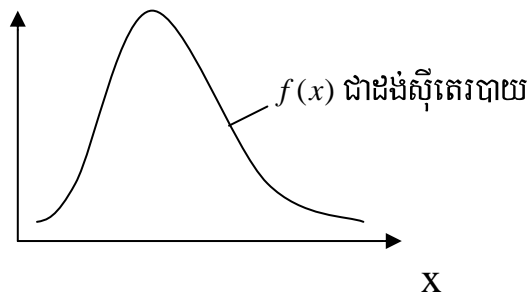


ប្រូបាប៊ីលីតេនៃអថេរចៃដន្យជាប់ (Probability of continuous random variables)

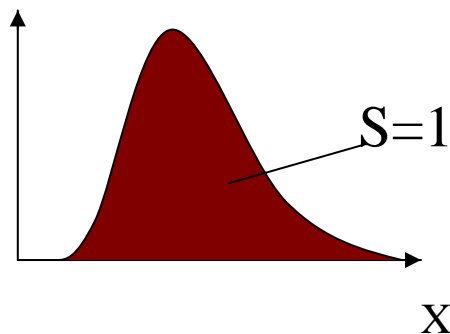
បើ X ជាអថេរចៃដន្យជាប់ (continuous random variables) នោះវាមានដង់ស៊ីតេ
របាយ (density distribution) $f(x), x \in \mathfrak{R}$ ដែល $F(x) = \int_{-\infty}^{+\infty} f(t)dt$ ហើយ $F(x)$ ជា
អនុគមន៍របាយ (distribution function) ។

លក្ខណៈសំគាល់នៃដង់ស៊ីតេរបាយ :

១- $f(x) \geq 0 \forall x \in \mathfrak{R}$ បានន័យថាខ្សែកោងដង់ស៊ីតេរបាយជានិច្ចជាកាលស្ថិតនៅខាងលើខ្សែអ័ក្ស
អាប៊ស៊ីស ។



២- $\int_{-\infty}^{+\infty} f(t)dt = 1$ បានន័យថាក្រលាផ្ទៃដែលខ័ណ្ឌដោយខ្សែកោងដង់ស៊ីតេរបាយ និងអ័ក្សអាប៊ស៊ីស
ស្មើ 1 ។



៣- $P(a \leq X \leq b)$ ជាក្រលាផ្ទៃដែលបង្កើតឡើងដោយខ្សែកោងដងស៊ីតេរេបាយនឹងអ័ក្សអាប់ស៊ីស ខ័ណ្ឌដោយបន្ទាត់ឈរកាត់តាមចំនុចអាប់ស៊ីស a និង b ។

ចំណាំ : នៅក្នុងករណីដែល X ជាអថេរចៃដន្យជាប់ នោះ :

១- $P(a < X < b) = P(a \leq X < b) = P(a < X \leq b) = P(a \leq X \leq b) = F(b) - F(a)$

២- $P(X = a) = P(a \leq X \leq a) = F(a) - F(a) = 0$

**សង្ឃឹមគណិត និង វារ្យង់នៃអថេរចៃដន្យ
(Expectation and variance of random variables)**

១- សង្ឃឹមគណិតនៃអថេរចៃដន្យដាច់ :

X ជាអថេរចៃដន្យដាច់ដែលអាចទទួលតំលៃ x_1, x_2, \dots, x_n ហើយដែលមានប្រូបាប៊ីលីតេរៀង $P(X=x_i)=p_i, i=1,2,3,\dots,n$ ។ សង្ឃឹមគណិតនៃអថេរចៃដន្យ X កំណត់ដូចខាងក្រោម :

$$E(X) = \sum_{i=1}^n x_i p_i$$

តំលៃសង្ឃឹមគណិតជាចំនួនពិត μ ។ នៅក្នុងស្ថិតិកេច្រើនប្រើនិមិត្តសញ្ញា μ ជំនួសតំលៃ $E(X)$ ។

២- វារ្យង់និងតំលាតគំរូ (Standard deviation) នៃអថេរចៃដន្យដាច់ :

បើ X ជាអថេរចៃដន្យដាច់ដែលអាចទទួលតំលៃ x_1, x_2, \dots, x_n នោះវារ្យង់នៃ X កំណត់ដោយ :

$$V(x) = \sum_{i=1}^n (x_i - E(X))^2 p_i$$

ដែល $p_i = P(X = x_i), i = 1, 2, \dots, n$ ។

រឹសការេនៃវារ្យង់អោយឈ្មោះថា standard deviation ដែលតាងដោយ $\sigma(X)$:

$$\sigma(X) = \sqrt{V(X)}$$

ជាទូទៅវារ្យង់នៃអថេរចៃដន្យ X កំណត់ដោយ :

$$V(X) = E(X - E(X))^2 = E(X^2) - E^2(X)$$

3- សង្ឃឹមគណិត និង វារ្យង់នៃអថេរចៃដន្យជាប់ (continuous random variables)

នៅក្នុងករណីដែល X ជាអថេរចៃដន្យជាប់ (continuous random variables)

នោះសង្ឃឹមគណិតនិងវារូងនៃ X កំណត់ដោយ :

$$E(X) = \int_{-\infty}^{+\infty} xf(x)dx$$

$$V(X) = \int_{-\infty}^{+\infty} (x - E(X))^2 f(x)dx$$

ដែល $f(x)$ ជាដង់ស៊ីតេរបាយនៃ X ។

៤- ករណីដែលពុំស្គាល់របាយប្រូបាប៊ីលីតេ :

