) +
  labs(title = "Unvaccinated Respondents Health Insurance Status by Race/Ethnicity",
       x = NULL,
       fill = NULL) +
  theme_minimal() +
  theme(axis.text.x = element_blank(),
        axis.title.x = element_blank(),
        panel.grid = element_blank(),
        legend.position = "right")
```



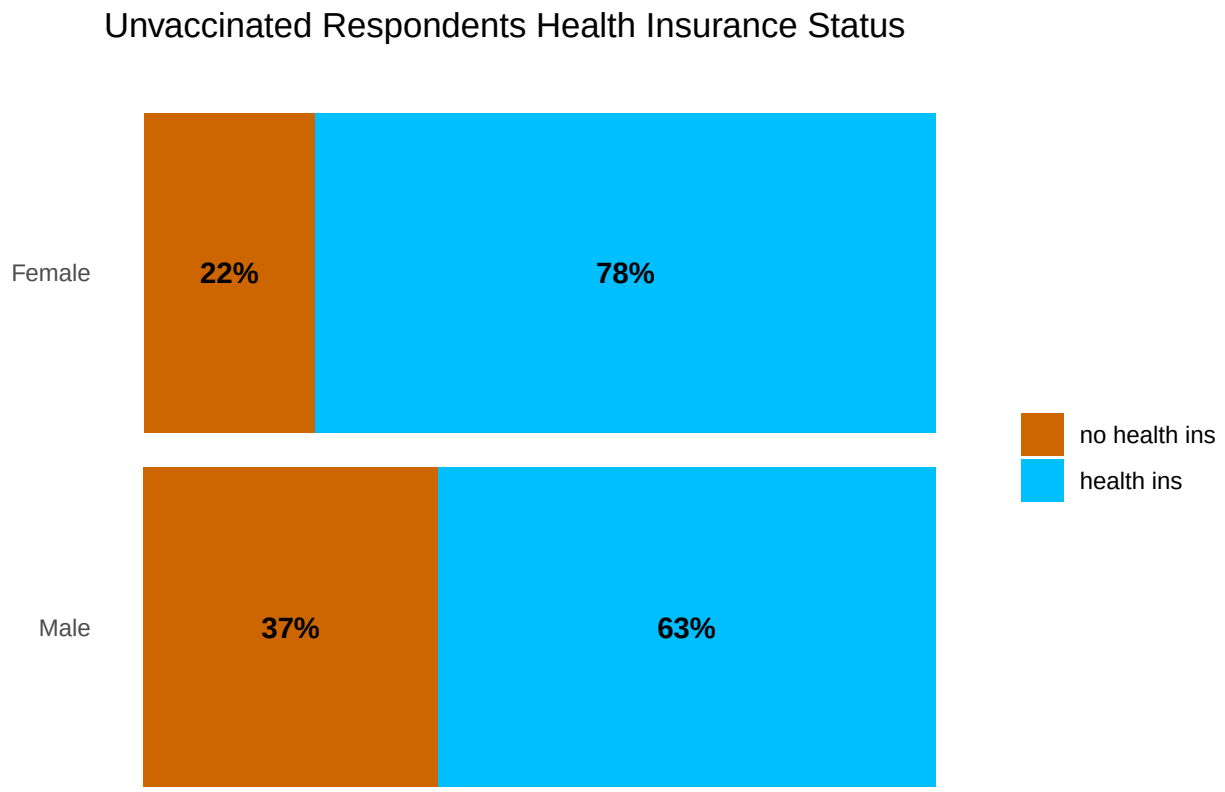

Data manipulation for exploratory plot of Unvaccinated respondents by gender and insurance status Creating dataset summary for unvaccinated respondents by gender and health insurance status Recoding the dataset column variable health ins and created new dataset name in the process

```
Unvaxed_by_ins_by_gender=
pulse_surv%>%
  select(vax_stat1, gender, health_ins)%>%
```

```
  filter(vax_stat1=="no_vax")%>%
  filter(!is.na(health_ins))%>%
  filter(!is.na(gender))%>%
  filter(health_ins!="M")%>%
  group_by(gender, health_ins)%>%
  count(name="n_answers")%>%
  group_by(gender)%>%
  mutate(percent_answers=n_answers/sum(n_answers))%>%
  ungroup()%>%
  mutate(percent_answers_label = percent(percent_answers, accuracy=1))




Unvaxed_by_ins_by_gender2=
  Unvaxed_by_ins_by_gender%>%
  mutate(health_ins=recode(health_ins,
                           "1"="health ins",
                           "0"="no health ins"))%>%
  mutate(gender=fct_relevel(gender,
                       "Male",
                        "Female"))
```

Exploratory plot of Unvaccinated respondents by gender and insurance status

```
Unvaxed_by_ins_by_gender2%>%
  ggplot(aes(x = gender,
             y = percent_answers,
             fill = health_ins)) +
  geom_col() +
  geom_text(aes(label = percent_answers_label),
            position = position_stack(vjust = 0.5),
            color = "black",
            fontface = "bold") +
  coord_flip() +
  scale_x_discrete() +
  scale_fill_manual(breaks =c("no health ins", "health ins"),#changes the order of the legend text list
                    values = c(
                      "no health ins"= "darkorange3",
                      "health ins"= "deepskyblue"
                    )) +
  labs(title = "Unvaccinated Respondents Health Insurance Status",
       x = NULL,
       fill = NULL) +
  theme_minimal() +
  theme(axis.text.x = element_blank(),
        axis.title.x = element_blank(),
        panel.grid = element_blank(),
        legend.position = "right")
```




Data manipulation for exploratory plot of Unvaccinated Respondents by education Creating dataset summary for unvaccinated respondents by education Recoding the unvaccinated dataset summary column variable education

```
Unvaxed_by_ed=
  pulse_surv%>%
  select(vax_stat1, educ1)%>%
  filter(vax_stat1=="no_vax")%>%
  filter(!is.na(educ1))%>%
  filter(educ1!="NA")%>%
  group_by(educ1)%>%
  count(name="n_answers")%>%
  mutate(percent_answers=n_answers/201)%>%
  mutate(percent_answers_label = percent(percent_answers, accuracy=1))




