PSC 400
SYRACUSE UNIVERSITY

# DATA ANALYTICS FOR POLITICAL SCIENCE

## GETTING STARTED WITH R

# OBSERVATION

- **Integers (single data point)**

```
integer1 <- 8
integer2 <- 6 - 8
```

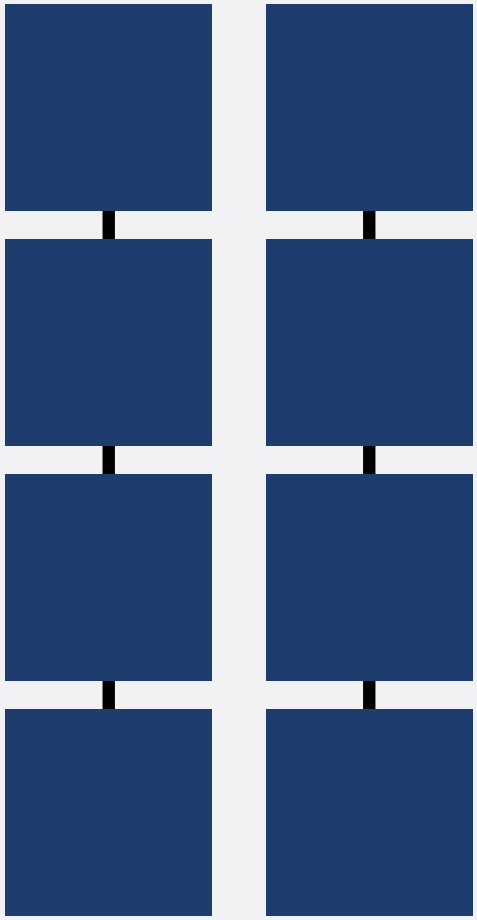
# VECTOR

- Vector (several data points)

```
vector1 <- c(3, -5, 8, 12, 0)
vector2 <- c(integer1, integer2)
```

# DATA FRAME

- Data Frame (several vectors)



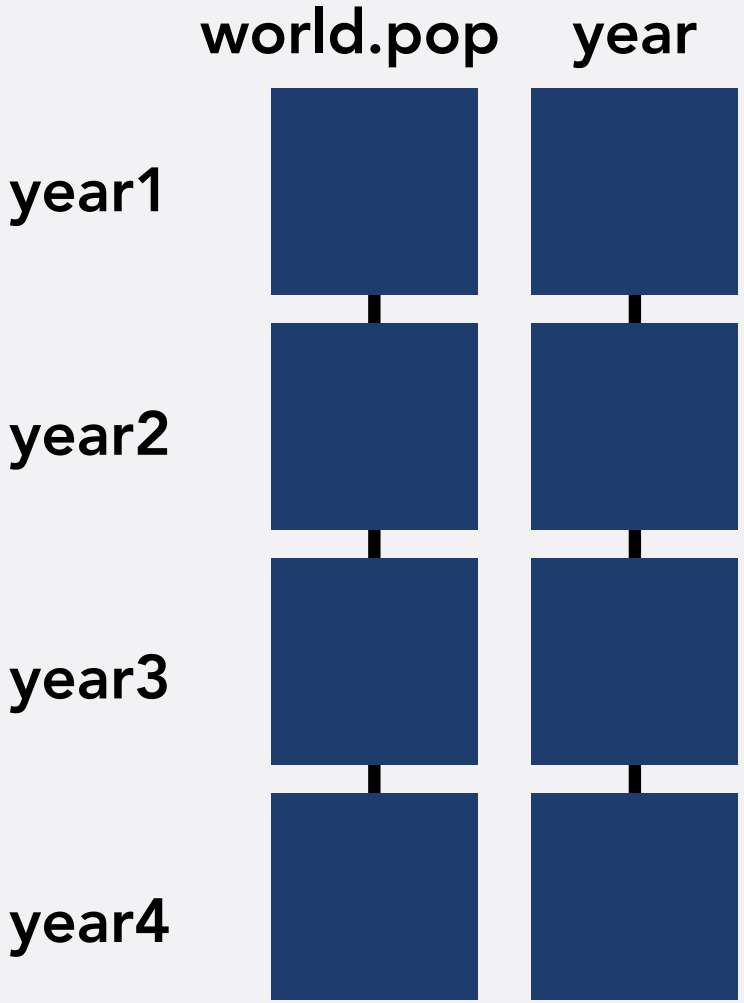
```
world.pop <- c(2525779,
3026003, 3691173, 4449049,
5320817, 6127700, 6916183)

years <- seq(from=1950,
to=2010, by=10)

data <-
data.frame(world.pop, years)
```

