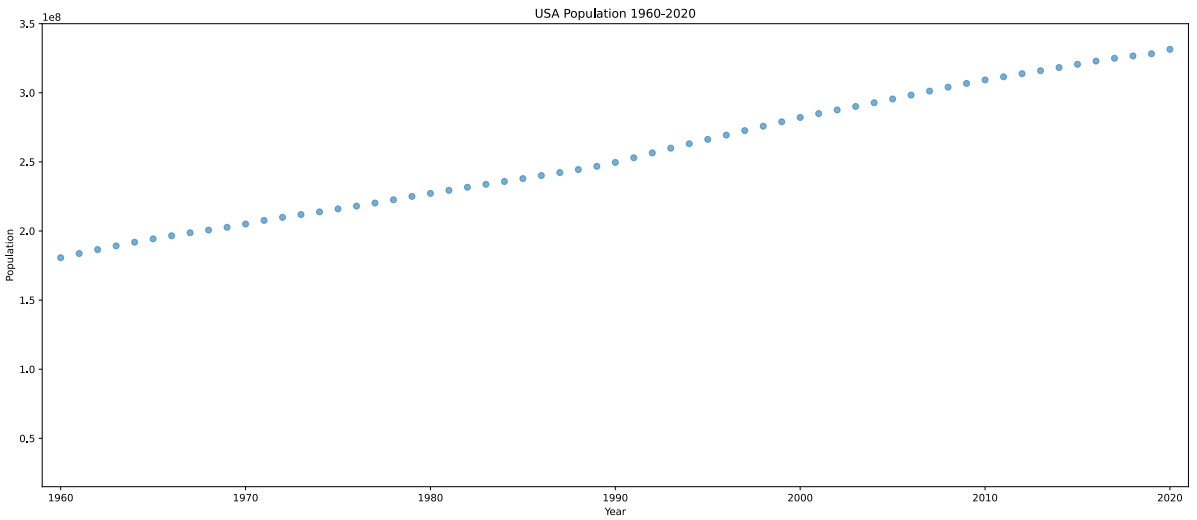
In [ ]:
```
import pandas as pd
import matplotlib.pyplot as mplib
from sklearn.linear_model import LinearRegression
```

In [ ]:
```
Data = pd.read_csv('USA_Population_1960-2020.csv')
```

In [ ]:
```
Year = pd.DataFrame(Data,columns=["Year"])
Population = pd.DataFrame(Data,columns=["Population"])
```

In [ ]:
```
mplib.figure(figsize=(20.5,8.5))
mplib.scatter(Year,Population,alpha=0.6)
mplib.title("USA Population 1960-2020")
mplib.xlabel("Year")
mplib.ylabel("Population")
mplib.ylim(15000000,350000000)
mplib.xlim(1959,2021)
mplib.style.use("fivethirtyeight")
```



In [ ]:
```
LinearRegression  = LinearRegression()
LinearRegression.fit(Year,Population)
LinearRegression.coef_ # returns slope coeficent theta_subcript_1
```

Out[ ]:
```
array([[2566194.82797462]])
```

In [ ]:
```
LinearRegression.intercept_
```

Out[ ]:
```
array([-4.85168057e+09])
```

In [ ]:
```
LinearRegression.score(Year,Population)
```

Out[ ]:
```
0.9965160200489581
```

In [ ]:
```
def population(year):
    population = -4.85168057e+09 +(2566194.82797462 * year)
    return int(population)
```

In [ ]:
```
print(population(2021))
```

```
334599177
```