### Presenting Data

- Suppose we collapse ages into 5 mutually exclusive and exhaustive categories:

<table>
<thead>
<tr>
<th>Age Class</th>
<th>Number of Individuals (Freq.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>60-64</td>
<td>5</td>
</tr>
<tr>
<td>65-69</td>
<td>17</td>
</tr>
<tr>
<td>70-74</td>
<td>12</td>
</tr>
<tr>
<td>75-79</td>
<td>12</td>
</tr>
<tr>
<td>80-84</td>
<td>2</td>
</tr>
<tr>
<td>85-89</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cumulative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age Class</td>
</tr>
<tr>
<td>------------</td>
</tr>
<tr>
<td>60-64</td>
</tr>
<tr>
<td>65-69</td>
</tr>
<tr>
<td>70-74</td>
</tr>
<tr>
<td>75-79</td>
</tr>
<tr>
<td>80-84</td>
</tr>
<tr>
<td>85-89</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

### Frequency Histogram

- [Frequency Histogram](#)
Relative Frequency Histogram

Cumulative Relative Frequency Histogram

Numerical and Graphical Summaries

- Dichotomous and categorical
  - Tables: Frequencies and relative frequencies
  - Graphs: Bar charts (freq. or relative freq.)

- Ordinal
  - Tables: Frequencies, relative frequencies, cumulative frequencies and cumulative relative frequencies
  - Graphs: Histograms or bar charts
Nominal and Ordinal Variables

• Communication: Bar Charts

Numerical and Graphical Summaries

• Continuous
  – Graphs: Histogram, box plot
  – Tables: Frequencies, relative frequencies, cumulative frequencies and cumulative relative frequencies
  – Statistics: Mean, standard deviation, minimum, maximum, range, median, quartiles, interquartile range

Continuous Variables

• Communication: Histograms
Continuous Variables

• Measures of spread: outliers

• Outliers
  \[ x > \text{upper quartile} + 1.5 \times IQR \]
  \[ x < \text{lower quartile} + 1.5 \times IQR \]

Continuous Variables: Central Tendency

• Measures of central tendency: mode

Q10: The last time you partied/socialized how many drinks of alcohol did you have?

- 0
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10

4. What is the mode?

Most commonly students drank 3 drinks the last time they partied/socialized?

Continuous Variables

• Measures of spread: range

• Range:
  - easy to compute (largest observation – smallest observation)
  - Sensitive to extreme values
  - Larger n, the larger the range tends to be
  - Difficult to compare ranges from different size data sets
Continuous Variables

• Relationship between mean and median helps describe the symmetry of the data

![Relationship between mean and median](image)

![Frequency Histogram for Systolic BP](image)

![Frequency Polygon for Systolic BP](image)