Unvaxed_by_ed2=
  Unvaxed_by_ed%>%
  mutate(educ1=fct_relevel(educ1,
                           "8th or less",
                           "some hs",
                           "hs_GED",
```

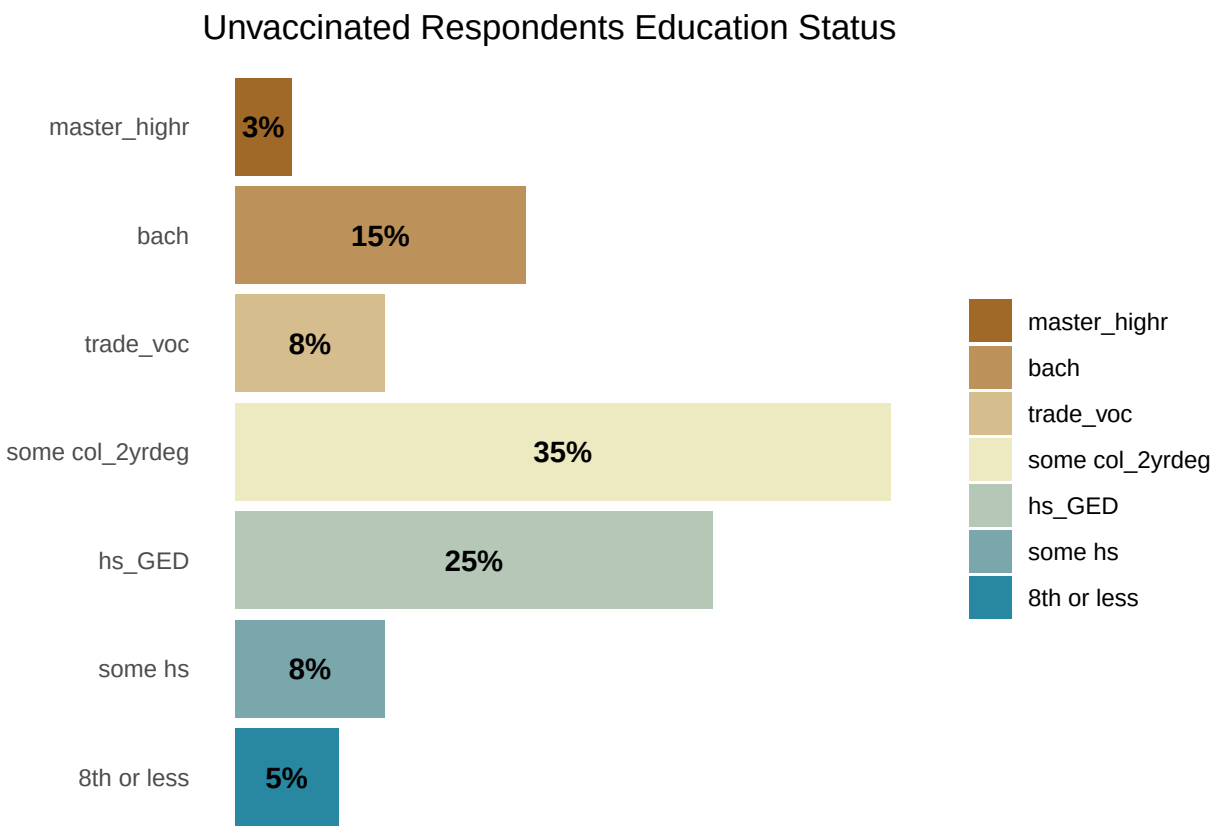
```
                          "some col_2yrdeg",
                          "trade_voc",
                          "bach",
                          "master_highr"))



head(Unvaxed_by_ed, 5)
```

```
## # A tibble: 5 x 4
## # Groups:   educ1 [5]
##   educ1           n_answers percent_answers percent_answers_label
##   <fct>               <int>           <dbl> <chr>
## 1 8th or less            11          0.0547 5%
## 2 bach                   31          0.154  15%
## 3 hs_GED                 51          0.254  25%
## 4 master_highr            6          0.0299 3%
## 5 some col_2yrdeg        70          0.348  35%
```

Exploratory plot of Unvaccinated Respondents by education

```
Unvaxed_by_ed2%>%
  ggplot(aes(x = educ1,
             y = percent_answers,
             fill = educ1)) +
  geom_col() +
  geom_text(aes(label = percent_answers_label),
            position = position_stack(vjust = 0.5),
            color = "black",
            fontface = "bold") +
  coord_flip() +
  scale_x_discrete() +
  scale_fill_manual(breaks =c("master_highr",
                          "bach",
                          "trade_voc",
                          "some col_2yrdeg",
                          "hs_GED",
                           "some hs",
                           "8th or less"),#changes the order of the legend text listing
                    values = c(
                      "master_highr"= "#A16928FF",
                      "bach"= "#BD925AFF",
                      "trade_voc"= "#D6BD8DFF",
                      "some col_2yrdeg"= "#EDEAC2FF",
                      "hs_GED"= "#B5C8B8FF",
                      "some hs"= "#79A7ACFF",
                      "8th or less"= "#2887A1FF"
                    )) +
  labs(title = "Unvaccinated Respondents Education Status",
       x = NULL,
       fill = NULL) +
  theme_minimal() +
  theme(axis.text.x = element_blank(),
```

