1.1 Jerry and Susan have a joint bank account. Jerry goes to the bank 20% of the days. Susan goes there 30% of the days. Together they are at the bank 8% of the days.

a. Susan was at the bank last Monday. What's the probability that Jerry was there too?

b. Last Friday, Susan wasn't at the bank. What's the probability that Jerry was there?

c. Last Wednesday at least one of them was at the bank. What is the probability that both of them were there?

Event A = On that day Jerry will be at the bank

Event B = On the day Susan is at the bank

$P(A) = 0.2$, $P(B) = 0.3$, $P(A \cap B) = 0.08$

a. Last Monday Susan is at the bank, so the probability that Jerry was there too
   $P(A \mid B) = P(A \cap B) / P(B) = 0.08 / 0.3 = 0.267$

b. Last Friday Susan wasn't at the bank, so the probability that Jerry was there is,
   $P(A|B') = P(A) - P(A \cap B) / 1 - P(B) = 0.2 - 0.08 / 1 - 0.3 = 0.1714$

c. Last Wednesday at least one of them was at the bank, then the probability of both of them at the bank is,
   $P[(A \cap B) \mid (A \cup B)] = P[(A \cap B) \cap (A \cup B)] / P(A \cup B) = P(A \cap B) / P(A \cup B)$
   $P(A \cup B) = P(A)+P(B)-P(A \cap B) = 0.2 + 0.3 - 0.08 = 0.42$
   Probability $= 0.08 / 0.42 = 0.19$

1.2 Harold and Sharon are studying for a test. Harold's chances of getting a "B" are 80%. Sharon's chances of getting a "B" are 90%. The probability of at least one of them getting a "B" is 91%.
a. What is the probability that only Harold gets a "B"?
b. What is the probability that only Sharon gets a "B"?
c. What is the probability that both won't get a "B"?

Event H = Harold's chances of getting "B"

Event S = Sharon's chances of getting "B"
$P(H) = 0.8$, $P(S) = 0.9$, $P(H \cup S) = 0.91$
$P(H \cap S) = P(H) + P(S) - P(H \cup S) = 0.8 + 0.9 - 0.91 = 0.79$

a. Probability that only Harold gets a "B",
   $P(H \cap S') = P(H) - P(H \cap S) = 0.8 - 0.79 = 0.01$

b. Probability that only Sharon gets a "B",
   $P(H' \cap S) = 0.9 - 0.79 = 0.11$

c. Probability that both won't get "B",

   $P(H \cup S)' = 1 - P(H \cup S) = 1 - 0.91 = 0.09$

1.3 Jerry and Susan have a joint bank account. Jerry goes to the bank 20% of the days. Susan goes there 30% of the days. Together they are at the bank 8% of the days. Are the events "Jerry is at the bank" and "Susan is at the bank" independent?

Event A = On the day Jerry will be at the bank
Event B = On the day Susan is at the bank

$P(A) = 0.2$, $P(B) = 0.3$, $P(A \cap B) = 0.08$

$P(A \cap B) = P(A) * P(B) = 0.2 * 0.3 = 0.06 \neq 0.08$
Here, $P(A \cap B) \neq P(A) * P(B)$
So, these two events are not independent.

1.4 You roll 2 dice.

a. **Are the events "the sum is 6" and "the second die shows 5" independent?**
b. **Are the events "the sum is 7" and "the first die shows 5" independent?**

a. Event A = The sum is 6
   Event B = The second die shows 5
   $P(A) = 5 / 36$
   $P(B) = 1 / 36$
   P(The sum is 6 & The second die shows 5) = $P(A \cap B) = 1 / 36$
   $P(A \cap B) \neq P(A) * P(B)$
   So, these events are not independent.

b. Event A = The sum is 7
   Event B = The first die shows 5
   $P(A) = 6 / 36 = 1/6$
   $P(B) = 6 / 36 = 1/6$
   P(The sum is 7 & The first die shows 5) = $P(A \cap B) = 1 / 36 = P(A) * P(B)$
   Here, $P(A \cap B) = P(A) * P(B)$. So, these events are independent.

1.5 An oil company is considering drilling in either TX, AK and NJ. The company may operate in only one state. There is 60% chance the company will choose TX and 10% chance - NJ. There is 30% chance of finding oil in TX, 20% - in AK, and 10% - in NJ.

