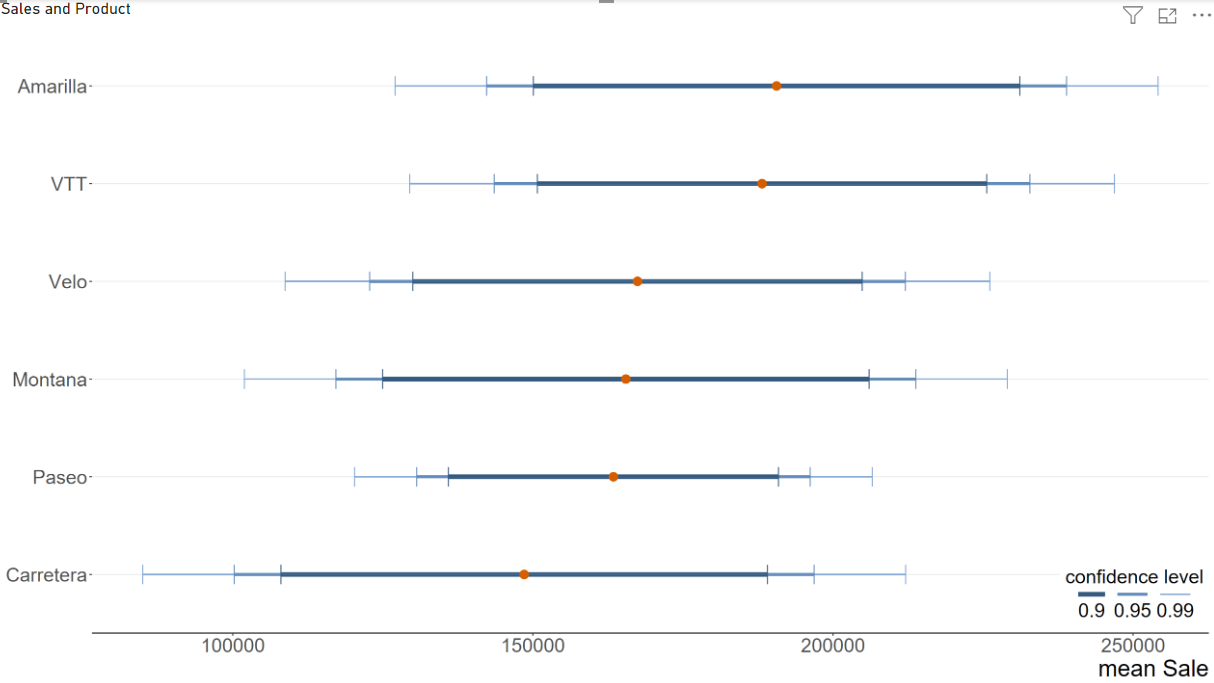
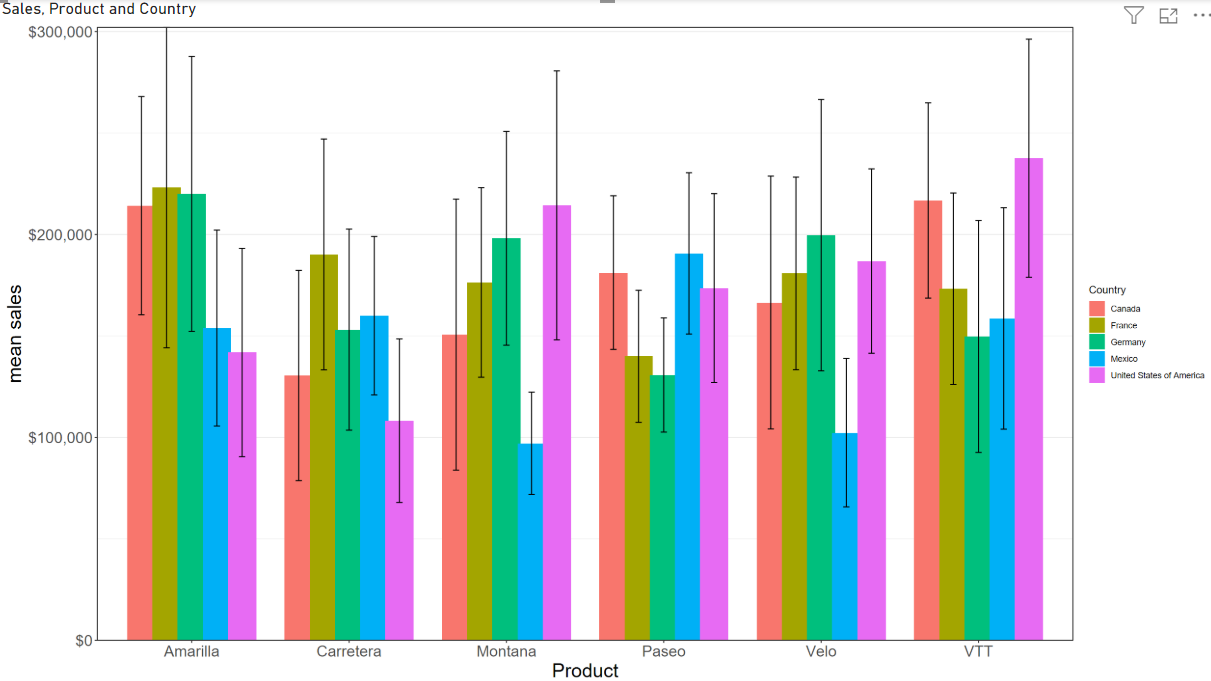
**BÀI TẬP THỰC HÀNH 6**

**Bài tập 1:** Tiếp tục ví dụ 2 trong bài Thực Hành 7 (dùng **Sample** data của Power BI, vẽ biểu đồ khoảng tin cậy cho trung bình Sales theo từng mức của Product).

1. Hãy vẽ các khoảng tin cậy **99%, 95%, và 90%** chồng lên nhau để được đồ thị như sau:



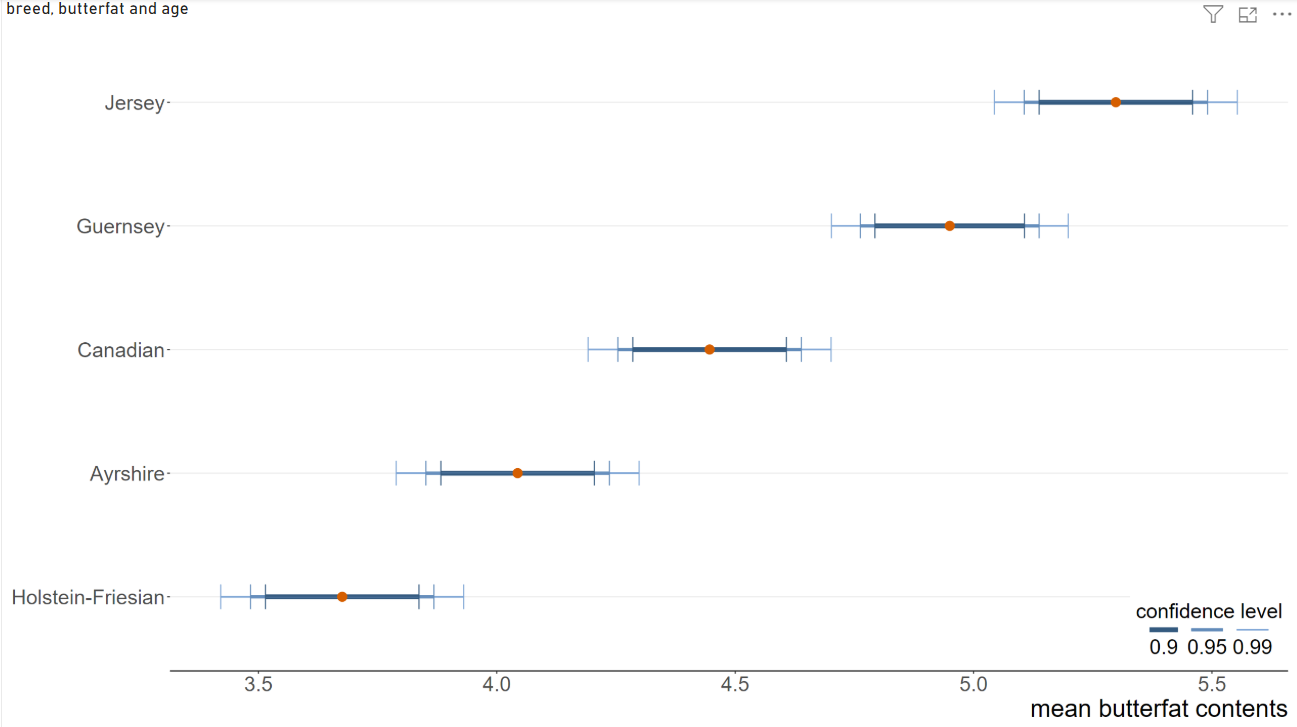
1. Vẽ biểu đồ cột nhóm kèm error bars cho trung bình **Sales** theo **Product và Country.** Với KTC là **[mean – 1\*se; mean + 1\*se]** để được đồ thị như sau:



1. Thực hiện biểu đồ như ở câu 2, nhưng đối với tỷ lệ Product theo Country.
2. Sử dụng kết hợp hàm binom.confint() trong R package binom, với hàm geom\_errorparh() hoặc geom\_errorbar() để trực quan ước lượng tỷ lệ và khoảng tin cậy có hiệu chỉnh.

**Bài tập 2:** Load data **cows.csv** và thực hiện các yêu cầu sau:

1. Vẽ biểu đồ cột kèm error bars cho hàm lượng bơ sữa trung bình (**butterfat**) theo từng giống bò **(breed).** Với KTC là **[mean – 2\*se; mean + 2\*se]**
2. Vẽ biểu đồ khoảng tin cậy 99%, 95% và 90% chồng lên nhau cho hàm lượng bơ sữa trung bình (**buterrfat)** theo từng giống bò **(breed).**



**Bài tập 3:** Load file **economics\_data.csv** và thực hiện các yêu cầu sau:

1. Từ bảng dữ liệu **economics\_data** hãy vẽ đồ thị đường biểu diễn số người thất nghiệp **(Unemploy)** theo **năm (date),** và vẽ đường xu hướng kèm vùng tin cậy theo LOESS và Splines.(cubic spline, B spline) với k lần lượt là 3, 5, 10, 30.

**Bài tập 4**: Dữ liệu china\_gdp.csv từ <https://www.kaggle.com/code/ibrahimbahbah/non-linear-regression-tutorial/notebook>

1. Dùng Scatter plot để biểu diễn dữ liệu.
2. Tìm ra đường biểu diễn phù hợp với biểu đồ ở câu 1.
3. Vẽ dải tin cậy 95% được xác định theo phương pháp bootstrap percentiles.