CSC 370

Assignment 4

Question 1

The number of sorted sublists at the end of phase 1 is $= \dfrac{10,000,000}{320} = 31250$

The number of sorted sublists at the end of phase 2 is $= \dfrac{31250}{319} \approx 98$
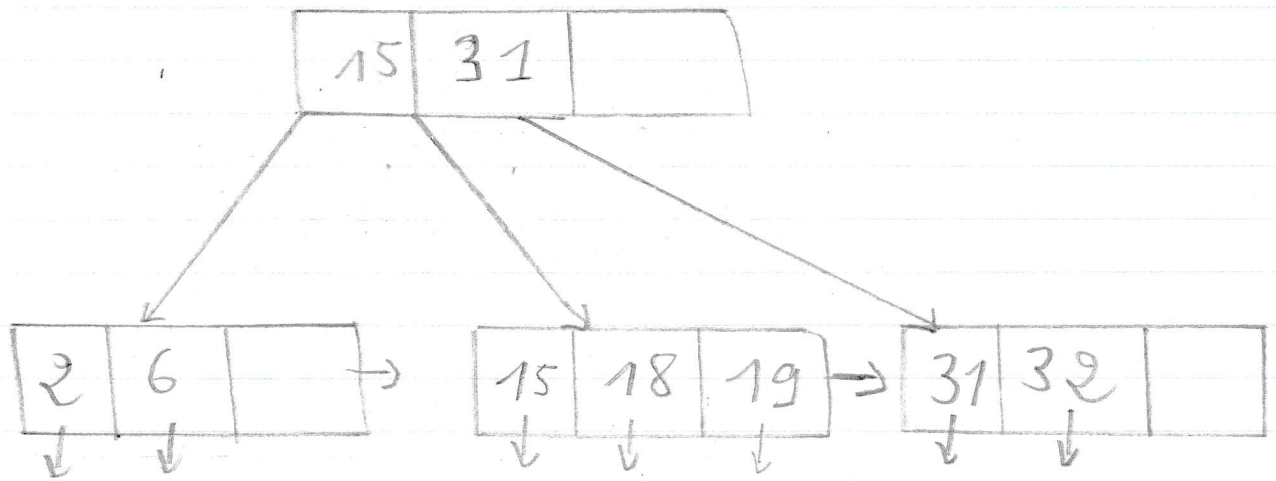
These sublists need to be merged in turn in another phase, phase 3.

Since in each phase, we read 10,000,000 blocks and write 10,000,000; so, the total number of I/O is $3 \times 2 \times 10,000,000 =$
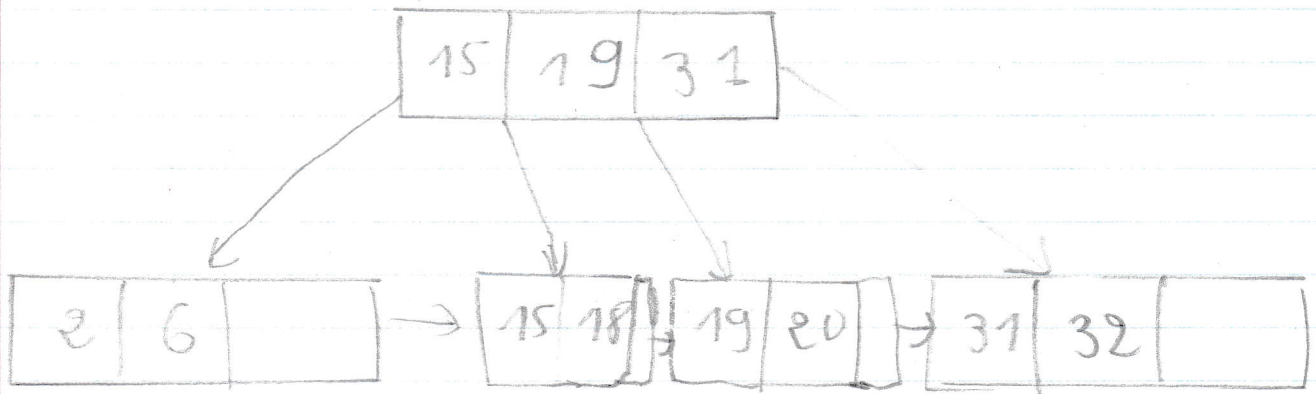
$$\boxed{60,000,000}$$

