## Nâng cao doanh thu cho rạp phim

thông qua phân phối suất chiếu và chiến lược khuyến mại



## What We Do?

**Expect & Outcome** 

Xác định những yếu tố dân tới hành vi chi tiêu của khách hàng.

Đưa ra những chiến lược phù hợp nâng cao doanh thu **Dataset** 

Cinema hall ticket and customer behavior
Nguồn: Kaggle

#### Marketing Strategy

Đưa khuyến nghị phân phối suất chiếu phù hợp với nhu cầu khách hàng.
Thêm các combo hợp lý để tăng lượng khách nhóm mua vé.

## Action Plan

Phase 1

Làm sạch dữ liệu



Phase 2

Khám phá dữ liệu



Phase 3

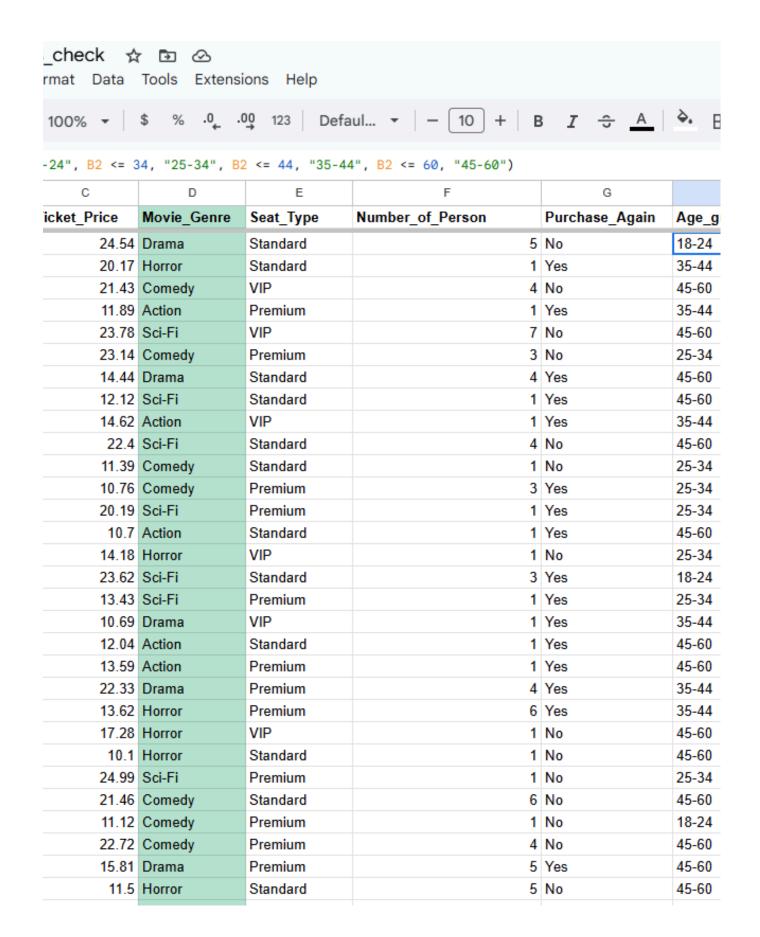
Trực quan hóa dữ liệu, đưa ra những khuyến nghị cần thiết.

## Phase 1

#### Phần mềm sử dụng: Googlesheet

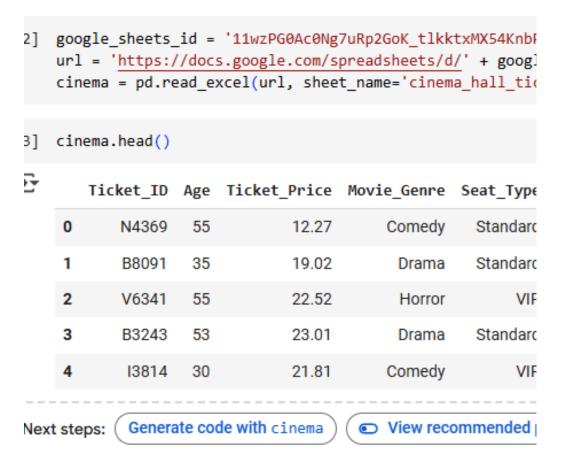
- không có duplicate, không giá trị trống.
- thay đổi giá trị 'Alone' thành '1' tại cột 'Number\_of\_person'
- tạo cột 'Age\_group', và 'total\_spend'

## Data cleaning



## Khám phá dữ liệu

#### Phần mềm sử dụng: python



- Load dữ liệu qua collab.
- Không sử dụng hai cột 'Age\_group' và 'Total\_spend'.
- In ra những hàng đầu

```
Ticket Price Number of Person
                     1440.000000
      1440.000000
                                        1440.000000
                       17.383972
                                           2.881944
         38.904861
                        4.397087
         12.426293
                                          2.192591
        18.000000
                       10.010000
                                          1.000000
                       13.640000
         28.000000
                                          1.000000
50%
                       17.280000
         39.000000
                                           2.000000
         50.000000
                       21.392500
                                          5.000000
max
         60.000000
                       24.990000
                                           7.000000
(1440, 7)
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1440 entries, 0 to 1439
Data columns (total 7 columns):
    Column
                       Non-Null Count Dtype
    Ticket ID
                                       obiect
                       1440 non-null
                       1440 non-null
                                       int64
    Ticket Price
                       1440 non-null
                                       float64
    Movie Genre
                       1440 non-null
                                       object
                       1440 non-null
    Seat Type
                                       object
                      1440 non-null
    Number of Person
    Purchase_Again
                       1440 non-null
                                       object
dtypes: float64(1), int64(2), object(4)
memory usage: 78.9+ KB
Ticket ID
```

 Kiểm tra thông số tổng quan dựa trên các lệnh: info, decribe, isnull, shape → Movie Genre: ['Comedy' 'Drama' 'H Seat Type: ['Standard' 'VIP' 'Pre Purchase Again ['No' 'Yes'] Number of Person [7 1 3 6 4 2 5]

 Kiểm tra các giá trị độc nhất băng lệnh unique()

#### → Movie\_Genre

Action 310 Horror 299 Comedy 287 Drama 286 Sci-Fi 258

Name: count, dtype: int64

Total number of person: 4150

Total number of ticket price: 25032.92

Average number of ticket price: 17.383

Number of purchase again (yes): 707
Number of purchase again (no): 733

 Thông số cơ bản băng các lệnh: value\_count, sum, len.

#### Kết luận sơ bộ:

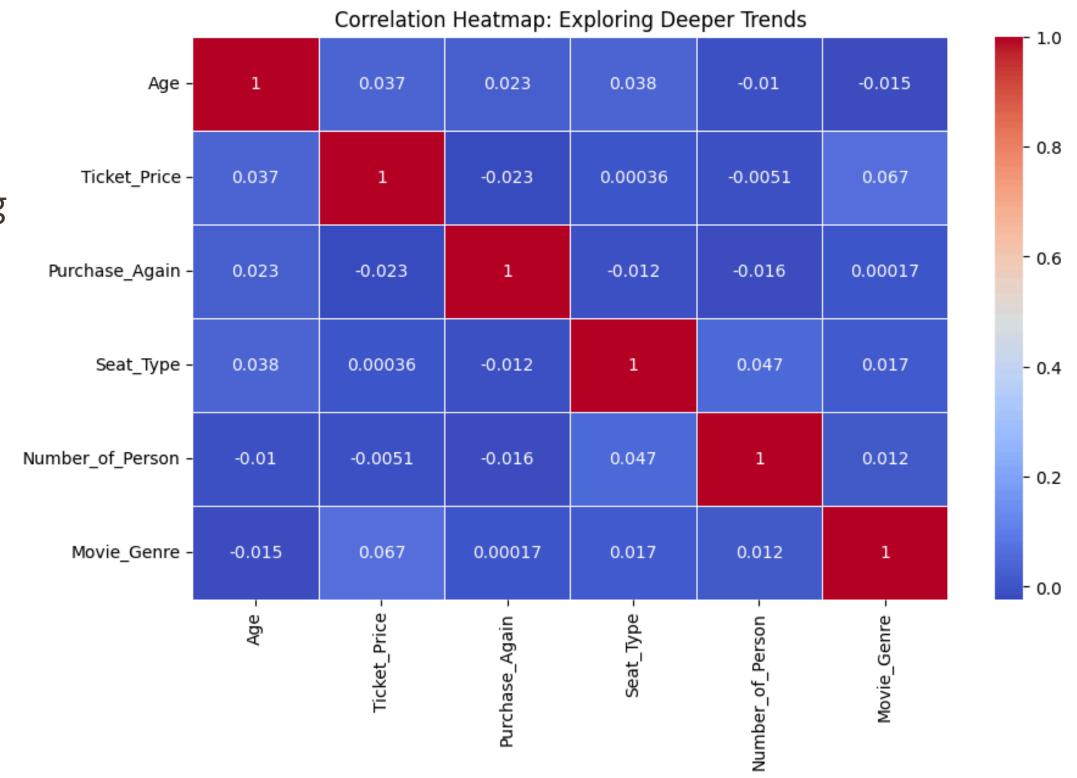
- Action, Horror và Comedy là 3 thể loại dân đầu.
- Có hơn 50% lượng khách không mua vé lại.

## Correlation

**HEATMAP** 

#### Nhận xét:

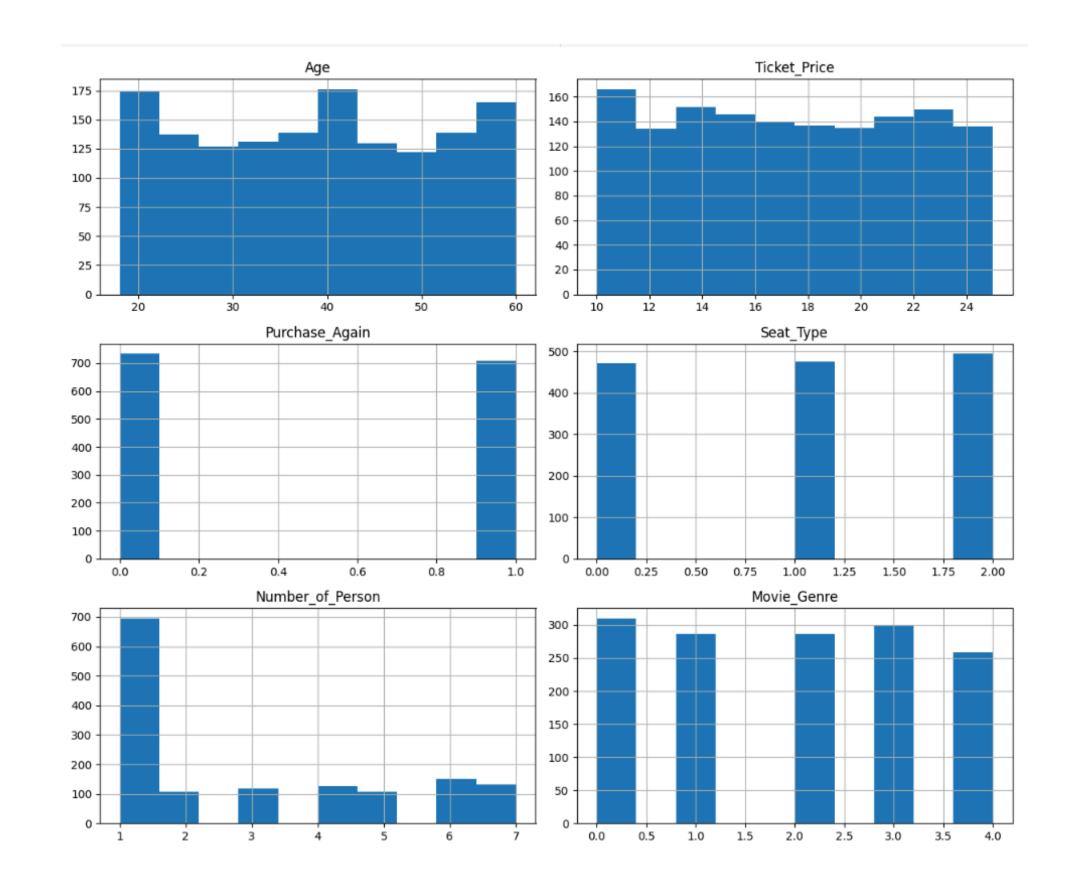
- Age & Seat\_type (0.038): tương quan thấp, không phụ thuộc vào nhau
- Ticket\_price & Purchase\_again (-0.023): giá vé không thúc đẩy tới độ trung thành.
- Movie\_genre & Ticket\_price (0.067): một vài thể loại có giá vé cao hơn.
- Number\_of\_person & Purchase\_again (-0.016): không ảnh hưởng độ trung thành.



### Correlation

### HISTOGRAM Nhận xét:

- Các nhóm tuổi phân bố rõ rệt, có thể tạo ra các khuyển mại dựa theo nhóm độ tuổi.
- Giá vé giao động khá đồng đều > có thể đưa ra chiến lược giá biến động hoặc cố định tùy loại.
- Tỉ lệ không mua lại vé cao hơn.
- Các loại ghế có tỉ lệ khá đồng đều, cần phân tích sâu để đưa kết luận.
- Phần lớn khách hàng đi xem phim một mình, các nhóm rất ít.
- Môi thế loại phim có lượng khách khác nhau.
- Tổng kết: rạp phim có lượng khách đa dạng dựa theo độ tuổi và thể loại phim.
- Khuyễn nghị: nên phân tích sâu vào sở thích khách hàng, tối ưu giá vé cho từng nhóm khách, và tìm ra vẫn đề ảnh hưởng tới quyết định mua lại vé.



## Regression

giữa GIÁ VÉ với HẠNG GHẾ & PURCHASE\_AGAIN

#### Nhận xét:

 r-square thấp > không phù hợp để dự đoán giá vé.

		ion Results			
Ticket	Price	R-squared:			
_		-			
				0.7797	
-				0.505	
•				-4174.2	
		_		8356.	
	1436	BIC:		8377.	,
	3				
nonr	obust				
========			========	=========	
coef	std er	r t	P> t	[0.025	0.975]
9.4807	0.06	3 149.912	0.000	9.357	9.605
3.2661	0.16	19.635	0.000	2.940	3.592
2.9523	0.16	5 17.784	0.000	2.627	3.278
3.2623	0.16	4 19.917	0.000	2.941	3.584
4.8333	0.120	9 40.362	0.000	4.598	5.068
4.6473	0.12	1 38.466	0.000	4.410	4.884
			-,-		
	1.781	Cond. No.			
_	Least So Wed, 19 Mar 21: nonr coef 9.4807 3.2661 2.9523 3.2623 4.8333 4.6473	OLS Least Squares Wed, 19 Mar 2025 21:04:47 1440 1436 3 nonrobust  coef std err  9.4807 0.063 3.2661 0.166 2.9523 0.166 3.2623 0.166 4.8333 0.126 4.6473 0.123	Least Squares F-statistic: Wed, 19 Mar 2025 Prob (F-statis 21:04:47 Log-Likelihood 1440 AIC: 1436 BIC: 3 nonrobust  coef std err t  9.4807 0.063 149.912 3.2661 0.166 19.635 2.9523 0.166 17.784 3.2623 0.164 19.917 4.8333 0.120 40.362 4.6473 0.121 38.466  1317.636 Durbin-Watson: 0.000 Jarque-Bera (J 0.022 Prob(JB): 1.781 Cond. No.	OLS Adj. R-squared: Least Squares F-statistic: Wed, 19 Mar 2025 Prob (F-statistic): 21:04:47 Log-Likelihood: 1440 AIC: 1436 BIC: 3 nonrobust  coef std err t P> t   9.4807 0.063 149.912 0.000 3.2661 0.166 19.635 0.000 2.9523 0.166 17.784 0.000 3.2623 0.164 19.917 0.000 4.8333 0.120 40.362 0.000 4.6473 0.121 38.466 0.000  1317.636 Durbin-Watson: 0.000 Jarque-Bera (JB): 0.022 Prob(JB): 1.781 Cond. No.	OLS Adj. R-squared: -0.000 Least Squares F-statistic: 0.7797 Wed, 19 Mar 2025 Prob (F-statistic): 0.505 21:04:47 Log-Likelihood: -4174.2 1440 AIC: 8356. 1436 BIC: 8377. 3 nonrobust  Coef std err t P> t  [0.025  9.4807 0.063 149.912 0.000 9.357 3.2661 0.166 19.635 0.000 2.940 2.9523 0.166 17.784 0.000 2.627 3.2623 0.164 19.917 0.000 2.941 4.8333 0.120 40.362 0.000 4.598 4.6473 0.121 38.466 0.000 4.410  1317.636 Durbin-Watson: 1.947 0.000 Jarque-Bera (JB): 89.282 0.022 Prob(JB): 4.10e-20

#### Notes

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The smallest eigenvalue is 3.2e-28. This might indicate that there are strong multicollinearity problems or that the design matrix is singular.

## Logistic

giữa PURCHASE AGAIN với HẠNG GHẾ, ĐỘ TUỔI, THỂ LOẠI PHIM

#### Nhận xét:

- giá trị Pseudo R-squared rất thấp > mô hình không phù hợp để dự đoán khả năng mua lại vé.
- Các hệ số có giá trị rất nhỏ và độ lệch chuẩn rất lớn, có thể là do vấn đề về tỷ lệ dữ liệu hoặc đa cộng tuyến

Optimization terminated successfully.

Current function value: 0.689150

Iterations 8

Logit Regression Results

=======================================						==
Dep. Variable:	Purchase_Again_Yes		No. Observations:		1440	
Model:	Logit		Df Residuals:		1432	
Method:	MLE		Df Model:		7	
Date:	Wed, 19 Ma	ar 2025	Pseudo R-squ.		0.0055	33
Time:	26	0:50:55	Log-Likelihoo	d:	-992.	38
converged:	True		LL-Null:		-997.90	
Covariance Type:	nor	nrobust	LLR p-value:		0.1368	
=======================================	coef	std err	z	P> z	[0.025	0.975]
const	-0.2286	3.78e+06	-6.05e-08	1.000	-7.41e+06	7.41e+06
Seat_Type_Premium	-0.1142	3.78e+06	-3.02e-08	1.000	-7.41e+06	7.41e+06
Seat_Type_Standard	0.0536	3.78e+06	1.42e-08	1.000	-7.41e+06	7.41e+06
Seat_Type_VIP	-0.1680	3.78e+06	-4.44e-08	1.000	-7.41e+06	7.41e+06
Age	0.0031	0.004	0.727	0.467	-0.005	0.011
Movie_Genre_Action	0.1598	0.169	0.943	0.346	-0.172	0.492
Movie_Genre_Comedy	-0.0294	0.173	-0.170	0.865	-0.368	0.309
Movie_Genre_Drama	0.3331	0.173	1.928	0.054	-0.006	0.672
Movie_Genre_Horror	0.2586	0.171	1.510	0.131	-0.077	0.594

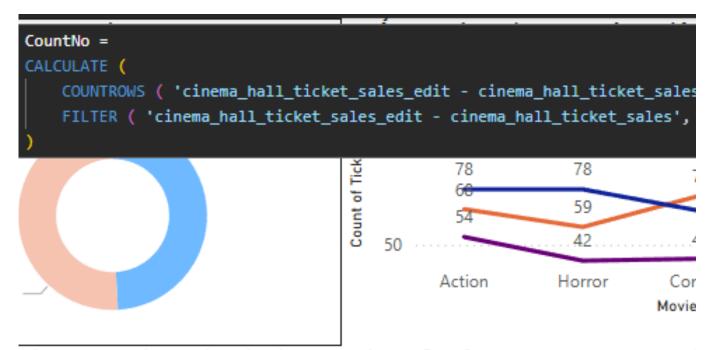
### Kết luận sơ bộ:

- Khó có thể dự đoán hành vi mua lại vé của khách hàng.
- Nhưng có thể đi sâu vào các thể loại ưa thích, cũng như hạng ghế và độ tuổi để đưa ra phương án điều chỉnh.

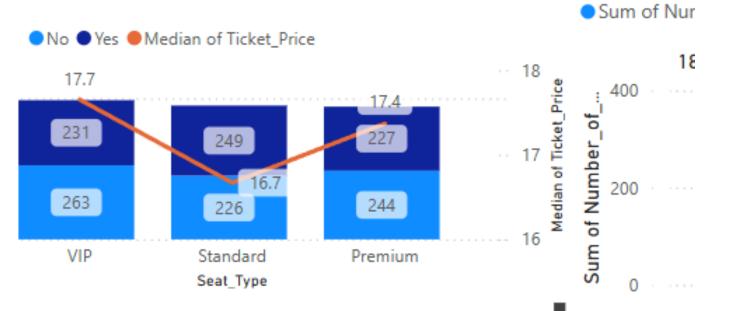
#### Phần mềm sử dụng: PowerBl

• Tạo 2 measure: CountNo, CountYes

## TRỰC QUAN HÓA DƯ LIỆU



hách mua vé lại và giá vé trung bình của từng loại Lượng khá



pình của mỗi hạng ghế không có sự chênh lệch quá lớn > Vậy đây guyên nhân chính dẫn tới việc không mua lại vé của khách hàng. ¿ẽ có những thể loại phim ưa thích khác nhau.

a, và Action là 3 thể loại có lượng khách hàng quay lại cao nhất. Trong à thể loại khó giữ chân khách hàng nhất. Sci-fi là thể loại có lượng nhất nhưng ổn định của nhóm từ 35 đổ lên.



## Tổng quan

120

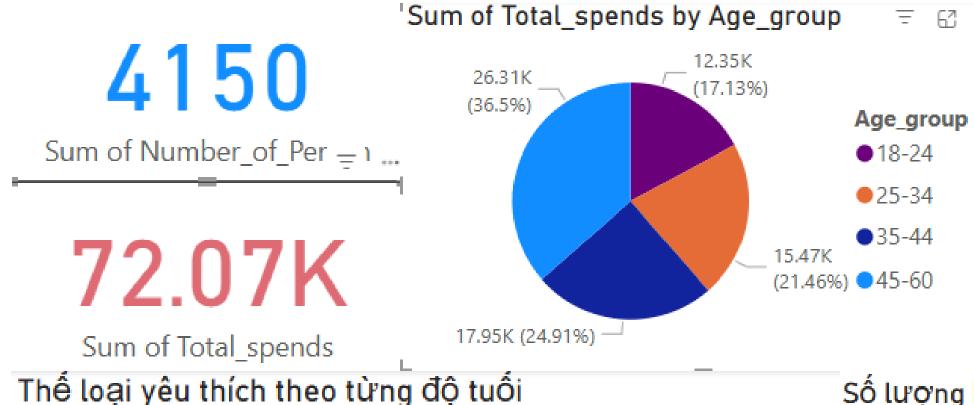
78

Count of Ticket\_ID

100

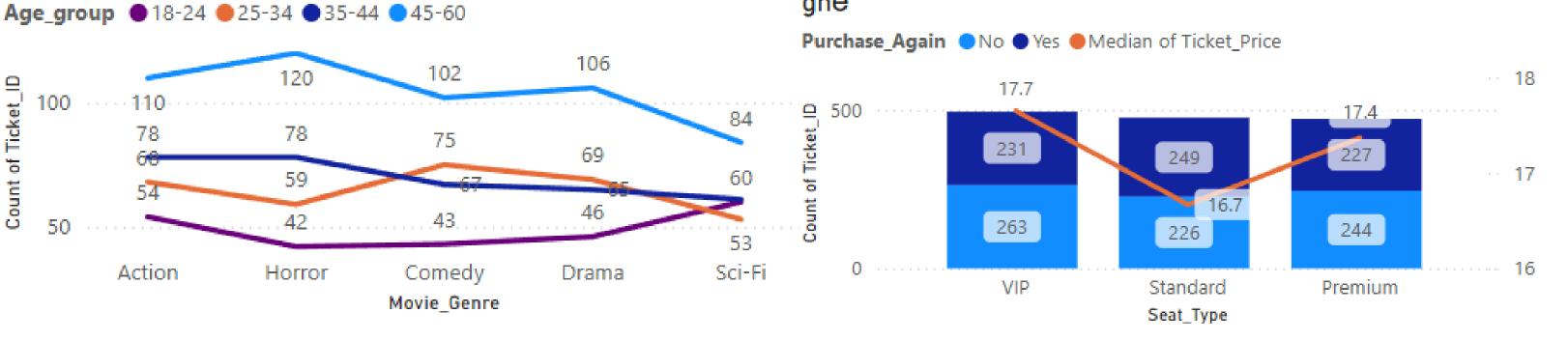
50

Action

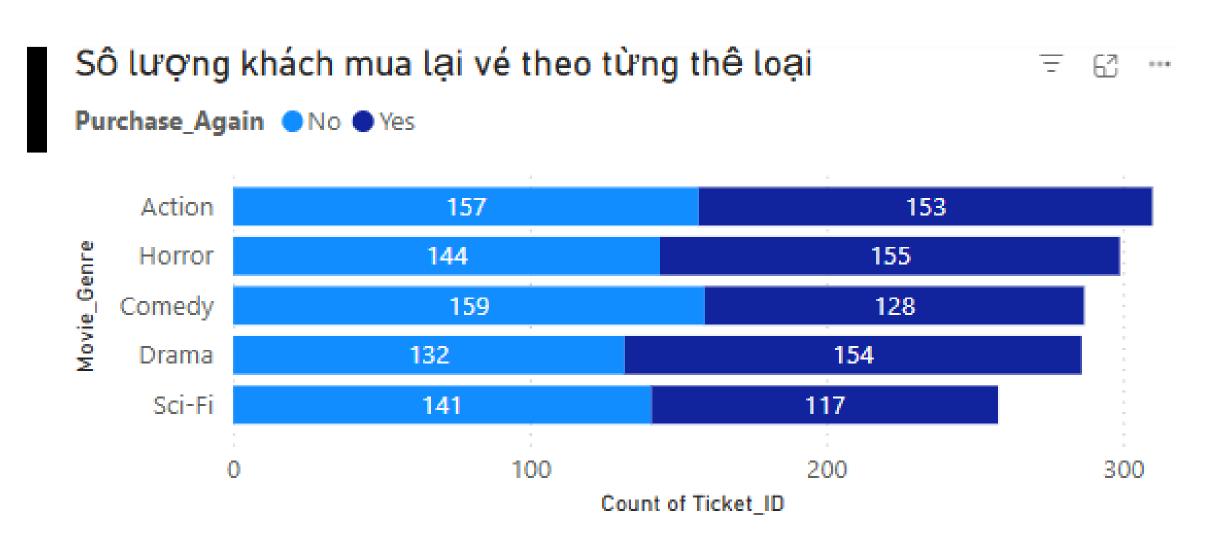


- Nhóm khách hàng (45–60) chiếm % doanh thu cao nhất, nhóm (18-24) có doanh thu thấp nhất.
- Môi nhóm khách có những thể loại ưa thích khác nhau: (45-60) thích action, horror. (35-44) thích action, horror. (25-34) thích comedy, drama. (18-24): thích sci-fi, action.
- Giá vé trung bình các hạng ghế không chênh lệch nhiều, nhưng lượt giữ chân cao nhất thuộc về hạng Standard > do giá cả phải chăng. Nhưng hạng VIP lại có lượng bán cao nhất > khách hàng có nhu cầu trải nghiệm tốt.

Số lượng khách mua vé lại và giá vé trung bình của từng loại ghê

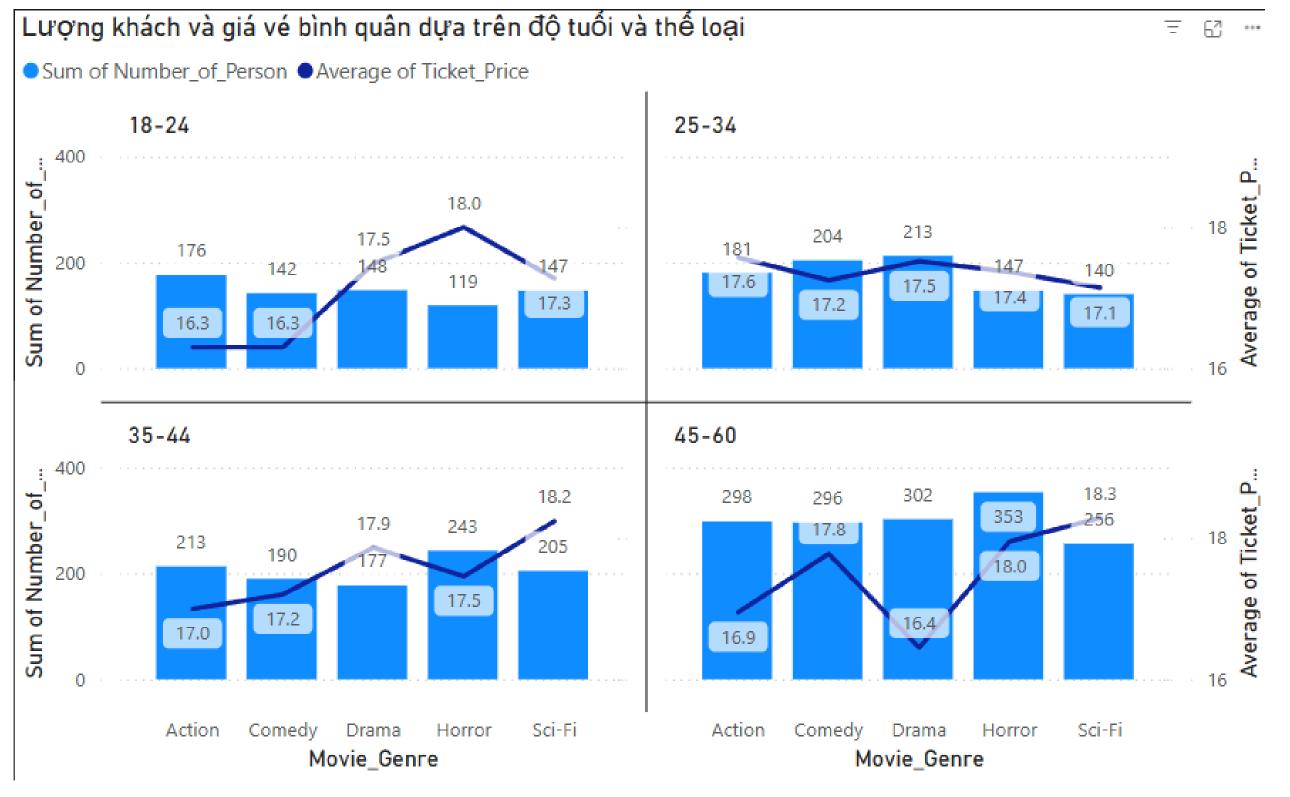


## Phân bổ suất chiếu



- Action, Horror và comedy là 3 thể loại có lượt mua cao nhất.
- Tuy nhiên, Comedy lại là thể loại khó giữ chân người xem nhất.
- Sci-fi là thể loại ít người xem, nhưng như đã biết, đây là thể loại thu hút khách hàng độ tuổi (18– 24) > xét về lâu dài, để tạo thói quen xem phim lâu dài, vẫn nên tạo ra những chương trình khuyến mại nhất định để giữ chân nhóm khách này.

## Phân bổ suất chiếu

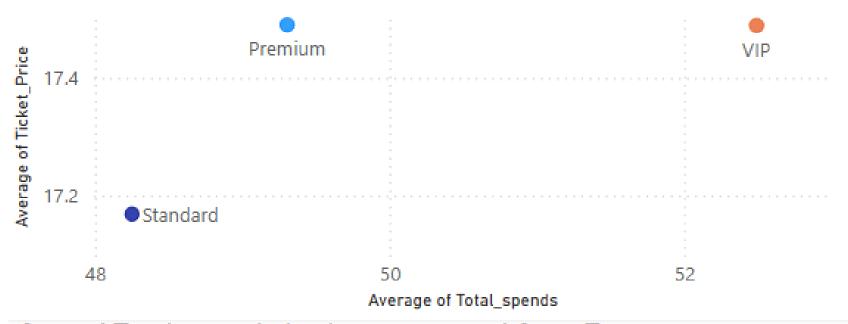


#### Suggestion:

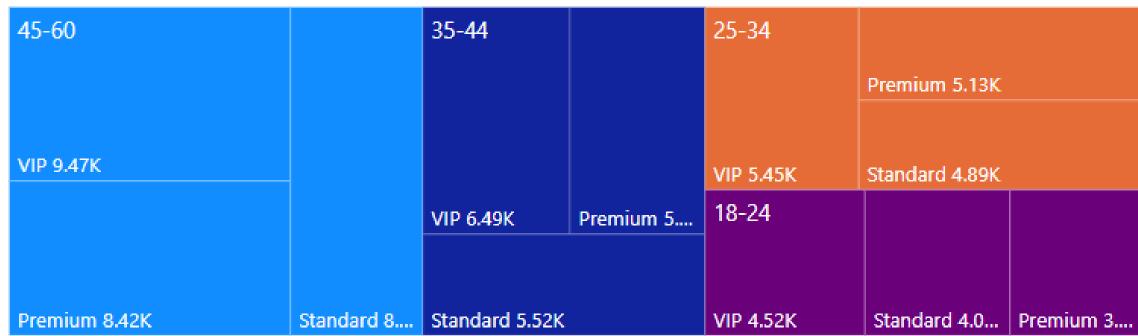
- Tăng suất chiếu các thể loại Horror, Drama, và Action.
- Tung các gói ưu đãi nâng hạng ghế để kích thích khách hàng nhóm, đối với các thể loại trên.
- Cân băng suất chiếu và giá vé của sci-fi. Đi kèm những combo bán kèm nhân vật ưa thích, hoặc liên kết tổ chức các sự kiện kèm theo để tăng doanh thu.
- Giảm giá và suất chiếu đối với thể loại Comedy.
- Ngoài ra, ưu tiên săp xếp thể loại Action, Horror vào các rạp chất lượng cao. Drama vào các rạp lớn.

### Promotion

Average of Total\_spends and Average of Ticket\_Price by Seat\_Type



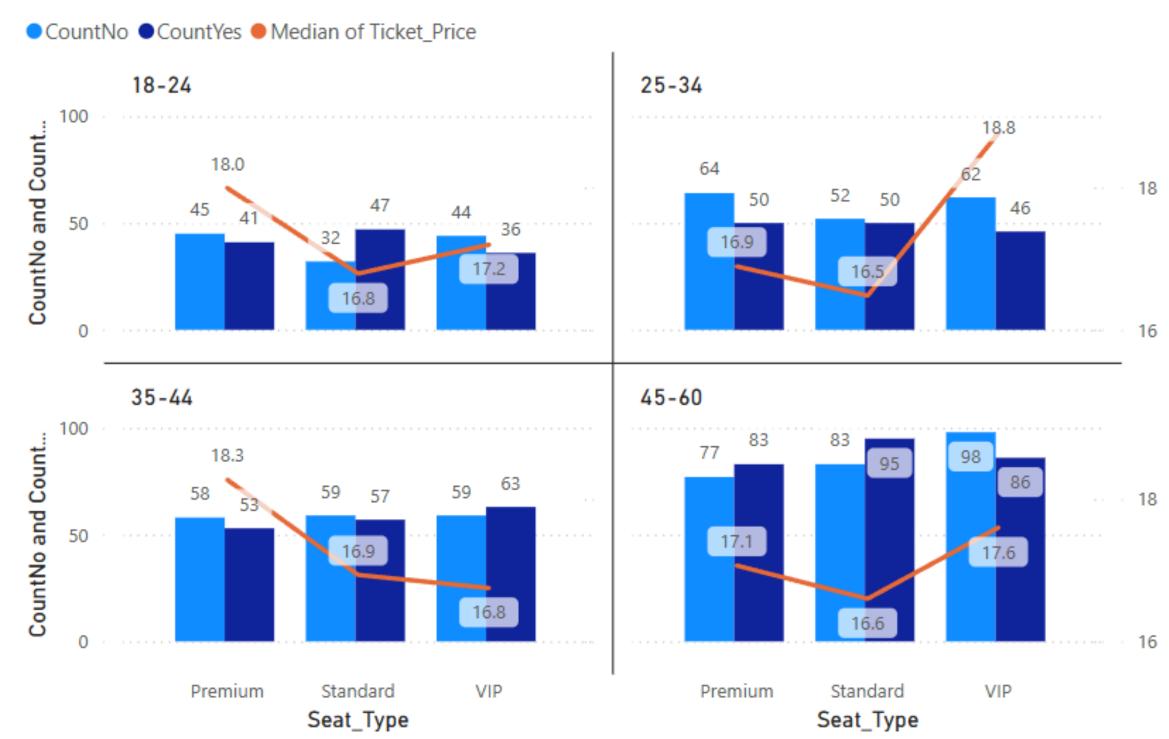
#### Sum of Total\_spends by Age\_group and Seat\_Type



- (18-24) mua nhiều hạng ghế
   Premium, nhưng tỉ lệ giữ chân
   khách thấp với hạng ghế này. Tỉ lệ
   giữ chân cao nhất của nhóm này
   thuộc về hạng Standard > có lẽ là
   do giá cả.
- (24-34) có tỉ lệ giữ chân khách thấp. Chi tiêu nhiều vào ghế Premium. > có lẽ do yêu cầu của nhóm cao, nhưng chi tiêu có tính toán.
- (35-44) là nhóm khách có tỉ lệ giữ chân cân băng, có thể xem là nhóm dẽ chịu. Tỉ lệ giữ chân cao thuộc về hạng VIP > nhóm khách yêu cầu trải nghiệm tốt.
- (45-60) là nhóm chi tiêu nhiều nhất, với mức độ giữ chân trung bình. Chi tiêu tập trung vào hạng VIP và Standard > chứng tỏ họ chú trọng giải trí, nhưng cũng cần kinh tế. Muốn giữ chân lâu dài cần có promotion có giá trị lâu dài.

### Promotion

CountNo, CountYes and Median of Ticket\_Price by Seat\_Type and Age\_group



#### Suggestion:

- Tạo ra các combo nhóm tiện ích đối với nhóm khách lửa tuổi (18-24).
- Tạo khuyến mại nâng hạng ghế cho nhóm (24–34).
- Nâng cao chất lượng ghế VIP: và thêm dịch vụ kèm theo, ví dụ: phục vụ tại rạp, nước miên phí, chăn ấm, vv...
- Đổi với nhóm khách (45-60) cần đưa ra chiến lược lâu dài: vơ tích điểm trên giao dịch đổi dịch vụ hoặc up hạng ghế miên phí dựa trên số lượng vé đã mua. Giới thiệu cho khách sự khác biệt về chất lượng dịch vụ khách biệt hạng VIP và hạng thường, và những khuyến mại cho các hạng premium và VIP (vì avarage của 2 hạng này không chênh lệch quá cao)

# Thank You!

