**Bash operator - ETL project with Airflow**

**Scenario:** a project to de-congest the national highways by analyzing the road traffic data from different toll plazas. Each highway is operated by a different toll operator with a different IT setup that uses different file formats. Here collected data available in different formats and consolidate it into a single file.

developed an Apache Airflow DAG that will:

* Extract data from a csv file
* Extract data from a tsv file
* Extract data from a fixed-width file
* Transform the data
* Load the transformed data into the staging area
* After python script is done, save it.
* Submit a DAG
  + Run below commands
  + Export AIRFLOW\_HOME=/home/project/airflow
  + Cp ETL\_toll\_data.py $AIRFLOW\_HOME/dags
* Verify if DAG is submitted or not
  + Airflow dags list|grep “ETL”
* If can not find DAG then run below command to check errors, resolve error delete python file from dag folder and again copy using above cp command.
  + Airflow dags list-import-errors