In terms of time, $6 \cdot 10^7 \times 16 \text{ ms} = 96 \cdot 10^7 \text{ ms}$
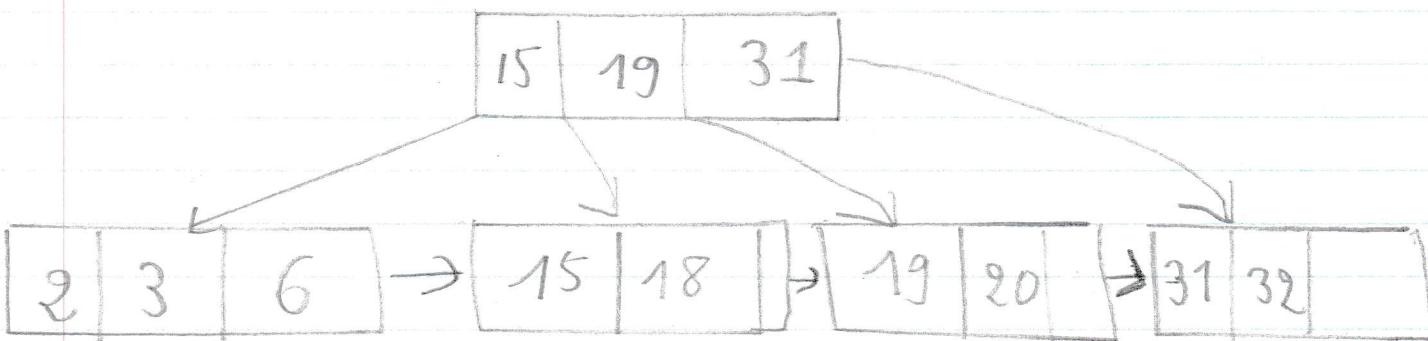
Insert 19

```
        ┌────┬────┬────┐
        │ 15 │ 31 │    │
        └────┴────┴────┘
       ╱      │        ╲
      ╱       │         ╲
┌───┬───┬───┐  ┌────┬────┬────┐  ┌────┬────┬────┐
│ 2 │ 6 │   │→ │ 15 │ 18 │ 19 │→ │ 31 │ 32 │    │
└───┴───┴───┘  └────┴────┴────┘  └────┴────┴────┘
  ↓   ↓          ↓    ↓    ↓        ↓    ↓
```

Insert 20

```
            ┌────┬────┬────┐
            │ 15 │ 19 │ 31 │
            └────┴────┴────┘
           ╱    │      │     ╲
          ╱     │      │      ╲
┌───┬───┬───┐  ┌────┬────┐  ┌────┬────┐  ┌────┬────┬────┐
│ 2 │ 6 │   │→ │ 15 │ 18 │→ │ 19 │ 20 │→ │ 31 │ 32 │    │
└───┴───┴───┘  └────┴────┘  └────┴────┘  └────┴────┴────┘
```

Insert 3

```
            ┌────┬────┬────┐
            │ 15 │ 19 │ 31 │
            └────┴────┴────┘
           ╱    │      │     ╲
          ╱     │      │      ╲
┌───┬───┬───┐  ┌────┬────┐  ┌────┬────┐  ┌────┬────┬────┐
│ 2 │ 3 │ 6 │→ │ 15 │ 18 │→ │ 19 │ 20 │→ │ 31 │ 32 │    │
└───┴───┴───┘  └────┴────┘  └────┴────┘  └────┴────┴────┘
```

