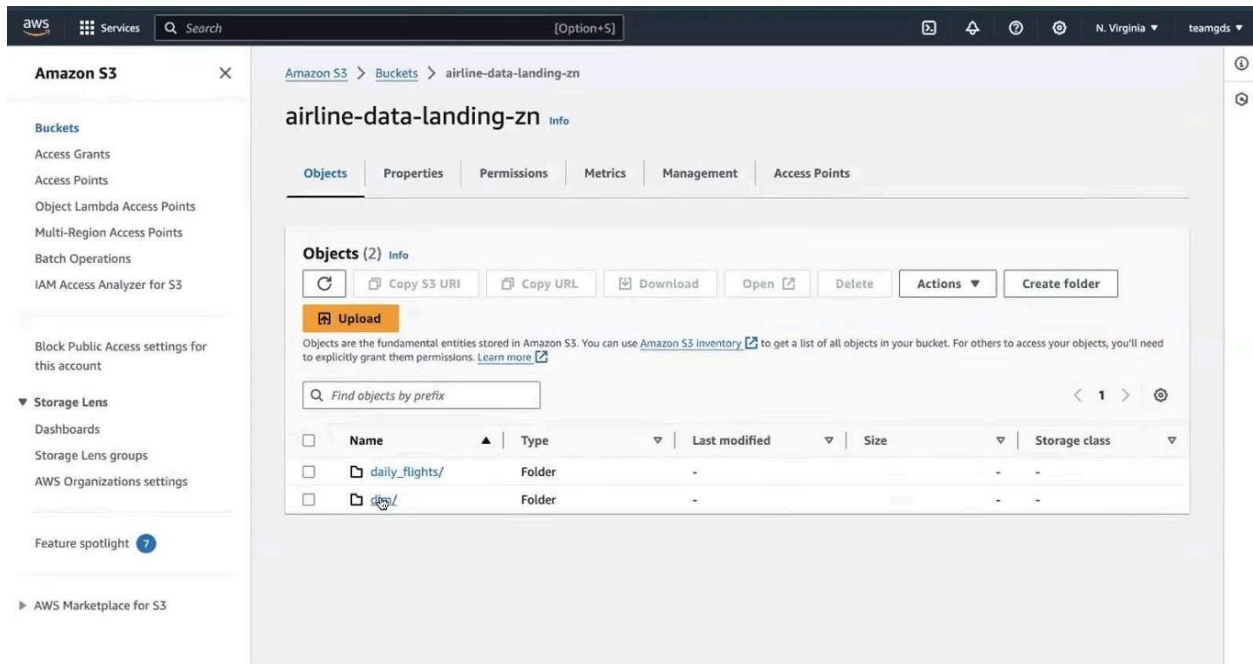


# Aviation Data Flow Management:

## Amazon S3:



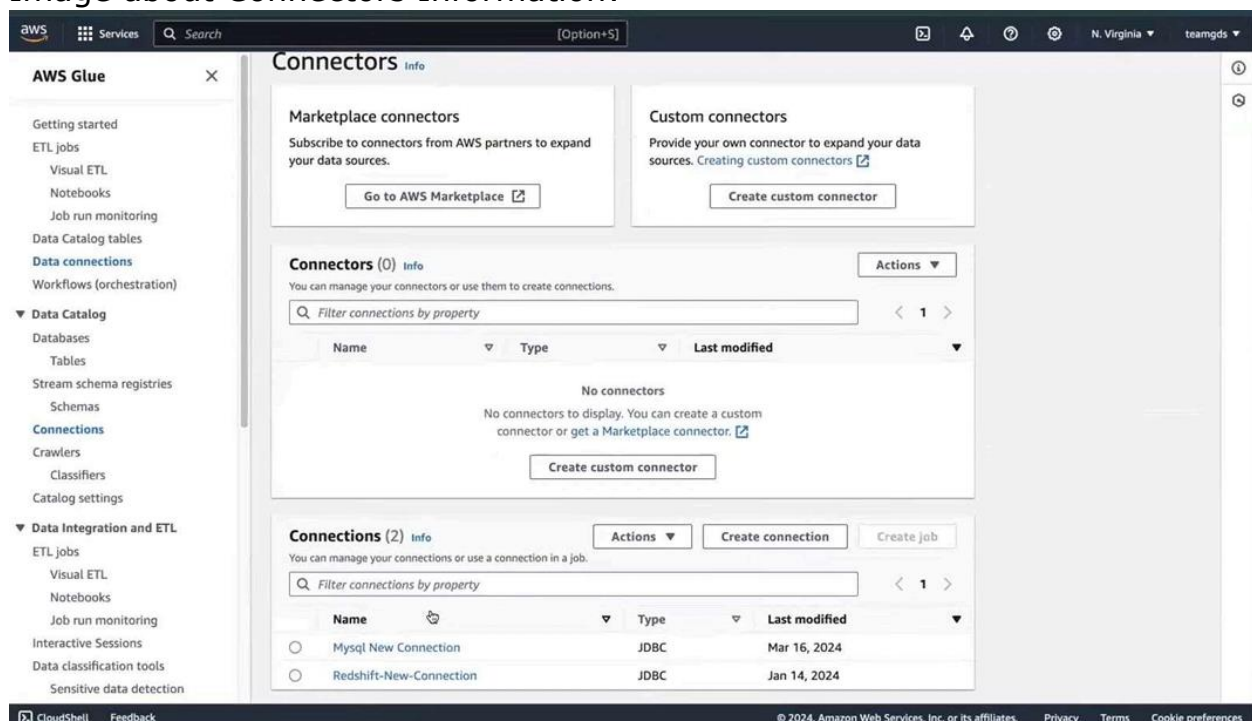
How does the airports data look like?

```
airports.csv X
Airline_Data_Ingestion_Project > airports.csv
1 airport_id,city,state,name
2 10165,Adak Island,AK,Adak
3 10299,Anchorage,AK,Ted Stevens Anchorage International
4 10304,Aniak,AK,Aniak Airport
5 10754,Barrow,AK,Wiley Post/Will Rogers Memorial
6 10551,Bethel,AK,Bethel Airport
7 10926,Cordova,AK,Merle K Mudhole Smith
8 14709,Deadhorse,AK,Deadhorse Airport
9 11336,Dillingham,AK,Dillingham Airport
10 11630,Fairbanks,AK,Fairbanks International
11 11997,Gustavus,AK,Gustavus Airport
12 12523,Juneau,AK,Juneau International
13 12819,Ketchikan,AK,Ketchikan International
14 10245,King Salmon,AK,King Salmon Airport
15 10170,Kodiak,AK,Kodiak Airport
16 13970,Kotzebue,AK,Ralph Wien Memorial
17 13873,Nome,AK,Nome Airport
18 14256,Petersburg,AK,Petersburg James A Johnson
19 14828,Sitka,AK,Sitka Rocky Gutierrez
20 12807,St. Mary's,AK,St. Mary's Airport
21 11445,Unalaska,AK,Unalaska Airport
22 15841,Wrangell,AK,Wrangell Airport
23 15991,Yakutat,AK,Yakutat Airport
24 10599,Birmingham,AL,Birmingham-Shuttlesworth International
25 11308,Dothan,AL,Dothan Regional
26 12217,Huntsville,AL,Huntsville International-Carl T Jones Field
27 13422,Mobile,AL,Mobile Regional
28 13277,Montgomery,AL,Montgomery Regional
29 15919,Fayetteville,AR,Northwest Arkansas Regional
30 11778,Fort Smith,AR,Fort Smith Regional
31 12992,Little Rock,AR,Adams Field
32 15401,Texarkana,AR,Texarkana Regional-Webb Field
33 11695,Flagstaff,AZ,Flagstaff Pulliam
34 11869,Grand Canyon,AZ,Grand Canyon National Park
35 14107,Phoenix,AZ,Phoenix Sky Harbor International
36 15376,Tucson,AZ,Tucson International
37 16218,Yuma,AZ,Yuma MCAS/Yuma International
```

How does the daily flight information data look like?

```
flights.csv x
Airline_Data_Ingestion_Project > flights.csv
1 Carrier,OriginAirportID,DestAirportID,DepDelay,ArrDelay
2 DL,11433,13303,-3,1
3 DL,14869,12478,0,-8
4 DL,14057,14869,-4,-15
5 DL,15016,11433,28,24
6 DL,11193,12892,-6,-11
7 DL,10397,15016,-1,-19
8 DL,15016,10397,0,-1
9 DL,10397,14869,15,24
10 DL,10397,10423,33,34
11 DL,11278,10397,323,322
12 DL,14107,13487,-7,-13
13 DL,11433,11298,22,41
14 DL,11298,11433,40,20
15 DL,11433,12892,-2,-7
16 DL,10397,12451,71,75
17 DL,12451,10397,75,57
18 DL,12953,10397,-1,10
19 DL,11433,12953,-3,-10
20 DL,10397,14771,31,38
21 DL,13204,10397,8,25
22 DL,13487,12889,-3,-2
23 DL,12889,13487,7,16
24 DL,10397,14027,0,19
25 DL,14027,10397,13,25
26 DL,13244,10397,4,13
27 DL,12953,11697,12,-8
28 DL,10397,12892,15,7
29 DL,14869,14747,-7,-6
30 DL,14747,14869,15,18
31 DL,10397,10423,0,56
32 DL,10423,10397,56,49
33 DL,11292,12953,-5,6
34 DL,14869,14893,-5,-14
35 DL,12953,11697,-3,21
36 DL,14869,14908,-6,-14
37 DL,12892,15304,21,-2
```

Image about Connectors Information:



## Glue Crawler Information:

**Step 1: Set crawler properties**

Set crawler properties

Name	Description	Tags
temp-crawler	-	-

**Step 2: Choose data sources and classifiers**

Data sources (1) Info

The list of data sources to be scanned by the crawler.

Type	Data source	Parameters
JDBC	dev/airlines/airports_dim	-

**Step 3: Configure security settings**

Configure security settings

IAM role	Security configuration	Lake Formation configuration
glue-role	-	-

**Step 4: Set output and scheduling**

Set output and scheduling

Database	Table prefix - optional	Schedule
airline_datamart	-	On demand

Cancel Previous Create crawler

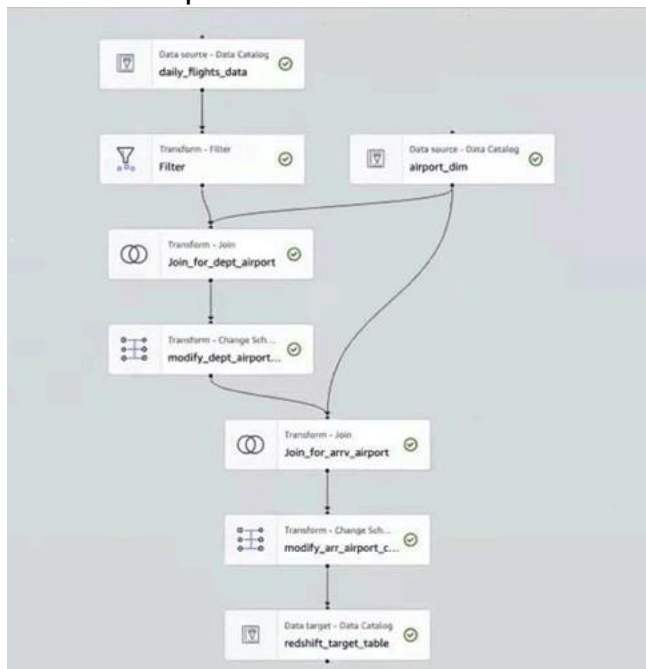
List of Columns dropped after performing 1<sup>st</sup> join operation:

