WE CAN CREATE NUMBER OF PARTITIONS WE WANT
BY LOADING THE DATA INTO THE SAME FILE IT CREATES
NEW FILE IN ITS SUB DIRECTORY WHICH IS POINTING TO HDFS
****

we need to create a new AVSC file by copying same properties of file in local system
cp orders.avsc sandeep.avsc
---
view sandeep.avsc
Now we need to modifie this avsc file to improve query performance
1.we need to add a partition column

```
{
  "type" : "record",
  "name" : "CHANGE NAME",
  "doc" : "HIVE AVRO PARTITINED TABLE CHANGE NAME",
  "fields" : [ {
    "name" : "order_id",
    "type" : [ "null", "int" ],  REMOVE COMMA HERE
    "default" : null,
    "columnName" : "order_id",                    DELETE THAT IN STARS
    "sqlType" : "4"  ReMoVe UpTo HeRe       FROM COMA TO SQL TYPE
  }, {
    "name" : "order_date",
    "type" : [ "null", "long" ],
    "default" : null,
    "columnName" : "order_date",
    "sqlType" : "93"
  }, {
    "name" : "order_customer_id",
    "type" : [ "null", "int" ],
    "default" : null,
    "columnName" : "order_customer_id",
    "sqlType" : "4"
  }, {
    "name" : "order_status",
    "type" : [ "null", "string" ],
    "default" : null,
```

```
   "columnName" : "order_status",
   "sqlType" : "12"
  } ],
  "tableName" : "CHANGE NAME"
}
```

==========

Now before creating the table make sure that the table which you are going to take refrence
is existing in the same DATABASE(HIVE) which you are going to create TABLE as you can see in our DATABASE(HIVE) we have orders and thats why we are taking refrence of orders
while we are LOADING the DATA

==========

Now create **table** name should be same as mentioned in .AVSC file

==========

```
CREATE TABLE record_sandeep (
order_id int,
order_date bigint,
order_customer_id int,
order_status string
)
PARTITIONED BY (order_month string)
STORED AS AVRO
LOCATION 'hdfs:///user/hive/warehouse/retail_stage.db/orders_part_avro'
TBLPROPERTIES
('avro.schema.url'='hdfs://quickstart.cloudera/user/cloudera/avscData/sandeep.avsc'
);
```

==========

ADDING MANUVAL PARTITION

```
alter table record_sandeep add partition (order_month='2014-01');
```

==========

INSERTING DATA INTO A PARTITION

---

insert into table record_sandeep partition (order_month='2014-01')
select * from orders where from_unixtime(cast(substr(order_date, 1, 10) as int))
like '2014-01%';

+++++

here orders already exsits in our current database we are taking refrence to our data

==========

validate

==========

dfs -ls /user/hive/warehouse/retail_stage.db/orders_part_avro/*
dfs -ls /user/hive/warehouse/retail_stage.db/orders_part_avro/