



Typically, the wet season in Peru is from **November to March** with warmer temperatures and intermittent showers. There is little to almost no rain in the Andes region during the months of May to October. This is probably the best time for long, extensive hikes.

Puerto Rico's rainy season lasts from **April to November** and the driest season is December to March.

Important parameters:

1. Temperature
2. Humidity
3. Precipitation/Rainfall
4. Vegetation cover
5. Wind Speed
6. Dew point and Specific Humidity are highly correlated. We choose only Specific Humidity.
7. Precipitation amount and Reanalysis\_ppt\_amt.
8. (Max\_air\_temperature, Min\_air\_temperature) and reanalysis\_tdtr\_k. We can use just tdtr\_k.
9. For NDVI – we take average of the four directions.

|  |  |  |
| --- | --- | --- |
| Feature 1 | Feature 2 | Corr |
| reanalysis\_specific\_humidity\_g\_per\_kg | reanalysis\_dew\_point\_temp\_k | 0.997051 |
| reanalysis\_dew\_point\_temp\_k | reanalysis\_specific\_humidity\_g\_per\_kg | 0.997051 |
| reanalysis\_tdtr\_k | reanalysis\_max\_air\_temp\_k | 0.918578 |
| reanalysis\_max\_air\_temp\_k | reanalysis\_tdtr\_k | 0.918578 |
| reanalysis\_air\_temp\_k | reanalysis\_avg\_temp\_k | 0.901777 |
| reanalysis\_avg\_temp\_k | reanalysis\_air\_temp\_k | 0.901777 |
| reanalysis\_tdtr\_k | station\_diur\_temp\_rng\_c | 0.881176 |
| station\_diur\_temp\_rng\_c | reanalysis\_tdtr\_k | 0.881176 |
| ndvi\_ne | ndvi\_nw | 0.850902 |
| ndvi\_nw | ndvi\_ne | 0.850902 |
| reanalysis\_max\_air\_temp\_k | station\_diur\_temp\_rng\_c | 0.834263 |
| station\_diur\_temp\_rng\_c | reanalysis\_max\_air\_temp\_k | 0.834263 |
| ndvi\_se | ndvi\_sw | 0.820924 |
| ndvi\_sw | ndvi\_se | 0.820924 |
| reanalysis\_tdtr\_k | reanalysis\_min\_air\_temp\_k | 0.815511 |
| reanalysis\_min\_air\_temp\_k | reanalysis\_tdtr\_k | 0.815511 |

Steps in Analysis:

1. Data cleaning
2. EDA
3. Basic model with Linear Regression.

EDA analysis inferences:

1. A plot of NDVI vs Total cases for both cities shows that Peru has less number of cases while having higher NDVI than Puerto Rico which has higher no. of cases for lower NDVI.
2. Humidity does have an impact on the number of total cases (although weak). When compared over the years we could see that the number of cases have decreased with time and we suspect that it could be because of general improvement in healthcare. It has an effect within a certain range of humidity.
3. We are trying to analyze based on the months of the year instead of weeks.