how to extract text from generate Tablefetcha and input it to executeSQL

<https://community.hortonworks.com/questions/138406/how-to-extract-text-from-generatetablefetcha-and-i.html>

I understand I need to use ExtractText but it does not output anything to executeSQL processor. How do I do that?

I have not added anything additional in the config of the ExtractText processor but mapped both matched and unmatched queue to the executSQL processor.

WHat do I pass in SQL select query in executeSQL processor?

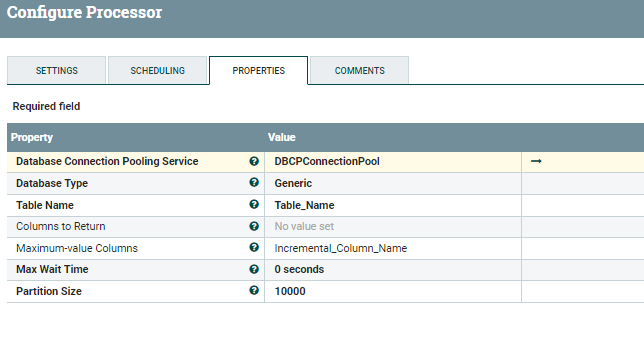
you can do this in 2 methods

1. by using **GenerateTableFetch** Processor

2. byusing **QueryDatabaseTable** Processor

**Method1:-**

**GenerateTableFetch** processor gives incremental sql statements based on partition size



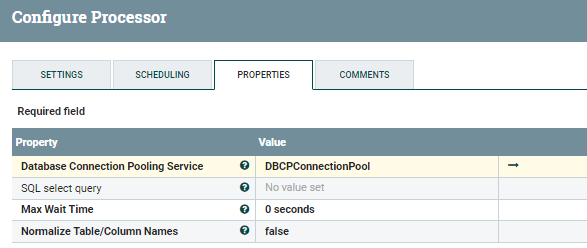
In my configuration i have given PartitionSize 10000 so the output of the table fetch processor will be

1. SELECT \* FROM Table\_Name ORDER BY Incremental\_column\_name LIMIT 10000
2. If you are using **NiFi 1.3** then the **Fetch statements** would be in **correct syntax** as i'm using **1.1** i **don't** have any **DatabaseType** property i need to select only **Generic** it will result the above fetch statements.
3. i need to use **replace text processor** to correct these fetch statements before passing statements to ExecuteSql processor.
4. But in new versions we are having Specific Database Type which will generates correct syntax fetch statements, you dont need to use replace text processors.
5. Just connect **Success** relation of **Tablefetch** processor to **ExecuteSQL**.

**ExecuteSQL Processor:-**

Drag and drop an executesql processor and don't mention any sql statements in it, as this processor executes all the sql statements that got generated by GenerateTableFetch processor.

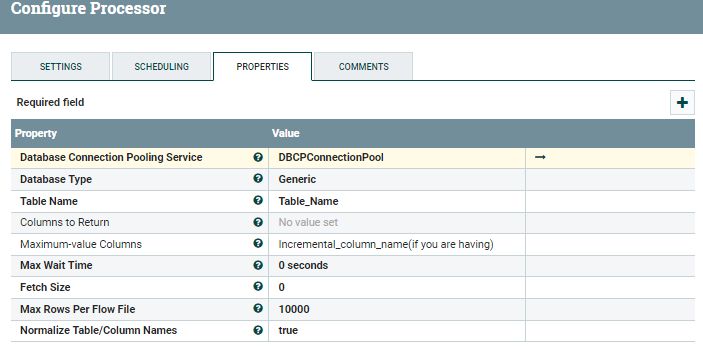
In nifi every executesql,querydatabase processor results default format is avro.



**Method2:-**

**QueryDatabaseTable** Processor does the **both two steps(generating table fetch and executing sql)** and **stores the state** as GenerateTableFetch processor does and as **output** we will have **data in Avro format**(default in NiFi)

**Configs:-**



As a result from both the methods above we will have Avro format data and having not more than 10000(we mentioned in Processor) records in them.

We cannot extract text from Avro format so we need to convert avro format to Json.

