

Из задания которое разбирали на семинаре взять запрос и использовать его в скаловском скрипте, через спарк создать вторую таблицу.
И сделать через GROUP_CONCAT список как на семинаре, Сделайте формат времени 11.01.23 22.29 (без секунд) и через case сократите названия статусов (Например, Зарегистрирован 3, Закрыт --> 3T).

SQL запрос №1:

```
SELECT `ticket`, FROM_UNIXTIME(`StatusTime`) `StatusTime`,
      IFNULL((LEAD(`StatusTime`) OVER(PARTITION BY `ticket` ORDER BY `StatusTime`) - `StatusTime`) /
3600, 0) AS Timers,
      CASE WHEN `status` IS NULL THEN @PREF1
      ELSE @PREF1:=`status`
      END `status`,
      CASE WHEN `group` IS NULL THEN @PREF2
      ELSE @PREF2:=`group`
      END `group`
FROM
(SELECT `ticket`, `StatusTime`, `status`,
IF(ROW_NUMBER() OVER(PARTITION BY `ticket` ORDER BY `StatusTime`) = 1 AND `Destination` IS NULL,
", `group`) `group`
FROM
(SELECT DISTINCT a.objectid `ticket`, a.restime `StatusTime`, `status`, `group`, `Destination`,
(SELECT @PREF1:= ""),
(SELECT @PREF2:= "")
FROM
(SELECT DISTINCT a.objectid, a.restime FROM spark.w3t5v1 AS a
WHERE fieldname IN ('status' AND 'GNAME2')) AS a
LEFT JOIN
      (SELECT DISTINCT `objectid`, `restime`, `fieldvalue` 'status' FROM spark.w3t5v1
      WHERE fieldname IN ('status')) AS a1
ON a.objectid = a1.objectid AND a.restime = a1.restime
LEFT JOIN
      (SELECT DISTINCT `objectid`, `restime`, `fieldvalue` 'group', 1 `Destination` FROM spark.w3t5v1
      WHERE fieldname IN ('GNAME2')) AS a2
```

ON a.objectid = a2.objectid AND a.restime = a2.restime) AS b1

ORDER BY 1, 2) AS b2

spark.w3t5v1: 34 строк (exact)						
#	objectid	restime	desc	fieldvalue	fieldname	
1	15	1 653 305 822	'Группа исполнителя' изменена ...	ЦМИК\СДиА	GNAME2	
2	15	1 653 305 822	Зона ответственности' изменен н...	МГТС	Resp_zone	
3	15	1 653 305 822	'Статус' изменен на:Зарегистри...	Зарегистрирован	Status	
4	15	1 653 305 823	'Статус' изменен на:Назначен	Назначен	Status	
5	15	1 653 306 133	'Статус' изменен на:В работе	В работе	Status	
6	15	1 653 320 199	'Группа исполнителя'измененан...	ЭЛС\АВР	GNAME2	
7	15	1 653 320 199	'Статус' изменен на:Назначен	Назначен	Status	
8	15	1 653 322 363	'Статус' изменен на:В работе	В работе	Status	
9	15	1 653 322 906	'Группа исполнителя'измененан...	ЦАВРИР\ФГК АВР	GNAME2	
10	15	1 653 305 822	Зона ответственности' изменен н...	Клиент	Resp_zone	
11	15	1 653 322 906	'Статус' изменен на:Назначен	Назначен	Status	
12	15	1 653 324 076	'Группа исполнителя'измененан...	ЦАВРИР\ОСПД\ГРОТИФС	GNAME2	
13	15	1 653 324 088	'Статус' изменен на:В работе	В работе	Status	
14	15	1 653 328 511	'Статус' изменен на:Назначен	Назначен	Status	
15	15	1 653 328 511	'Группа исполнителя'измененан...	ЦМИК\СДиА	GNAME2	
16	15	1 653 328 591	'Группа исполнителя'измененан...	ЭЛС\АВР	GNAME2	
17	15	1 653 328 726	'Статус' изменен на:В работе	В работе	Status	
18	15	1 653 364 688	'Группа исполнителя'измененан...	ЭЛС\Север	GNAME2	
19	15	1 653 364 688	'Статус' изменен на:Назначен	Назначен	Status	
20	15	1 653 365 165	'Статус' изменен на:В работе	В работе	Status	
21	15	1 653 365 268	'Статус' изменен на:Решен	Решен	Status	
22	15	1 653 365 279	'Статус' изменен на:Исследован...	Исследование ситуации	Status	
23	15	1 653 365 287	'Статус' изменен на:Закрыт	Закрыт	Status	
24	16	1 659 305 822	Зона ответственности' изменен н...	МГТС	Resp_zone	
25	16	1 659 305 822	'Статус' изменен на:Зарегистри...	Зарегистрирован	Status	
26	16	1 659 305 922	Группа исполнителя'измененан...	ОТЭ	GNAME2	
27	16	1 659 305 932	'Статус' изменен на:Назначен	Назначен	Status	
28	16	1 659 306 922	'Статус' изменен на:В работе	В работе	Status	
29	16	1 659 306 922	Группа исполнителя'измененан...	ОТПМР	GNAME2	
30	16	1 659 306 922	Зона ответственности' изменен н...	Клиент	Resp_zone	
31	16	1 659 310 922	Группа исполнителя'измененан...	ЭЛС	GNAME2	
32	16	1 659 311 922	'Статус' изменен на:Решен	Решен	Status	
33	16	1 659 311 950	Группа исполнителя'измененан...	ОТЭ	GNAME2	
34	16	1 659 311 950	'Статус' изменен на:Закрыт	Закрыт	Status	

