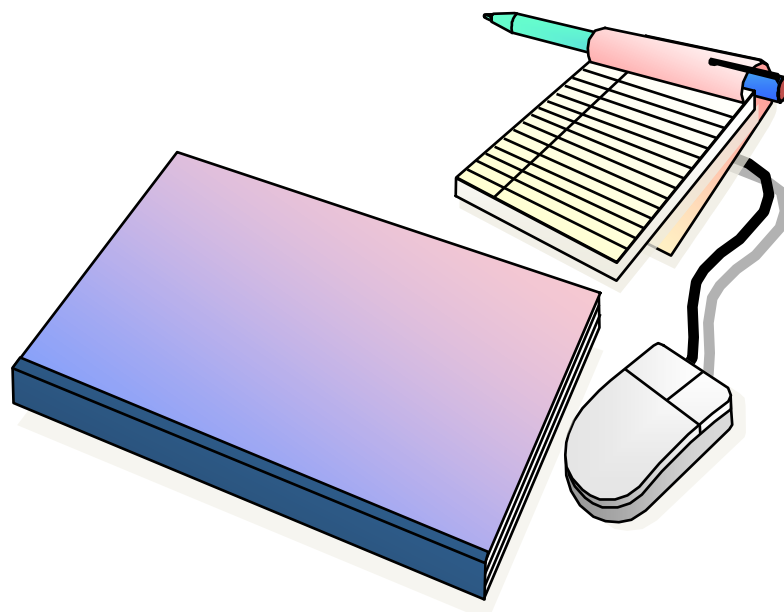


Top20



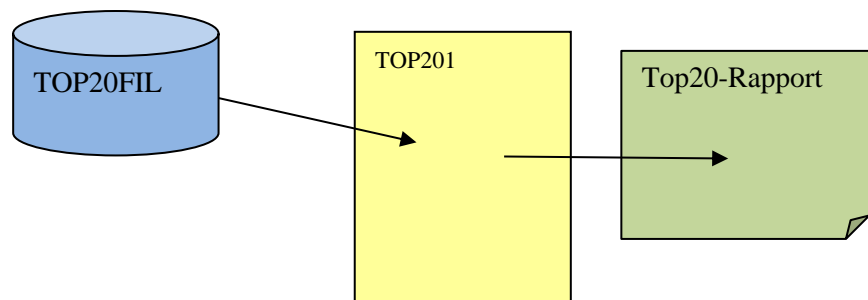
Innehållsförteckning

Top20 – ord i artikel.....	3
Uppgift 1 – Top20fil	3
Uppgift 2 – Top201 – rapportprogram	4
Uppgift 2 – Top202 – rapportprogram	6

Top20 – ord i artikel

I denna övning kommer du att skriva ett program som läser en fil som innehåller en tidningsartikel. Programmet kommer att sammanställa statistik över de tjugo mesta använda orden i artikeln och skriva en rapport.

Uppgift 1 – Top20fil



- Innehållet i filen Dokument finns på www.cobolskolan.se/CA01 under rubriken Top20 – Filer – Top20Fil1 och Top20Fil2 och har följande utseende:

Fält	Typ	Längd	Innehåll
Textrad	X	80	Löpande text från tidningsartikel

- Ladda ned innehållet till denna fil till en medlem i biblioteket `userid.CA01.COBDATA` och ge medlemmen namnet `TOP20FI1` resp `TOP20FI2`.
- Du kan även kopiera någon tidningsartikel från nätet och placera i motsvarande medlem.

Uppgiften är avslutad.

Uppgift 2 – Top201 – rapportprogram

- Du skall skriva ett program med namnet TOP201. Programmet skall läsa samtliga rader från filen TOP20FIL.
- Programmet skall extrahera samtliga ord i varje rad i filen, och placera dessa i en tabell.
- Tabellraderna skall ha plats för ett ord och en räknare som anger antal förekomster av ordet. Antag att varje ord i texten som finns i filen TOP20FIL är maximerat till 30 tecken.
- Du kan maximera antalet rader i denna tabell till 1000. Detta förutsätter då att det inte finns fler än 1000 olika ord i en artikel.
- För varje läst rad så skall tabellen kompletteras, om ordet inte finns sedan tidigare i tabellen, eller så skall antalet justeras.
- Innan tabellen fylls med ord, så skall specialtecken som t.ex. kommatecken, kolon, semikolon etc. elimineras från raden.
- Innehållet av tabellen kan påminna om denna:

Ordtext	Ordantal
Detta	1
Är	2
Ett	1
Antal	2
Ord	1

- När samtliga rader är lästa så skall programmet fylla en ny tabell, som skall ha plats för de 20 mest frekvent använda orden, sorterad på antal förekomster.
- Du skall inte använda sorteringsprogrammet, utan sorteringen skall göras med COBOL-kod.
- Programmet skall sedan skriva ut innehållet i denna Top20-tabell i en rapport.
- Rapporten kan påminna om denna:

```
*** Topp 20 ord i artikel ***
Pos  Antal  Dokument-ord  Procent
(01)  28    att           0,08
(02)  26    som           0,08
(03)  18    och           0,05
(04)  18    det           0,05
(05)  18    är            0,05
(06)  18    i             0,05
(07)  16    har           0,05
(08)  16    inte          0,05
(09)  14    Det           0,04
(10)  12    var           0,04
(11)  11    I             0,03
(12)  11    en            0,03
(13)  10    för           0,03
(14)  10    på            0,03
(15)   9    regionerna    0,03
(16)   9    av            0,03
(17)   9    om            0,03
(18)   7    testning     0,02
(19)   7    regeringen   0,02
(20)   7    vad           0,02
```

Det finns totalt 336 olika ord i denna artikel.

*** S L U T på rapport ***

Kompilera och korrigera tills du har en felfri kompilering.

Exekvera programmet och kontrollera resultatet.

Exekvera programmet med olika innehåll i filen och jämför resultaten.

Uppgiften är avslutat.

Uppgift 2 – Top202 – rapportprogram

- Du skall skriva ett program med namnet TOP202. Används programmet TOP201 som mall.
- Programmet skall nu använda COBOL verbet SORT för att sortera fram de 20 mest använda orden i filen TOP20FIL.

- Kompilera och korrigerar tills du har en felfri kompilering.
- Exekvera programmet och kontrollera resultatet.
- Jämför med resultatet från programmet TOP201.
- Exekvera programmet med olika innehåll i filen och jämför resultaten.

Uppgiften är avslutat.