Welcome to the user manual for the Data Analysis Application. This application helps users 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)
2. To install PyCharm Community Edition and required packages, follow these steps:
   1. Download PyCharm Community Edition for Windows from <https://www.jetbrains.com/edu-products/download/other-PCE.html>.

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

Description automatically generated

* 1. Double-click the downloaded file to start the installation process.
  2. Follow the installation wizard prompts to complete the installation, choosing to install for all users or just for yourself and selecting the installation location.
  3. Open PyCharm Community Edition.
  4. Click the Windows Start button to locate PyCharm Community in the JetBrains folder.
  5. Run PyCharm Community and click "Agree" to the Terms and Conditions.

Graphical user interface, text, application, email

Description automatically generated

* 1. If prompted with a Windows Security Alert message, select "Allow access."

Graphical user interface, text, application

Description automatically generated

* 1. From the Welcome to PyCharm initial menu, select the middle folder icon "Open."

Graphical user interface, application, Teams

Description automatically generated

* 1. In the Open File or Project menu, locate the saved **C964\_Capstone-Diabetes-ML-Predicator\_031923 project,** which will look something like the figure on the next page below:

Graphical user interface, text

Description automatically generated

* 1. Next select the location where you saved the Capstone project: **C964\_Capstone-Diabetes-ML-Predicator\_031923** which will look something like the figure below on the next page

Text

Description automatically generated

* 1. Click "Yes" to trust the author of the project,
  2. The project is now loaded into the PyCharm editor, but first, install the following packages:
     1. To install the **tkinter** package, go to the "**File**" menu and select "**Settings**".
        1. In the left-hand pane of the Settings/Preferences window, click "Project: "**C964\_Capstone-Diabetes-ML-Predicator\_031923**" (or "Project Interpreter" if you don't have a project open yet).
        2. In the right-hand pane, click the [ **+** ] button to add a new package.
     2. In the "**Available Packages**" window, type "**tkinter**" in the search bar and click the checkbox next to it in the search results.
        1. Click the "**Install Package**" button at the bottom right of the window.
        2. Wait for the package to install.
     3. Repeat steps **l.i. – l.ii**. for the following packages:
        1. **pandas,**
        2. **plotly,**
        3. **scikit-learn**, and
        4. **webbrowser**
     4. Once all packages have been installed, import them into your code and use them in your PyCharm projects.

1. (Optional) Download and install Python 3.6 or later from the official Python website (<https://www.python.org/downloads/>).

**To use the Data Analysis Application:**

1. Open the application using PyCharm.
2. The Data Analysis window will appear.
   1. If the "Data Analysis" screen does not "automatically" appear
   2. Check the Windows 10 taskbar
   3. You should see an icon like this: 
   4. Click the icon to get the "**Data Analysis"** screen as shown below:

Graphical user interface, text, application

Description automatically generated

1. Next, click the "**Browse**" button and select a CSV file ("**diabetes\_data.csv**") for analysis.

Graphical user interface, text, application, email

Description automatically generated

1. The selected file's path will be displayed in the text field next to the "**Browse**" button.
2. Click the "Analyze" button to generate charts and metrics for the selected file.
3. You may be prompted to agree to which application you will use to view the HTML-based data
   1. If you are prompted, you will do something like the figure below:

Graphical user interface, application, Word

Description automatically generated

* 1. Select the default web browser you are accustomed to using
  2. When you do that four HTML tabs will load on your browser
  3. If you see less than 4 HTML browser tabs close the browser and select the "**Analyze**" button again from the "**Data Analysis**" screen
  4. You should now see all four browser tabs:

1. Pie chart of the target variable and scatter matrix of all features will be displayed in separate browser tabs.

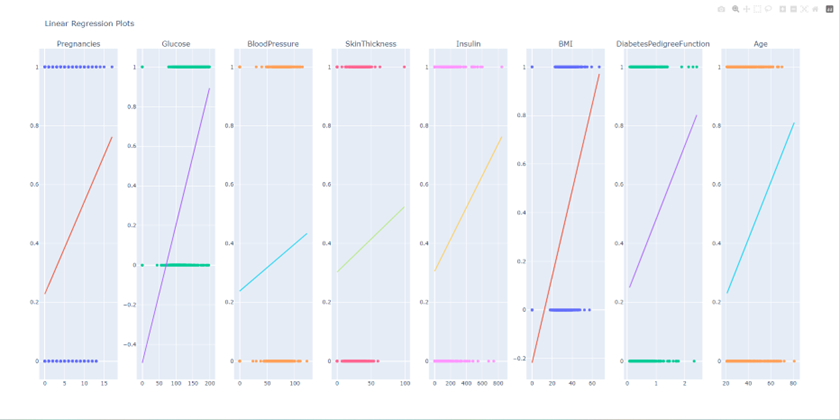
Chart, pie chart

Description automatically generated-

Diagram, application

Description automatically generated

1. Linear regression plots and histograms for all features will also appear in separate browser tabs.



Chart

Description automatically generated

1. The "MSE" and "R2 Score" values for the linear regression model will be displayed below the "Analyze" button.

The authors hope this user manual helps you use the Data Analysis Application effectively. If you have any questions or issues, please get in touch with us.