**ConvertAvroToJSON processor:-**

in new version of NiFi there is Convert record processor which will convert Avro format data to CSV directly but in my version i dont have this processor.

This processor will converts all the avro to json data once we get json data then you can store the json data directly to HDFS... etc.

If you are thinking to store the data as text format not as Json then you need to Split the json array as individual records

**Split Array using SplitJson Processor:-**

We are having array of json message so use splitjson processor and property for **JsonPath Expression** as

1. $.\*

It will splits the json array to individual records so each flowfile will have one record data in it.

**Input to SplitJson:-**

1. [{"id":1,"first\_name":"Jeanette","last\_name":"Penddreth","email":"jpenddreth0@census.gov","gender":"Female"},{"id":2,"first\_name":"Jte","last\_name":"ght","email":"ddreth0@census.gov","gender":"male"}]

**Output:-**

**Flowfile1:-**

1. {"id":1,"first\_name":"Jeanette","last\_name":"Penddreth","email":"jpenddreth0@census.gov","gender":"Female"}

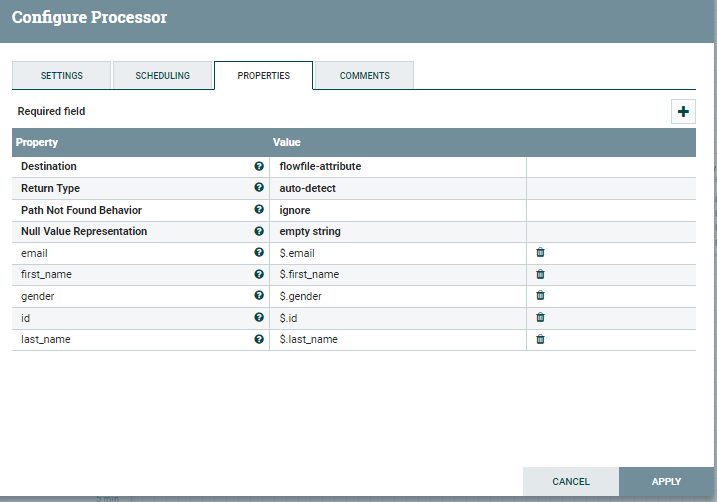
**flowfile2:-**

1. {"id":2,"first\_name":"Jte","last\_name":"ght","email":"ddreth0@census.gov","gender":"male"}

Once we splitted all the json array then use

**EvaluateJsonPath processor configs:-**

use the Destination property to **flowfile-attribute**,then add the list required of attributes that you need to be extracted as attributes from the content.

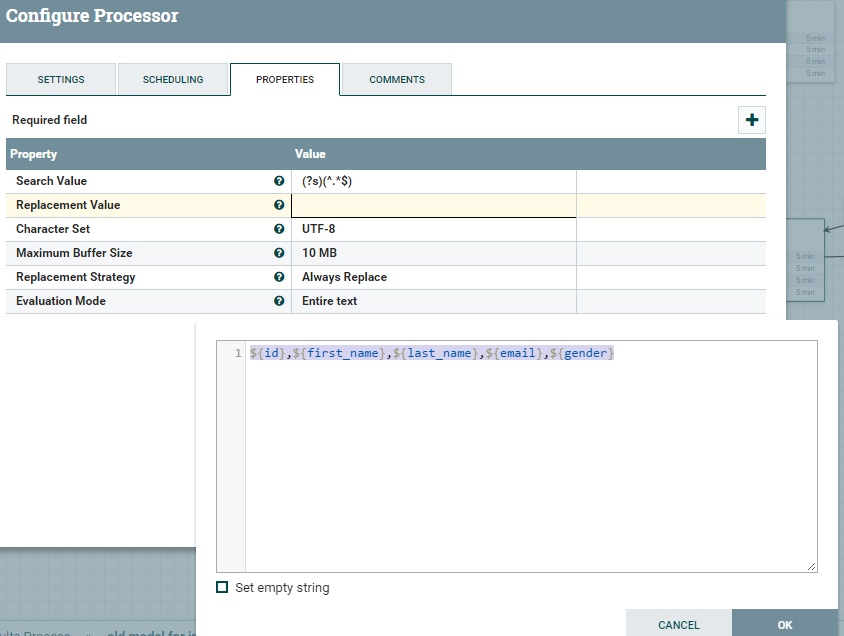


in this screenshot i am extracting all the contents as attributes to the flow file, but you can extract **only the required content values** as attributes to the flow file.

