**CTEC 298-101 Fall 2024: Data Visualization and Python Deliverables**

**Title:**

**Implementation and Analysis of Data Visualizations using Python**

**Authors:**

**Marquis Jones (Group Leader)**

**Class:**

**CTEC 298-101**

**Professor Bemley**

**Session:**

**Fall 2024**

## Introduction

In this paper, I will provide a detailed description of the various deliverables and Python code used in CTEC 298-101 Fall 2024. The primary focus of the project is to apply Python for data analysis and visualization, specifically through Jupyter Notebooks. The assignment includes working with datasets, performing data cleaning and wrangling, and creating several plots including bar plots, pie charts, histograms, scatter plots, stack plots, and multiplots. These plots serve as both the deliverables for the class and as visual representations for the data presented in the paper. This paper also covers the process and tools involved in completing the tasks, providing evidence for all deliverables.

In the CTEC 128 course, I worked on various assignments that focused on the foundational concepts of data analysis and visualization using tools like Python, Jupyter Notebooks, and Pandas. The class was structured to be collaborative, with group work being a significant component. We were divided into teams, and each member was assigned specific tasks that contributed to the final deliverables, such as research papers, presentations, and reports. Unfortunately, due to limited access to past assignments, I am unable to provide specific examples or assignments from this course. However, I can describe how the course operated in a group format, where the work was divided amongst the members, and each individual contributed their part to the collective goals of the team. The focus of the assignments included cleaning and transforming data, applying Python libraries for plotting, and ultimately presenting our findings in a clear and structured manner.

## Summaries of CTEC 128 Papers

**Summary 1 - Group Leader (Marquis Jones):**
The first section of CTEC 128 involved learning the basics of Python programming, including syntax, variables, and data structures. The assignments focused on understanding how to write Python code and organize it effectively for data analysis tasks. Emphasis was placed on working with libraries such as pandas, matplotlib, and numpy, which were essential for the data wrangling and visualization tasks required in CTEC 298. The paper submissions also involved using Jupyter Notebooks for interactive coding, ensuring that all steps were documented and executed efficiently.

**Summary 2 - Member 1:**
In CTEC 128, the course materials were focused on learning the fundamentals of data science and Python programming. One key takeaway was the integration of pandas for data manipulation and matplotlib for data visualization. The first group assignment revolved around using Python to clean and prepare datasets, followed by generating basic plots such as histograms and pie charts. While I did not participate in the team for CTEC 128, the skills acquired during that course have directly contributed to the work presented in CTEC 298.

**Summary 3 - Member 2:**
CTEC 128 allowed for the exploration of Python libraries and introduced fundamental data science concepts. One of the most impactful lessons was understanding how to visualize data effectively using plots and charts. By utilizing pandas for data wrangling and matplotlib for plotting, we learned how to manipulate and analyze datasets to extract meaningful insights. The group assignments were collaborative and helped reinforce the coding skills needed for more complex analysis in CTEC 298.

**Description of CTEC 298 Material Submitted**

In CTEC 298, the focus was on applying the skills from CTEC 128 to more complex datasets and visualizations. As the sole member of the group, I was responsible for completing and submitting all materials. This included the creation of multiple data visualizations, data wrangling, and the documentation of each step taken during the process. The goal was to create meaningful insights from the dataset while ensuring that the visualizations were clear, accurate, and informative.

**Description of the CTEC 298 Plot Deliverables**

**Description 1 - Group Leader (Marquis Jones):**
In this assignment, I created six distinct plots using Python. These plots include bar plots, pie charts, histograms, scatter plots, stack plots, and multiplots. For each plot, I provided a detailed explanation of the data cleaning process and the final dataset used. The Python code used to generate these plots was thoroughly documented in Jupyter Notebooks, which were organized in

the GitHub and Google Drive repositories. Additionally, I included the original dataset and final dataset, ensuring that every step of the data wrangling process was visible and reproducible.

**Description 2 - Member 1:**
For the plot deliverables, I learned the importance of data wrangling, as it directly impacts the quality of the visualizations. By using pandas and numpy for cleaning and transforming the datasets, I was able to generate meaningful plots. Each plot was accompanied by a detailed explanation of the steps taken to clean the data, followed by the corresponding code used to produce the visualizations.

**Description 3 - Member 2:**
The deliverables for CTEC 298 included several Python-generated plots. I was responsible for understanding the various plot types and ensuring that the data was processed and formatted properly. The use of Jupyter Notebooks allowed for clear documentation and easy reproduction of the steps involved in each plot creation. Python code and final plots were uploaded to GitHub and Google Drive for easy access and reference.

**Linked Files**
*** file:///Users/bigquis/Downloads/Multiplot-Executable.html ***
This link gives you the provided code examples focus on creating various plots using Python, including bar plots, pie charts, histograms, scatter plots, stack plots, and multiplots. Each plot type is constructed by first preparing the data, which involves cleaning and transforming it using Python libraries like Pandas. The code demonstrates how to visualize the data using Matplotlib, a powerful plotting library. The steps typically involve importing necessary libraries, preparing datasets, creating specific plots, and displaying them in a clear, informative manner. In the case of the multiplot, the code combines multiple plots in a single figure, allowing for a comparison of different visualizations side by side. These steps are integral to transforming raw data into visual insights that help inform decision-making, all while providing a clear, step-by-step explanation of the Python commands used in the process.

**Summary/Conclusion**

In conclusion, the work presented in this paper showcases the progression of skills learned from CTEC 128 to CTEC 298. By working with datasets, cleaning the data, and generating various types of plots, I was able to gain valuable insights into both data wrangling and visualization techniques. The visualizations created in Python were integral to the assignment and provided a means of communicating complex data in an accessible and clear manner. The experience gained from CTEC 298 will prove invaluable as I continue to work with data analysis and visualization tools in future projects.

**References**

Python Software Foundation. (2024). *Python documentation*. Retrieved from https://www.python.org/doc/

Matplotlib Team. (2024). *Matplotlib documentation*. Retrieved from https://matplotlib.org/stable/contents.html

Pandas Development Team. (2024). *Pandas documentation*. Retrieved from https://pandas.pydata.org/pandas-docs/stable/

Waskom, M., Botvinnik, O., O'Kane, D., Hobson, P., Lueck, L., & Sea, K. (2024). *Seaborn: Statistical data visualization*. Retrieved from https://seaborn.pydata.org/