AWK Programlama Dili

Nedir?

- AWK, Alfred **A**ho, Peter **W**einberger ve Brain **K**ernighan tarafından 1978 yılında geliştirilmiş ve bir çok UNIX sürümünde (özellikle System V, version 3.1'den sonrakilerde) kendisine yer bulmuş bir programlama dilidir.
- Dilin açık kaynak kodlu GNU versiyonu 1986 yılında Richard Stallman'ın tavsiyesi ile Paul Rubin ve Jay Fenlason tarafından yazılmıştır.
- GAWK'ın yazılması aşamasında ve sonrasında Brain Kernighan projeye fikirleri ve müdahaleleri ile destek olmuştur.
- Zaman içerisinde MS Windows ve Mac üzerinde çalışan versiyonları ve AWK kodunu C, C++ ve Perl kodlarına dönüştüren araçlar programcılar tarafından geliştirilmiştir.

Nedir? (devam)

- AWK genel olarak metinleri kolayca işlemenize ve raporlar oluşturabilmenize olanak sağlayan data-driven bir programlama dilidir. Awk ile,
 - Küçük kişisel veritabanlarınızı yönetebilir, formatlı raporlar oluşturabilir,
 - Veriler üzerinde aritmetik ve string operasyonlarını gerçekleştirebilir,
 - Genel programlama yapılarını kullanabilir (kıyas operatörleri, döngüler v.s.),
 - Kabukta bir komutun çıktısını on-the-fly işleyebilir ve başka bir komutun girdisi olacak şekilde formatlayabilir,
 - Doğru yerde kullandığınızda C, Pascal gibi dillerle -nispeten- zor olan operasyonları tek satırda gerçekleştirebilirsiniz.

Nasıl Çalışır?

- AWK yorumlanan bir dildir, yani AWK kodları tek başına çalıştırılabilir bir uygulamaya dönüştürülemez (!), AWK scriptlerinin bir sistemde çalışabilmesi için o sistemde AWK yorumlayıcısının bulunması gerekir.
- AWK'nin çalışma şekli genel olarak şöyledir:
 - Bir dosyayı ya da kendisine yönlendirilen bir girdiyi (örneğin bir komut çıktısını) satır satır okur,
 - Her bir girdi satırını alanlara ayırır,
 - İstediğimiz durumlarla eşleşen satırlar üzerinde istediğimiz işlemleri gerçekleştirir.

Nasıl Çalışır? (devam)

• İster komut satırından yazdığımız AWK programı ile:

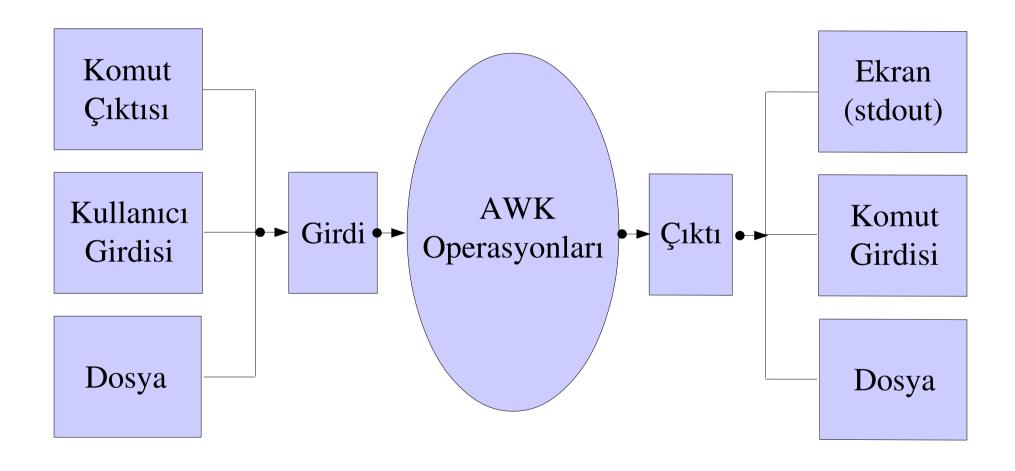
```
]$ awk [options] 'program_kodu' dosyalar
(awk -F : '/evreniz/ {if($5>1000)print $1 $2}' kullanici_dosyalari)
```

• İstersek de daha önce bir dosyaya yazmış olduğumuz AWK programı ile çalıştırabiliriz:

```
]$ awk [options] -f script.awk dosyalar (awk -f rutin_control.awk /etc/passwd)
```

• *info awk* ya da *man awk* ile seçenekler ve kullanım ile ilgili ayrıntılı bilgi alınabilir.

