PO	Date 5 03 202
Topi	28
	Experiment-9
Q	Working with datagrame
1	load Werdetails. joon file from data (local) folde
	Working with dataframe load wordetails. json file from data (local) folde to Databrick file System.
X	breate the databrame from usurdetails. ison
->	df = spark. read. json ("dlys:/Filestore/ tables/
	usudetails. json ")
	df. Snow()
79	df 1 = spark, read, format ("json"), load ("dlys:/ FileStore/tables/userdetails, json")
	df1. show()
3	Perint the schema in a true townat
7	brint the schema in a true format df. print Schema
	The second secon
4	Select only the first name column. df. select ("first name"), show()
5	df. select ("first name"), show()
1	Select the people whose id is greater than 10. a = df. filter (df ['Id'] > 10)
7	a = af. filter (af (1a) = 19)
	a. show()
	Teacher's Signature

JAINA.

POI	
Pol	
Topic	Page No. 25
6	df. gewoupky ("gender"). count(), show()
0	working with parquet format datagrame. import SBL Context package: from pyspark. sql import SBL Context
7	from kysparek. sql import SOLContext
2	Mate Salantert
2	splantext = SBLContext (SC)
3	Load 'episodes parquet' file into Databricks 'tables folder.
7	
4	Load the episodes parquet file and vulate datagrame.
->	episodeDF = Spark, read, parquet ("dlys: / Filestore) table/episodes, parquet")
5	Print schema of the dataset. episode DF. printschema ()
611)	Show the contents of the dataframe
1	Count total number of records. episodeDF. count()
	Teacher's Signature

PE	PU°
Top	ic
8	Filter the records where numbers of doctor. are greater than =. filter_episode = episode of. filter ("doctor > 5")
	filter_episode = episode pf. filter ("doctor > 5")
	pisplay filtered data using shows
ij	Save the above data in Patabricks tables folder.
	filter episode twite parquet ("alfs: 1 File store) tables / episode file")
<u> </u>	Verify the output in Databricks. myfile = spark. read. parquet ("dbfs: 1 File store 1 table) episode-file") myfile. show ()
0	Lucreking with arres format databrame. Import S&LContext package. broon pyspark. SQL import S&L Context
~	create Sallontext. Sallontext = Sallontext (SC)
3	Load 'lpisodes avro' bile from local folder into Databricks tables folder.
	Teacher's Signature

POPU°
Topic
y load the episodes auro file and create datagrame
-> episode_ouro = spark.read.format ("com. databuicts.
. spark anow'). load ("dfs:/filedtore/table) epiòods pur
5 Print schema of dataset
5 Print schema of dataset. -> episode avro, printschema ()
& Show the Contents of dataset
5 Show the Contents of dataset. -> Episode_awro.show()
) Character in the contraction of the contraction o
7 / Dunt total munches al moderned
7 Court total number of records. - episode-aww.count D
- Count
S. Cillan III a sur again de l'III anno 100 al a a a action de
8 Filter the records where no. of doctor are greater
than 3.
-> episode-filter = episode_awro, filter ("doctor>z")
2 Display feltered data using anow function () episode filter show()
- episode filter show()
Teacher's Signature

Outputs: 3(a)

11	OUT 11				(7	- 354	71 2 12
	city	email		first name	gender	id	ip address
	Miami	wheel & the	mble com	wayne	Male	da l	73.12.20.13
	Xiejea	oal len sa		Anthony	Hale	2	62.40.195, 120
			last ram	e race	timesto	am	120
		j rajin (izzi)	Bele:	Wata Rican	14634	- In-	2
			Allen	Comparche	1442		

DU+(3) 200+.

1-- cly: string (nullable = tome)

1-- amount; storing (nullable = tours)

1- first_name storing (nullable = true)

1 -- gender: string Churable = touch

1- id: long (nuelable = +rue)

1 -- ipaddress: string (melable = true)

1 -- last-name: storing (nullable = true)

1- race: strong (nullable = true)

1- timestamp: Storing chellable = true

OU+[4]

first rame
Wayne
Anthony
Evic
Timmy
Diana

Teacher's Signature

out[5]:		12 100	41.15	(1100)		generalis felicinates esterio		
-			lin	utrame.	gen	den	id	Ch add
city	emouil		υ,		0		1.0	1000
Madison	lnewyojiath	is, como no	33 16	avory s	Ma	u	18	3 (1136.
iden	awoud ayout	ulle con	8	nily	Fam	ale	13	12 219
0		last ra	M	nace	January and and a second	+ Pm	esta	nd
		needy		Venegue	lan			2688
1. 35 1.	53 194 June 6	wood	3, 1, 1	Count		145	1819	2245

OUT[6]:

1	gender	count
and the same of	Female 6	20 3
	Hale	453

1-- title: storing (nullable = true)

1-- air-date: Storing (mullable = true)
1-- doctors integer (mulable = true)

	ATTN 1885-1871	The state of the s	2011-1-10
04[6]	title	I our date	doctor
<u>}</u> .	The Eleventh hour		The state of the s
	The Doctors wife	14 Mgy 2011	
	Howar of pang Rock	3244 1072	4
0010	The Hysterious	6 Acpt 386	1 3/6 76
	Rose	26 Han - 2005	2
	Castrolana	4 Jan. 1382	

out(1):

Teacher's Signature

Out (0).

title	airdase	doctor
The Eleventh Howe	zapril to to	Marin Marin
The Doctor's wife	14 May 2011	1 1 1
The Hysterious	6 dept. 1386	65
Rose	26 May . 2005	9

90

out(s) most

1-- title: Storing (millable = tome) =

1- air dat e strong (nullable - true

OUT(6)

OUT(6)	Title	air-date	doctor	
5-1)	The Eleventh Hour.	3 April 2010	do 11 12	tis
-	The Doctors wife	Ly May2011		
	ille this to the factorion.	shork ared	= -1010	e igeaware a

out(2) 8 (Home and Islandian in)

Out [9]

Ī	The state of the s		SILES FAIR
	Title of	an air-date is do cho	The second state
	Rosa	26 Haran 2005 9	Some managements
	Castrolana	1 Den 182 5	

and the contest of the contest of

we will be a comment with it the a

Teacher's Signature