In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

In [7]:
```
path=r"C:\Users\Mrityunjay\Desktop\Data science naresh it\Class notes by me\EDA_\ba
bank_df=pd.read_csv(path,sep=";")
bank_df
```

Out[7]:

| | age | job | marital | education | default | balance | housing | loan | contact | da |
|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 30 | unemployed | married | primary | no | 1787 | no | no | cellular | |
| **1** | 33 | services | married | secondary | no | 4789 | yes | yes | cellular | |
| **2** | 35 | management | single | tertiary | no | 1350 | yes | no | cellular | |
| **3** | 30 | management | married | tertiary | no | 1476 | yes | yes | unknown | |
| **4** | 59 | blue-collar | married | secondary | no | 0 | yes | no | unknown | |
| **...** | ... | ... | ... | ... | ... | ... | ... | ... | ... | |
| **4516** | 33 | services | married | secondary | no | -333 | yes | no | cellular | |
| **4517** | 57 | self-employed | married | tertiary | yes | -3313 | yes | yes | unknown | |
| **4518** | 57 | technician | married | secondary | no | 295 | no | no | cellular | |
| **4519** | 28 | blue-collar | married | secondary | no | 1137 | no | no | cellular | |
| **4520** | 44 | entrepreneur | single | tertiary | no | 1136 | yes | yes | cellular | |

4521 rows × 17 columns

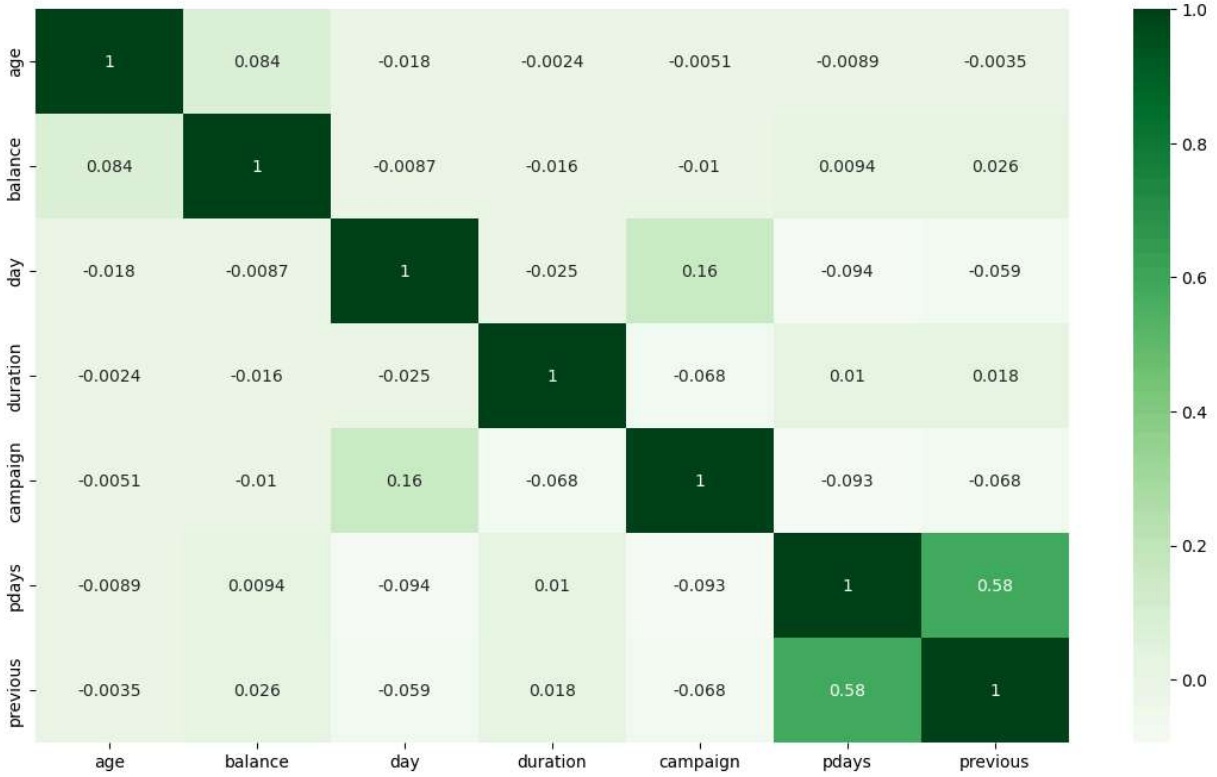
In [11]:
```
dfs=bank_df.corr(numeric_only=True)
dfs
```