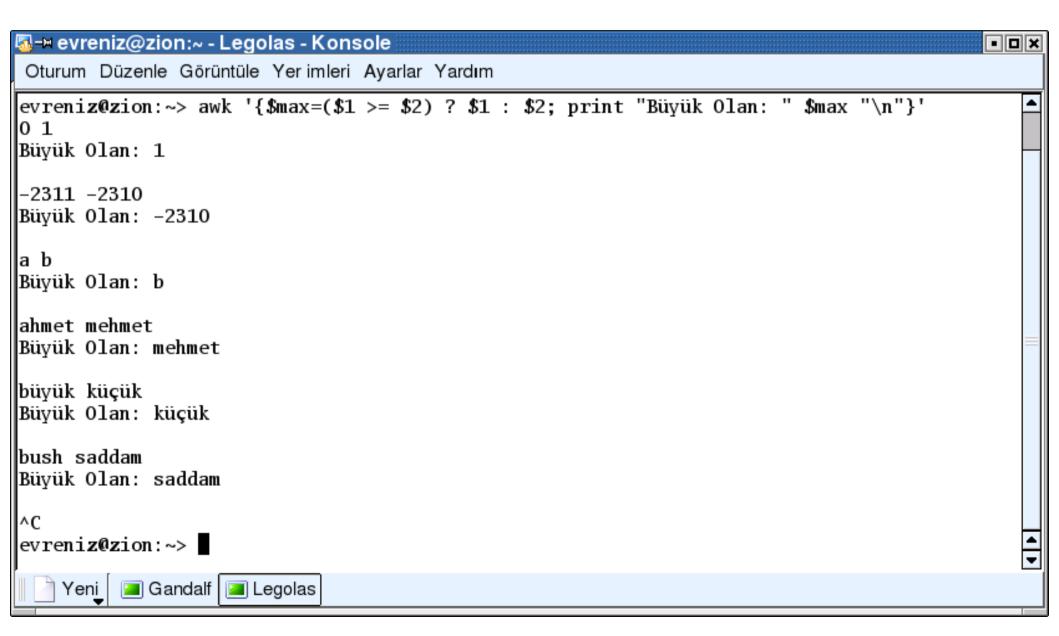
Nasıl Çalışır? (devam)



