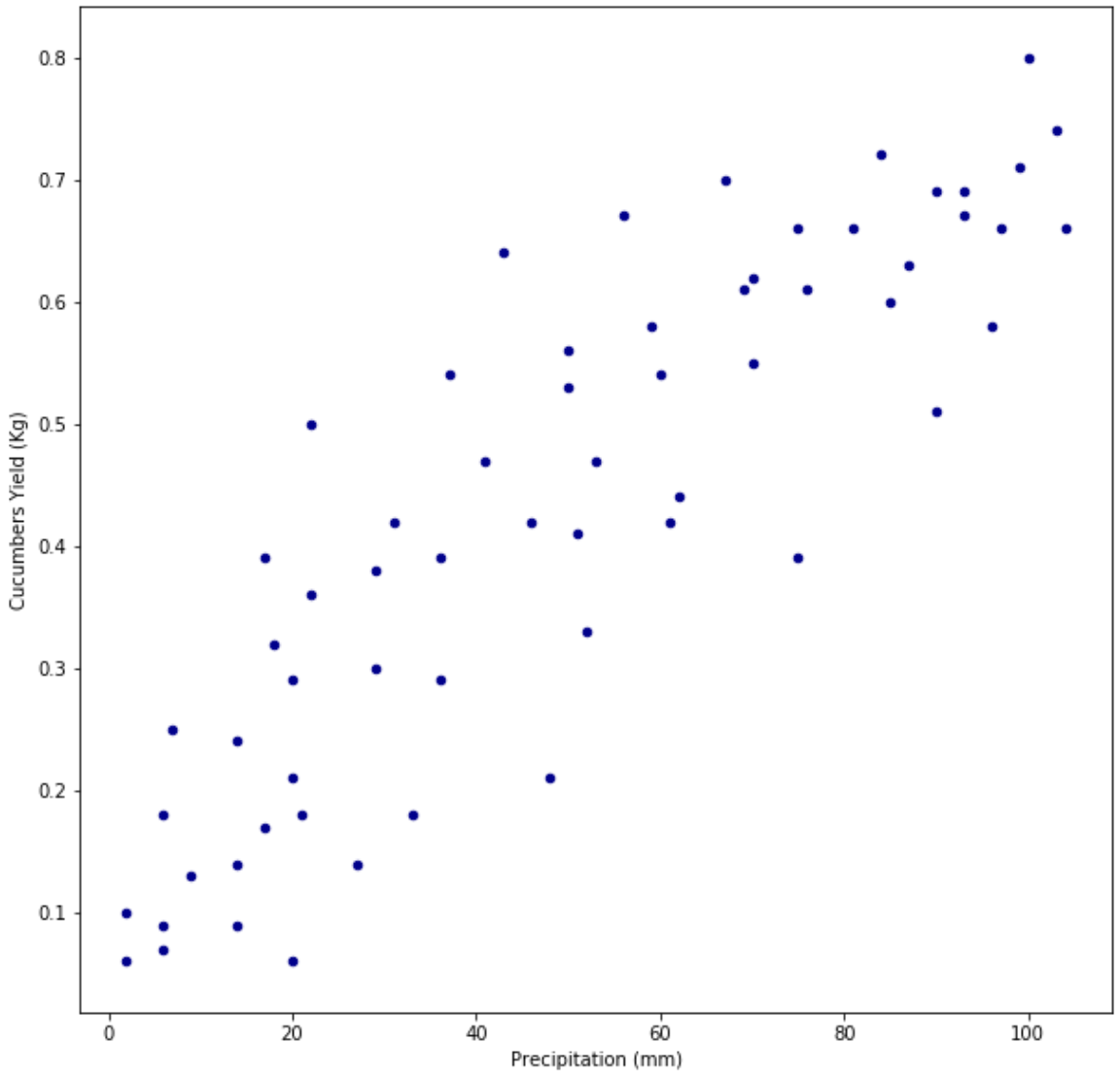
In [1]:
```
# Read Cucumber Data
import pandas as pd
from scipy import stats
import numpy as np
import matplotlib.pyplot as plt

data = pd.read_excel('TestData.xlsx', sheet_name=1)
data
```

Out[1]:

| | Precipitation (mm) | Cucumbers Yield (Kg) |
|---|---|---|
| 0 | 22 | 0.36 |
| 1 | 6 | 0.09 |
| 2 | 93 | 0.67 |
| 3 | 62 | 0.44 |
| 4 | 84 | 0.72 |
| ... | ... | ... |
| 57 | 97 | 0.66 |
| 58 | 33 | 0.18 |
| 59 | 20 | 0.06 |
| 60 | 96 | 0.58 |
| 61 | 61 | 0.42 |

62 rows × 2 columns

In [2]:
```
ax1 = data.plot.scatter(x="Precipitation (mm)",y="Cucumbers Yield (Kg)",c='DarkBlue', figsize=(10,10))
```

