**User Manual for C964 Capstone - Data Analysis Application**

Welcome to the user manual for the Data Analysis Application. This application helps you analyze data stored in CSV files, providing various charts and metrics for better data understanding.

* System Requirements:
* Windows 10 machine
* Python 3.6 or later installed
* Required libraries: tkinter, pandas, plotly, scikit-learn, and webbrowser

**Environment Installation**:

1. (Preferred) Download PyCharm Community Edition for a consistent development environment (<https://www.jetbrains.com/edu-products/download/other-PCE.html>)
   1. A step-by-step guide to installing PyCharm Community Edition, along with the additional required packages:
      1. First, go to the PyCharm Community Edition download page at <https://www.jetbrains.com/edu-products/download/other-PCE.html>.
      2. Click the "**Download**" button for the **PyCharm Community Edition** for **Windows**

Graphical user interface, text, application, chat or text message

Description automatically generated

* + 1. Once the download is complete, double-click the downloaded file to begin the installation process.
    2. Follow the prompts in the installation wizard to complete the installation.
       1. You can choose to install **PyCharm** for all users or just for yourself,
       2. And also choose the installation location.
    3. Once PyCharm installation is complete, open **PyCharm Community Edition**.
       1. Next, click on Windows Start button
       2. If you installed PyCharm Community Edition Locate the Jet
    4. To install the tkinter package, go to the "File" menu and select "Settings" (or "Preferences" on a Mac).
    5. In the left-hand pane of the Settings/Preferences window, click "Project: <your project name>" (or "Project Interpreter" if you don't have a project open yet).
    6. In the right-hand pane, click the "+" button to add a new package.
    7. In the "Available Packages" window, type "tkinter" in the search bar and click the checkbox next to "tkinter" in the search results.
    8. Click the "Install Package" button at the bottom right of the window.
    9. Wait for the package to install (this may take a few minutes).
    10. To install the pandas package, repeat steps 6-10, but this time search for "pandas" in the "Available Packages" window and install that package instead.
    11. To install the plotly package, repeat steps 6-10, but this time search for "plotly" in the "Available Packages" window and install that package instead.
    12. To install the scikit-learn package, repeat steps 6-10, but this time search for "scikit-learn" in the "Available Packages" window and install that package instead.
    13. Once all of the packages are installed, you should be able to use them in your PyCharm projects by importing them into your code.

1. (Optional) Download and install Python 3.6 or later from the official Python website (<https://www.python.org/downloads/>).
   1. If you intend to uses Python and IDLE go to step 3
2. If you should opt in to use Python and Pyhton’s onwn IDLE the you will need to do the following:
   1. Open the command prompt and type the following command to install the required libraries:
      1. pip install tkinter
      2. pip install pandas
      3. pip install plotly
      4. pip install scikit-learn

Using the Data Analysis Application:

* Open the application using Python or PyCharm.
* The Data Analysis window will appear. Click the "Browse" button and select a CSV file (e.g., "diabetes\_data.csv") for analysis.
* The selected file's path will be displayed in the text field next to the "Browse" button.
* Click the "Analyze" button to generate charts and metrics for the selected file.
* A pie chart of the target variable and a scatter matrix of all features will be displayed in separate browser tabs.
* Linear regression plots and histograms for all features will also appear in separate browser tabs.
* The "MSE" and "R2 Score" values for the linear regression model will be displayed below the "Analyze" button.
* We hope this user manual helps you effectively use the Data Analysis Application. If you have any questions or issues, please feel free to contact us.

Graphical user interface, text, application

Description automatically generated

Graphical user interface, text, application

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

C:\<some\_path>\<some\_other\_path\PycharmProjects\C964\_Capstone-Diabetes-ML-Predicator\_031923\diabetes\_data.csv

Graphical user interface, application, Word

Description automatically generated



Chart, pie chart

Description automatically generated

Diagram, application

Description automatically generated

A picture containing text, sky

Description automatically generated

Chart

Description automatically generated