```
        axis.title.x = element_blank(),
        panel.grid = element_blank(),
        legend.position = "right")
```




Data manipulation for exploratory plot of unvaccinated respondents by income Creating dataset summary for unvaccinated by income Releveling of the income variable

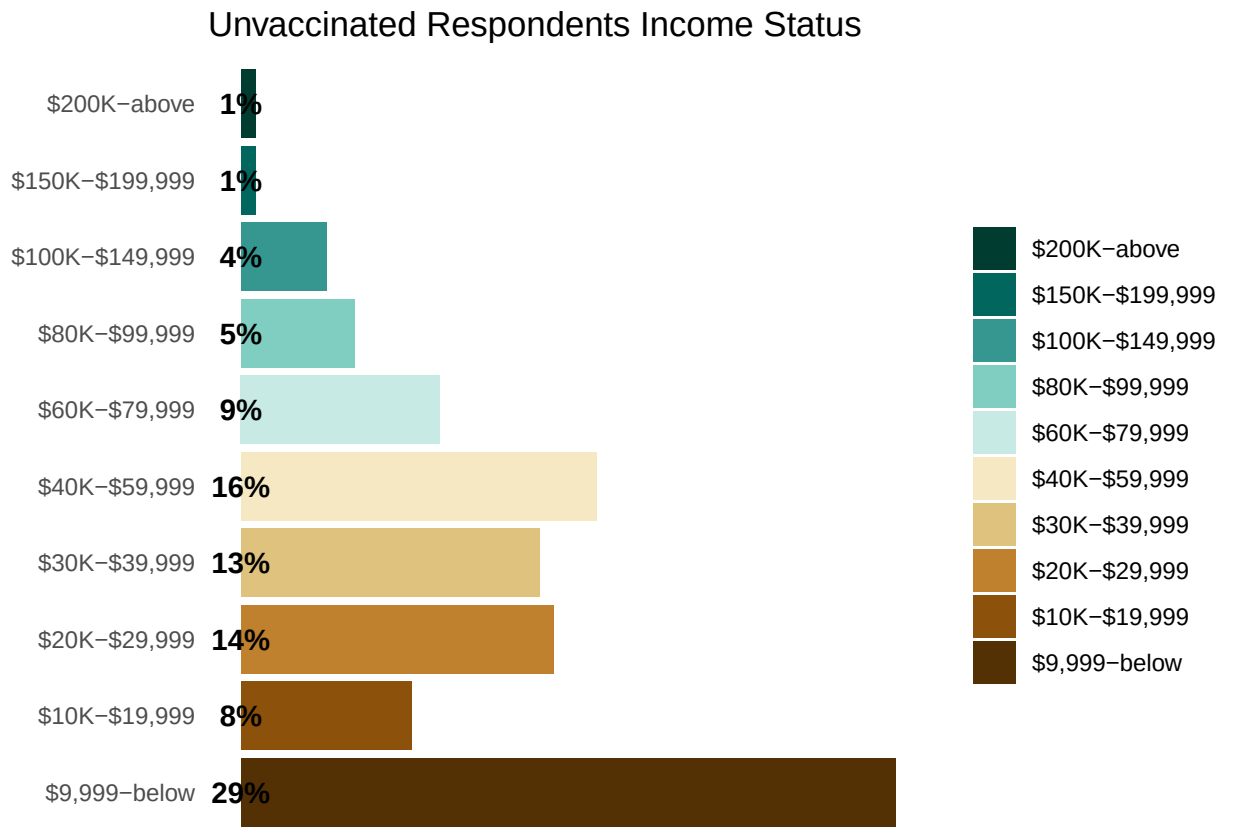
```
Unvaxed_by_inc=
  pulse_surv%>%
  select(vax_stat1, inc1)%>%
  filter(vax_stat1=="no_vax")%>%
  filter(!is.na(inc1))%>%
  filter(inc1!="NA")%>%
  group_by(inc1)%>%
  count(name="n_answers")%>%
  mutate(percent_answers=n_answers/156)%>% #some respondents refused to provide income
  mutate(percent_answers_label = percent(percent_answers, accuracy=1))
```

```
Unvaxed_by_inc2=
  Unvaxed_by_inc%>%
  mutate(inc1=fct_relevel(inc1,
                          "$9,999-below",
                          "$10K-$19,999",
                          "$20K-$29,999",
                          "$30K-$39,999",
                          "$40K-$59,999",
                          "$60K-$79,999",
                          "$80K-$99,999",
                          "$100K-$149,999",
                          "$150K-$199,999",
                          "$200K-above" ))
```

Exploratory plot of Unvaccinated respondents by income

```
Unvaxed_by_inc2%>%
  ggplot(aes(x = inc1,
             y = percent_answers,
             fill = inc1)) +
  geom_col() +
  geom_text(aes(label = percent_answers_label),
            position = position_stack(vjust = 0),
            color = "black",
            fontface = "bold",
            size=4) +
  coord_flip() +
  scale_x_discrete() +
  scale_fill_manual(breaks =c( "$200K-above",
                               "$150K-$199,999",
                               "$100K-$149,999",
                               "$80K-$99,999",
                               "$60K-$79,999",
                               "$40K-$59,999",
                               "$30K-$39,999",
                               "$20K-$29,999",
                               "$10K-$19,999",
                               "$9,999-below"),#changes the order of the legend text listing
                    values = c("$200K-above"= "#003C30FF",
                               "$150K-$199,999"= "#01665EFF",
                               "$100K-$149,999"= "#35978FFF",
                               "$80K-$99,999"= "#80CDC1FF",
                               "$60K-$79,999"= "#C7EAE5FF",
                               "$40K-$59,999"= "#F6E8C3FF",
                               "$30K-$39,999"= "#DFC27DFF",
                               "$20K-$29,999"= "#BF812DFF",
                               "$10K-$19,999"= "#8C510AFF",
                               "$9,999-below"= "#543005FF"
                    )) +
  labs(title = "Unvaccinated Respondents Income Status",
       x = NULL,
       fill = NULL) +
  theme_minimal() +
```

```
  theme(axis.text.x = element_blank(),
        axis.title.x = element_blank(),
        panel.grid = element_blank(),
        legend.position = "right")
```




There are four additional plots that will be generated. Those plots are comparisons between unvaccinated and vaccinated respondents that reveal the perception of trust for:

1. Their doctor/healthcare provider for information of COVID-19
2. Local state officials for information of COVID-19
3. Friends and family for information of COVID-19
4. Religious leaders for information on COVID-19

The four (4) sources of trust selections listed above were selected out of # of___ trusted sources questions as a result of limited time and limited r programming skills. Moreover, the rationale used for seleccting the four trusted sources above were decided upon after analyzing the plot for unvaccinated respondents by health insurance status. Based on the diverging chart, there are a considerable proportion of respondents that are insured yet opted not to obtain the COVID-19 vaccine. The assumption being that those with health insurance would both trust their doctor and obtain the COVID-19 vaccine. Local state officials, friends and family and religious leaders were selected as a proxy for political, family and social constructs. A more in depth analysis in its entirety would provide a more robust and rigorous analysis.

The scale to rank trust is 1 to 4 with the following codes: 1=Great deal 2=Fair amount 3= Not much 4= Not at all

The following r code generated will include the summary dataset for each source of trust question and the additional data manipulation used to generate the final plot.

## Trusted Source-Doctor/Healthcare provider for information of COVID-19

Data manipulation Recoding levels and renaming the trust in doctor variable mydr Creating the data summary corresponding to vaccination status and responses to doctor/healthcare provider

