Incremental Data Load in Glue

1. created a s3 bucket where customer subscription csv data is added on daily basis (filename with date)

	A	B	C	D	E	F	G
1	Customer	FirstName	LastName	PlanType	MonthlyC	SubscriptionDate	
2	1201	Angela	Evans	Postpaid	48	12-01-2022	
3	1202	Angelina	Hall	Prepaid	29	15-03-2022	
4	1203	Anita	Collins	Postpaid	56	20-04-2022	
5	1204	Anna	Edwards	Prepaid	27	25-06-2022	
6	1205	Anne	Walker	Postpaid	50	12-08-2022	
7	1206	Anthony	White	Prepaid	33	22-09-2022	
8	1207	Antonio	Turner	Postpaid	49	20-02-2023	
9	1208	April	Phillips	Prepaid	30	18-03-2023	
10	1209	Archer	Perez	Postpaid	56	10-05-2023	
11	1210	Aria	Campbell	Prepaid	28	21-06-2023	

2. ran a crawler over s3 bucket to create metadata table on the top of it.
3. created a visual etl job where I chose glue catalog table(for s3) as source, formed dataframe from it and enabled job bookmark.

```python
from awsglue.utils import getResolvedOptions
from pyspark.context import SparkContext
from awsglue.context import GlueContext
from awsglue.job import Job

## @params: [JOB_NAME]
args = getResolvedOptions(sys.argv, ['JOB_NAME'])

sc = SparkContext()
glueContext = GlueContext(sc)
spark = glueContext.spark_session
job = Job(glueContext)
job.init(args['JOB_NAME'], args)

subsDf = glueContext.create_dynamic_frame.from_catalog(
    database="telecom-datalog",
    table_name="gds_telecom_data_glue",
    transformation_ctx="s3_input_new"
    )

subsDf.printSchema()
sparkSubsDf = subsDf.toDF()
sparkSubsDf.show()
print(sparkSubsDf.count())

job.commit()
```

4. only new or updated data will be added to the dataframe in incremental manner