Use **ReplaceText Processor**:-

in this processor keep all the attributes in it to extract only the values of the attributes. we are having all the listed attributes below as flowfile attributes so i am extracting the values from flowfile.

change the **Replacement Strategy** property to **AlwaysReplace.**



**Output from this processor:-**

**Flowfile1 Content:-**

1. 1,Jeanette,Penddreth,jpenddreth0@census.gov,Female

Right now we dont have any json data, if you want to merge the contents of flow files then use

**MergeContent processor :-**

Configure this processor to how many flowfiles you want to merge, increase Minimum Group Size to some number **Default it is 0 B** then this processor will wait to reach Minimum group size and merges all the contents flowfiles into one.

***Note:-***

MergeContent causes memory issues follow instructions in the below link, there are lot of questions on community you can refer to them before configuring this processor

<https://community.hortonworks.com/questions/64337/apache-nifi-merge-content.html>

*i have reached max attachements for this answer,i'm unable to attach the flow screenshot,i'm going to attach flow ss in comments*.

**Hope this Helps...!**

[generate-table-fetch.png](https://community.hortonworks.com/storage/attachments/40544-generate-table-fetch.png) (21.3 kB)

[executesql.png](https://community.hortonworks.com/storage/attachments/40545-executesql.png) (16.0 kB)

[query-database-table.png](https://community.hortonworks.com/storage/attachments/40546-query-database-table.png) (25.8 kB)

[evaluate.png](https://community.hortonworks.com/storage/attachments/40548-evaluate.png) (26.8 kB)

[replacetext-config.png](https://community.hortonworks.com/storage/attachments/40549-replacetext-config.png) (37.7 kB)

[Comment](https://community.hortonworks.com/questions/138406/how-to-extract-text-from-generatetablefetcha-and-i.html)

puja

**People who voted for this**

Close

1 Hide 1 · [Share](https://community.hortonworks.com/questions/138406/how-to-extract-text-from-generatetablefetcha-and-i.html)

Top of Form



[Align text to the leftCenter textAlign text to the rightJustify text](https://community.hortonworks.com/questions/138406/how-to-extract-text-from-generatetablefetcha-and-i.html)

[Insert TableAdd Row AboveAdd Row BelowAdd Column LeftAdd Column RightAdd HeaderDelete HeaderDelete ColumnDelete RowDelete Table](https://community.hortonworks.com/questions/138406/how-to-extract-text-from-generatetablefetcha-and-i.html)

[Normal Text*Quote*Code**Header 1Header 2Header 3Header 4Header 5**](https://community.hortonworks.com/questions/138406/how-to-extract-text-from-generatetablefetcha-and-i.html)

[Insert linkRemove link](https://community.hortonworks.com/questions/138406/how-to-extract-text-from-generatetablefetcha-and-i.html)

