|  |  |
| --- | --- |
| A picture of a winding road and trees  Python for Data science | Joseph Mulingwa Kithome  Python for data science |

**Python for Data Science**

* Excel
* python
* R
* Tableau

**Python libraries**

1. Pandas-provide a fast, flexible and expressive way of working with structure.s
2. Numpy-used to work with arrays.
3. Matplotlib- is a plotting library
4. Seaborn-plotting library based on matplotlib

used for visualization

**Data Analysis**

**Steps involved are;**

* Data wrangling-gather all the data you need for the analysis. Gather from spreadsheets, databases, web.
* Assess to identify problems (Quality and structures).
* Explore (Exploratory Data Analysis)
* understand the structure of your data faster than statistics.

**Analysis**

* Analyze the interesting questions.
* Draw conclusions
* Drawing conclusion from summarizing data with descriptive statistics.
* Visualization

Communicate and Justify conclusions. This can be done by exploratory analysis.

->explanatory

->exploratory

**Things to keep in mind**

* know techniques for extracting, cleaning, exploring, analysis, and visualizing
* learn inferential and descriptive statistics
* learn how to look for outliers
* get skills in A/B testing’s, hypothesis.