

MONT 105Q -- Environmental Mathematics

The Correlation Coefficient and Its Geometric Meaning

April 11, 2016

Given a scatter plot, the correlation coefficient r is a way to quantify how close the points (x, y) are to lying along a single straight line. The coefficient takes values in the range -1 to 1 :

A value of $r = -1$: indicates that the points all lie along a line of negative slope

$r = 0$: indicates no linear relationship at all

$r = +1$: indicates that the points all lie along a line of positive slope

Values between 0 and 1 or between -1 and 0 indicate the "strength" of the tendency toward a linear relation between x and y . Here are some illustrative examples:

```
randomize( );
```

1460387638

(1)

```
r1 := rand(1..100000000) :  
      10000000
```

```
CloudPlot := proc(a, b, p)  
local xs, ys, i, theta, s, C;
```

```
with(Statistics) :
```

```
xs := [ ];
```

```
ys := [ ];
```

```
for i to 1000 do
```

```
theta := evalf(2·Pi·r1( ));
```

```
s := r1( );
```

```
xs := [op(xs), s·a·cos(theta) - (1 - p)·s·b·sin(theta)];
```

```
ys := [op(ys), s·b·cos(theta) + (1 - p)·s·a·sin(theta)];
```

```
end do:
```

```
C := Statistics[Correlation](⟨seq(xs[i], i = 1..1000)⟩, ⟨seq(ys[i], i = 1..1000)⟩, ignore);
```

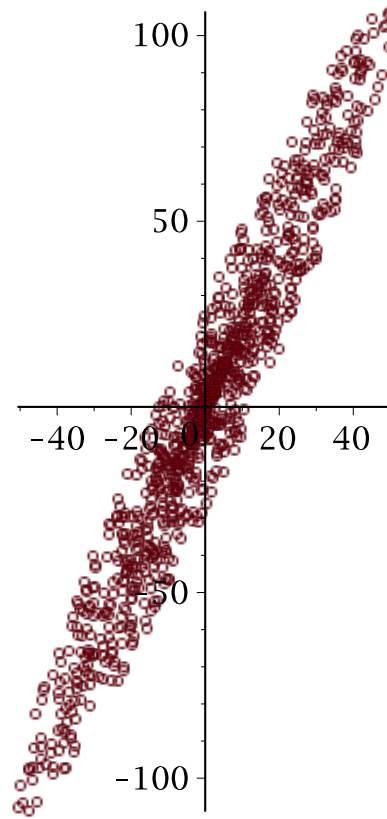
```
print("Correlation coefficient is r = ", C);
```

```
plot([seq([xs[i], ys[i]], i = 1..1000)], style = point, symbol = circle, scaling  
      = constrained);
```

```
end proc:
```