spark.w3t5v1a: 23 строк (exact)						
#	tiket	StatusTime	Timers	status	group	
1	15	2022-05-23 16:37:02	0,0003	Зарегистрирован	ЦМИК\СДиА	
2	15	2022-05-23 16:37:03	0,0861	Назначен	ЦМИК\СДиА	
3	15	2022-05-23 16:42:13	3,9072	В работе	ЦМИК\СДиА	
4	15	2022-05-23 20:36:39	0,6011	Назначен	ЭЛС\АВР	
5	15	2022-05-23 21:12:43	0,1508	В работе	ЭЛС\АВР	
6	15	2022-05-23 21:21:46	0,325	Назначен	ЦАВРИР\ФГК АВР	
7	15	2022-05-23 21:41:16	0,0033	Назначен	ЦАВРИР\ОСПД\ГРОТИФС	
8	15	2022-05-23 21:41:28	1,2286	В работе	ЦАВРИР\ОСПД\ГРОТИФС	
9	15	2022-05-23 22:55:11	0,0222	Назначен	ЦМИК\СДиА	
10	15	2022-05-23 22:56:31	0,0375	Назначен	ЭЛС\АВР	
11	15	2022-05-23 22:58:46	9,9894	В работе	ЭЛС\АВР	
12	15	2022-05-24 08:58:08	0,1325	Назначен	ЭЛС\Север	
13	15	2022-05-24 09:06:05	0,0286	В работе	ЭЛС\Север	
14	15	2022-05-24 09:07:48	0,0031	Решен	ЭЛС\Север	
15	15	2022-05-24 09:07:59	0,0022	Исследование ситуации	ЭЛС\Север	
16	15	2022-05-24 09:08:07	0,0	Закрыт	ЭЛС\Север	
17	16	2022-08-01 03:17:02	0,0278	Зарегистрирован		
18	16	2022-08-01 03:18:42	0,0028	Зарегистрирован	ОТЭ	
19	16	2022-08-01 03:18:52	0,275	Назначен	ОТЭ	
20	16	2022-08-01 03:35:22	1,1111	В работе	ОТПМР	
21	16	2022-08-01 04:42:02	0,2778	В работе	ЭЛС	
22	16	2022-08-01 04:58:42	0,0078	Решен	ЭЛС	
23	16	2022-08-01 04:59:10	0,0	Закрыт	ОТЭ	

```
SELECT `tiket`, GROUP_CONCAT(' ',
CASE
    WHEN DATE(`StatusTime`) = CURDATE() THEN DATE_FORMAT(`StatusTime`, '%Y-%m-%d
%H:%i:%s')
ELSE DATE_FORMAT(`StatusTime`, '%d-%m-%Y %H:%i')
END,
', ',
CASE
    WHEN `status` = 'Зарегистрирован' THEN 'З'
    WHEN `status` = 'Назначен' THEN 'Н'
    WHEN `status` = 'В работе' THEN 'ВР'
    WHEN `status` = 'Закрыт' THEN 'ЗТ'
    WHEN `status` = 'Исследование ситуации' THEN 'ИС'
    WHEN `status` = 'Решен' THEN 'Р'
ELSE `status`
END,
' -> ', `group`, ' ' SEPARATOR '\r\n') AS 'new format'
FROM spark.w3t5a
GROUP BY 1
```

