(1)

department	age	salary	status	count
sales	31_35	46K_50K	senior	30
sales	26_30	26K_30K	junior	40
sales	31_35	31K_35K	junior	40
systems	21_25	46K_50K	junior	20
systems	31_35	66K_70K	senior	5
systems	26_30	46K_50K	junior	3
systems	41_45	66K_70K	senior	3
marketing	36_40	46K_50K	senior	10
marketing	31_35	41K_45K	junior	4
secretary	46_50	36K_40K	senior	4
secretary	26_30	26K_30K	junior	6

C_1 = Junior
C_2 = Senior

Data to be classified:
X = (dep = systems, age = 26_30, salary = 46_50

1. Given a data tuple having the values "systems", "26_30", and "46K_50K" for the attributes department, age, and salary, respectively, what would a naive Bayesian classification of the status according to the data above? Notice that *Count* column is **NOT** an attribute. It just tells how many times a row occurs in our database and *status* is our target variable.

$P(C_i)$: $P(junior) = \frac{40+40+20+3+4+6}{165} = \frac{113}{165} = 0.684848$

$P(Senior) = \frac{30+5+3+10+4}{165} = \frac{52}{165} = 0.315152$

$P(X|C_i)$:

$P(dep=systems \mid junior) = \frac{20+3}{113} = \frac{23}{113} = 0.20354$

$P(systems \mid senior) = \frac{8}{52} = 0.153846$

$P(26\_30 \mid junior) = \frac{49}{113} = 0.433628$

$P(26\_30 \mid senior) = \frac{0}{52} = 0$

$P(46k-50k \mid junior) = \frac{23}{113} = 0.20354$

$P(46k-50k \mid Senior) = \frac{40}{52} = 0.769231$

$P(X \mid junior) = 0.20354 \times 0.433628 \times 0.20354$
$= 0.017965$

$P(X \mid senior) = 0.153846 \times 0 \times 0.769231 = 0$

$$P(X \mid C_i) \cdot P(C_i) = P(X \mid junior) \cdot P(junior) \approx 0.012303$$

$$P(X \mid senior) \cdot P(senior) = 0$$

X Belongs to class "junior"