There are 3 big data files provided to you. File names are as follows

* Demographic\_Statistics\_By\_Zip\_Code
* Grocery\_Stores
* Popular\_Baby\_Names

Create ingest, snapshot, latest and analysis databases

* In the ingest maintain the raw data
* In the snapshot maintain the partitioned data
* In latest create a view pointing to the latest partition in the snapshot db
* Make sure to store the result in the analysis DB in the form of view

Naming conventions for the DB’s are as follows

* Username\_ingest
* Username\_snapshot
* Username\_latest
* Username\_analysis

Ex : ajay\_ingest, ajay\_snapshot

In Analysis DB the views created should have the names as per the question numbers

Ex: question 1 view name will be **v1** , question 2 view name will be **v2…etc**

Questions :

* **Demographic\_Statistics\_By\_Zip\_Code**

1. For all the Jurisdiction name whose COUNT GENDER TOTAL is greater than 40, find the male count, which are greater than the female count for the same jurisdiction name .
2. Select the Jurisdiction name whose COUNT GENDER TOTAL is 0 and 1. For selected rows , replace 0 with 1 and 1 with 0
3. For all the Jurisdiction name between 11111 and 11367, which are having 100 PERCENT PUBLIC ASSISTANCE TOTAL, get the total count of both female and male , only for those records having percentage of female between 30% and 60%

* **Popular\_Baby\_Names**

1. For Child name which are repeated more than once,get the ethnicity if it is a male child, child first name if female kid
2. find the number of occurences of each name in the table and if the number of occurences are between 2 and 5 , blank the name
3. for ASIAN AND PACIFIC ISLANDER ethnicity whatever male baby names which starts with ‘A’

extract only those with the lowest rank

* **Grocery\_Stores**

1. select the zip codes which are going to have more than one supermarkets
2. for supermarkets which are not having more than one branches find the supermakets in the next zip code
3. find the zipcodes for the super markets which starts with ‘G’ and has only full supermarket
4. Create a shell script to execute all te above queries one by one and store the result in a directory each