Practical - 7

Write a program to construct different types of shingles for given document.

Installation of required packages before executing program:-

```
install.packages("tm")
require("tm")
install.packages("devtools")
```

Source Code :-

```
readinteger <- function()
{
        n <- readline(prompt="Enter value of k-1: ")
        k<-as.integer(n)
        u1 <- readLines("c:/msc/r-corpus/File1.txt")
        Shingle<-0
        i <-0
        while(i<nchar(u1)-k+1){
        Shingle[i] <- substr(u1, start=i, stop=i+k)
        print(Shingle[i])
        i=i+1
  }
}
if(interactive()) readinteger()
```

OutPut:-

```
> if(interactive()) readinteger()
Enter value of k-1: 2
character(0)
[1] "thi"
[1] "his"
[1] "is "
[1] "s i"
[1] " is"
[1] "is "
[1] "s a"
[1] " a "
[1] "a t"
[1] " te"
[1] "tex"
[1] "ext"
[1] "xt."
[1] "t. "
[1] ". i"
[1] " it"
[1] "it "
[1] "t i"
[1] " is"
[1] "is "
[1] "s s"
[1] " sh"
[1] "sho"
[1] "hor"
[1] "ort"
[1] "rt,"
[1] "t, "
[1] ", a"
[1] " an"
[1] "and"
[1] "nd "
[1] "d i"
[1] " it"
[1] "it "
[1] "t i"
[1] " is"
[1] "is "
[1] "s o"
[1] " on"
[1] "one"
[1] "ne "
[1] "e l"
[1] " li"
[1] "lin"
[1] "ine"
[1] "ne."
[1] "e. "
[1] ". i"
[1] " it"
[1] "it "
[1] "t i"
[1] " is"
[1] "is "
[1] "s n"
[1] " no"
[1] "not"
[1] "ot "
[1] "t a"
```

```
[1] " ab"
[1] "abo"
[1] "bou"
[1] "out"
[1] "ut "
[1] "t w"
[1] " wh"
[1] "wha"
[1] "hal"
[1] "ale"
[1] "les"
[1] "es "
[1] "s o"
[1] " or"
[1] "or "
[1] "r r"
[1] " ro"
[1] "rom"
[1] "oma"
[1] "man"
[1] "anc"
[1] "nce"
[1] "ce."
[1] "e. "
[1] ". t"
[1] " th"
[1] "thi"
[1] "his"
[1] "is "
[1] "s i"
[1] " is"
[1] "is "
[1] "s o"
[1] " on"
[1] "onl"
[1] "nly"
[1] "ly "
[1] "y a"
[1] " a "
[1] "a t"
[1] " te"
[1] "tex"
[1] "ext"
[1] "xt."

>
```

OutPut:-

```
> if(interactive()) readinteger()
Enter value of k-1: 3
character(0)
[1] "this"
[1] "his "
[1] "is i"
[1] "s is"
[1] " is "
[1] "is a"
[1] "s a "
[1] " a t"
[1] "a te"
```

```
[1] " tex"
[1] "text"
[1] "ext."
[1] "xt. "
[1] "t. i"
[1] ". it"
[1] " it "
[1] "it i"
[1] "t is"
[1] " is "
[1] "is s"
[1] "s sh"
[1] " sho"
[1] "shor"
[1] "hort"
[1] "ort,"
[1] "rt, "
[1] "t, a"
[1] ", an"
[1] " and"
[1] "and "
[1] "nd i"
[1] "d it"
[1] " it "
[1] "it i"
[1] "t is"
[1] " is "
[1] "is o"
[1] "s on"
[1] " one"
[1] "one "
[1] "ne l"
[1] "e li"
[1] " lin"
[1] "line"
[1] "ine."
[1] "ne. "
[1] "e. i"
[1] ". it"
[1] " it "
[1] "it i"
[1] "t is"
[1] " is "
[1] "is n"
[1] "s no"
[1] " not"
[1] "not "
[1] "ot a"
[1] "t ab"
[1] " abo"
[1] "abou"
[1] "bout"
[1] "out "
[1] "ut w"
[1] "t wh"
[1] " wha"
```

```
[1] "whal"
[1] "hale"
[1] "ales"
[1] "les "
[1] "es o"
[1] "s or"
[1] " or "
[1] "or r"
[1] "r ro"
[1] " rom"
[1] "roma"
[1] "oman"
[1] "manc"
[1] "ance"
[1] "nce."
[1] "ce. "
[1] "e. t"
[1] ". th"
[1] " thi"
[1] "this"
[1] "his "
[1] "is i"
[1] "s is"
[1] " is "
[1] "is o"
[1] "s on"
[1] " onl"
[1] "only"
[1] "nly "
[1] "ly a"
[1] "y a "
[1] " a t"
[1] "a te"
[1] " tex"
[1] "text"
[1] "ext."
>
```

OutPut:-

```
> if(interactive()) readinteger()
Enter value of k-1: 4
character(0)
[1] "this "
[1] "his i"
[1] "is is"
[1] "s is "
[1] " is a"
[1] "is a "
[1] "s a t"
[1] " a te"
[1] "a tex"
[1] " text"
[1] "text."
[1] "ext. "
```

```
[1] "xt. i"
[1] "t. it"
[1] ". it "
[1] " it i"
[1] "it is"
[1] "t is "
[1] " is s"
[1] "is sh"
[1] "s sho"
[1] " shor"
[1] "short"
[1] "hort,"
[1] "ort, "
[1] "rt, a"
[1] "t, an"
[1] ", and"
[1] " and "
[1] "and i"
[1] "nd it"
[1] "d it "
[1] " it i"
[1] "it is"
[1] "t is "
[1] " is o"
[1] "is on"
[1] "s one"
[1] " one "
[1] "one l"
[1] "ne li"
[1] "e lin"
[1] " line"
[1] "line."
[1] "ine. "
[1] "ne. i"
[1] "e. it"
[1] ". it "
[1] " it i"
[1] "it is"
[1] "t is "
[1] " is n"
[1] "is no"
[1] "s not"
[1] " not "
[1] "not a"
[1] "ot ab"
[1] "t abo"
[1] " abou"
[1] "about"
[1] "bout "
[1] "out w"
[1] "ut wh"
[1] "t wha"
[1] " whal"
[1] "whale"
[1] "hales"
[1] "ales "
[1] "les o"
[1] "es or"
[1] "s or "
[1] " or r"
[1] "or ro"
[1] "r rom"
[1] " roma"
```

```
[1] "roman"
[1] "omanc"
[1] "mance"
[1] "ance."
[1] "nce. "
[1] "ce. t"
[1] "e. th"
[1] ". thi"
[1] " this"
[1] "this "
[1] "his i"
[1] "is is"
[1] "s is "
[1] " is o"
[1] "is on"
[1] "s onl"
[1] " only"
[1] "only "
[1] "nly a"
[1] "ly a "
[1] "y a t"
[1] " a te"
[1] "a tex"
[1] " text"
[1] "text."

>
```