In [34]:

```
import seaborn as sns
import matplotlib.pyplot as plt
import pandas as pd
```

In [36]:

```
file_path = r"C:\Users\vinee\Downloads\myexcel - myexcel.csv.csv"
dataset= pd.read_csv(file_path)
```

In [37]:

```
print(dataset.columns)
```

```
Index(['Name', 'Team', 'Number', 'Position', 'Age', 'Height', 'Weight',
       'College', 'Salary'],
      dtype='object')
```

1.How many are there in each team and the percentage splitting with respect to the total employees.

In [38]:

```
# number of employee in each team
team_count = dataset['Team'].value_counts()
team_count
```

Out[38]:

```
New Orleans Pelicans      19
Memphis Grizzlies         18
Utah Jazz                 16
New York Knicks           16
Milwaukee Bucks           16
Brooklyn Nets             15
Portland Trail Blazers    15
Oklahoma City Thunder     15
Denver Nuggets            15
Washington Wizards        15
Miami Heat                15
Charlotte Hornets         15
Atlanta Hawks             15
San Antonio Spurs         15
Houston Rockets           15
Boston Celtics            15
Indiana Pacers            15
Detroit Pistons           15
Cleveland Cavaliers       15
Chicago Bulls             15
Sacramento Kings          15
Phoenix Suns              15
Los Angeles Lakers        15
Los Angeles Clippers      15
Golden State Warriors     15
Toronto Raptors           15
Philadelphia 76ers        15
Dallas Mavericks          15
Orlando Magic             14
Minnesota Timberwolves    14
Name: Team, dtype: int64
```

In [39]:

```
# the percentage splitting with respect to the total employees
total_employees = len(dataset)
team_percentage = (team_count / total_employees) * 100
a = pd.DataFrame({'Number of Employees': team_count,'Percentage of Employees': team_p
print("percentage splitting with respect to the total employees")
print(a)
```

```
percentage splitting with respect to the total employees
                        Number of Employees  Percentage of Employees
New Orleans Pelicans                     19                4.148472
Memphis Grizzlies                        18                3.930131
Utah Jazz                                16                3.493450
New York Knicks                          16                3.493450
Milwaukee Bucks                          16                3.493450
Brooklyn Nets                            15                3.275109
Portland Trail Blazers                   15                3.275109
Oklahoma City Thunder                    15                3.275109
Denver Nuggets                           15                3.275109
Washington Wizards                       15                3.275109
Miami Heat                               15                3.275109
Charlotte Hornets                        15                3.275109
Atlanta Hawks                            15                3.275109
San Antonio Spurs                        15                3.275109
Houston Rockets                          15                3.275109
Boston Celtics                           15                3.275109
Indiana Pacers                           15                3.275109
Detroit Pistons                          15                3.275109
Cleveland Cavaliers                      15                3.275109
Chicago Bulls                            15                3.275109
Sacramento Kings                         15                3.275109
Phoenix Suns                             15                3.275109
Los Angeles Lakers                       15                3.275109
Los Angeles Clippers                     15                3.275109
Golden State Warriors                    15                3.275109
Toronto Raptors                          15                3.275109
Philadelphia 76ers                       15                3.275109
Dallas Mavericks                         15                3.275109
Orlando Magic                            14                3.056769
Minnesota Timberwolves                   14                3.056769
```

2.Segregate the employees w.r.t different positions.

In [40]:

```
#the employees by their positions
pos = dataset.groupby('Position')
pos_count = pos.size()
b = pd.DataFrame({'Number of Employees': pos_count})
print("Employees w.r.t different Positions:")
print(b)
```

```
Employees w.r.t different Positions:
          Number of Employees
Position
C                          79
PF                        100
PG                         92
SF                         85
SG                        102
```

3.Find from which age group most of the employees belong to.

