

下面结合实际的例子介绍主键索引和普通索引

```
create table workers
(
    id int(11) not null auto_increment comment '员工工号',
    name varchar(16) not null comment '员工名字',
    sales int(11) default null comment '员工销售业绩',
    primary key (id)
) engine InnoDB
AUTO_INCREMENT = 10
default charset = utf8;

insert into workers(id, name, sales)
values (1, '江南', 12744);
insert into workers(id, name, sales)
values (3, '今何在', 14082);
insert into workers(id, name, sales)
values (7, '路明非', 14738);
insert into workers(id, name, sales)
values (8, '吕归尘', 7087);
insert into workers(id, name, sales)
values (11, '姬野', 8565);
insert into workers(id, name, sales)
values (15, '凯撒', 8501);
insert into workers(id, name, sales)
values (20, '绘梨衣', 7890);
insert into workers(id, name, sales)
values (21, '西泽尔', 16634);
```

我们现在在自己的测试数据库中创建一个包含销售员信息的测试表workers

包含id(主键),name,sales三个字段, 指定表的存储引擎为InnoDB。

然后插入8条数据

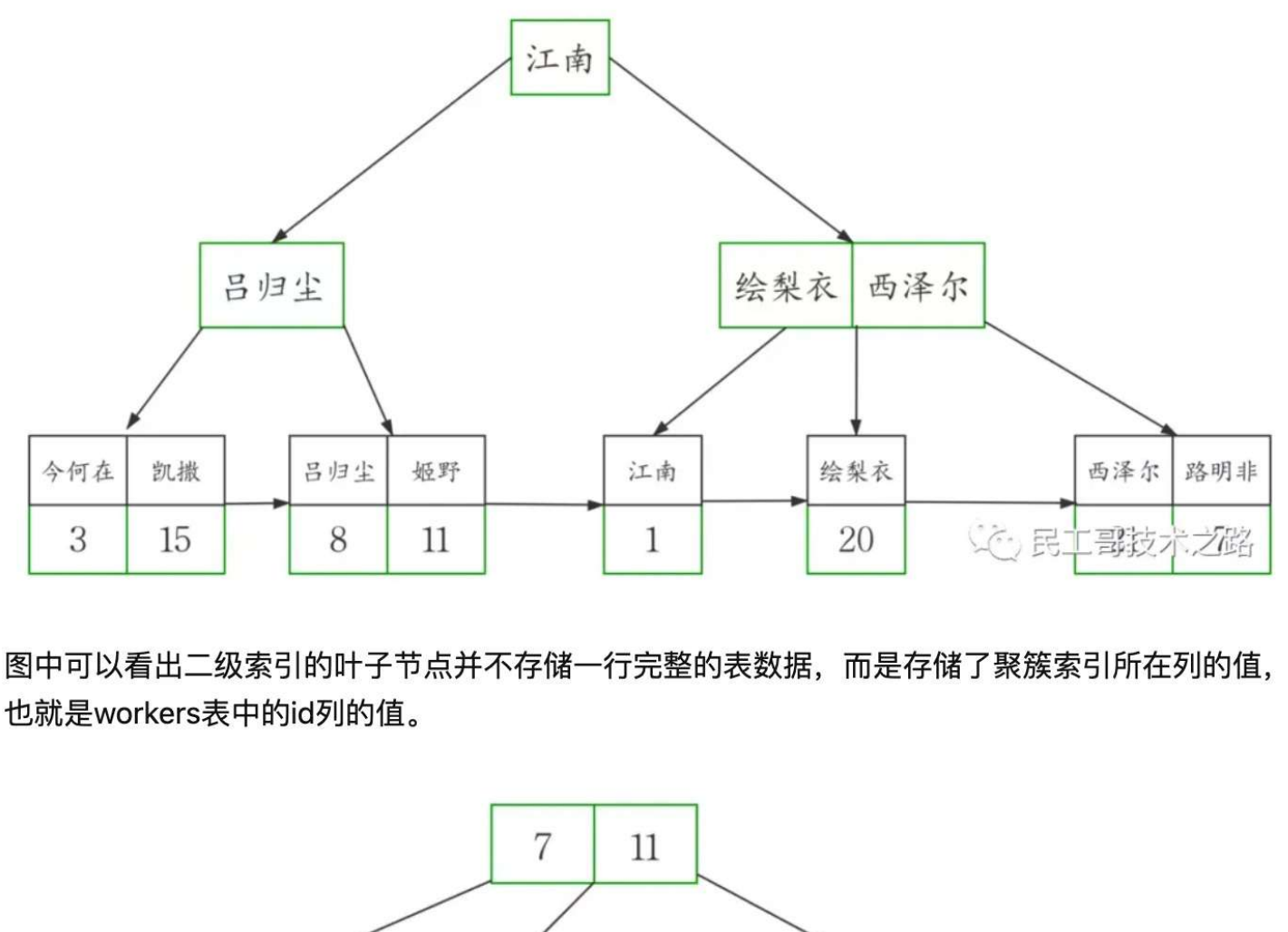
	id	name	sales
1	1	江南	12744
2	3	今何在	14082
3	7	路明非	14738
4	8	吕归尘	7087
5	11	姬野	8565
6	15	凯撒	8501
7	20	绘梨衣	7890
8	21	西泽尔	16634

这个例子当中, workers表的聚簇索引建立在字段id上

为了准确模拟, 我们先把主键id插入b+tree得到下图



然后在此图基础上, 我画出了高清图。



从图中可以看到, 聚簇索引的每个叶子节点存储了一行完整的表数据, 叶子节点间采用单向链表按id列递增连接, 可以方便的进行顺序检索。

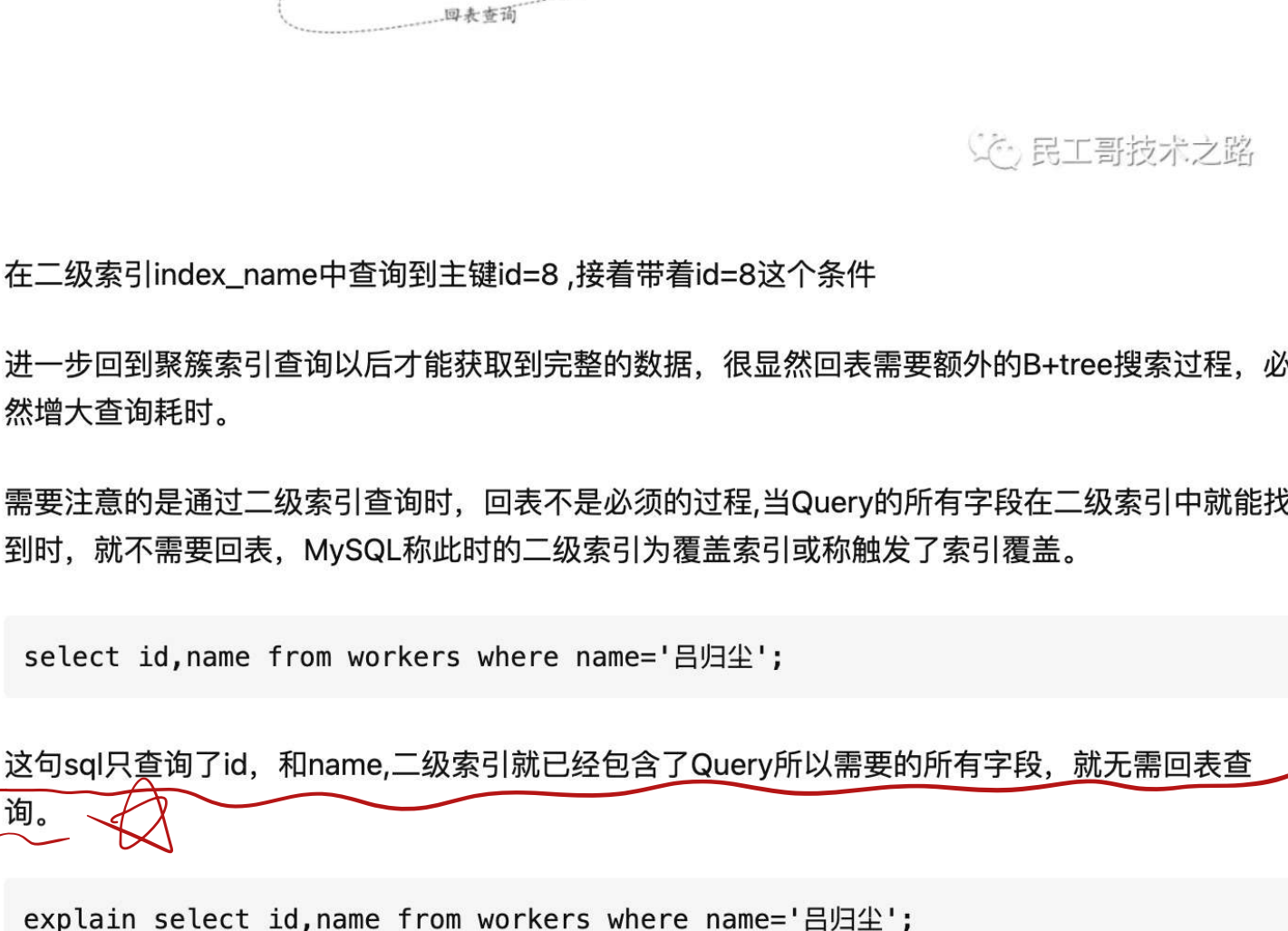
InnoDB表要求必须有聚簇索引, 默认在主键字段上建立聚簇索引, 在没有主键字段的情况下, 表的第一个NOT NULL的唯一索引将被建立为聚簇索引, 在前两者都没有的情况下, InnoDB将自动生成一个隐式自增id列并在此列上创建聚簇索引。

接着来看二级索引

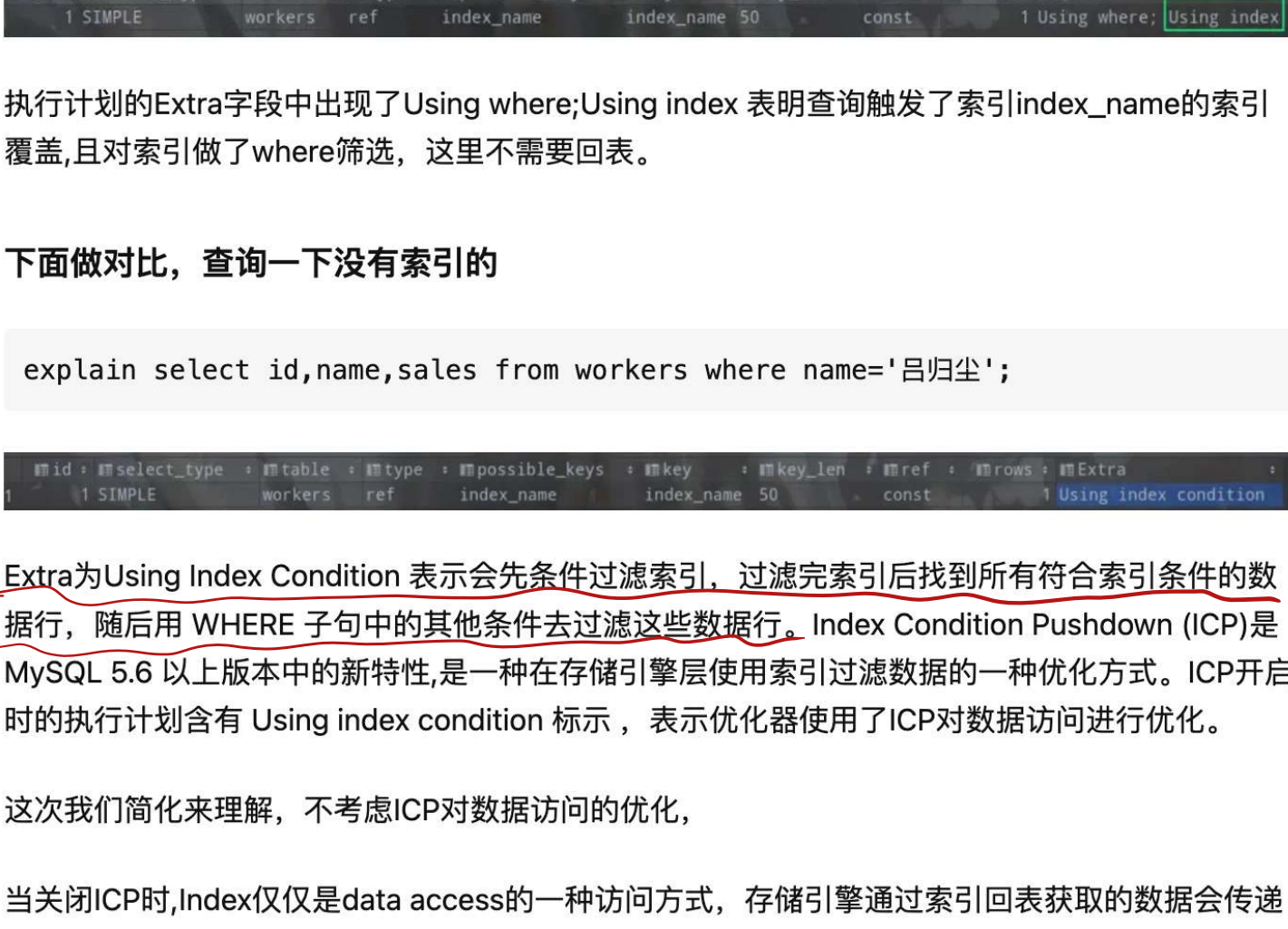
还以刚才的workers表为例

我们在name字段上添加二级索引index_name

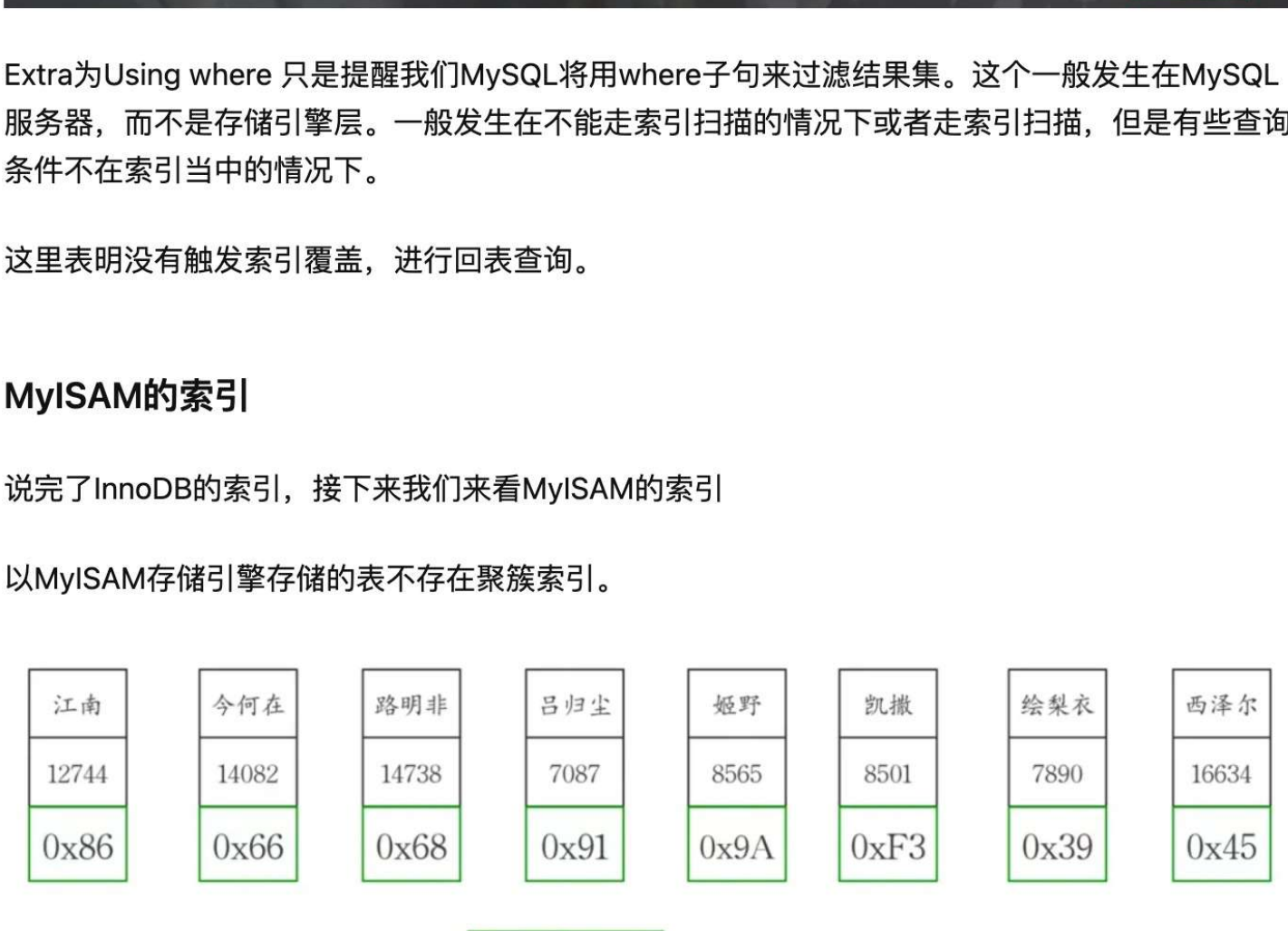
```
alter table workers add index index_name(name);
```



同样我们画出了二级索引index_name的B+tree示意图



图中可以看出二级索引的叶子节点并不存储一行完整的表数据, 而是存储了聚簇索引所在列的值, 也就是workers表中的id列的值。



这两张示意图中B+tree的度设置为3, 这也主要是为了方便演示。

实际的B+tree索引中, 树的度通常会大于100。

说了聚簇索引和二级索引 肯定要提到回表查询。

由于二级索引的叶子节点不存储完整的表数据, 所以当通过二级索引查询到聚簇索引的列值后, 还需要回到局促索引也就是表数据本身进一步获取数据。

比如说我们要在workers表中查询 名叫吕归尘的人

```
select * from workers where name='吕归尘';
```

这条sql通过name='吕归尘'的条件



在二级索引index_name中查询到主键id=8, 接着带着id=8这个条件

进一步回到聚簇索引查询以后才能获取到完整的数据, 很显然回表需要额外的B+tree搜索过程, 必然增大查询耗时。

需要注意的是通过二级索引查询时, 回表不是必须的过程, 当Query的所有字段在二级索引中就能找到时, 就不需要回表, MySQL称此时的二级索引为覆盖索引或称触发了索引覆盖。

```
select id,name from workers where name='吕归尘';
```

这句sql只查询了id, 和name, 二级索引就已经包含了Query所以需要的所有字段, 就无需回表查询。

```
explain select id,name from workers where name='吕归尘';
```

使用explain查看此条sql的执行计划

```
id | select_type | table | type | possible_keys | key | key_len | ref | rows | extra
1 | SIMPLE | workers | ref | index_name | index_name | 50 | const | 1 | Using where; Using index;
```

执行计划的Extra字段中出现了Using where; Using index 表明查询触发了索引index_name的索引覆盖, 且对索引做了where筛选, 这里不需要回表。

下面做对比, 查询一下没有索引的

```
explain select id,name,sales from workers where name='吕归尘';
```

```
id | select_type | table | type | possible_keys | key | key_len | ref | rows | extra
1 | SIMPLE | workers | ref | index_name | index_name | 50 | const | 1 | Using where;
```

Extra为Using Index Condition 表示会先条件过滤索引, 过滤完索引后找到所有符合索引条件的数据行, 随后用 WHERE 子句中的其他条件去过滤这些数据行。Index Condition Pushdown (ICP)是MySQL 5.6 以上版本中的新特性, 是一种在存储引擎层使用索引过滤数据的一种优化方式。ICP开启时的执行计划含有 Using index condition 标示, 表示优化器使用了ICP对数据访问进行优化。

这次我们简化来理解, 不考虑ICP对数据访问的优化,

当关闭ICP时, index仅仅是data access的一种访问方式, 存储引擎通过索引回表获取的数据会传递到MySQL Server 层进行WHERE条件过滤。

Extra为Using where 只是提醒我们MySQL将用where子句来过滤结果集。这个一般发生在MySQL服务器, 而不是存储引擎层。一般发生在不能走索引扫描的情况下或者走索引扫描, 但是有些查询条件不在索引当中的情况下。

这里表明没有触发索引覆盖, 进行回表查询。

MyISAM的索引

说完了InnoDB的索引, 接下来我们来看MyISAM的索引

以MyISAM存储引擎存储的表不存在聚簇索引。

MyISAM索引B+tree示意图

MyISAM表中的主键索引和非主键索引的结构是一样的, 从上图我们可以看到

他们的叶子节点是不存储表数据的, 节点中存放的是表数据的地址, 所以MyISAM表可以没有主键。

MyISAM表的数据和索引是分开的, 是单独存放的。

MyISAM表中的主键索引和非主键索引的区别仅在于主键索引B+tree上的key必须符合主键的限制,

非主键索引B+tree上的key只要符合相应字段的特性就可以了。

索引字段特性角度看索引

主键索引

- 建立在主键字段上的索引
- 一张表最多只有一个主键索引
- 索引列值不允许为null
- 通常在创建表的时候一起创建

唯一索引

- 建立在UNIQUE字段上的索引就是唯一索引
- 一张表可以有多个唯一索引, 索引列值允许为null

我们演示创建索引

```
create table persons
(
    id int(11) not null auto_increment comment '主键id',
    eno int(11) comment '工号',
    eid int(11) comment '身份证号',
    veid int(11) comment '虚拟身份证号',
    name varchar(16) comment '名字',
    primary key (id) comment '主键索引',
    UNIQUE key (eno) comment 'eno唯一索引',
    UNIQUE key (eid) comment 'eid唯一索引'
) engine = InnoDB
auto_increment = 1000
default charset = utf8;

alter table persons
add unique index index_veid (veid) comment 'veid唯一索引';
```

通过show index from persons;命令我们看到已经成功创建了三个唯一索引。

普通索引

主键索引和唯一索引对字段的要求是要求字段为主键或unique字段,

而那些建立在普通字段上的索引叫做普通索引, 既不要求字段为主键也不要求字段为unique。

前缀索引

前缀索引是指对字符类型字段的前几个字符或对二进制类型字段的前几个bytes建立的索引, 而不是在整个字段上建索引。

例如, 可以对persons表中的name(varchar(16))字段 中name的前5个字符建立索引。

```
create index index_name on persons (name(5)) comment '前缀索引';
show index from persons;
```


前缀索引可以建立在类型为

- char
- varchar
- binary
- varbinary

的列上, 可以大大减少索引占用的存储空间, 也能提升索引的查询效率。

索引列的个数角度看索引

- 建立在单个列上的索引为单列索引
 - 上文演示的都是单列索引
- 建立在多列上的称为联合索引 (复合索引)

演示一下联合索引

```
create index index_id_name on workers(id,name) comment '组合索引';
```

这条语句在我们演示表workers中建立id, name这两个字段的联合索引。

借助show index命令查看索引的详细信息 操作后结果如下:

虽然详细信息当中列出了两条关于联合索引的条目, 但并不表示联合索引是建立了多个索引, 联合索引是一个索引结构, 这两个条目表示的是组合索引中字段的详细信息, 按建立索引时的书写顺序排列。

同样我们来看下联合索引的B+tree示意图

从图中看到

组合索引的非叶子节点保存了两个字段的值作为B+tree的key值, 当B+tree上插入数据时, 先按字段id比较, 在id相同的情况下按name字段比较。