Insert 4

```
                    ┌────┬────┬────┐
                    │ 19 │    │    │
                    └────┴────┴────┘
              ┌──────────┘         └──────────┐
     ┌────┬────┬────┐              ┌────┬────┬────┐
     │ 4  │ 15 │    │              │ 31 │    │    │
     └────┴────┴────┘              └────┴────┴────┘
        │    │      └────┐       ┌────┘      └────┐
```

```
┌───┬───┬───┐   ┌───┬───┬───┐   ┌────┬────┬───┐   ┌────┬────┬───┐   ┌────┬────┬───┐
│ 2 │ 3 │   │ → │ 4 │ 6 │   │ → │ 15 │ 18 │   │ → │ 19 │ 20 │   │   │ 31 │ 32 │   │
└───┴───┴───┘   └───┴───┴───┘   └────┴────┴───┘   └────┴────┴───┘   └────┴────┴───┘
```

Insert 5

```
                    ┌────┬────┐
                    │ 19 │    │
                    └────┴────┘
              ┌──────┘         └──────────┐
     ┌────┬────┬────┐              ┌────┬────┐
     │ 4  │ 15 │    │              │ 31 │    │
     └────┴────┴────┘              └────┴────┘
      │    │      └────┐         ┌────┘     └────┐
```

```
┌───┬───┬───┐   ┌───┬───┬───┐   ┌────┬────┐   ┌────┬────┐   ┌────┬────┐
│ 2 │ 3 │   │ → │ 4 │ 5 │ 6 │ → │ 15 │ 18 │ → │ 19 │ 20 │ → │ 31 │ 32 │
└───┴───┴───┘   └───┴───┴───┘   └────┴────┘   └────┴────┘   └────┴────┘
```

Insert 40

```
                    ┌────┐
                    │ 19 │
                    └────┘
           ┌──────┬──────┐              ┌───────────┐
           │  4   │  15  │              │    31     │
           └──────┴──────┘              └───────────┘
         ╱       │       ╲            ╱              ╲
┌──────┬──────┐ ┌────┬────┬────┐ ┌──────┬──────┐ ┌──────┬──────┐ ┌──────┬──────┬──────┐
│  2   │  3   │→│ 4  │ 5  │ 6  │→│  15  │  18  │→│  19  │  20  │→│  31  │  32  │  40  │
└──────┴──────┘ └────┴────┴────┘ └──────┴──────┘ └──────┴──────┘ └──────┴──────┴──────┘
```

Insert 41

```
                    ┌────┐
                    │ 19 │
                    └────┘
         ╱                        ╲
┌──────┬──────┐              ┌──────┬──────┐
│  4   │  15  │              │  31  │  40  │
└──────┴──────┘              └──────┴──────┘
 ╱     │      ╲            ╱       │       ╲
┌────┬────┐ ┌────┬────┬────┐ ┌──────┬──────┐ ┌──────┬──────┐ ┌──────┬──────┐ ┌──────┬──────┐
│ 2  │ 3  │→│ 4  │ 5  │ 6  │→│  15  │  18  │→│ 19 │ 20 │→│ 31 │ 32 │→│ 40 │ 41 │
└────┴────┘ └────┴────┴────┘ └──────┴──────┘ └──────┴──────┘ └──────┴──────┘ └──────┴──────┘
```

Insert 42

```
                    ┌────┐
                    │ 19 │
                    └────┘
         ╱                        ╲
┌──────┬──────┐              ┌──────┬──────┐
│  4   │  15  │              │  31  │  40  │
└──────┴──────┘              └──────┴──────┘
 ╱     │      ╲            ╱       │       ╲
┌────┬────┐ ┌────┬────┬────┐ ┌──────┬──────┐ ┌──────┬──────┐ ┌──────┬──────┐ ┌──────┬──────┬──────┐
│ 2  │ 3  │→│ 4  │ 5  │ 6  │→│  15  │  18  │→│ 19 │ 20 │→│ 31 │ 32 │→│ 40 │ 41 │ 42 │
└────┴────┘ └────┴────┴────┘ └──────┴──────┘ └──────┴──────┘ └──────┴──────┘ └──────┴──────┴──────┘
```
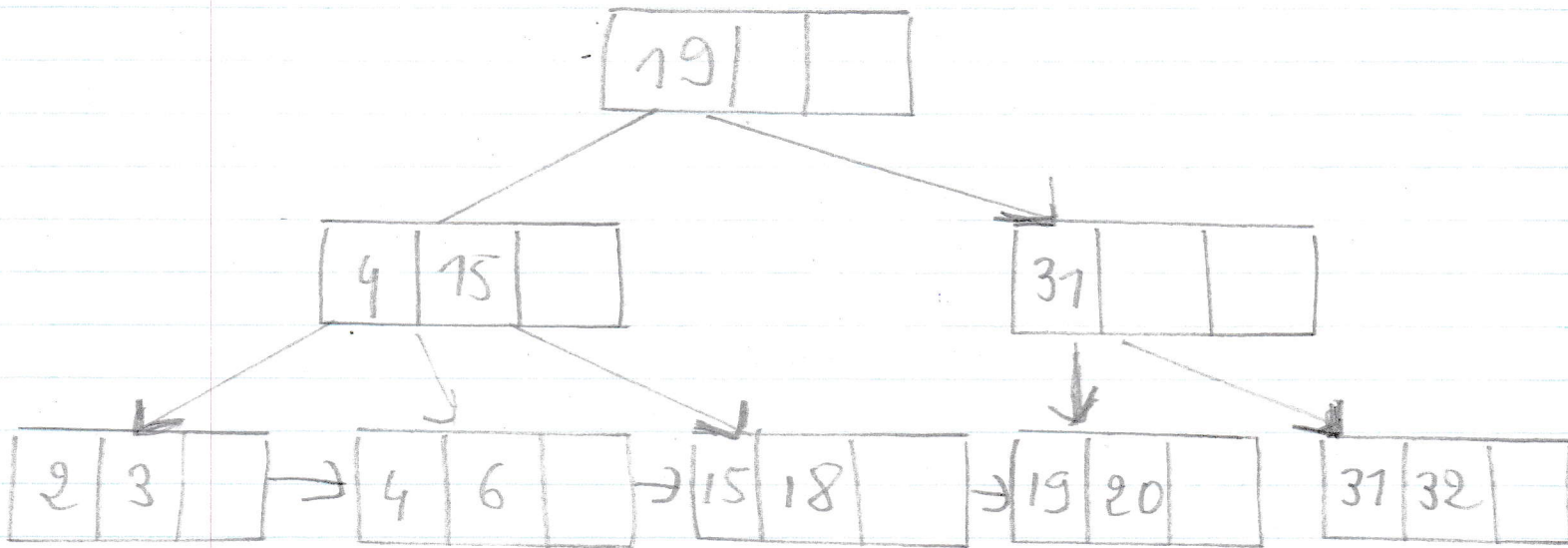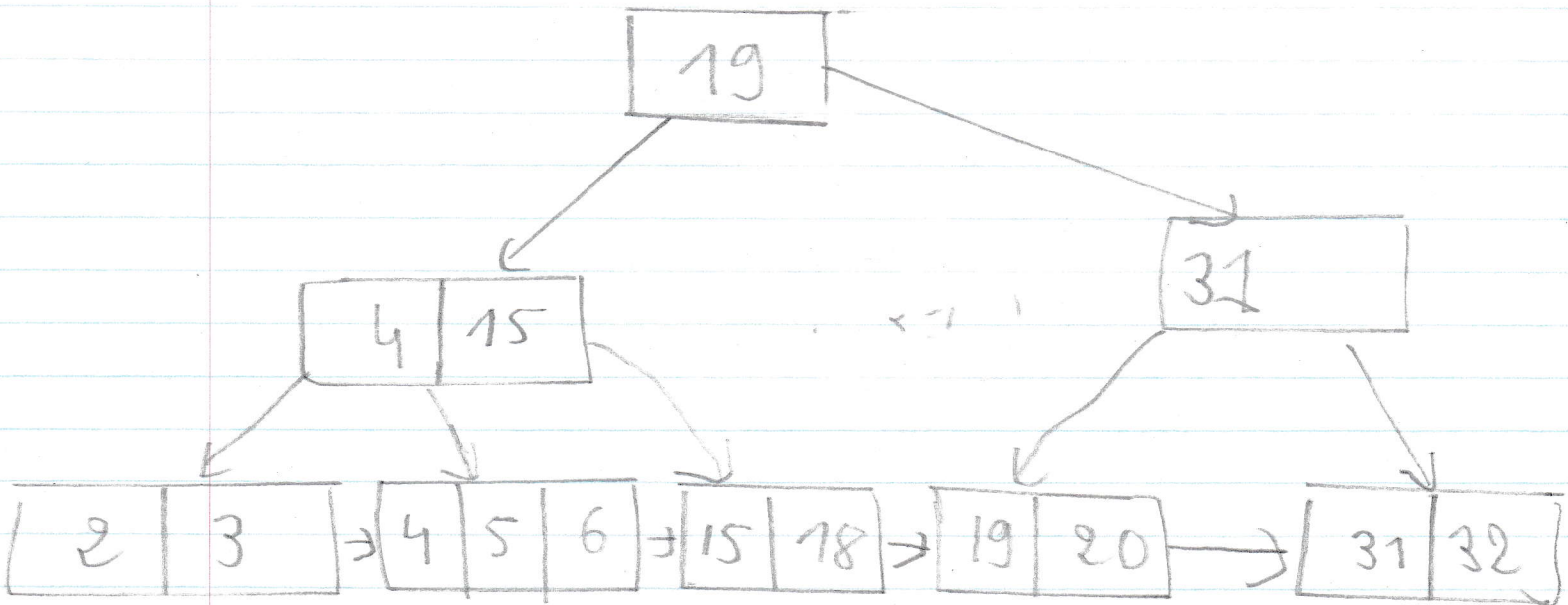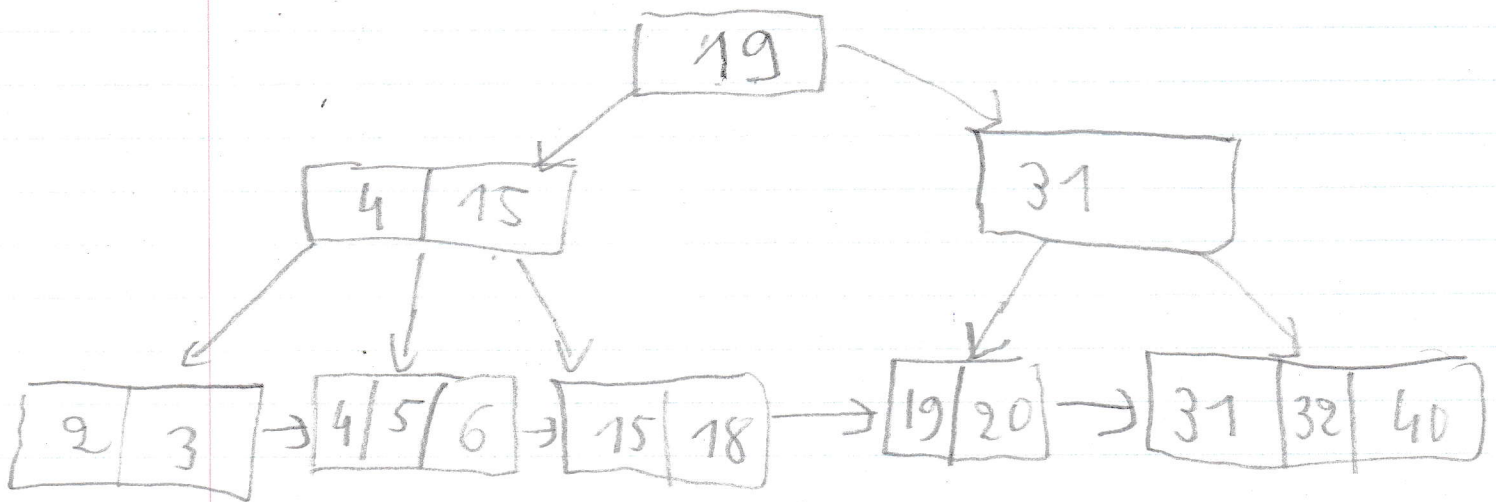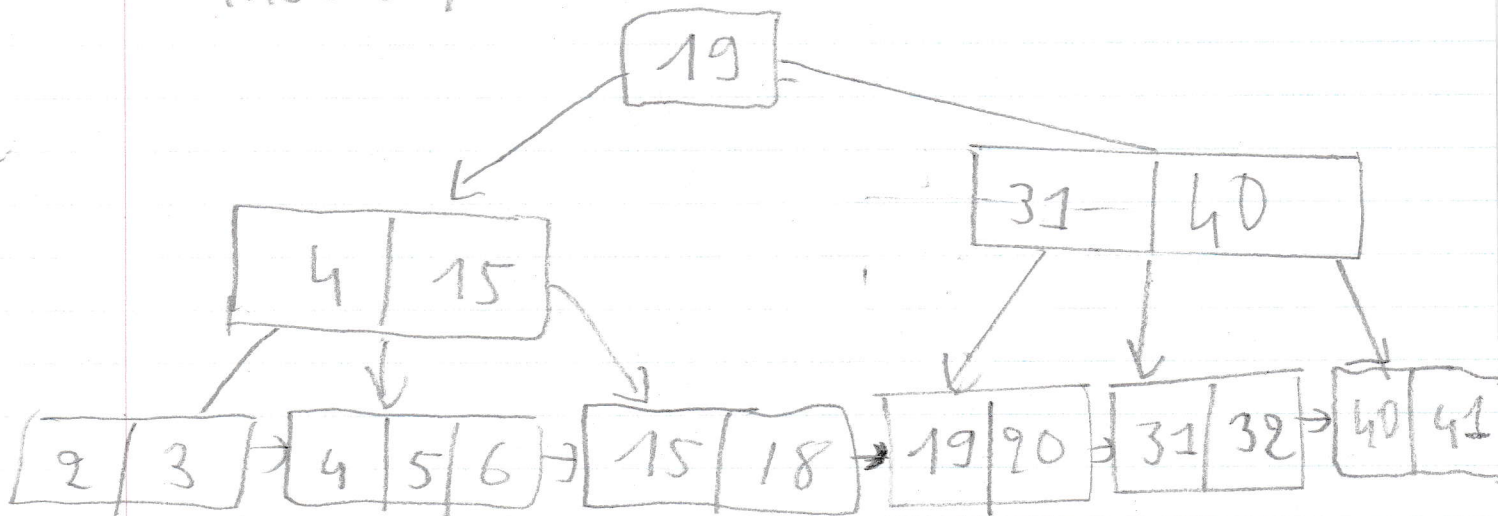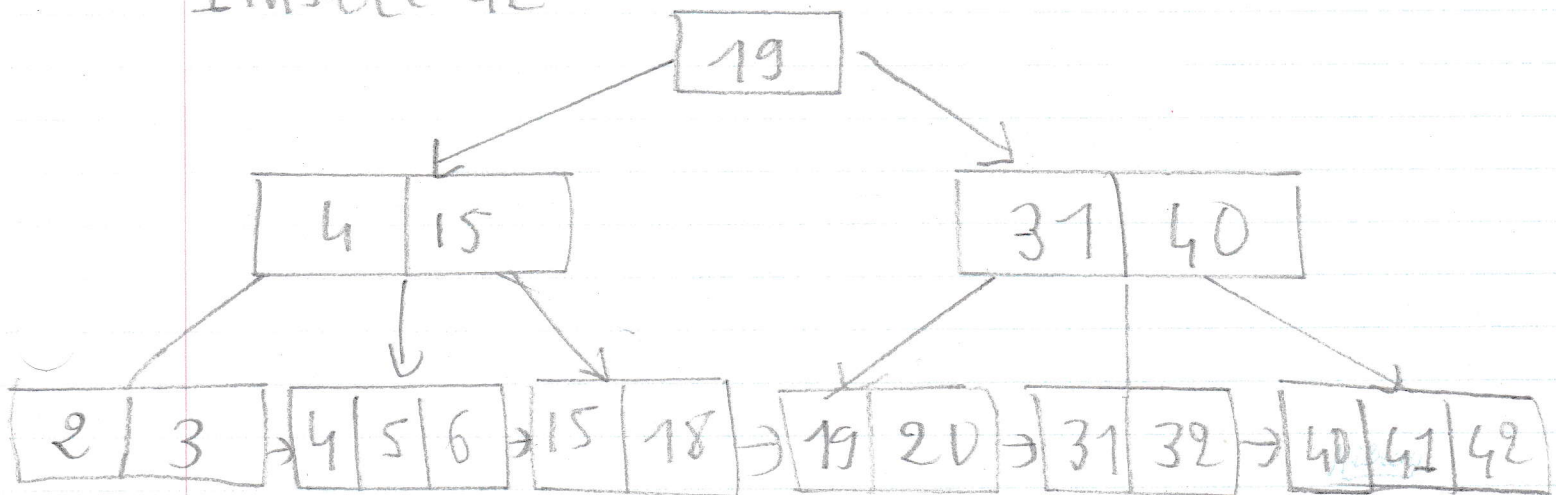
Question3

