# Disparity of Vaccination Rate in Chicago by Zipcode

# Key Questions

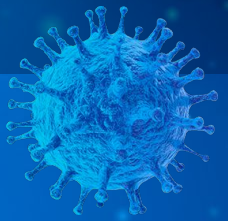
1. Which zip codes in Chicago have the highest income per capita?
2. Which zip codes in Chicago have the highest percentage of people who have received at least one COVID vaccine?
3. Is there a relationship between areas with a high income per capita and the percentage of people who have received at least one COVID vaccine dose?

# Hypothesis

**Null Hypothesis:** There is not a correlation between the income per capita by zip code and the percentage of people who have received one COVID vaccine by zip code

**Alternate Hypothesis:** There is a correlation between the income per capita by zip code and the percentage of people who have received one COVID vaccine by zip code

# Data



United States® Census Bureau

CHICAGO DATA PORTAL

# Jupyter Notebook Demonstration

```
from matplotlib.cm import inferno
from matplotlib.colors import to_hex
import json
import gmaps

#import api key
from config import gkey

#import data
# Store filepath in a variable
census_data = "chicago_census_data.csv"
vaccine_data = "COVID-19_Vaccinations_by_ZIP_Code.csv"
area_data = "chicago_areas.csv"
census_df = pd.read_csv(census_data, encoding="ISO-8859-1")
vaccine_df = pd.read_csv(vaccine_data,encoding="ISO-8859-1")
area_df = pd.read_csv(area_data,encoding="ISO-8859-1")
```

In [28]:
```
#cleaning census so Zip Code column matches Vaccine CSV Zip Code
census_df = census_df.rename(columns={"Zipcode":"Zip Code"})
census_df.dtypes
```

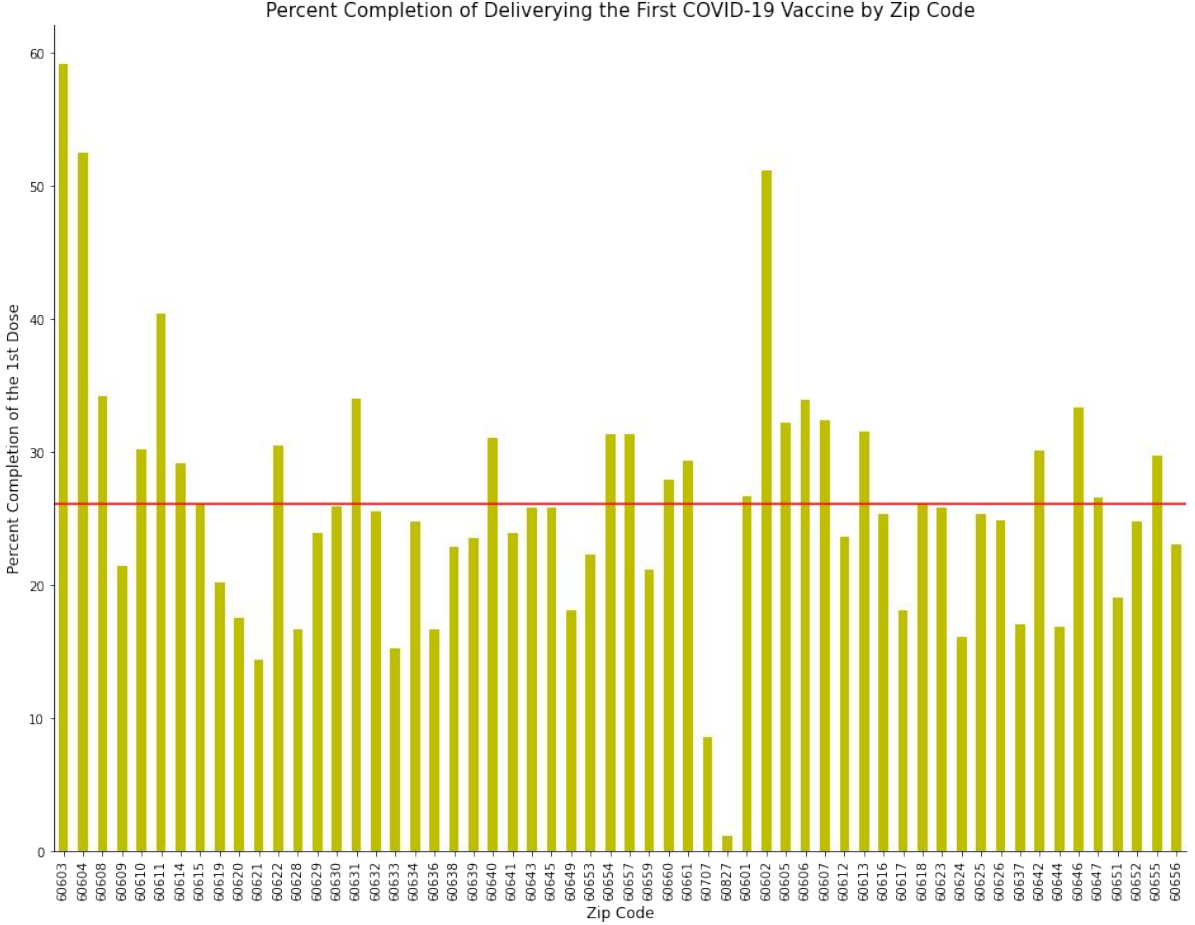
Out[28]:
```
Zip Code               int64
Population             int64
Median Age           float64
Household Income       int64
Per Capita Income      int64
Poverty Count          int64
Poverty Rate         float64
dtype: object
```