```
pulse_surv=
  pulse_surv%>%
  mutate(mydr_r=recode(mydr,
                           "1"="Great deal",
                           "2"="Fair amount",
                           "3"="Not much",
                           "4"="Not at all",
                           "M"="NA"))




dr_by_vaxstat=
  pulse_surv%>%
  select(vax_stat1, mydr_r)%>%
  filter(mydr_r!= "NA")%>%
  filter(!is.na(vax_stat1))%>%
  group_by(vax_stat1, mydr_r)%>%
  count(name="n_answers")%>%
  group_by(vax_stat1)%>%
  mutate(percent_answers=n_answers/sum(n_answers))%>%
  ungroup()%>%
  mutate(percent_answers_label = percent(percent_answers, accuracy=1))




head(dr_by_vaxstat, 5)
```

```
## # A tibble: 5 x 5
##   vax_stat1 mydr_r      n_answers percent_answers percent_answers_label
##   <fct>     <chr>           <int>           <dbl> <chr>
## 1 no_vax    Fair amount        69           0.350 35%
## 2 no_vax    Great deal         47           0.239 24%
## 3 no_vax    Not at all         36           0.183 18%
## 4 no_vax    Not much           45           0.228 23%
## 5 vax       Fair amount       365           0.324 32%
```

Data manipulation Adjust the percentages with the negative feedback as negatives to position to one side Convert the negative label to absolute values-create a new dataset Reorder the bars Create new dataset for plot

```
dr_by_vaxstat_diverging=
dr_by_vaxstat%>%
```

```
  mutate(percent_answers=if_else(mydr_r %in% c("Great deal", "Fair amount"),
                                 percent_answers, -percent_answers))%>%
  mutate(percent_answers_label=percent(percent_answers, accuracy=1))
```

```
dr_by_vaxstat_labels=
  dr_by_vaxstat_diverging%>%
  mutate(percent_answers_label =abs(percent_answers))%>% #converts the negative values to absolute
  mutate(percent_answers_label= percent(percent_answers_label, accuracy =1))
```

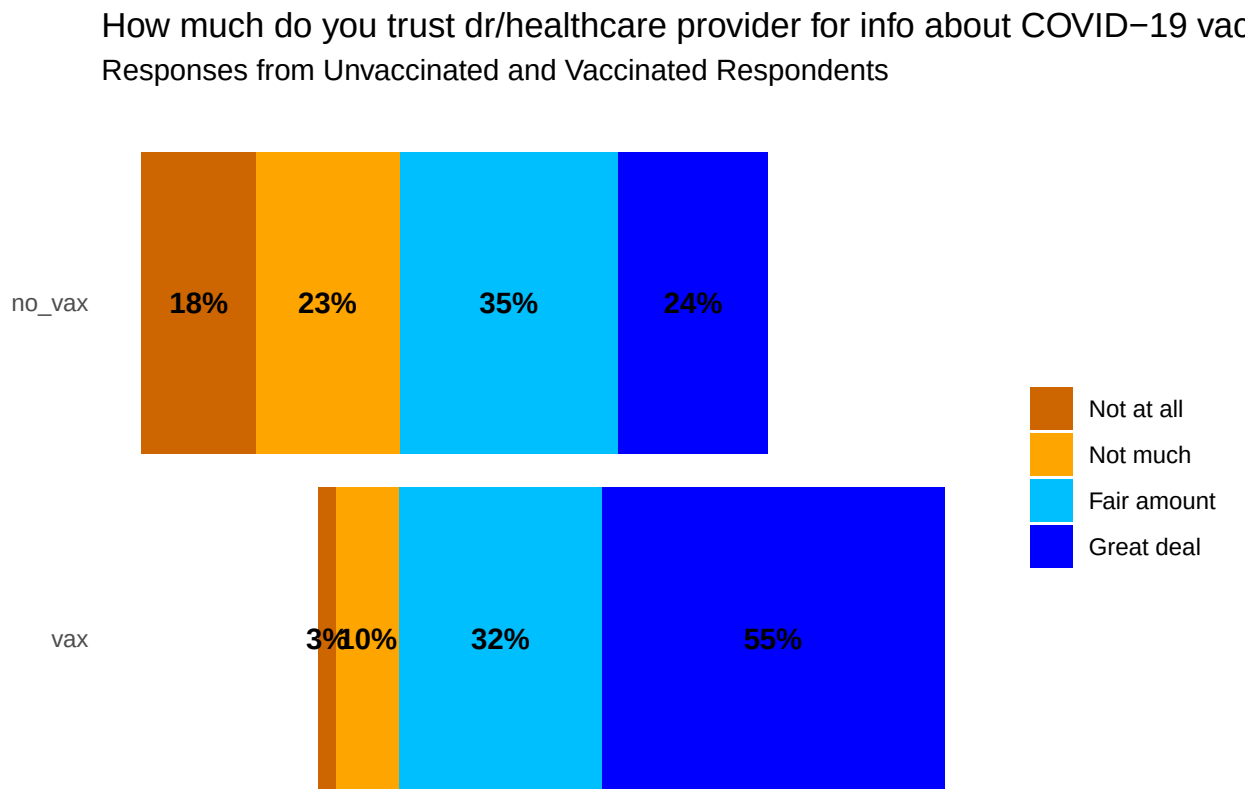
```
dr_by_vaxstat_labels_diverging_right_order=
dr_by_vaxstat_labels%>%
  mutate(mydr_r=fct_relevel(mydr_r,
                            "Not at all", "Not much",
                            "Great deal", "Fair amount"))%>%
  mutate(vax_stat1=fct_relevel(vax_stat1,
                            "vax",
                            "no_vax"))
```

## How much do you trust your Doctor/Healthcare provider for information of COVID-19

Plot Unvaccinated and Vaccinated responses

```
dr_by_vaxstat_labels_diverging_right_order%>%
ggplot(aes(x = vax_stat1,
           y = percent_answers,
           fill = mydr_r)) +
  geom_col() +
  geom_text(aes(label = percent_answers_label),
            position = position_stack(vjust = 0.5),
            color = "black",
            fontface = "bold") +
  coord_flip() +
  scale_x_discrete() +
  scale_fill_manual(breaks=c("Not at all", "Not much", "Fair amount", "Great deal"),
                    values=c(

                      "Not at all"= "darkorange3",
                      "Not much"="orange",
                      "Fair amount"= "deepskyblue",
                      "Great deal"= "blue"
                      )) +
  labs(title = "How much do you trust dr/healthcare provider for info about COVID-19 vaccine?",
```

```
      subtitle = "Responses from Unvaccinated and Vaccinated Respondents",
      x = NULL,
      fill = NULL) +
  theme_minimal() +
  theme(axis.text.x = element_blank(),
        axis.title.x = element_blank(),
        panel.grid = element_blank(),
        legend.position = "right")
```




## Trusted Source-state and local officials for information of COVID-19

Data manipulation Recoding levels and renaming the trust in state and local officials variable stat_loca_off Creating the data summary corresponding to vaccination status and responses to state and local officials

```
pulse_surv=
  pulse_surv%>%
  mutate(stat_loca_off_r=recode(stat_loca_off,
                        "1"="Great deal",
                        "2"="Fair amount",
                        "3"="Not much",
                        "4"="Not at all",
                        "M"="NA"))
```

```
stat_local_off_by_vax=
pulse_surv%>%
  select(vax_stat1, stat_loca_off_r)%>%
  filter(stat_loca_off_r!= "NA")%>%
  filter(!is.na(vax_stat1))%>%
  group_by(vax_stat1, stat_loca_off_r)%>%
  count(name="n_answers")%>%
  group_by(vax_stat1)%>%
  mutate(percent_answers=n_answers/sum(n_answers))%>%
  ungroup()%>%
  mutate(percent_answers_label = percent(percent_answers, accuracy=1))
```

```
head(stat_local_off_by_vax, 5)
```

```
## # A tibble: 5 x 5
##   vax_stat1 stat_loca_off_r n_answers percent_answers percent_answers_label
##   <fct>     <chr>               <int>           <dbl> <chr>
## 1 no_vax    Fair amount            41          0.209  21%
## 2 no_vax    Great deal             15          0.0765 8%
## 3 no_vax    Not at all             68          0.347  35%
## 4 no_vax    Not much               72          0.367  37%
## 5 vax       Fair amount           502          0.450  45%
```

Data manipulation: Reset responses so that negative responses are shown as negative values and therefore placed on the left side of the diverging barplot. Convert the negative values into absolute so that values on plot do not reveal a negative sign for correct placement Relevel the categories for state_loca_off and vax_stat1 variables

```
stat_local_off_by_vax_diverging=
stat_local_off_by_vax%>%
  mutate(percent_answers=if_else(stat_loca_off_r %in% c("Great deal", "Fair amount"),
                                 percent_answers, -percent_answers))%>%
  mutate(percent_answers_label=percent(percent_answers, accuracy=1))
```

```
stat_local_off_by_vax_diverging_labels=
stat_local_off_by_vax_diverging%>%
  mutate(percent_answers_label =abs(percent_answers))%>% #converts the negative values to absolute
  mutate(percent_answers_label= percent(percent_answers_label, accuracy =1))
```

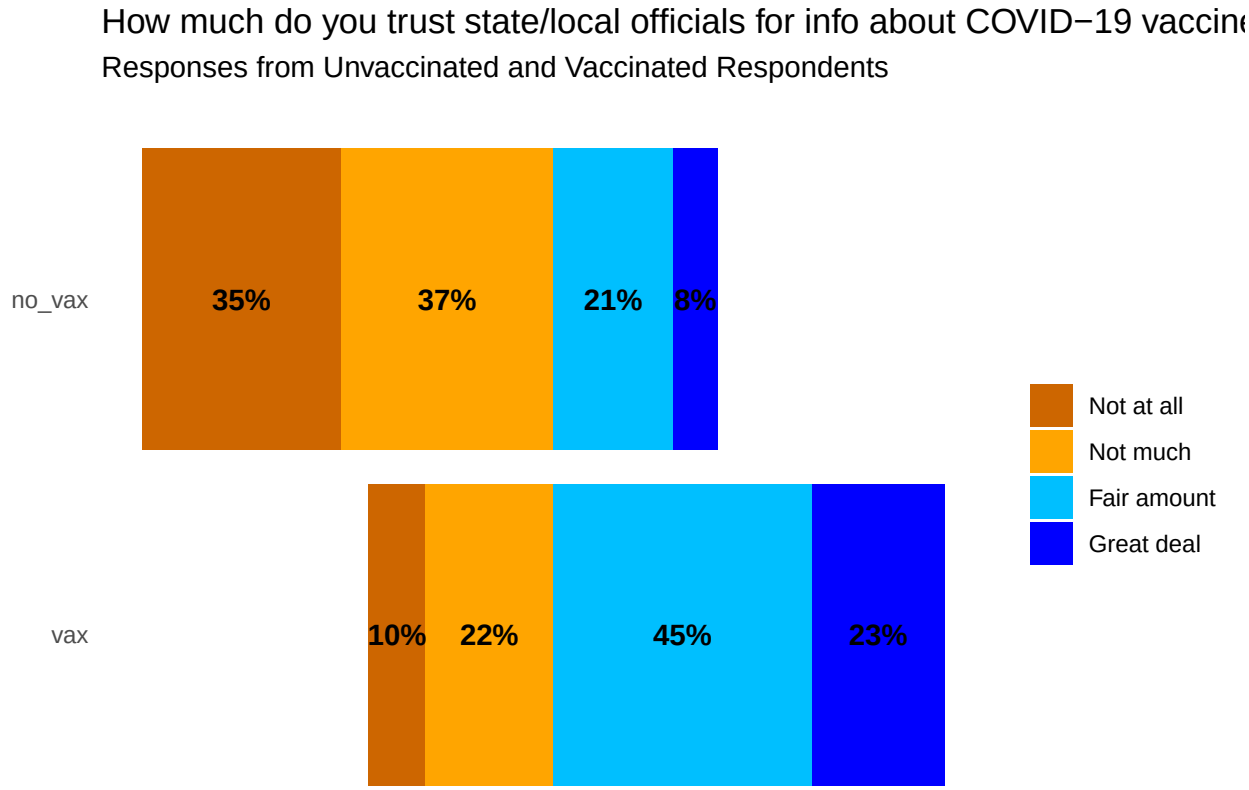
```
stat_local_off_by_vax_diverging_labels_right_order=
stat_local_off_by_vax_diverging_labels%>%
  mutate(stat_loca_off_r=fct_relevel(stat_loca_off_r,
```

```
                              "Not at all", "Not much",
                              "Great deal", "Fair amount"))%>%
  mutate(vax_stat1=fct_relevel(vax_stat1,
                               "vax",
                               "no_vax"))
```

## How much do you trust state and local officials for information about the COVID-19 vaccine?

Plot Unvaccinated and Vaccinated responses

```
stat_local_off_by_vax_diverging_labels_right_order%>%
  ggplot(aes(x = vax_stat1,
             y = percent_answers,
             fill = stat_loca_off_r)) +
  geom_col() +
  geom_text(aes(label = percent_answers_label),
            position = position_stack(vjust = 0.5),
            color = "black",
            fontface = "bold") +
  coord_flip() +
  scale_x_discrete() +
  scale_fill_manual(breaks=c("Not at all", "Not much", "Fair amount", "Great deal"),
                    values=c(

                      "Not at all"= "darkorange3",
                      "Not much"="orange",
                      "Fair amount"= "deepskyblue",
                      "Great deal"= "blue"
                    )) +
  labs(title = "How much do you trust state/local officials for info about COVID-19 vaccine?",
       subtitle = "Responses from Unvaccinated and Vaccinated Respondents",
       x = NULL,
       fill = NULL) +
  theme_minimal() +
  theme(axis.text.x = element_blank(),
        axis.title.x = element_blank(),
        panel.grid = element_blank(),
        legend.position = "right")
```




## Trusted Source-Friends and family for information of COVID-19

Data manipulation: Recoding levels and renaming the trust in friends and family variable famfrnd Creating the data summary corresponding to vaccination status and responses to friends and family variable

```
pulse_surv=
  pulse_surv%>%
  mutate(famfrnd_r=recode(famfrnd,
                                "1"="Great deal",
                                "2"="Fair amount",
                                "3"="Not much",
                                "4"="Not at all",
                                "M"="NA"))



famfrnd_by_vax=
  pulse_surv%>%
  select(vax_stat1, famfrnd_r)%>%
  filter(famfrnd_r!= "NA")%>%
  filter(!is.na(vax_stat1))%>%
  group_by(vax_stat1, famfrnd_r)%>%
  count(name="n_answers")%>%
```

```
  group_by(vax_stat1)%>%
  mutate(percent_answers=n_answers/sum(n_answers))%>%
  ungroup()%>%
  mutate(percent_answers_label = percent(percent_answers, accuracy=1))



head(famfrnd_by_vax, 5)
```

```
## # A tibble: 5 x 5
##   vax_stat1 famfrnd_r   n_answers percent_answers percent_answers_label
##   <fct>     <chr>           <int>           <dbl> <chr>
## 1 no_vax    Fair amount        56           0.28  28%
## 2 no_vax    Great deal         43           0.215 22%
## 3 no_vax    Not at all         39           0.195 20%
## 4 no_vax    Not much           62           0.31  31%
## 5 vax       Fair amount       469           0.419 42%
```

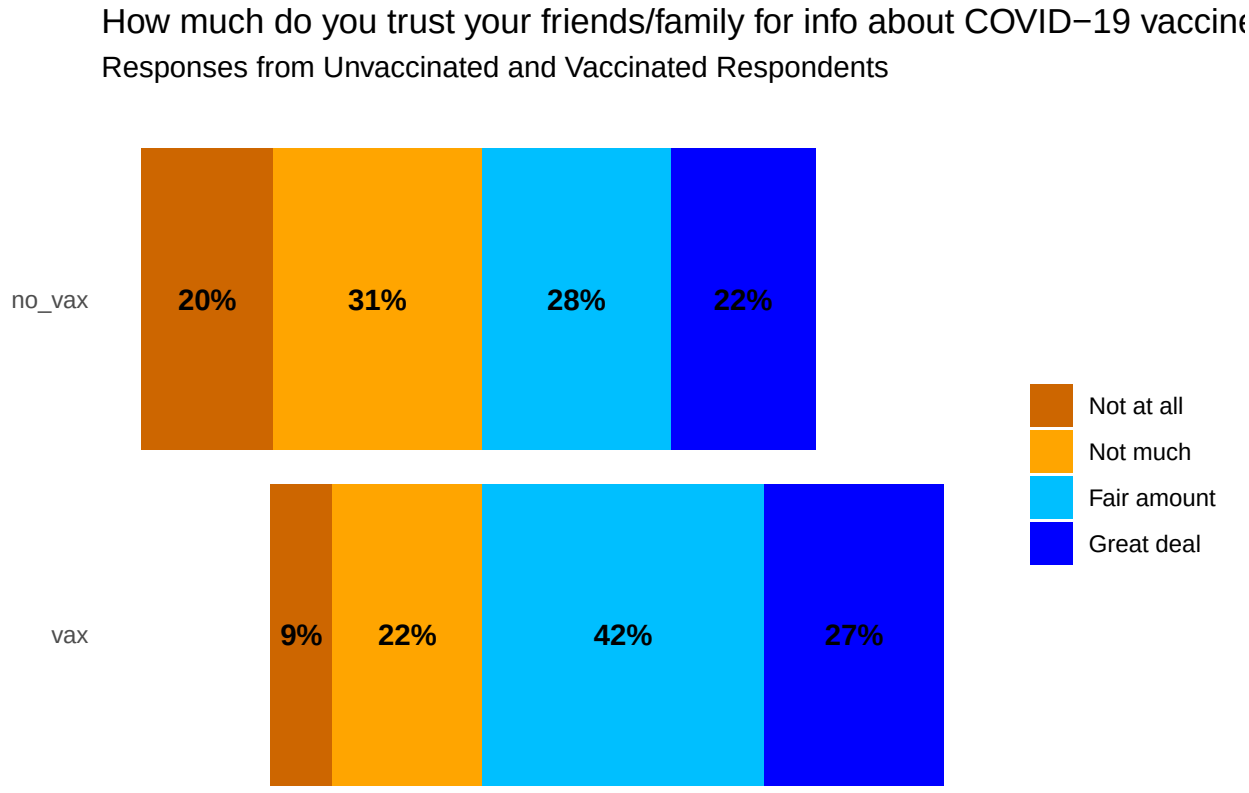
Data manipulation Redistribute negative and positive responses on corresponding sides Convert negative values to absolute values so that negative symbols do not appear on graph Relevel famfrnd_r variable and vax_stat1 variable to make sure order on graph is displayed accordingly Create new dataset

```
famfrnd_by_vax_diverging=
famfrnd_by_vax%>%
  mutate(percent_answers=if_else(famfrnd_r %in% c("Great deal", "Fair amount"),
                                 percent_answers, -percent_answers))%>%
  mutate(percent_answers_label=percent(percent_answers, accuracy=1))




famfrnd_by_vax_diverging_diverging_labels=
famfrnd_by_vax_diverging%>%
  mutate(percent_answers_label =abs(percent_answers))%>% #converts the negative values to absolute
  mutate(percent_answers_label= percent(percent_answers_label, accuracy =1))




famfrnd_by_vax_diverging_diverging_labels_right_order=
famfrnd_by_vax_diverging_diverging_labels%>%
  mutate(famfrnd_r=fct_relevel(famfrnd_r,
                                    "Not at all", "Not much",
                                    "Great deal", "Fair amount"))%>%
  mutate(vax_stat1=fct_relevel(vax_stat1,
                               "vax",
                               "no_vax"))
```

#How much do you trust Your friends and family for information about the COVID-19 vaccine?

Plot Unvaccinated and Vaccinated responses