ក- $E(X) = \bar{X} = \frac{1}{n} \sum_{i=1}^n x_i$ -sample mean

ខ- $V(X) = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X})^2} = s$ - sample standard deviation

ឧទាហរណ៍ :

យើងមានទិន្នន័យដូចខាងក្រោម

| profit(\$) | probability |
|------------|-------------|
| 25000 | 0.25 |
| 30000 | 0.30 |
| 35000 | 0.20 |
| 40000 | 0.15 |
| 45000 | 0.10 |

$E(\text{profit})=?$

$$\sum_{i=1}^5 x_i p_i = 25000 * 0.25 + 30000 * 0.30 + 35000 * 0.20 + 40000 * 0.15 + 45000 * 0.10$$

$=32750$ ដោយគណនាតាមតារាងរបាយប្រូបាប៊ីលីតេ ។

ក្នុងករណីដែលតារាងរបាយប្រូបាប៊ីលីតេពុំស្គាល់ $E(X)$ កំណត់ដោយ :

$$E(\text{profit}) = \frac{1}{5} \sum_{i=1}^5 x_i = \frac{25000 + 30000 + 35000 + 40000 + 45000}{5} = 35000$$

ក្នុងករណីដែល n ធំ $\sum_{i=1}^n x_i p_i \cong \frac{1}{n} \sum_{i=1}^n x_i$ ។

$S\text{dandard.deviation}(\text{profit})=?$

$$\text{Std.deviation} = \sqrt{\sum_{i=1}^5 (x_i - E(X))^2 * p_i} = 6417.749$$

ក្នុងករណីដែលពុំស្គាល់តារាងរបាយប្រូបាប៊ីលីតេ std.deviation កំណត់ដោយ :

$$\text{std.deviation} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X})^2 = 7905.694$$

ក្នុងករណីដែល n ធំ $\sqrt{\sum_{i=1}^n (x_i - E(X))^2 * p_i} \cong \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X})^2$

៥-មេគុណបំរែបំរួល (coefficient of variation):

ជាទូទៅមេគុណបំរែបំរួលកំណត់ដោយ :

$$c.v\% = \frac{\sigma}{E(X)} * 100$$

នៅក្នុងការអនុវត្តផ្នែកសេដ្ឋកិច្ច coefficient of variation គឺជា Risk ។ កាលណា តំលៃ coefficient of variation ធំបានសេចក្តីថាគំរោងពីអថេរចៃដន្យនិមួយៗមានតំលៃធំ ។ កាលណា coefficient of variation មានតំលៃតូចជាង 15% បានសេចក្តីថា Risk តូច របាយអថេរ នៃការអង្កេតមានលក្ខណៈស្មើសាច់ (មានភាពប្រែប្រួលតិចពីអថេរមួយទៅអថេរមួយ) យើងក៏អាចនិយាយ ផងដែរថា standard deviation តូច ។

ឧទាហរណ៍:

សហគមន៍ “optigestion” មានគំរោងទិញភាគហ៊ុននៅក្នុងសហគ្រាសមួយកន្លែងដែលផលិត micro-informatics ។ គេបានពិនិត្យមើលលើតារាងលិខិតក្នុងរយៈពេលមួយឆ្នាំ ដែលមានស្រាប់នៅក្នុង សហគ្រាសនេះ ។

| Output % | Probability p _i |
|----------|-------------------------------|
| 28.0 | 0.05 |
| 23.0 | 0.21 |
| 18.6 | 0.34 |
| 15.0 | 0.22 |
| 12.0 | 0.10 |
| 8.0 | 0.08 |

១-តើសហគមន៍ optigestion អាចនឹងសង្ឃឹមថាទទួលបាន output ប៉ុន្មានដែរ ?

២-គណនាគំណាត់គំរូ (standard deviation) និងមេគុណបំរែបំរួល (cv) នៃទិន្នផល (output)?

៣- សហគមន៍នេះបានសិក្សាផងដែរទៅ លើការទិញកំចីប័ណ្ណរដ្ឋាភិបាល (Government obligation) ដែលមានទិន្នផលថេរ 12% ថេរ ។ គណនាមេគុណបំរែបំរួល (C.V) ?

៥- តើការទិញប័ណ្ណភាគហ៊ុនប្រភេទណាដែលមាន Risk ធំជាងគេ ?

ចំណើយ:

ដោយប្រើ program S-plus ដូចខាងក្រោម :

```

function()
{
  output <- c(28, 23.4, 18.6, 15, 12, 8)
  prob <- c(0.05, 0.21, 0.34, 0.22, 0.1, 0.08)
  output.square <- output^2
  prod <- (output^2) * prob
  Table <- cbind(output, output.square, prob, prod)
  expectation <- sum(output * prob)
  std.dev <- sqrt(sum((output - expectation)^2 * prob))
  C.V <- (std.dev/expectation) * 100
  return(Table, expectation, std.dev, C.V)
}

```

យើងទទួលបានចំណើយដូចខាងក្រោម :

```

$Table:
      output output.square  prob  prod
[1,]  28.0      784.0    0.05 39.2000
[2,]  23.4      547.56    0.21 114.9876
[3,]  18.6      345.96    0.34 117.6264
[4,]  15.0      225.00    0.22  49.5000
[5,]  12.0      144.00    0.10  14.4000
[6,]   8.0       64.00    0.08   5.1200

