# Apply ETL pipeline and data engineering on uber data

-By Rahul B

## Steps to perform data engineering:

1)Get the uber dataset

2)Create ipynb file and import the uber dataset

3)Create ER diagram in draw.io website (Include ER-Entity relationship model which contains relationship between columns)

4)Make an analysis of uber data and assign name for values and save it as .ipynb file

5)Create Google Cloud Platform account and login with bank details and we will receive free cloud storage access for 300 dollars

5)Create instance and import the uber dataset into the instance and change the access to public and by the way the URL will be generated()By using this we can download the dataset from website

6)Click compute engine, create new instance in it

7)Setup for mage installation

Open SSH in instance column and enter the following commands

i)sudo apt-get install python3-distutils

ii)sudo apt-get install python3-apt

iii)sudo apt-get install wget

iv)wget https://bootstrap.pypa.io/get-pip.py

v)sudo python3 get-pip.py

8)Create new instance magma and Copy the external IP address(35.244.17.11) get port number 6789-tcp port, edit firewall rule and make changes such as IPV4 range change it to 0.0.0.0/0 then save and type URL as : 35.244.17.11/6789/pipeline

9)Enter this website: 35.244.17.11/6789/pipeline

10)Click load data and select python and click API then write as load_uber_data then create uber_transformation and finally uber_bq_load make use of api from cloud service into code of these file(load_uber_data->uber_transformation->uber_bq_load)

11)Create key and download it as JSON and past JSON file values in io_config.yaml in mage website

12)Install some google cloud packages by entering command:

i)sudo pip3 install google-cloud

ii)sudo pip3 install google-cloud-bigquery

12)Write sql query by including column relation as we designed er model

13)Open lookerstudio.google.com/u/0/navigation/reporting website and create the dashboard and click big query by google option and select big query file which you created in google cloud platform then create the dashboard by making visualizations.

THANK YOU!!!

Support us on: https://www.youtube.com/@mrherotechspot3831

https://github.com/rahul13289