

Q1. (a)

(i)	<u>Doc 1</u>	<u>Doc 2</u>	<u>Doc 3</u>
Covid	1	1	1
cases	1	0	1
in	1	1	1
India	1	1	0
are	1	0	0
diminishing	1	0	0
2020	0	1	0
Cumulative	0	1	0
vaccine	0	1	0
coverage	0	1	0
exceeds	0	1	0
173.42	0	1	0
crosses	0	1	0
may	0	0	1
still	0	0	1
surge	0	0	1
July	0	0	1

(ii) are \rightarrow 1 (stop word)
Covid \rightarrow 1, 2, 3
cases \rightarrow 1, 2, 3
in \rightarrow 1, 2, 3 (stop word)
India \rightarrow 1, 2
diminishing \rightarrow 1
2020 \rightarrow 2
cumulative \rightarrow 2
vaccine \rightarrow 2
coverage \rightarrow 2
exceeds \rightarrow 2
173.42 \rightarrow 2
cores \rightarrow 2
may \rightarrow 3 (stop word)
still \rightarrow 3 (stop word)
Surge \rightarrow 3
July \rightarrow 3

Q1.(b) ~~Surge~~ {heave, soar, tide, ascend} \cup {surging, surged}
 $=$ {surge, heave, soar, tide, ascend, surging, surged}

~~Q1.(c)~~ Q1.(c) Ans

\rightarrow helps in normalisation
 \rightarrow semantic analysis

Cons \rightarrow may change word meaning due to stemming

180810062 (Radhika Patwari)

Q1. (d) omicron OR vaccine =
$$\begin{array}{r} 30000 \\ 2\ 33\ 312 \\ \hline 3,23,312 \end{array}$$

(A)

(B) India OR lockdown =
$$\begin{array}{r} 400,530 \\ 161\ 658 \\ \hline 5,62,188 \end{array}$$

(C) delta OR variant =
$$\begin{array}{r} 107313 \\ 87009 \\ \hline 1,94,922 \end{array}$$

(A) and (B) =
$$\begin{array}{r} 323312 \\ \hline 562188 \end{array} = (D)$$

(B) and (C) =
$$1,94,922$$

(A) and (C) =
$$1,94,922$$

(D) and (C) =
$$1,94,922$$

First do, (A) and (B), then (D) and (C)