Lesson Plan 2/1
Thursday, February 1, 2018  9:23 AM

- Admin
  - Class notes linked to in Calendar
  - Any issues, comments?
- Hector’s Data Science Corner
  - Talk at the HCIL: https://hcil.umd.edu/events/event/bbl-bahador-saket-georgia-tech/
  - Data Skeptic: https://dataskeptic.com/
- Something I missed from last time
  - Categorical attributes and ‘factors’
- More operations
  - Sorting: arrange
  - Mutate
  - Summarize
  - Group by
- A couple other R programming issues
  - Vectors
  - Attributes are vectors
  - Defining functions
- Plotting
  - The ggplot triumvirate:
    - Data
    - Mapping
    - Representation
  - Mappings
  - Representations
  - Frequently Used Plots
    - Scatter plot
    - Bar graph
    - Histogram
    - Boxplot
- Rmarkdown
  - Markdown
  - Code chunks
- Categorical attributes and `factors`

- More operations
  - Sorting: arrange
  - Mutate
  - Summarize
  - Group by

- A couple other R programming issues
  - Vectors
    - Attributes are vectors
  - Defining functions

- Plotting
  - The ggplot triumvirate:
    - Data
    - Mapping
    - Representation
  - Frequently Used Plots
    - Scatter plot
    - Bar graph
    - Histogram
    - Boxplot

- Rmarkdown
  - Markdown
  - Code chunks
  - Notebook
  - Metadata
  - Rendering
"factor" in R: Categorical Data

Turns a string vector → vector of categorical values
possible values of a factor, are called "levels"

Operations:
Attributes: select
Entities: slice, filter, sample.x
Pipe: `%>%`

Ordering: Arrange qualities by attribute value

Summary:

Grouping:

Create new: `mutate`

Attributes

---

ggplot: grammar of graphics

```
CENTRAL
EASTERN
```
1) Data: Entities & attributes
2) Mapping: Data values $\rightarrow$ graphical characteristics
3) Representation: Geometrically representing characteristics

What is `aes`, aesthetic characteristics?