1. **What's the probability of finding oil?**
2. **The company decided to drill and found oil. What is the probability that they drilled in TX?**

Event A = Company choose TX

Event B = Company choose AK

Event C = Company choose NJ

Event D = Find oil in TX

Event E = Find oil in AK

Event F = Find oil in NJ
P(A) = 0.6, P(C) = 0.1, P(B) = 1 – 0.6 – 0.1 = 0.3, P(D) = 0.3, P(E) = 0.2, P(F) = 0.1

1. P (Finding oil) = 0.6 * 0.3 + 0.3 * 0.2 + 0.1 * 0.1 = 0.25
2. P (Drill and find oil in TX) = 0.6 * 0.3 / 0.25 = 0.72

1.6 The following slide shows the survival status of individual passengers on the Titanic. Use this information to answer the following questions.

a. What is the probability that a passenger did not survive?

b. What is the probability that a passenger was staying in the first class?

c. Given that a passenger survived, what is the probability that the passenger was staying in the first class?

d. Are survival and staying in the first class independent?

e. Given that a passenger survived, what is the probability that the passenger was staying in the first class and the passenger was a child?

f. Given that a passenger survived, what is the probability that the passenger was an adult?

g. Given that a passenger survived, are age and staying in the first class independent?

a. Event A = Passenger did not survive
   P(A) = (1490 – 673) / (2201 – 885) = 0.620
b. Event B = Person in 1^{st} Class
   P(B) = 325 / (2201 – 885) = 0.246
c. Event C = Passenger is survived and that was an adult
   P(C) = 203 / (711 – 212) = 0.406
d. Event D = All Survived
   Event E = Staying in 1^{st} Class
   P(D) = 711 / 2201 = 0.323
   P(E) = 325 / 2201 = 0.147
   P (D ∩ E) = 203 / 2201 = 0.092
   Hence, P (D ∩ E) ≠ P(D) * P(E). Therefore, these events are not independent.
e. Event F = Passenger is survived, Staying in 1^{st} Class and that was child
   P(F) = 6 / (711 – 212) = 0.012
f. Event G = Passenger is survived and that was an adult

P(G) = (654 – 212) / (711 – 212) = 0.885

g. Event H = Passenger are age and Survived
Event I = Staying in 1^{st} class and Survived
P (Adult and Survived) = (442 / 499) = 0.886
P (Child and Survived) = (57 / 499) = 0.114

P(H) = P (Adult and Survived) + P (Child and Survived)
= 0.886 + 0.114
= 1
P(I) = 203 / 499 = 0.406

P (Adult Survived & Staying in 1^{st} class) = 197 / 499 = 0.394
P (Child Survived & Staying in 1^{st} class) = 6 / 499 = 0.012
P(H∩I) = P (Adult Survived & Staying in 1^{st} class) + P (Child Survived & Staying in 1^{st} class)
= 0.394 + 0.012 = 0.406
P(H∩I) = P(H) * P(I) = 1 * 0.406 = 0.406
Hence, P(H∩I) = P(H) * P(I).
Therefore, these two events are independent.

1.7 A developer claims that her app can distinguish AI-generated documents from human generated ones. To assess its performance, we have submitted 1000 AI-generated and 1000 human-generated documents to the app.

• The app misclassified 70 human-generated documents as AI-generated

• and 30 AI generated documents as human- generated.

Build the confusion matrix for the above app and calculate the following:

Accuracy, precision, recall and F1

True Positive (TP): The number of human-generated documents correctly classified as human-generated.

True Negative (TN): The number of AI-generated documents correctly classified as AI-generated.

False Positive (FP): The number of AI-generated documents misclassified as human-generated.

False Negative (FN): The number of human-generated documents misclassified as AI-generated.

Total human-generated documents: 1000

Total AI-generated documents: 1000

	Predicted AI-Generated	Predicted Human-Generated
Actual AI-Generated	TN = 970	FP = 30
Actual Human-Generated	FN = 70	TP = 930

Accuracy = TP + TN / Total = (930 + 970) / 2000 = 0.95

Precision = TP / (TP + FP) = 930 / (930 + 30) = 0.968

Recall = TP / (TP + FN) = 930 / (930 + 70) = 0.93

F1 = 2 * Precision * Recall / (Precision + Recall)

= 2 * 0.968 * 0.93 / (0.968 + 0.93)

= 0.948