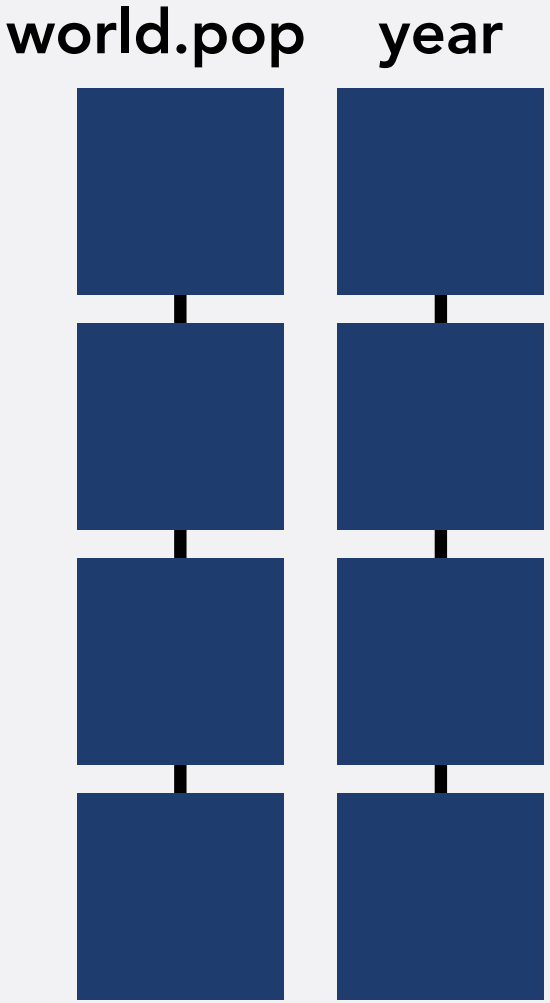
# DATA FRAME

- Rows: Different observations/cases

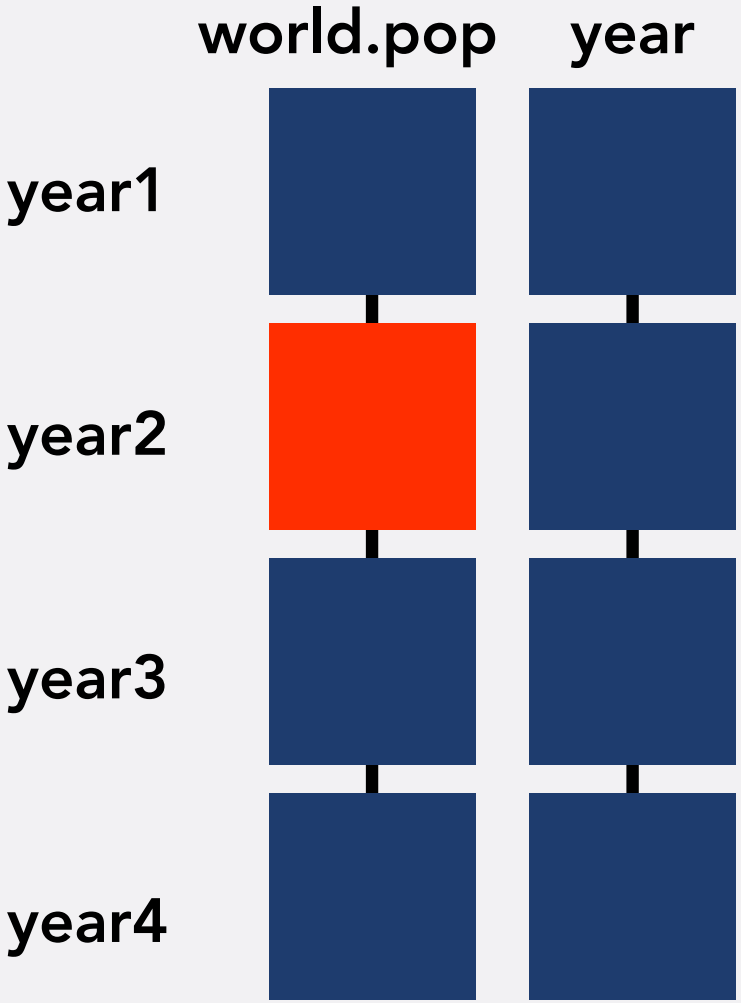


# DATA FRAME

- Columns: different variables



# DATA FRAME

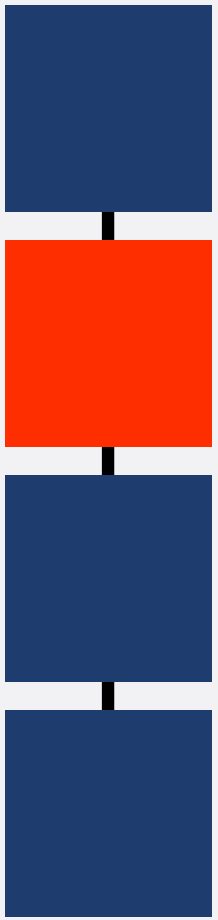



```
data[2, 1]

data[2, "world.pop"]
```

