```
(*
Instructor：Hiroyuki Akama
 Statistical  Analyses*)
```

In[160]:=
```
<< ANOVA`
<< HierarchicalClustering`
<< Histograms`
<< HypothesisTesting`
<< LinearRegression`
<< MultivariateStatistics`
<< StatisticalPlots`
```

General::obspkg :
Histograms` is now obsolete. The legacy version being loaded may conflict with current Mathematica functionality. See the Compatibility Guide for updating information. ≫

BarChart3D::shdw : Symbol BarChart3D appears in multiple contexts {BarCharts`, System`}; definitions in context BarCharts` may shadow or be shadowed by other definitions. ≫

Histogram3D::shdw : Symbol Histogram3D appears in multiple contexts {Histograms`, System`}; definitions in context Histograms` may shadow or be shadowed by other definitions. ≫

General::obspkg :
LinearRegression` is now obsolete. The legacy version being loaded may conflict with current Mathematica functionality. See the Compatibility Guide for updating information. ≫

In[167]:=
```
? <<
```

<< *name* reads in a file, evaluating each expression in it, and returning the last one. ≫

In[168]:=
```
dummydata1 = Table[RandomReal[], {i, 1, 100}]
```

Out[168]=
```
{0.128726, 0.467632, 0.44974, 0.847347, 0.420269, 0.847157, 0.435793, 0.401635, 0.20875,
 0.0823396, 0.883464, 0.650875, 0.171611, 0.152437, 0.232594, 0.202863, 0.793541,
 0.679996, 0.5981, 0.745931, 0.613366, 0.151293, 0.110339, 0.654964, 0.816853, 0.798738,
 0.284973, 0.722395, 0.0971221, 0.617709, 0.658894, 0.730312, 0.167853, 0.90637, 0.197484,
 0.845529, 0.240102, 0.211059, 0.0502255, 0.131545, 0.804643, 0.424158, 0.0303365,
 0.309669, 0.399737, 0.877938, 0.595326, 0.409811, 0.720729, 0.533928, 0.474595,
 0.59092, 0.655296, 0.0865104, 0.00790366, 0.804715, 0.756363, 0.202581, 0.394989,
 0.498079, 0.234343, 0.34621, 0.715015, 0.0666734, 0.73511, 0.560279, 0.728437, 0.923816,
 0.653285, 0.453134, 0.998845, 0.514393, 0.678287, 0.614422, 0.196597, 0.656616,
 0.0739574, 0.384497, 0.224065, 0.698815, 0.336951, 0.927486, 0.160045, 0.149327,
 0.0824671, 0.160478, 0.161564, 0.724479, 0.884009, 0.764644, 0.0240731, 0.00660753,
 0.644755, 0.909073, 0.350382, 0.574691, 0.18821, 0.0154269, 0.757284, 0.727519}
```

In[169]:=
```
Mean[dummydata1]
Median[dummydata1]
? Mean
? Median
```

Out[169]= 0.469924

Out[170]= 0.471113

Mean[*list*] gives the statistical mean of the elements in *list*.
Mean[*dist*] gives the mean of the symbolic distribution *dist*. ≫

Median[*list*] gives the median of the elements in *list*.
Median[*dist*] gives the median of the symbolic distribution *dist*. ≫

In[173]:=
```
StandardDeviation[dummydata1]
?StandardDeviation
```

Out[173]= 0.285072

StandardDeviation[*list*] gives the sample standard deviation of the elements in *list*.
StandardDeviation[*dist*] gives the standard deviation of the symbolic distribution *dist*. ≫