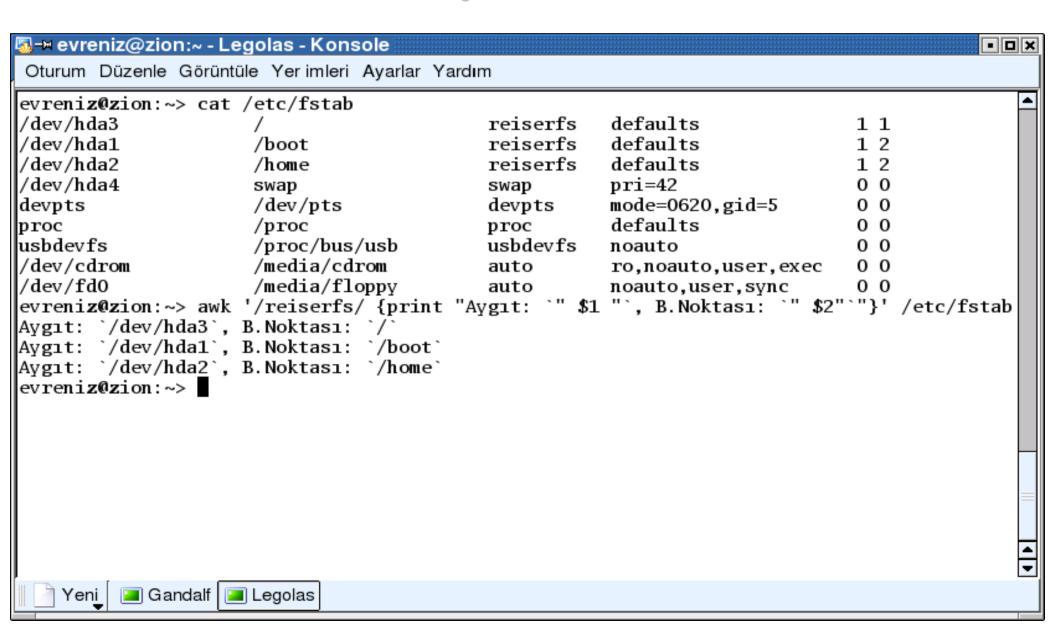
Komut Çıktısı (örnek)

```
■-M evreniz@zion:~
                                                                                       - 0 ×
Dosya Düzenle Görüntüle Terminal Git Yardım
evreniz@zion:~> date
Paz Ara 14 18:58:54 EET 2003
evreniz@zion:~> date | gawk '{print $6 " Yılındayız.."}'
2003 Yılındayız..
evreniz@zion:\sim ls -l | sort --key=5 -ir | gawk '{if(i==5)exit; print $8 ", " $5; i++;}'
Quake.3.Arena.Linux.iso, 491132928
wps.zip, 33851434
anjuta-1.0.2.tar.gz, 6238726
100_0017.MOV, 2517216
COMU-DatabaseBackUp-15.01.03, 2279763
evreniz@zion:~>
```

Kullanıcı Girdisi (örnek)



$Dosyadan \ (\"{o}\textit{rnek})$



Pattern ve Action

• AWK betikleri, kalıp (pattern) ve eylem (action)'lerden meydana gelirler.

```
]$ awk 'pattern {action} pattern {action} pattern {action} ... '
```

- Eğer eylemin başında bir kalıp yoksa eylem tüm kayıtlar için gerçekleştirilir.
- Eğer kalıbın ardından bir eylem yoksa tüm uygun kayıtlar hiç bir işleme tabi tutulmadan ekrana yazılır.
- Kalıplar düzenli ifadelerden, mantıksal kıyas ifadelerinden ya da virgül ile birbirini takip eden kalıp dizilerinden oluşabilir. Ayrıca BEGIN ve END adında iki özel kalıp vardır.

Pattern ve Action

```
BEGIN {eylem; eylem; ...}
BEGIN {eylem; eylem; ...}
                                    /kelime/ {eylem; eylem; ...}
                                     $1!~/regexp/{eylem; ...}
bütün kayıtlar için
ana hesaplama
                                     1 = 25  {eylem; eylem; ...}
ve operasyonlar
                                     $1 > $5 {eylem; eylem; ...}
                                     END {eylem; eylem; ...}
END {eylem; eylem; ...}
```

BEGIN ve END

 BEGIN ve END özel tanımlanmış iki kalıptır. BEGIN kalıbından sonra gelen eylem ilk girdi satırının işlenmesinden hemen önce, END kalıbından sonra tanımlanmış eylem ise son girdi satırının işlenmesinden hemen sonra çalışır.

```
BEGIN {
FS=":";
counter=0;
bin:x:1:1:bin:/bin/bash
daemon:x:2:2:Daemon:/sbin:/bin/bash
...
/root/ {
counter++;
} awk -f script.awk /etc/passwd
root kelimesi geçen 1 satıra rastlandı
]$

END {
print "root kelimesi geçen " counter " satıra rastlandı"
```

Built-In Değişkenler

 AWK çalışma zamanında kimi bilgilere kolaylıkla erişim için bir takım özel değişkenler ile beraber çalışır.