```

```
$expectation:
[1] 17.778
```

```
$std.dev:
[1] 4.977622
```

\$C.V:
[1] 27.99877

យោងទៅលើចំណែកខាងលើយើងបាន :

$$១- E(X) = \sum_{i=1}^6 x_i p_i = 17.778$$

$$២- V(X) = E(X - E(X))^2 = E(X^2) - E^2(X) \Rightarrow \sigma(X) = 4.977622$$

$$C.V\% = \frac{\sigma(X)}{E(X)} * 100 = 27.998$$

៣-ដោយ Government obligation មានទិន្នផលជាក់លាក់ 12% ថេរ នោះ
standard deviation ស្មើ 0 នាំអោយ C.V=0% ។

៤-

| Investment | Risk |
|--|-------------|
| Investment នៅក្នុងសហគ្រាស Micro-informatics | C.V=28.998% |
| Government obligation | C.V=0% |

៦- លក្ខណៈសំគាល់នៃសម្បូរគណិត និងហ វារ្យង់ :

ក- សម្បូរគណិត:

$$E(X+c)=E(X)+c \quad \text{ដែល } c \text{ ជាតំលៃថេរ}$$

$$E(aX)=aE(X) \quad \text{ដែល } a \text{ ជាតំលៃថេរ}$$

$$E(aX+c)=aE(X)+c$$

ខ- វារ្យង់ :

$$V(X+c)=V(X) \quad \text{ដែល } c \text{ ជាចំនួនថេរ}$$

$$V(aX)=a^2V(X)$$

$$V(aX+c)=a^2V(X)$$

គ- តំលាតគុំរ (standard deviation)

$$\sigma(X + c) = \sigma(X)$$

$$\sigma(aX) = a\sigma(X)$$

$$\sigma(aX + c) = a\sigma(X)$$

៧- **វិសមភាព Tchebychev :**

សន្មតិថា X ជាអថេរចៃដន្យដែលមានច្បាប់ប្រូបាប៊ីលីតេមួយកំណត់ មានសង្ឃឹមគណិត $E(X)$ និង វារ្យង់ $V(X) = \sigma^2(X)$ ។

វិសមភាព Tchebychev កំណត់ដូចខាងក្រោម :

$$P[E(X) - k\sigma(X) \leq X \leq E(X) + k\sigma(X)] \geq 1 - \frac{1}{k^2}$$

ដែល k ជាចំនួនថេរ ។

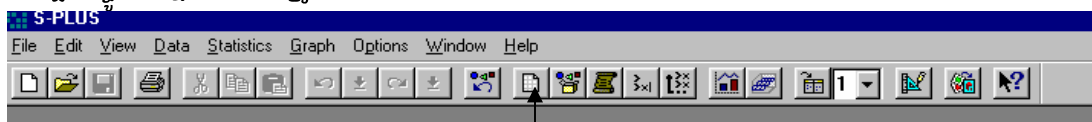
វិសមភាពនេះផ្ទៀងផ្ទាត់នូវគ្រប់ច្បាប់ប្រូបាប៊ីលីតេនៃអថេរចៃដន្យ (ដាច់ ឬ ជាប់) ក្នុងលក្ខខណ្ឌដែល យើងអាចកំណត់បាននូវតំលៃ $E(X)$ និង តំលៃ $\sigma(X)$ ។

Spplus និងការវិភាគទិន្នន័យ(data analysis)

Spplus ជា software មួយប្រភេទដូច SPSS ដែរ ប៉ុន្តែវាមានលក្ខណៈវិភាគស៊ីជម្រៅជាង SPSS និង ជា programming language ។ Spplus ជាប្រភេទ software ដែលផ្តល់លក្ខណៈងាយស្រួលក្នុងការបង្កើត graphics ។

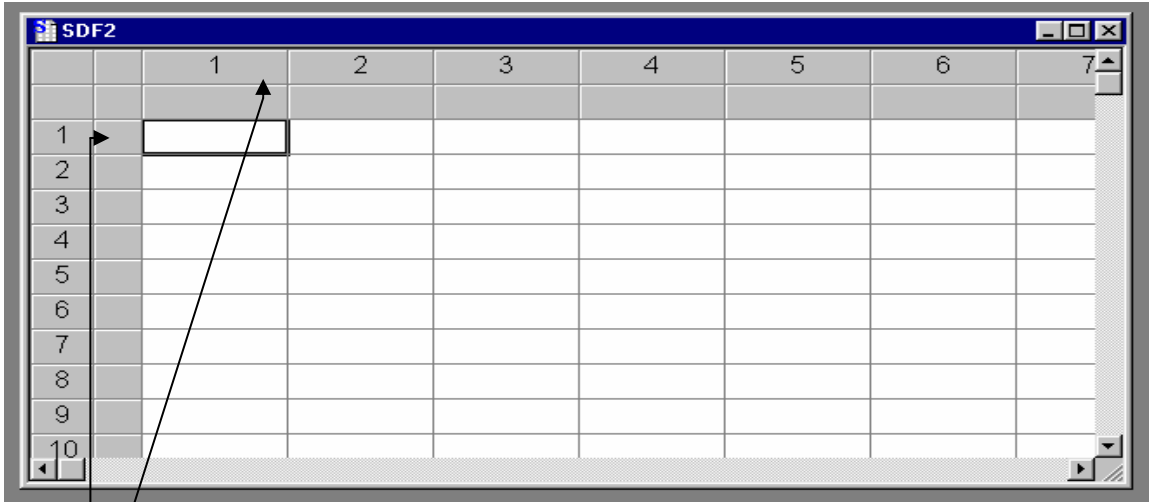
១- **ការញូសទិន្នន័យ (data input) :**

ដើម្បីបញ្ចូល ទិន្នន័យ យើងត្រូវហៅ data editor window :

