

3.3 – Data Displays: Categorical versus Quantitative

Data Display # 1 – Categorical

– categorical variable (in a display) → data that can be displayed using names or labels

Ex: 1.) blonde, brunette, red, black, etc. → category = hair color

2.) red, yellow, green, purple, etc. → category = favorite color

3.) collie, shepherd, terrier, labrador, etc. → category = dog breeds

▪ **pie chart** – a type of categorical data display which uses a circle divided into sectors where each “slice” represents a portion of the whole (percents %)

▪ **bar graph** – a type of categorical data display which uses numerical labels / titles that’s represented by rectangles of equal width (note: between each “bar” – there are gaps)

Example 1: Express the following data as a pie chart and a bar chart.

NOT A HISTOGRAM!

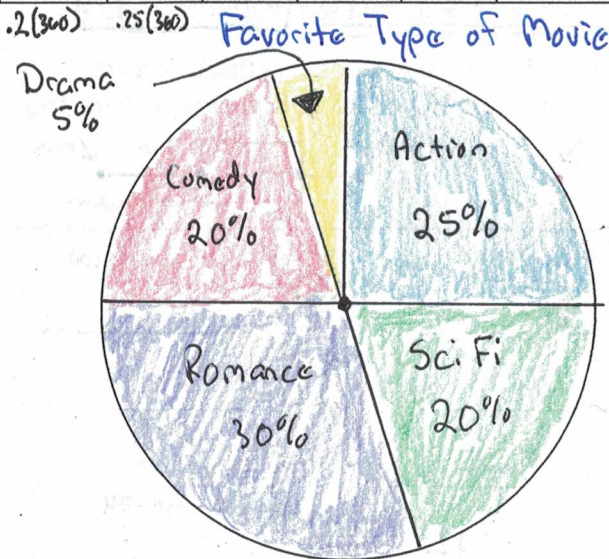
Carly did a survey with her friends about their favorite type of movies:

Class Survey: Favorite Type of Movie				
Comedy	Action	Romance	Drama	SciFi
4	5	6	1	4

Total = 20

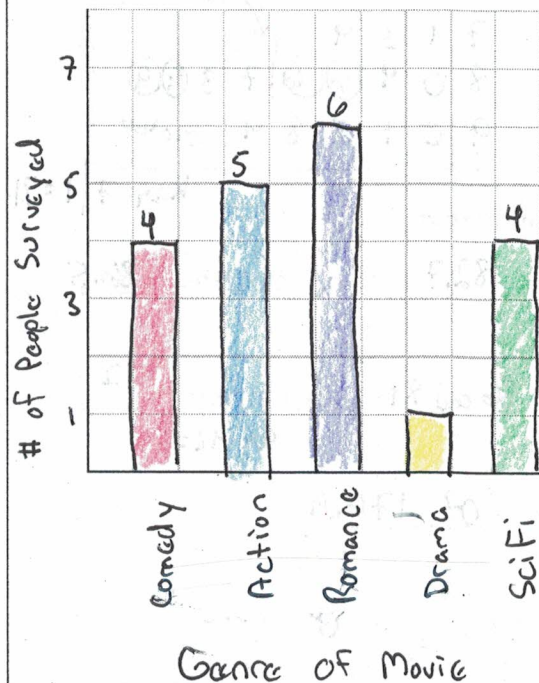
Data Display # 1 – Categorical: Pie Chart

Comd	Act	Rom	Dram	SciFi	Total
4	5	6	1	4	Total # = 20
$\frac{4}{20} = 20\%$	$\frac{5}{20} = 25\%$	$\frac{6}{20} = 30\%$	$\frac{1}{20} = 5\%$	$\frac{4}{20} = 20\%$	Total % = 100%
72°	90°	108°	18°	72°	Total Deg = 360°



Data Display # 1 – Categorical: Bar Graph

Title: Favorite Type of Movie



* All graphs/charts - need a title!!
Pie charts must be accurate!

Data Display # 2 – Quantitative

– **quantitative variable (in a display)** → data that can be displayed using numerical values

Ex: 1.) 1.2 million, 4.3 million, 2.5 million, etc. → category = population / salary

2.) 8 lb, 9 lb, 12 lb, 27 lb, etc. → category = weights

3.) 36 inches, 23 inches, 48 inches, etc. → category = lengths / heights

• **stem-and-leaf plot** – a type of quantitative data display that is organized from LEAST TO GREATEST and separated into 2 columns (mainly useful for organizing long list of numbers)

• **box-and-whisker plot** – a type of quantitative data display that is organized in to quartiles (4 equal parts) and shows the “spread” (min/max, median, range (and outliers)) of a set of data

Example 2: Express the following data as a stem-and-leaf plot and a box-and-whisker plot.

The class scores on a 50-item test are shown in the table below.

71 ✓	95 ✓	84 ✓	98 ✓	88 ✓	74 ✓
90 ✓	89 ✓	86 ✓	42 ✓	99 ✓	86 ✓
91 ✓	73 ✓	66 ✓	87 ✓	89 ✓	80 ✓

Data Display # 2 – Quantitative: Stem-Leaf Plot	Data Display # 2 – Quantitative: Box-Whisker Plot														
<p style="text-align: center; color: blue;">Class Scores</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th style="width: 50%;">Steam</th> <th style="width: 50%;">Leaf</th> </tr> </thead> <tbody> <tr> <td style="text-align: center;">4</td> <td style="text-align: center;">2</td> </tr> <tr> <td style="text-align: center;">5</td> <td></td> </tr> <tr> <td style="text-align: center;">6</td> <td style="text-align: center;">6</td> </tr> <tr> <td style="text-align: center;">7</td> <td style="text-align: center;">1 3 4</td> </tr> <tr> <td style="text-align: center;">8</td> <td style="text-align: center;">0 4 <u>6</u> <u>6</u> 7 8 <u>9</u> <u>9</u></td> </tr> <tr> <td style="text-align: center;">9</td> <td style="text-align: center;">0 1 5 8 9</td> </tr> </tbody> </table> <p style="text-align: right; color: red;">median</p> <p style="text-align: right; color: green;">modes</p> <p style="text-align: right; color: blue;">key 7 1 = 71</p>	Steam	Leaf	4	2	5		6	6	7	1 3 4	8	0 4 <u>6</u> <u>6</u> 7 8 <u>9</u> <u>9</u>	9	0 1 5 8 9	<p>minimum (min) → the lowest # that is <u>not</u> an outlier</p> <p>maximum (max) → the highest # that is <u>not</u> an outlier</p> <p>lower quartile (LQ / Q1) → median of lower half of data</p> <p>upper quartile (UQ / Q3) → median of upper half of data</p> <p>interquartile range (IQR) → range of the middle half of data and contains <u>50%</u> of data set: IQR = UQ – LQ</p> <p>outlier → an element of a set of data that's at least 1.5 IQR less than the LQ or 1.5 IQR greater than the UQ</p> <p>Min = <u>66</u> LQ = <u>74</u></p> <p>Max = <u>99</u> UQ = <u>90</u></p> <p>Median = <u>86.5</u> IQR = <u>16</u></p> <p style="text-align: right; color: green;">UQ - LQ 90 - 74</p> <p>Any outliers? <u>42</u></p>
Steam	Leaf														
4	2														
5															
6	6														
7	1 3 4														
8	0 4 <u>6</u> <u>6</u> 7 8 <u>9</u> <u>9</u>														
9	0 1 5 8 9														
<p>Find the following:</p> <p>a.) mean: <u>82.7</u> b.) median = <u>86.5</u></p> <p>c.) mode: <u>86 and 89</u> d.) range = <u>57</u></p> <p style="text-align: right; color: green;">99 - 42 =</p> <p>e.) variance: <u>0² 176.9</u></p> <p>f.) standard deviation: <u>σ 13.3</u></p>	<p style="text-align: center;">Lower Boundary LQ - 1.5(IQR) 74 - 1.5(16) 59</p> <p style="text-align: center;">Upper Boundary UQ + 1.5(IQR) 90 + 1.5(16) 114</p> <p style="text-align: right; color: red;">Any # greater than 114 is an outlier</p>														