[illegible]

```
/*
chcp 65001 && spark-shell -i
C:\Users\Esdusu\Desktop\JreJre\ETL\HomeWork\ETL\Work#3\Task_5\Task_1_v2.scala --
conf "spark.driver.extraJavaOptions=-Dfile.encoding=utf-8"
*/
```

```

import org.apache.spark.internal.Logging
import org.apache.spark.sql.functions.{col, collect_list, concat_ws}
import org.apache.spark.sql.{DataFrame, SparkSession}
import org.apache.spark.sql.expressions.Window
import scala.io.Source

sc.setLogLevel("ERROR")

val login =
Source.fromFile("C:/Users/Esdesu/Desktop/JreJre/ETL/config.txt").getLines.toList

val t1 = System.currentTimeMillis()

if(1==1){
  var df = spark.read.option("delimiter",",")
    .option("inferSchema", "true")
    .option("header", "true")
    .format("excel")
    .load("C:/Users/Esdesu/Desktop/JreJre/ETL/HomeWork/ETL/Work#3/Task_5/s3.xlsx")

  val df1 = df
    df1.write.format("jdbc").option("url", login(0))
      .option("driver", login(1)).option("dbtable", "w3t5v2")
      .mode("overwrite").save()
  df1.show()

  val df2 = spark.read.format("jdbc").option("url", login(0))
    .option("driver", login(1))
    .option("dbtable", "w3t5v2")
    .load()

  val df_group = df2.distinct().where(col("fieldname") === "GNAME2")
    .select("objectid", "restime", "fieldvalue")
    .withColumnRenamed("fieldvalue", "Group")
    .withColumn("Destination", lit("1").cast("integer"))

  val df_status = df2.distinct().where(col("fieldname") === "Status")
    .select("objectid", "restime", "fieldvalue")
    .withColumnRenamed("fieldvalue", "Status")

  val df_sg = df2.filter((col("fieldname") isin ("status", "GNAME2")))
    .select("objectid", "restime").distinct()

  val df_inner = df_sg.as("a")
    .join(df_status.as("a1"),col("a.objectid") === col("a1.objectid") &&
col("a.restime") === col("a1.restime"),"left")
    .join(df_group.as("a2"),col("a.objectid") === col("a2.objectid") &&
col("a.restime") === col("a2.restime"),"left")
    .select(col("a.objectid"),col("a.restime"),col("a1.Status"),col("a2.Group"),col("a2.Destination"))

```

```

        .withColumnRenamed("objectid", "Tiket")
        .withColumnRenamed("restime", "StatusTime")
        .distinct()

    val df_outer =
df_inner.select(col("Tiket"),col("StatusTime"),col("Status"),when(row_number().over(
Window.partitionBy(col("Tiket"))
        .orderBy(col("StatusTime"))) === 1 &&
col("Destination").isNull,"").otherwise(col("Group")).alias("Group"),col("Destina
tion"))

    val df_result =
df_outer.select(col("Tiket"),from_unixtime(col("StatusTime")).alias("StatusTime")
,((lead(col("StatusTime"), 1)
        .over(Window.partitionBy(col("Tiket")).orderBy(col("StatusTime"))) -
col("StatusTime")) / 3600).alias("Timers"),last(col("Status"), true)
        .over(Window.partitionBy(col("Tiket")).orderBy(col("StatusTime")))
        .alias("Status"),last(col("Group"),
true).over(Window.partitionBy(col("Tiket")).orderBy(col("StatusTime")))
        .alias("Group"),col("Destination"))
        .withColumn("Timers", coalesce(col("Timers"), lit(0)))
        .withColumn("Timers", round(col("Timers"), 4))

df_result.write.format("jdbc").option("url", login(0))
        .option("driver", login(1)).option("dbtable", "w3t5v2a")
        .mode("overwrite").save()
df_result.show()

val df3 = spark.read.format("jdbc").option("url", login(0))
        .option("driver", login(1))
        .option("dbtable", "w3t5v2a")
        .load()

val df3_concat = df3.groupBy("Tiket")
        .agg(concat_ws(".\r\n", collect_list(concat_ws("
",when(date_format(col("StatusTime"), "yyyy-MM-dd") ===
current_date(),date_format(col("StatusTime"), "yyyy-MM-dd HH:mm:ss"))
        .otherwise(date_format(col("StatusTime"), "dd-MM-yyyy HH:mm")),
concat_ws(" -> ",when(col("Status") === "Зарегистрирован", "З")
        .when(col("Status") === "Назначен", "Н")
        .when(col("Status") === "В работе", "ВР")
        .when(col("Status") === "Закрыт", "ЗТ")
        .when(col("Status") === "Исследование ситуации", "ИС")
        .when(col("Status") === "Решен", "Р"),col("Group")))))
        .alias("new format"))
        .withColumn("new format", concat(col("new format"),lit(".")))
        .withColumn("Tiket",col("Tiket"))

df3_concat.write.format("jdbc").option("url", login(0))
        .option("driver", login(1)).option("dbtable", "w3t5v2b")
        .mode("overwrite").save()

```

```

df3_concat.show()

println("Work 3, Task 5, Successful Load and Save")

}

val s0 = (System.currentTimeMillis() - t1)/1000
val s = s0 % 60
val m = (s0/60) % 60
val h = (s0/60/60) % 24
println("%02d:%02d:%02d".format(h, m, s))
System.exit(0)

```

Филтр баз данн | Филтр таблиц | Хост: 127.0.0.1 | База данных: spark | Таблица: w3t5v2 | Данные | Запрос* | Запрос #3*

spark.w3t5v2: 34 строк (exact)

#	objectid	restime	desc	fieldvalue	fieldname
1	15	1 653 305 822	'Группа исполнителя' изменена ...	ЦМИК\СДиА	GNAME2
2	15	1 653 305 822	Зона ответственности' изменен н...	МГТС	Resp_zone
3	15	1 653 305 822	'Статус' изменен на:Зарегистри...	Зарегистрирован	Status
4	15	1 653 305 823	'Статус' изменен на:Назначен	Назначен	Status
5	15	1 653 306 133	'Статус' изменен на:В работе	В работе	Status
6	15	1 653 320 199	'Группа исполнителя'измененан...	ЭЛС\АВР	GNAME2
7	15	1 653 320 199	'Статус' изменен на:Назначен	Назначен	Status
8	15	1 653 322 363	'Статус' изменен на:В работе	В работе	Status
9	15	1 653 322 906	'Группа исполнителя'измененан...	ЦАВРиР\ФГК АВР	GNAME2
10	15	1 653 305 822	Зона ответственности' изменен н...	Клиент	Resp_zone
11	15	1 653 322 906	'Статус' изменен на:Назначен	Назначен	Status
12	15	1 653 324 076	'Группа исполнителя'измененан...	ЦАВРиР\ОСПД\ГРОТиФС	GNAME2
13	15	1 653 324 088	'Статус' изменен на:В работе	В работе	Status
14	15	1 653 328 511	'Статус' изменен на:Назначен	Назначен	Status
15	15	1 653 328 511	'Группа исполнителя'измененан...	ЦМИК\СДиА	GNAME2
16	15	1 653 328 591	'Группа исполнителя'измененан...	ЭЛС\АВР	GNAME2
17	15	1 653 328 726	'Статус' изменен на:В работе	В работе	Status
18	15	1 653 364 688	'Группа исполнителя'измененан...	ЭЛС\Север	GNAME2
19	15	1 653 364 688	'Статус' изменен на:Назначен	Назначен	Status
20	15	1 653 365 165	'Статус' изменен на:В работе	В работе	Status
21	15	1 653 365 268	'Статус' изменен на:Решен	Решен	Status
22	15	1 653 365 279	'Статус' изменен на:Исследован...	Исследование ситуации	Status
23	15	1 653 365 287	'Статус' изменен на:Закрыт	Закрыт	Status
24	16	1 659 305 822	Зона ответственности' изменен н...	МГТС	Resp_zone
25	16	1 659 305 822	'Статус' изменен на:Зарегистри...	Зарегистрирован	Status
26	16	1 659 305 922	Группа исполнителя'измененан...	ОТЭ	GNAME2
27	16	1 659 305 932	'Статус' изменен на:Назначен	Назначен	Status
28	16	1 659 306 922	'Статус' изменен на:В работе	В работе	Status
29	16	1 659 306 922	Группа исполнителя'измененан...	ОТПМР	GNAME2
30	16	1 659 306 922	Зона ответственности' изменен н...	Клиент	Resp_zone
31	16	1 659 310 922	Группа исполнителя'измененан...	ЭЛС	GNAME2
32	16	1 659 311 922	'Статус' изменен на:Решен	Решен	Status
33	16	1 659 311 950	Группа исполнителя'измененан...	ОТЭ	GNAME2
34	16	1 659 311 950	'Статус' изменен на:Закрыт	Закрыт	Status