In [29]:
```
#getting rid of 'Unknown' values in vaccine df
vaccine_df = vaccine_df[vaccine_df["Zip Code"] != "Unknown"]
```

In [30]:
```
#changing Zip Code from Object to int
vaccine_df["Zip Code"]=vaccine_df["Zip Code"].astype(int)
```

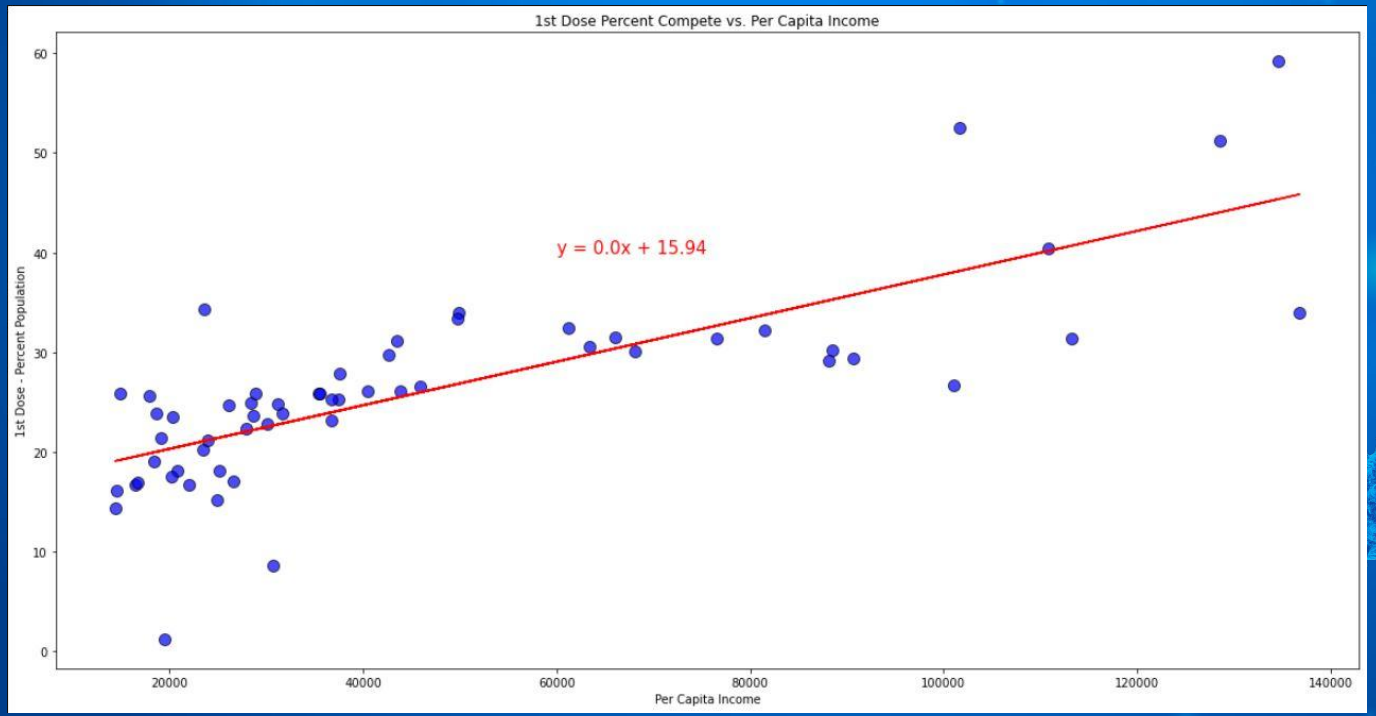
In [31]:
```
#merging vaccine data and census data
merge_df = pd.merge(vaccine_df, census_df, on="Zip Code")

#display df
merge_df.head()
```


