

## Assignment 01

Roblem Statement:  Perform the following operations using Python on any open source dataset:  () Import all required python bloodies.  2) Read an openiousce data From web. Provide a clear description of data & this source.  3) load the data sit into pondas clasta France.  5) Data preprocessing checking for muscing values in data using pandas ismul(), described).  Function to get some initial statistics.  6) Outa formatting of data france.  6) Outa formatting of data france.  6) Outa formatting of data normalization—  8 runcerise the type of variable by checking the data types (i.e. characters, numeric, integers fector & logical) of the variables in the data set. If you ables are not in the correct data type; apply proper conversions.  Checkers.  Checkers & logical) of the variables in the data frape; apply proper conversions.		Landren 3.8 fringles netebook
Froblem Statement:  Perform the following operations using Python on any open source dataset:  () Import all required python bloavies.  2) Read as openious ce data from web. Provide a clear description of data of the source.  3) load the data shouts pendas obsta frame.  3) load the data shouts pendas obsta frame.  3) load the data shouts pendas obsta frame.  4) loats preprocessing checking for unising values in data using pandas is null(), described)  -function to get some initial atotatics.  3) Check dimensions of data frame.  6) lata formatting of data frame.  6) lata formatting of data normalization—  summerice the type of variable by checking the data types (i.e. character, numeric, integer factor of logical) of the variables in the data set, if wariables are not in the consent data type, apply proper conversions.  Checking the data woughing of his importance in the data science cycle  20 understand data woughing of his measurement of data		Title: Data Wrongling-I
Perform the following operations using Python on any open source datasets:  () Import all required python characles.  () Read an open required python web. Provide a clear description of data of this source.  () Load the data sub into pandas data frame.  () Data preprocessing checking for muscing values in data using pandas senull(), describe()  () Function to get some invital statistics.  () Check dimensions of data frame.  () Data formatting of data frame.  () Data formatting of data frame.  () Data formatting of data normalization—  () Summerise the type of variable by checking the data types (i.e. character, numeric, integer factor of logical) of the variables in the data set, of the variables in the data type, apply proper conversions.  () Chectives:  () Operation of data wongling of life importance in the data rience cycle.		The course of th
Perform the following operations using Python on any open source datasets:  () Import all required python characles.  () Read an open required python web. Provide a clear description of data of this source.  () Load the data sub into pandas data frame.  () Data preprocessing checking for muscing values in data using pandas senull(), describe()  () Function to get some invital statistics.  () Check dimensions of data frame.  () Data formatting of data frame.  () Data formatting of data frame.  () Data formatting of data normalization—  () Summerise the type of variable by checking the data types (i.e. character, numeric, integer factor of logical) of the variables in the data set, of the variables in the data type, apply proper conversions.  () Chectives:  () Operation of data wongling of life importance in the data rience cycle.		Roblem Statement
open source datasets:  () Import all required python bloraries.  () Read an open source data from web. Provide a clear description of data & It's source.  () load the data sit into pendas clota frame.  () pata preprocessing checking for mussing values in data using pandas isnull(), describe()  -function to get some insteal statistics.  () Choit dimensions of data frame.  () pata formatting & data normalization—  sumerise the type of variable by checking the data types (i.e. character, numeric, integers fortos & logical) of the variables in the data sets of type, apply proper conversions.  () Chectives:  To understand data woongling to his importance in the data rience cycle.		
i) Import all required python bloaries.  3 Read an openiousce data from web. Provide a clear description of data & It's source.  3 load the data sib into pendas data frame.  5 Data preprocessing checking for missing values in data wing pandas ismul(), describe() function to get some initial statistics.  5 Chart dimensions of data frame.  6 Pata formatting & data normalization—  8 runexise the type of variable by checking the data types (i.e. chasacter, numeric, integer facts & logical) of the variables in the data sets of year apply proper conversions.  Chective:  pro understand data woongling of lis importance in the data science cycle.	da da da qu	open source dataset:
Read an openiousce data from web. Provide a clear description of data & It's source  3 load the data sib into pondas clota frame.  5 load the data sib into pondas clota frame.  5 load preprocessing checking for musting values in data wing pandas ismull(), describe()  -function to get some initial statistics.  6 Check dimensions of data frame.  6 lota formatting & data normalization—  8 unexise the type of variable by checking the data types (i.e. chasaites, numeric, integes factor & logical) of the variables in the data sets of type, apply proper conversions.  Checking  The data science cycle  276 understand data woongling at & IV's importance in the data science cycle	ssorts t	1) Import all required python libraries.
clear description of data & It's source  3 load the data sib into pendas data frame.  5 pata preprocessing directing for missing values in data using pandas senul(), describe()  function to get some initial statistics.  5 check dimensions of data frame.  6 pata formatting & data normalization—  8 rumexise the type of variable by checking the data types (i.e. character, numeric, integer  fatos & logical) of the variables in the data  set if variables are not in the correct data  type, apply proper conversions.  Checking to understand data coording of lis importance in  the data science cycle:  25 o understand & implement data preprocessing & data	71	2) Read an opensousce data from web. Provide a
3) load the data sib into pondas data frame.  5) Pata preprocessing cheeting for muscing values in data wing pandas ionul(), describe()  -function to get some instral statistics.  6) Cheet dimensions of data frame.  6) Pata formatting of data normalization—  8 numerise the type of variable by cheeting the data types (i.e. chasactes, numeric, integer  fortos of logical) of the variables by the data  set off variables are not in the correct data  type; apply proper conversions.  Chiedros:  pro understand data woongling of this importance in  the data science cycle:  2) To understand of implement data preprocessing of data	Lauloun	clear description of data & It's source
speak preprocessing checking for musting values  in data using pardas is null(), describe()  function to get some invital statistics.  special dimensions of data frame.  special formatting & data normalization—  summerise the type of variable by checking  the data types (i.e. chasacter, numeric, integers  factors & logical) of the variables by the data  set if variables are not in the correct data  type; apply proper conversions.  Chiefives:  pro understand data wongling of his importance in  the data science cycle.		3> load the data set into pondas docta frame.
in data wing pardar is null(), describe()  function to get some initial statistics.  Should dimensions of data frame.  Should formatting of data normalization—  summerize the type of variable by checking  the data types (i.e. character, numeric, integers  factor of logical) of the variables by the data  set of variables are not in the consect data  type, apply proper conversions.  Chectives:  To understand data wavegling at this importance in  the data science cycle.  2000 understand of implement data preprocessing & data		5) Pata preprocessing cheeking for uncerng values
School dimensions of data traine.  5) Pata formatting & data normalization—  8 numeriae the type of variable by checking  the data types (i.e. character, numeric, integer  fector & logical) of the variables in the data  Set str variables are not in the consect data  type, apply proper conversions.  Openives:  To understand data wangling at & its importance in  the data rience cycle.  2 To understand & implement data preprocessing & data		in data using pardas ionulic), describel
School dimensions of data traine.  5) Pata formatting & data normalization—  8 numeriae the type of variable by checking  the data types (i.e. character, numeric, integer  fector & logical) of the variables in the data  Set str variables are not in the consect data  type, apply proper conversions.  Openives:  To understand data wangling at & its importance in  the data rience cycle.  2 To understand & implement data preprocessing & data		function to get some initial statistics.
sumexise the type of variable by checking the data types (i.e. chasactes, numeric, integers factors & logical) of the variables in the data set strouvables are not in the consect data type, apply proper conversions.  Objectives: To understand data woongling & Livis importance in the data science cycle.  25 To understand & implement data preprocessing & data	- \$ condex	& Chort dimensions of data trans.
the data types (te. chasaites, numeric, integers factor & logical) of the variables by the data  set. It variables are not in the correct data  type, apply proper conversions.  Objectives:  To understand data waveling at & its importance in  the data science cycle.  25 o understand & implement data pre processing & data		6) Pata formatting & data normalization-
fector & logical) of the variables in the data  set. If variables are not in the correct data  type, apply proper conversions.  Objectives:  pro understand data wangling at & ils importance in  the data science cycle.  25 To understand & implement data preprocessing & data	- Sold	summeriae the type of variable by checking
Set. If variables are not in the correct data  type, apply proper conversions.  Objectives:  > To understand data wangling of this importance in  the data science cycle.  25 To understand & implement data preprocessing & data	enalysts	the data types (te. character, numeric, integer
Objectives:  To understand data wavefling of this importance in  the data science cycle.  25 To understand t implement data preprocessing & data		factor & logical) of the variables by the data
Objectives:  > To understand data wangling at & is importance in  the data science cycle.  25 To understand & implement data preprocessing & data	1.4	
Objectives:  To understand dota wavgling at its importance in  The data science cycle:  25 To understand & implement data preprocessing & data	LABIU	Type, apply proper conversions,
the data science cycle.  25 To understand & implement data preprocessing & data		
25 To understand & implement data proprocessing & data	ED WHA	Opertives:
25 To understand & implement data preprocessing & data	stal.	pro unquitang acid wongery 2 2 113 importance in
formatting functions wing pandas python library	HISO	The days syence agele.
tomasting tanonem wing powers grown wary.		25 to understand i implement adia preprocessing & adia
CONTRACTOR OF PROPERTY OF THE	in the	tomassing tanonous wing powers grown wary.
Sollings of Moodings Organizate	32.012	Sollinge & Mardiner Property to
Soffware & Mardware Regurrents:		
Processos - Intel. 1-5. 8th gen.		added and a constant of the
64-bit windows operating system		64-WF WINDOWS operating system



	kythan 3.8 frupytha notebaok.
	T-pullanoRII phot offit
	Theory:
	Boblery Stotement
פון פון	Pata wrongling:
	is It howolve a a process which involves taking
	data from it's original state to a format where
n obru	data from it's original state to a format where we can perform meaningful analysis on it.
	2) In practice there are 3 common tasks involved
מעופ	a stab Marthe process: who set that
	i) data cleaning
1)	1) data pantormation
. 10	no data ensichment
	3) data cleaning: If involves removing null values f
	atoung data in correct data tripos
DULI TON	45 data proportion at our 1/2 involved classes inc. 44.0
BALVI CAL	attitude of data as per devoustream analysis
Delacks dell	NI TORSIOURY 9217 AS CLOSEDOL & STOR
up deda	Pandas: is tan oen sold nearly at the
	plandas is an open-source library that provides
	high pertormance data manipulation in Python.
	processing such as
ni oo	resmuturing, cleaning & merging etc
	3) verila cul ditterent took available for onta
ptoh ) p	processing such as numpy scipy
prord	processing such as numpy, scipy  Panday is preferred because of its speed,
	simplicity & expressiveness, compared to other
	tools. The manual remains of the
	MIRPHR 2-1 Jahn D. August IV



1101,1 0112	
	Methodology:
	12 Suport the Required libraries
	3> Read 4 load csv into pandas datafrane
	\$ Preproces the data it involves handling misning
	values & related tasks.
	The functions at this step:
	y smill
	as begging to dropnacy
	11) describe():
	& Data formatting & normalization - checks for
	approprife data-types
	The functions used in this step:
	s) lonvert categorical variables into quontitative variables
	functions used, get-durances
	Tangrees user garage
	Carrier on.
	In this assignment lab, he has successfully
	implements data preprocessing, formatting &
hit in	implements data preprocessing, formatting & namalization using Pandas python library.
-	