In[175]:=
```
Variance[dummydata1]
%%^2 == %
?Variance
```

Out[175]= 0.0812662

Out[176]= $\{\text{Null}^2\}$ == 0.0812662

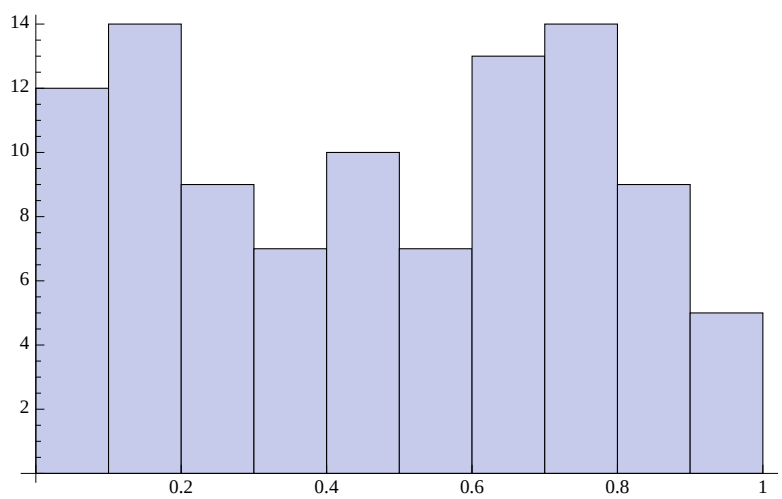
Variance[*list*] gives the sample variance of the elements in *list*.
Variance[*dist*] gives the variance of the symbolic distribution *dist*. ≫

In[178]:=
```
?Histogram
```

Histogram[{ x_1, x_2, … }] generates a histogram of the univariate data { x_1, x_2, … }.
Histogram[{ f_1, f_2, … }, FrequencyData −> True] generates a histogram of the univariate frequency data { f_1, f_2, … }, where f_i is the frequency with which the original data occurs in category i. ≫

In[179]:=
```
Histogram[dummydata1]
```

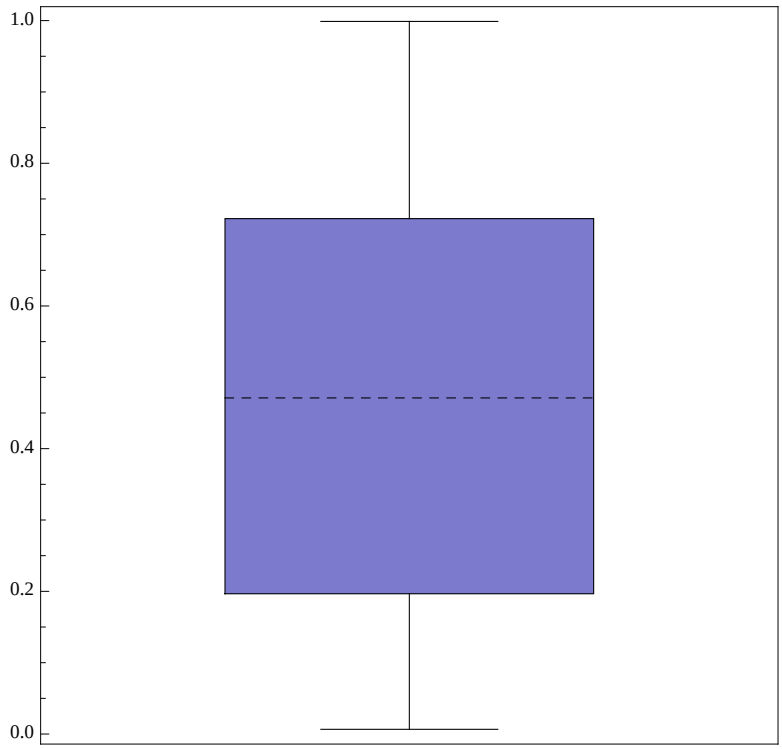
Out[179]=



In[180]:=
```
? BoxWhiskerPlot
BoxWhiskerPlot[dummydata1, BoxMedianStyle → Dashed]
```

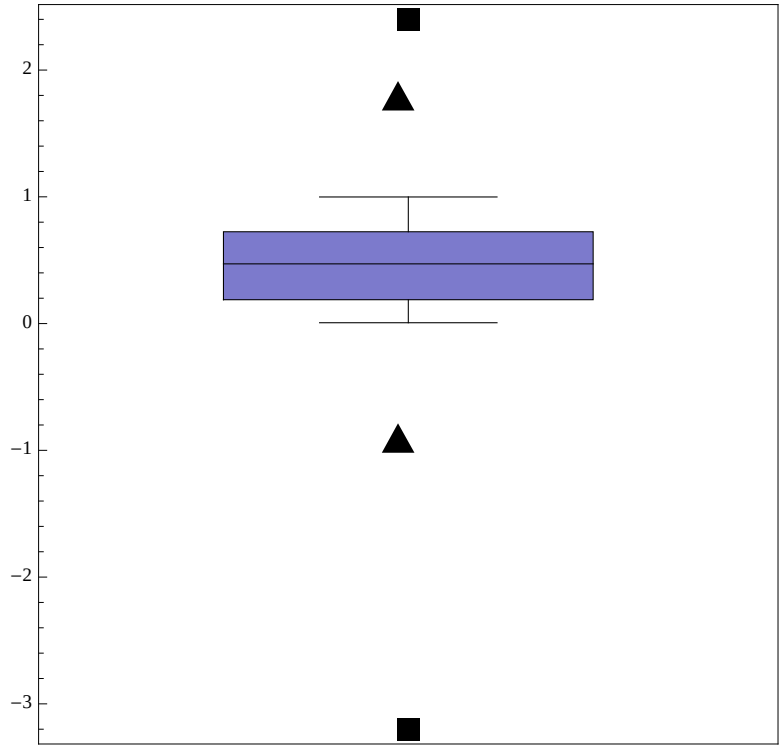
BoxWhiskerPlot[data] creates a box–and–whisker plot of the given vector of data .
BoxWhiskerPlot[data $_1$, data $_2$, …] creates multiple box–and–whisker plots.
BoxWhiskerPlot[matrix] also creates multiple
box–and–whisker plots, with each column of the matrix used for a separate box. ≫

Out[181]=



In[182]:=
```
dummydata2 = Join[dummydata1, {2.4, 1.8, -0.9, -3.2}];
BoxWhiskerPlot[dummydata2, BoxOutliers → Automatic, BoxOutlierMarkers → {▲, ■}]
```

Out[183]=



In[184]:=
```
? MeanTest
```

MeanTest[list, μ_0] performs a test with null hypothesis $\mu = \mu_0$. ≫

In[185]:=
```
MeanTest[Table[RandomReal[{0.45, 0.55}], {i, 1, 100}],
 1, SignificanceLevel → .05, TwoSided → True]
MeanTest[Table[RandomReal[{0.499, 0.501}], {i, 1, 100}],
 0.5, SignificanceLevel → .05, TwoSided → True]
```

Out[185]= {TwoSidedPValue → 2.7715×10^{-126}, Reject null hypothesis at significance level → 0.05}

Out[186]= {TwoSidedPValue → 0.481685, Fail to reject null hypothesis at significance level → 0.05}

In[187]:=
```
? MeanDifferenceTest
```

MeanDifferenceTest[list_1, list_2, $\Delta\mu_0$] performs a test with null hypothesis $\mu_1 - \mu_2 = \Delta\mu_0$. ≫

In[188]:=
```
diffsample1 = Table[RandomReal[{0.45, 0.55}], {i, 1, 100}];
diffsample2 = Table[RandomReal[{0.45, 0.55}], {i, 1, 100}];
MeanDifferenceTest[diffsample1, diffsample2, 0, SignificanceLevel → .05]
```

Out[190]= {OneSidedPValue → 0.0179426, Reject null hypothesis at significance level → 0.05}

In[191]:=
```
? VarianceTest
```

VarianceTest[list, ${\sigma_0}^2$] performs a test with null hypothesis $\sigma^2 = {\sigma_0}^2$. ≫

In[192]:=
```
VarianceTest[Table[RandomReal[{0.45, 0.55}], {i, 1, 100}], 1, SignificanceLevel → .05]
```

Out[192]= {OneSidedPValue → 1.54372×10^{-133}, Reject null hypothesis at significance level → 0.05}

In[193]:=
```
Variance[Table[RandomReal[{0.45, 0.55}], {i, 1, 100}]]
```

Out[193]= 0.000873018

In[194]:=
```
VarianceTest[Table[RandomReal[{0.45, 0.55}], {i, 1, 100}], 0.0008, SignificanceLevel → .05]
```

Out[194]= {OneSidedPValue → 0.437336, Fail to reject null hypothesis at significance level → 0.05}

In[195]:=
```
? VarianceRatioTest
```

VarianceRatioTest[list $_1$, list $_2$, r] performs a test with null hypothesis $\sigma_1{}^2 / \sigma_2{}^2 = r$. ≫

In[196]:=
```
VarianceRatioTest[diffsample1, diffsample2, 1, SignificanceLevel → .05]
```

Out[196]= {OneSidedPValue → 0.268984, Fail to reject null hypothesis at significance level → 0.05}

In[197]:=
```
VarianceRatioTest[diffsample1,
 Table[RandomReal[], {i, Length[diffsample2]}], 1, SignificanceLevel → .05]
```

Out[197]= {OneSidedPValue → 2.37418×10^{-68}, Reject null hypothesis at significance level → 0.05}

