

3.

(a) How many players have an unknown birthdate?

```
SELECT COUNT(playerID)
FROM Master
WHERE birthYear = '' or birthMonth = '' or birthDay = '';
```

Without any index:

id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
1	SIMPLE	Master	NULL	ALL	NULL	NULL	NULL	NULL	18931	27.10	Using where

Only primary key:

Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Visible	Expression
Master	0	PRIMARY	1	playerID	A	18582	NULL	NULL	NULL	BTREE			YES	NULL

id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
1	SIMPLE	Master	NULL	ALL	PRIMARY	NULL	NULL	NULL	18582	27.10	Using where

Add index:

Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Visible	Expression
Master	1	birthYear	1	birthYear	A	167	NULL	NULL	YES	BTREE			YES	NULL
Master	1	birthMonth	1	birthMonth	A	13	NULL	NULL	YES	BTREE			YES	NULL
Master	1	birthDay	1	birthDay	A	32	NULL	NULL	YES	BTREE			YES	NULL

id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
1	SIMPLE	Master	NULL	index_merge	birthYear,birthMonth,birthDay	birthYear,birthMonth,birthDay	5,5,5	NULL	883	100.00	Using union(birthYear,birthMonth,birthDay); Using where

Thus, we can see the approximated number of rows to scan when adding index is much lower than that of only primary and without any indexes.

(b) How many people are in the Hall of Fame? What fraction of each category of person are in the Hall Of Fame? Are more people in the Hall Of Fame alive or dead? Does this vary by category?

How many people are in the Hall of Fame?

```
SELECT COUNT(DISTINCT playerID)
FROM HallOfFame;
```

Without any index:

id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
1	SIMPLE	HallOfFame	NULL	ALL	NULL	NULL	NULL	NULL	4136	100.00	NULL

With primary key only:

Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Visible	Expression
HALLOFFAME	0	PRIMARY	1	playerID	A	1260	NULL	NULL	BTREE				YES	NULL
HALLOFFAME	0	PRIMARY	2	yearid	A	4024	NULL	NULL	BTREE				YES	NULL
HALLOFFAME	0	PRIMARY	3	votedBy	A	4156	NULL	NULL	BTREE				YES	NULL
id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra			
1	SIMPLE	HallOfFame	NULL	index	PRIMARY	PRIMARY	1538	NULL	4156	100.00	Using index			

With Foreign key and primary key:

Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Visible	Expression
HALLOFFAME	0	PRIMARY	1	playerID	A	1260	NULL	NULL	BTREE				YES	NULL
HALLOFFAME	0	PRIMARY	2	yearid	A	4024	NULL	NULL	BTREE				YES	NULL
HALLOFFAME	0	PRIMARY	3	votedBy	A	4156	NULL	NULL	BTREE				YES	NULL
HALLOFFAME CREATE TABLE `HALLOFFAME` (`playerID` varchar(255) NOT NULL, `yearid` int NOT NULL, `votedBy` varchar(255) NOT NULL, `ballots` int DEFAULT NULL, `needed` int DEFAULT NULL, `votes` int DEFAULT NULL, `inducted` varchar(255) DEFAULT NULL, `category` varchar(255) DEFAULT NULL, `needed_note` varchar(255) DEFAULT NULL, PRIMARY KEY (`playerID`,`yearid`,`votedBy`), CONSTRAINT `halloffame_ibfk_1` FOREIGN KEY (`playerID`) REFERENCES `Master` (`playerID`)) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci														

id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
1	SIMPLE	HallOfFame	NULL	index	PRIMARY	PRIMARY	2048	NULL	4156	100.00	Using index

Add indexes:

Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Visible	Expression
HALLOFFAME	1	PLAYERID	1	playerID	A	1260	NULL	NULL	BTREE				YES	NULL
id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra			
1	SIMPLE	HallOfFame	NULL	range	PLAYERID	PLAYERID	767	NULL	1261	100.00	Using index for group-by (scanning)			

What fraction of each category of person are in the Hall Of Fame? Are more people in the Hall Of Fame alive or dead? Does this vary by category?

```

SELECT t1.category, t2.dead, t3.alive, t1.total, t1.fraction
FROM
(SELECT DISTINCT category,COUNT(DISTINCT playerID)/(SELECT COUNT(DISTINCT playerID) FROM HallOfFame) AS fraction, COUNT(DISTINCT playerID) AS total
FROM HallOfFame
group by category) t1
INNER JOIN
(SELECT DISTINCT category,Count(DISTINCT Master.playerID) AS dead
FROM HallOfFame INNER JOIN Master
ON (HallOfFame.playerID = Master.playerID)
WHERE deathDay != '' and deathMonth != '' and deathYear != ''
group by category) t2
ON t1.category = t2.category
INNER JOIN
(SELECT DISTINCT category, Count(DISTINCT Master.playerID) AS alive
FROM HallOfFame INNER JOIN Master
ON (HallOfFame.playerID = Master.playerID)
WHERE deathDay = '' and deathMonth = '' and deathYear = ''
group by category) t3
ON t1.category = t3.category;

```

Without any index:

	id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
	1	PRIMARY	<derived2>	NULL	ALL	NULL	NULL	NULL	t1.category	4136	100.00	Using where
	1	PRIMARY	<derived5>	NULL	ref	<auto_key0>	<auto_key0>	768	t1.category	10	100.00	NULL
	1	PRIMARY	<derived4>	NULL	ref	<auto_key0>	<auto_key0>	768	t1.category	1380	100.00	NULL
	5	DERIVED	Master	NULL	ALL	NULL	NULL	NULL		18931	0.10	Using where; Using temporary; Using filesort
	5	DERIVED	HallofFame	NULL	ALL	NULL	NULL	NULL		4136	10.00	Using where; Using join buffer (hash join)
	4	DERIVED	HallofFame	NULL	ALL	NULL	NULL	NULL		4136	100.00	Using temporary; Using filesort
	4	DERIVED	Master	NULL	ALL	NULL	NULL	NULL		18931	7.29	Using where; Using join buffer (hash join)
	2	DERIVED	HallofFame	NULL	ALL	NULL	NULL	NULL		4136	100.00	Using filesort
	3	SUBQUERY	HallofFame	NULL	ALL	NULL	NULL	NULL		4136	100.00	NULL

Only with primary key:

Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Visible	Expression
HALLOFFAME	0	PRIMARY	1	playerID	A	1268	NULL	NULL		BTREE			YES	NULL
HALLOFFAME	0	PRIMARY	2	yearid	A	4024	NULL	NULL		BTREE			YES	NULL
HALLOFFAME	0	PRIMARY	3	votedBy	A	4156	NULL	NULL		BTREE			YES	NULL

Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Visible	Expression
MASTER	0	PRIMARY	1	playerID	A	18961	NULL	NULL		BTREE			YES	NULL

	id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
	1	PRIMARY	<derived5>	NULL	ALL	NULL	NULL	NULL		62	100.00	Using where
	1	PRIMARY	<derived4>	NULL	ref	<auto_key0>	<auto_key0>	768	t3.category	98	100.00	NULL
	1	PRIMARY	<derived2>	NULL	ref	<auto_key0>	<auto_key0>	768	t3.category	41	100.00	NULL
	5	DERIVED	Master	NULL	ALL	PRIMARY	PRIMARY	NULL		18961	0.10	Using where; Using temporary; Using filesort
	5	DERIVED	HallofFame	NULL	ref	PRIMARY	PRIMARY	767	baseball.Master.playerID	3	100.00	NULL
	4	DERIVED	HallofFame	NULL	ALL	PRIMARY	NULL	NULL		4156	100.00	Using filesort
	4	DERIVED	Master	NULL	eq_ref	PRIMARY	PRIMARY	767	baseball.HallofFame.playerID	1	72.90	Using where
	2	DERIVED	HallofFame	NULL	ALL	NULL	NULL	NULL		4156	100.00	Using filesort
	3	SUBQUERY	HallofFame	NULL	index	PRIMARY	PRIMARY	1538	NULL	4156	100.00	Using index

With primary key and foreign key:

Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Visible	Expression
Haloffame	0	PRIMARY	1	playerID	A	1269	NULL	NULL		BTREE			YES	NULL
Haloffame	0	PRIMARY	2	yearid	A	4024	NULL	NULL		BTREE			YES	NULL
Haloffame	0	PRIMARY	3	votedBy	A	4156	NULL	NULL		BTREE			YES	NULL

Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Visible	Expression
Master	0	PRIMARY	1	playerID	A	18582	NULL	NULL		BTREE			YES	NULL

```

| halloffame | CREATE TABLE `halloffame` (
`playerID` varchar(255) NOT NULL,
`yearid` int NOT NULL,
`votedBy` varchar(255) NOT NULL,
`ballots` int DEFAULT NULL,
`needed` int DEFAULT NULL,
`votes` int DEFAULT NULL,
`inducted` varchar(255) DEFAULT NULL,
`category` varchar(255) DEFAULT NULL,
`needed_note` varchar(255) DEFAULT NULL,
PRIMARY KEY (`playerID`, `yearid`, `votedBy`),
CONSTRAINT `halloffame_ibfk_1` FOREIGN KEY (`playerID`) REFERENCES `Master`(`playerID`)
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci |

```

id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows			Extra
									rows	filtered	Extra	
1	PRIMARY	<derived5>	NULL	ALL	NULL	NULL	NULL	NULL	61	100.00	NULL	
1	PRIMARY	<derived4>	NULL	ALL	NULL	NULL	NULL	NULL	3029	10.00	Using where; Using join buffer (hash join)	
1	PRIMARY	<derived2>	NULL	ALL	NULL	NULL	NULL	NULL	4156	10.00	Using where; Using join buffer (hash join)	
5	DERIVED	Master	NULL	ALL	PRIMARY	PRIMARY	NULL	NULL	18582	0.10	Using where; Using temporary; Using filesort	
5	DERIVED	Halloffame	NULL	ref	PRIMARY	PRIMARY	1022	baseball_pkfk.Master.playerID	3	100.00	NULL	
4	DERIVED	Halloffame	NULL	ALL	PRIMARY	PRIMARY	NULL	NULL	4156	100.00	Using filesort	
4	DERIVED	Master	NULL	eq_ref	PRIMARY	PRIMARY	1022	baseball_pkfk.Halloffame.playerID	1	72.90	Using where	
2	DERIVED	Halloffame	NULL	ALL	NULL	NULL	NULL	NULL	4156	100.00	Using filesort	
3	SUBQUERY	Halloffame	NULL	index	PRIMARY	PRIMARY	2048	NULL	4156	100.00	Using index	

Add index on category:

mysql> SHOW INDEXES IN HallOffFame;														
Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Visible	Expression
Halloffame	0	PRIMARY	1	playerID	A	1260	NULL	NULL	BTREE				YES	NULL
Halloffame	0	PRIMARY	2	yearid	A	4024	NULL	NULL	BTREE				YES	NULL
Halloffame	0	PRIMARY	3	votedby	A	4156	NULL	NULL	BTREE				YES	NULL
Halloffame	1	PLAYERID	1	playerID	A	1260	NULL	NULL	BTREE				YES	NULL
Halloffame	1	Category	1	category	A	4	NULL	NULL	YES	BTREE			YES	NULL

mysql> SHOW INDEXES IN Master;														
Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Visible	Expression
MASTER	0	PRIMARY	1	playerID	A	18961	NULL	NULL	BTREE				YES	NULL

+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+													
id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra		
1	PRIMARY	<derived5>	NULL	ALL	NULL	NULL	NULL	NULL	62	100.00	Using where		
1	PRIMARY	<derived2>	NULL	ref	cauto_key0>	cauto_key0>	768	t2.category	12	100.00	Using temporary		
1	PRIMARY	<derived4>	NULL	ref	cauto_key0>	cauto_key0>	768	t3.category	30	100.00	NULL		
5	DERIVED	Master	NULL	ALL	PRIMARY	PRIMARY	NULL	NULL	18862	0.10	Using where; Using temporary; Using filesort		
5	DERIVED	Halloffame	NULL	ref	PRIMARY,PLAYERID,Category	PRIMARY,PLAYERID,Category	767	baseball.Master.playerID	3	100.00	NULL		
4	DERIVED	Halloffame	NULL	index	PRIMARY,PLAYERID,Category	PRIMARY,PLAYERID,Category	767	baseball.Master.playerID	4156	100.00	Using index		
4	DERIVED	Master	NULL	eq_ref	PRIMARY	PRIMARY	767	baseball.Halloffame.playerID	1	72.90	Using where		
2	DERIVED	Halloffame	NULL	range	Category	Category	1535	NULL	1277	100.00	Using index for group-by (scanning)		
3	SUBQUERY	Halloffame	NULL	range	PRIMARY,PLAYERID,Category	PLAYERID	767	NULL	1261	100.00	Using index for group-by (scanning)		

- (c) What are the names and total pay (individually) of the three people with the three largest total salaries? What category are these people (Players? Managers? Other?)? What are the top three by category?

What are the names and total pay (individually) of the three people with the three largest total salaries?

```

SELECT t2.nameFirst, t2.nameLast, t2.nameGiven, total

FROM

(SELECT DISTINCT playerID, SUM(salary) AS total
FROM Salaries
group by playerID
) t1

INNER JOIN

(
SELECT playerID, nameFirst, nameLast, nameGiven
FROM Master
) t2

ON (t1.playerID = t2.playerID)

Order BY total DESC
LIMIT 3;
    
```

Without any index:

ID	Select_Type	Table	Partitions	Type	Possible_Keys	Key	Key_Len	Ref	Rows	Filtered	Extra
1	PRIMARY	Master	NULL	ALL	NULL	NULL	NULL	baseball.Master.playerID	18952	100.00	Using temporary; Using filesort
1	PRIMARY	<derived2>	NULL	ref	<auto_key0>	<auto_key0>	767	baseball.Master.playerID	18	100.00	NULL
2	DERIVED	Salaries	NULL	ALL	NULL	NULL	NULL	NULL	26112	100.00	Using temporary

Only with primary key:

Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Visible	Expression
Salaries	0	PRIMARY	1	yearID	A	31	NULL	NULL	BTREE				YES	NULL
Salaries	0	PRIMARY	2	teamID	A	928	NULL	NULL	BTREE				YES	NULL
Salaries	0	PRIMARY	3	playerID	A	26648	NULL	NULL	BTREE				YES	NULL
Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Visible	Expression
MASTER	0	PRIMARY	1	playerID	A	18961	NULL	NULL	BTREE				YES	NULL

ID	Select_Type	Table	Partitions	Type	Possible_Keys	Key	Key_Len	Ref	Rows	Filtered	Extra
1	PRIMARY	<derived2>	NULL	ALL	NULL	NULL	NULL	NULL	26644	100.00	Using filesort
1	PRIMARY	Master	NULL	eq_ref	PRIMARY	PRIMARY	767	t1.playerID	1	100.00	NULL
2	DERIVED	Salaries	NULL	ALL	PRIMARY	NULL	NULL	NULL	26644	100.00	Using temporary

With pk and fk:

Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Visible	Expression
Salaries	0	PRIMARY	1	teamID	A	43	NULL	NULL	BTREE				YES	NULL
Salaries	0	PRIMARY	2	yearID	A	936	NULL	NULL	BTREE				YES	NULL
Salaries	0	PRIMARY	3	playerID	A	28571	NULL	NULL	BTREE				YES	NULL
Salaries	1	playerID	1	playerID	A	5224	NULL	NULL	BTREE				YES	NULL
Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Visible	Expression
Master	0	PRIMARY	1	playerID	A	18582	NULL	NULL	BTREE				YES	NULL

ID	Select_Type	Table	Partitions	Type	Possible_Keys	Key	Key_Len	Ref	Rows	Filtered	Extra
1	PRIMARY	<derived2>	NULL	ALL	NULL	NULL	NULL	NULL	2	100.00	Using filesort
1	PRIMARY	Master	NULL	eq_ref	PRIMARY	PRIMARY	1022	t1.playerID	1	100.00	NULL
2	DERIVED	Salaries	NULL	index	PRIMARY,playerID	playerID	1022	NULL	1	100.00	NULL

With index:

All the needed index has been in with pk and fk.

What category are these people (Players? Managers? Other?)? What are the top three by category?

```

SELECT t2.nameFirst, t2.nameLast, t2.nameGiven, total
FROM
(SELECT DISTINCT playerID, SUM(salary) AS total
FROM Salaries
WHERE playerID IN (SELECT playerID FROM Master) AND playerID NOT IN (SELECT playerID FROM Managers)
group by playerID
) t1
INNER JOIN
(
SELECT playerID, nameFirst, nameLast, nameGiven
FROM Master
) t2
ON (t1.playerID = t2.playerID)
Order BY total DESC
LIMIT 3;

```

Players Only:

Without any index:

ID	Select_Type	Table	Partitions	Type	Possible_Keys	Key	Key_Len	Ref	Rows	Filtered	Extra
1	PRIMARY	Master	NULL	ALL	NULL	NULL	NULL	baseball.Master.playerID	18552	100.00	Using temporary; Using filesort
1	PRIMARY	<derived2>	NULL	ref	<auto_key0>	<auto_key0>	767	baseball.Master.playerID	18	100.00	NULL
2	DERIVED	<subquery3>	NULL	ALL	NULL	NULL	NULL	NULL	NULL	100.00	Using temporary
2	DERIVED	Salaries	NULL	ALL	NULL	NULL	NULL	NULL	26784	10.00	Using where; Using join buffer (hash join)
2	DERIVED	<subquery3>	NULL	eq_ref	<auto_distinct_key>	<auto_distinct_key>	768	<subquery3>.playerID	1	100.00	Using where; Not exists
3	MATERIALIZED	Master	NULL	ALL	NULL	NULL	NULL	NULL	18552	100.00	NULL
4	MATERIALIZED	Managers	NULL	ALL	NULL	NULL	NULL	NULL	3436	100.00	NULL

With pk:

```

+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Table | Non_unique | Key_name | Seq_in_index | Column_name | Collation | Cardinality | Sub_part | Packed | Null | Index_type | Comment | Index_comment | Visible | Expression |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Salaries | 0 | PRIMARY | 1 | yearID | A | 31 | NULL | NULL | BTREE | | YES | NULL |
| Salaries | 0 | PRIMARY | 2 | teamID | A | 920 | NULL | NULL | BTREE | | YES | NULL |
| Salaries | 0 | PRIMARY | 3 | playerID | A | 26648 | NULL | NULL | BTREE | | YES | NULL |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
3 rows in set (0.00 sec)

mysql> SHOW INDEXES IN Managers;
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Table | Non_unique | Key_name | Seq_in_index | Column_name | Collation | Cardinality | Sub_part | Packed | Null | Index_type | Comment | Index_comment | Visible | Expression |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Managers | 0 | PRIMARY | 1 | playerID | A | 698 | NULL | NULL | BTREE | | YES | NULL |
| Managers | 0 | PRIMARY | 2 | yearID | A | 3385 | NULL | NULL | BTREE | | YES | NULL |
| Managers | 0 | PRIMARY | 3 | inseason | A | 3436 | NULL | NULL | BTREE | | YES | NULL |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
3 rows in set (0.01 sec)

mysql> SHOW INDEXES IN Master;
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Table | Non_unique | Key_name | Seq_in_index | Column_name | Collation | Cardinality | Sub_part | Packed | Null | Index_type | Comment | Index_comment | Visible | Expression |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| MASTER | 0 | PRIMARY | 1 | playerID | A | 18961 | NULL | NULL | BTREE | | YES | NULL |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| id | select_type | table | partitions | type | possible_keys | key | key_len | ref | rows | filtered | Extra |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| 1 | PRIMARY | <derived2> | NULL | ALL | NULL | NULL | NULL | 26644 | 100.00 | Using filesort | |
| 1 | PRIMARY | Master | NULL | eq_ref | PRIMARY | PRIMARY | 767 | t1.playerID | 1 | 100.00 | NULL |
| 2 | DERIVED | Salaries | NULL | ALL | PRIMARY | NULL | NULL | 26644 | 100.00 | Using temporary |
| 2 | DERIVED | Master | NULL | eq_ref | PRIMARY | PRIMARY | 767 | baseball.Salaries.playerID | 1 | 100.00 | Using index |
| 2 | DERIVED | <subquery4> | NULL | eq_ref | <auto_distinct_key> | <auto_distinct_key> | 768 | baseball.Salaries.playerID | 1 | 100.00 | Using where; Not exists |
| 4 | MATERIALIZED | Managers | NULL | index | PRIMARY | PRIMARY | 775 | NULL | 3436 | 100.00 | Using index |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

```

With pk and fk:

```

+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Table | Non_unique | Key_name | Seq_in_index | Column_name | Collation | Cardinality | Sub_part | Packed | Null | Index_type | Comment | Index_comment | Visible | Expression |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Salaries | 0 | PRIMARY | 1 | teamID | A | 43 | NULL | NULL | BTREE | | YES | NULL |
| Salaries | 0 | PRIMARY | 2 | yearID | A | 936 | NULL | NULL | BTREE | | YES | NULL |
| Salaries | 0 | PRIMARY | 3 | playerID | A | 28571 | NULL | NULL | BTREE | | YES | NULL |
| Salaries | 1 | playerID | A | 5224 | NULL | NULL | BTREE | | YES | NULL |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
4 rows in set (0.00 sec)

mysql> SHOW INDEXES IN Managers;
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Table | Non_unique | Key_name | Seq_in_index | Column_name | Collation | Cardinality | Sub_part | Packed | Null | Index_type | Comment | Index_comment | Visible | Expression |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Managers | 0 | PRIMARY | 1 | playerID | A | 698 | NULL | NULL | BTREE | | YES | NULL |
| Managers | 0 | PRIMARY | 2 | yearID | A | 3385 | NULL | NULL | BTREE | | YES | NULL |
| Managers | 0 | PRIMARY | 3 | inseason | A | 3436 | NULL | NULL | BTREE | | YES | NULL |
| Managers | 1 | yearID | A | 146 | NULL | NULL | BTREE | | YES | NULL |
| Managers | 1 | yearID | A | 276 | NULL | NULL | BTREE | | YES | NULL |
| Managers | 1 | yearID | A | 2835 | NULL | NULL | BTREE | | YES | NULL |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
6 rows in set (0.00 sec)

mysql> SHOW INDEXES IN Master;
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Table | Non_unique | Key_name | Seq_in_index | Column_name | Collation | Cardinality | Sub_part | Packed | Null | Index_type | Comment | Index_comment | Visible | Expression |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Master | 0 | PRIMARY | 1 | playerID | A | 18582 | NULL | NULL | BTREE | | YES | NULL |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| id | select_type | table | partitions | type | possible_keys | key | key_len | ref | rows | filtered | Extra |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| 1 | PRIMARY | <derived2> | NULL | ALL | NULL | NULL | NULL | 2 | 100.00 | Using filesort | |
| 1 | PRIMARY | Master | NULL | eq_ref | PRIMARY | PRIMARY | 1022 | t1.playerID | 1 | 100.00 | NULL |
| 2 | DERIVED | Salaries | NULL | ALL | PRIMARY,playerID | NULL | NULL | 1 | 100.00 | Using temporary |
| 2 | DERIVED | Master | NULL | eq_ref | PRIMARY | PRIMARY | 1022 | baseball_pkfk.Salaries.playerID | 1 | 100.00 | Using index |
| 2 | DERIVED | Managers | NULL | ref | PRIMARY | PRIMARY | 1022 | baseball_pkfk.Salaries.playerID | 1 | 100.00 | Using where; Not exists; Using index |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

```

With index:

All the needed index has been in with pk and fk.

Managers Only:

```

SELECT t2.nameFirst, t2.nameLast, t2.nameGiven, total

FROM

(SELECT DISTINCT playerID, SUM(salary) AS total
FROM Salaries
WHERE playerID IN (SELECT playerID FROM Managers)
GROUP BY playerID
) t1

INNER JOIN

(
SELECT playerID, nameFirst, nameLast, nameGiven
FROM Master

) t2

ON (t1.playerID = t2.playerID)

Order BY total DESC
LIMIT 3;

```

Without any index:

	id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
1	PRIMARY		Master	NULL	ALL	NULL	<auto_key0>	NULL	NULL	18931	100.00	Using temporary; Using filesort
1	PRIMARY		<derived2>	NULL	ALL	<auto_key0>	<auto_key0>	767	baseball.Master.playerID	1	100.00	NULL
2	DERIVED		<subquery3>	NULL	ALL	NULL	NULL	NULL	NULL	1	100.00	Using temporary
2	DERIVED		Salaries	NULL	ALL	NULL	NULL	NULL	NULL	26117	100.00	Using where; Using join buffer (hash join)
3	MATERIALIZED		Managers	NULL	ALL	NULL	NULL	NULL	NULL	3436	100.00	NULL

With pk:

Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Visible	Expression
Salaries	0	PRIMARY	1	yearID	A	21	NULL	NULL	BTREE				YES	NULL
Salaries	0	PRIMARY	2	teamID	A	928	NULL	NULL	BTREE				YES	NULL
Salaries	0	PRIMARY	3	playerID	A	26648	NULL	NULL	BTREE				YES	NULL
3 rows in set (0.00 sec)														
mysql> SHOW INDEXES IN Managers;														
Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Visible	Expression
Managers	0	PRIMARY	1	playerID	A	698	NULL	NULL	BTREE				YES	NULL
Managers	0	PRIMARY	2	yearID	A	3385	NULL	NULL	BTREE				YES	NULL
Managers	0	PRIMARY	3	inseason	A	3436	NULL	NULL	BTREE				YES	NULL
3 rows in set (0.00 sec)														
mysql> SHOW INDEXES IN Master;														
Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Visible	Expression
MASTER	0	PRIMARY	1	playerID	A	18961	NULL	NULL	BTREE				YES	NULL
	id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra		
1	PRIMARY	<derived2>	NULL	ALL	NULL	NULL	PRIMARY	NULL	NULL	26648	100.00	Using filesort		
1	PRIMARY	Master	NULL	eq_ref	PRIMARY	PRIMARY	767	t1.playerID	1	100.00	NULL			
2	DERIVED	Salaries	NULL	ALL	PRIMARY	NULL	NULL	NULL	NULL	26648	100.00	Using temporary		
2	DERIVED	<subquery3>	NULL	eq_ref	<auto_distinct_key>	<auto_distinct_key>	767	baseball.Salaries.playerID	1	100.00	NULL			
3	MATERIALIZED	Managers	NULL	index	PRIMARY	PRIMARY	775	NULL	NULL	3436	100.00	Using index		

With pk and fk:

```

+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Table | Non_unique | Key_name | Seq_in_index | Column_name | Collation | Cardinality | Sub_part | Packed | Null | Index_type | Comment | Index_comment | Visible | Expression |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Salaries | 0 | PRIMARY | 1 | teamID | A | 43 | NULL | NULL | BTREE | | YES | NULL |
| Salaries | 0 | PRIMARY | 2 | yearID | A | 936 | NULL | NULL | BTREE | | YES | NULL |
| Salaries | 0 | PRIMARY | 3 | playerID | A | 28571 | NULL | NULL | BTREE | | YES | NULL |
| Salaries | 1 | playerID | 1 | playerID | A | 5224 | NULL | NULL | BTREE | | YES | NULL |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
4 rows in set (0.00 sec)

mysql> SHOW INDEXES IN Managers;
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Table | Non_unique | Key_name | Seq_in_index | Column_name | Collation | Cardinality | Sub_part | Packed | Null | Index_type | Comment | Index_comment | Visible | Expression |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Managers | 0 | PRIMARY | 1 | playerID | A | 698 | NULL | NULL | BTREE | | YES | NULL | |
| Managers | 0 | PRIMARY | 2 | yearID | A | 3385 | NULL | NULL | BTREE | | YES | NULL |
| Managers | 0 | PRIMARY | 3 | inseason | A | 3436 | NULL | NULL | BTREE | | YES | NULL |
| Managers | 1 | yearID | 1 | yearID | A | 146 | NULL | NULL | BTREE | | YES | NULL |
| Managers | 1 | yearID | 2 | lgID | A | 276 | NULL | NULL | YES | BTREE | | YES | NULL |
| Managers | 1 | yearID | 3 | teamID | A | 2835 | NULL | NULL | YES | BTREE | | YES | NULL |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
6 rows in set (0.00 sec)

mysql> SHOW INDEXES IN Master;
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Table | Non_unique | Key_name | Seq_in_index | Column_name | Collation | Cardinality | Sub_part | Packed | Null | Index_type | Comment | Index_comment | Visible | Expression |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Master | 0 | PRIMARY | 1 | playerID | A | 18582 | NULL | NULL | BTREE | | YES | NULL |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
1 row in set (0.00 sec)

+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| id | select_type | table | partitions | type | possible_keys | key | key_len | ref | rows | filtered | Extra |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| 1 | PRIMARY | <derived> | NULL | ALL | NULL | NULL | NULL | NULL | 2 | 100.00 | Using filesort |
| 1 | PRIMARY | Master | NULL | eq_ref | PRIMARY | PRIMARY | 1022 | t1.playerID | 1 | 100.00 | NULL |
| 2 | DERIVED | Salaries | NULL | ALL | PRIMARY,playerID | NULL | NULL | NULL | 1 | 100.00 | Using temporary |
| 2 | DERIVED | Managers | NULL | ref | PRIMARY | PRIMARY | 1022 | baseball_pkfk.Salaries.playerID | 4 | 100.00 | Using index; FirstMatch(Salaries) |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

```

With index:

All the needed index has been in with pk and fk.

Other:

```

SELECT t2.nameFirst, t2.nameLast, t2.nameGiven, total
FROM
(SELECT DISTINCT playerID, SUM(salary) AS total
FROM Salaries
WHERE playerID NOT IN (SELECT playerID FROM Managers) AND playerID NOT IN (SELECT playerID FROM Master)
GROUP BY playerID
) t1
INNER JOIN
(
SELECT playerID, nameFirst, nameLast, nameGiven
FROM Master
) t2
ON (t1.playerID = t2.playerID)
ORDER BY total DESC
LIMIT 3;

```

Without any index:

```

+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| id | select_type | table | partitions | type | possible_keys | key | key_len | ref | rows | filtered | Extra |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| 1 | PRIMARY | Master | NULL | ALL | NULL | <auto_key0> | <auto_key0> | NULL | 18931 | 100.00 | Using temporary; Using filesort |
| 1 | PRIMARY | <derived> | NULL | ref | <auto_key0> | <auto_key0> | 73 | baseball.Master.playerID | 18 | 100.00 | NULL |
| 2 | DERIVED | Salaries | NULL | ALL | NULL | <auto_distinct_key> | <auto_distinct_key> | NULL | 2673 | 100.00 | Using temporary |
| 2 | DERIVED | <subquery3> | NULL | eq_ref | <auto_distinct_key> | <auto_distinct_key> | 768 | baseball.Salaries.playerID | 1 | 100.00 | Using where; Not exists |
| 2 | DERIVED | <subquery4> | NULL | eq_ref | <auto_distinct_key> | <auto_distinct_key> | 768 | baseball.Salaries.playerID | 1 | 100.00 | Using where; Not exists |
| 3 | MATERIALIZED | Managers | NULL | ALL | NULL | <auto_key0> | <auto_key0> | NULL | 3436 | 100.00 | NULL |
| 4 | MATERIALIZED | Master | NULL | ALL | NULL | <auto_key0> | <auto_key0> | NULL | 18931 | 100.00 | NULL |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

```

With pk:

```
| Table | Non_unique | Key_name | Seq_in_index | Column_name | Collation | Cardinality | Sub_part | Packed | Null | Index_type | Comment | Index_comment | Visible | Expression |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Salaries | 0 | PRIMARY | 1 | yearID | A | 31 | NULL | NULL | BTREE | | YES | NULL |
| Salaries | 0 | PRIMARY | 2 | teamID | A | 920 | NULL | NULL | BTREE | | YES | NULL |
| Salaries | 0 | PRIMARY | 3 | playerID | A | 26648 | NULL | NULL | BTREE | | YES | NULL |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
3 rows in set (0.00 sec)

mysql> SHOW INDEXES IN Managers;
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Table | Non_unique | Key_name | Seq_in_index | Column_name | Collation | Cardinality | Sub_part | Packed | Null | Index_type | Comment | Index_comment | Visible | Expression |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Managers | 0 | PRIMARY | 1 | playerID | A | 698 | NULL | NULL | BTREE | | YES | NULL |
| Managers | 0 | PRIMARY | 2 | yearID | A | 3385 | NULL | NULL | BTREE | | YES | NULL |
| Managers | 0 | PRIMARY | 3 | inseason | A | 3436 | NULL | NULL | BTREE | | YES | NULL |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
3 rows in set (0.01 sec)

mysql> SHOW INDEXES IN Master;
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Table | Non_unique | Key_name | Seq_in_index | Column_name | Collation | Cardinality | Sub_part | Packed | Null | Index_type | Comment | Index_comment | Visible | Expression |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| MASTER | 0 | PRIMARY | 1 | playerID | A | 18961 | NULL | NULL | BTREE | | YES | NULL |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
4 rows in set (0.00 sec)

+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| id | select_type | table | partitions | type | possible_keys | key | key_len | ref | rows | filtered | Extra |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| 1 | PRIMARY | <derived2> | NULL | ALL | NULL | NULL | NULL | NULL | 26648 | 100.00 | Using filesort |
| 3 | PRIMARY | Master | NULL | eq_ref | PRIMARY | PRIMARY | 767 | t1.playerID | 1 | 100.00 | NULL |
| 2 | DERIVED | Salaries | NULL | ALL | PRIMARY | NULL | NULL | NULL | 26648 | 100.00 | Using temporary |
| 2 | DERIVED | <subquery3> | NULL | eq_ref | <auto_distinct_key> | <auto_distinct_key> | 768 | baseball.Salaries.playerID | 1 | 100.00 | Using where; Not exists |
| 2 | DERIVED | <subquery4> | NULL | eq_ref | <auto_distinct_key> | <auto_distinct_key> | 768 | baseball.Salaries.playerID | 1 | 100.00 | Using where; Not exists |
| 3 | MATERIALIZED | Managers | NULL | index | PRIMARY | PRIMARY | 775 | NULL | 3436 | 100.00 | Using index |
| 4 | MATERIALIZED | Master | NULL | index | PRIMARY | PRIMARY | 767 | NULL | 18962 | 100.00 | Using index |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
```

With pk and fk:

```
| Table | Non_unique | Key_name | Seq_in_index | Column_name | Collation | Cardinality | Sub_part | Packed | Null | Index_type | Comment | Index_comment | Visible | Expression |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Salaries | 0 | PRIMARY | 1 | teamID | A | 43 | NULL | NULL | BTREE | | YES | NULL |
| Salaries | 0 | PRIMARY | 2 | yearID | A | 936 | NULL | NULL | BTREE | | YES | NULL |
| Salaries | 0 | PRIMARY | 3 | playerID | A | 28571 | NULL | NULL | BTREE | | YES | NULL |
| Salaries | 1 | playerID | 1 | playerID | A | 5224 | NULL | NULL | BTREE | | YES | NULL |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
4 rows in set (0.00 sec)

mysql> SHOW INDEXES IN Managers;
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Table | Non_unique | Key_name | Seq_in_index | Column_name | Collation | Cardinality | Sub_part | Packed | Null | Index_type | Comment | Index_comment | Visible | Expression |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Managers | 0 | PRIMARY | 1 | playerID | A | 698 | NULL | NULL | BTREE | | YES | NULL | |
| Managers | 0 | PRIMARY | 2 | yearID | A | 3385 | NULL | NULL | BTREE | | YES | NULL |
| Managers | 0 | PRIMARY | 3 | inseason | A | 3436 | NULL | NULL | BTREE | | YES | NULL |
| Managers | 1 | yearID | 1 | yearID | A | 146 | NULL | NULL | BTREE | | YES | NULL |
| Managers | 1 | yearID | 2 | lgID | A | 276 | NULL | NULL | YES | BTREE | | YES | NULL |
| Managers | 1 | yearID | 3 | teamID | A | 2835 | NULL | NULL | YES | BTREE | | YES | NULL |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
6 rows in set (0.00 sec)

mysql> SHOW INDEXES IN Master;
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Table | Non_unique | Key_name | Seq_in_index | Column_name | Collation | Cardinality | Sub_part | Packed | Null | Index_type | Comment | Index_comment | Visible | Expression |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Master | 0 | PRIMARY | 1 | playerID | A | 18582 | NULL | NULL | BTREE | | YES | NULL |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
4 rows in set (0.00 sec)

--> LIMIT 3;

+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| id | select_type | table | partitions | type | possible_keys | key | key_len | ref | rows | filtered | Extra |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| 1 | PRIMARY | <derived2> | NULL | ALL | NULL | NULL | NULL | NULL | 2 | 100.00 | Using filesort |
| 1 | PRIMARY | Master | NULL | eq_ref | PRIMARY | PRIMARY | 1022 | t1.playerID | 1 | 100.00 | NULL |
| 2 | DERIVED | Salaries | NULL | ALL | PRIMARY,playerID | NULL | NULL | NULL | 1 | 100.00 | Using temporary |
| 2 | DERIVED | Managers | NULL | ref | PRIMARY | PRIMARY | 1022 | baseball_pkfk.Salaries.playerID | 4 | 100.00 | Using where; Not exists; Using index |
| 2 | DERIVED | Master | NULL | eq_ref | PRIMARY | PRIMARY | 1022 | baseball_pkfk.Salaries.playerID | 1 | 100.00 | Using where; Not exists; Using index |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
```

With index:

All the needed index has been in with pk and fk.

(d) What is the average number of Home Runs a player has?

```
-- SELECT (SUM(t1.player HR) / COUNT(t1.playerID)) AS HR
SELECT AVG(player HR) AS HR
FROM
(SELECT DISTINCT playerID, SUM(HR) AS player HR
FROM Batting
group by playerID) t1;
```

Without any index:

id select_type table partitions type possible_keys key key_len ref rows filtered Extra
1 PRIMARY <derived2> NULL ALL NULL NULL NULL NULL 102125 100.00 NULL
2 DERIVED Batting NULL ALL NULL NULL NULL NULL 102125 100.00 Using temporary

With pk:

Table Non_unique Key_name Seq_in_index Column_name Collation Cardinality Sub_part Packed Null Index_type Comment Index_comment Visible Expression
Batting 0 PRIMARY 1 playerID A 20287 NULL NULL BTREE YES NULL
Batting 0 PRIMARY 2 yearID A 94489 NULL NULL BTREE YES NULL
Batting 0 PRIMARY 3 stint A 102351 NULL NULL BTREE YES NULL

id select_type table partitions type possible_keys key key_len ref rows filtered Extra
1 PRIMARY <derived2> NULL ALL NULL NULL NULL NULL 102070 100.00 NULL
2 DERIVED Batting NULL index PRIMARY PRIMARY 775 NULL 102070 100.00 NULL

With pk and fk:

Table Non_unique Key_name Seq_in_index Column_name Collation Cardinality Sub_part Packed Null Index_type Comment Index_comment Visible Expression
Batting 0 PRIMARY 1 playerID A 21261 NULL NULL BTREE YES NULL
Batting 0 PRIMARY 2 yearID A 102176 NULL NULL BTREE YES NULL
Batting 0 PRIMARY 3 stint A 102176 NULL NULL BTREE YES NULL
Batting 1 yearID 1 yearID A 147 NULL NULL BTREE YES NULL
Batting 1 yearID 2 lgID A 317 NULL NULL YES BTREE YES NULL
Batting 1 yearID 3 teamID A 2625 NULL NULL YES BTREE YES NULL

id select_type table partitions type possible_keys key key_len ref rows filtered Extra
1 PRIMARY <derived2> NULL ALL NULL NULL NULL NULL 102176 100.00 NULL
2 DERIVED Batting NULL index PRIMARY,yearID PRIMARY 1030 NULL 102176 100.00 NULL

With index:

Table Non_unique Key_name Seq_in_index Column_name Collation Cardinality Sub_part Packed Null Index_type Comment Index_comment Visible Expression
Batting 0 PRIMARY 1 playerID A 21261 NULL NULL BTREE YES NULL
Batting 0 PRIMARY 2 yearID A 102176 NULL NULL BTREE YES NULL
Batting 0 PRIMARY 3 stint A 102176 NULL NULL BTREE YES NULL
Batting 1 yearID 1 yearID A 147 NULL NULL BTREE YES NULL
Batting 1 yearID 2 lgID A 317 NULL NULL YES BTREE YES NULL
Batting 1 yearID 3 teamID A 2625 NULL NULL YES BTREE YES NULL
Batting 1 PLAYERID 1 playerID A 18729 NULL NULL BTREE YES NULL

id select_type table partitions type possible_keys key key_len ref rows filtered Extra
1 PRIMARY <derived2> NULL ALL NULL NULL NULL NULL 102176 100.00 NULL
2 DERIVED Batting NULL index PRIMARY,yearID,PLAYERID PRIMARY 1030 NULL 102176 100.00 NULL

(e) If we only count players who got at least 1 Home Run, what is the average number of HomeRuns a player has?

```
SELECT AVG(plaver HR) AS HR
FROM
(SELECT DISTINCT Battina.playerID, SUM(Battina.HR) AS plaver HR
FROM Battina
group by Battina.playerID) t1
WHERE plaver HR >= 1;
```

Without any index:

id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
1	PRIMARY	<derived2>	NULL	ALL	NULL	NULL	NULL	NULL	101939	100.00	NULL
2	DERIVED	Batting	NULL	ALL	NULL	NULL	NULL	NULL	101939	100.00	Using temporary

With pk:

Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Visible	Expression
Batting	0	PRIMARY	1	playerID	A	20287	NULL	NULL	BTREE				YES	NULL
Batting	0	PRIMARY	2	yearID	A	94489	NULL	NULL	BTREE				YES	NULL
Batting	0	PRIMARY	3	stint	A	102351	NULL	NULL	BTREE				YES	NULL

id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
1	PRIMARY	<derived2>	NULL	ALL	NULL	NULL	NULL	NULL	102211	100.00	NULL
2	DERIVED	Batting	NULL	index	PRIMARY	PRIMARY	775	NULL	102211	100.00	NULL

With pk and fk:

Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Visible	Expression
Batting	0	PRIMARY	1	playerID	A	21261	NULL	NULL	BTREE				YES	NULL
Batting	0	PRIMARY	2	yearID	A	102176	NULL	NULL	BTREE				YES	NULL
Batting	0	PRIMARY	3	stint	A	102176	NULL	NULL	BTREE				YES	NULL
Batting	1	yearID	1	yearID	A	147	NULL	NULL	BTREE				YES	NULL
Batting	1	yearID	2	lgID	A	317	NULL	NULL	YES	BTREE			YES	NULL
Batting	1	yearID	3	teamID	A	2625	NULL	NULL	YES	BTREE			YES	NULL

id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
1	PRIMARY	<derived2>	NULL	ALL	NULL	NULL	NULL	NULL	102176	100.00	NULL
2	DERIVED	Batting	NULL	index	PRIMARY,yearID	PRIMARY	1030	NULL	102176	100.00	NULL

With index:

Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Visible	Expression
Batting	0	PRIMARY	1	playerID	A	21261	NULL	NULL	BTREE				YES	NULL
Batting	0	PRIMARY	2	yearID	A	102176	NULL	NULL	BTREE				YES	NULL
Batting	0	PRIMARY	3	stint	A	102176	NULL	NULL	BTREE				YES	NULL
Batting	1	yearID	1	yearID	A	147	NULL	NULL	BTREE				YES	NULL
Batting	1	yearID	2	lgID	A	317	NULL	NULL	YES	BTREE			YES	NULL
Batting	1	yearID	3	teamID	A	2625	NULL	NULL	YES	BTREE			YES	NULL
Batting	1	PLAYERID	1	playerID	A	18729	NULL	NULL	BTREE				YES	NULL
Batting	1	HR	1	HR	A	40	NULL	NULL	YES	BTREE			YES	NULL

id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
1	PRIMARY	<derived2>	NULL	ALL	NULL	NULL	NULL	NULL	102176	100.00	NULL
2	DERIVED	Batting	NULL	index	PRIMARY,yearID,PLAYERID,HR	PRIMARY	1030	NULL	102176	100.00	NULL

(f) If we define a player as a good batter if they have more than the average number of HomeRuns, and a player is a good Pitcher if they have more than the average number of ShutOut games, then how many players are both good batters and good pitchers?

```

SELECT COUNT(t6.playerID) AS Good Player
FROM
((SELECT t1.playerID
FROM
(SELECT DISTINCT playerID, SUM(Battino.HR) AS player HR FROM Battino group by Battino.playerID) t1
WHERE (t1.player HR >
(SELECT AVG(t2.player HR) AS HR FROM (SELECT DISTINCT Battino.playerID, SUM(Battino.HR) AS player HR FROM Battino group by Battino.playerID) t2)) t3
INNER JOIN
(SELECT t4.playerID
FROM
(SELECT DISTINCT playerID, SUM(Pitching.SH0) AS player SH0 FROM Pitching group by Pitching.playerID) t4
WHERE (t4.player SH0 >
(SELECT AVG(t5.player SH0) AS SH0 FROM (SELECT DISTINCT Pitching.playerID, SUM(Pitching.SH0) AS player SH0 FROM Pitching group by Pitching.playerID) t5)
ON t3.playerID = t6.playerID);

```

Without any index:

id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
1	PRIMARY	<derived7>	NULL	ALL	NULL	NULL	NULL	NULL	44628	33.33	Using where
1	PRIMARY	<derived3>	NULL	ref	<auto_key1>	<auto_key1>	767	t4.playerID	10	33.33	Using where
8	SUBQUERY	<derived9>	NULL	ALL	NULL	NULL	NULL	NULL	44628	100.00	NULL
9	DERIVED	Pitching	NULL	ALL	NULL	NULL	NULL	NULL	44628	100.00	Using temporary
7	DERIVED	Pitching	NULL	ALL	NULL	NULL	NULL	NULL	44628	100.00	Using temporary
4	SUBQUERY	<derived5>	NULL	ALL	NULL	NULL	NULL	NULL	102125	100.00	NULL
5	DERIVED	Batting	NULL	ALL	NULL	NULL	NULL	NULL	102125	100.00	Using temporary
3	DERIVED	Batting	NULL	ALL	NULL	NULL	NULL	NULL	102125	100.00	Using temporary

With pk:

Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Visible	Expression
Batting	0	PRIMARY	1	playerID	A	28287	NULL	NULL	BTREE				YES	NULL
Batting	0	PRIMARY	2	yearID	A	94489	NULL	NULL	BTREE				YES	NULL
Batting	0	PRIMARY	3	stint	A	102351	NULL	NULL	BTREE				YES	NULL
3 rows in set (0.00 sec)														
mysql> SHOW INDEXES IN Pitching;														
Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Visible	Expression
Pitching	0	PRIMARY	1	playerID	A	8684	NULL	NULL	BTREE				YES	NULL
Pitching	0	PRIMARY	2	yearID	A	41081	NULL	NULL	BTREE				YES	NULL
Pitching	0	PRIMARY	3	stint	A	44668	NULL	NULL	BTREE				YES	NULL
3 rows in set (0.00 sec)														
id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra			
1	PRIMARY	<derived7>	NULL	ALL	NULL	NULL	NULL	NULL	44707	33.33	Using where			
1	PRIMARY	<derived3>	NULL	ref	<auto_key1>	<auto_key1>	767	t4.playerID	10	33.33	Using where			
8	SUBQUERY	<derived9>	NULL	ALL	NULL	NULL	NULL	NULL	44707	100.00	NULL			
9	DERIVED	Pitching	NULL	index	PRIMARY	PRIMARY	775	NULL	44707	100.00	NULL			
7	DERIVED	Pitching	NULL	index	PRIMARY	PRIMARY	775	NULL	44707	100.00	NULL			
4	SUBQUERY	<derived5>	NULL	ALL	NULL	NULL	NULL	NULL	102421	100.00	NULL			
5	DERIVED	Batting	NULL	index	PRIMARY	PRIMARY	775	NULL	102421	100.00	NULL			
3	DERIVED	Batting	NULL	index	PRIMARY	PRIMARY	775	NULL	102421	100.00	NULL			

With pk and fk:

Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Visible	Expression
Batting	0	PRIMARY	1	playerID	A	21261	NULL	NULL	BTREE				YES	NULL
Batting	0	PRIMARY	2	yearID	A	182276	NULL	NULL	BTREE				YES	NULL
Batting	0	PRIMARY	3	stint	A	182276	NULL	NULL	BTREE				YES	NULL
Batting	1	yearID	1	yearID	A	147	NULL	NULL	BTREE				YES	NULL
Batting	1	yearID	2	lgID	A	317	NULL	YES	BTREE				YES	NULL
Batting	1	yearID	3	teamID	A	2625	NULL	NULL	YES	BTREE			YES	NULL
6 rows in set (0.00 sec)														
mysql> SHOW INDEXES IN Pitching;														
Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Visible	Expression
Pitching	0	PRIMARY	1	playerID	A	8184	NULL	NULL	BTREE				YES	NULL
Pitching	0	PRIMARY	2	yearID	A	44456	NULL	NULL	BTREE				YES	NULL
Pitching	0	PRIMARY	3	stint	A	47436	NULL	NULL	BTREE				YES	NULL
Pitching	1	yearID	1	yearID	A	182	NULL	NULL	BTREE				YES	NULL
Pitching	1	yearID	2	lgID	A	349	NULL	NULL	YES	BTREE			YES	NULL
Pitching	1	yearID	3	teamID	A	2818	NULL	NULL	YES	BTREE			YES	NULL

id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
1	PRIMARY	<derived7>	NULL	ALL	NULL	NULL	NULL	47436	33.33	Using where	
1	PRIMARY	<derived3>	NULL	ALL	NULL	NULL	NULL	102176	3.33	Using where; Using join buffer (hash join)	
8	SUBQUERY	<derived9>	NULL	ALL	NULL	NULL	NULL	47436	100.00	NULL	
9	DERIVED	Pitching	NULL	index	PRIMARY,yearID	PRIMARY	1030	NULL	47436	100.00	NULL
7	DERIVED	Pitching	NULL	index	PRIMARY,yearID	PRIMARY	1030	NULL	47436	100.00	NULL
4	SUBQUERY	<derived5>	NULL	ALL	NULL	NULL	NULL	102176	100.00	NULL	
5	DERIVED	Batting	NULL	index	PRIMARY,yearID	PRIMARY	1030	NULL	102176	100.00	NULL
3	DERIVED	Batting	NULL	index	PRIMARY,yearID	PRIMARY	1030	NULL	102176	100.00	NULL

With index:

mysql> SHOW INDEXES IN Batting;														
Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Visible	Expression
Batting	0	PRIMARY	1	playerID	A	21261	NULL	NULL	BTREE			YES	NULL	
Batting	0	PRIMARY	2	yearID	A	102176	NULL	BTREE				YES	NULL	
Batting	0	PRIMARY	3	stint	A	102176	NULL	BTREE				YES	NULL	
Batting	1	yearID	1	yearID	A	147	NULL	BTREE				YES	NULL	
Batting	1	yearID	2	teamID	A	317	NULL	YES	BTREE			YES	NULL	
Batting	1	yearID	3	teamID	A	2628	NULL	YES	BTREE			YES	NULL	
Batting	1	HR	1	HR	A	49	NULL	BTREE				YES	NULL	
Batting	1	PLAYERID	1	playerID	A	18729	NULL	BTREE				YES	NULL	

mysql> SHOW INDEXES IN Pitching;														
Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Visible	Expression
Pitching	0	PRIMARY	1	playerID	A	8184	NULL	NULL	BTREE			YES	NULL	
Pitching	0	PRIMARY	2	yearID	A	44456	NULL	BTREE				YES	NULL	
Pitching	0	PRIMARY	3	stint	A	47436	NULL	BTREE				YES	NULL	
Pitching	1	yearID	1	yearID	A	152	NULL	BTREE				YES	NULL	
Pitching	1	yearID	2	lgID	A	349	NULL	YES	BTREE			YES	NULL	
Pitching	1	yearID	3	teamID	A	2818	NULL	YES	BTREE			YES	NULL	
Pitching	1	SHO	1	SHO	A	15	NULL	YES	BTREE			YES	NULL	
Pitching	1	PLAYERID	1	playerID	A	9213	NULL	BTREE				YES	NULL	

id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
1	PRIMARY	<derived7>	NULL	ALL	NULL	NULL	NULL	47436	33.33	Using where	
1	PRIMARY	<derived3>	NULL	ALL	NULL	NULL	NULL	102176	3.33	Using where; Using join buffer (hash join)	
8	SUBQUERY	<derived9>	NULL	ALL	NULL	NULL	NULL	47436	100.00	NULL	
9	DERIVED	Pitching	NULL	index	PRIMARY,yearID,SHO,PLAYERID	PRIMARY	1038	NULL	47436	100.00	NULL
7	DERIVED	Pitching	NULL	index	PRIMARY,yearID,SHO,PLAYERID	PRIMARY	1038	NULL	47436	100.00	NULL
4	SUBQUERY	<derived5>	NULL	ALL	NULL	NULL	NULL	102176	100.00	NULL	
5	DERIVED	Batting	NULL	index	PRIMARY,yearID,HR,PLAYERID	PRIMARY	1038	NULL	102176	100.00	NULL
3	DERIVED	Batting	NULL	index	PRIMARY,yearID,HR,PLAYERID	PRIMARY	1038	NULL	102176	100.00	NULL

Conclusion:

For the results in question a and the first part of question b, after we add some important index, the rows to scan decrease a lot!

For the results in question b, we can get the importance of primary key, and it can dramatically improve our performance sometimes, and based on the primary key, we add some index, then get a better performance.

For the results in question c, we can see the importance of foreign key, sometimes when we add foreign key, the mysql also add some important index for us which could improve our performance because our inner join operation and group by operation can be improved.

For the results in question d, e, f, we can see the queries and get that all the inner queries need to scan the table because of the aggregate function sum; thus, there is no performance improvement by adding indexes, primary key, foreign key.