The screenshot shows the PostgreSQL pgAdmin interface. The left sidebar displays the database structure, including the 'information_schema' database and the 'spraw' table. The main window shows the 'spraw' table with columns 'id', 'titel', and 'new format'. The data is filtered to show rows where 'id' is 1 or 2. The table size is 160.0 KB. The 'new format' column contains various dates and times, some with additional information in parentheses.

#	id	titel	new format
1	13	23-05-2022 16:37, 3 ->	ЦММК/Сл.В.А. 23-05-2022 16:37, Н -> ЦММК/Сл.В.А. 23-05-2022 16:42, ВР -> ЦММК/Сл.В.А. 23-05-2022 20:36, Н -> ЗПС/АВР. 23-05-2022 21:12, ВР -> ЗПС/АВР. 23-05-2022 21:21, Н -> ЦАВР/П/ВТК. АВР. 23-05-2022 21:41, Н -> ...
2	16	01-09-2022 03:17, 3 ->	01-09-2022 03:18, 3 -> ОТС 01-09-2022 03:18, Н -> ОТС 01-09-2022 03:25, ВР -> ОТС 01-09-2022 04:42, ВР -> ЗПС 01-09-2022 04:58, П -> ЗПС 01-09-2022 04:58, СТ -> ОТС

```
Командная строка
Microsoft Windows [Version 10.0.19045.4291]
(c) Корпорация Майкрософт (Microsoft Corporation). Все права защищены.

C:\Users\Esdesu>chcp 65001 && spark-shell -i C:\Users\Esdesu\Desktop\JreJre\ETL\HomeWork\ETL\Work#3\Task_5\Task_1_v2.scala --conf "spark.driver.extraClassPath=C:\Users\Esdesu\Desktop\JreJre\ETL\HomeWork\ETL\Work#3\Task_5\Task_1_v2\lib\*.jar"
Active code page: 65001
Using Spark's default log4j profile: org/apache/spark/log4j-defaults.properties
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).
Spark context Web UI available at http://Alicorn-113-St-527-BSF-13:4040
Spark context available as 'sc' (master = local[*], app id = local-1713964525355).
Spark session available as 'spark'.
ERROR StatusLogger Log4j2 could not find a logging implementation. Please add log4j-core to the classpath. Using SimpleLogger to log to the console

+-----+-----+-----+-----+-----+
|objectid|  restime|      desc|  fieldvalue|fieldname|
+-----+-----+-----+-----+-----+
|15|1653305822|'Группа исполните...|      ЦМИК\СДИА|  GNAME2|
|15|1653305822|Зона ответственнос...|      МГТС|Resp_zone|
|15|1653305822|'Статус' изменен ...|Зарегистрирован|  Status|
|15|1653305823|'Статус' изменен ...|Назначен|  Status|
|15|1653306133|'Статус' изменен ...|В работе|  Status|
|15|1653320199|'Группа исполните...|      ЭЛС\АВР|  GNAME2|
|15|1653320199|'Статус' изменен ...|Назначен|  Status|
|15|1653322363|'Статус' изменен ...|В работе|  Status|