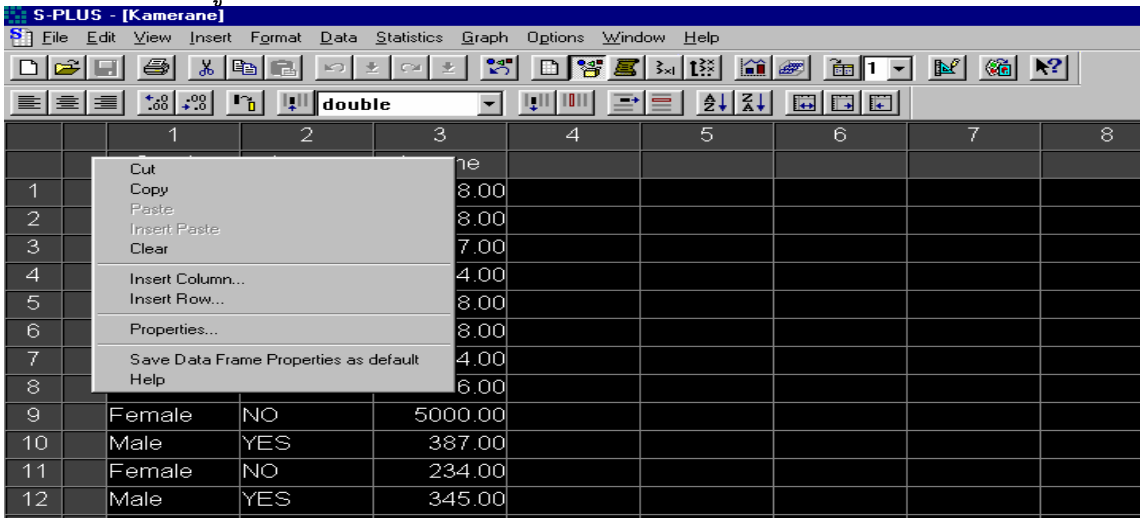


new data frame

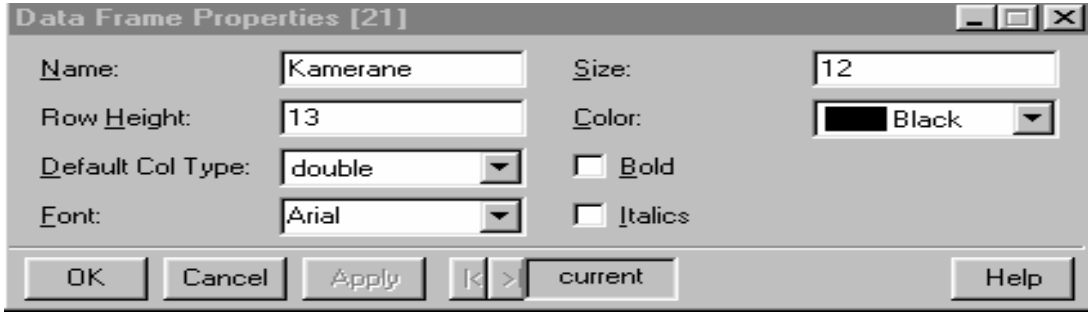
ចុចទៅលើ button ដែលមានឈ្មោះថា new data frame យើងទទួលបាន data editor window ដែលមានរាងដូចខាងក្រោម :



ចុចពីរដងទៅលើ ថតខាងលើដើម្បីបញ្ចូល variable name ។
 ដូចគ្នានេះដែរចុចពីរដងទៅលើថតនៃជួរដេកនេះដើម្បីបញ្ចូល name of observation
 ប៉ុន្តែត្រូវបញ្ចូលទិន្នន័យនៅជួរឈរសិន ។



