Name: Key

1) Representations of Categorical Data

What's your favorite color?
When asked this question, the most popular color named was blue.

Symbols

Key:
Male = M Female = F
Blue = B Not Blue = N

Sample size = 200

$P(B) = \frac{84}{200}$

$P(M) = \frac{64}{200}$

$P(F|B) = \frac{48}{84}$

$P(B|F) = \frac{48}{136}$

$P(M \cap B) = \frac{36}{200}$

$P(M \cup B) = \frac{112}{200}$

2-Way Table

	Blue	Not Blue	Total
Male	36	28	64
Female	48	88	136
Total	84	116	200

Dependent
Is color preference independent of gender?

How do you know?

$P(A) = \text{Male } \frac{64}{200} = .32$

$P(A|B) = P(A)$

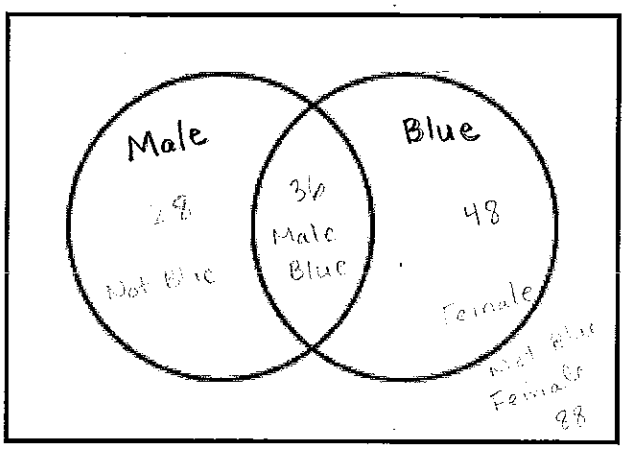
$\frac{P(A \cap B)}{P(B)} \quad \frac{36}{84} = .43$

$P(B) = \text{Blue } \frac{84}{200} = .42$

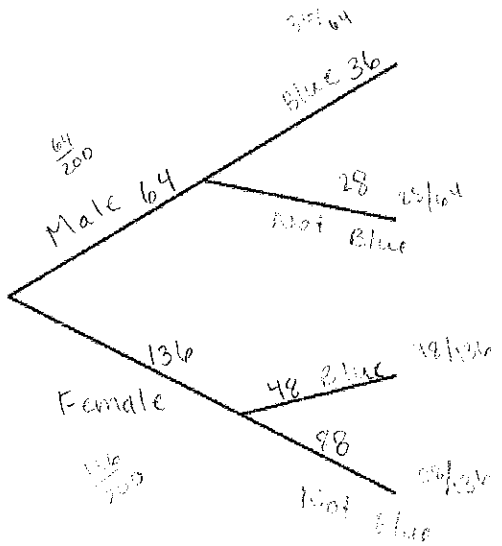
$P(B|A) = P(B)$

$\frac{P(A \cap B)}{P(A)} \quad \frac{36}{64} = .56$

Venn Diagram



Tree Diagram



2) Representations of Categorical Data

Are you a lefty or a righty?

Symbols

Key:
Male = M Female = F
Lefty = L Righty = R

Sample size = 100

P(L) = $\frac{50}{100}$

P(M) = $\frac{90}{100}$

P(F) = $\frac{10}{100}$

P(L|F) = $\frac{3}{10}$

P(L|M) = $\frac{47}{90}$

In this sample are there equal proportions of males and females who are left handed?
Explain.

$\frac{3}{10}$ $\frac{47}{90}$ no

2-Way Table

	Lefty	Righty	Total
Male	47	43	90
Female	3	7	10
Total	50	50	100

Is "handedness" independent from gender? Dependent no

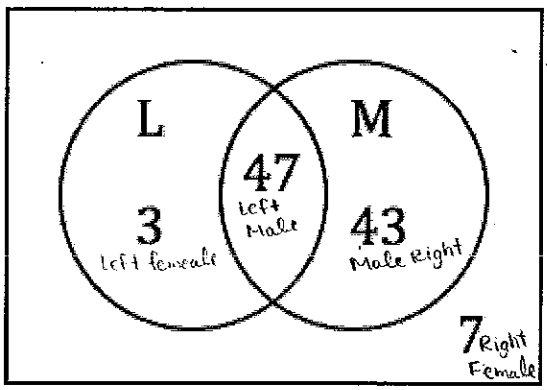
How do you know?

P(L) = $\frac{50}{100}$ P(M) = $\frac{90}{100}$

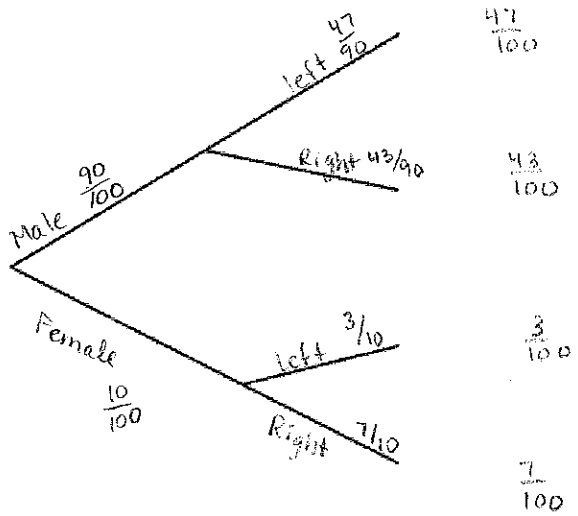
P(L|M) P(M|L)

$\frac{P(L \cap M)}{P(M)} = \frac{47}{90}$ $\frac{P(L \cap M)}{P(L)} = \frac{47}{50}$

Venn Diagram



Tree Diagram



3) Representations of Categorical Data

Do you eat breakfast or not?

Symbols

Key:

Male = M Female = F

East Breakfast = E Doesn't Eat Breakfast = D

Sample size = 685

P(E) = 351/685

P(E|M) = 247/425

P(E∩M) = 247/685

P(E|F) = 104/260

P(E∩F) = 104/685

2-Way Table

	Eats	Doesn't	Total
Male	247	178	425
Female	104	156	260
Total	351	334	685

Dependent

Is eating breakfast (or not) independent from gender? P(M) = 425/685 = .62 P(E) = 351/685 = .51

How do you know?

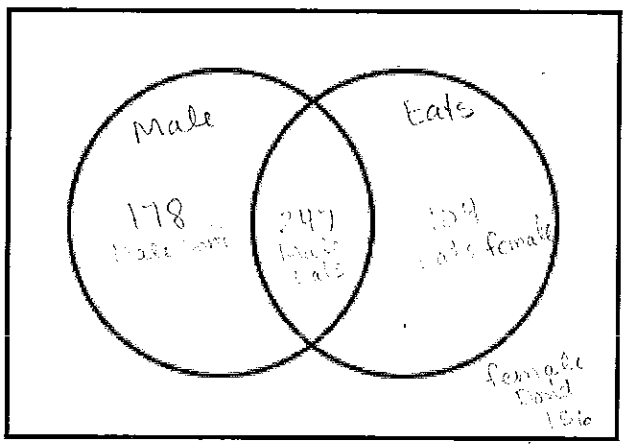
P(M|E) = P(M∩E)/P(E)

247/351 = .7

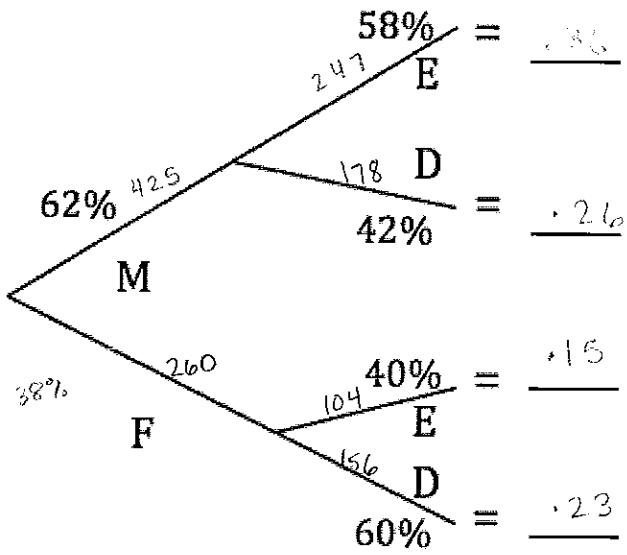
P(E|M) = P(M∩E)/P(M)

247/425 = .58

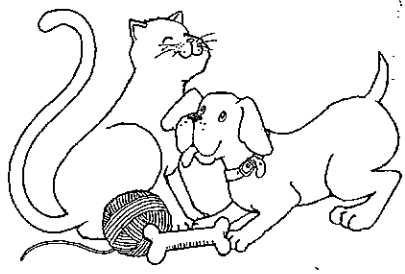
Venn Diagram



Tree Diagram



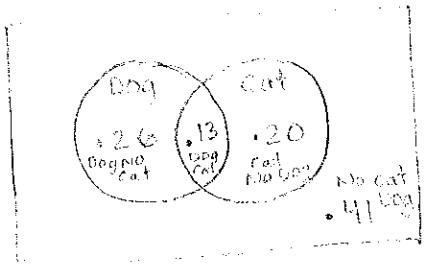
4) Representations of Categorical Data



The Humane Society likes to keep track of the percent of people who are dog and cat owners. Some people own only dogs, some only own cats, and some own both.

The Humane Society reports that 13% of households own both dogs and cats, 33% own cats, 39% own dogs, and 41% don't own either.

Is ownership of dogs vs. cats independent? Justify our response with appropriate representations of the date and numerical calculations.



$P(D) = \frac{39}{100} = .39$

$P(D|C) = \frac{P(D \cap C)}{P(C)}$

$\frac{.13}{.33}$

Same

$\frac{39}{100} = \frac{13}{33} = .39$

$P(C) = \frac{33}{100} = .33$

$P(C|D) = \frac{P(D \cap C)}{P(D)}$

$\frac{.13}{.39} = .33$

Same

$\frac{33}{100} = \frac{13}{39} = .33$

cats vs dogs is independent

	own cat	don't own cat	total
own dog	.13	.26	.39
don't own dog	.20	.41	.61
total	.33	.67	100