```
famfrnd_by_vax_diverging_diverging_labels_right_order%>%
  ggplot(aes(x = vax_stat1,
             y = percent_answers,
             fill = famfrnd_r)) +
  geom_col() +
  geom_text(aes(label = percent_answers_label),
            position = position_stack(vjust = 0.5),
            color = "black",
            fontface = "bold") +
  coord_flip() +
  scale_x_discrete() +
  scale_fill_manual(breaks=c("Not at all", "Not much", "Fair amount", "Great deal"),
                    values=c(

                      "Not at all"= "darkorange3",
                      "Not much"="orange",
                      "Fair amount"= "deepskyblue",
                      "Great deal"= "blue"
                    )) +
  labs(title = "How much do you trust your friends/family for info about COVID-19 vaccine?",
       subtitle = "Responses from Unvaccinated and Vaccinated Respondents",
       x = NULL,
       fill = NULL) +
  theme_minimal() +
  theme(axis.text.x = element_blank(),
        axis.title.x = element_blank(),
        panel.grid = element_blank(),
        legend.position = "right")
```




## Trusted Source-religious leaders for information of COVID-19

Data manipulation Recoding levels and renaming the trust in religious leaders variable reli_ldrs Creating the data summary corresponding to vaccination status and responses to religious leaders

```
pulse_surv=
  pulse_surv%>%
  mutate(reli_ldrs_r=recode(reli_ldrs,
                            "1"="Great deal",
                            "2"="Fair amount",
                            "3"="Not much",
                            "4"="Not at all",
                            "M"="NA"))




reli_ldrs_by_vax=
pulse_surv%>%
  select(vax_stat1, reli_ldrs_r)%>%
  filter(reli_ldrs_r!= "NA")%>%
  filter(!is.na(vax_stat1))%>%
  group_by(vax_stat1, reli_ldrs_r)%>%
```

```
  count(name="n_answers")%>%
  group_by(vax_stat1)%>%
  mutate(percent_answers=n_answers/sum(n_answers))%>%
  ungroup()%>%
  mutate(percent_answers_label = percent(percent_answers, accuracy=1))



head(reli_ldrs_by_vax, 5)
```

```
## # A tibble: 5 x 5
##   vax_stat1 reli_ldrs_r n_answers percent_answers percent_answers_label
##   <fct>     <chr>           <int>           <dbl> <chr>
## 1 no_vax    Fair amount        45           0.228 23%
## 2 no_vax    Great deal         27           0.137 14%
## 3 no_vax    Not at all         61           0.310 31%
## 4 no_vax    Not much           64           0.325 32%
## 5 vax       Fair amount       317           0.284 28%
```

Data manipulation Redistribute negative and positive responses on corresponding sides Convert negative values to absolute values so that negative symbols do not appear on graph Relevel reli_ldrs_r variable and vax_stat1 variable to make sure order on graph is displayed accordingly Create new dataset

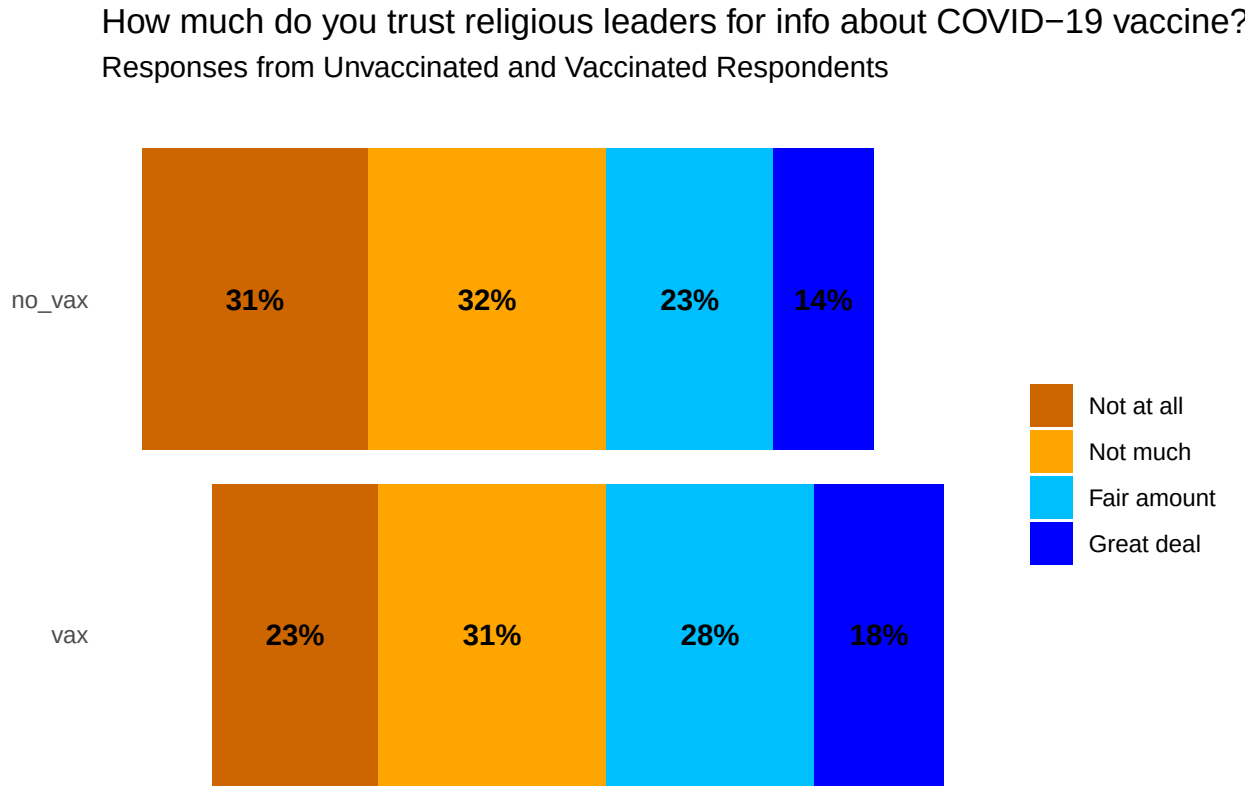
```
reli_ldrs_by_vax_diverging=
reli_ldrs_by_vax%>%
  mutate(percent_answers=if_else(reli_ldrs_r %in% c("Great deal", "Fair amount"),
                                 percent_answers, -percent_answers))%>%
  mutate(percent_answers_label=percent(percent_answers, accuracy=1))