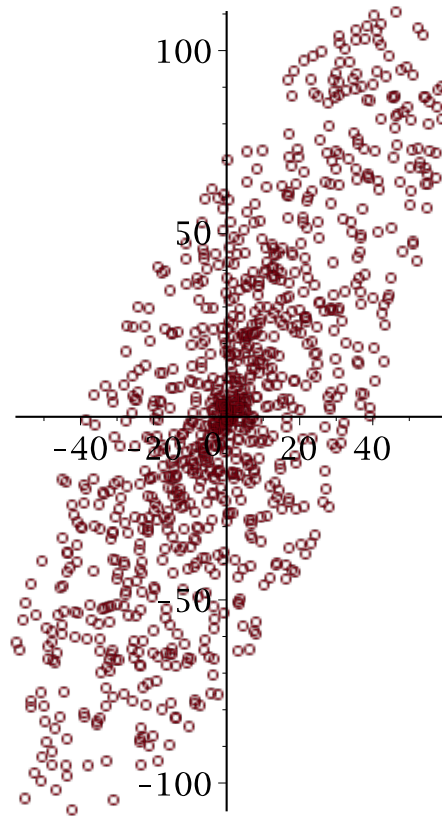
```
CloudPlot(5, 11, .9);
```

"Correlation coefficient is r = ", 0.963413425338172



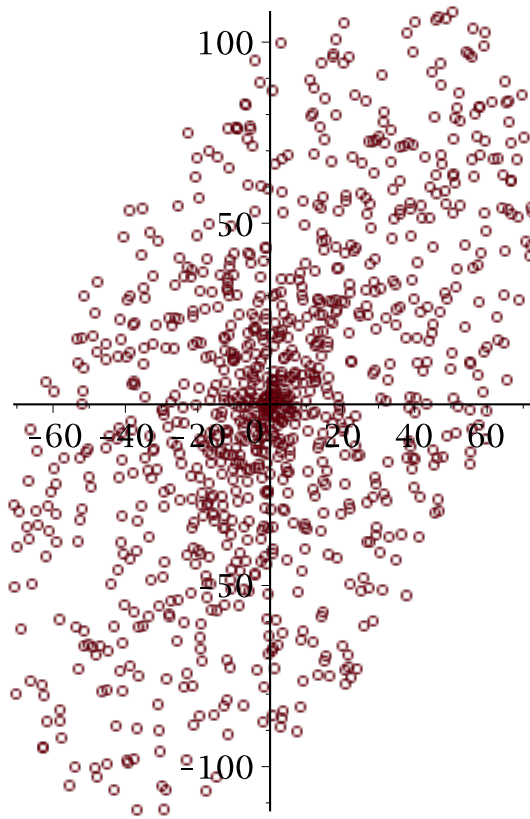
CloudPlot(5, 11, .7);

"Correlation coefficient is r = ", 0.751370183529856



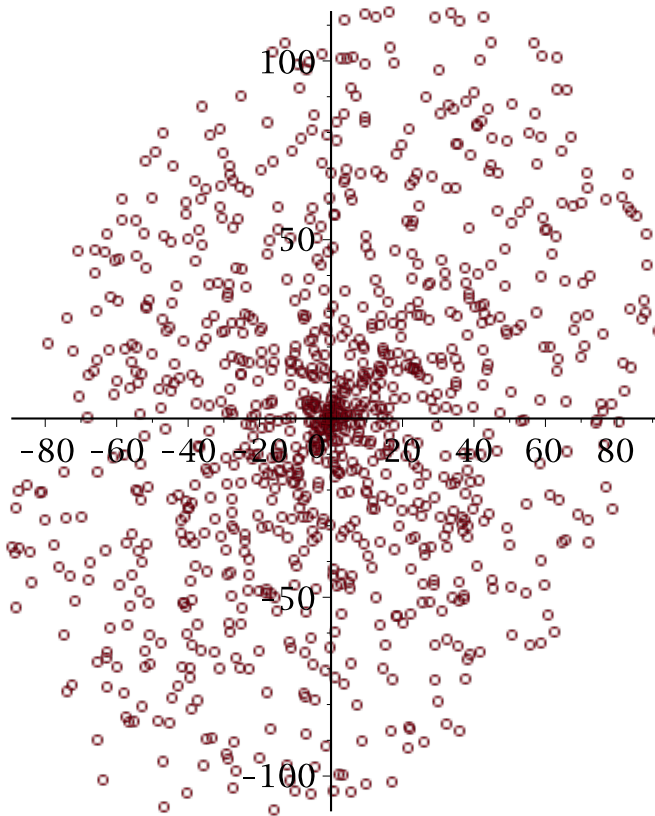
CloudPlot(5, 11, .5);