Change Schema (Apply mapping)			
Source key	Target key	Data type	Drop
date			<input checked="" type="checkbox"/>
originairportid			<input checked="" type="checkbox"/>
depdelay	dep_delay	bigint	<input type="checkbox"/>
arrdelay	arr_delay	bigint	<input type="checkbox"/>
destairportid	destairport	long	<input type="checkbox"/>
carrier	carrier	string	<input type="checkbox"/>
airport_id			<input checked="" type="checkbox"/>
city	dep_city	string	<input type="checkbox"/>
name	dep_airpor	string	<input type="checkbox"/>
state	dep_state	string	<input type="checkbox"/>

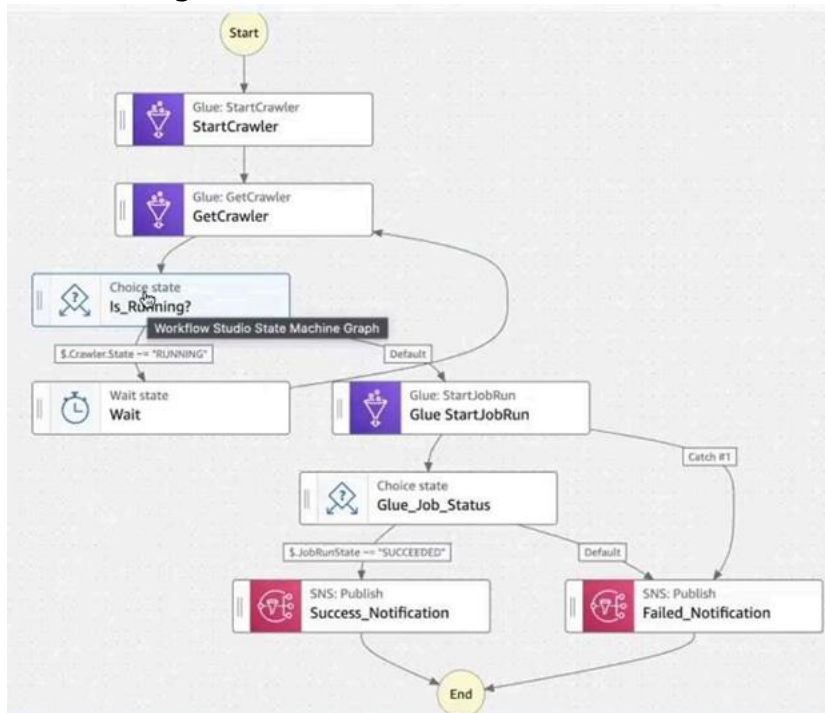
List of Columns dropped after performing 2<sup>nd</sup> join operation:

Change Schema (Apply mapping)			
Source key	Target key	Data type	Drop
dep_delay	dep_delay	bigint	<input type="checkbox"/>
arr_delay	arr_delay	bigint	<input type="checkbox"/>
destairportid			<input checked="" type="checkbox"/>
carrier	carrier	string	<input type="checkbox"/>
dep_city	dep_city	string	<input type="checkbox"/>
dep_airport	dep_airpor	string	<input type="checkbox"/>
dep_state	dep_state	string	<input type="checkbox"/>
airport_id			<input checked="" type="checkbox"/>
city	arr_city	string	<input type="checkbox"/>
name	arr_airport	string	<input type="checkbox"/>
state	arr_state	string	<input type="checkbox"/>

Glue Job Pipeline:



## Event Bridge Patterns:



## Step Function Pipeline:

### Event pattern [Info](#)

#### Event pattern

Write an event pattern in JSON. You can test the event pattern against the sample event. You can also go to pre-defined pattern.

Prefix matching ▼

Insert

☒ Content-based filter syntax

```
1 {
2   "source": ["aws.s3"],
3   "detail-type": ["Object Created"],
4   "detail": {
5     "bucket": {
6       "name": ["airline-data-landing-zn"]
7     },
8     "object": {
9       "key": [{
10        "suffix": ".csv"
11      }]
12    }
13  }
14 }
```

✓ JSON is valid

Copy

{ } Prettify

Event pattern form

Test pattern

Sample State Machine Image while execution:

