Strathmore University
@iLabAfrica Centre

DSA 8102 - DATA MINING, STORAGE AND RETRIEVAL

(Guideline and Evaluation – Term Project)

Facilitator: Dr. Kennedy Senagi

Contacts: ksenagi@strathmore.edu

Evaluation:

CONTENT	SCORE
CAT 1 • Exploratory data analytics exercise [already submitted]	10%

CAT2 – Deadline: 25th July 2024
• Introduction [10 marks]
  ○ Background [1 paragraph of at least 6 lines]
  ○ Research Problem [1 paragraph of at least 6 lines]
  ○ Objectives [At least 3 SMART points]
  ○ Hypothesis [At least 2 points]
• Methodology, Results and Discussion
  ○ Data description [1 paragraph of at least 10 lines with variables described in a table] [10 marks]
  ✓ Source of data
  ✓ Period collected (year and month/day)
  ✓ How it was collected
  ✓ Under what conditions was it collected
  ○ Exploratory data analytics [with visualizations and interpretations; **new hind insights**] [10 marks]
  ✓ Descriptive analytics
  ✓ Diagnostics analytics
  ○ Data cleaning/pre-treatment for machine learning purpose. [10 marks]
  ○ Predictive data analytics [with data science and machine learning models; **new foresights**] [15 marks]
• Conclusion [5 marks]
	60%

EXAM: Deadline: 1 day before exam presentations • A detailed and improved version of all above	30%

Submission requirements:

- A unique dataset should be identified and used for this work.
- Work should be documented on a Jupyter Notebook - https://jupyter.org/
- Programming language: Python >3.0
- Presentations: Jupyter Notebook with project/scripts demo

Disclaimer:

- Late submissions will not be accepted; will be penalized.
- Plagiarism will not be accepted.
- Think critically, far, and wide.