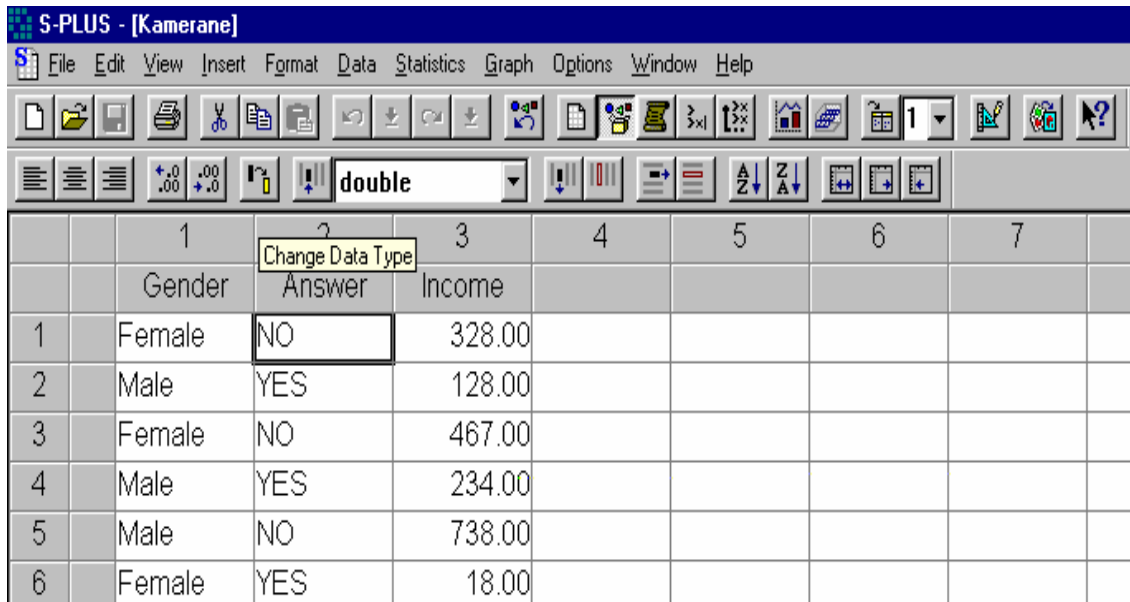
នៅពេលបញ្ចូលទិន្នន័យចប់ចូរសង្កេតមើលឈ្មោះទិន្នន័យនៅផ្នែកខាងលើនៃ data editor window ដែលអោយដោយ default របស់ computer ។ យើងអាចដូរឈ្មោះទិន្នន័យនេះដោយ គ្រាន់តែ select data editor window ហើយចុច mouse ខាងស្តាំ select យក properties ហើយដូរឈ្មោះ ដែលយើងចង់បាន :



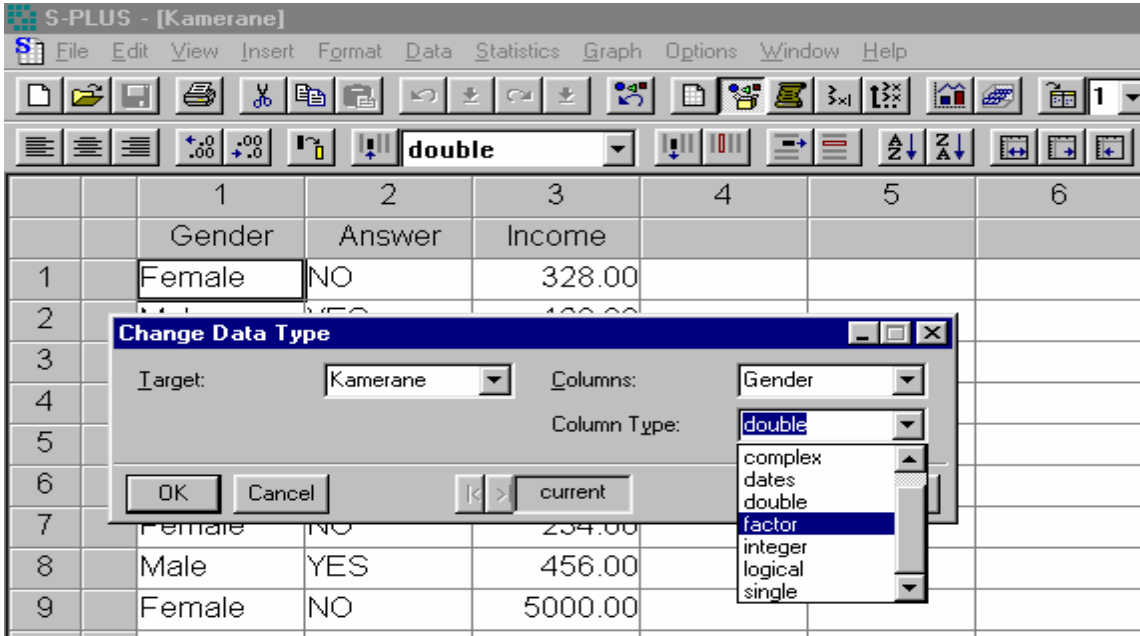
ពេលនេះយើងអាចប្តូរឈ្មោះ data editor window នៅក្នុងថត Name តាមដែលយើងចង់បាន ។

២-លក្ខណៈនៃទិន្នន័យ (Data type) :

ទិន្នន័យដែលយើងប្រើប្រាស់ក្នុងកញ្ចប់មានលក្ខណៈជា numeric, string, dates,.....និងមានមួយចំនួនទៀត ដែលយើងអាចប្រើបាននៅក្នុង change data type button ដែលមានរូបរាងដូចខាងក្រោម:



មុននឹងបញ្ចូលទិន្នន័យយើងត្រូវសង្កេតទៅលើ data type សិន តើ វាជា numeric , string ឬប្រភេទទិន្នន័យបែបណា ។ ឧទាហរណ៍នៅក្នុងតារាងខាងលើនៅ column ទីមួយទិន្នន័យមានលក្ខណៈ ជា string ដូច្នេះនៅក្នុងការបញ្ចូលទិន្នន័យនេះ យើងត្រូវបញ្ចូល variable name រួច change data type :



បន្ទាប់មក select យក factor ព្រោះថាទិន្នន័យរបស់យើងមាន factor ពីរ គឺ Female និង Male ។ ដូចគ្នានេះដែរចំពោះ column ទីពីរ ប៉ុន្តែ column ទី 3 ទិន្នន័យមានលក្ខណៈជា numerical data ដូច្នេះការបញ្ចូលព័ត៌មានត្រូវបានផ្លាស់ប្តូរ ការ change data type ទេ ។