In[198]:=
```
numbering = # + 1 & /@ Table[RandomInteger[5], {i, 1, 104}] // Sort // Split // Flatten
lenseq = Length[#] & /@ Split[Sort[numbering]]
```

Out[198]=
```
{1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,
 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4,
 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6}
```

Out[199]= {16, 23, 13, 18, 16, 18}

In[200]:=
```
PartitionByLengthSequence[list_, lengthseq_] := Module[{mylist}, mylist = list;
   Map[Block[{a}, a = Take[mylist, #]; mylist = Drop[mylist, #]; a] &, lengthseq]];
```

In[201]:=
```
dummydata2typea = PartitionByLengthSequence[dummydata2, lenseq]
```

Out[201]=
```
{{0.128726, 0.467632, 0.44974, 0.847347, 0.420269, 0.847157, 0.435793, 0.401635,
  0.20875, 0.0823396, 0.883464, 0.650875, 0.171611, 0.152437, 0.232594, 0.202863},
 {0.793541, 0.679996, 0.5981, 0.745931, 0.613366, 0.151293, 0.110339, 0.654964,
  0.816853, 0.798738, 0.284973, 0.722395, 0.0971221, 0.617709, 0.658894,
  0.730312, 0.167853, 0.90637, 0.197484, 0.845529, 0.240102, 0.211059, 0.0502255},
 {0.131545, 0.804643, 0.424158, 0.0303365, 0.309669, 0.399737, 0.877938,
  0.595326, 0.409811, 0.720729, 0.533928, 0.474595, 0.59092},
 {0.655296, 0.0865104, 0.00790366, 0.804715, 0.756363, 0.202581, 0.394989, 0.498079, 0.234343,
  0.34621, 0.715015, 0.0666734, 0.73511, 0.560279, 0.728437, 0.923816, 0.653285, 0.453134},
 {0.998845, 0.514393, 0.678287, 0.614422, 0.196597, 0.656616, 0.0739574, 0.384497,
  0.224065, 0.698815, 0.336951, 0.927486, 0.160045, 0.149327, 0.0824671, 0.160478},
 {0.161564, 0.724479, 0.884009, 0.764644, 0.0240731, 0.00660753, 0.644755, 0.909073,
  0.350382, 0.574691, 0.18821, 0.0154269, 0.757284, 0.727519, 2.4, 1.8, -0.9, -3.2}}
```

In[202]:=
```
dummydata2typeb = Transpose[{numbering, dummydata2}]
Split[dummydata2typeb, #1[[1]] == #2[[1]] &]
```

General::spell1 :
New symbol name "dummydata2typeb" is similar to existing symbol "dummydata2typea" and may be misspelled. ≫