Percent Completion of Deliverying the First COVID-19 Vaccine by Zip Code
Percent Completion of the 1st Dose
0
10
20
30
40
50
60
Zip Code
60603 60604 60608 60609 60610 60611 60614 60615 60619 60620 60621 60622 60628 60629 60630 60631 60632 60633 60634 60636 60638 60639 60640 60641 60643 60645 60649 60653 60654 60657 60659 60660 60661 60707 60827 60601 60602 60605 60606 60607 60612 60613 60616 60617 60618 60623 60624 60625 60626 60637 60642 60644 60646 60647 60651 60652 60655 60656

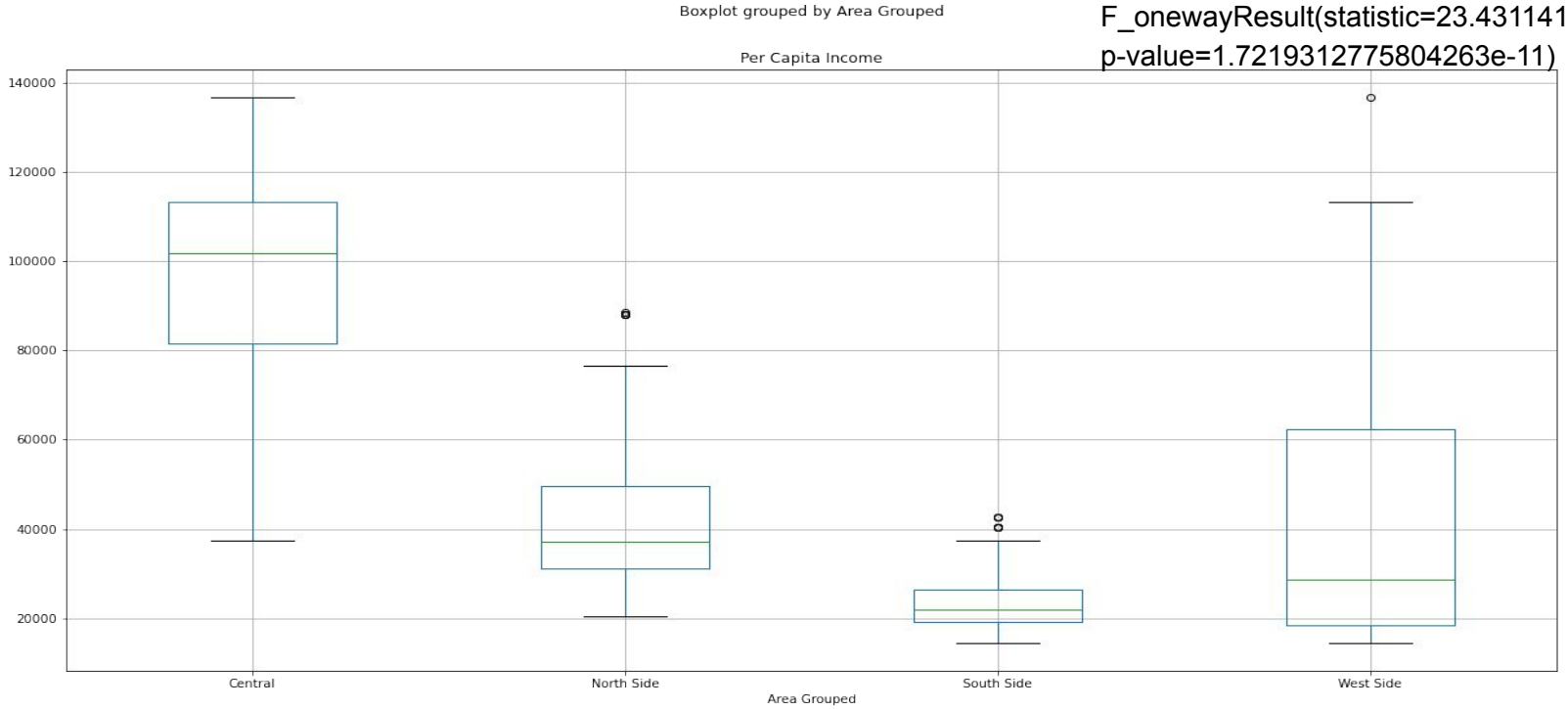
| | 60827 (Riverdale) | 60603 (Loop) |
|---|---|---|
| Population | 27,946 | 493 |
| Population density | 3,985 people per sq mi | 3402 |
| Housing Units | 11,457 | 749 |
| **Median Home Value** | $100,300 | $643,800 |
| **Land Area** | 7.01 sq mi | 0.14 sq mi |
| **Water Area** | 0.47 sq mi | 0 sq mi |
| **Occupied Housing Units** | 9637 | 187 |
| **Median Household Income** | $33,108 | $107,419 |

