Written questions

Question 1

1. If you use same keystream, C1 XOR C2 = P1 XOR K XOR P2 XOR K = P1 XOR P2.
2. Given we do not know the key and P1 and P2, we can obtain plaintexts P1 and P2 by C1 XOR C2 together to eliminate the key. Then the result is a matter of identifying patterns. Since it is only 52 english letters, there is no special character like spaces. We do this by detecting a pattern within the data.

Question 2

a) If an attacker controls the exit, middle and guard, they can perform many different attacks. They can see unencrypted traffic leaving the circuit from the exit relay, observe the entire circuit and perform traffic analysis to link the source and destination. Attackers can manipulate the plaintext as well.
b) Periodic relay changes means relays in the circuit are changed at intervals. 2 effects of periodic relay changes on user privacy. Firstly, changing relays periodically makes it harder for an attacker to correlate a user's activity over time. Secondly, malicious relays may lose their positions periodically, making it more challenging for them to consistently target a specific user.
c) If email provider colludes with adversary controlling relays, all connections to the email account can be linked to the user, as the collusion allows the adversary to associate the user's actions with their identity. Thus, they can link the entire pathway to the user compromising anonymity.
d) Advantage of not changing circuits is users experience less disruption and certain applications benefit from a stable circuit. Disadvantage is increased vulnerability as it long term tracking and correlation is easier for adversaries, they can potentially map out the encryption of data passing through the tunnel.
e) Choosing a small set of guard relays when the user first joins provides a balance between security and performance. It helps prevent a single malicious relay from always being the first hop, improving security. Changing middle and exit relays also add another layer of protection.
f) Tor do not delay or re order packets when forwarding them. Attackers can use circuit fingerprinting attack. In circuit fingerprinting attack, adversary controls the entry node. the adversary injects specific patterns into the user's traffic and these patterns can be introduced by manipulating packet timings. Since there is no reordering or delayed packets, the injected patterns will maintain the same order as they pass through. Since the adversary also controls the exit node, he can observe the traffic leaving the tor network. By analyzing the traffic and patterns, the adversary can correlate them with the injected patterns and can successfully link the source (user) and destination of the communication
g) The onion service can publish content anonymously without being dependent on any company for things like domain name registration. Normally you need domain name and credentials and pay for it. but for domain name in onion service you don't need to pay fees or give your credential information. you can use your computer to host website files so that

you don’t need to share your domain name or host service. Also your real ip address is hidden due to the tor relays.

Question 3

a) Output of queries with certain conditions will differ based on whether Charlie’s record is included or not. Tracker here is the difference in the output of the sum(Grade) query with and without charlies record

tracker:

This tracker is used to measure the difference in query results when Charlie’s record is included or excluded

SELECT SUM(Grade) FROM Student WHERE Gender = ‘M’ AND Postal_Code = ‘G3R 4S2’

Query 1 include charlies record

SELECT SUM(Grade) FROM Student WHERE Gender = ‘M’ AND Postal_Code = ‘G3R 4S2’ AND Birthdate = ‘0222’

Query 2 no Charlie’s record

SELECT SUM(Grade) FROM Student WHERE Gender = ‘M’ AND Postal_Code = ‘G3R 4S2’ AND NOT (Birthdate = ‘0222’) (Name = ‘Charlie’)

Query 3 Random

SELECT SUM(Grade) FROM Student WHERE Gender = 'F' AND Postal_Code = 'N2R 3M5'

Take result 1 – result2 > k, we can infer charlie’s grade is included in result1 but not in result2 because there are same number of males and females. Therefore the grade difference is the d

b) To find Natalie’s grade,

Tracker query

SELECT COUNT(*) FROM Student WHERE Gender = 'F' AND Name != 'Natalie';

Query counts all females excluding natalie.

Query 1 Exclude Natalie's record

SELECT COUNT(*) FROM Student WHERE Gender = 'F' AND Name != 'Natalie';

Query 2 combine count of females (excluding natalie)

SELECT COUNT(*) FROM Student WHERE Gender = 'F' AND Name != 'Natalie' AND Grade > 80;

This query counts the number of females excluding Natalie with a grade greater than 80

Query 3

SELECT COUNT(*) FROM Student WHERE Gender = 'F' AND Name != 'Natalie' AND Grade > 80 UNION SELECT COUNT(*) FROM Student WHERE Gender = 'F' AND Name != 'Natalie';

This query combines the count of females with a grade greater than 80 (excluding natalie) with the count of all females excluding natalie.

The count difference between query 2 and query 1 gives an indication of the number of females with a grade greater than 80 excluding Natalie.

The tracker attack using query counts provides information about the number of females with a grade greater than 80. Excluding natalie.

But if the count difference in query 3 is greater than 0 this suggest there are females with a grade greater than 80 and we can infer natalie might have a grade greater than 80.

This query combines the count of females excluding Natalie with the count of females

c) K anonymous, each combination of values in the columns must appear at least k times in the table.

Name Birthdate Gender Postal Code

* 091* F G9Q 3X2

* 113* F G9Q 3X2

* 043* M H4A 5A6

* 092* F Y1R 4J4

* 052* F H4A 5A6

* 100* F H4A 5A6

* 103* M H4A 5A6

* 123* M H4A 5A6

* 052* M Y1R 4J4

* 121* M Y1R 4J4

* 071* F G9Q 3X2

* 112* F H4A 5A6

* 072* M H4A 5A

For each unique combination of table 2 values:

(, 091, F): Appears 2 times

(, 113, F): Appears 2 times

(, 043, M): Appears 1 time

(, 092, F): Appears 1 time

(, 052, F): Appears 2 times

(, 100, F): Appears 1 time

(, 103, M): Appears 1 time

(, 123, M): Appears 1 time

(, 052, M): Appears 1 time

(, 121, M): Appears 1 time

(, 071, F): Appears 1 time

(, 112, F): Appears 1 time

(, 072, M): Appears 1 time

The minimum count is 1, which is less than 3. Therefore, Table 2 is not 3-anonymous.