FILENAME Yürürlükteki dosyanın adı

FS Girdi alan ayracı (öntanımlı olarakSPACE)

RS Girdi kayıt ayracı (öntanımlı olarak \n)

NF Yürürlükteki kayıdın alan sayısı

NR Yürürlükteki kayıdın satır numarası

OFS Çıktı alan ayracı (öntanımlı olarak SPACE)

ORS Çıktı kayıt ayracı (öntanımlı olarak \n)

\$0 Girdi kaydının tümü

\$n Girdi kaydının n. alanı.

Operatörler

- Mantıksal Kıyas Operatörleri
 - && (Mantiksal VE)
 - || (Mantıksal VEYA)
 - ! (Mantıksal DEĞİL)
- Aritmetik Operatörler
 - *, /, +, (Çarpma, Bölme, Toplama, Çıkarma)
 - % (Modül)
 - ^ (Üs)

Operatörler (devam)

- İlişkisel Operatörler
 - < (Küçük)
 - > (Büyük)
 - = (Eşit)
 - != (Eşit Değil)
 - <=, >= (Küçük Eşit, Büyük Eşit)
 - ~ (regexp ile Uygun [x ~ /regexp/])
 - !~ (regexp ile Uygun Değil [x !~ /regexp/])

Kontrol İfadeleri

• if - else

```
if ( durum ) { yapılacaklar } else { yapılacaklar }
```

```
if ( x % 2 == 0 )
    print "x çift sayıdır"
else
    print "x tek sayıdır"
```

```
]$ awk 'if( $1 % 2 == 0) print $1 " çift sayıdır"; else print $1 " tek sayıdır"
```

Kontrol İfadeleri (devam)

while

```
while ( durum ) { yapılacaklar }
```

```
i=1
while ( i <= 3 ){
    print i
    i++
}</pre>
```

```
]$ awk 'BEGIN {i=3; while (i \le 3) {print i; i++}}'
```

Kontrol İfadeleri (devam)

• for

```
for ( başlangıç_durumu ; koşul ; artış ) { yapılacaklar }
```

```
for(i = 0; i <= 100; i += 5){
    print i
}
```

```
]$ awk 'BEGIN {for(i = 0; i \le 100; i += 5) print i}'
```

Kontrol İfadeleri (devam)

Ayrıca AWK,

- continue
- break
- next
- exit

akış kontrol ifadelerini de destekler.

Built-In Fonksiyonlar

- AWK'nin beraberinde gelen bir çok önceden tanımlanmış fonksiyon mevcuttur. Bu hazır fonksiyonlar da programcıya kolaylık sağlar..
 - Numerik Fonksiyonlar: int(x), sqrt(x), exp(x), log(x), sin(x), cos(x), atan2(x, y), rand(), srand(x), ...
 - Karakter Dizileri ile İş Yapan Fonksiyonlar: index(in, find), length(string), match(string, regexp), split(string, array, fieldsep), sub(regexp, replacement, target), substr(string, start, length), tolower(string), toupper(string), ...
 - Girdi/Çıktı Fonksiyonları: system(command), ...

Fonksiyon Tanımlama

• AWK'de -elbette- fonksiyon tanımlamak ve bunu çağırmak mümkündür, fonksiyonlar şu şekilde tanımlanır ve çağırılır:

```
function fonksiyon_adı (parametre1, parametre2, ...){

fonksiyon gövdesi
...
return değer
}
```

```
fonksiyon_adı (parametre1, "para metre 2", parametre3, ...)
```

Örnekler

A. Murat Eren <meren@comu.edu.tr>, http://zion.comu.edu.tr/~evreniz

Örnek 1

⊠ -⊭ evren	iz@zion:~ - G	andalf - Konsole			- D X
Oturum Düzenle Görüntüle Yer imleri Ayarlar Yardım					
root	1749 2.8	5.2 64636 13384 ?	S	18:41	2:08 /usr/X11R6/bin/X vt7 -auth /var/lib/xdm/authdir/authfiles/A:0-hF
root	1750 0.0		S	18:41	0:00 -:0
evreniz	1780 0.0		S	18:41	0:00 /bin/sh /usr/X11R6/bin/kde
evreniz	1796 0.0		S	18:41	0:00 /bin/sh /usr/X11R6/bin/kde 0:00 /opt/gnome/bin/medusa-idled Ornek bir ps -aux
evreniz	1819 0.0		S	18:41	0:00 kdeinit: Kunning
evreniz	1822 0.0		S	18:41	0:00 kdeinit: dcopservernosid 0:00 kdeinit: klauncher C1Kt1S1
evreniz	1825 0.0		S	18:41	J
evreniz	1828 0.0		S	18:41	0:00 kdeinit: kded
evreniz	1834 0.0		S	18:42	0:01 /opt/kde3/bin/artsd -F 10 -S 4096 -s 5 -m artsmessage -l 3 -f
evreniz	1845 0.0		S	18:42	0:00 kdeinit: knotify
evreniz	1846 0.0		S	18:42	0:00 kwrapper ksmserver
evreniz	1848 0.0		S	18:42	0:00 kdeinit: ksmserver
evreniz	1849 0.0		S	18:42	0:03 kdeinit: kwin -session 11c1ff6111000106389896500000101660000_107
evreniz	1851 0.0		S	18:42	0:00 kdeinit: kwrited
evreniz	1853 0.0 1856 0.0		S	18:42 18:42	0:03 kdeinit: kdesktop 0:02 kdeinit: kicker
evreniz evreniz	1863 0.0		S S	18:42	0:02 kdeinit: kicker 0:00 susewatcher -caption SuSE Update Checker -icon kinternet.png -mi
evreniz	1864 0.0		S	18:42	0:00 susewatcher -caption suse opdate checker -icon kinternet.phg -mil 0:00 kgpg
wwwrun	1870 0.0		S	18:42	0:00 kgpg 0:00 /usr/sbin/httpd -f /etc/httpd/httpd.conf
evreniz	1875 0.0		S	18:42	0:00 /dsi/sbin/httpu =1 /etc/httpu/httpu/com1 0:00 kdeinit: kmix -session 11c1ff6111000106389897300000101660008_107
evreniz	1877 0.0		S	18:42	0:00 kget -session 1053fc353c000106680735800000017810034_1070957561_2
evreniz	1883 0.0		S	18:42	0:02 sim -session 105f2b1091000106853402200000017810034_1070957561_12
evreniz	1884 0.0		S	18:42	0:02 kdeinit: konquerorpreload
evreniz	1886 0.0		S	18:42	0:00 kalarındlogin
evreniz		5.1 27760 13092 ?	S	18:42	0:03 extensionproxyconfigfile taskbar_panelextensionrccallbacki
evreniz		19.3 120236 49536 ?	S	18:44	1:28 /opt/OpenOffice.org/program/soffice.bin /home/evreniz/Desktop/AW
evreniz		19.3 120236 49536 ?	S	18:44	0:00 /opt/OpenOffice.org/program/soffice.bin /home/evreniz/Desktop/AW
evreniz		19.3 120236 49536 ?	S	18:44	0:00 /opt/OpenOffice.org/program/soffice.bin /home/evreniz/Desktop/AW
evreniz		19.3 120236 49536 ?	S	18:44	0:00 /opt/OpenOffice.org/program/soffice.bin /home/evreniz/Desktop/AW
evreniz	1938 0.0	19.3 120236 49536 ?	S	18:44	0:00 /opt/OpenOffice.org/program/soffice.bin /home/evreniz/Desktop/AW
evreniz	1939 0.0	19.3 120236 49536 ?	S	18:44	0:00 /opt/OpenOffice.org/program/soffice.bin /home/evreniz/Desktop/AW
evreniz	1960 0.1	4.9 30860 12748 ?	S	18:47	0:05 kdeinit: konsole
evreniz	1961 0.0	0.6 4768 1664 pts/1	S	18:47	0:00 /bin/bash
evreniz		10.5 44464 27052 ?	S	19:19	0:21 kdeinit: konquerorsilent
evreniz		2.1 24656 5416 ?	S	19:19	0:00 kdeinit: kio_file file /tmp/ksocket-evreniz/klauncherux6Jwb.slav
evreniz	2139 0.0	4.4 27256 11508 ?	S	19:19	0:00 kdeinit: kio_uiserver
evreniz	2227 0.0		S	19:52	0:00 kdeinit: kio_http http /tmp/ksocket-evreniz/klauncherux6Jwb.slav
evreniz	2229 0.0		S	19:52	0:00 kdeinit: kio_http http /tmp/ksocket-evreniz/klauncherux6Jwb.slav_
evreniz		2.5 22336 6456 ?	S	19:53	0:00 kdeinit: kio_http http /tmp/ksocket-evreniz/klauncherux6Jwb.slav
evreniz		0.2 2512 656 pts/1	R	19:55	0:00 ps aux
evreniz@zion:~>					
Yeni Gandalf					

