Question 1

- 20000 tuples
- 600 data pages
- Prefix: 20 bytes
- Full: 30 bytes
- 200 different types
- 50 producers
- Rid has 10 bytes
- Pointer has 6 bytes
- Leaf pages are filled about 70%
- Index page has 4000 Bytes

1) We have $200 * 50 = 10000$ possible different values. Then as its uniformly distributed we have 10000 data entries
   The number of rids per data entry is $\left(\frac{number\ of\ tuples}{diff\ values}\right) = \frac{20000}{10000} = 2$
   The average length of a data entry is $size\ of\ key + nb)(rids * size(rids) = 30 * 2 + 2 * 10 = 80$

2) The size of a index entry is $20 + 20 + 6 = 46$. Then the average number of index entry per intermediate page is $\frac{4000}{46} = 86$ and the average number of data entry per page is $\frac{0.7*4000}{80} = 35$. If we have a tree of height 2 then we can cover at most $86 * 86 = 7396$ different cases then we need a tree of height 3 as $86 * 86 * 86 = 636056 \gg 10000$.
   Number of leaf pages is $\frac{10000}{35} = 285$ and as the root is full it has 86 data entry so the there is 87 intermediate pages.