

6.4.4 Find the Correlation Between Latitude and Percent Cloudiness

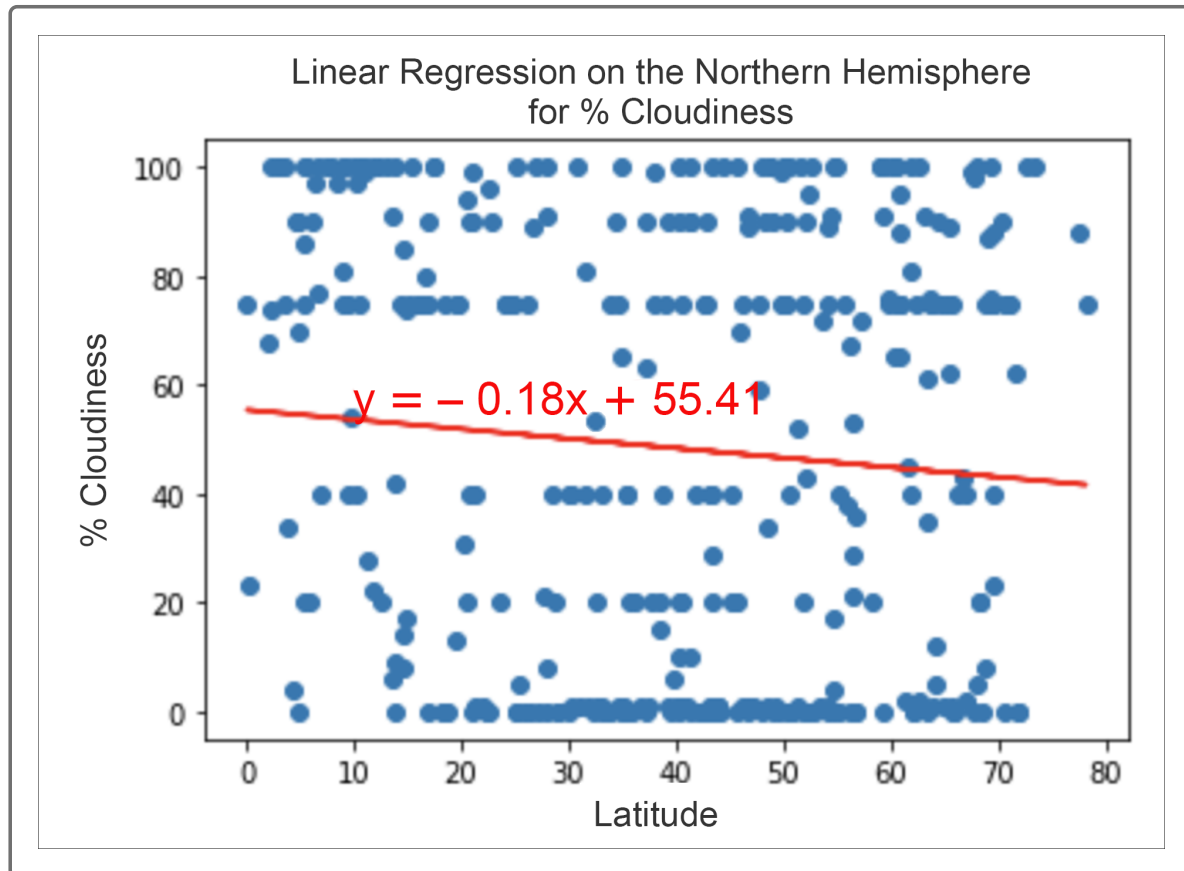
You're almost done creating the linear regression line and plots. This time, create the linear equation plot on the latitude and percent humidity for the Northern and Southern Hemispheres. Stay focused so you don't make any mistakes when you refactor the code.

Being able to write a function and call it with the appropriate data is a valuable skill. Let's knock out another regression line and plot. This time we'll get the data for the percent cloudiness for the Northern and Southern Hemispheres.

Perform Linear Regression on the Percent Cloudiness for the Northern Hemisphere

SKILL DRILL

1. Refactor the code we have been using for linear regression lines and plots to create the x- and y-values for the percent cloudiness and latitudes on the Northern Hemisphere DataFrame.
2. Call the `plot_linear_regression` function with the correct arguments to create the linear regression line and plot for percent cloudiness in the Northern Hemisphere. It should look like the following plot.



[Retake](#)

Perform Linear Regression on the Percent Cloudiness for the Southern Hemisphere

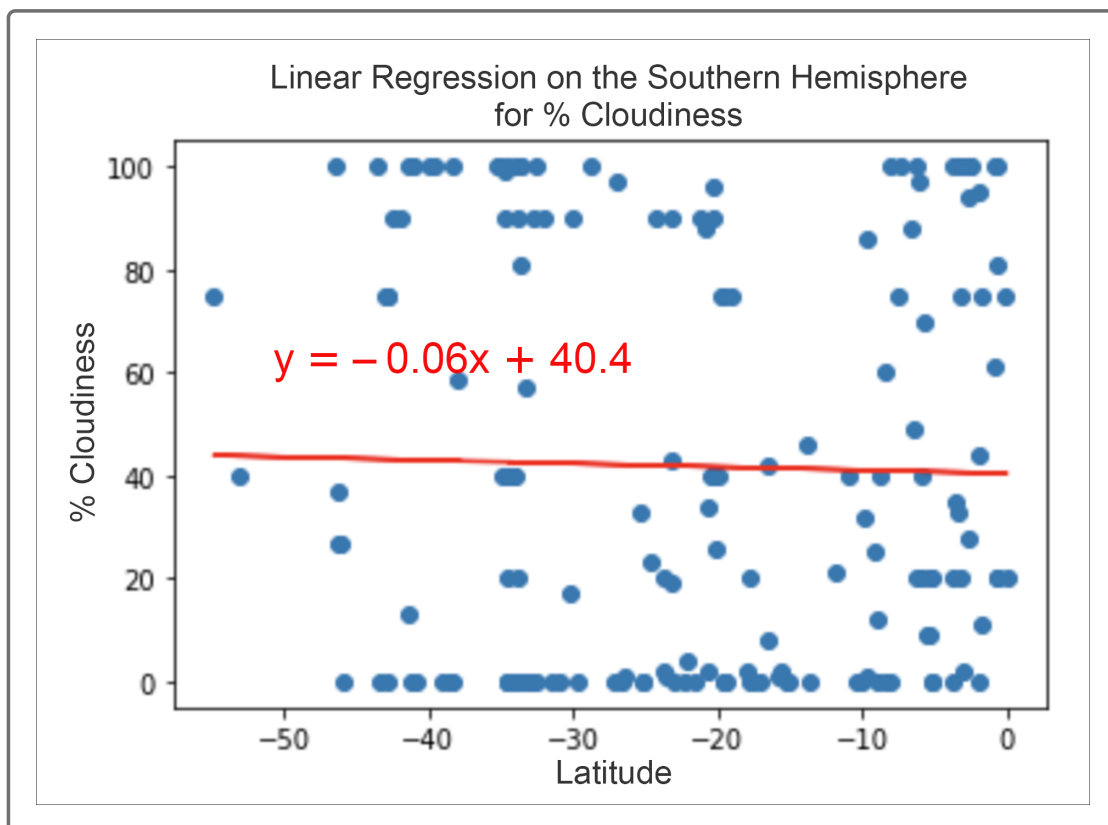
Now we'll create the linear regression line and plot for percent cloudiness in the Southern Hemisphere.

Generate the x-value equal to the latitude column and the y-value equal to the cloudiness column from the `southern_hemi_df` DataFrame. Call the `plot_linear_regression` function. Be sure to edit the `title`, `y_label`, and `text_coordinates` for the percent cloudiness scatter plot.

Add the code to a new cell and run it to generate the linear regression and plot the data.

```
# Linear regression on the Southern Hemisphere
x_values = southern_hemi_df["Lat"]
y_values = southern_hemi_df["Cloudiness"]
# Call the function.
plot_linear_regression(x_values, y_values,
                      'Linear Regression on the Southern Hemisphere \n
                      for % Cloudiness', '% Cloudiness', (-50, 60))
```

The scatter plot with the regression line and equation should look like the following.



FINDING

The correlation between the latitude and percent cloudiness is very low because the r-value is less than -0.09 for the Northern Hemisphere and less than -0.02 for the Southern Hemisphere for the plots shown here. This means that cloudiness is unpredictable due to changing weather patterns that can increase or decrease percent cloudiness. Check the r-values for your plots.

© 2020 - 2022 Trilogy Education Services, a 2U, Inc. brand. All Rights Reserved.