

Categorical Data in Scatter Plots

Learn / Courses / Exploratory Data Analysis in Python

← Course Outline →

Daily XP 415

Exercise

Categorical data in scatter plots

In the video, we explored how men's education and age at marriage related to other variables in our dataset, the `divorce` DataFrame. Now, you'll take a look at how *women's* education and age at marriage relate to other variables!

Your task is to create a scatter plot of each woman's age and income, layering in the categorical variable of education level for additional context.

The `divorce` DataFrame has been loaded for you, and `woman_age_marriage` has already been defined as a column representing an estimate of the woman's age at the time of marriage. `pandas` has been loaded as `pd`, `matplotlib.pyplot` has been loaded as `plt`, and Seaborn has been loaded as `sns`.

Instructions 100 XP

- Create a scatter plot that shows `woman_age_marriage` on the x-axis and `income_woman` on the y-axis; each data point should be colored based on the woman's level of education, represented by `education_woman`.

Take Hint (-30 XP)

script.py

Light Mode

```
1 # Create the scatter plot
2
3 plt.show()
```

↺ Run Code Submit Answer

IPython Shell Slides

In [1]:

Question

Create a scatter plot that shows `woman_age_marriage` on the x-axis and `income_woman` on the y-axis; each data point should be colored based on the woman's level of education, represented by `education_woman`.

Explanation of the Question

This task involves creating a scatter plot to explore the relationship between a woman's age at marriage and her income, while adding an additional layer of context by coloring points based on education level.

Answer

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

sns.scatterplot(
    data=divorce,
    x="woman_age_marriage",
    y="income_woman",
    hue="education_woman",
    palette="viridis"
)
plt.title("Age at Marriage vs Income (Colored by Education Level)")
plt.xlabel("Woman's Age at Marriage")
plt.ylabel("Income of Woman")
plt.legend(title="Education Level")
plt.show()
```

Explanation of the Answer

The code uses Seaborn's scatterplot function to plot `woman_age_marriage` on the x-axis and `income_woman` on the y-axis. The `hue` parameter adds a categorical dimension, coloring points based on `education_woman`. The `palette` ensures distinctive colors for clarity. The plot is further enhanced with a title, axis labels, and a legend.