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

```
df=pd.read_csv('china_gdp.csv')
df.head()
```

	Year	Value
0	1960	5.918412e+10
1	1961	4.955705e+10
2	1962	4.668518e+10
3	1963	5.009730e+10
4	1964	5.906225e+10

Next steps: Generate code with df | New interactive sheet

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 55 entries, 0 to 54
Data columns (total 2 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   Year    55 non-null     int64
 1   Value   55 non-null     float64
dtypes: float64(1), int64(1)
memory usage: 1012.0 bytes
```

```
x = df["Year"].values
y = df["Value"].values
```

```
def logistic(x, L, x0, k):
    return L / (1 + np.exp(-k * (x - x0)))
```

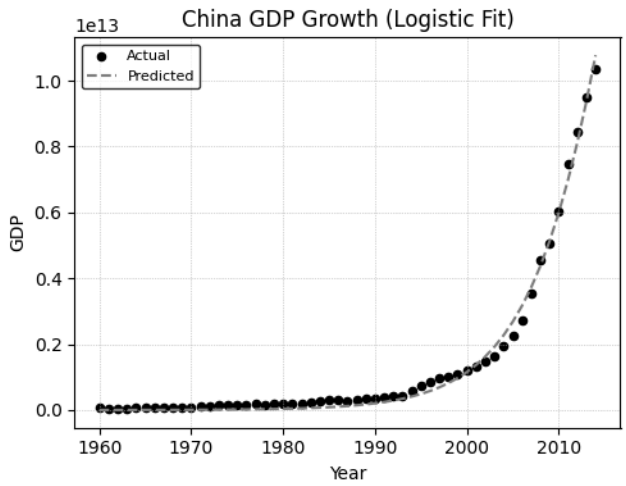
```
from scipy.optimize import curve_fit
popt, _ = curve_fit(logistic, x, y, p0=[1e13, 1990, 0.03])
y_pred = logistic(x, *popt)
```

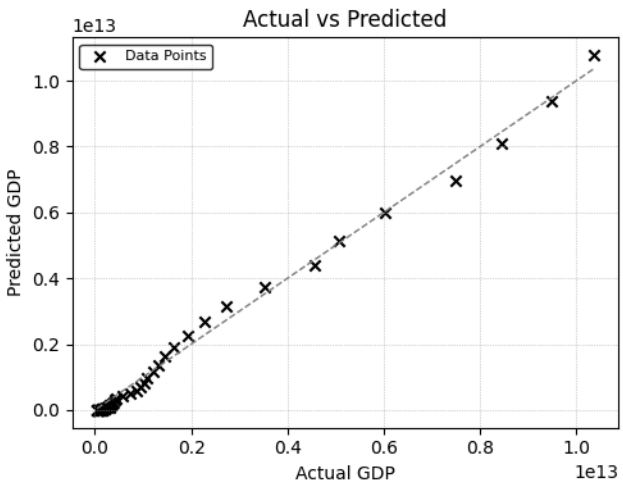
```
from sklearn.metrics import r2_score
r2 = r2_score(y, y_pred)
print(f"R² Score: {r2:.4f}")
```

```
R² Score: 0.9938
```

```
plt.figure(figsize=(10,4))


plt.subplot(1,2,1)
plt.scatter(x, y, marker='o', color='black', label='Actual', s=20)
plt.plot(x, y_pred, linestyle='--', color='gray', label='Predicted', linewidth=1.5)
plt.xlabel("Year")
plt.ylabel("GDP")
plt.title("China GDP Growth (Logistic Fit)")
plt.legend(edgecolor='black', fontsize=8)
plt.grid(True, linestyle=':', linewidth=0.5)


plt.subplot(1,2,2)
plt.scatter(y, y_pred, marker='x', color='black', label='Data Points')
plt.plot([y.min(), y.max()], [y.min(), y.max()], linestyle='--', color='gray', linewidth=1)
plt.xlabel("Actual GDP")
plt.ylabel("Predicted GDP")
plt.title("Actual vs Predicted")
plt.legend(edgecolor='black', fontsize=8)
plt.grid(True, linestyle=':', linewidth=0.5)

plt.tight_layout()
```

```
plt.show()
```


China GDP Growth (Logistic Fit)
1e13
Actual
Predicted
GDP
Year
1960
1970
1980
1990
2000
2010
0.0
0.2
0.4
0.6
0.8
1.0


Actual vs Predicted
1e13
Data Points
Predicted GDP
Actual GDP
1e13
0.0
0.2
0.4
0.6
0.8
1.0