

```
17 val file = new File(filePath)
18 val fis = new FileInputStream(file)
19 prop.load(fis)
20
21 val login = prop.getProperty("login")
22 val password = prop.getProperty("password")
23
24 var misqlCon = s"jdbc:mysql://localhost:3306/spark?user=$login&password=$password"
25 var driver = "com.mysql.cj.jdbc.Driver"
26
27 val t1 = System.currentTimeMillis()
28
29 if (t1 > 0) {
30     var df = spark.read.option("delimiter", ";").
31         .option("header", "true")
32         .csv("C:/Users/Esdesu/Desktop/3re3re/ETL/Homework/ETL/work2/Task_1/fifa_s2.csv")
33
34     val df1 = df
35     df1.write.format("jdbc").option("url", misqlCon)
36         .option("driver", driver).option("dbtable", "w2t1")
37         .mode("overwrite").save()
38     df1.show()
39
40     val columns_null = df1.select(df1.columns.map(c => count(when(col(c).isNull || col(c) === "" || col(c).isNull, c)).alias(c):_*))
41     columns_null.show()
42
43     val df2 = df1.na.drop()
44     df2.write.format("jdbc").option("url", misqlCon)
45         .option("driver", driver).option("dbtable", "w2t1a")
46         .mode("overwrite").save()
47     df2.show()
48
49     val columns = df2.columns.map(c => sum(col(c).isNull.cast("integer")).alias(c))
50     val dataf = df2.agg(columns.head, columns.tail:_)
51     val t = dataf.columns.map(c => dataf.select(lit(c).alias("col_name"), col(c).alias("null_count")))
52     val dataf_agg_col = t.reduce((df2, dataf) => df2.union(dataf))
53     dataf_agg_col.show()
54
55     val df3 = df2
56         .withColumn("Name", lower(col("Name")))
57         .withColumn("Nationality", lower(col("Nationality")))
58         .withColumn("Club", lower(col("Club")))
59         .withColumn("Preferred Foot", lower(col("Preferred Foot")))
60         .withColumn("Position", lower(col("Position")))
61         .dropDuplicates()
62     df3.write.format("jdbc").option("url", misqlCon)
63         .option("driver", driver).option("dbtable", "w2t1b")
64         .mode("overwrite").save()
65     df3.show()
66 }
67
68
69 val s0 = (System.currentTimeMillis() - t1)/1000
70 val s = s0 % 60
71 val m = (s0/60) % 60
72 val h = (s0/60/60) % 24
73 println("%02d:%02d:%02d".format(h, m, s))
74 System.exit(0)
75
```

8205	Braga	35	Portugal	70	70	CD Aves	750	5	Right	1	3	LCB	2017	01/01/2019	5.916666667	165	1600
9014	A. Robben	34	Netherlands	84	84	FC Bayern München	15500	110	Left	4	4	RW	2009	01/01/2019	5.916666667	176	25600
19521	J. Beheere	35	England	61	61	Cambridge United	120	2	Right	1	2	ST	2017	01/01/2019	6.333333333	203	210
20209	Ricardo Costa	37	Portugal	74	74	CD Tondel	750	4	Right	2	2	LCB	2017	01/01/2019	6.333333333	181	1600
20801	Cristiano Ronaldo	39	Portugal	86	86	Juventus	77000	405	Right	5	5	ST	2018	01/01/2022	6.166666667	183	127100
27532	D. de Vries	37	Netherlands	66	66	Celtic	40	7	Right	1	1	GC	2016	01/01/2019	6.25	212	721
45983	S. von Bergen	35	Switzerland	73	73	BSC Young Boys	950	12	Right	2	2	LCB	2013	01/01/2022	6.1	174	1400
46175	M. Rosenberg	35	Sweden	73	73	Malmö FF	1800	6	Right	1	3	RS	2014	01/01/2020	6.1	179	2400
47201	S. Protzo	35	Belgium	76	76	Lazio	1800	20	Right	2	1	GC	2018	01/01/2021	6.163	163	3100
52400	M. Lehmann	35	Germany	72	72	1. FC Köln	725	10	Right	2	2	CB	2012	01/01/2019	5.833333333	165	1100
52679	J. Murphy	36	Republic of Ireland	62	62	Bury	40	1	Left	1	1	GC	2017	01/01/2019	6.166666667	194	70
53110	C. Gordon	35	Scotland	74	74	Celtic	1300	20	Left	2	1	GC	2014	01/01/2020	6.333333333	183	2400
53363	A. Collin	33	England	62	62	Carlisle United	150	2	Right	1	1	GC	2018	01/01/2019	6.25	172	263
53749	L. Grant	39	England	74	74	Manchester United	1300	39	Right	1	1	GC	2018	01/01/2020	6.333333333	183	2500
54000	S. Bwester	37	England	63	63	Burton Albion	20	1	Right	1	1	GC	2016	01/01/2019	6.083333333	179	35
54037	T. Berni	35	Italy	64	64	Inter	110	5	Right	1	1	GC	2017	01/01/2019	6.083333333	176	187
101800	R. Green	38	England	72	72	Chelsea	210	28	Right	2	1	GC	2018	01/01/2019	6.166666667	203	389
102056	S. Padoin	34	Italy	71	71	Cagliari	675	18	Right	2	3	RB	2016	01/01/2019	5.833333333	161	1100
108000	S. Lichtsteiner	34	Switzerland	79	79	Arsenal	3600	94	Right	3	3	RB	2018	01/01/2021	6.1	168	6000

only showing top 20 rows

	col_name	null_count
ID		0
Name		0
Age		0
Nationality		0
Overall		0
Potential		0
Club		0
Value		0
Wage		0
Preferred Foot		0
International Rep...		0
Skill Moves		0
Position		0
Joined		0
Contract Valid Until		0
Height		0
Weight		0
Release Clause		0

ID	Name	Age	Nationality	Overall	Potential	Club	Value	Wage	Preferred Foot	International Reputation	Skill Moves	Position	Joined	Contract Valid Until	Height	Weight	Release Clause
117766	g. bong	30	cameroon	74	74	brighton & hove a...	4100	35	left	1	3	lb	2015	01/01/2019	6.166666667	163	7700
184599	jaime rowser	27	spain	71	71	córdoba cf	2500	7	left	1	3	rw	2017	01/01/2020	5.75	170	3000
206551	roberto torres	29	spain	72	72	ca osasuna	3000	9	right	1	3	cm	2011	01/01/2020	5.833333333	159	4900
223402	e. tasdemir	22	turkey	71	78	bursaspor	2800	9	left	1	3	lb	2014	01/01/2019	5.75	154	6200
233125	m. hartel	22	germany	69	76	1. fc union berlin	1500	14	right	1	3	lw	2017	01/01/2020	5.75	152	2700
233164	a. muric	19	montenegro	61	78	manchester city	450	3	right	1	1	gc	2017	01/01/2020	6.5	179	1100
238516	l. estigland	38	norway	61	81	brighton & hove a...	475	3	right	1	2	cb	2018	01/01/2021	6.1	176	1400
240873	genaro	20	spain	61	71	sevilla fc	400	2	right	1	2	cm	2017	01/01/2021	6.1	157	920
240926	n. phillips	21	england	61	74	liverpool	425	9	right	1	2	cb	2017	01/01/2021	6.166666667	154	1100
241651	v. gyökeres	20	sweden	62	79	brighton & hove a...	675	6	right	1	3	st	2018	01/01/2020	6.166666667	168	1700
244502	a. paloczic	24	serbia	70	75	cd nacional	2300	6	right	1	3	cm	2018	01/01/2021	5.916666667	146	5100
117505	linares	35	spain	69	69	cf reus deportiu	475	4	right	1	2	st	2018	01/01/2019	5.916666667	172	736
1178518	r. nainggolan	30	belgium	85	85	inter	35500	105	right	3	3	cm	2018	01/01/2022	5.75	170	60400
191740	ander herrera	28	spain	81	81	manchester united	17500	140	right	3	3	cm	2014	01/01/2019	6.1	154	32400
197242	g. krychowiak	28	poland	78	79	lokomotiv moscow	9500	1	right	3	3	cm	2016	01/01/2021	6.083333333	187	20700
200714	h. hope	24	barbados	68	65	carlisle united	280	2	right	1	2	st	2017	01/01/2019	5.833333333	160	511
212244	c. lenz	23	germany	66	72	1. fc union berlin	750	7	left	1	3	lb	2016	01/01/2019	5.916666667	174	1400
214440	d. melo	24	chile	69	76	cd palestino	1100	2	right	1	1	gc	2011	01/01/2020	5.833333333	168	1700
227676	m. kennedy	21	northern ireland	53	63	charlton athletic	100	1	right	1	2	rm	2014	01/01/2019	5.916666667	168	195
234014	jorginobauer	34	brazil	68	68	chapecoense	450	7	right	1	3	cm	2018	01/01/2021	5.833333333	152	855

only showing top 20 rows