# Income per Capita Vs. 1st Dose Percentage



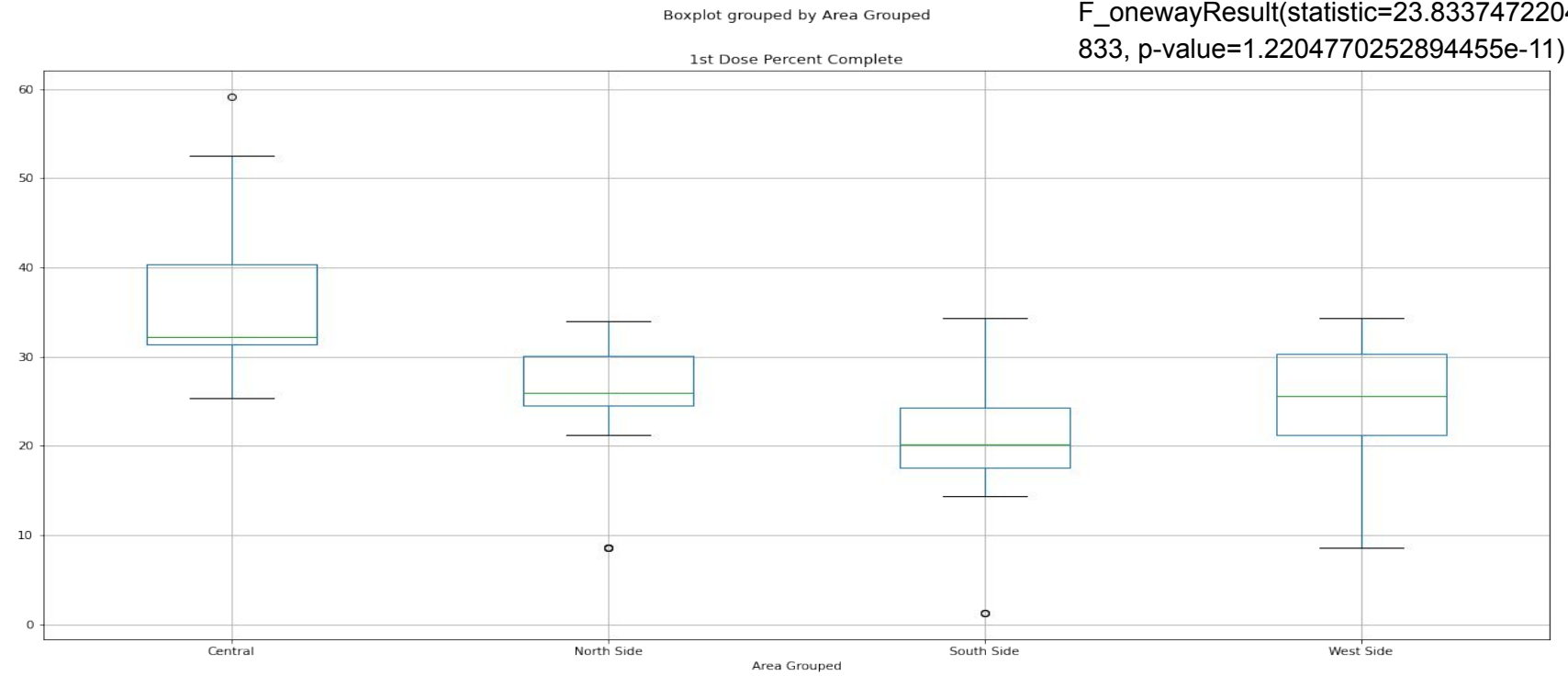

The r-value is:
0.7675260542636421

# Per Capita Income by Chicago Area



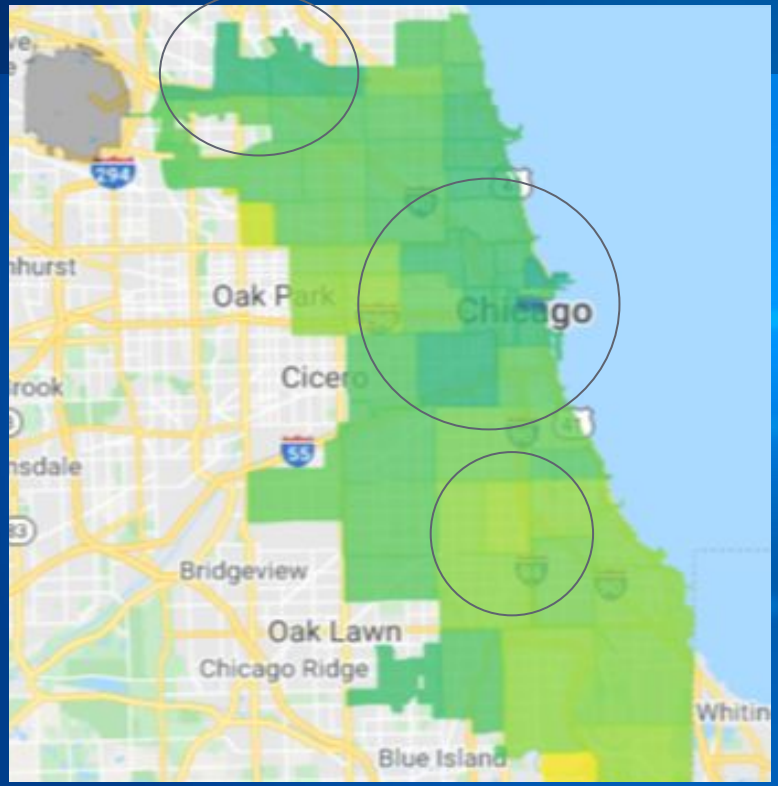

F_onewayResult(statistic=23.4311418406542,
p-value=1.7219312775804263e-11)