Örnek 1 (devam)

`ps -aux` çıktısı, sistemde çalışan süreçlerin ve o süreçlere ait kimin tarafından ne zaman çalıştırıldığı, ne kadar işlemci ve bellek kullandığı gibi kimi bilgileri verir. Aşağıdaki tek satırlık komut ile, root kullanıcısı dışındaki kullanıcıların çalıştırdığı süreçleri öldürülebilir:

```
]# ps aux | awk '$1 !~ /root/ {if(NR > 1) print $2}' | xargs kill -9
```

Aynı çıktıdan, en çok işlemci kullanan süreç de kolaylıkla elde edilebilir:

```
| Sps aux | awk |
| BEGIN {max=0}
| {if($3>max) {user=$1; max=$3; surec=$2; cmd=$11;}}
| END {print "\n"cmd " komutu ile " user " tarafından çalıştırılmış olan " surec " id`li süreç, işlemcinin yüzde " max " kadarlık kısmını kullanmaktadır. \n"}
| **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company **Total Company
```

Bu komutun bilgisayarımdaki çıktısı şu oldu:

/usr/X11R6/bin/X komutu ile root tarafından çalıştırılmış olan 1749 id`li süreç, işlemcinin yüzde 2.6 kadarlık kısmını kullanmaktadır.

Örnek 2

Aşağıdaki AWK kodu bir metin içerisinde geçen tüm kelimelerin geçiş sıklıklarını gösterir:

```
#!/usr/bin/awk -f
 for (i=1; i \le NF; i++)
  kelime=tolower($i)
  if(index(kelime, ",") || index(kelime, ".")){
     kelime=substr(kelime, 0, (length(kelime)-1))
 freq[kelime]++
END {
 for (kelime in freq)
  if(freq[kelime]>4)
     printf "\"%s\" %d\n", kelime, freq[kelime]
```

Bu betik freq.awk adı ile kaydedilip herhangi bir dosya için,

]\$./freq.awk makale

biçiminde kullanabilir.

Örnek 2 (devam)

"lisans" 6

```
evreniz@zion:~> ./freq.awk makale.1 | sort -key=2 -nr
                                                                 evreniz@zion:~> ./freq.awk makale.2 | sort --key=2 -nr
"ve" 62
                                                                 "bir" 82
"bir" 49
                                                                 "ve" 57
"ile" 26
                                                                 "yazılım" 24
"bu" 19
                                                                 "ile" 24
"daha" 14
                                                                 "kaynak" 23
"da" 14
                                                                 "özgür" 21
"Internet" 12
                                                                 "bu" 21
"de" 11
                                                                 "cok" 17
                                                                 "linux" 17
"toplumsal" 10
"en" 10
                                                                 "için" 17
"sekilde" 9
                                                                 "da" 17
"bilgisayar" 9
                                                                 "olarak" 13
"için" 8
                                                                 "gnu" 10
"azalma" 8
                                                                 "acık" 10
"çok" 7
                                                                 "gibi" 9
"haline" 7
                                                                 "tarafından" 8
"olarak" 6
                                                                 "sahip" 8
"olan" 6
                                                                 "olan" 7
"kadar" 6
                                                                 "de" 7
"arama" 6
                                                                 "bilgisayar" 7
                                                                 "üzerinde" 6
"ya" 5
"son" 5
                                                                 "yeni" 6
                                                                 "sistem" 6
```

