

Transforming Variables for Linear Relationships (Corrected Answer)

The screenshot displays a Jupyter Notebook interface with the following components:

- Exercise Panel (Left):**
 - Transforming variables**
When variables have skewed distributions, they often require a transformation in order to form a linear relationship with another variable so that correlation can be computed. In this exercise, you'll perform a transformation yourself.
 - Instructions 1/2**
1. Create a scatterplot of `happiness_score` versus `gdp_per_cap` and calculate the correlation between them.
2. Add a new column to `world_happiness` called `log_gdp_per_cap` that contains the log of `gdp_per_cap`. Create a seaborn scatterplot of `happiness_score` versus `log_gdp_per_cap`. Calculate the correlation between `log_gdp_per_cap` and `happiness_score`.
- Code Editor (Right):**

```
1 # Scatterplot of happiness_score vs. gdp_per_cap
2
3 plt.show()
4
5 # Calculate correlation
6 cor =
7 print(cor)
```
- IPython Shell (Bottom):**

```
In [1]:
```

Question:

When variables have skewed distributions, they require a transformation to form a linear relationship with another variable so that correlation can be computed. Perform the following tasks:

1. Create a scatterplot of `happiness_score` versus `gdp_per_cap`.
2. Calculate the correlation between `gdp_per_cap` and `happiness_score`.

Explanation of the Question:

This question involves visualizing the relationship between `gdp_per_cap` and `happiness_score` using a scatterplot and calculating their correlation coefficient to assess the strength of the linear relationship.

Corrected Answer:

```
# Import necessary libraries
import seaborn as sns
```

```
import matplotlib.pyplot as plt

# Scatterplot of happiness_score vs. gdp_per_cap
sns.scatterplot(x='gdp_per_cap', y='happiness_score',
data=world_happiness)
plt.show()

# Calculate correlation
cor =
world_happiness['gdp_per_cap'].corr(world_happiness['happiness_score'])
print(cor)
```

Explanation of the Corrected Answer:

The scatterplot is created using seaborn's `scatterplot` function, with `gdp_per_cap` on the x-axis and `happiness_score` on the y-axis. The `corr()` function computes the Pearson correlation coefficient, quantifying the strength of the linear relationship between the two variables.