Student performance

October 6, 2024

[9]:
```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

[2]:
```
df = pd.read_csv("Student_Performance.csv")
```

[5]:
```
df.isnull().mean()*100
```

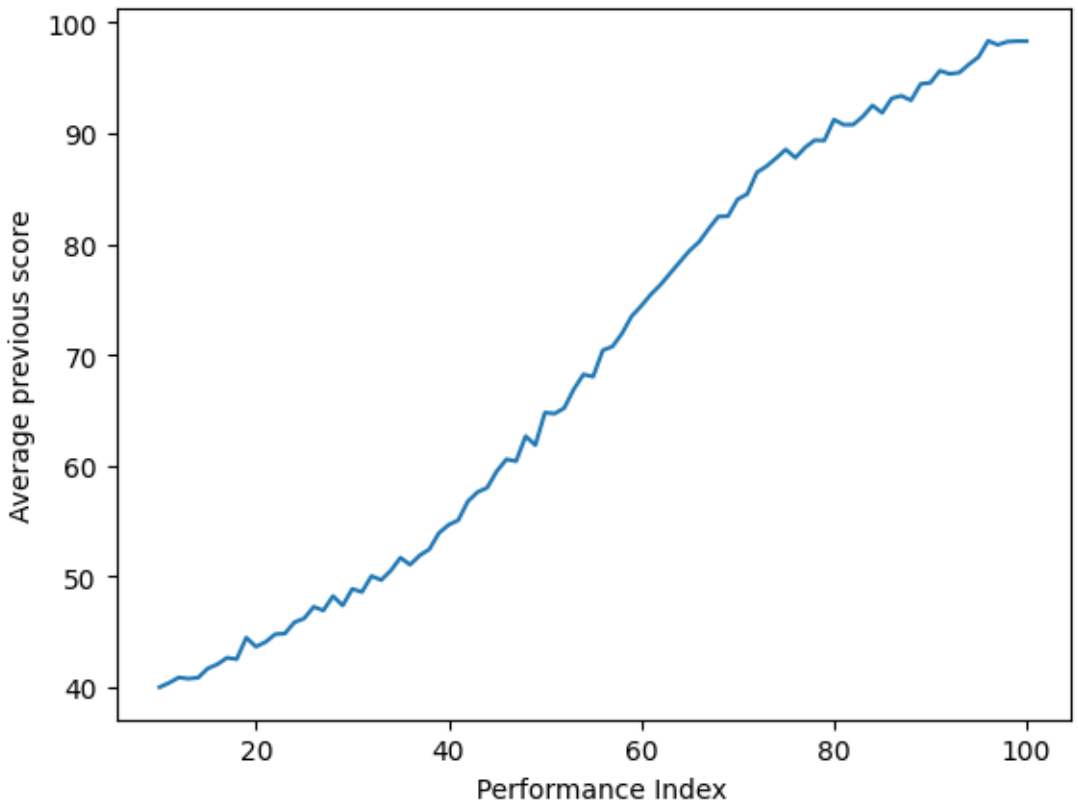
[5]:
```
Hours Studied                       0.0
Previous Scores                     0.0
Extracurricular Activities          0.0
Sleep Hours                         0.0
Sample Question Papers Practiced    0.0
Performance Index                   0.0