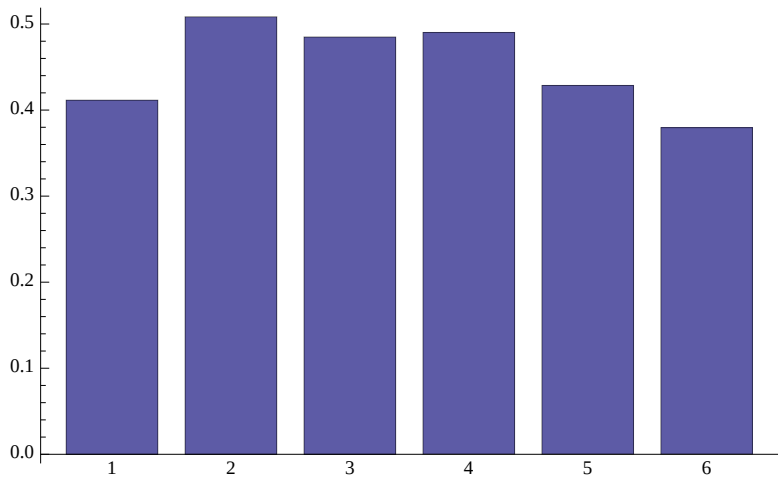
Out[202]=
```
{{1, 0.128726}, {1, 0.467632}, {1, 0.44974}, {1, 0.847347}, {1, 0.420269}, {1, 0.847157},
 {1, 0.435793}, {1, 0.401635}, {1, 0.20875}, {1, 0.0823396}, {1, 0.883464}, {1, 0.650875},
 {1, 0.171611}, {1, 0.152437}, {1, 0.232594}, {1, 0.202863}, {2, 0.793541}, {2, 0.679996},
 {2, 0.5981}, {2, 0.745931}, {2, 0.613366}, {2, 0.151293}, {2, 0.110339}, {2, 0.654964},
 {2, 0.816853}, {2, 0.798738}, {2, 0.284973}, {2, 0.722395}, {2, 0.0971221}, {2, 0.617709},
 {2, 0.658894}, {2, 0.730312}, {2, 0.167853}, {2, 0.90637}, {2, 0.197484}, {2, 0.845529},
 {2, 0.240102}, {2, 0.211059}, {2, 0.0502255}, {3, 0.131545}, {3, 0.804643}, {3, 0.424158},
 {3, 0.0303365}, {3, 0.309669}, {3, 0.399737}, {3, 0.877938}, {3, 0.595326},
 {3, 0.409811}, {3, 0.720729}, {3, 0.533928}, {3, 0.474595}, {3, 0.59092}, {4, 0.655296},
 {4, 0.0865104}, {4, 0.00790366}, {4, 0.804715}, {4, 0.756363}, {4, 0.202581},
 {4, 0.394989}, {4, 0.498079}, {4, 0.234343}, {4, 0.34621}, {4, 0.715015}, {4, 0.0666734},
 {4, 0.73511}, {4, 0.560279}, {4, 0.728437}, {4, 0.923816}, {4, 0.653285}, {4, 0.453134},
 {5, 0.998845}, {5, 0.514393}, {5, 0.678287}, {5, 0.614422}, {5, 0.196597}, {5, 0.656616},
 {5, 0.0739574}, {5, 0.384497}, {5, 0.224065}, {5, 0.698815}, {5, 0.336951},
 {5, 0.927486}, {5, 0.160045}, {5, 0.149327}, {5, 0.0824671}, {5, 0.160478},
 {6, 0.161564}, {6, 0.724479}, {6, 0.884009}, {6, 0.764644}, {6, 0.0240731},
 {6, 0.00660753}, {6, 0.644755}, {6, 0.909073}, {6, 0.350382}, {6, 0.574691}, {6, 0.18821},
 {6, 0.0154269}, {6, 0.757284}, {6, 0.727519}, {6, 2.4}, {6, 1.8}, {6, -0.9}, {6, -3.2}}
```

Out[203]=
```
{{{1, 0.128726}, {1, 0.467632}, {1, 0.44974}, {1, 0.847347}, {1, 0.420269}, {1, 0.847157},
  {1, 0.435793}, {1, 0.401635}, {1, 0.20875}, {1, 0.0823396}, {1, 0.883464},
  {1, 0.650875}, {1, 0.171611}, {1, 0.152437}, {1, 0.232594}, {1, 0.202863}},
 {{2, 0.793541}, {2, 0.679996}, {2, 0.5981}, {2, 0.745931}, {2, 0.613366}, {2, 0.151293},
  {2, 0.110339}, {2, 0.654964}, {2, 0.816853}, {2, 0.798738}, {2, 0.284973}, {2, 0.722395},
  {2, 0.0971221}, {2, 0.617709}, {2, 0.658894}, {2, 0.730312}, {2, 0.167853}, {2, 0.90637},
  {2, 0.197484}, {2, 0.845529}, {2, 0.240102}, {2, 0.211059}, {2, 0.0502255}},
 {{3, 0.131545}, {3, 0.804643}, {3, 0.424158}, {3, 0.0303365}, {3, 0.309669},
  {3, 0.399737}, {3, 0.877938}, {3, 0.595326}, {3, 0.409811},
  {3, 0.720729}, {3, 0.533928}, {3, 0.474595}, {3, 0.59092}},
 {{4, 0.655296}, {4, 0.0865104}, {4, 0.00790366}, {4, 0.804715}, {4, 0.756363}, {4, 0.202581},
  {4, 0.394989}, {4, 0.498079}, {4, 0.234343}, {4, 0.34621}, {4, 0.715015}, {4, 0.0666734},
  {4, 0.73511}, {4, 0.560279}, {4, 0.728437}, {4, 0.923816}, {4, 0.653285}, {4, 0.453134}},
 {{5, 0.998845}, {5, 0.514393}, {5, 0.678287}, {5, 0.614422}, {5, 0.196597}, {5, 0.656616},
  {5, 0.0739574}, {5, 0.384497}, {5, 0.224065}, {5, 0.698815}, {5, 0.336951},
  {5, 0.927486}, {5, 0.160045}, {5, 0.149327}, {5, 0.0824671}, {5, 0.160478}},
 {{6, 0.161564}, {6, 0.724479}, {6, 0.884009}, {6, 0.764644}, {6, 0.0240731},
  {6, 0.00660753}, {6, 0.644755}, {6, 0.909073}, {6, 0.350382}, {6, 0.574691}, {6, 0.18821},
  {6, 0.0154269}, {6, 0.757284}, {6, 0.727519}, {6, 2.4}, {6, 1.8}, {6, -0.9}, {6, -3.2}}}
```

