

# BÁO CÁO THỰC HÀNH BUỔI 9

## Nhóm 6

[COMPANY NAME] [Company address]

# Thành viên

- **Vũ Bạch Long 21139141**

**Fig1:**



**BIỂU ĐỒ THỂ HIỆN TIPS TRÊN TỔNG SỐ BILL GIỮA CÁC KHÁCH HÀNG HÚT THUỐC VÀ KHÁCH HÀNG KHÔNG HÚT THUỐC**

	total_bill	tip	sex	smoker	day	time	size
0	16.99	1.01	Female	No	Sun	Dinner	2
1	10.34	1.66	Male	No	Sun	Dinner	3
2	21.01	3.50	Male	No	Sun	Dinner	3
3	23.68	3.31	Male	No	Sun	Dinner	2
4	24.59	3.61	Female	No	Sun	Dinner	4

Dữ liệu bao gồm:

Total bill, tip, sex, smoker, day, time, size

Nhận xét:

Vào giờ trưa lượng khách hàng ít hơn vào buổi tối

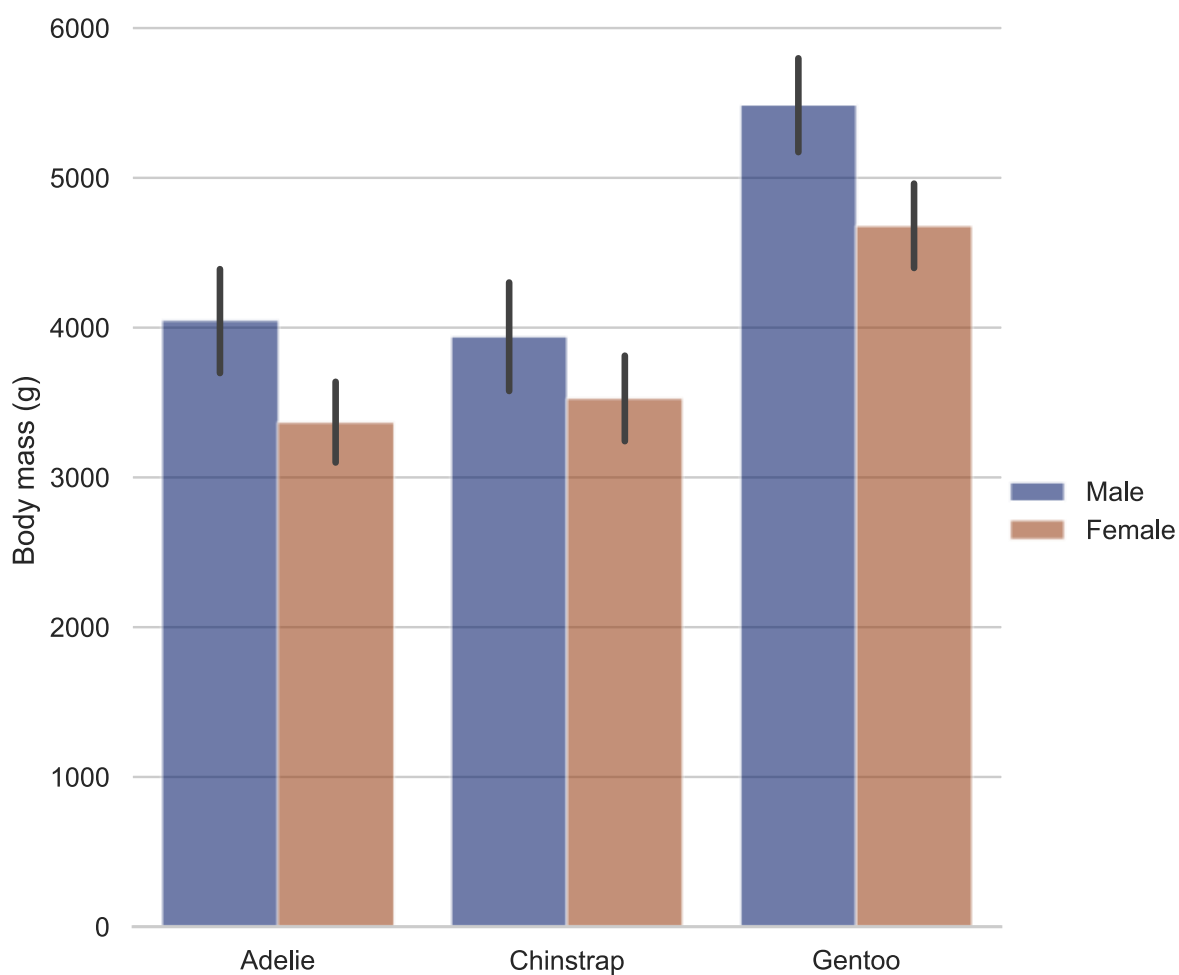
Tổng số bill thấp nhất khoảng 8, cao nhất khoảng 52

Tips thấp nhất khoảng 1, cao nhất khoảng 7

Các khách hàng hút thuốc có tổng số bill cao hơn các khách hàng không hút thuốc

Các khách hàng không hút thuốc trả tiền tips nhiều hơn các khách hàng hút thuốc

**Fig2:**



**BIỂU ĐỒ THỂ HIỆN TỈ LỆ ĐỰC CÁI VÀ CHÊNH LỆCH CÂN NẶNG  
GIỮA CÁC GIỐNG CHIM CÁNH CỤT**

```

In [10]: penguins = penguins[penguins['species'] != 'Gentoo']
In [11]: penguins
Out[11]:
  species  island  bill_length_mm  bill_depth_mm  flipper_length_mm  body_mass_g  sex
0  Adelie  Torgersen           39.1           18.7           181.0        3750.0  Male
1  Adelie  Torgersen           39.5           17.4           186.0        3800.0  Female
2  Adelie  Torgersen           40.3           18.0           195.0        3250.0  Female
3  Adelie  Torgersen           NaN            NaN            NaN            NaN    NaN
4  Adelie  Torgersen           36.7           19.3           193.0        3450.0  Female

```

Dữ liệu bao gồm:

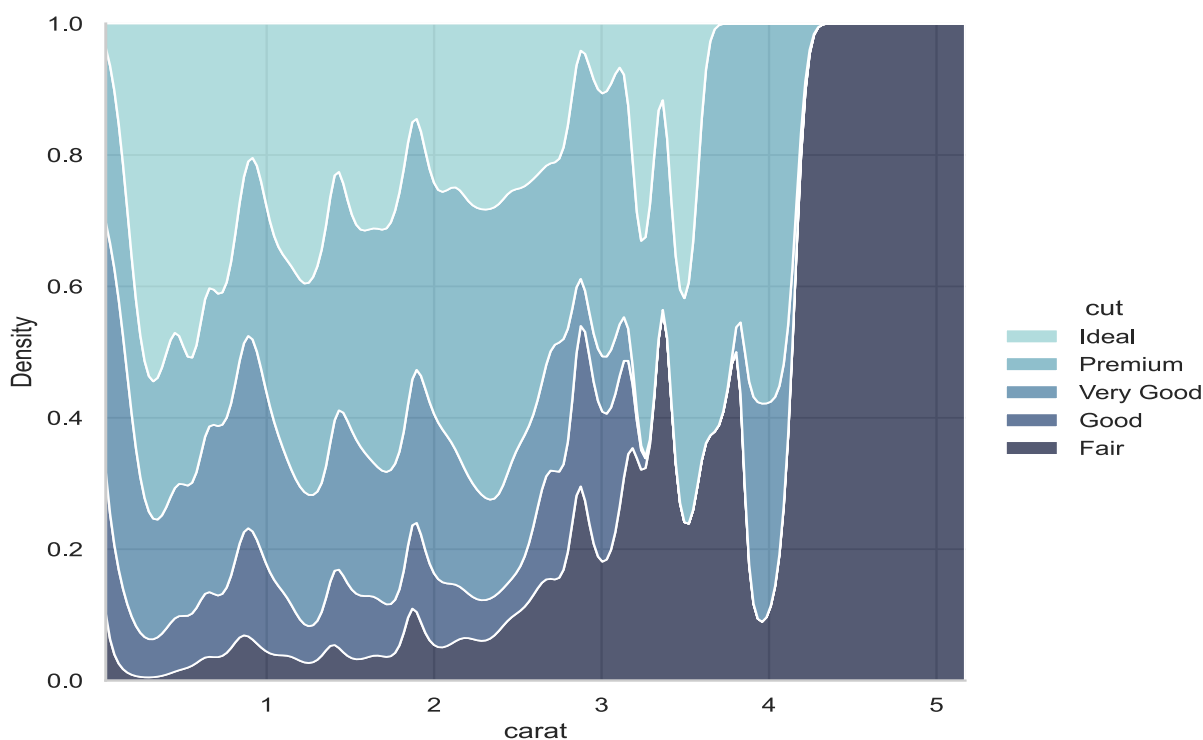
species, island, bill length, bill depth, flipper length, body mass, sex

Nhận xét:

Chim cánh cụt Gentoo có kích thước lớn nhất, cân nặng của chim trống nặng hơn chim mái ở cả ba giống chim.

Hai loài Adelie và Chinstrap có chênh lệch khối lượng không quá lớn.

**Fig3:**



**BIỂU ĐỒ PHÂN PHỐI ƯỚC TÍNH MẬT ĐỘ HẠT NHÂN CÓ ĐIỀU KIỆN  
CỦA KIM CƯƠNG THEO KHỐI LƯỢNG**

	carat	cut	color	clarity	depth	table	price	x	y	z
0	0.23	Ideal	E	SI2	61.5	55.0	326	3.95	3.98	2.43
1	0.21	Premium	E	SI1	59.8	61.0	326	3.89	3.84	2.31
2	0.23	Good	E	VS1	56.9	65.0	327	4.05	4.07	2.31
3	0.29	Premium	I	VS2	62.4	58.0	334	4.20	4.23	2.63
4	0.31	Good	J	SI2	63.3	58.0	335	4.34	4.35	2.75

Dữ liệu bao gồm:

Carat, cut, color, clarity, depth, table, price, x, y, z

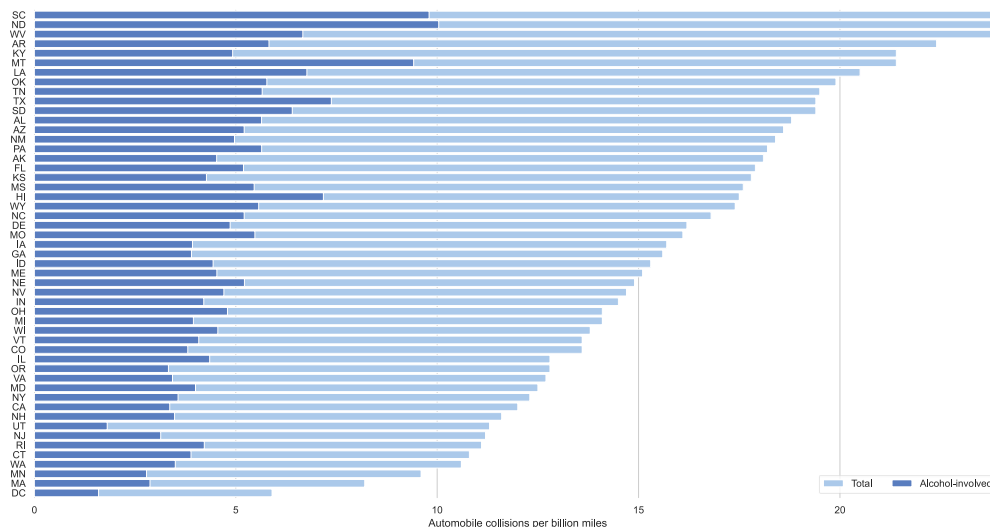
Nhận xét:

Mật độ kim cương chất lượng cao giảm dần theo khối lượng tăng dần

Mật độ chất lượng kim cương phân bố không đều theo khối lượng

Kim cương chất lượng Fair có mật độ cao nhất

**Fig4:**



**BIỂU ĐỒ THANH NGANG THỂ HIỆN CHÊNH LỆCH GIỮA TỔNG SỐ VỤ VA CHẠM XE CỘ VÀ VA CHẠM DO RƯỢU BIA TẠI CÁC TIỂU BANG CỦA MỸ TRÊN MỘT TRIỆU DẶM**

	total	speeding	alcohol	not_distracted	no_previous	ins_premium	ins_losses	abbrev
40	23.9	9.082	9.799	22.944	19.359	858.97	116.29	SC
34	23.9	5.497	10.038	23.661	20.554	688.75	109.72	ND
48	23.8	8.092	6.664	23.086	20.706	992.61	152.56	WV
3	22.4	4.032	5.824	21.056	21.280	827.34	142.39	AR
17	21.4	4.066	4.922	16.692	16.264	872.51	137.13	KY

Dữ liệu bao gồm:

Tatal, speeding, alcohol, not\_distracted, no\_previous, ins\_premium, ins\_losses, abbrev

Nhận xét:

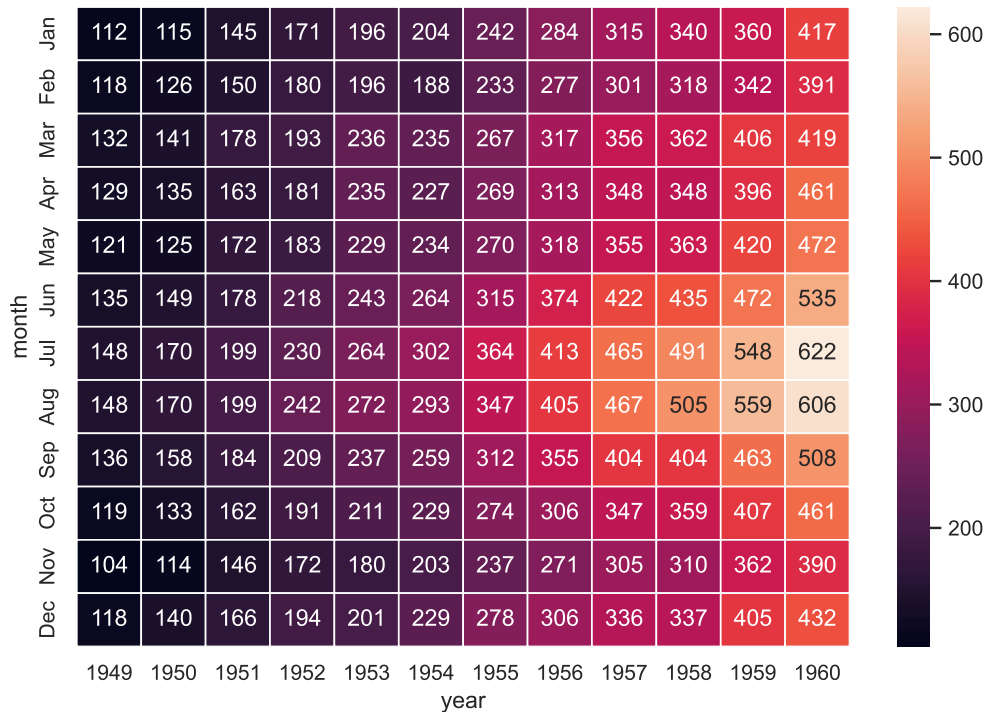
Bang South Carolina có số vụ va chạm nhiều nhất

Bang Washington có số vụ va chạm ít nhất

Số vụ va chạm do rượu bia của bang North Dakota nhiều nhất

Số vụ va chạm do rượu bia của bang Washington ít nhất

Fig5:



### BIỂU ĐỒ NHIỆT THỂ HIỆN SỐ HÀNH KHÁCH TRÊN MỖI CHUYẾN BAY MỖI THÁNG TỪ NĂM 1949 ĐẾN NĂM 1960

```
flights = flights_long.pivot(
  index="year",
  columns="month",
  values="passengers"
)
```

	year	month	passengers
0	1949	Jan	112
1	1949	Feb	118
2	1949	Mar	132
3	1949	Apr	129
4	1949	May	121

Dữ liệu bao gồm:

Year, month, passenger

Nhận xét:

Số hành khách trên các chuyến bay mỗi tháng tăng dần theo từng năm

Vào năm 1960, từ tháng 6 đến tháng 9 số lượng hàng khách lớn



Lượng hành khách những tháng mùa hè nhiều hơn các tháng còn lại