dtype: float64
```

[6]:
```
df.columns
```

[6]:
```
Index(['Hours Studied', 'Previous Scores', 'Extracurricular Activities',
       'Sleep Hours', 'Sample Question Papers Practiced', 'Performance Index'],
      dtype='object')
```

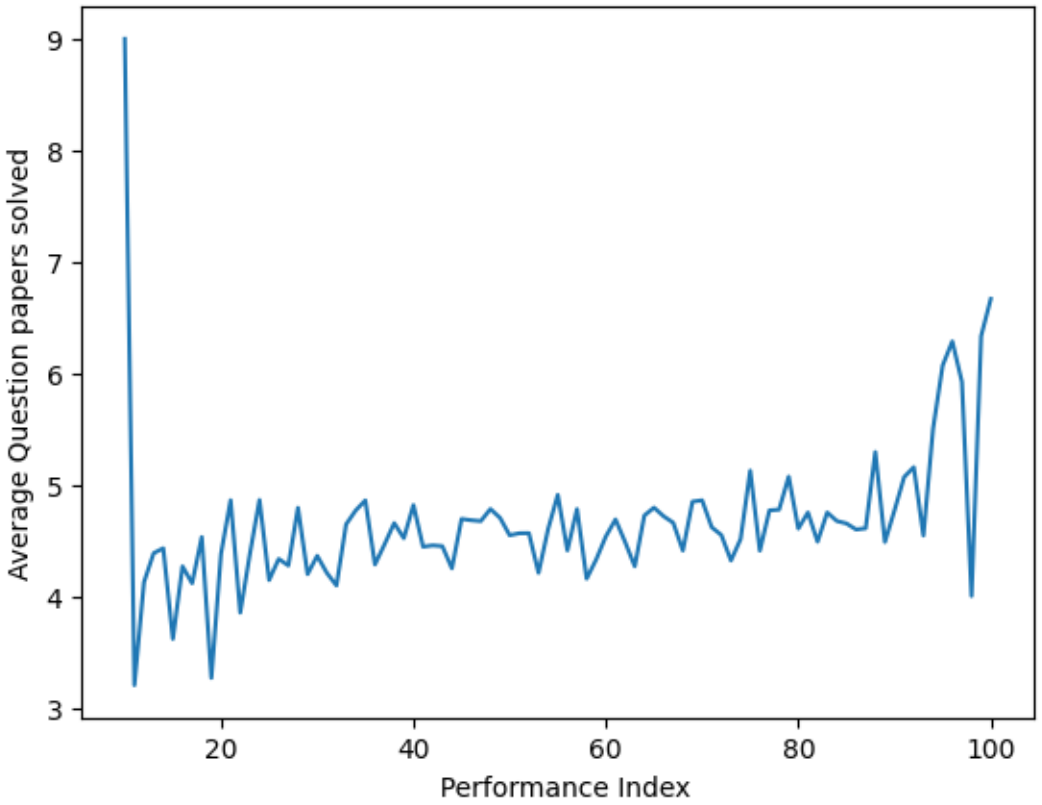
[49]:
```
previous = df.groupby("Performance Index")["Previous Scores"].mean()
plt.plot(previous.index, previous.values)
plt.ylabel("Average previous score")
plt.xlabel("  Performance Index")
#once a topper always a topper