In[204]:=
```
? BarChart
BarChart[Mean[#] & /@ dummydata2typea]
```

BarChart[{y_1, y_2, …}] generates a bar chart of the values y_1, y_2, ….
BarChart[{$list_1$, $list_2$, …}] generates a bar chart of the data in the lists. ≫

Out[205]=



In[206]:=
```
? ANOVA
ANOVA[dummydata2typeb, PostTests → Bonferroni, SignificanceLevel → .05]
```

ANOVA[*data*] performs a one-way analysis of variance.
ANOVA[*data*, *model*, *vars*] performs an
analysis of variance for *model* as a function of the categorical variables *vars*. ≫

Out[207]=
```
{ANOVA →          DF    SumOfSq     MeanSq       FRatio      PValue
          Model   5     0.242783    0.0485566    0.166316    0.974367
          Error   98    28.6114     0.291954                          ,
          Total   103   28.8542

              All       0.452812
              Model[1]  0.411452
              Model[2]  0.508398
   CellMeans → Model[3]  0.484872, PostTests → {Model → Bonferroni {}}}
              Model[4]  0.490152
              Model[5]  0.428578
              Model[6]  0.379595
```

In[208]:=
```
dummydata3 = Join[Transpose[{Table[1, {i, 25}], Table[RandomReal[{0, 1}], {i, 1, 25}]}],
   Transpose[{Table[2, {i, 25}], Table[RandomReal[{0.3, 1.3}], {i, 1, 25}]}],
   Transpose[{Table[3, {i, 25}], Table[RandomReal[{0.6, 1.6}], {i, 1, 25}]}],
   Transpose[{Table[4, {i, 25}], Table[RandomReal[{0.9, 1.9}], {i, 1, 25}]}]];
```

In[209]:=
```
ANOVA[dummydata3, PostTests → {Tukey, Bonferroni}, SignificanceLevel → .05]
BarChart[#[[2]] & /@ (Mean[#] & /@ Split[dummydata3, #1[[1]] == #2[[1]] &]),
 BarValues → True, PlotRange → {0, 1.6}, BarStyle → {Red, Green, Blue, Yellow}]
```

Out[209]=

ANOVA →

	DF	SumOfSq	MeanSq	FRatio	PValue
Model	3	11.9012	3.96707	41.4348	2.88058×10^{-17}
Error	96	9.1913	0.0957427		
Total	99	21.0925			

,

CellMeans →

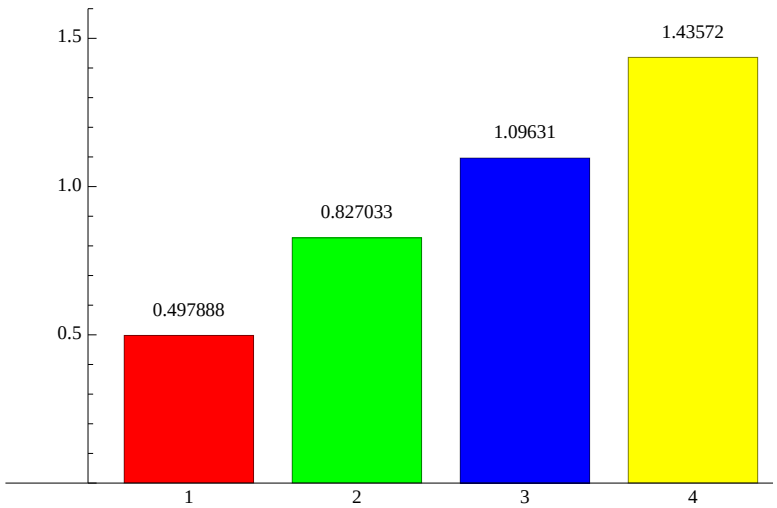
All	0.964239
Model[1]	0.497888
Model[2]	0.827033,
Model[3]	1.09631
Model[4]	1.43572

PostTests → {Model →

Bonferroni	{{1, 2}, {1, 3}, {2, 3}, {1, 4}, {2, 4}, {3, 4}}
Tukey	{{1, 2}, {1, 3}, {2, 3}, {1, 4}, {2, 4}, {3, 4}}

}}

Out[210]=

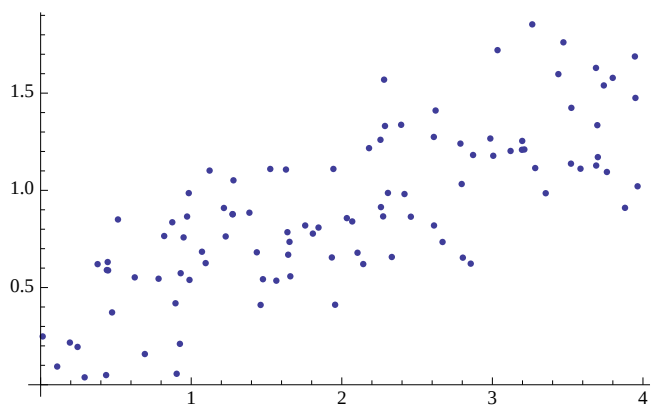


In[211]:=
```
dummydata4 =
 Join[Transpose[{Table[RandomReal[{0, 1}], {i, 25}], Table[RandomReal[{0, 1}], {i, 1, 25}]}],
  Transpose[{Table[RandomReal[{1, 2}], {i, 25}], Table[RandomReal[{0.3, 1.3}], {i, 1, 25}]}],
  Transpose[{Table[RandomReal[{2, 3}], {i, 25}], Table[RandomReal[{0.6, 1.6}], {i, 1, 25}]}],
  Transpose[{Table[RandomReal[{3, 4}], {i, 25}], Table[RandomReal[{0.9, 1.9}], {i, 1, 25}]}]]
```

Out[211]=
```
{{0.439649, 0.589965}, {0.292252, 0.0381561}, {0.972625, 0.865285}, {0.783227, 0.545396},
 {0.434834, 0.0497202}, {0.874311, 0.835761}, {0.245276, 0.194622}, {0.378319, 0.62042},
 {0.895606, 0.419111}, {0.445289, 0.631189}, {0.513565, 0.850192}, {0.988593, 0.539306},
 {0.903431, 0.0565045}, {0.0136069, 0.24855}, {0.109996, 0.0937107}, {0.983695, 0.985757},
 {0.924854, 0.210415}, {0.692075, 0.158052}, {0.19434, 0.216995}, {0.625014, 0.552365},
 {0.448126, 0.588363}, {0.949382, 0.757919}, {0.474449, 0.37213}, {0.820579, 0.765079},
 {0.930026, 0.573716}, {1.27438, 0.87768}, {1.27552, 0.876569}, {1.6389, 0.784772},
 {1.22891, 0.762915}, {1.38641, 0.885002}, {1.62926, 1.10701}, {1.07145, 0.684239},
 {1.64401, 0.668852}, {1.47631, 0.542887}, {1.46088, 0.410545}, {1.43562, 0.681307},
 {1.56524, 0.535368}, {1.95561, 0.411519}, {1.93392, 0.654817}, {1.525, 1.10993},
 {1.09589, 0.625728}, {1.65817, 0.557807}, {1.28049, 1.05156}, {1.21716, 0.909291},
 {1.84531, 0.808609}, {1.75695, 0.81925}, {1.80782, 0.777482}, {1.9442, 1.11007},
 {1.12211, 1.10166}, {1.65279, 0.735035}, {2.66907, 0.734363}, {2.27451, 0.865958},
 {2.61237, 0.8193}, {2.06927, 0.840036}, {2.45851, 0.864342}, {2.85624, 0.622937},
 {2.25681, 1.25996}, {2.62294, 1.41059}, {2.41637, 0.981002}, {2.18092, 1.21708},
 {2.78781, 1.24045}, {2.80378, 0.653719}, {2.30637, 0.986633}, {2.14245, 0.620801},
 {2.79617, 1.03258}, {2.10439, 0.679035}, {2.28085, 1.56931}, {2.98635, 1.26656},
 {2.87178, 1.18178}, {2.61123, 1.27461}, {2.2868, 1.33109}, {2.33216, 0.65714},
 {2.25996, 0.913783}, {2.39412, 1.33713}, {2.03345, 0.857099}, {3.43819, 1.59777},
 {3.76044, 1.09444}, {3.88076, 0.909996}, {3.52452, 1.42436}, {3.94637, 1.68859},
 {3.52188, 1.13725}, {3.28445, 1.11466}, {3.19643, 1.20866}, {3.74008, 1.53965},
 {3.94996, 1.47538}, {3.26481, 1.85371}, {3.58518, 1.11119}, {3.70027, 1.17095},
 {3.68935, 1.12757}, {3.69686, 1.33514}, {3.19815, 1.25411}, {3.4717, 1.7613},
 {3.21233, 1.21025}, {3.68797, 1.62972}, {3.79926, 1.57856}, {3.12086, 1.20253},
 {3.35417, 0.985163}, {3.96425, 1.02085}, {3.00581, 1.17782}, {3.03423, 1.72112}}
```

