baseball.md 2/7/2020

PRIMARY KEYS AND FOREIGN KEYS

When Primary keys and Foreign keys where enabled query a and b finished in 0.02 seconds and 0.18 seconds respectively. The results of the two queries are shown below.

NO FOREIGN KEYS

When Primary keys where enabled and Foreign keys where disabled query a and b finished in 0.03 seconds and 0.17 seconds respectively. The results of the two queries are shown below. As the queries finished so quickly it is hard to gauge if this affects the performance or not.

```
+-----+
| count(playerID) |
+------+
| 458 |
+-----+
1 row in set (0.03 sec)

+----+
| playerID | totalPay |
+-----+
| belleal01 | 37417830 |
| kershcl01 | 33000000 |
| lackejo01 | 31950000 |
+------+
3 rows in set (0.17 sec)
```

NO PRIMARY KEYS AND FOREIGN KEYS

When both Primary keys where and Foreign keys where disabled query a and b finished in 0.04 seconds and 14.47 seconds respectively. It can be seen that query b took significantly longer when compared to the

baseball.md 2/7/2020

query with Primary Keys and Foreign Keys enabled (0.18 sec). As such it can be concluded that Primary keys improve performance significantly.

```
+-----+
| count(playerID) |
+-----+
| 458 |
+-----+
1 row in set (0.04 se
+-----+
| playerID | totalPay |
+-----+
| belleal01 | 37417830 |
| kershcl01 | 33000000 |
| lackejo01 | 31950000 |
+-----+
3 rows in set (1 min 14.47 sec)
```

EFFECTS OF INDEXES

I added indexes on yearID for tables Salaries, Managers, Appearances and indexes on birthyear, birthmonth, birthcountry, deathcountry on table Master. Query a and b took 0.04 seconds and 0.16 seconds respectively. It can be seen that the effect on performance is negligible with these chosen indexes for this dataset as it is very similiar to the results with Primary and Foreign keys. These attributes where chosen as they seem to have numerous tuples with the same values, so it seemed bucketing them would be logical i.e many players born in the same year.