Out[11]:

| | age | balance | day | duration | campaign | pdays | previous |
|---|---|---|---|---|---|---|---|
| **age** | 1.000000 | 0.083820 | -0.017853 | -0.002367 | -0.005148 | -0.008894 | -0.003511 |
| **balance** | 0.083820 | 1.000000 | -0.008677 | -0.015950 | -0.009976 | 0.009437 | 0.026196 |
| **day** | -0.017853 | -0.008677 | 1.000000 | -0.024629 | 0.160706 | -0.094352 | -0.059114 |
| **duration** | -0.002367 | -0.015950 | -0.024629 | 1.000000 | -0.068382 | 0.010380 | 0.018080 |
| **campaign** | -0.005148 | -0.009976 | 0.160706 | -0.068382 | 1.000000 | -0.093137 | -0.067833 |
| **pdays** | -0.008894 | 0.009437 | -0.094352 | 0.010380 | -0.093137 | 1.000000 | 0.577562 |
| **previous** | -0.003511 | 0.026196 | -0.059114 | 0.018080 | -0.067833 | 0.577562 | 1.000000 |

In [17]:
```
#heatmap for seborn package
plt.figure(figsize=(14,8))
sns.heatmap(dfs,annot=True,cmap="Greens")
```

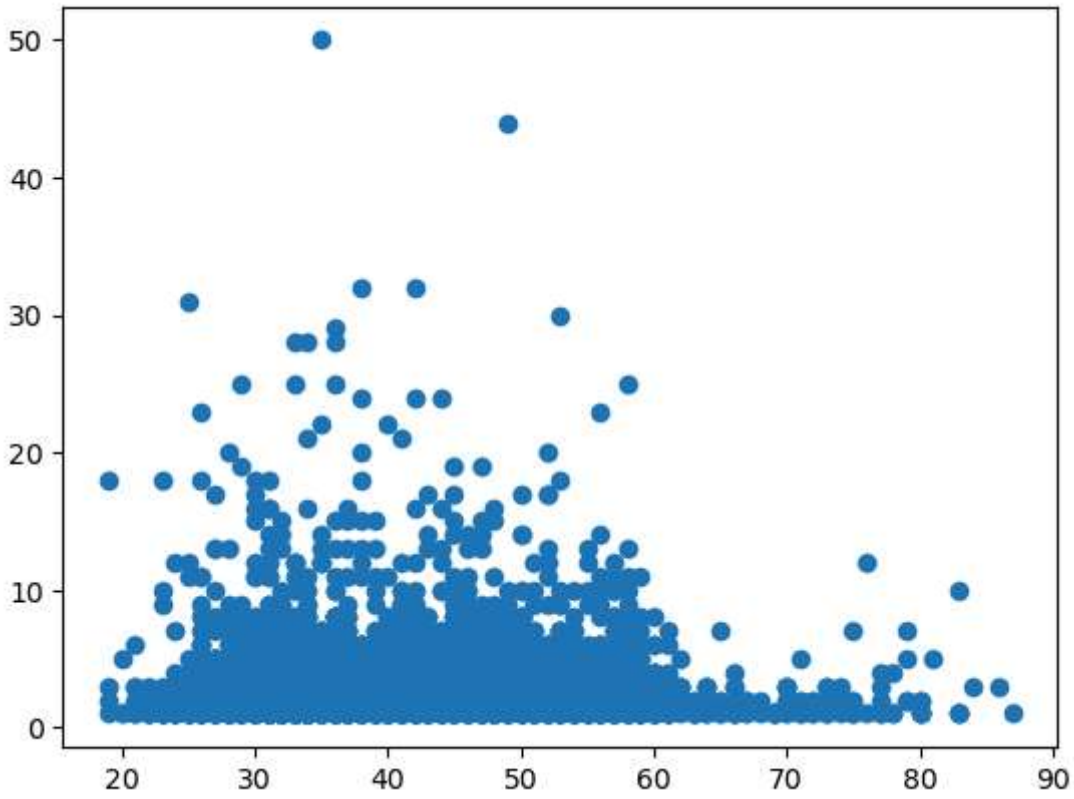
Out[17]: <Axes: >




In [21]:
```
#sactter plot
indx=bank_df["age"]
colm=bank_df["campaign"]

plt.scatter(indx,colm)
```

Out[21]: <matplotlib.collections.PathCollection at 0x186b7b5f5c0>




In [ ]: