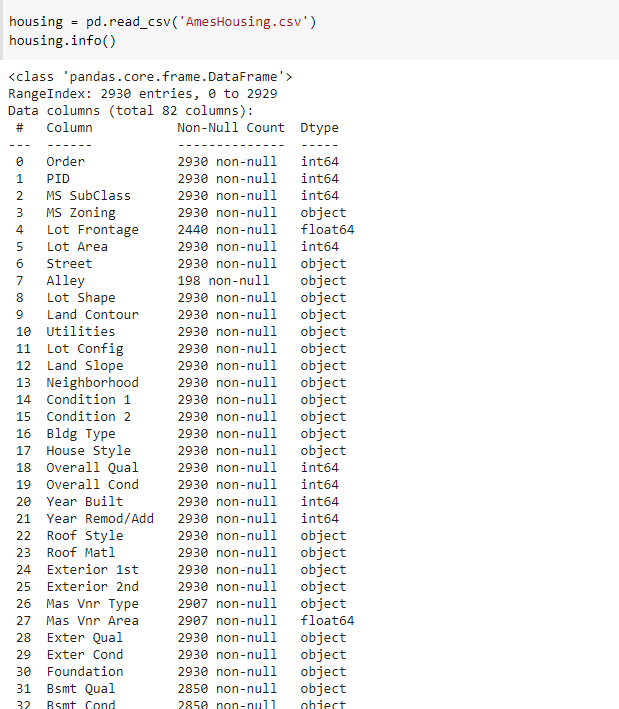
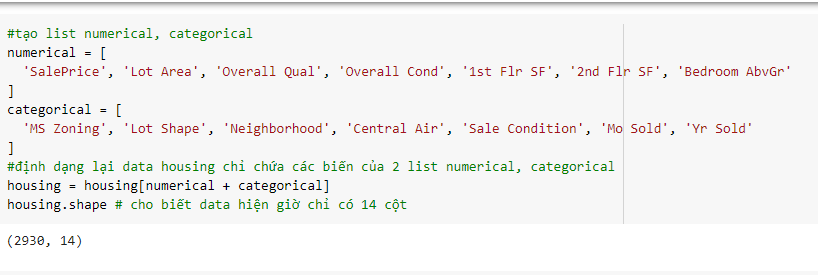
Khai báo thư viện



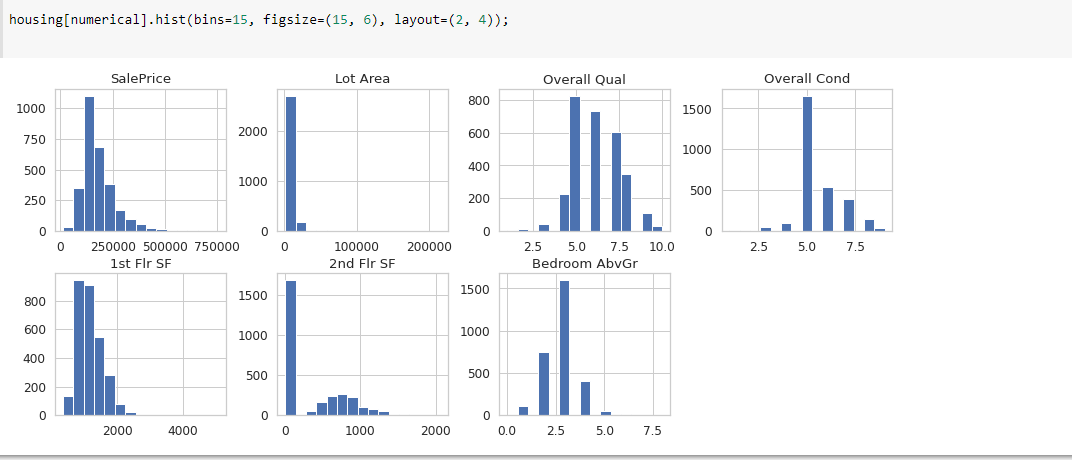
Chuẩn bị dữ liệu, xem thông tin dữ liệu





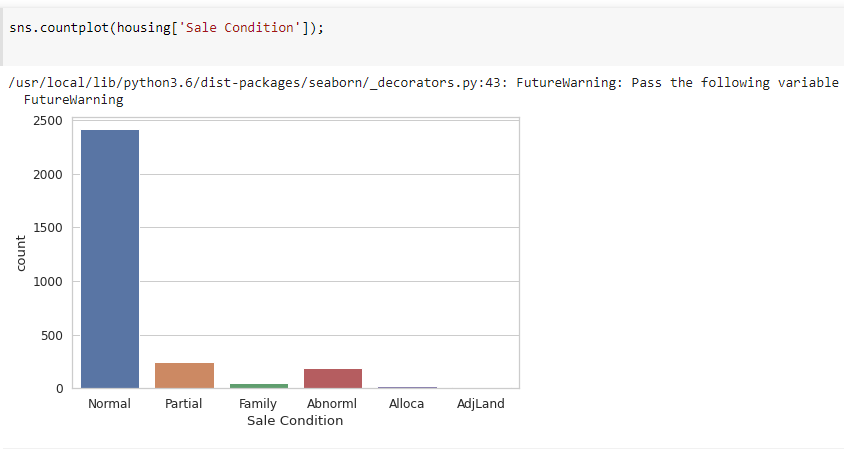


Nhận xét: Từ biểu đồ, ta thấy rất ít ngôi nhà có giá dưới 100000, hầu hết những ngôi nhà có già từ 100000 đến 200000, và rất ít ngôi nhà được bán với giá trên 400000

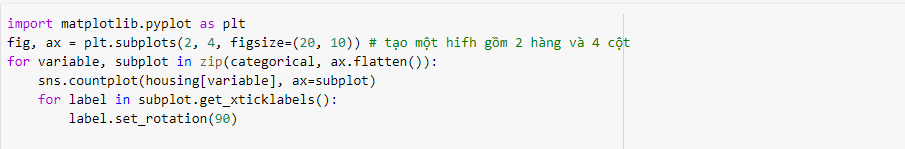


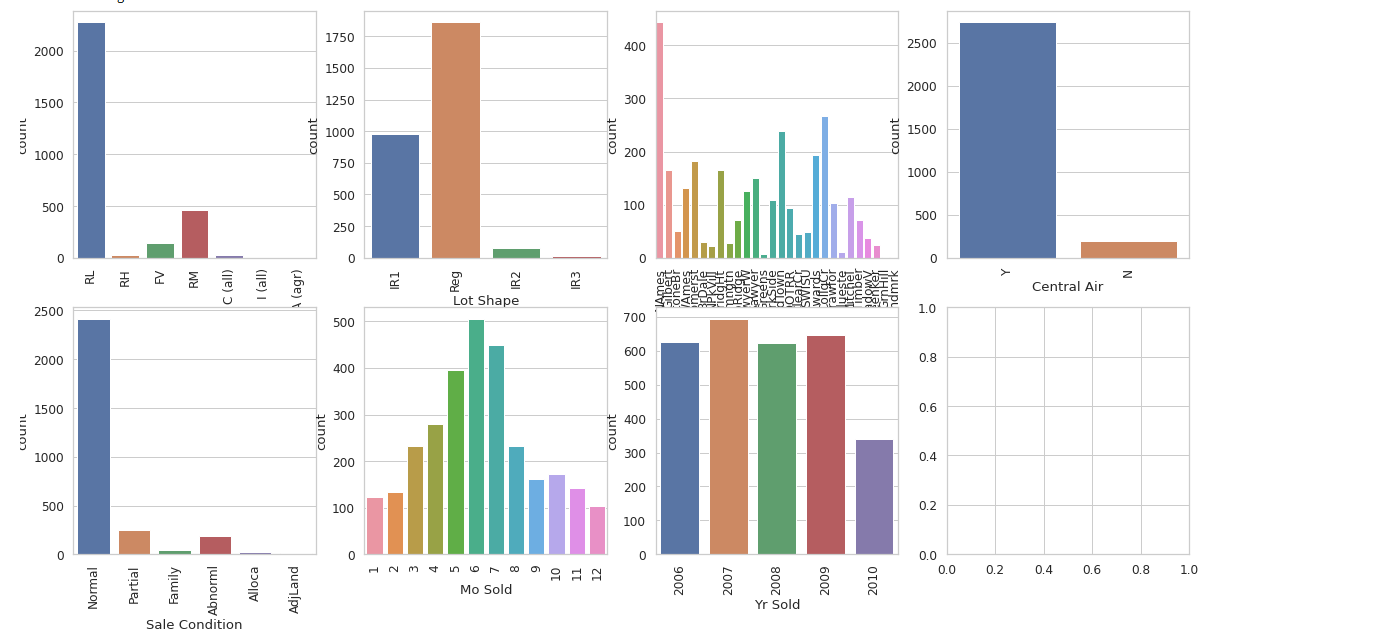
Nhận xét:

* diện tích của tầng một(1st flr SF) bị lệch rất nhiểu sang phải
* hầu hết các ngôi nhà không có tầng 2(2nd flr SF)
* hầu hết có 3 phòng ngủ (Bedroom abvGr)
* hầu hết các ngồi nhà đã được bán với overall qual là 5 và overall cond là 5 hoặc cao hơn



Nhận xét: hầu hết các ngôi nhà được bán trong tình trạng bình thường, rất ít được bán trong tình trang adjland(mua đất liền kề), alloca(2 bất động sản liên kết với nhau bằng chứng từ riêng) và family(bán giữa các thành viên trong gia đình)

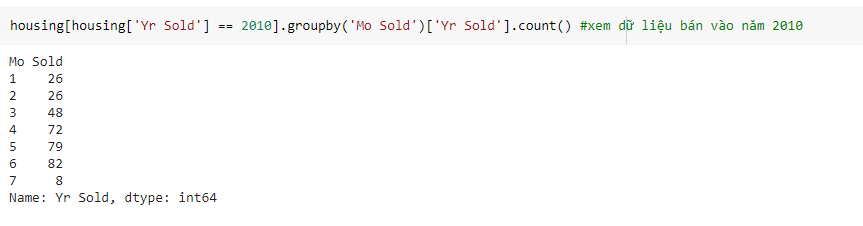


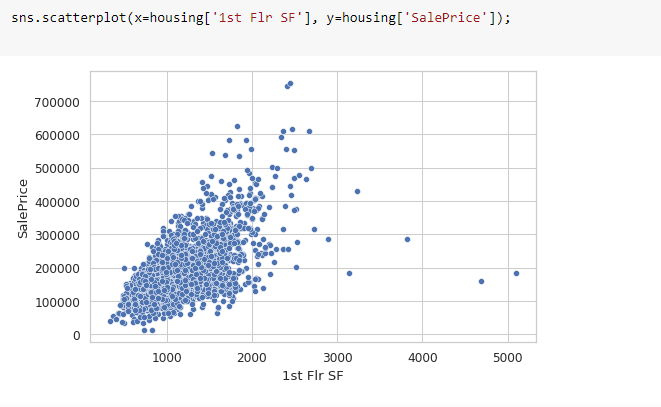


Nhận xét:

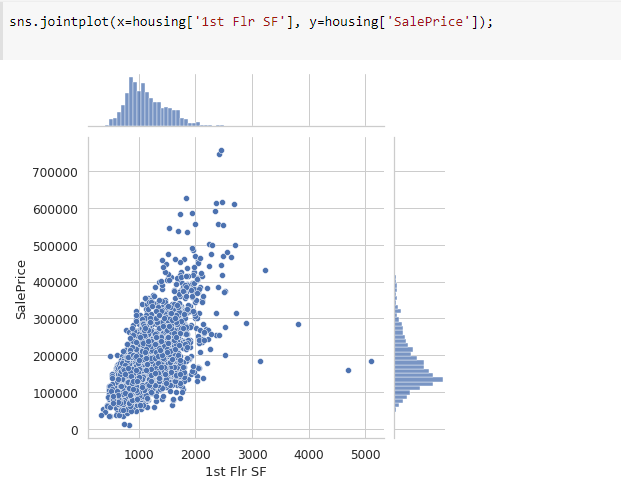
Hầu hết các ngôi nhà đều có phân loại phân vùng RL, có hình dạng lô đất thông thường và có CentralAir, những ngôi nhà thường được bán vào mùa hè(MoSold), được bán nhiều nhất ở khu vực NAmes(), và có sự sụt giảm số lượng bán vào năm 2010(Yr Sold)

Xem dữ iệu bán vào năm 2010

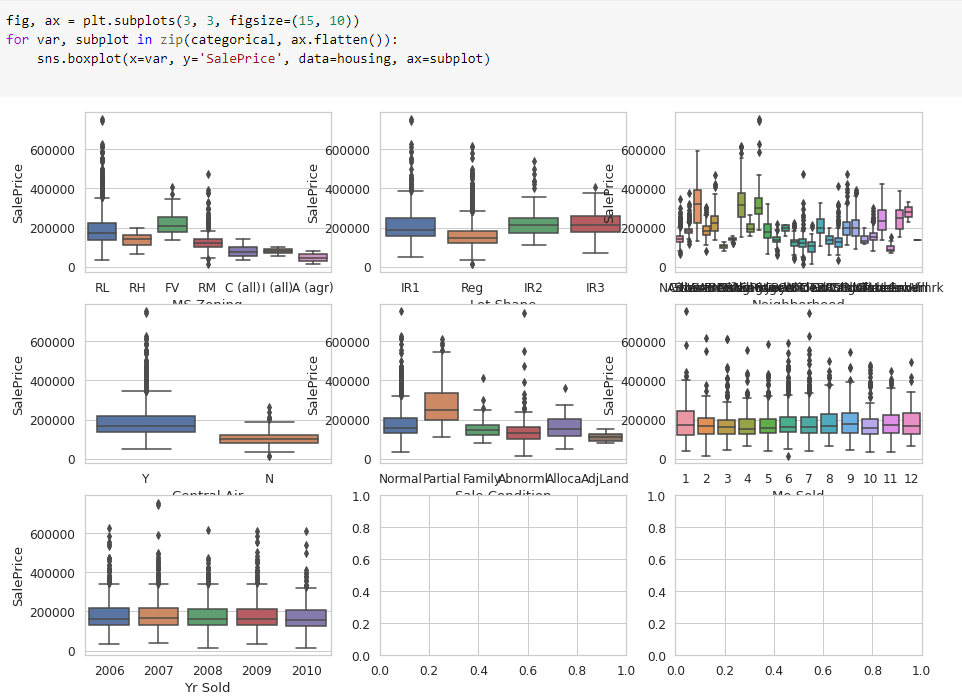




Nhận xét: từ biểu đồ phân tán , có thể thấy mối quan hệ cùng chiều giữa 1st Flr SF và saleprice. Giá trị tầng 1 của ngôi nhà càng lớn thì giá bán càng cao

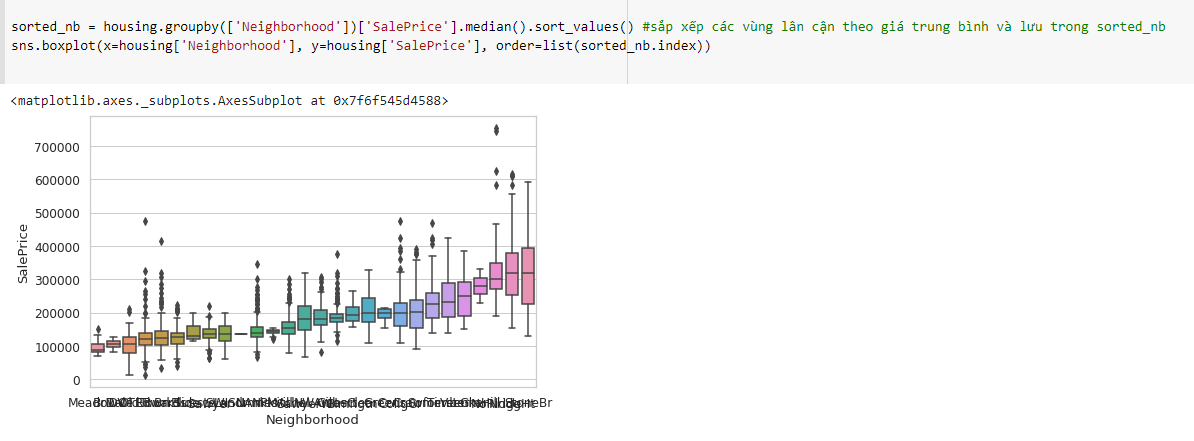


Nhận xét: biết được mối quan hệ giữa 2 biến, biết cách của từng biến được phân phối riêng lẻ



Nhận xét:

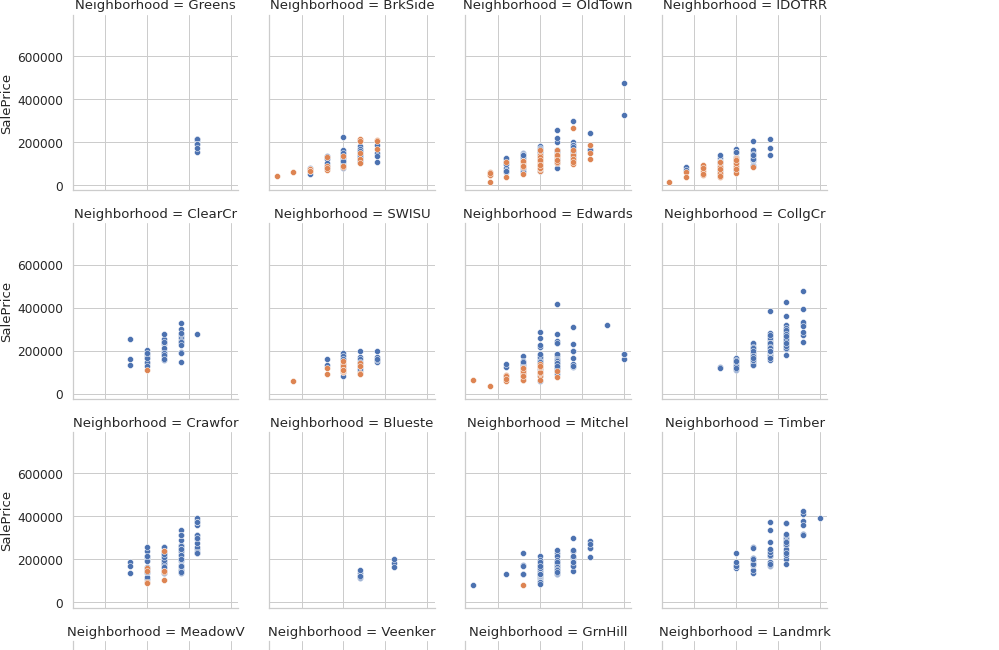
* Nhà ở phân khu FV có giá bán trung bình cao hơn các phân khu khác
* Nhà có central Air, partical có điều kiện bán một phần(nhà chưa hoàn thành khi được đánh giá lần cuối)
* Có ít sự khác biệt về giá bán giữa các ngôi nhà có lotShapes, hoặc giữa MoSold và YrSold



Nhận xét: ở những con phố rẻ nhất, nhà đc bán với giá trung bình khoảng 100000, ở những khu phố đắt nhất, nhà đc bán với giá khoảng 300000. đối với một số vùng lân cận, sự phân tán giữa các mức giá là rất thấp, nghĩa là tất cả các mức giá đều gần nhau. trong khu phố đắt đỏ nhất NridgHt, có sự phân tán lớn trong việc phân bổ giá

Sns.FacetGrid biểu diễn sự phân bố của một biến hoặc mối quan hệ giữa nhiều biến riêng biệt trong các tập con của bộ dữ liệu





Nhận xét:

- Đối với mỗi vùng lân cận, chúng ta có thể thấy mối quan hệ giữa Tổng thể và Giá bán.

Chúng tôi cũng đã thêm một biến phân loại khác CentralAir vào tham số màu (tùy chọn) - các điểm màu cam tương ứng với các ngôi nhà không có CentralAir. Như bạn thấy, những ngôi nhà này có xu hướng bán với giá thấp hơn.