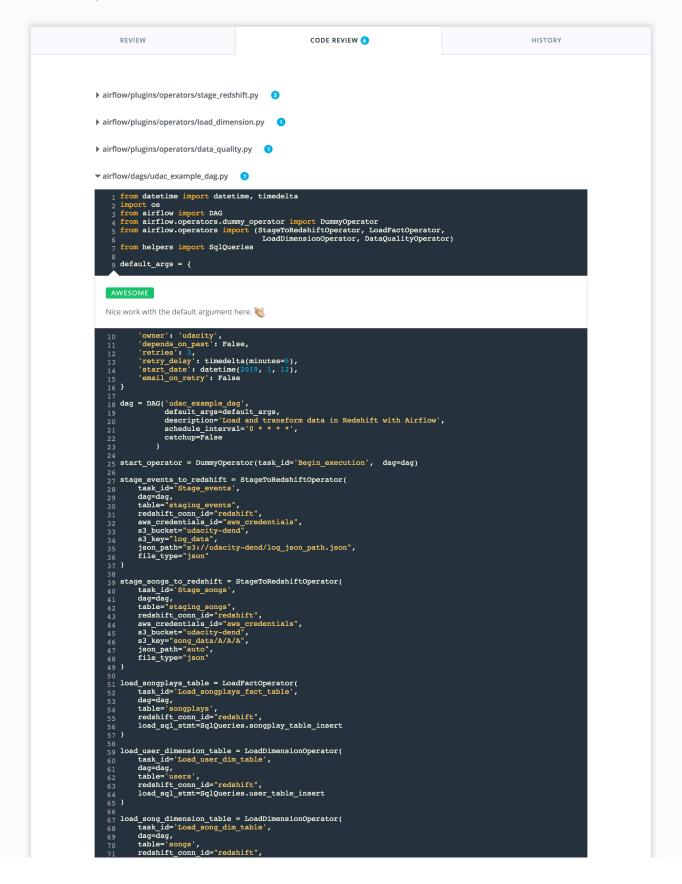


Return to "Data Engineering Nanodegree" in the classroom

☐ DISCUSS ON STUDENT HUB

## Data Pipelines with Airflow



```
load_sql_stmt=SqlQueries.song_table_insert
     73 )
74
75 load_artist_dimension_table = LoadDimensionOperator(
76 task_id='Load_artist_dim_table',
77 dag=dag,
78 table='artists',
79 redshift_conn_id="redshift",
80 load_sql_stmt=SqlQueries.artist_table_insert
81 )
82
83 load_time_dimension_table = LoadDimensionOperator(
84 task_id='Load_time_dim_table',
85 dag=dag,
86 table='time',
                    task_id= load_time_dim_table ,
dag=dag,
table='time',
redshift_conn_id="redshift",
load_sql_stmt=SqlQueries.time_table_insert
     90
91 run_quality_checks = DataQualityOperator(
92 task_id='Run_data_quality_checks',
93 dag=dag,
94 tables=['songplays', 'users', 'songs', 'artists', 'time'],
95 redshift_conn_id="redshift"
     97
98 end_operator = DummyOperator(task_id='Stop_execution', dag=dag)
99
99
   100 #
101 # Task ordering for the defined DAG tasks
102 #
   103
104 start_operator >> stage_events_to_redshift
105 start_operator >> stage_songs_to_redshift
   107 stage_events_to_redshift >> load_songplays_table 108 stage_songs_to_redshift >> load_songplays_table
   109
110 load_songplays_table >> load_user_dimension_table
111 load_songplays_table >> load_song_dimension_table
112 load_songplays_table >> load_artist_dimension_table
113 load_songplays_table >> load_time_dimension_table
    115 load_user_dimension_table >> run_quality_checks
   116 load song dimension table >> run quality checks
117 load_artist_dimension_table >> run_quality_checks
118 load_time_dimension_table >> run_quality_checks
     119
120 run_quality_checks >> end_operator
▶ airflow/plugins/operators/load_fact.py
airflow/plugins/operators/_init_.py
airflow/plugins/helpers/sql_queries.py
▶ airflow/plugins/helpers/_init_.py
▶ airflow/plugins/_init_.py
▶ airflow/create_tables.sql
```

Rate this review