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

In [2]:
```
data = pd.read_csv("Position_Salaries.csv")
```

In [3]:
```
data
```

Out[3]:

	Position	Level	Salary
0	Business Analyst	1	45000
1	Junior Consultant	2	50000
2	Senior Consultant	3	60000
3	Manager	4	80000
4	Country Manager	5	110000
5	Region Manager	6	150000
6	Partner	7	200000
7	Senior Partner	8	300000
8	C-level	9	500000
9	CEO	10	1000000

In [5]:
```
y = data["Salary"]
```

In [6]:
```
#method two
data.columns
```

Out[6]:
```
Index(['Position', 'Level', 'Salary'], dtype='object')
```

In [8]:
```
#method two
y1 = data.iloc[:, 2].values
```

In [10]:
```
y1
```

Out[10]:
```
array([  45000,   50000,   60000,   80000,  110000,  150000,  200000,
        300000,  500000, 1000000], dtype=int64)
```

In [ ]:

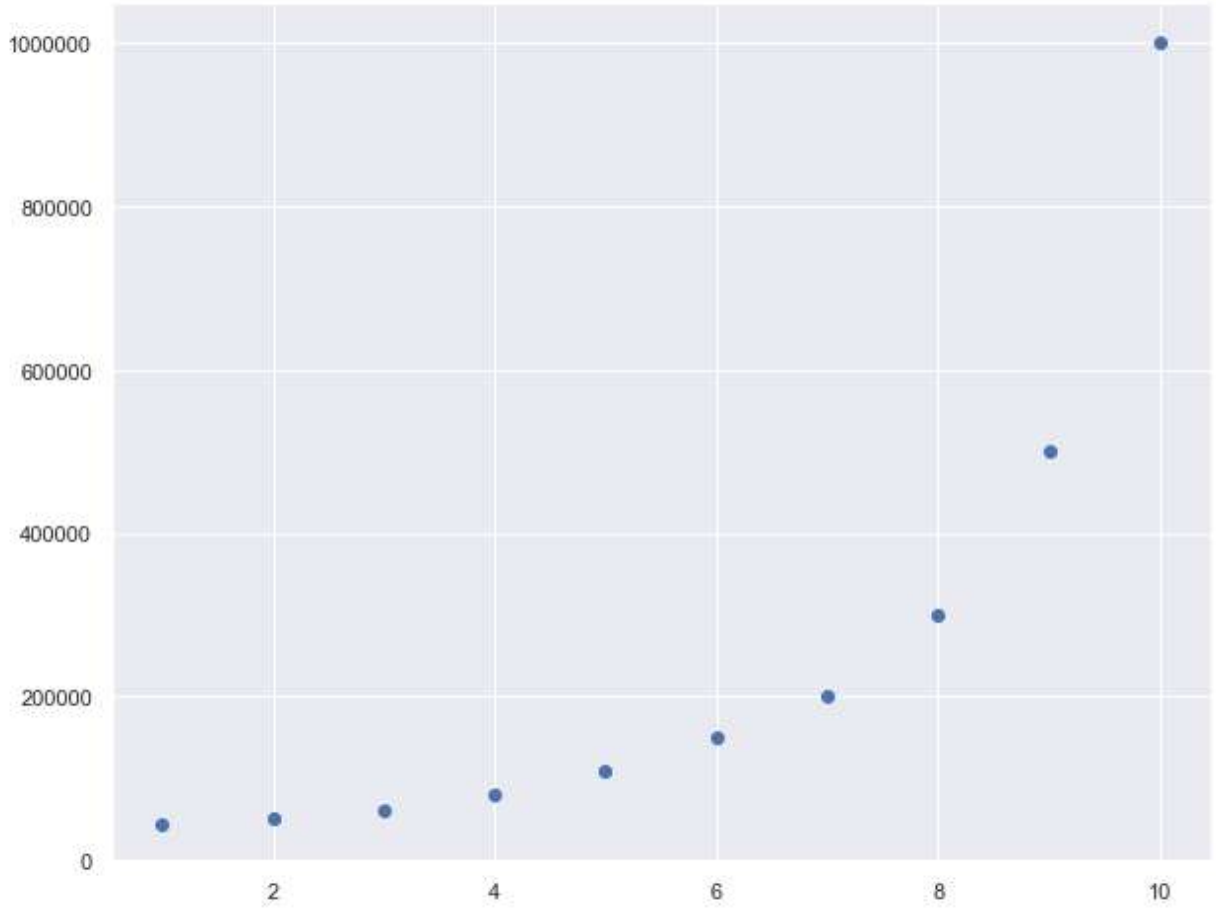
In [ ]:

In [32]:
```
x = data.iloc[:, 1]
```

In [33]:
```
x
```

Out[33]:
```
0     1
1     2
2     3
3     4
4     5
5     6
6     7
7     8
8     9
9    10
Name: Level, dtype: int64
```

