

## Other operations lab: indexes and SQLite performance

1. Download the file 'campaign-2008-all.zip' file using the Google Drive link on our iLearn page, and unzip the file (about 110 MB). Also, download 'campaign.sql' from our iLearn page.
2. Load the data into sqlite:

```
sqlite> .read campaign.sql
```

You may need to modify campaign.sql so that it looks in the right place for the .csv file.

3. The data will take a while to load. It consists of a single table. Look at the number of rows in this table
4. Run this query and note approximately how long it takes to run  

```
select avg(contb_receipt_amt) from campaign where contbr_nm="SMITH, JAMES";
```
5. Look at the query plan for the query  

```
explain query plan  
select avg(contb_receipt_amt) from campaign where contbr_nm="SMITH, JAMES";
```
6. Create an index for column contbr\_nm. Note how long this takes.
7. Re-run step 5. What has changed?
8. Re-run the query in step 4. Do you see a difference in performance?
9. For this query, look at the query plan, run the query, and note how long it takes to run:  

```
select * from campaign where contbr_zip = "93933" and cand_nm="Paul, Ron";
```
10. Would it be better to create an index on contbr\_zip or cand\_nm? Think about this. Then try creating an index for cand\_nm:  

```
create index cand_index on campaign(cand_nm);
```

Run the query of question 10 again and note any difference. Now delete the index you just created:

```
drop index cand_index;
```

and create an index on contbr\_zip, and run the query of item 10 again. Does the index make any difference?

## Appendix:

### Creating an index in SQLite:

```
sqlite> create index trans_index on campaign(tran_id);
```

### Listing and deleting indexes:

```
sqlite> .index
trans_index
sqlite>
sqlite> drop index trans_index;
```

### Using analyze:

```
sqlite> analyze;
sqlite>
```

### Looking at SQLite's internal statistics table:

```
sqlite> select * from sqlite_stat1;
contribution,,4085665
contributor,,1535236
candidate,sqlite_autoindex_candidate_1,"20 1"
```