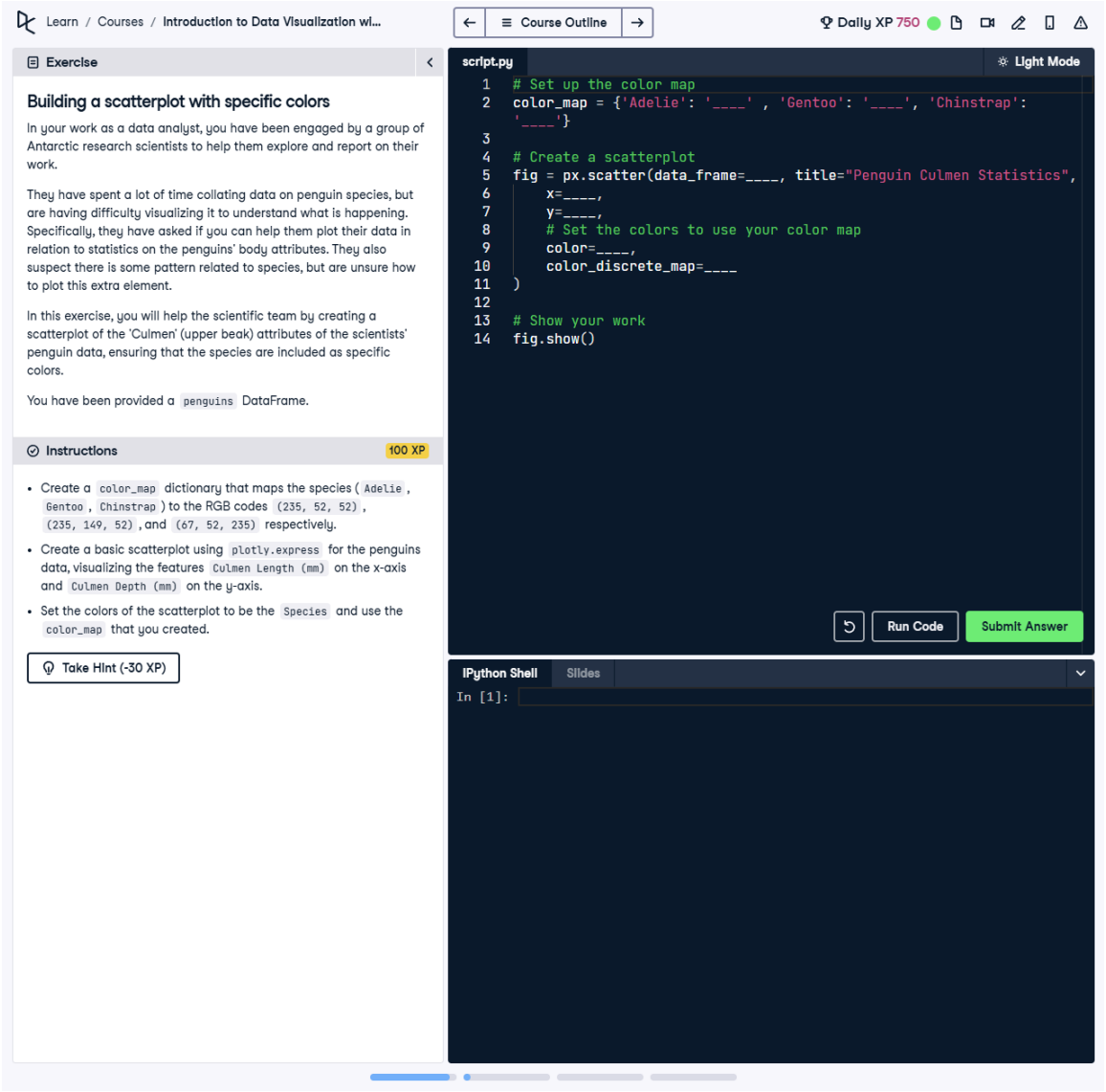
Building a Scatterplot with Specific Colors



Question:

As part of your work as a data analyst, you have been engaged by a group of Antarctic research scientists to help them explore and report on their work.

They have spent a lot of time collating data on penguin species, but are having difficulty visualizing it to understand what is happening. Specifically, they have asked if you can help them plot their data in relation to statistics on the penguins' body attributes. They also asked that the colors are set per species, as they are unsure how to plot this extra element.

In this exercise, you will help the scientist team by creating a scatterplot of

the 'Culmen' (upper beak) attributes of the statistics' penguin data, ensuring that the species are included as specific colors.

Instructions:
1. Create a `color_map` dictionary that maps the species (`Adelie`, `Gentoo`, `Chinstrap`) to the RGB codes:
   - Adelie: (235, 52, 52)
   - Gentoo: (235, 149, 52)
   - Chinstrap: (67, 52, 235)
2. Create a scatterplot using `plotly.express` for the penguins data, visualizing the features `Culmen Length (mm)` on the x-axis and `Culmen Depth (mm)` on the y-axis.
3. Set the colors of the scatterplot to be the `Species` and use the `color_map` that you created.
4. Display the plot using `fig.show()`.

Explanation of the Question:

This task involves creating a scatterplot of penguin data with the culmen length on the x-axis and the culmen depth on the y-axis. Each species is assigned a specific color using a color map dictionary. The `color` parameter is set to `Species` to distinguish the species in the plot visually.

Answer:

```
# Set up the color map
color_map = {
    'Adelie': 'rgb(235, 52, 52)',
    'Gentoo': 'rgb(235, 149, 52)',
    'Chinstrap': 'rgb(67, 52, 235)'
}

# Create a scatterplot
fig = px.scatter(
    data_frame=penguins,
    title='Penguin Culmen Statistics',
    x='Culmen Length (mm)',
    y='Culmen Depth (mm)',
    # Set the colors to use your color map
    color='Species',
    color_discrete_map=color_map
)

# Show your work
fig.show()
```

Explanation of the Answer:

The `color_map` dictionary defines RGB colors for each penguin species. The `color_discrete_map` parameter in Plotly Express is used to apply these colors to the `Species` column. The x-axis represents the culmen length, and the y-axis represents the culmen depth. The `fig.show()` function displays the scatterplot. This plot visually distinguishes the species based on their respective colors and attributes.