In [34]:
```
plt.figure(figsize = (10,8))
plt.scatter(x, y1)
plt.show()
```



In [35]:
```
import seaborn as sns
```

In [36]:
```
sns.set()
```

In [37]:
```
from sklearn.linear_model import LinearRegression
```

In [38]:
```
model = LinearRegression()
```

In [40]:
```
y = pd.DataFrame(y)
x = pd.DataFrame(x)
```

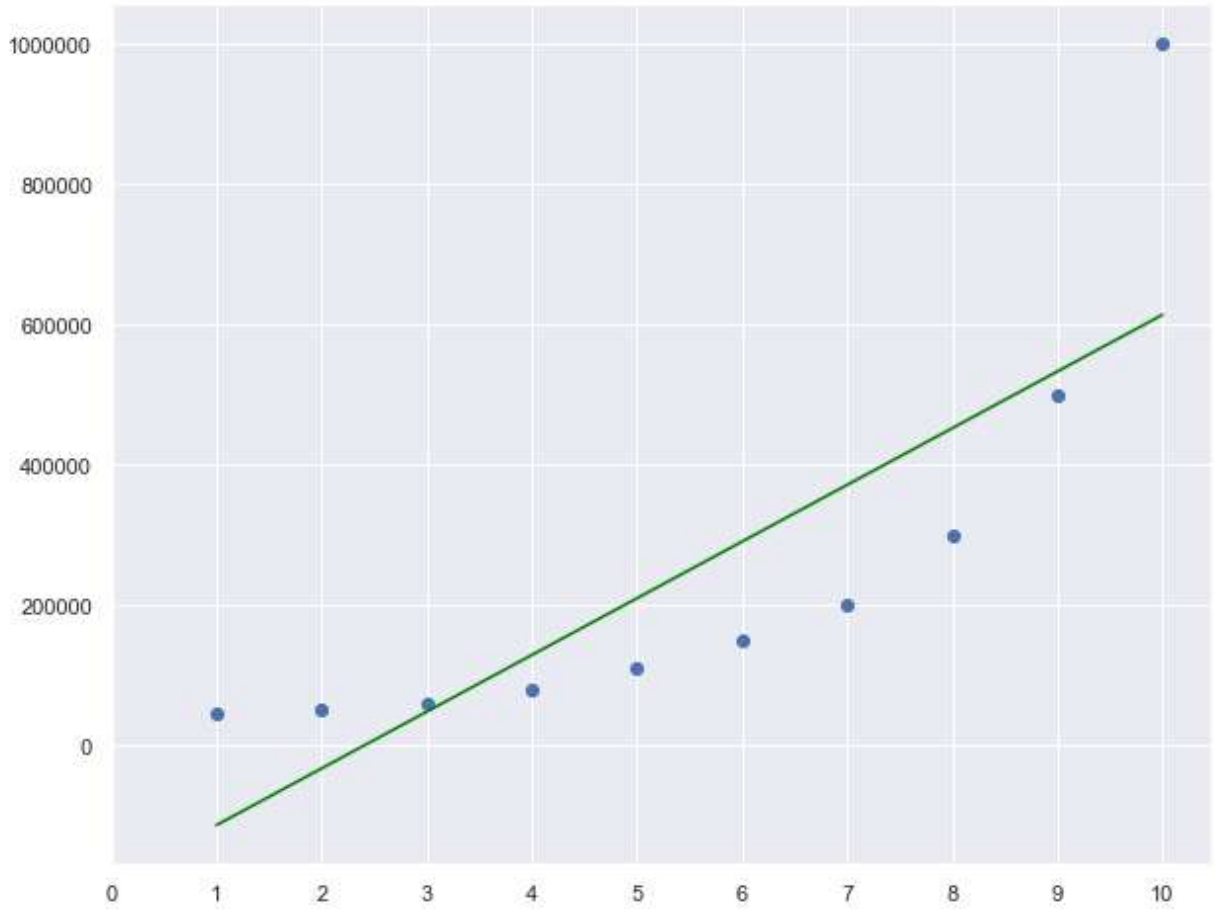
In [41]:
```
model.fit(x,y)
```

Out[41]:
```
LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None, normalize=False)
```