In[212]:=
```
ListPlot[dummydata4]
```

Out[212]=



In[213]:=
```
? Regress
Regress[dummydata4, {1, x}, x]
```

Regress[data, funs, vars] finds a least-squares fit to a list of data as a linear combination of the functions funs of variables vars. >>

Out[214]=

ParameterTable →

	Estimate	SE	TStat	PValue
1	0.330021	0.0548019	6.02206	3.00794×10^{-8}
x	0.280248	0.0236768	11.8364	1.33755×10^{-20}

RSquared → 0.588407, AdjustedRSquared → 0.584207, EstimatedVariance → 0.0700872,

ANOVATable →

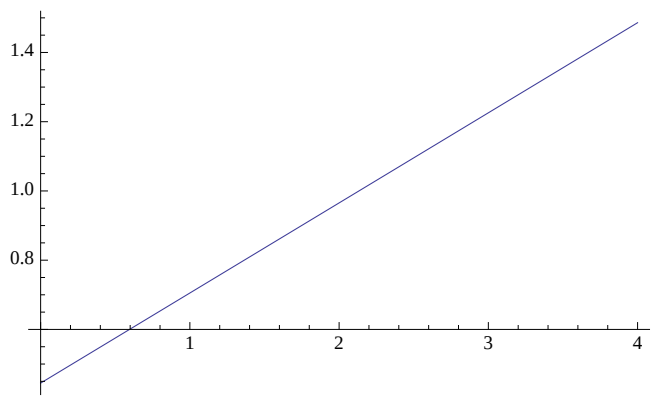
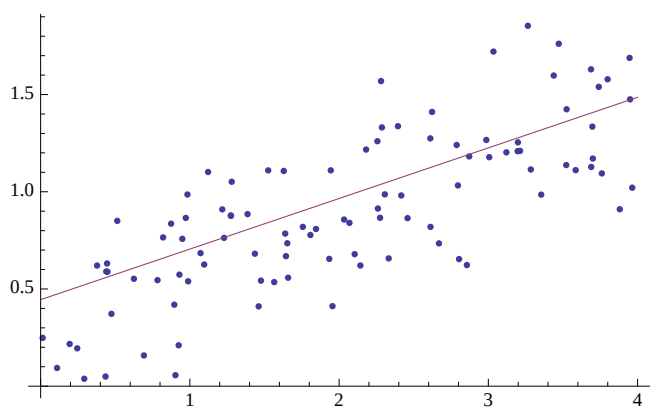
	DF	SumOfSq	MeanSq	FRatio	PValue
Model	1	9.81918	9.81918	140.099	1.33755×10^{-20}
Error	98	6.86855	0.0700872		
Total	99	16.6877			

In[215]:=
```
Regress[dummydata4, {1, x}, x, RegressionReport → {BestFit}]
```

Out[215]= {BestFit → 0.330021 + 0.280248 x}

In[216]:=
```
ListPlot[
 Transpose[{{0, 1, 2, 3, 4}, 0.4449249158794369` + 0.26025177235938046` # & /@ {0, 1, 2, 3, 4}}],
 Joined → True]
ListPlot[{dummydata4, Transpose[{{0, 1, 2, 3, 4},
    0.4449249158794369` + 0.26025177235938046` # & /@ {0, 1, 2, 3, 4}}]}, Joined → {False, True}]
```

Out[216]=



Out[217]=