Assignment - 4

PANDAS

What is Pandas?

Pandas is a python library for working with data sets. It has functions for analyzing, cleaning, exploring, and manipulating data

The name "Pandas" has a refrence to both "Panel Data", and "Python Data Analysis" and was created by wes Mckinney in 2008.

Why Use Pandas?

Pandas allows us to analyze big data and make conclusions based on statistical theories

Pandas can clean messy data sets, and make them readable and relevant. Relevant data is very important in data science

Functions

1. Read Data

```
import pandas as pd
df = pd.read_csv("heat_failure.csv")
```

2. Head & Trial

df.head() & df.Tail.

3. shape size and Info

df.shape() - total no of some & then columns

df.size() - returns the number of some times number of columns in the data frame

df.info() - foo is such as resource from Range Index Data columns and then data type of each columns