In [41]:

```
age_group_count = dataset['Age'].value_counts()
age_groups = [(18, 21), (21, 24), (24, 27), (27, 30), (30, 33), (33, 36)]
age_grp= ['18-21', '21-24', '24-27', '27-30', '30-33', '33-36']
# using cut func to Categorize the employees into age groups
d = pd.cut(dataset['Age'], bins=len(age_groups), labels=age_grp)
#no_of eomplyees in each age group
age_group_counts = d.value_counts()
print("Number of Employees in  Age Group:")
age_group_counts
```

```
Number of Employees in  Age Group:
```

Out[41]:

```
21-24    170
24-27    100
27-30     80
18-21     66
30-33     29
33-36     13
Name: Age, dtype: int64
```

In [42]:

```
#  age group with most of employees
max_age = age_group_counts.idxmax()
print("Age Group with Most Employees:", max_age)
```

```
Age Group with Most Employees: 21-24
```

4.Find out under which team and positions,spending in terms of salary is high

In [43]:

```
team_spending = dataset.groupby(['Team', 'Position'])['Salary'].sum()
#team and position with the highest spending
high_spend = team_spending.idxmax()
high_spend
print("Team with Highest Spending:", high_spend[0])
print("Position with Highest Spending:", high_spend[1])
```

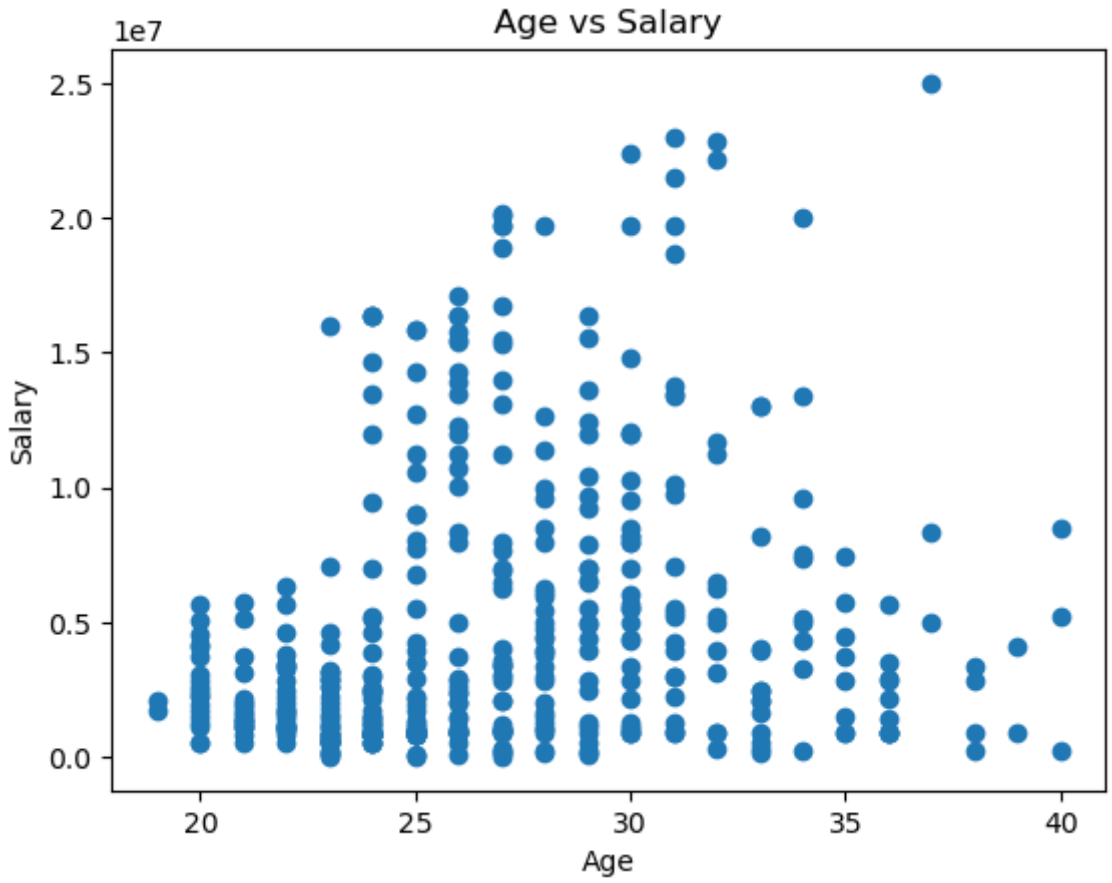
```
Team with Highest Spending: Los Angeles Lakers
Position with Highest Spending: SF
```

5. Find if there is any correlation between age and salary,represent it visually.

In [45]:

```
import matplotlib.pyplot as plt
correlation = dataset['Age'].corr(dataset['Salary'])
print("Correlation between Age and Salary:",correlation)
plt.scatter(dataset['Age'], dataset['Salary'])
plt.xlabel('Age')
plt.ylabel('Salary')
plt.title('Age vs Salary')
plt.show()
```

```
Correlation between Age and Salary: 0.21400941226570974
```



In [ ]: