ASSIGNMENT -4

EXTRACT , TRANSFORM , LOAD

Hitesh Kumar - PES1201701511

Nishchal M N -PES1201701523

Suraj M M - PES1201701551

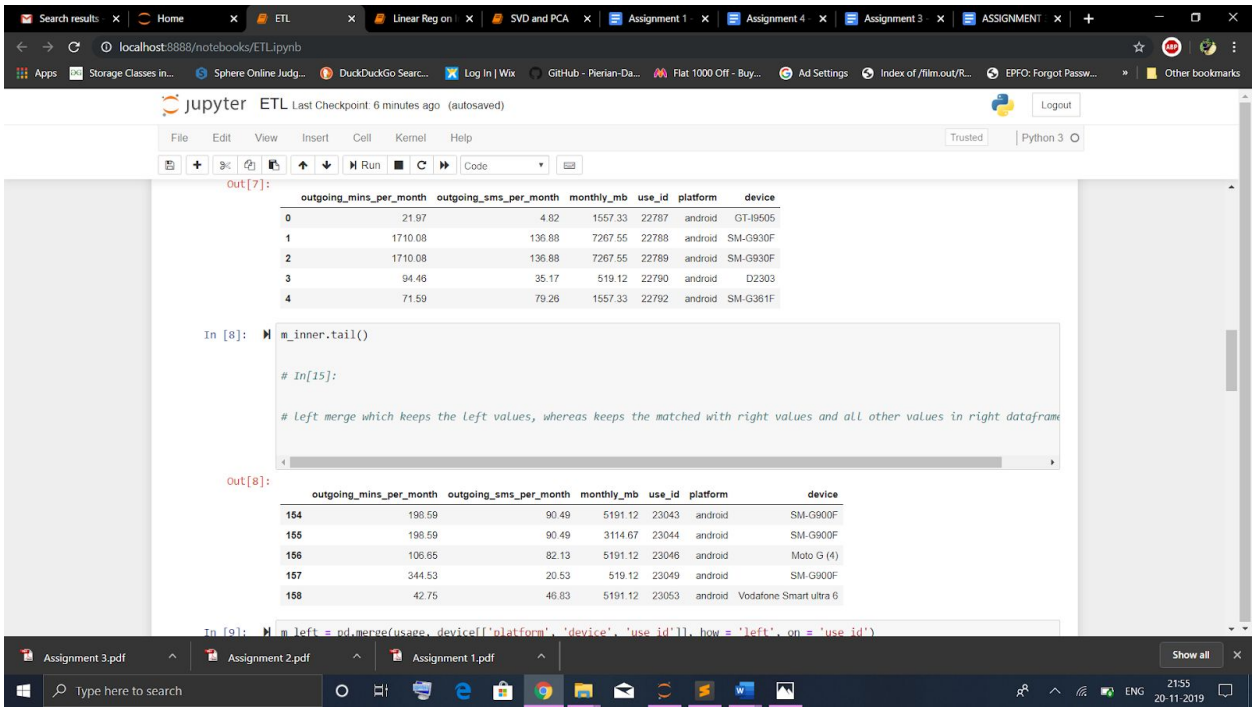
Vidhan Mishra - PES1201700055

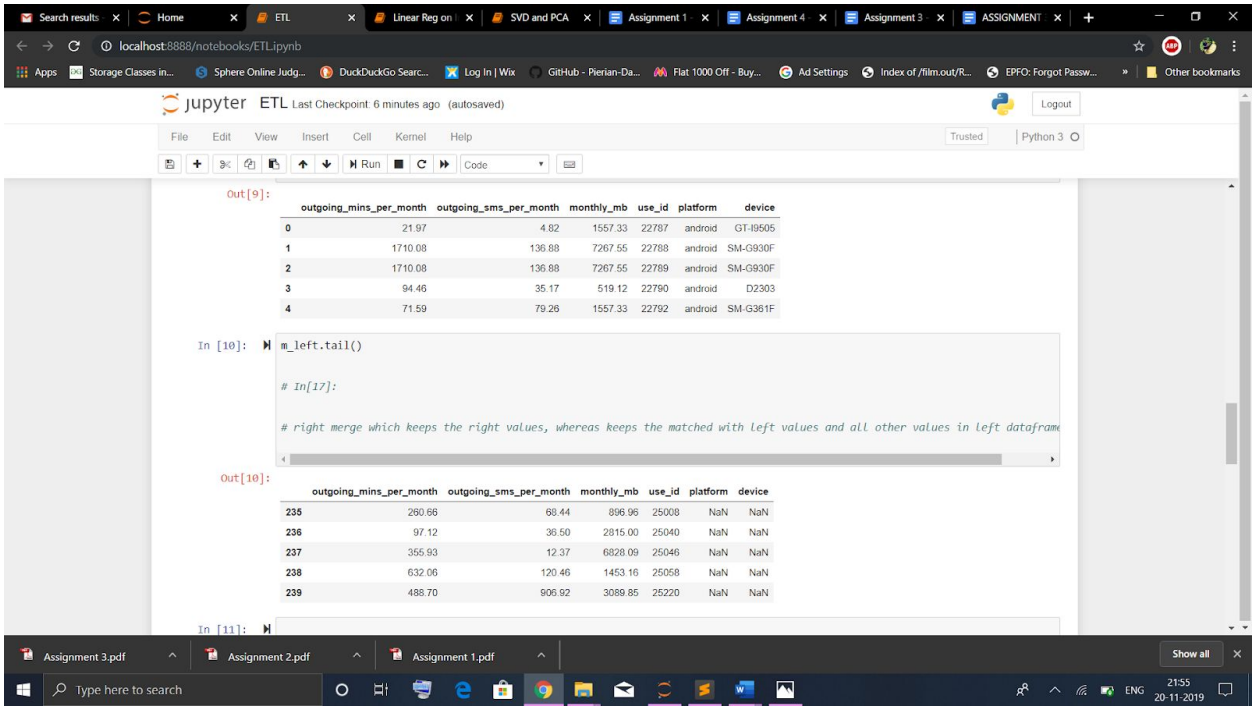
ETL is short for extract, transform, load, three database functions that are combined into one tool to pull data out of one database and place it into another database.

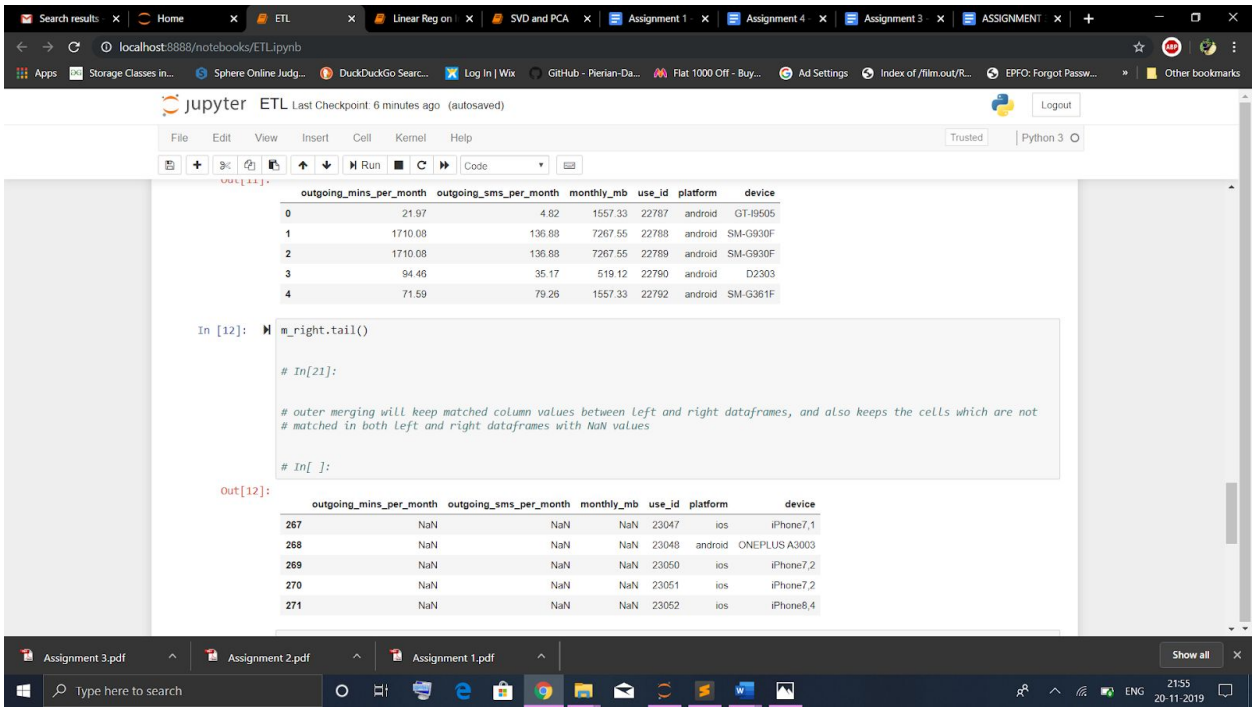
Extract is the process of reading data from a database. In this stage, the data is collected, often from multiple and different types of sources.

Transform is the process of converting the extracted data from its previous form into the form it needs to be in so that it can be placed into another database. Transformation occurs by using rules or lookup tables or by combining the data with other data.

Load is the process of writing the data into the target database.

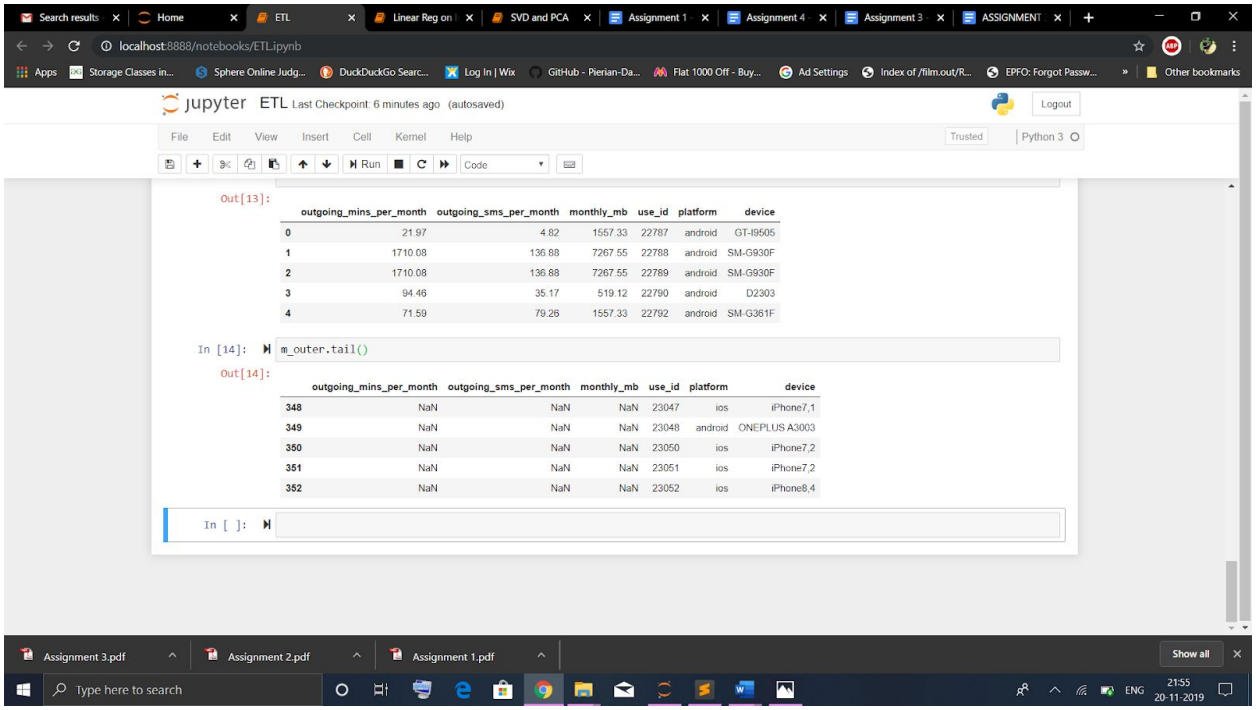

Search results
Home
ETL
Linear Reg on
SVD and PCA
Assignment 1
Assignment 4
Assignment 3
ASSIGNMENT
localhost:8888/notebooks/ETL.ipynb
jupyter ETL Last Checkpoint: 6 minutes ago (autosaved)
Logout
File Edit View Insert Cell Kernel Help
Trusted
Python 3
Run
Code
Out[7]:
outgoing_mins_per_month outgoing_sms_per_month monthly_mb use_id platform device
0 21.97 4.82 1557.33 22787 android GT-I9505
1 1710.08 136.88 7267.55 22788 android SM-G930F
2 1710.08 136.88 7267.55 22789 android SM-G930F
3 94.46 35.17 519.12 22790 android D2303
4 71.59 79.26 1557.33 22792 android SM-G361F
In [8]:
m_inner.tail()
# In[15]:
# left merge which keeps the left values, whereas keeps the matched with right values and all other values in right datafram
Out[8]:
outgoing_mins_per_month outgoing_sms_per_month monthly_mb use_id platform device
154 198.59 90.49 5191.12 23043 android SM-G900F
155 198.59 90.49 3114.67 23044 android SM-G900F
156 106.65 82.13 5191.12 23046 android Moto G (4)
157 344.53 20.53 519.12 23049 android SM-G900F
158 42.75 46.83 5191.12 23053 android Vodafone Smart ultra 6
In [9]:
m_left = pd.merge(usage, device['platform', 'device', 'use_id'], how = 'left', on = 'use_id')
Assignment 3.pdf
Assignment 2.pdf
Assignment 1.pdf
Show all
Type here to search
ENG
21:55
20-11-2019


Search results
Home
ETL
Linear Reg on
SVD and PCA
Assignment 1
Assignment 4
Assignment 3
ASSIGNMENT
localhost:8888/notebooks/ETL.ipynb
jupyter ETL Last Checkpoint: 6 minutes ago (autosaved)
Logout
File Edit View Insert Cell Kernel Help
Trusted
Python 3
Run
Code
Out[9]:
outgoing_mins_per_month outgoing_sms_per_month monthly_mb use_id platform device
0 21.97 4.82 1557.33 22787 android GT-I9505
1 1710.08 136.88 7267.55 22788 android SM-G930F
2 1710.08 136.88 7267.55 22789 android SM-G930F
3 94.46 35.17 519.12 22790 android D2303
4 71.59 79.26 1557.33 22792 android SM-G361F
In [10]:
m_left.tail()
# In[17]:
# right merge which keeps the right values, whereas keeps the matched with left values and all other values in left datafram
Out[10]:
outgoing_mins_per_month outgoing_sms_per_month monthly_mb use_id platform device
235 260.66 68.44 896.96 25008 NaN NaN
236 97.12 36.50 2815.00 25040 NaN NaN
237 355.93 12.37 6828.09 25046 NaN NaN
238 632.06 120.46 1453.16 25058 NaN NaN
239 488.70 906.92 3089.85 25220 NaN NaN
In [11]:
Assignment 3.pdf
Assignment 2.pdf
Assignment 1.pdf
Show all
Type here to search
ENG
21:55
20-11-2019



	outgoing_mins_per_month	outgoing_sms_per_month	monthly_mb	use_id	platform	device
0	21.97	4.82	1557.33	22787	android	GT-I9505
1	1710.08	136.88	7267.55	22788	android	SM-G930F
2	1710.08	136.88	7267.55	22789	android	SM-G930F
3	94.46	35.17	519.12	22790	android	D2303
4	71.59	79.26	1557.33	22792	android	SM-G361F

	outgoing_mins_per_month	outgoing_sms_per_month	monthly_mb	use_id	platform	device
267	NaN	NaN	NaN	23047	ios	iPhone7,1
268	NaN	NaN	NaN	23048	android	ONEPLUS A3003
269	NaN	NaN	NaN	23050	ios	iPhone7,2
270	NaN	NaN	NaN	23051	ios	iPhone7,2
271	NaN	NaN	NaN	23052	ios	iPhone8,4



	outgoing_mins_per_month	outgoing_sms_per_month	monthly_mb	use_id	platform	device
0	21.97	4.82	1557.33	22787	android	GT-I9505
1	1710.08	136.88	7267.55	22788	android	SM-G930F
2	1710.08	136.88	7267.55	22789	android	SM-G930F
3	94.46	35.17	519.12	22790	android	D2303
4	71.59	79.26	1557.33	22792	android	SM-G361F

	outgoing_mins_per_month	outgoing_sms_per_month	monthly_mb	use_id	platform	device
348	NaN	NaN	NaN	23047	ios	iPhone7,1
349	NaN	NaN	NaN	23048	android	ONEPLUS A3003
350	NaN	NaN	NaN	23050	ios	iPhone7,2
351	NaN	NaN	NaN	23051	ios	iPhone7,2
352	NaN	NaN	NaN	23052	ios	iPhone8,4