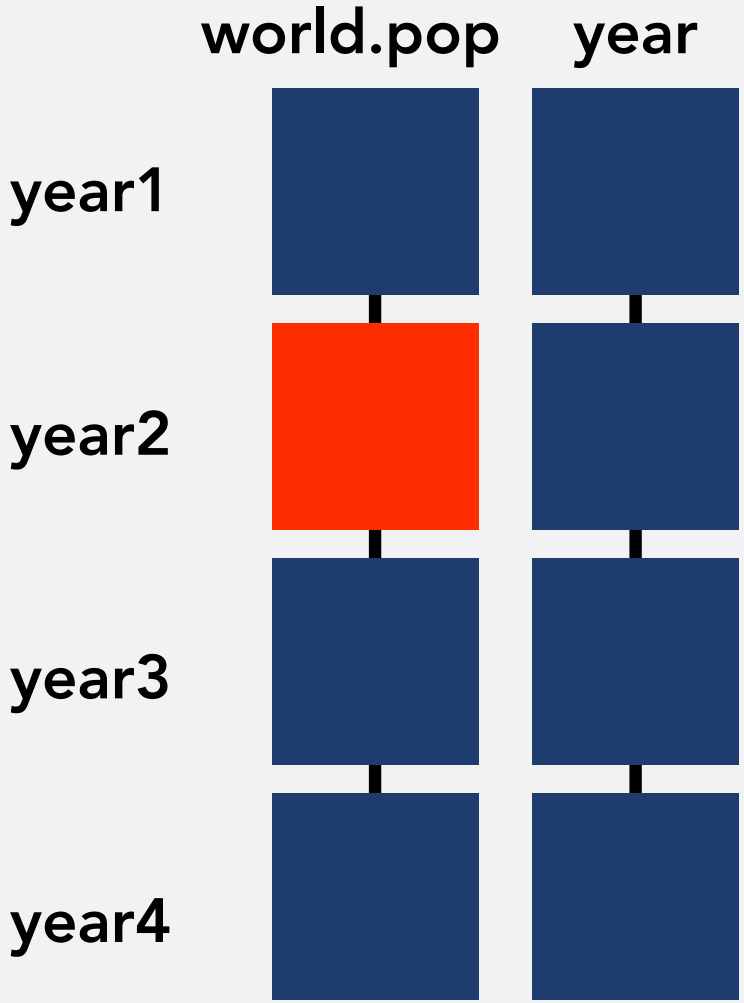
# VECTOR

world.pop



```
world.pop[2]
```

# DATA FRAME




```
data[2, 1]

data[2, "world.pop"]

data$world.pop[2]
```

# IN-CLASS EXERCISE

- Bias in Self-Reported Turnout
- Download and open file “turnout.csv”

| *Variable* | *Description* |
|---|---|
| year | election year |
| ANES | ANES estimated turnout rate |
| VEP | voting eligible population (in thousands) |
| VAP | voting age population (in thousands) |
| total | total ballots cast for highest office (in thousands) |
| felons | total ineligible felons (in thousands) |
| noncitizens | total noncitizens (in thousands) |
| overseas | total eligible overseas voters (in thousands) |
| osvoters | total ballots counted by overseas voters (in thousands) |

# IN-CLASS EXERCISE

| *Variable* | *Description* |
|---|---|
| year | election year |
| ANES | ANES estimated turnout rate |
| VEP | voting eligible population (in thousands) |
| VAP | voting age population (in thousands) |
| total | total ballots cast for highest office (in thousands) |
| felons | total ineligible felons (in thousands) |
| noncitizens | total noncitizens (in thousands) |
| overseas | total eligible overseas voters (in thousands) |
| osvoters | total ballots counted by overseas voters (in thousands) |

- What are the data dimensions?
- How many observations are there?
- Summarize the variables
- What is the range of years covered?

# IN-CLASS EXERCISE

| *Variable* | *Description* |
|---|---|
| year | election year |
| ANES | ANES estimated turnout rate |
| VEP | voting eligible population (in thousands) |
| VAP | voting age population (in thousands) |
| total | total ballots cast for highest office (in thousands) |
| felons | total ineligible felons (in thousands) |
| noncitizens | total noncitizens (in thousands) |
| overseas | total eligible overseas voters (in thousands) |
| osvoters | total ballots counted by overseas voters (in thousands) |

- Calculate turnout rate based on voting eligible population (VEP)
  - What difference do you observe?

# IN-CLASS EXERCISE

| *Variable* | *Description* |
|---|---|
| year | election year |
| ANES | ANES estimated turnout rate |
| VEP | voting eligible population (in thousands) |
| VAP | voting age population (in thousands) |
| total | total ballots cast for highest office (in thousands) |
| felons | total ineligible felons (in thousands) |
| noncitizens | total noncitizens (in thousands) |
| overseas | total eligible overseas voters (in thousands) |
| osvoters | total ballots counted by overseas voters (in thousands) |

- **Compute the difference between the VEP and ANES estimates of turnout**
  - **What is the range of the differences?**

# IN-CLASS EXERCISE

| *Variable* | *Description* |
|---|---|
| year | election year |
| ANES | ANES estimated turnout rate |
| VEP | voting eligible population (in thousands) |
| VAP | voting age population (in thousands) |
| total | total ballots cast for highest office (in thousands) |
| felons | total ineligible felons (in thousands) |
| noncitizens | total noncitizens (in thousands) |
| overseas | total eligible overseas voters (in thousands) |
| osvoters | total ballots counted by overseas voters (in thousands) |

- **Plot the difference between the VEP and ANES estimates of turnout over time**