**Nominal data**

-north, south, east, west

-suburb, city, town, rural

-no order to it, not clear what is “best” or “worst”; “first” or “last”

-non-incremental differentiation

**Ordinal data**

-incremental, shows an order

-position of bars in a bar graph do tell you something—there is an order to it

-position gives information

-symmetry/bias—use a box plot; good for ordinal data

-box plot show you bias in distribution of the data

* Categorical data are the output of the classification process.
* A nominal classification related to education can be type of school funding: public or private; while an ordinal one can be: elementary, middle, high, college and graduate school level.

**Other notes:**

\*\*use his notes from Week 2 to go through the code

* Wants very good, simple plots about something interesting
* Exclude=nothing
  + Missing values: whenever you present a plot, people will add totals; missing values means that there is not sufficient data—should maybe name as unknown or re-name to something better so the totals are accurate
* Run a frequency table first
  + Prop table will show relative frequencies (multiply by 100 to get it in percent’s)
* Pie chart
  + Criticism
    - Need to use color to make a difference, color should be used with care
    - Difficult to see difference in chunks of similar size
    - Harder with more and more variables
    - Try to avoid
* Bar chart
  + Categorical, nominal data---use bar charts
  + Data frame with three columns—location, count, percentage
* Frequency table has to be a data frame
  + Frequency table is ggplot (start with that, then can make it more complex)
  + Base—telling ggplot what data frame you want to use and what values you want to plot
* Turning frequency table into data frame
* Every time he prepares material—will get one report for every class
  + Clone report and then upload it in R
  + Code—open with Github desktop
  + See where it is saved in folder (show in folder)
  + Then in R—open the 3 dots on the bottom right panel and fit in, click open
    - Click on the rmd extension file and you will have the code
  + **Write questions with a hashtag (#) in the code in R**
    - **Professor will come by and we can ask the questions**

**PROJECT**

Prepare a plot of one categorical variable of your choosing

Ex. How many people are getting free lunch in school

-turn numerical data into a plot—cute it to create labels (ranges)

\*\*Prepare a plot using the table above\*\*

-reach out to him with any issues—make the basic bar chart, add things as you need

-need to have it by next week

This line of code was critical:

```{r}

newtable=table(eduwa$Free.LunchGroup)

```

```{r}

newtabledf=as.data.frame(newtable)

```

```{r}

base=ggplot(data=newtabledf,aes(x=Var1,y=Freq))

```

```{r}

plot1 = base + geom\_bar(fill ="gray",

stat = 'identity')

plot1

```

**Basic Steps to Prepare a Visual in ggplot2**

* Turn the frequency table into a data frame
  + You can rename data frame columns, add percents,