"Correlation coefficient is r = ", 0.489636268125680



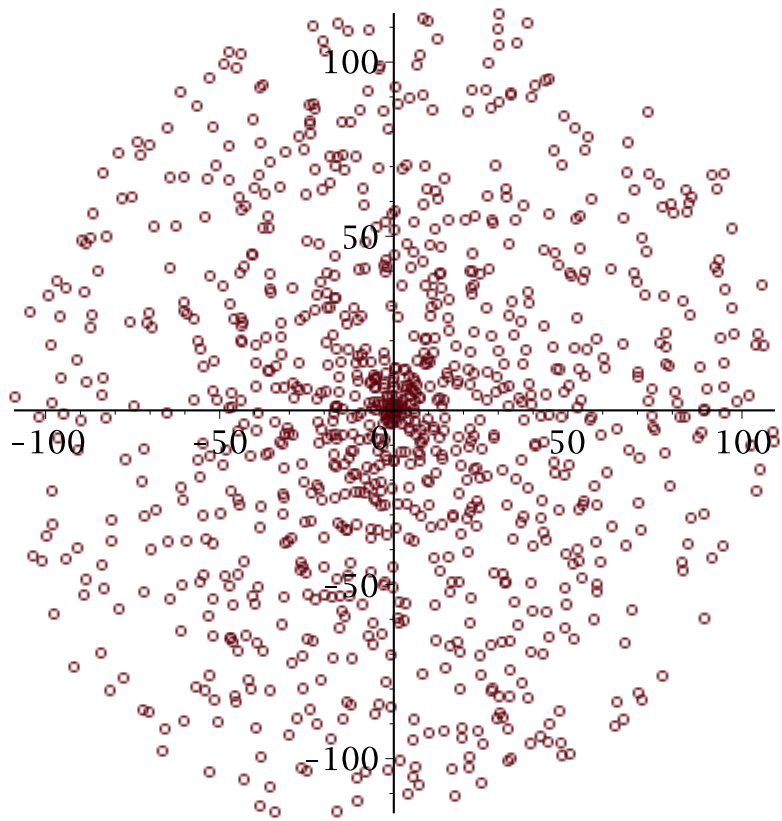
CloudPlot(5, 11, .3);

"Correlation coefficient is r = ", 0.228691233149027



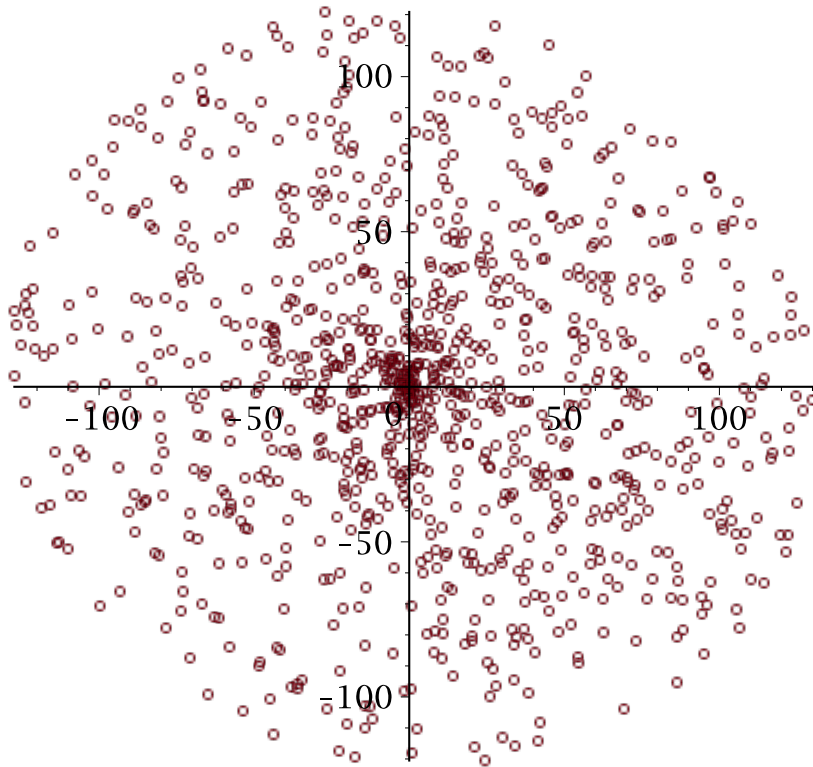
CloudPlot(5, 11, .1);