reli_ldrs_by_vax_diverging_labels=
reli_ldrs_by_vax_diverging%>%
  mutate(percent_answers_label =abs(percent_answers))%>% #converts the negative values to absolute
  mutate(percent_answers_label= percent(percent_answers_label, accuracy =1))




reli_ldrs_by_vax_diverging_labels_right_order=
reli_ldrs_by_vax_diverging_labels%>%
  mutate(reli_ldrs_r=fct_relevel(reli_ldrs_r,
                                 "Not at all", "Not much",
                                 "Great deal", "Fair amount"))%>%
  mutate(vax_stat1=fct_relevel(vax_stat1,
                                 "vax",
                                 "no_vax"))
```

## How much do you trust religious leaders for information about the COVID-19 vaccine?

Plot Unvaccinated and Vaccinated responses

```
plot_reli=
reli_ldrs_by_vax_diverging_labels_right_order%>%
  ggplot(aes(x = vax_stat1,
             y = percent_answers,
             fill = reli_ldrs_r)) +
  geom_col() +
  geom_text(aes(label = percent_answers_label),
            position = position_stack(vjust = 0.5),
            color = "black",
            fontface = "bold") +
  coord_flip() +
  scale_x_discrete() +
  scale_fill_manual(breaks=c("Not at all", "Not much", "Fair amount", "Great deal"),
                    values=c(

                      "Not at all"= "darkorange3",
                      "Not much"="orange",
                      "Fair amount"= "deepskyblue",
                      "Great deal"= "blue"
                    )) +
  labs(title = "How much do you trust religious leaders for info about COVID-19 vaccine?",
       subtitle = "Responses from Unvaccinated and Vaccinated Respondents",
       x = NULL,
       fill = NULL) +
  theme_minimal() +
  theme(axis.text.x = element_blank(),
        axis.title.x = element_blank(),
        panel.grid = element_blank(),
        legend.position = "right")
```

```
plot_reli
```




## Limitations

Possible limitations to the data is insufficient sample size for unvaccinated respondents with race/ethnicity/age not being a representative sample. Focus of sample population is area residents living in community areas of the city that are underserved. Some survey respondents resided in areas of the city outside the focus areas. There was missing data and refusal to report for income (Missing 20 and 248 Refusals). This data will be omitted from this analysis. Doing so will present bias. Gender included 68 nonbinary, transgender, queer, two-spirit combination. This data was omitted in order to simplify the analysis between binary genders. The process of data manipulation in r programming presents challenges largely due to novice level understanding and novice level experience. Moreover, in order to visualize data, in particular, likert scale data, datasets must be structured in ways that were unfamiliar to me. The skill level required combined with the dataset structure and uncertainty of the manner the analysis should be structured presented personal challenges that prevented me from rigorously exploring, manipulating and visualizing the data. More time and additional/frequent interaction with r programming would present the opportunity to do a more comprehensive analysis. The omission of non binary gender data was done as a simplification of the data analysis process and in no way represents the methodology that would be taken under different circumstances. This analysis was an exercise and practice in performing r programming functions. Full recognition is acknowledged that data omitted presents biases and requires a more robust methodology for analysis. Lastly, self reporting surveys, although insightful, are subject to social desireability bias, non-response bias and other types of survey bias.

## Conclusions

When examining responses between vaccinated and unvaccinated respondents about perceptions of safety of the COVID-19 vaccine, perceptions for the unvaccinated respondents revealed a larger proportion of uncertainty over the safety of the vaccine as compared to vaccinated respondents. Moreover, when analyzing the health insurance rates for unvaccinated respondents, 50% or more of the race/ethnicity groups reported having health insurance with White (93%) and Black(78%) unvaccinated respondents indicating the highest percentage of health insurance status followed by Hispanic(50%) and Asian(50%). When looking at the gender break down for unvaccinated respondents by health innsurance status, 78% of unvaccinated female respondents reported having having health insurance and 63% of unvaccinated male respondents reported having health insurance. Over 60% of unvaccinated respondents reported having at least 2 years of college or more, and 36% of unvaccinated respondents reported earning $40K or more. When looking at trust perceptions for respondents own doctor/health care provider perceptions between unvaccinated and vaccinated respondents, 41% of unvaccinated respondents reported not having trust for their own doctor/health care provider while 13% of vaccinated respondents reported not having trust for their own doctor/health care provider. This observation is of particular interest and merits further examination into health care experiences of underserved populations in Chicago.

For trust perceptions among the unvaccinated and vaccinated respondents with state/local officials, 72% of unvaccinated respondents did not trust state/local officials for information about COVID-19 compared to 32% of vaccinated respondents. For trust perceptions between unvaccinated and vaccinated with friends and family, vaccinated respondents trust friends anf family 19% points higher than unvaccinated respondents. When unvaccinated and vaccinated respondents reported trust perceptions for religious leaders, vaccinated respondents trusted religious leaders more than unvaccinated respondents by a 9% point difference.

Much investment has been made in the dissemination of vaccination information via different media channels using doctors, politicians and religious leaders. The efforts made encouraging COVID-19 vaccination through billboards, television commericals, etc. are methods that keep the topic top of mind, however, trust perceptions among the unvaccinated is a population segment that deserves further research. Some scholars maintain that public health strategies that embark on media campaigns, compulsory vaccination requirements and stigmatization through the negative categorizing of vaccine hesitancy are strategically problematic (Kitta 2015). Information alone cannot solve this problem, instead, research that understands health choice as a concept reliant on cultural concerns and influences that make up risk for communities and individuals should be explored. Lastly, what is deemed scientific fact and the way it is interpreted and spread can depend on contemporary rumors and gossip, approaches that investigate vaccination fear and risk should be considered as part vaccine policy to better understand the factors that affect vaccine trust in general.

## External References:

Kitta, A. Goldberg, D.S.(2017). The significance of Folklore for Vaccine Policy: Discarding the Deficit Model. Critical Public Health, Volume 27, No. 4, 506-514.

Dube, E. Vivion, M. MacDonald N.E.(2015) Vaccine Hesitancy, Vaccine Refusal and the Anti-Vaccine Movement: Influence, Impact and Implications. Expert Rev.Vaccines, Volume 14(1), 99-117.