|15|1653322906|'Группа исполните...|ЦАВРиР\ФГК АВР|  GNAME2|
|15|1653305822|Зона ответственнос...|      Клиент|Resp_zone|
|15|1653322906|'Статус' изменен ...|Назначен|  Status|
|15|1653324076|'Группа исполните...|ЦАВРиР\ОСПД\ГРОТИФС|  GNAME2|
|15|1653324088|'Статус' изменен ...|В работе|  Status|
|15|1653328511|'Статус' изменен ...|Назначен|  Status|
|15|1653328511|'Группа исполните...|      ЦМИК\СДИА|  GNAME2|
|15|1653328591|'Группа исполните...|      ЭЛС\АВР|  GNAME2|
|15|1653328726|'Статус' изменен ...|В работе|  Status|
|15|1653364688|'Группа исполните...|      ЭЛС\Север|  GNAME2|
|15|1653364688|'Статус' изменен ...|Назначен|  Status|
|15|1653365165|'Статус' изменен ...|В работе|  Status|
+-----+-----+-----+-----+-----+
only showing top 20 rows

+-----+-----+-----+-----+-----+
|Tiket|      StatusTime|Timers|      Status|      Group|Destination|
+-----+-----+-----+-----+-----+
|15|2022-05-23 16:37:02|3.0E-4|Зарегистрирован|      ЦМИК\СДИА|1|
|15|2022-05-23 16:37:03|0.0861|Назначен|      ЦМИК\СДИА|null|
|15|2022-05-23 16:42:13|3.9072|В работе|      ЦМИК\СДИА|null|
|15|2022-05-23 20:36:39|0.6011|Назначен|      ЭЛС\АВР|1|
|15|2022-05-23 21:12:43|0.1508|В работе|      ЭЛС\АВР|null|
|15|2022-05-23 21:21:46|0.325|Назначен|ЦАВРиР\ФГК АВР|1|
|15|2022-05-23 21:41:16|0.0033|Назначен|ЦАВРиР\ОСПД\ГРОТИФС|1|
|15|2022-05-23 21:41:28|1.2286|В работе|ЦАВРиР\ОСПД\ГРОТИФС|null|
|15|2022-05-23 22:55:11|0.0222|Назначен|      ЦМИК\СДИА|1|
|15|2022-05-23 22:56:31|0.0375|Назначен|      ЭЛС\АВР|1|
|15|2022-05-23 22:58:46|9.9894|В работе|      ЭЛС\АВР|null|
|15|2022-05-24 08:58:08|0.1325|Назначен|      ЭЛС\Север|1|
|15|2022-05-24 09:06:05|0.0286|В работе|      ЭЛС\Север|null|
|15|2022-05-24 09:07:48|0.0031|Решен|      ЭЛС\Север|null|
|15|2022-05-24 09:07:59|0.0022|Исследование ситу...|      ЭЛС\Север|null|
|15|2022-05-24 09:08:07|0.0|Закрит|      ЭЛС\Север|null|
|16|2022-08-01 03:17:02|0.0278|Зарегистрирован||null|
|16|2022-08-01 03:18:42|0.0028|Зарегистрирован|      ОТЭ|1|
|16|2022-08-01 03:18:52|0.275|Назначен|      ОТЭ|null|
|16|2022-08-01 03:35:22|1.1111|В работе|      ОТПМР|1|
+-----+-----+-----+-----+-----+
only showing top 20 rows

+-----+-----+
|Tiket|      new format|
+-----+-----+
|16|01-08-2022 03:17,...|
|15|23-05-2022 16:37,...|
+-----+-----+

Work 3, Task 5, Successful Load and Save
00:00:16

C:\Users\Esdesu>
```