Örnek 3

evreniz@zion:~> cat ogrenciler.BM113 Öğrenci Adı:Numarası:Ödev 1:Ödev 2:Ödev 3:Vize Notu Sanem Çelik:030401016:+:+:74 Gökçen Eraslan:030401008:+:-:+:92 Atacan Tufan:030401003:-:-:+:72 Kerem Ilica:030401010:+:+:-:66 İşbaran Akçayır:030401004:+:+:+:55 Aysun Özdemir:030401017:-:-:+:57 Selda Yurtdaş:030401005:+:+:54 Eray Kaplan:020401038:-:+:+:49 Fatih Şen:010401023:-:-:80 Aykut Aksoy:030401021:+:+:+:86 Rafet Cambaz:030401010:+:+:+:80 Şule Gök:030401009:-:+:+:55 Mevlüt Uşan:0304010022:+:-:-:69

Elimizde alanları birbirinden `:` ile, kayıtları da birbirinden \n` ile ayrılmış küçük bir veritabanı olduğunu farzedelim (benim var:)). Bu veritabanındaki bilgileri zaman zaman html'ye dönüştürüp kişilere duyurmak istiyor olalım.

Örnek 3 (devam)

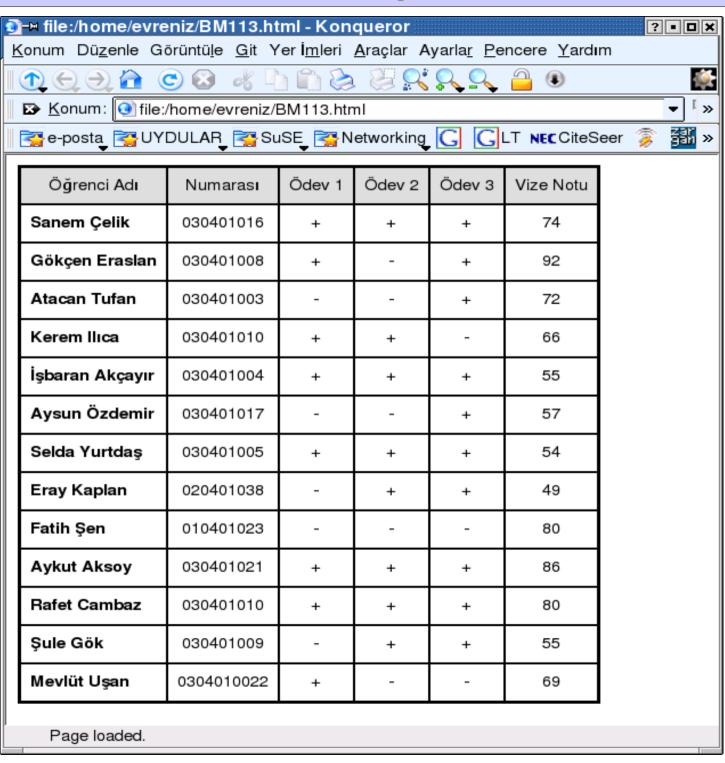
```
evreniz@zion:~> cat convert2html.awk
#!/usr/bin/gawk -f
BEGIN {
    FS =":";
    print "<TABLE cellpadding=\"8pt\" BORDER=\"2pt\" CELLSPACING=\"0pt\""
    print "bgcolor=\"\#ffffff\" bordercolor=\"\#000000\">\n";
(NR==1){
    print "<TR bgcolor=\"\#dfdfdf\">\n"
    for(i=1; i \le NF; i++)print "<TD><center>"<math>i"</center></TD>\n";
    print "</TR>\n"
(NF>0 && NR>1){
    printf "<TR>\n"
    for(i=1; i \le NF; i++)
         if (i==1) print "<TD align=left>"$i"</TD>\n";
         else print "<TD align=center>"$i"</TD>\n";
    print "</TR>\n";
END {
    print "</TABLE>\n";
```

