Assignment (1 Data Analysis

Title: Inis Data Analysis Perte of Completion: Problem Statement: - Download the Iris Flower Dataset
Or any other dataset into a
Pataframe use Python / R and Perform
following Jatatrame use Kython B and Perform
following

Thow many features are there & what

Ore their types?

2) Compute & display summary satisfies
for each feature avoidable in the dataset.

3) Data Visubization: - Create a histogram
for each feature in the dataset
to illustrate the feature distributions.

Plat each histogram.

4) Create a box plot for each feature
in the dataset. All of the boxplots
Should be combined into a single plot.
Compare distribution & identify outliers Learning Objective:) Understand dataframes and its features

a) Analyse Iris dataset. Learning Outlome: Students will be able to Malyze different datasets. Software Hardware Requirement: - OS (Linux), Python,

	Date: / /	
	SITISTAL STREET	
id.	Theory:-	
	Population May 11=5x	
	Libraries Med:	
	1) pandas x = x most deposit	
	a) Numpy	
	a) matpollib	
340	manufaction of the second of t	
	Mathematical Model: -	17
	het 8 be the system set:-	المدا
	S= 28; e; X; Y; Fme; PD; NDD; FC; Sc4 where	Dataset
	is loaded into the datatrane	
	S= Start State	0 1
	e = end state ie summany statistics for each	featur
	is captured	
	x = set of inputs	
	y = set of outputs	13
	DD -> Doleminis tic Data	4)
WE !	NDD > non deterministic Data	ALCOHOLD TO SERVE
	fc → Failure lase	
V	fc → Failure lase	
V	fc → Failure lase	1CI
٧	fc → Failure lase	19,
V &-	fc → Failure lase	19, 2
V	Data set is bollection of data Data set is bollection of data Data analysis is a process of inspecting, cleansing dransforming & modelling data with the goal of discovering useful information, information conclusion described making	19,
٧ ٠	Data set is bollection of data Data set is bollection of data Data analysis is a process of inspecting, cleansing dransforming & modelling data with the goal of discovering useful information, information conclusion described making	ig, & are the
V &.	fc → Failure lase	are the

Page:			7
Date:	1	1	

No. of Concession, Name of Street, or other Designation, Name of Street, or other Designation, Name of Street, Online of	
	Nean
	Sum of data entities divided by no of entities
	Sum of dot a entities divided by no of entities. Population Mean $u = \sum x$
	N
	Sample Mean $\bar{x} = 5 x$
	Charactery de in le
	medical devication:
	Standard deviation: - measure variability and consistency of the sample or population
	$6 = \frac{5(x-u)^2}{2}$
	N
A	Q= 28: P; X; Y: 5mp; PD: NOD; FC; Soly where
	Variance:
	averaged equared deviation from the mean
	e = and hate is bungaous thatthis ar ear
	boundary M
	Pataset Used: Ins dataset.
	thought tox = 6
	-describe ()
	diver all the parameters like mean, stateviation, varietha
	to + failure lase
	· hist ()
	creates histogram
120	and analytic of more of the state of the sta
	donner () tolyxod.
	Plot a box plot
	discion unking.
	Start and vanious business had been
	And an analysis of the state of

(0,)

Page:
Date: / /

	Test cases			
	Input	Actual Output	Expected Ofp	Remark
	describe for column 1.	mean 5.84 mean 5.84 min 4.30 max 7.90	mean 5.84 Std 2 8.82 min 4.30 max 1.9	Paned
	plot histogram	plotted	plotted	Passed
1	Boxplot	plotted	plotted	Paused
	Conclusion: Thus	9 analyzed irin	dataset su	ecces fully.