៣-Descriptive analysis :

ការវិភាគទិន្នន័យមានលក្ខណៈស្តុកស្តួញទៅតាមប្រធានបទផ្សេងៗគ្នា ។ នៅពេលនេះយើងលើកយកតែ ការវិភាគដែលមានលក្ខណៈជា Descriptive analysis ។ ឧទាហរណ៍យើងមានទិន្នន័យដូចខាង

ក្រោម :

| | CUSTOMER | REVENUE | REGION | INDUSTRY |
|----|--------------------|------------|--------|------------|
| 1 | Regular customer | 3786.8670 | West | Government |
| 2 | Regular customer | 1733.6005 | South | Commercial |
| 3 | Preferred customer | 2125.7356 | West | Government |
| 4 | Preferred customer | 2259.1379 | West | Commercial |
| 5 | Preferred customer | 1586.6104 | North | Government |
| 6 | Regular customer | NA | West | Academic |
| 7 | Preferred customer | 1838.4591 | South | Government |
| 8 | Regular customer | 1846.7495 | West | Government |
| 9 | Regular customer | 1714.1120 | East | Academic |
| 10 | Preferred customer | 1717.5371 | South | Academic |
| 11 | Preferred customer | 4388.0708 | West | Government |
| 12 | Preferred customer | 12346.0000 | North | Commercial |
| 13 | Preferred customer | 3833.5526 | West | Government |
| 14 | Preferred customer | 2290.6734 | North | Commercial |
| 15 | Preferred customer | 4139.7650 | South | Commercial |
| 16 | Preferred customer | 2193.9056 | South | Academic |
| 17 | Preferred customer | 2938.3083 | South | Government |
| 18 | Regular customer | 3313.3158 | East | Commercial |
| 19 | Preferred customer | 3327.0176 | East | Academic |

BUILD BRIGH UNIVERSITY, FACULTY OF SCIENCE AND TECHNOLOGY.

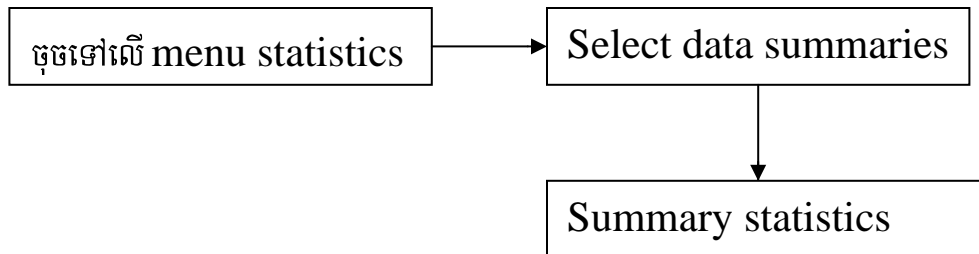
| | | | | |
|----|--------------------|------------|-------|------------|
| 20 | Regular customer | 1449.3379 | East | Government |
| 21 | Preferred customer | 2695.6158 | North | Academic |
| 22 | Preferred customer | 3272.2086 | South | Commercial |
| 23 | Regular customer | 4096.3530 | North | Academic |
| 24 | Regular customer | 4339.9974 | South | Academic |
| 25 | Regular customer | 2462.3752 | West | Government |
| 26 | Preferred customer | 3629.5420 | East | Commercial |
| 27 | Preferred customer | 465.1717 | West | Academic |
| 28 | Preferred customer | 1870.9232 | North | Commercial |
| 29 | Preferred customer | 1376.5828 | West | Academic |
| 30 | Regular customer | 3796.7922 | South | Government |
| 31 | Preferred customer | 4253.4813 | North | Academic |
| 32 | Preferred customer | 3057.0979 | South | Government |
| 33 | Preferred customer | 3394.7124 | West | Government |
| 34 | Regular customer | 1387.3440 | West | Government |
| 35 | Preferred customer | 2148.4856 | North | Commercial |
| 36 | Preferred customer | 2204.7234 | West | Commercial |
| 37 | Preferred customer | 2021.8557 | South | Academic |
| 38 | Regular customer | 123.0000 | East | Government |
| 39 | Preferred customer | 3301.1004 | West | Academic |
| 40 | Regular customer | 1855.5423 | South | Government |
| 41 | Preferred customer | 2885.1827 | East | Government |
| 42 | Preferred customer | 3353.5838 | East | Academic |
| 43 | Preferred customer | 2263.7547 | East | Commercial |
| 44 | Regular customer | 1620.7525 | West | Government |
| 45 | Regular customer | 3398.7307 | East | Academic |
| 46 | Preferred customer | 3987.9912 | West | Government |
| 47 | Preferred customer | 3808.7805 | East | Academic |
| 48 | Preferred customer | 3605.4659 | West | Academic |
| 49 | Regular customer | 3226.4753 | West | Commercial |
| 50 | Regular customer | 4356.6054 | South | Commercial |
| 51 | Regular customer | 2120.6041 | East | Academic |
| 52 | Preferred customer | 3943.9258 | North | Government |
| 53 | Regular customer | 3689.8859 | West | Commercial |
| 54 | Regular customer | 430.1497 | West | Commercial |
| 55 | Preferred customer | 2267.0971 | North | Government |
| 56 | Preferred customer | 1123.1291 | South | Academic |
| 57 | Preferred customer | 2384.9484 | West | Commercial |
| 58 | Regular customer | 5011.5981 | West | Academic |
| 59 | Preferred customer | 34566.0000 | West | Commercial |
| 60 | Preferred customer | 2742.6422 | West | Commercial |
| 61 | Preferred customer | 2608.7393 | North | Commercial |
| 62 | Preferred customer | 2261.5493 | West | Commercial |
| 63 | Regular customer | 2744.7494 | West | Academic |
| 64 | Regular customer | 4024.3463 | North | Commercial |
| 65 | Preferred customer | 4728.5242 | South | Commercial |
| 66 | Preferred customer | 3185.4868 | North | Academic |
| 67 | Regular customer | 2406.9697 | East | Commercial |
| 68 | Preferred customer | 5417.3652 | East | Commercial |
| 69 | Regular customer | 3875.5855 | East | Government |
| 70 | Preferred customer | 1902.2245 | East | Commercial |
| 71 | Regular customer | 836.1970 | East | Government |
| 72 | Regular customer | 3750.3950 | South | Academic |
| 73 | Preferred customer | 3418.8871 | East | Government |
| 74 | Regular customer | 2068.3873 | North | Commercial |
| 75 | Regular customer | 3810.8523 | West | Commercial |

Prepared by MEAK KAMERANE(Statistician, Data analyst, Mathematician).

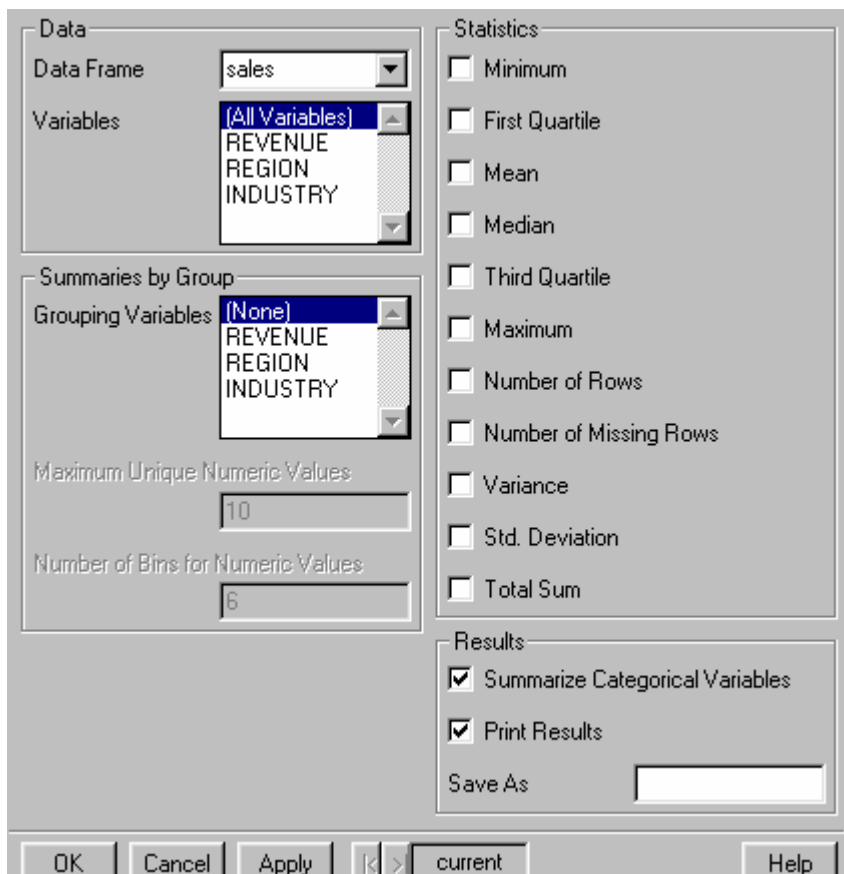
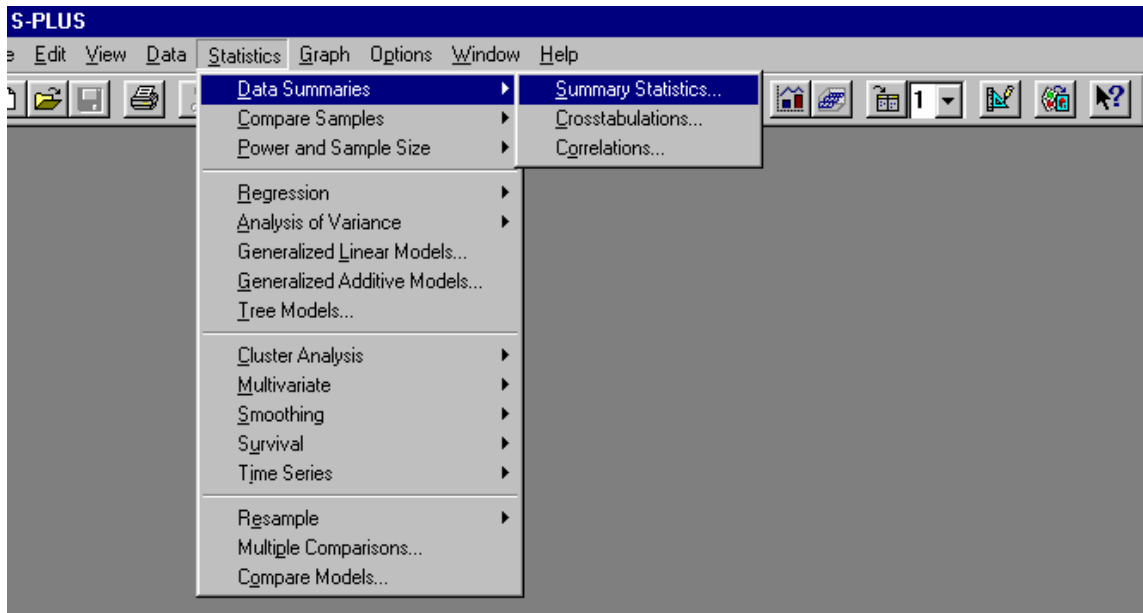
| | | | | |
|-----|--------------------|------------|-------|------------|
| 76 | Preferred customer | 432.0000 | South | Academic |
| 77 | Preferred customer | 430.4231 | South | Academic |
| 78 | Preferred customer | 718.5732 | North | Commercial |
| 79 | Preferred customer | 2570.7565 | North | Commercial |
| 80 | Preferred customer | 2671.3214 | West | Commercial |
| 81 | Preferred customer | 2241.6868 | South | Academic |
| 82 | Regular customer | 2409.5362 | South | Government |
| 83 | Regular customer | 54322.0000 | West | Government |
| 84 | Regular customer | 1547.8898 | East | Academic |
| 85 | Preferred customer | 2233.1350 | West | Government |
| 86 | Regular customer | 1283.8048 | South | Government |
| 87 | Preferred customer | 2658.7289 | North | Government |
| 88 | Regular customer | 2823.8458 | East | Government |
| 89 | Preferred customer | 336.6098 | East | Commercial |
| 90 | Preferred customer | 2120.1780 | East | Academic |
| 91 | Preferred customer | 2917.7317 | East | Government |
| 92 | Preferred customer | 2370.9864 | North | Commercial |
| 93 | Regular customer | 2696.9629 | East | Academic |
| 94 | Regular customer | 1975.5370 | East | Academic |
| 95 | Regular customer | 2704.2350 | East | Government |
| 96 | Regular customer | 3658.2319 | North | Government |
| 97 | Regular customer | 3501.6697 | East | Government |
| 98 | Preferred customer | 2895.8984 | North | Government |
| 99 | Regular customer | 1429.5769 | East | Commercial |
| 100 | Regular customer | 3913.8769 | South | Academic |

ដោយ column ទី 1,3,4 ទិន្នន័យមានលក្ខណៈជា string (factor) ដូច្នេះ យើងធ្វើ

descriptive analysis ទៅលើ columns ទាំងនេះ ដូចខាងក្រោម :



យើងទទួលបានតារាងដូចខាងក្រោម :



ជ្រើសរើសយក variable names ដែលយើងចង់ធ្វើការវិភាគ ហើយរក្សាទុកនូវផ្នែក ខាង ក្រោមគឺ Result នូវ summarize Categorical Variables និង Print Results ព្រោះទិន្នន័យមានលក្ខណៈជា string ដូច្នេះការវិភាគ ត្រូវធ្វើឡើងតាម frequency analysis ។ នៅក្នុងតារាងខាងលើយើង select Costumer បន្ទាប់មក select Region and Industry រួចចុច OK យើងទទួលបានចំណេញដូចខាងក្រោម :

*** Summary Statistics for data in: sales ***

CUSTOMER

Preferred customer:58
Regular customer:42

*** Summary Statistics for data in: sales ***

| REGION | INDUSTRY |
|----------|---------------|
| East:28 | Academic:31 |
| North:20 | Commercial:34 |
| South:22 | Government:35 |
| West:30 | |

ដូច្នេះយើងឃើញថា summary statistics បានរាប់ចំនួន factor និមួយៗ នៅក្នុង variable names ដែលយើងធ្វើការវិភាគ ។

Graphics :

ដោយអនុវត្តទៅលើទិន្នន័យខាងលើហើយដោយប្រើកម្មវិធី S-plus (programming) យើងអាចធ្វើ data summary និង graphical analysis ដូចខាងក្រោម :

```
function(sales)
{
  i <- 1:3
  analysis <- summary(sales[, i])
  par(mfg = c(1, 1, 2, 2))
  graph1 <- hist(sales[, 1], nclass = 30, xlab =
    "REVENUE", ylab = "NUMBER OF PERSON",
    main = "Descriptive analysis of revenue",
    )
  par(mfg = c(1, 2, 2, 2))
  graph2 <- barplot(summary(sales[, 2]), names =
    levels(factor(sales[, 2])), xlab =
    "Region", ylab = "Number of person",
    main = "Descriptive analysis of region"
    )
  text(graph2, summary(sales[, 2]) + 1, paste(
    "n=", summary(sales[, 2])))
  par(mfg = c(2, 1, 2, 1))
  graph3 <- pie(summary(sales[, 3]), names =
    levels(factor(sales[, 3])), col = 2:4,
    explode = 1:3, size = 1.3, rotate = F,
    main = "Descriptive analysis of person's job"
    )
  return(analysis)
}
```

output of data analysis :

| REVENUE | REGION | INDUSTRY |
|---------------|----------|---------------|
| Min.: 123 | East:28 | Academic:31 |
| 1st Qu.: 1890 | North:20 | Commercial:34 |
| Median: 2670 | South:22 | Government:35 |
| Mean: 3570 | West:30 | |
| 3rd Qu.: 3640 | | |
| Max.:54300 | | |