​



10 |6000characters neededcharacters leftcharacters exceeded



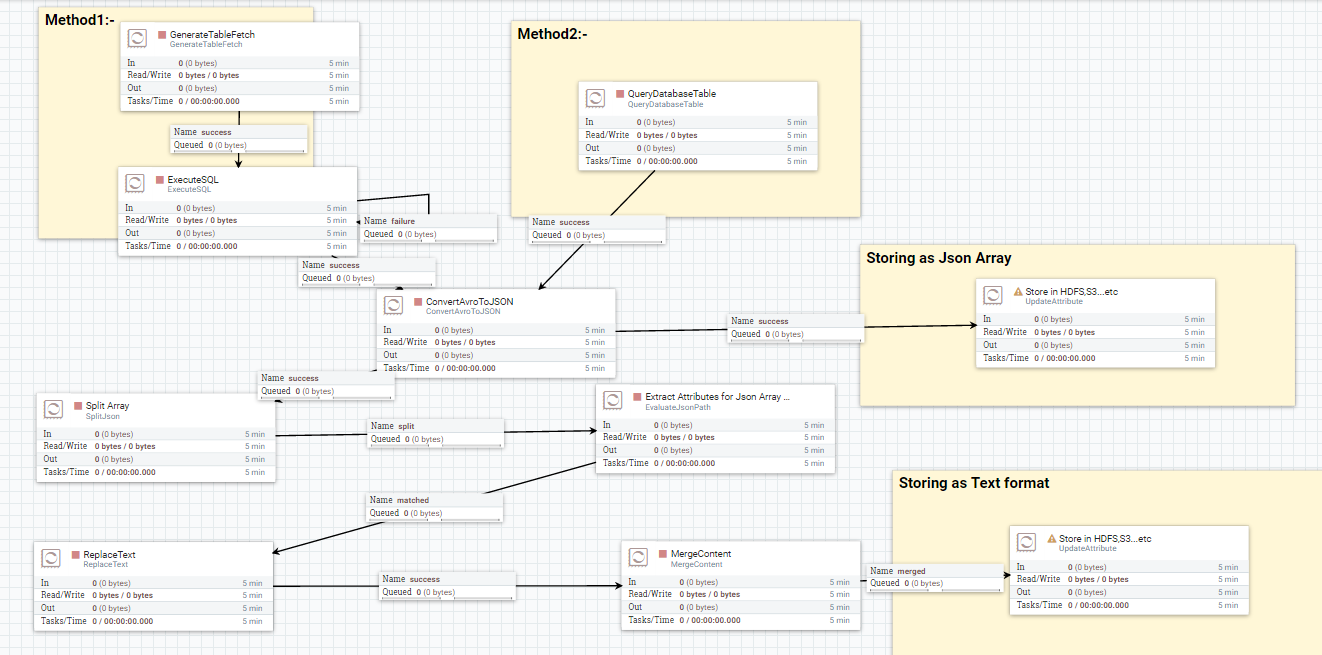
Bottom of Form

avatar image[Shu](https://community.hortonworks.com/users/18929/yaswanthmuppireddy.html) · [Sep 27, 2017 at 02:41 AM](https://community.hortonworks.com/comments/139383/view.html) 1

[Vote](https://community.hortonworks.com/questions/138406/how-to-extract-text-from-generatetablefetcha-and-i.html)[Share](https://community.hortonworks.com/questions/138406/how-to-extract-text-from-generatetablefetcha-and-i.html)

[@Simran Kaur](https://community.hortonworks.com/users/10486/simrank.html)

**FlowScreenshot:-**



This flow demonstrates how to do things with both method1 and method2. in place of update attributes in the screenshot you can use PutHDFS,PutS3..etc.

Keep in mind **don't connect** all the **available relations**(like success,failure,matched,unmatched,etc) to the **next processor**, take the above **screenshot as reference** just **connect the required relations** to **next processor** and for other relations that are **not shown** in this **screenshot** **Auto terminate** them in **processor Settings**.

[flow-table-fetch.png](https://community.hortonworks.com/storage/attachments/40550-flow-table-fetch.png) (140.2 kB)