In [45]:
```
model.predict(x)
```

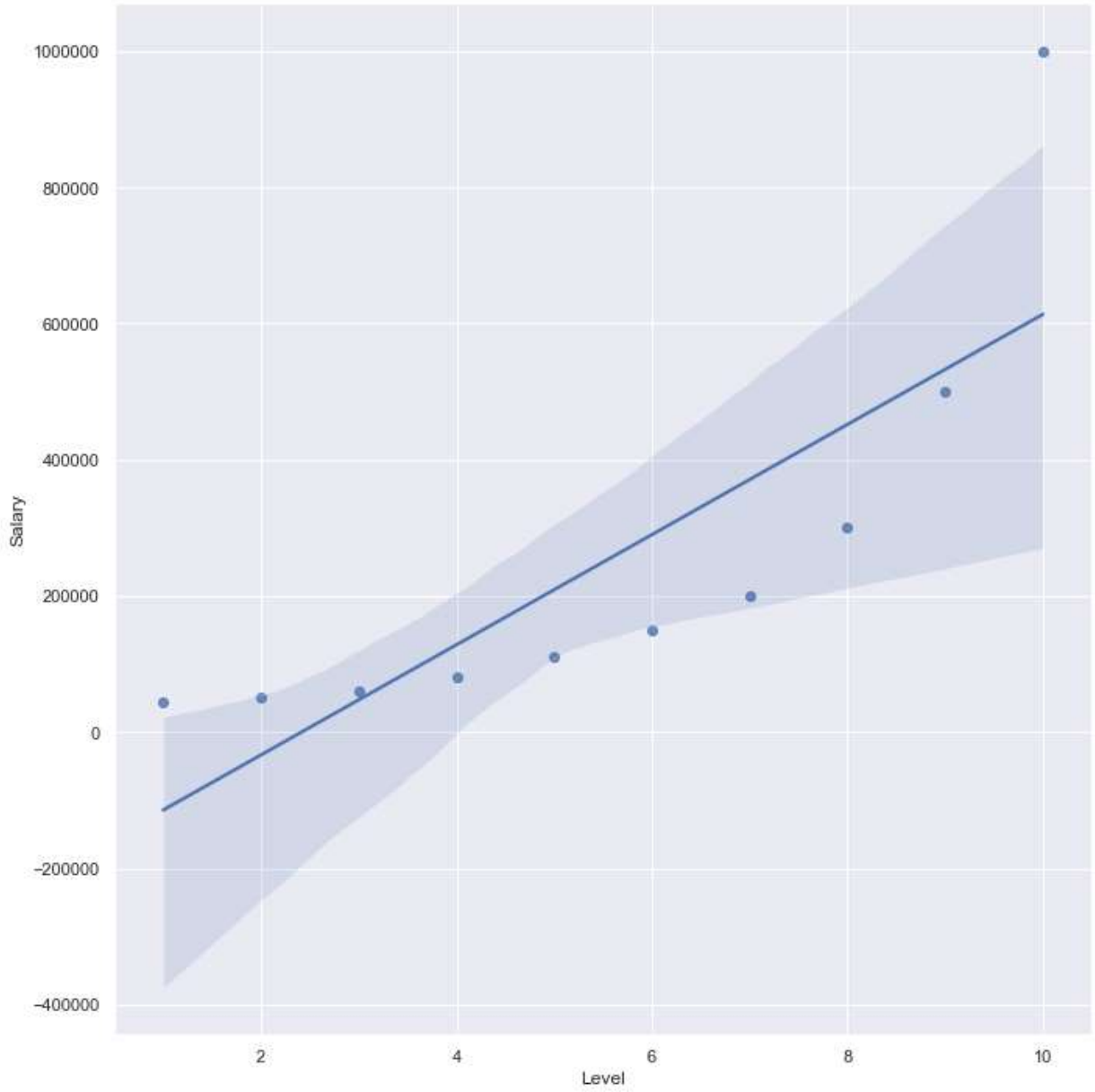
Out[45]:
```
array([[-114454.54545455],
       [ -33575.75757576],
       [  47303.03030303],
       [ 128181.81818182],
       [ 209060.60606061],
       [ 289939.39393939],
       [ 370818.18181818],
       [ 451696.96969697],
       [ 532575.75757576],
       [ 613454.54545455]])
```

In [49]:
```
plt.figure(figsize = (10,8))
plt.scatter(x, y)
plt.plot(x, model.predict(x), 'green')
plt.xticks(range(0, 11))
plt.show()
```



In [62]:
```
plt.figure(figsize = (10,8))
sns.lmplot(data = data, x = "Level", y = 'Salary', height = 10)
plt.show()
```

```
<Figure size 720x576 with 0 Axes>
```