"Correlation coefficient is r = ", 0.0318756828673633



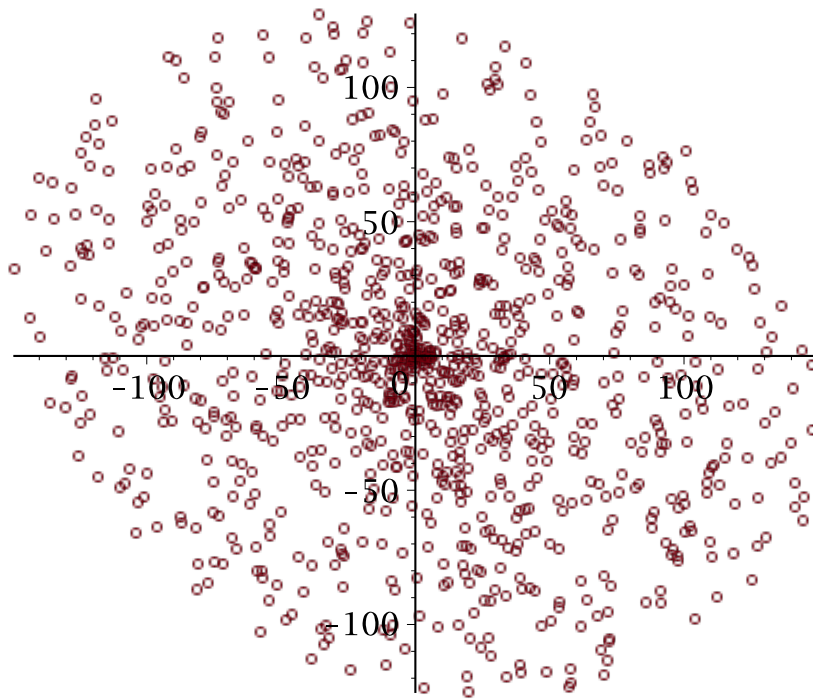
CloudPlot(5, 11, -.1);

"Correlation coefficient is r = ", -0.0991571074071821



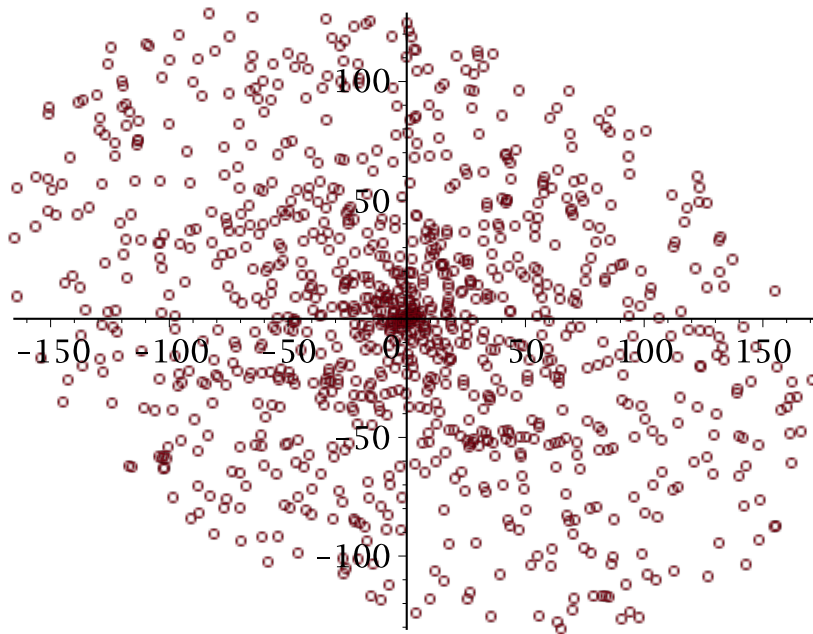
CloudPlot(5, 11, -0.3);

"Correlation coefficient is r = ", -0.211842333386389



CloudPlot(5, 11, -.5);

"Correlation coefficient is r = ", -0.253323814746533



CloudPlot(5, 11, -4);

"Correlation coefficient is r = ", -0.867400252785737

