

# Visualizing Variable Relationships - Updated Scatter Plot

Learn / Courses / Exploratory Data Analysis in Python

← Course Outline →

Daily XP 265

Exercise

### Visualizing variable relationships

In the last exercise, you may have noticed that a longer `marriage_duration` is correlated with having more children, represented by the `num_kids` column. The correlation coefficient between the `marriage_duration` and `num_kids` variables is 0.45.

In this exercise, you'll create a scatter plot to visualize the relationship between these variables. `pandas` has been loaded as `pd`, `matplotlib.pyplot` has been loaded as `plt`, and Seaborn has been loaded as `sns`. The dataset name is `divorce`.

Instructions

100 XP

- Create a scatterplot showing `marriage_duration` on the x-axis and `num_kids` on the y-axis.

Take Hint (-30 XP)

script.py

Light Mode

```
1 # Create the scatterplot
2 ----
3 plt.show()
```

↺ Run Code Submit Answer

IPython Shell

Slides

In [1]:

## Question

Create a scatter plot showing `marriage_duration` on the x-axis and `num_kids` on the y-axis.

## Explanation of the Question

This task requires visualizing the relationship between the variables `marriage_duration` and `num_kids` in the `divorce` dataset using a scatter plot. This visualization helps to explore trends and correlations.

## Answer

```
# Create the scatter plot
import seaborn as sns
import matplotlib.pyplot as plt

sns.scatterplot(data=divorce, x="marriage_duration", y="num_kids")
plt.title("Marriage Duration vs Number of Kids")
plt.xlabel("Marriage Duration (years)")
plt.ylabel("Number of Kids")
plt.grid(True)
plt.show()
```

## Explanation of the Answer

The code utilizes Seaborn's scatterplot function to plot marriage\_duration on the x-axis and num\_kids on the y-axis. Additional formatting such as titles, axis labels, and gridlines enhance the clarity of the plot.