In [63]:
```
from sklearn.preprocessing import PolynomialFeatures
```

In [80]:
```
model_poly = PolynomialFeatures(degree = 4)
```

In [81]:
```
p2 = model_poly.fit_transform(x)
```

In [82]:
```
p2
```

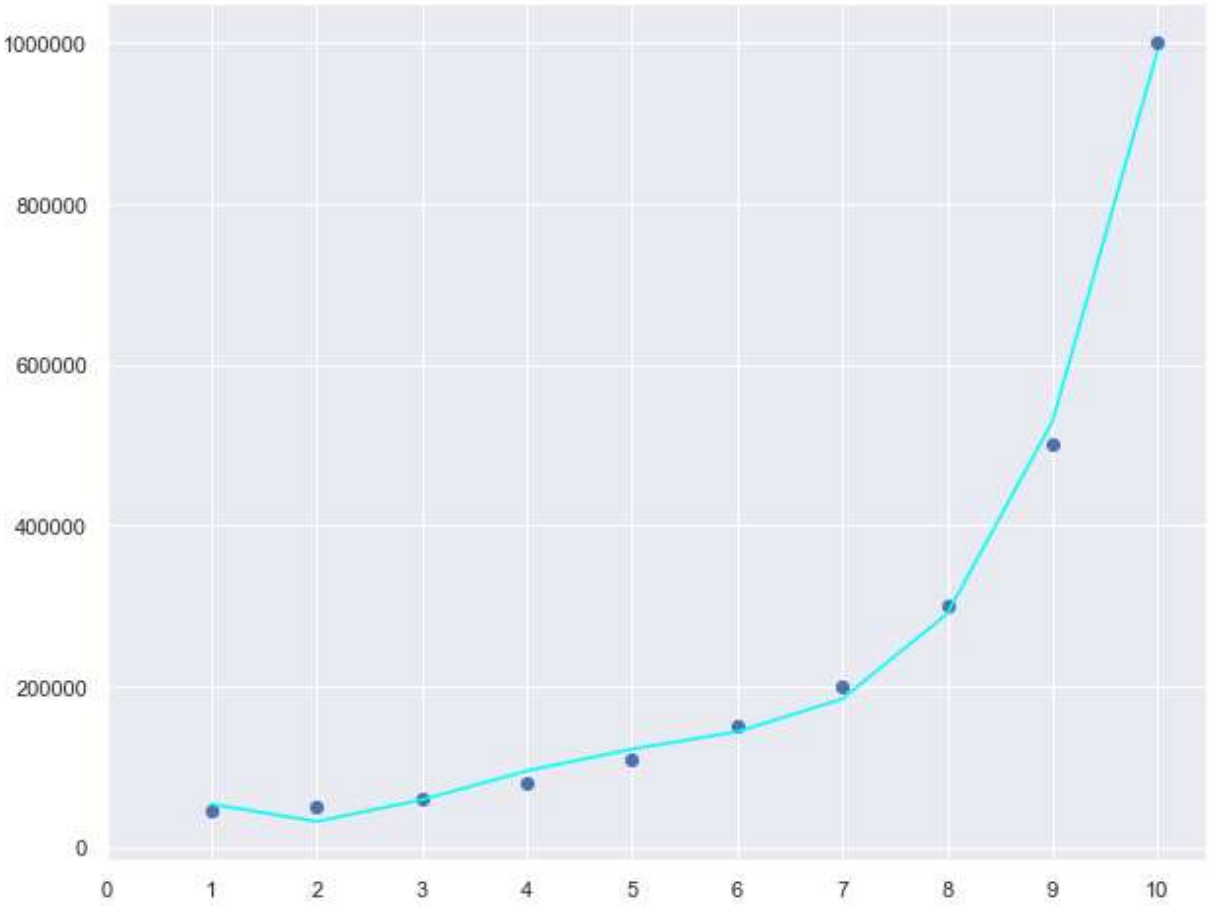
Out[82]:
```
array([[1.000e+00, 1.000e+00, 1.000e+00, 1.000e+00, 1.000e+00],
       [1.000e+00, 2.000e+00, 4.000e+00, 8.000e+00, 1.600e+01],
       [1.000e+00, 3.000e+00, 9.000e+00, 2.700e+01, 8.100e+01],
       [1.000e+00, 4.000e+00, 1.600e+01, 6.400e+01, 2.560e+02],
       [1.000e+00, 5.000e+00, 2.500e+01, 1.250e+02, 6.250e+02],
       [1.000e+00, 6.000e+00, 3.600e+01, 2.160e+02, 1.296e+03],
       [1.000e+00, 7.000e+00, 4.900e+01, 3.430e+02, 2.401e+03],
       [1.000e+00, 8.000e+00, 6.400e+01, 5.120e+02, 4.096e+03],
       [1.000e+00, 9.000e+00, 8.100e+01, 7.290e+02, 6.561e+03],
       [1.000e+00, 1.000e+01, 1.000e+02, 1.000e+03, 1.000e+04]])
```

In [83]:
```
model2 = LinearRegression()
```

In [84]:
```
model2.fit(p2, y)
```

Out[84]:
```
LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None, normalize=False)
```

In [85]:
```
plt.figure(figsize = (10,8))
plt.scatter(x,y)
plt.plot(x, model2.predict(model_poly.fit_transform(x)), 'cyan')
plt.xticks(range(0,11))
plt.show()
```



In [ ]:
```

```

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]: