

```
# Functions to read in the CSV table that contains all of the raw data.
# Before running these functions, make sure the file "all-data.csv" is
# in the local directory.
# Also, within the R environment, change the working directory to the directory
# that contains the data file using the toolbar menu:
# File -> Change dir
#
```

```
# Read the data from the csv file.
processors <- read.csv("all-data.csv")
```

```
#####
```

```
#
# This function returns the data from the desired column.
# Example: clock<-get_column("Fp2000","Processor.Clock..MHz.")
```

```
get_column <- function(x,y) {
# x = string with the name of the desired benchmark
# y = desired column
#
# Find the indices of all rows that have an entry for the
# indicated benchmark
benchmark <- paste(paste("Spec",x,sep=""),"..average.base.",
                    sep="")
ix <- !is.na(processors[,benchmark])
return(processors[ix,y])
}
```

```
#####
```

```
#####
# This function extracts the interesting data columns for the given benchmark
# program and returns a dataframe with these columns.
```

```
extract_data <- function(benchmark) {
temp <- paste(paste("Spec",benchmark,sep=""),"..average.base.", sep="")

# perf = the performance reported in the database
perf <- get_column(benchmark,temp)

#nperf = performance normalized to the overall range
max_perf <- max(perf)
min_perf <- min(perf)
range <- max_perf - min_perf
nperf <- 100 * (perf - min_perf) / range

clock <- get_column(benchmark,"Processor.Clock..MHz.")
threads <- get_column(benchmark,"Threads.core")
cores <- get_column(benchmark,"Cores")
TDP <- get_column(benchmark,"TDP")
transistors <- get_column(benchmark,"Transistors..millions.")
dieSize <- get_column(benchmark,"Die.size..mm.2.")
voltage <- get_column(benchmark,"Voltage..low.")
featureSize <- get_column(benchmark,"Feature.Size..microns.")
channel <- get_column(benchmark,"Channel.length..microns.")
F04delay <- get_column(benchmark,"F04.Delay..ps.")
Llicache <- get_column(benchmark,"L1..instruction...on.chip.")
```

```
Lldcache <- get_column(benchmark,"L1..data...on.chip.")
L2cache <- get_column(benchmark,"L2..on.chip.")
L3cache <- get_column(benchmark,"L3..on.chip.")

return(data.frame(nperf,perf,clock,threads,cores,TDP,transistors,dieSize,voltage,featureSize
,channel,F04delay,Llicache,Lldcache,L2cache,L3cache))

}
#####
```

```
# Extract a new data frame for each of the benchmark programs available in the data set.
```

```
int92.dat <- extract_data("Int1992")
fp92.dat <- extract_data("Fp1992")
int95.dat <- extract_data("Int1995")
fp95.dat <- extract_data("Fp1995")
int00.dat <- extract_data("Int2000")
fp00.dat <- extract_data("Fp2000")
int06.dat <- extract_data("Int2006")
fp06.dat <- extract_data("Fp2006")
```