```

[49]:
```
Text(0.5, 0, '  Performance Index')
```



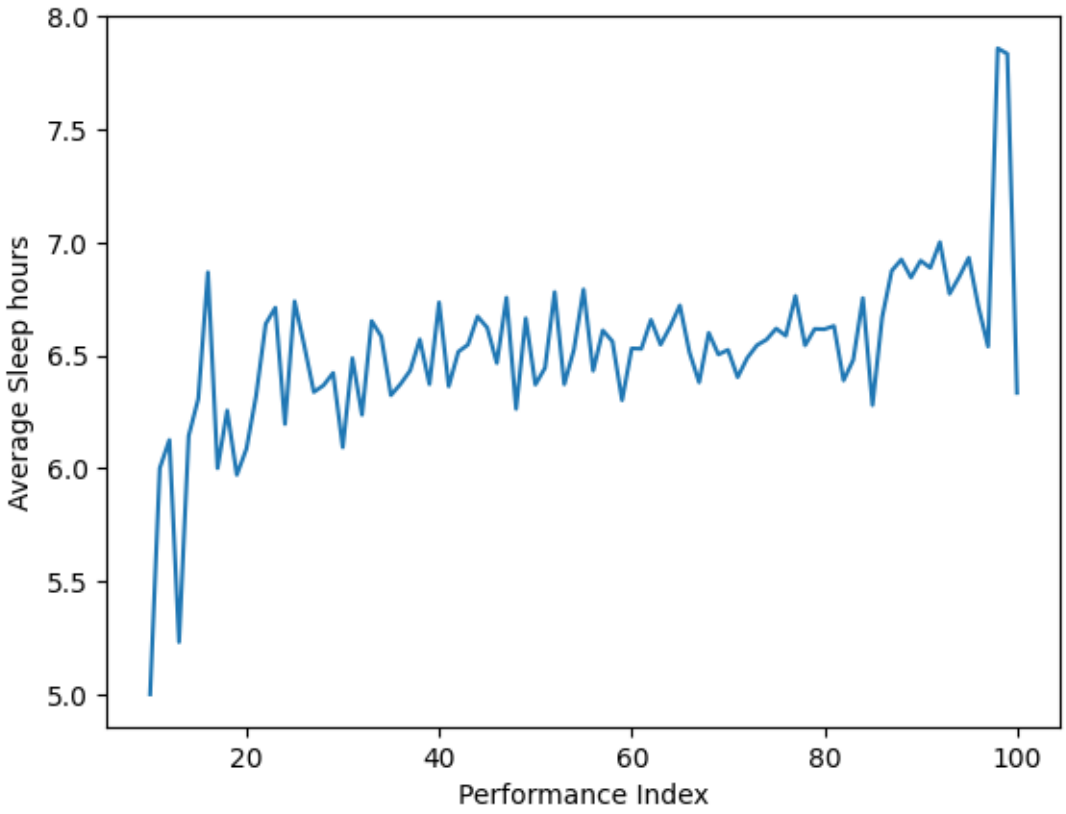
[46]:
```
questionpapersreverse = df.groupby("Performance Index")["Sample Question Papers␣
↪Practiced"].mean()
plt.plot(questionpapersreverse.index, questionpapersreverse.values)
plt.ylabel("Average Question papers solved")
plt.xlabel("  Performance Index")
#funny that topper haven't solved the question that much
```

[46]:
```
Text(0.5, 0, '  Performance Index')
```



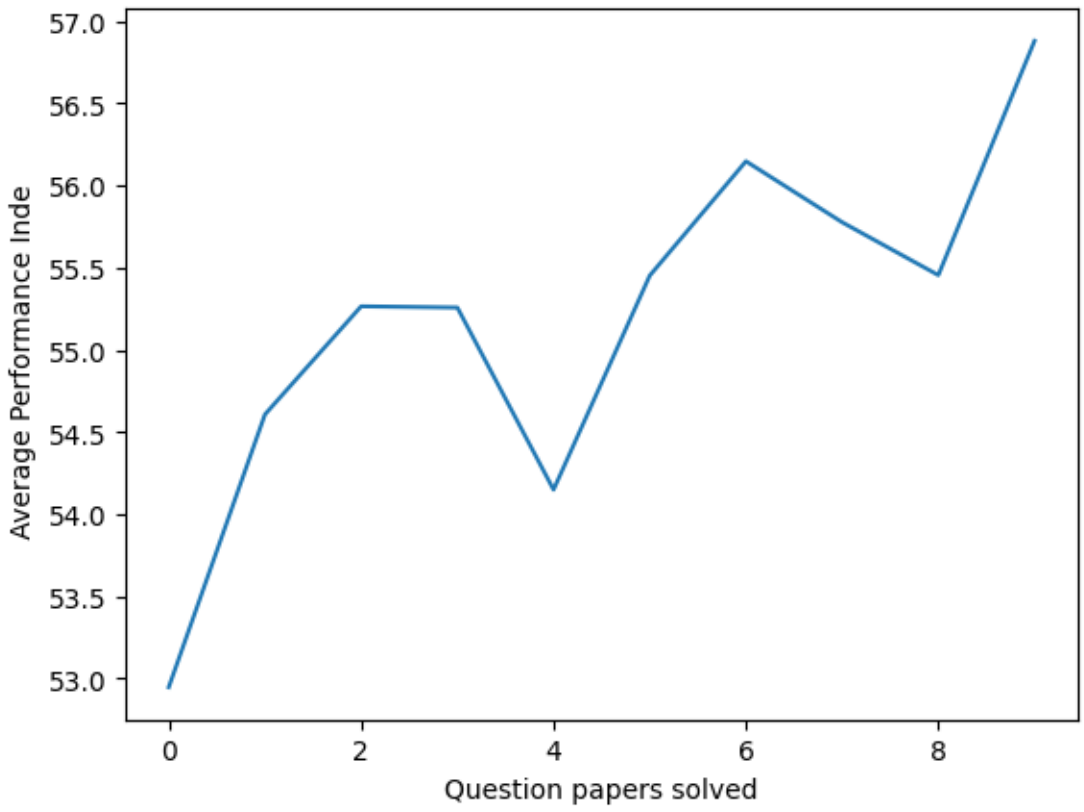
[52]:
```
reverse = df.groupby("Performance Index")["Sleep Hours"].mean()
plt.plot(reverse.index, reverse.values)
plt.ylabel("Average Sleep hours")
plt.xlabel("Performance Index")
# toppers usually sleep between 7.5 to 8 hours
```

[52]:
```
Text(0.5, 0, 'Performance Index')
```



[37]:
```
questionpapers = df.groupby("Sample Question Papers Practiced")["Performance␣
↪Index"].mean()
plt.plot(questionpapers.index, questionpapers.values)
plt.xlabel("Question papers solved")
plt.ylabel(" Average Performance Index")
```

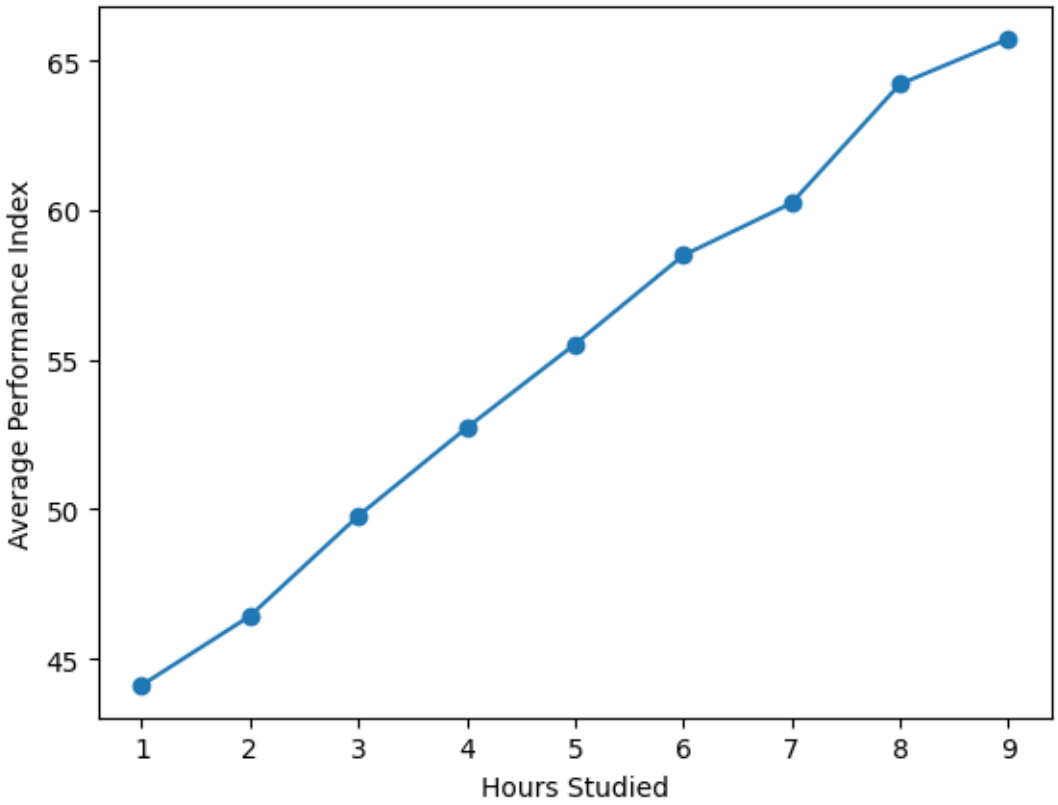
[37]:
```
Text(0, 0.5, ' Average Performance Inde')
```



[11]:
```
studyhours = df.groupby("Hours Studied")["Performance Index"].mean()
```

[29]:
```
plt.plot(studyhours.index, studyhours.values, marker='o')
plt.xlabel("Hours Studied")
plt.ylabel("Average Performance Index")
```

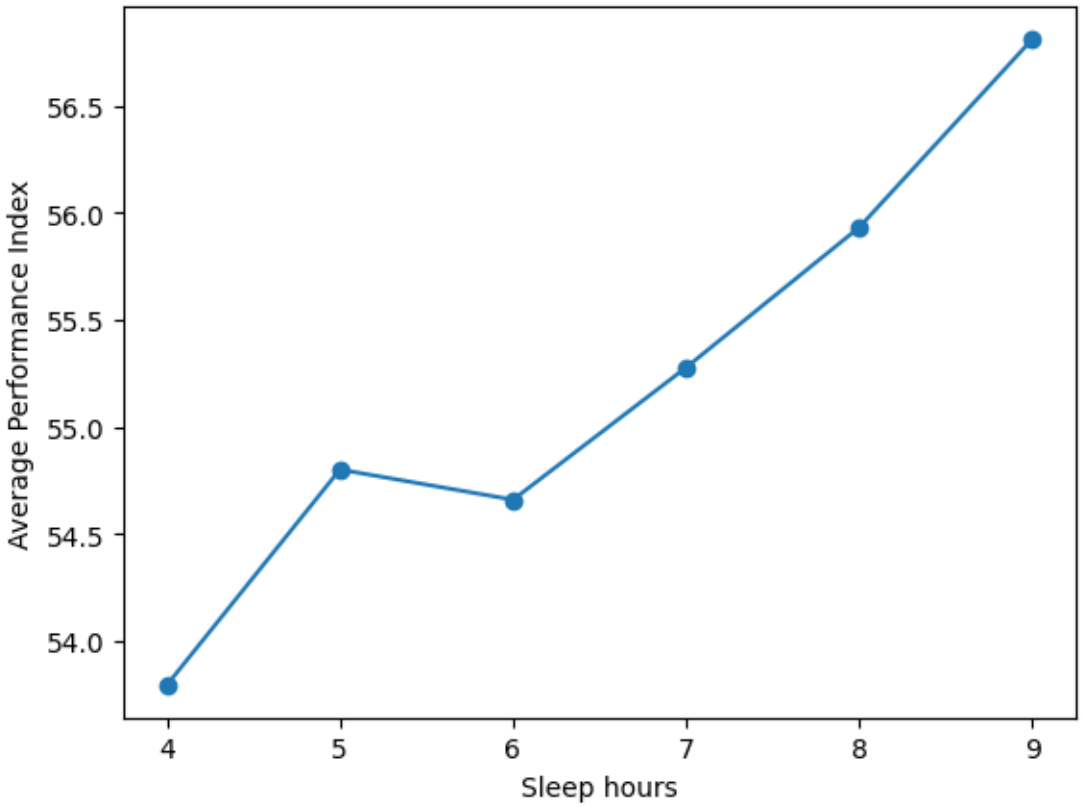
[29]:
```
Text(0, 0.5, 'Average Performance Index')
```



[19]:
```