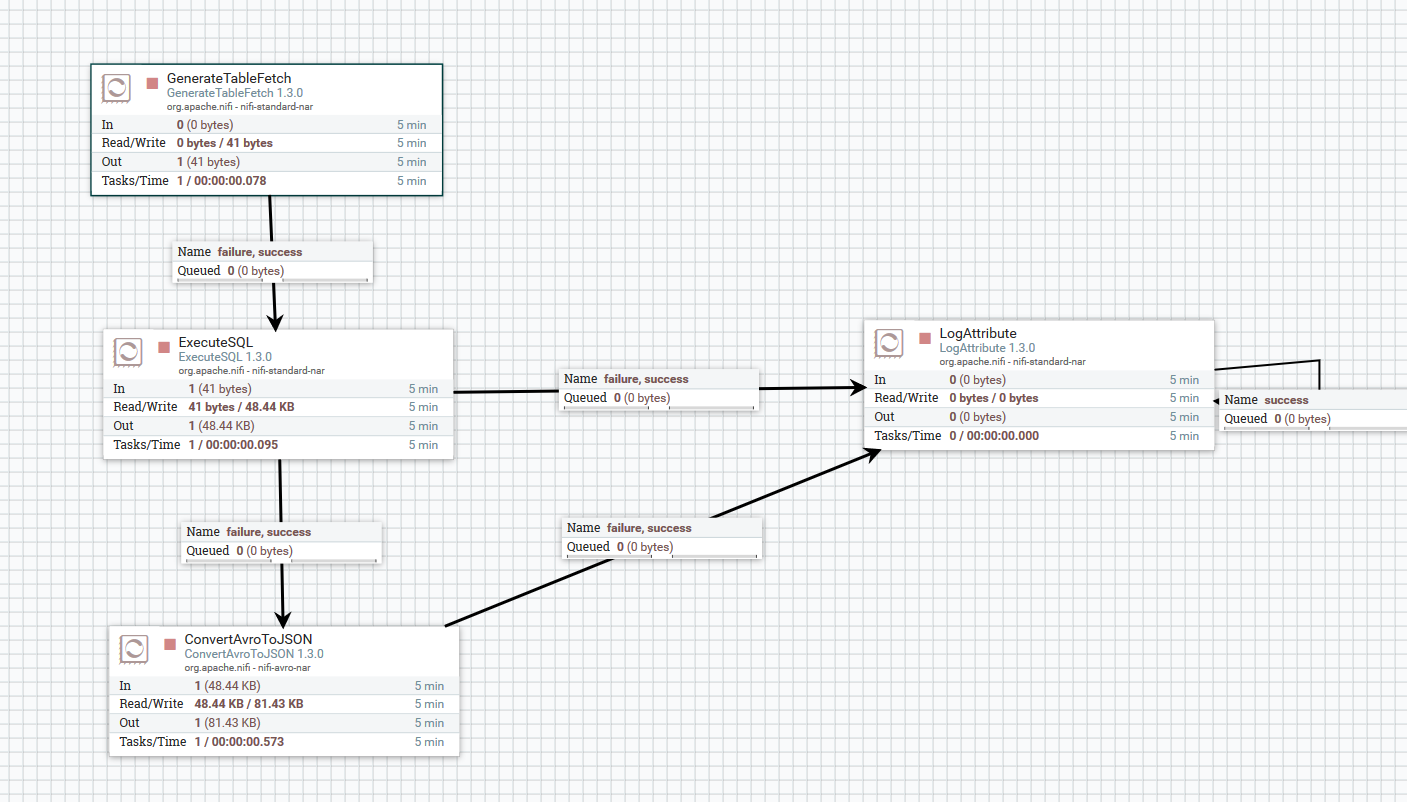
avatar image

**Answer** by [Adrian Oprea](https://community.hortonworks.com/users/12029/opreaadrian1.html) · Sep 26, 2017 at 11:32 PM

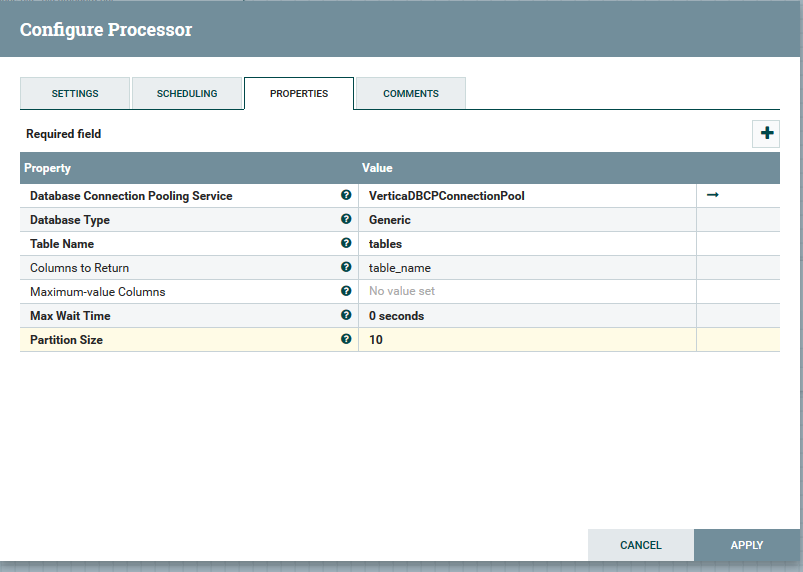
You don`t need to Extract the text .