# 1st Dose Percent by Chicago Areas

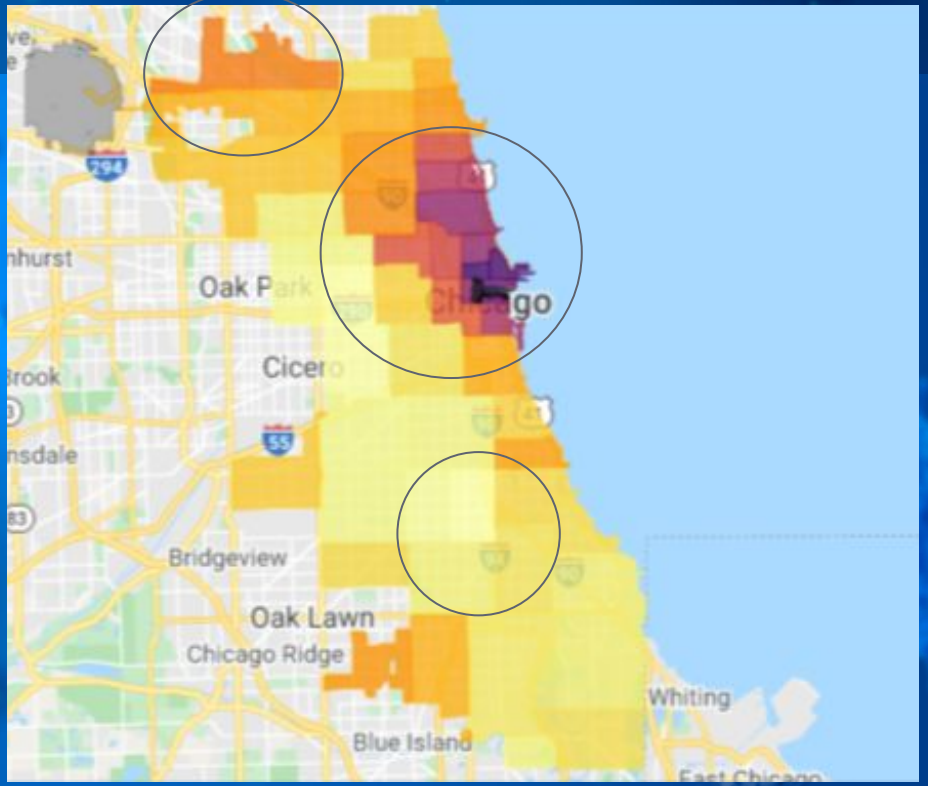



F_onewayResult(statistic=23.833747220401833, p-value=1.2204770252894455e-11)

% population 1 dose received




Income per capita



# Questions for Further Study

1. Access to vaccination facilities? Internet Access? Time availability?
2. Mentality towards the vaccine?