

**Managing & Programming Database MCDA5540**

**Master of Science in Computing and Data Analytics**

**Assignment-3**

**Submitted by**:

Caner Adil Irfanoglu (A00425840)

Sunil Padikar (A00428089)

Vinay Govindan (A00429120)

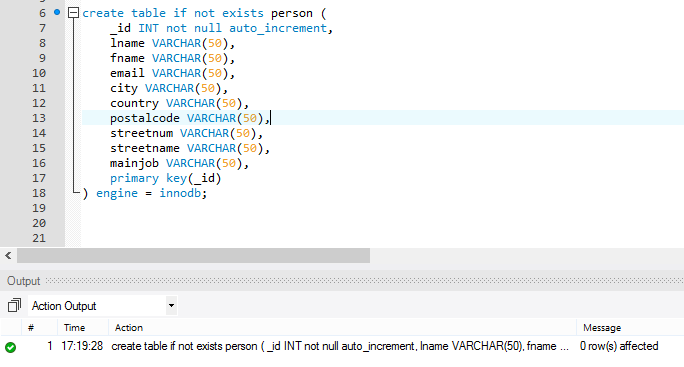
Diven Sambhwani (A00429660)

**Submitted to**:

TRISHLA SHAH

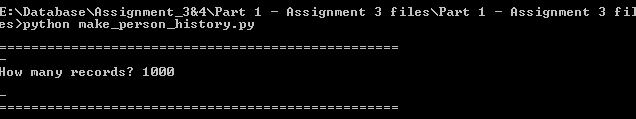
**Problem**

1. ***The MySQL table "person" is created in the script "make\_person\_table.sql" using 20,000 person lines from the file "persons.tsv".***

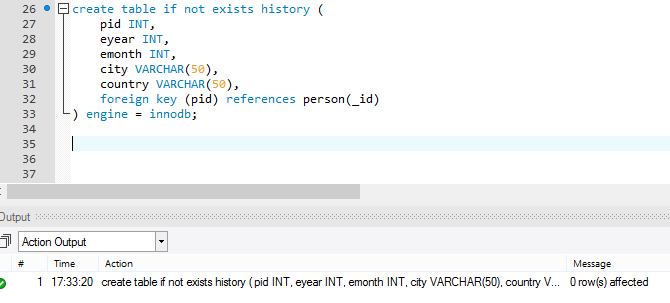




1. ***The python script "make\_person\_history.py" can be used to generate any number of events about 20,000 persons (technically about persons with id's between 1 and 20,000). If N is the number of events given as input, the script generates N event lines and saves them in the file "phistory.tsv". Note that the data files "cities.txt" and "countries.txt" are used by the script to generate city and country data.***



1. ***The MySQL table "history" is created in the script "make\_hist\_table.sql" using the event lines from the file "phistory.tsv".***





1. ***4. In particular, the test was to make a search in the join of the tables person and history on person id using history.city as a search attribute. The search was first done without an index on the search attribute history(city), and then after creating an index on history(city).***





***Effect of Indexing on Searching and Insertion***

|  |  |  |
| --- | --- | --- |
| **INSERTION** | | |
| **N** | **without index(sec)** | **with index(sec)** |
| 1,000 | 0.1 | 0.08 |
| 5,000 | 0.16 | 0.29 |
| 15,000 | 0.5 | 0.7 |
| 25,000 | 0.68 | 0.99 |
| 50,000 | 1.38 | 2.13 |
| 75,000 | 2.5 | 2.9 |
| 100,000 | 2.75 | 3.2 |
| 500,000 | 12.33 | 16.03 |
| 800,000 | 17.9 | 23.3 |
| 1,000,000 | 20.7 | 35.7 |

|  |  |  |
| --- | --- | --- |
| **SEARCHING** | | |
| **N** | **without index(sec)** | **with index(sec)** |
| 1,000 | 0.001 | 0 |
| 5,000 | 0.004 | 0 |
| 15,000 | 0.02 | 0 |
| 25,000 | 0.03 | 0 |
| 50,000 | 0.039 | 0 |
| 75,000 | 0.045 | 0 |
| 1,00,000 | 0.057 | 0 |
| 500,000 | 0.32 | 0 |
| 800,000 | 0.57 | 0 |
| 1,000,000 | 0.93 | 0.012 |