In [36]:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

In [37]:

```
#Reading dataset in jupyter
df = pd.read_csv('cars (1) (1).csv')
```

In [38]:

```
#To understand and see the nature of data
df.head()
```

Out[38]:

	Unnamed: 0	speed	dist
0	1	4	2
1	2	4	10
2	3	7	4
3	4	7	22
4	5	8	16

In [4]:

```
df
```

Out[4]:

	Unnamed: 0	speed	dist
0	1	4	2
1	2	4	10
2	3	7	4
3	4	7	22
4	5	8	16
5	6	9	10
6	7	10	18
7	8	10	26
8	9	10	34
9	10	11	17
10	11	11	28
11	12	12	14
12	13	12	20
13	14	12	24
14	15	12	28
15	16	13	26
16	17	13	34
17	18	13	34
18	19	13	46
19	20	14	26
20	21	14	36
21	22	14	60

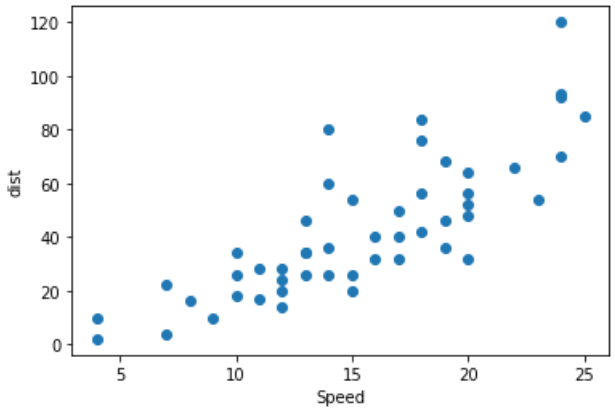
	Unnamed: 0	speed	dist
22	23	14	60
23	24	15	20
24	25	15	26
25	26	15	54
26	27	16	32
27	28	16	40
28	29	17	32
29	30	17	40
30	31	17	50
31	32	18	42
32	33	18	56
33	34	18	76
34	35	18	84
35	36	19	36
36	37	19	46
37	38	19	68
38	39	20	32
39	40	20	48
40	41	20	52
41	42	20	56
42	43	20	64
43	44	22	66
44	45	23	54
45	46	24	70
46	47	24	92
47	48	24	93
48	49	24	120
49	50	25	85

In [39]:

```
#scatter plot for eisting dataset
plt.scatter(df["speed"],df["dist"])
plt.xlabel('Speed')
plt.ylabel('dist')
```

Out[39]:

```
Text(0, 0.5, 'dist')
```



In [40]:

```
x = df["speed"].values
```

```
y = df["dist"].values
```

In [41]:

```
#reshaping data
x = x.reshape(-1,1)
y = y.reshape(-1,1)
```

In [42]:

```
from sklearn.model_selection import train_test_split
```

In [43]:

```
#splitting data into train and test set
x_train,x_test,y_train,y_test = train_test_split(x,y,test_size = 0.3)
```

In [44]:

```
len(x_train)
```

Out[44]:

```
35
```

In [33]:

```
len(x_test)
```

Out[33]:

```
15
```

In [10]:

```
x_train
```

Out[10]:

```
array([[10],
       [18],
       [24],
       [15],
       [15],
       [16],
       [ 8],
       [20],
       [14],
       [ 9],
       [15],
       [12],
       [24],
       [14],
       [12],
       [10],
       [14],
       [12],
       [13],
       [20],
       [17],
       [19],
       [19],
       [16],
       [18],
       [12],
       [ 4],
       [20],
       [ 4],
       [17],
       [17],
```

```
       [23],
       [24],
       [20],
       [ 7]], dtype=int64)
```

In [11]:

```
x_train.ndim
```

Out[11]:

```
2
```

In [45]:

```
 x_test
```

Out[45]:

```
array([[13],
       [ 7],
       [20],
       [ 9],
       [14],
       [22],
       [15],
       [20],
       [12],
       [13],
       [12],
       [23],
       [11],
       [18],
       [25]], dtype=int64)
```

In [13]:

```
from sklearn.linear_model import LinearRegression
model = LinearRegression()
```

In [46]:

```
#running regression model
model.fit(df[['speed']],df.dist)
```

Out[46]:

```
LinearRegression()
```

In [47]:

```
model.score(x_test,y_test)
```

Out[47]:

```
0.6191669680194729
```

In [48]:

```
#measuring accuracy of the model
model.score(x_test,y_test)
```

Out[48]:

```
0.6191669680194729
```

In [49]:

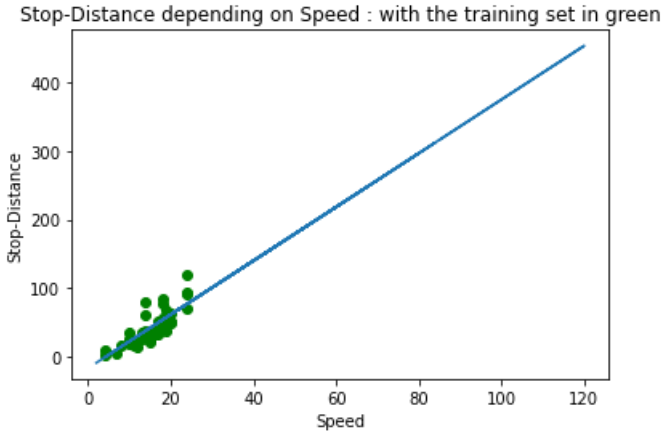
```
 model.predict(x_test)
```

Out[49]:

```
array([33.54221898,  9.94776642, 61.06908029, 17.81258394, 37.47462774,
       68.93389781, 41.4070365 , 61.06908029, 29.60981022, 33.54221898,
       29.60981022, 72.86630657, 25.67740146, 53.20426277, 80.73112409])
```

In [50]:

```
plt.scatter(x_train,y_train, color='green')
plt.plot(df.dist,model.predict(df[['dist']]))
plt.title('Stop-Distance depending on Speed : with the training set in green')
plt.xlabel('Speed')
plt.ylabel('Stop-Distance')
plt.show()
```



In [51]:

```
model.predict([[100]])
```

Out[51]:

```
array([375.66178102])
```

In [52]:

```
# y= mx+c
model.coef_
```

Out[52]:

```
array([3.93240876])
```

In [53]:

```
model.intercept_
```

Out[53]:

```
-17.57909489051095
```

In [54]:

```
y = 3.93240876*100 +-17.57909489051095
```

In [55]:

```
y
```

Out[55]:

```
375.6617811094891
```

In [ ]:

```
```