Assume for Maintenances:
 - a typ tuple is 160 bytes
 - a block can hold 100 tuples
 - we have 1000 blocks

Assume for Employees:
 - a tuple is 130 bytes
 - a block can hold 120 tuples
 - we have 50 blocks

Since the index on planned Id is clustering, and we can estimate the number of selected tuples to be $100,000/100 = 1000$, therefore the cost is $(1000/100) + 2 = 12$ I/Os

Cost of rating)s to Employees:
$50 + 25 = 75$ I/Os

- Scan Employees (50 blocks)
- Pipe

Cost of Employees (Scanning
and piping) is 50 I/0 S

- Cost of sorting Employees
using 2 PHHS is $2 \times 2 \times 25 = 100$ I/0S
- To join- merge, we need to scan
them, $10 + 25 = 35$ I/0S
The total cost of the plan is
$(12 + 50) + (100 + 35) = 197$ I/0 S

Question 4

$R(A, B, C, D)$ with FD's

$$AB \to C, \quad B \to D, \quad CD \to A, \quad AD \to B$$

Let compute $\{A,B\}^+, \{B\}^+, \{C,D\}^+$ and $\{A,D\}^+$

$$\{A,B\}^+ = \{A, B, C, D\}$$

$$\{B\}^+ = \{B, D\} \Rightarrow \text{violates BCNF}$$

$$\{C,D\}^+ = \{C, D, A, B\}$$

$$\{A,D\}^+ = \{A, D, B, C\}$$

Since $\{A,B\}^+, \{C,D\}^+, \{A,D\}^+$ give us all the attributes, there are not BCNF violation for $AB \to C, CD \to A$ and $AD \to B$.

On the other hand, $B \to D$ violates BCNF rule

Since $\{B\}^+ = \{B, D\}$

We can decompose R into:

$R_1(A, B, C), R_3(C, D, A), R_4(A, D, B)$

$R_2(B, D)$. But since $B \to D$ violate BCNF

we would like to decompose R_2, but
we cannot do that since R_2 has only
two attributes.
Thus the relations ~~R_1 and~~ R_2 remains
unchanged. So ~~R_1 (A, B, C) and~~ $R_2 (B, D)$ does not
change.

* Q_{4-2},

$R(A, B, C, D)$ with FD'S
$A \Rightarrow B$, $B \Rightarrow C$, $C \Rightarrow D$, $D \Rightarrow A$
Let's compute A^+, B^+, C^+ and D^+
$A^+ = \{A, B, C, D\}$
$B^+ = \{B, C, D, A\}$
$C^+ = \{C, D, A, B\}$
$D^+ = \{D, A, B, C\}$

There is no BCNF violation here since
we get all of the attributes for each function
dependency. Thus, $R_1 (A, B)$, $R_2 (B, C)$
$R_3 (C, D)$ and $R_4 (D, A)$