Your flow should look like this

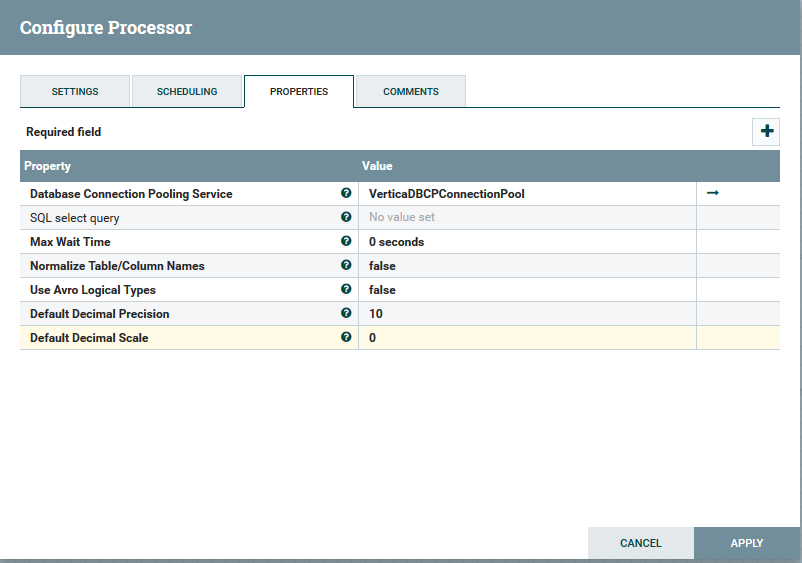
**GenerateTableFetch --> ExecuteSQL -->ConvertAvroToJson(optional) --> StoreResults**



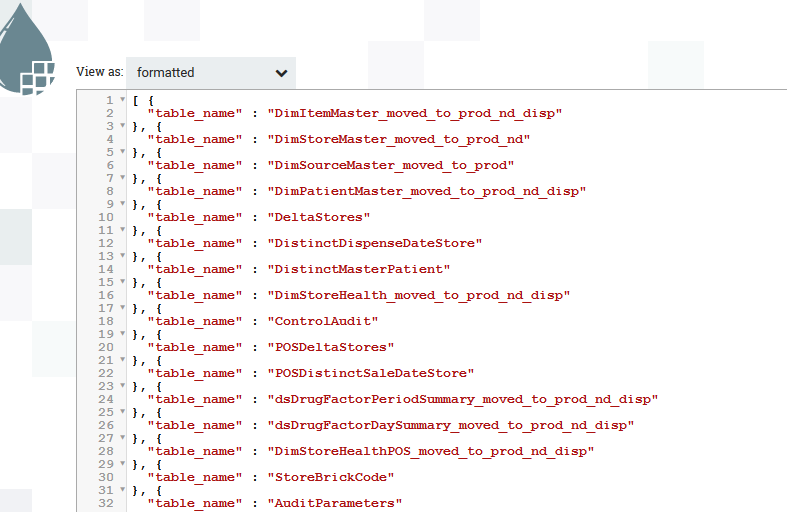
**Where Your GenerateTableFetch conf would be:**



**And the ExecuteSQL is:**



**You would get a Json/Avro Output Format:**



Then you can work on you Json file to extract it

[flow.png](https://community.hortonworks.com/storage/attachments/39572-flow.png) (125.9 kB)

[generatetable.png](https://community.hortonworks.com/storage/attachments/39573-generatetable.png) (31.4 kB)

[executesql.png](https://community.hortonworks.com/storage/attachments/39574-executesql.png) (31.2 kB)

[output.png](https://community.hortonworks.com/storage/attachments/39575-output.png) (34.9 kB)

[Comment](https://community.hortonworks.com/questions/138406/how-to-extract-text-from-generatetablefetcha-and-i.html)