**EDAV 5702 Community Contribution**

**Luwei Zhang**

**lz2815**

**Motivation:**

**Python** and **R** have been two popular programming languages for data scientists. However, data visualization in **Python** has long been referred to as painful and ugly compared to the handy tools available in **R**, specifically to **ggplot2**. However, as **Python** is more commonly used in machine learning modeling, it would be more efficient to conduct the exploratory analysis and visualization in the same environment as where the model is built. In other words, modeling in **Python** and visualization in **R** makes no sense.

To address this issue, this tutorial talks about how to use **ggplot** for visualization in **Python**. It introduces common functions with examples that have been covered in previous lectures of 5702 but in **Python** this time.

**Evaluation:**

Although the APIs are basically the same as in **R**, plotting in **Python** requires quite different syntaxes because it incorporates the **matplotlib.plotly** package. This tutorial gives some examples of such difference, and I learned more about the language structures for plotting of both **Python** and **R**. If time permitting, it would be great to include **Python** plotting with other R packages beyond **ggplot2** that have been introduced in the lecture, such as **vcd**, **ggmosaic**, etc.