

3 Rules of Data Analysis:

- Make a Picture
- Make a Picture
- Make a Picture

**Chapter 3:**

**Displaying & Describing Categorical Data**

In order to make a picture, we must put things in piles. We do this so we can see how the cases distribute across different categories. To do this, we just count the number of cases corresponding to each category and pile them up. Be cautious when making your picture, and follow the **area principle**. The area principle says that the area occupied by a part of the graph should correspond to the magnitude of the value it represents. Also, make sure the variables are **independent**; meaning the conditional distribution of one variable is the same for each category of the other. **Simpson's paradox** is another caution that should be taken as this is when averages taken across different groups appear to contradict the overall averages.

**Frequency Tables:** records the totals and the category names

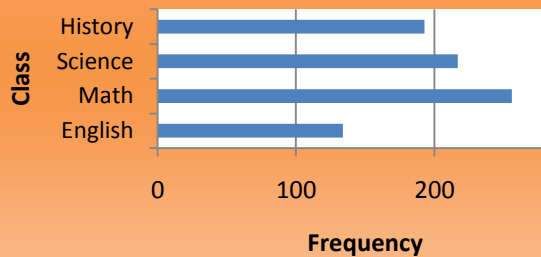
Class	Count
English	134
Math	256
Science	217
History	193

**Relative Frequency Table:** displays the **percentages** of the values in each category. We do this because we want to know the **proportion** of the data in each category, so we divide the counts by the total number of cases.

Class	Percentage
English	16.75
Math	32
Science	27.12
History	24.12

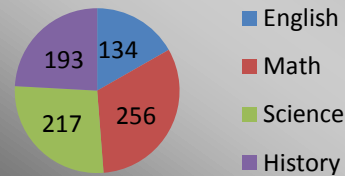
**Bar Chart:** displays the distribution of a categorical variable, showing the counts for each category next to each other for easy comparison.

**Student Distribution**



**Pie Chart:** Shows the whole group of cases as a circle. We slice the circle into pieces whose size is proportional to the fraction of the whole in each category.

**Count**



**Contingency Table:** shows how individuals are distributed along each variable, contingent on the value of the other variable.

	English	Math	Science	History	Total
Male	125	97	174	39	435
Female	9	159	43	154	365
Total	134	256	217	193	800

**Marginal distribution:** The counts/percentages are the totals found in the margins (last row/column) of the table.