sleephours = df.groupby("Sleep Hours")["Performance Index"].mean()
```

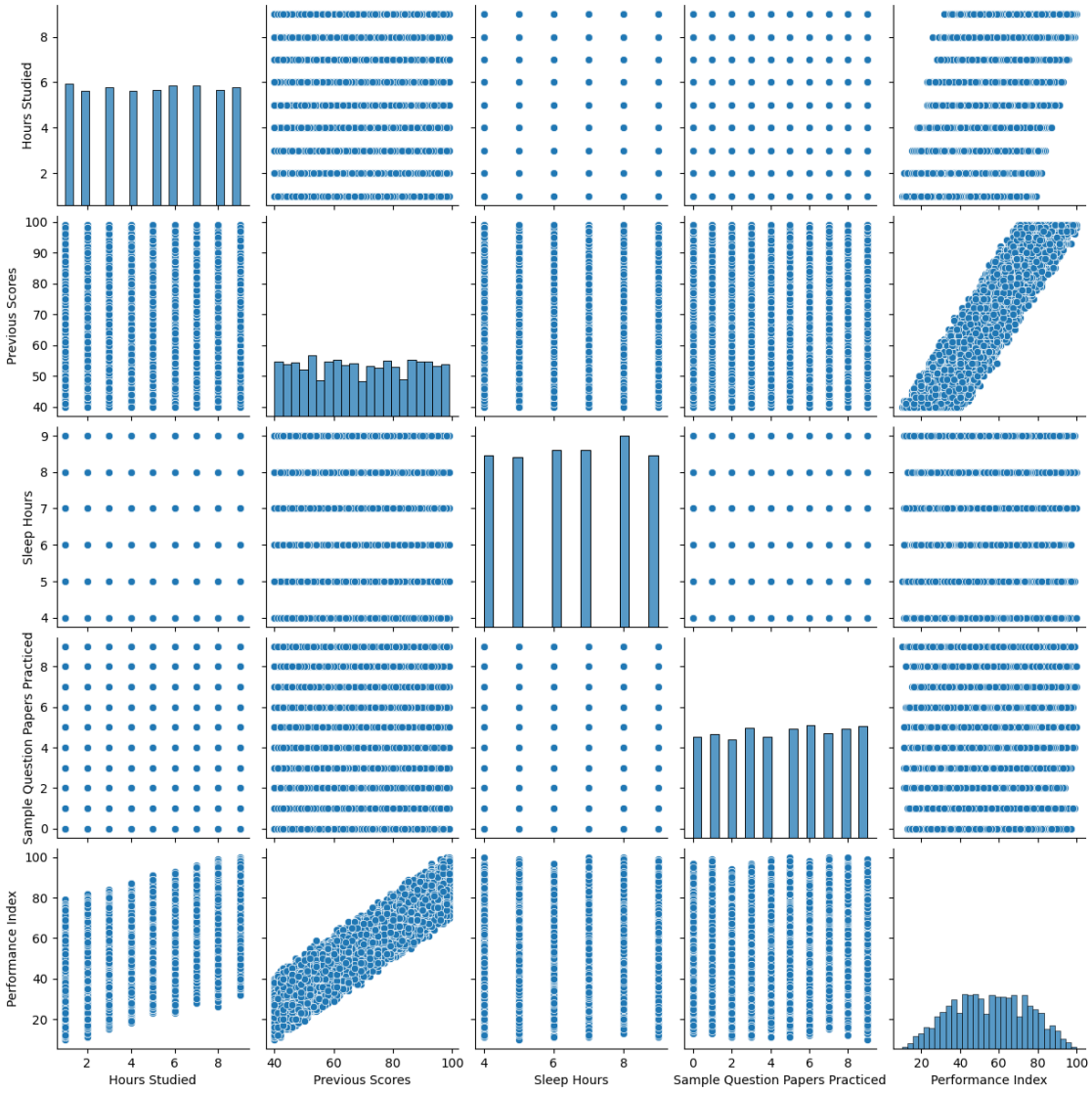
[28]:
```
plt.plot(sleephours.index, sleephours.values, marker='o')
plt.xlabel("Sleep hours")
plt.ylabel("Average Performance Index")
```

[28]:
```
Text(0, 0.5, 'Average Performance Index')
```



[8]:
```
sns.pairplot(df)
```

[8]: <seaborn.axisgrid.PairGrid at 0x251bac17440>


Hours Studied
Previous Scores
Sleep Hours
Sample Question Papers Practiced
Performance Index

[ ]: