

Báo cáo Talent 5
Tungdinh – 03.2021

1. Báo cáo T0 + 3:

1. Backup data

```
SELECT * INTO du_lieu_ban_hang_backup  
FROM du_lieu_ban_hang
```

2. Làm sạch dữ liệu

1. Giá trị null: khu_vuc, thoi_gian_trung_binh và so_cuoc_goi

```
DELETE FROM du_lieu_ban_hang_backup WHERE thoi_gian_trung_binh IS NULL
```

```
[75] 1 delete from du_lieu_ban_hang_backup where thoi_gian_trung_binh is null
```

(3 rows affected)

Total execution time: 00:00:01.167

```
DELETE FROM du_lieu_ban_hang WHERE khu_vuc IS NULL
```

```
1 delete from du_lieu_ban_hang_backup where khu_vuc is null
```

(1 row affected)

Total execution time: 00:00:00.173

```
DELETE FROM du_lieu_ban_hang_backup WHERE so_cuoc_goi IS NULL
```

2. Dữ liệu trùng:

```
SELECT ngay_ban, san_pham, khu_vuc, so_cuoc_goi, thoi_gian_trung_binh,  
so_luong_ban, camera_truoc, ram, the_nho, dung_luong_pin, man_hinh_rong, don_gia,  
COUNT(*)  
FROM du_lieu_ban_hang_backup  
GROUP BY ngay_ban, san_pham, khu_vuc, so_cuoc_goi, thoi_gian_trung_binh,  
so_luong_ban, camera_truoc, ram, the_nho, dung_luong_pin, man_hinh_rong, don_gia  
HAVING COUNT(*) > 1
```

```
1 select ngay_ban, san_pham, khu_vuc, so_cuoc_goi, thoi_gian_trung_binh, so_luong_ban, camera  
2 from du_lieu_ban_hang_backup  
3 group by ngay_ban, san_pham, khu_vuc, so_cuoc_goi, thoi_gian_trung_binh, so_luong_ban, came  
4 having count(*) > 1
```

(1 row affected)

Total execution time: 00:00:00.661

🔍 📄 📊 📈 📉

	ngay_ban	san_pham	khu_vuc	so_cuoc_goi	thoi_gian_trung_binh	so_luong_ban	ca
1	2018-06-13	BlackBerry KEY2	Bắc Ninh	154	NULL	2126	8

3. Dữ liệu bất thường: don_gia của san_pham VIVO V19 Neo: thay bằng giá trị median 649000

```
UPDATE du_lieu_ban_hang_backup SET don_gia = 6490000 WHERE san_pham = 'VIVO V19 Neo' AND don_gia = 6500000
```

```
[90] 1 update du_lieu_ban_hang_backup set don_gia = 6490000 where san_pham = 'VIVO V19 Neo'
      2 and don_gia = 6500000
```

(1 row affected)

Total execution time: 00:00:00.170

4. Typo: xảy ra với “bắc ninh” trong khu_vuc: đổi thành “Bắc Ninh”

```
UPDATE du_lieu_ban_hang_backup SET khu_vuc = N'Bắc Ninh' WHERE khu_vuc = N'bắc ninh' COLLATE sql_latin1_general_cp1_cs_as
```

```
[127] 1 set khu_vuc = N'Bắc Ninh' where khu_vuc = N'bắc ninh' collate sql_latin1_general_cp1_cs_as
```

(1 row affected)

Total execution time: 00:00:00.279

5. Datatype

3. Thiết kế