**HO CHI MINH UNIVERSITY OF SCIENCE**

**FACULTY OF INFORMATION**

**COURSE: DATA VISUALIZATION**



**Student: Trần Hoàng Khải - 21127620**

**Từ Phước Toàn - 21127705**

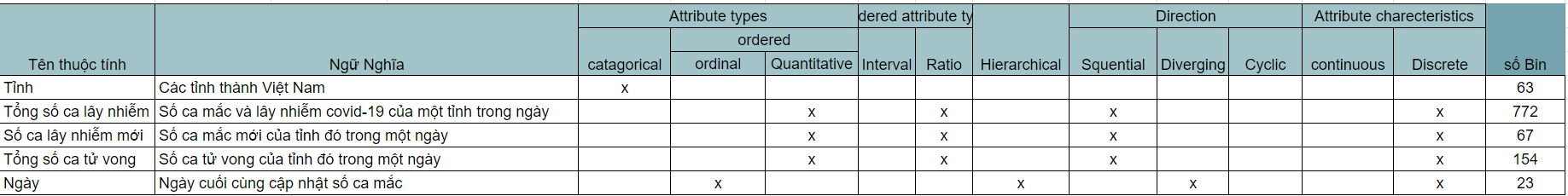
**Ng. Ngọc Diệu Trang - 20127361**

**Trần Đình Thiện - 20127632**

**Class: 21HTTT**

**Teacher: Tiết Gia Hồng**

1. **Data abstraction**

* Dataset: Thống kê số ca nhiễm, số ca lây nhiễm mới và tổng số ca tử vong COVID của các tỉnh thành mỗi ngày
* Loại dataset: Static
* Dataset type: Table
* Item: mỗi dòng là thông tin về tổng số ca nhiễm, số ca lây nhiễm mới, tổng số ca tử vong COVID của mỗi tỉnh thành trong một ngày

1. **Task abstraction**

* Câu Khai Thác

1. So sánh số ca tử vong theo ngày của 2 tỉnh thành Bình Dương và Thành Phố Hồ Chí Minh

Action: Query -> compare

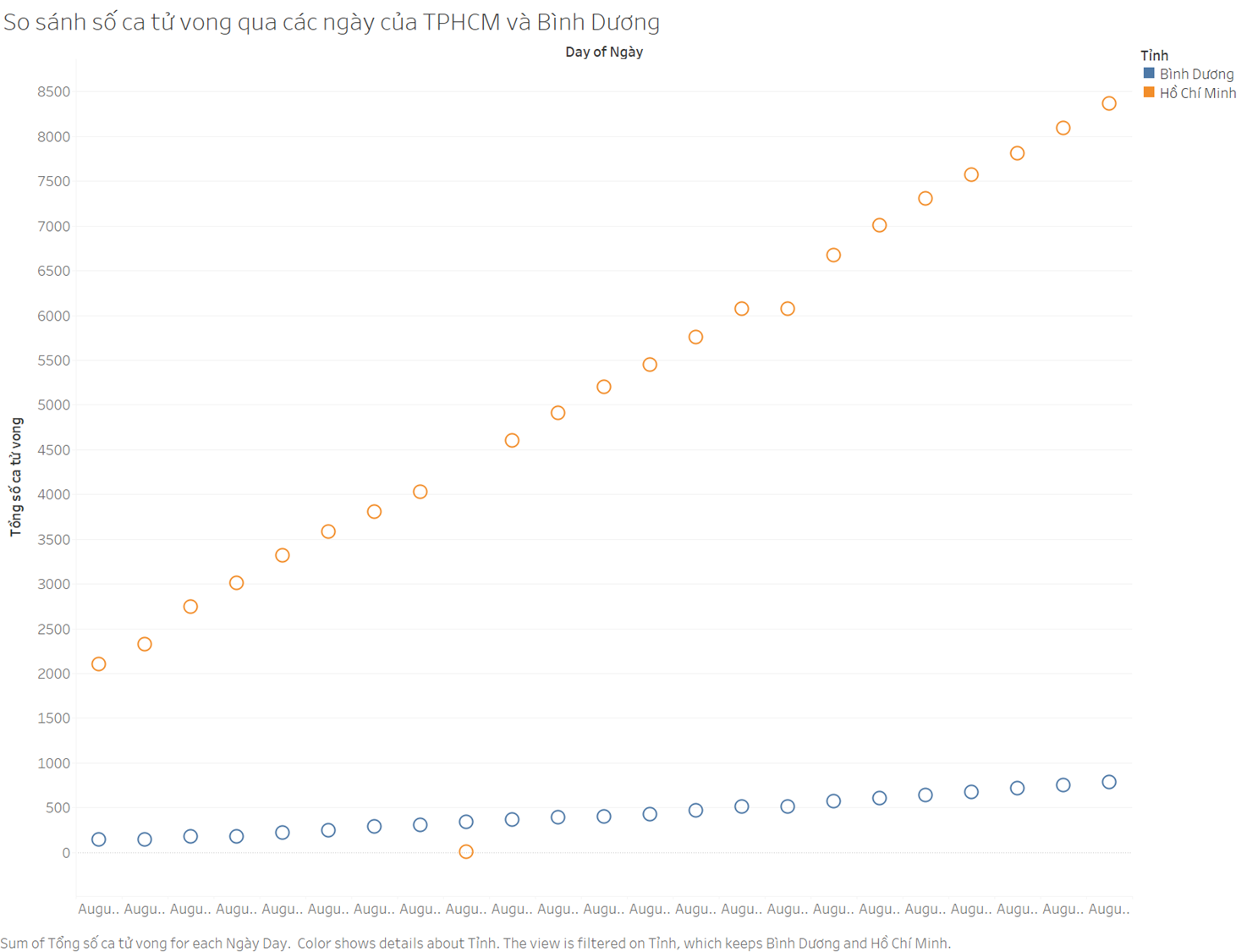
Target: Attributes -> many -> Similarity

1. Thống kê số ca nhiễm mới và tổng số ca lây nhiễm theo từng ngày của tháng 8/2021 của thành phố HCM

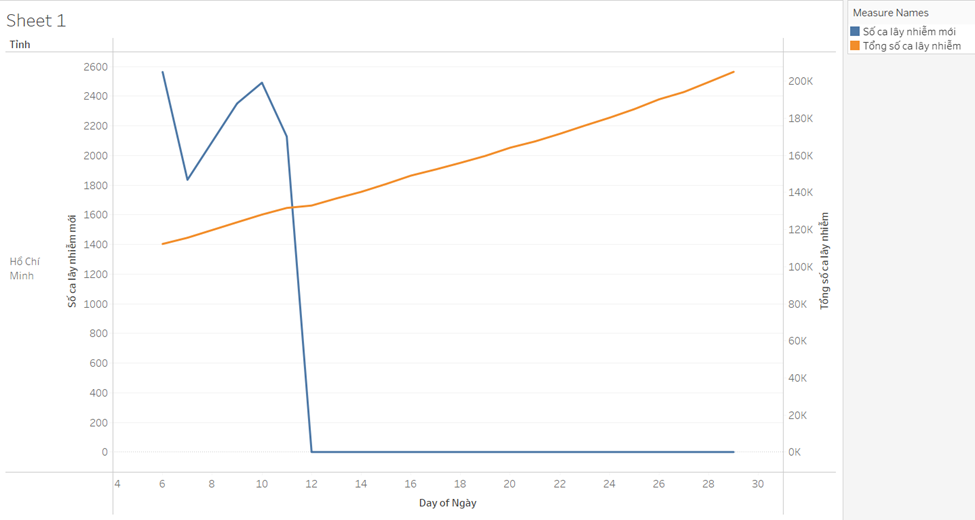
Action: Query -> compare

Target: Attributes->one -> extremes

1. **Tableau**
2. So sánh số ca tử vong theo ngày của 2 tỉnh thành Bình Dương và Thành Phố Hồ Chí Minh



1. Thống kê số ca nhiễm mới và tổng số ca lây nhiễm theo từng ngày của tháng 8/2021 của thành phố HCM.



1. **Đánh Giá**

* Biểu đồ 1:
* + Nguyên tắc biểu đạt: Biểu đồ trên biểu diễn được hết tất cả các dữ liệu và biểu diễn đầy các data có trong dataset bao gồm số ca tử vong theo ngày của tỉnh Bình Dương và TP. HCM
* + Hiệu quả: Độ chính xác, sai số không quá lớn, vẫn nhìn rõ được sự biến thiên của số ca tử vong theo từng ngày của 2 thành phố.

+ khả năng phân biệt: Dùng scater plot nên số ca tử vong của mỗi ngày được trong thành phố đó bằng những hình tròn

+ khả năng tách biệt: dùng 2 kênh màu cam và xanh nước biển để phân biệt 2 thành phố Bình Dương và TP.HCM

* Biểu đồ 2:
* + Nguyên tắc biểu đạt: Biểu đồ trên biểu diễn được hết tất cả các dữ liệu và biểu diễn đầy các data có trong dataset bao gồm số ca tử vong theo ngày của tỉnh Bình Dương và TP. HCM
* + Hiệu quả: Độ chính xác không cao, sai số đáng kể trong tháng 8/2021, vẫn nhìn rõ được sự biến thiên của số ca tử vong theo từng ngày của 2 thành phố.

+ khả năng phân biệt: Dùng line chart để phân biệt cho thuộc tính rời rạc nên không tạo được sự phân biệt rõ ràng trong một thuộc tính

+ khả năng tách biệt: dùng 2 kênh màu cam và xanh nước biển để phân biệt số ca nhiễm mới và tổng số ca lây nhiễm theo từng ngày của tháng 8/2021 của thành phố HCM.