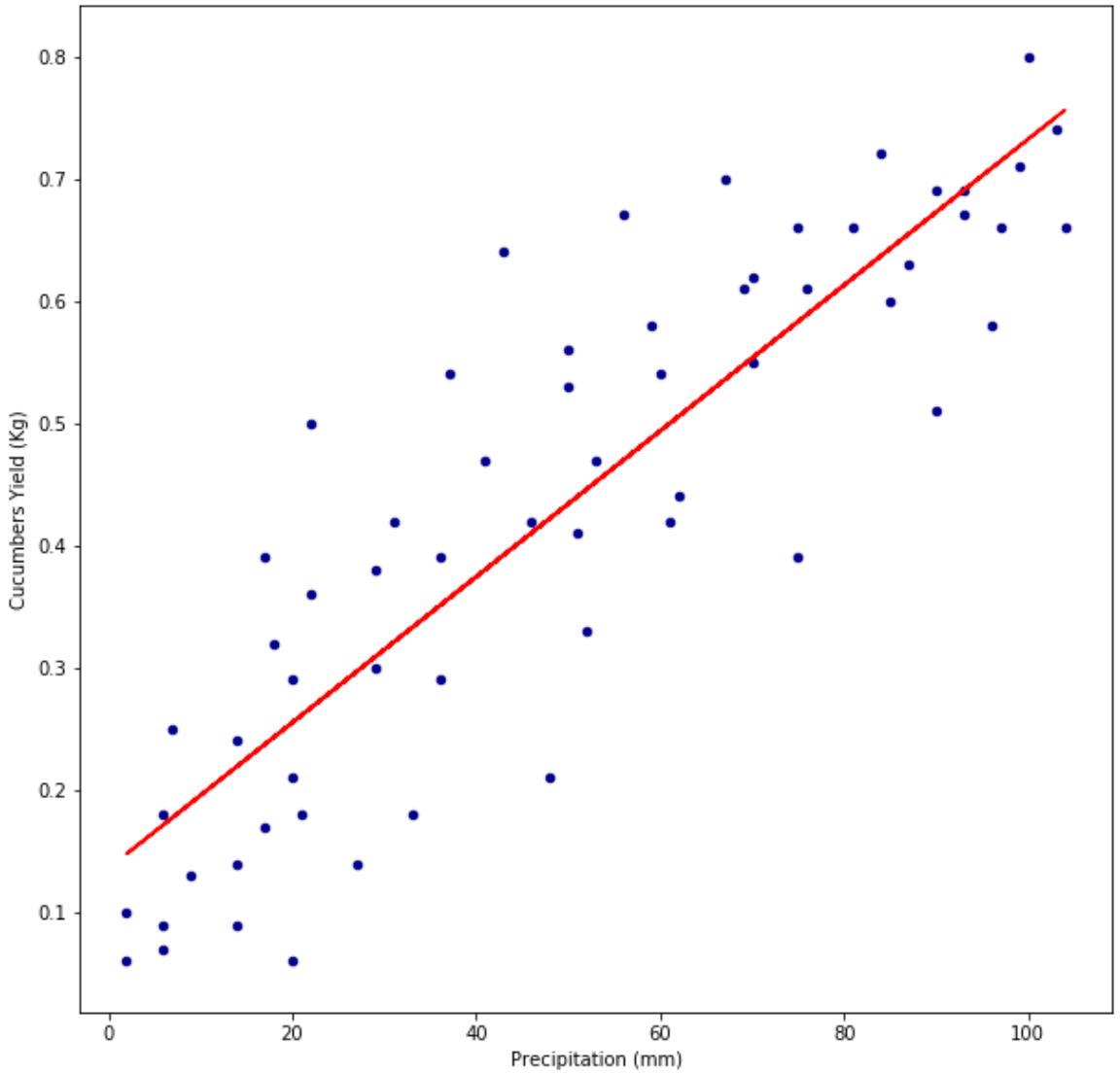



In [3]:
```
r, pvalue = stats.pearsonr(data["Precipitation (mm)"], data["Cucumbers Yield (Kg)"])
print("r:", r)
print("pvalue:", pvalue)
```

```
r: 0.8708738116424835
pvalue: 3.6363048335179605e-20
```

In [4]:
```
ax1 = data.plot.scatter(x="Precipitation (mm)",y="Cucumbers Yield (Kg)",c='DarkBlu
e', figsize=(10,10))
x = data["Precipitation (mm)"]
y = data["Cucumbers Yield (Kg)"]

regressStats = stats.linregress(x, y)
m = regressStats.slope
b = regressStats.intercept
plt.plot(x, m * x + b, color="red")
```

Out[4]:
```
[<matplotlib.lines.Line2D at 0x2bcf129add8>]
```




In [5]:
```
regressStats
```

Out[5]:
```
LinregressResult(slope=0.0059650438630961256, intercept=0.13581695441441483, rvalu
e=0.8708738116424838, pvalue=3.636304833517717e-20, stderr=0.00043462249343379037)
```

In [6]:
```
print("R-squared: %f" % regressStats.rvalue**2)
```

```
R-squared: 0.758421
```

In [ ]:
```

```