Bu minik scripti komut satırından ogrenciler.BM113 dosyasını formatlaması için,

]\$./convert2html.awk ogrenciler.BM113

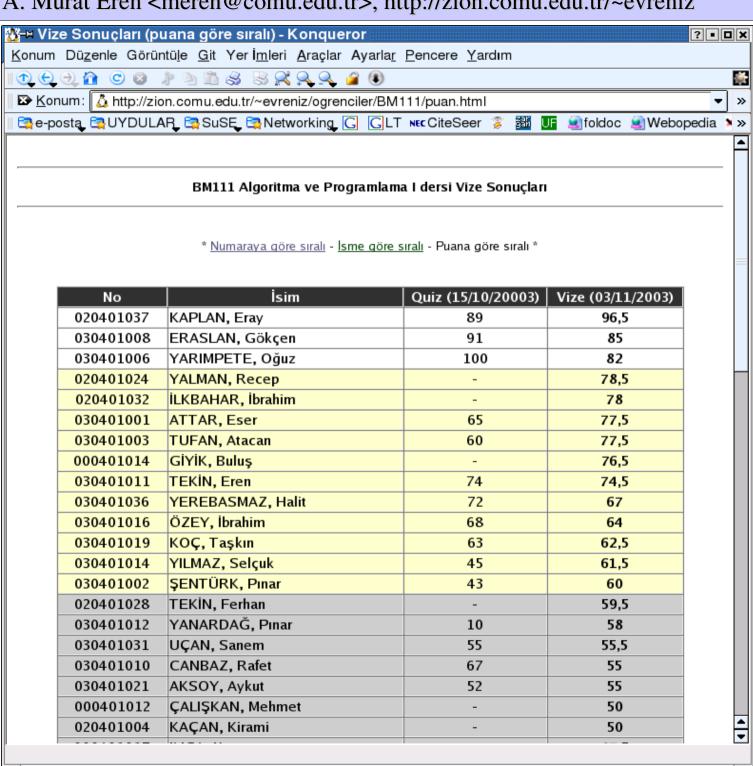
komutu ile cağırabiliriz. Ekrana göndereceği HTML kodlarını `>` ile BM113.html dosyasına yönlendirebiliriz...

A. Murat Eren <meren@comu.edu.tr>, http://zion.comu.edu.tr/~evreniz



Örnek 3 (sonuç)

A. Murat Eren <meren@comu.edu.tr>, http://zion.comu.edu.tr/~evreniz



Ornek 3 (güzel)

Elbette biraz daha uğraşarak daha güzel ve kullanışlı sonuçlar da almak mümkündür:)

Neden AWK?

- AWK Perl, C gibi dillere nazaran daha kolaydır.
- Sözdizimi daha anlaşılırdır, bu da yeni başlayanlar için büyük bir avantaj teşkil eder.
- Ufak tefek işlerinizi halledebilecek kadar AWK, zaten biliyor olabilirsiniz, belki tek ihtiyacınız olan bir kaç deneme yapmak.
- Ayrıntılarla boğuşmadan algoritmanızı implemente edebilmenin mutluluğunu yaşarsınız.
- AWK ile ufacık programlar yaratılabilir, bu sayede çok hızlı üretilmesi ve sonuç alınması gereken küçük işler için bire birdir.
- Temiz bir Perl ya da C kodu temiz olmayan bir AWK kodundan çok daha iyidir fakat, temiz olmayan bir Perl kodunun (ya da C kodunun) yanına hiç bir şey yaklaşamaz :)

Daha Fazlası?

- AWK için faydalı bir FAQ: http://www.faqs.org/faqs/computer-lang/awk/faq/
- GAWK Kullanım Klavuzu http://www.gnu.org/manual/gawk-3.0.3/gawk.html
- Effective GAWK Programming http://zion.comu.edu.tr/~evreniz/gawkbook/egp2002.pdf
- AWK Haber Grubu (25.800 mesaj) http://groups.google.com/groups?group=comp.lang.awk

Teşekkürler

A. Murat Eren meren~comu.edu.tr http://zion.comu.edu.tr/~evreniz