

The background is a gradient of teal and blue. It features several white medical crosses of varying sizes and two stylized dollar bills, one near the top left and one near the bottom left.

# Healthcare is expensive

*& it is getting more expensive*

## Singapore's healthcare spending is likely to continue rising; targeted support needed: Gan

© FRI, MAR 05, 2021 - 4:13 PM | UPDATED FRI, MAR 05, 2021 - 4:17 PM

**SHARON SEE** ✉ [sharons@sph.com.sg](mailto:sharons@sph.com.sg) 🐦 [@SharonSeeBT](https://twitter.com/SharonSeeBT)

SINGAPORE'S spending on healthcare has grown every year and is likely to continue rising, which is why the country needs to be prudent and ensure support is targeted at those in greater need, Health Minister Gan Kim Yong said on Friday.



Ministry of Health &gt; News Highlights

## MANAGING HEALTHCARE COST INCREASES

2ND NOV 2020

Name and Constituency of Member of Parliament

Ms Ng Ling Ling  
MP for Ang Mo Kio GRC

**Question No. 264**

To ask the Mii  
announced ra

**Answer**

1. Several factors were associated with more medical visits. The mean number of visits was 6.9 days, and the mean cost of healthcare was \$1,000. A



# Rising healthcare costs and universal health coverage in India: an analysis of national sample surveys, 1986-2014

**GUMBER, Anil**, LALITHA, N and DHAK, Biplab (2017). *Rising healthcare costs and universal surveys, 1986-2014*. Working Paper. India, Gujarat Institute of Development Research, Ahmedabad.



## ■ THE FISCAL CHALLENGE

## FINDING SOLUTIONS

## ■ WHAT WE'RE DOING

## ■ WHAT YOU CAN DO

**RESOURCE  
LIBRARY**

[Home](#) | [Blog](#) | [Why Are Americans Paying More For Healthcare?](#)

## HOW MUCH DOES THE UNITED STATES SPEND ON HEALTHCARE?

States has one of the highest costs of healthcare in the world. In the United States spent about \$3.6 trillion on healthcare, which averages \$11,000 per person. Relative to the size of the economy, healthcare costs have increased over the past few decades, from 5 percent of gross domestic product (GDP) in 1960 to 18 percent in 2018. The Centers for Medicare and Medicaid Services (CMS) project that by 2028, such costs will climb to \$16.2 trillion, or about \$18,000 per person, and will represent about 20 percent of GDP. However, those projections do not take into account the impacts of the

## THE EFFECTS OF RISING HEALTHCARE COSTS ON THE US ECONOMY

THE EFFECTS OF RISING HEALTHCARE COSTS ON THE US ECONOMY

**Author(s):** Elena Toader

**Subject(s):** Health and medicine and law

**Published by:** Addleton Academic Publishers

**Keywords:** healthcare cost growth; US economy; reform;

## Health coverage in India: an analysis of national sample



 Share  Print

## FROM THE BLOG

### ■ How Does the U.S. Healthcare System Compare to Other Countries?

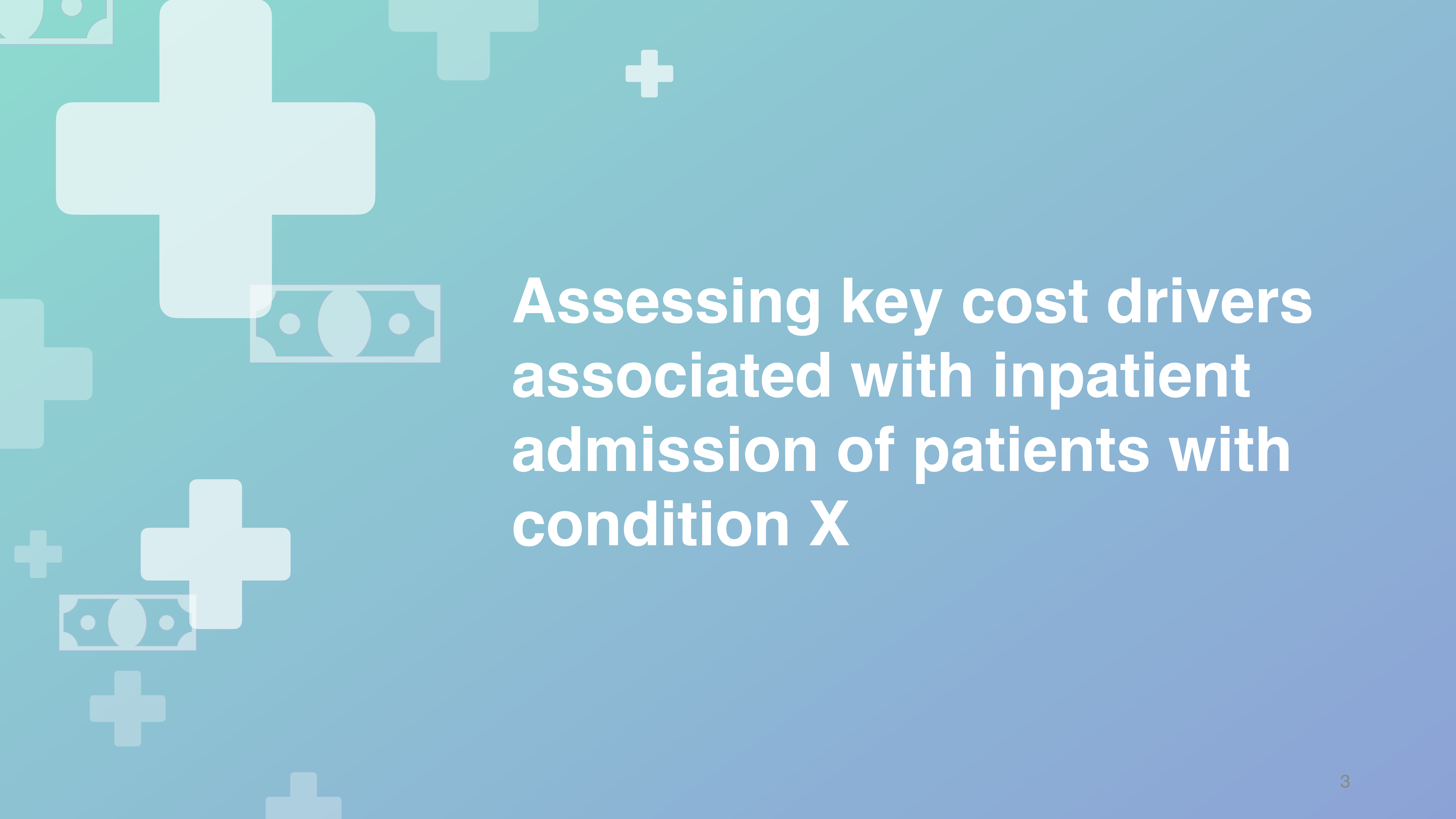


## Rising pressure of healthcare cost

By CECILIA KOK

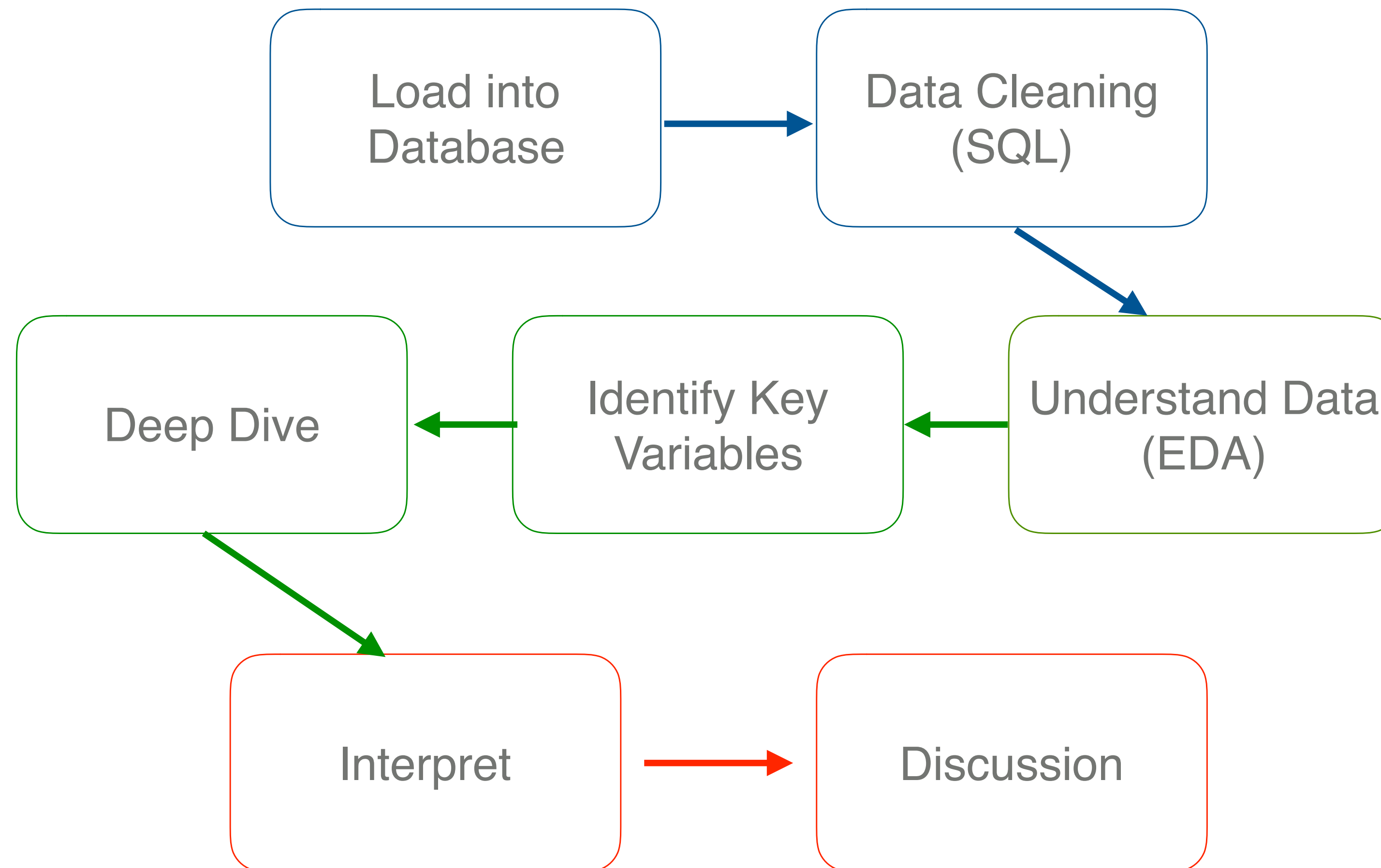
Increases in medical bills are outpacing the general inflation rate each year. That raises the question whether healthcare is reserved only for those who can afford it

*"I got the bill for my surgery. Now I know what those doctors were wearing masks for"*  
– American bureaucrat, James H. Boren (1925)

The background is a gradient of teal and blue. It features several white and light blue icons: a large cross in the top left, a smaller cross in the top center, a cross with a dollar bill inside it in the middle left, a cross with a dollar bill inside it in the bottom left, and several other smaller crosses and dollar bills scattered around.

# Assessing key cost drivers associated with inpatient admission of patients with condition X

# Project Flow





# Contents

- **Our Data**
  - New Variables & Clean Up
  - Length of Stay & Target Variable
  - Demographics
- **Which variable matters?**
  - Decision Tree for Feature Importance
  - Correlation Coefficient Matrix
- **Patient Level Data vs Admission Clinical Data**
  - Encounter per year
  - Medical History
  - Body Mass Index (BMI)
- **Model & Interpretation**
- **Use Case Discussion**
- **Given more time ...**



# Our Data

A summary of our dataset .

# Our Data

- Clinical and financial data of patients hospitalised for a certain (X) condition
- 3,400 inpatient admission
- 3,000 unique patients
- 5 year data (2011 - 2015)
- Per row per inpatient admission

## Clinical (Admission)

- Date of admission/discharge
- Pre-Op Medication
- Symptoms
- Lab Results
- Weight / Height

## Patient

- Demographics
- Medical History

## Financial

- Total Bill per admission

# New Variables

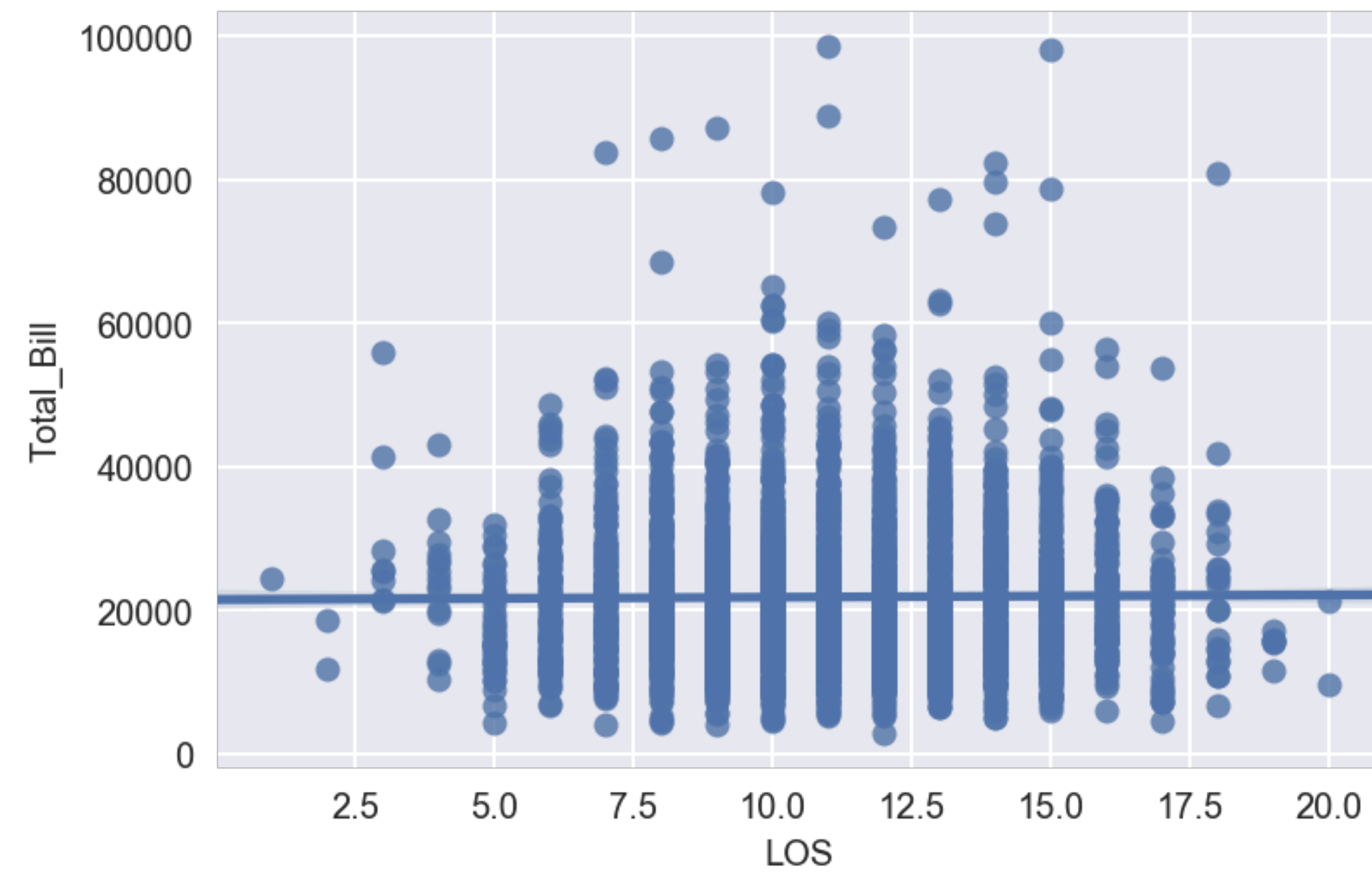
- Number of encounter per patient (in 5 years)
- Length of Stay
- Age
- BMI
- BMI Risk Level
- Year / Quarter of admission
- 30 Days Readmission
- Number of Medical History
- Number of Symptoms
- Number of Pre-Op Medications

# Clean-Up

- Standardising strings
- Converting into numeric
- Imputing null values



# Length of Stay



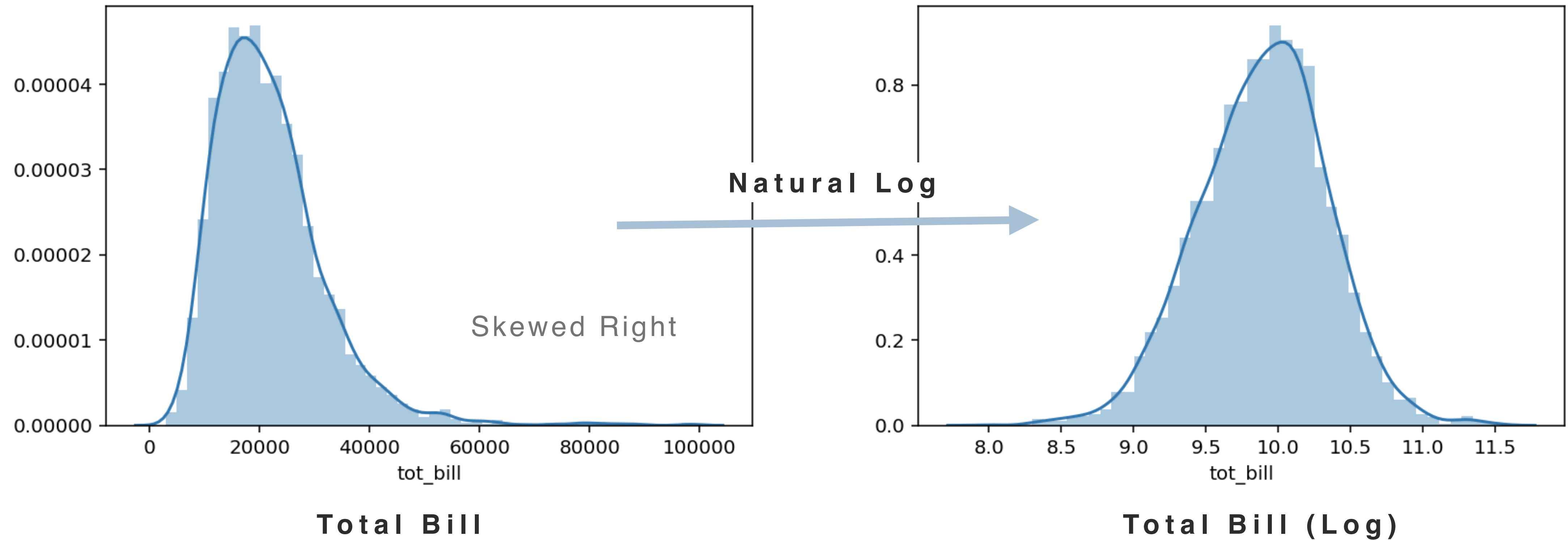
**Pearson Correlation Coefficient**

R = 0.0090

P-value = 0.5996

# Target Variable

Distribution



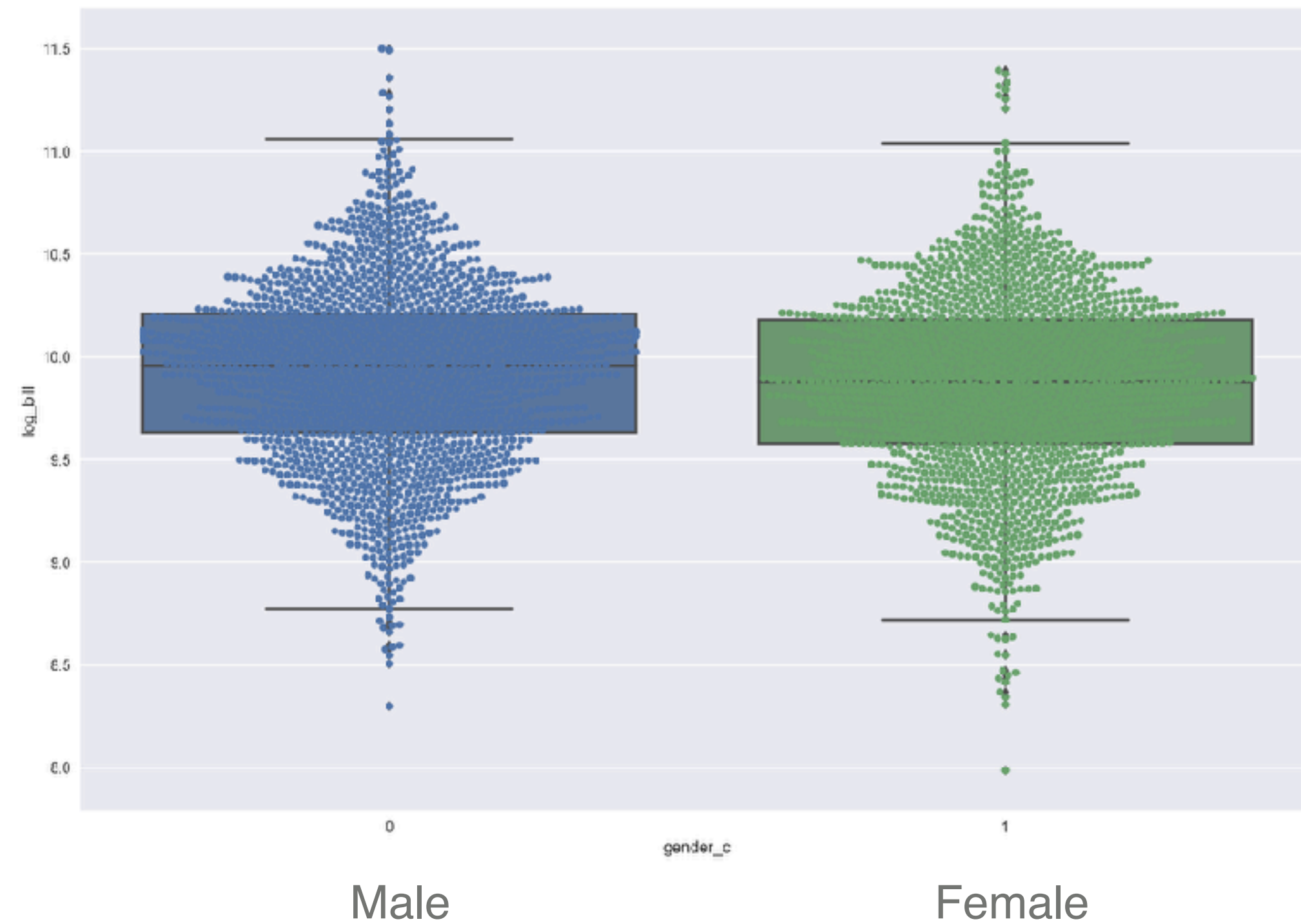
To ensure approximate normality

# Demographics

		Count (per patient)	Total Bill (per adm)	
			Mean	P-Value
<b>Total</b>		<b>3,000</b>	<b>21,859</b>	
<b>Gender</b>	Female (1)	1,497	21,273	<b>&lt;0.01*</b>
	Male (0)	1,503	22,446	
<b>Resident Status</b>	Singaporean (0)	2,392	20,211	<b>&lt;0.01*</b>
	PR (1)	465	24,370	
	Foreigner (2)	143	41,704	
<b>Race</b>	Indian (1)	295	23,682	<b>&lt;0.01*</b>
	Chinese (2)	1,915	19,118	
	Malay (3)	629	29,506	
	Others (4)	161	21,320	
<b>Age</b>	< 55	1,755	19,334	<b>&lt;0.01*</b>
	> 55	1,252	25,398	

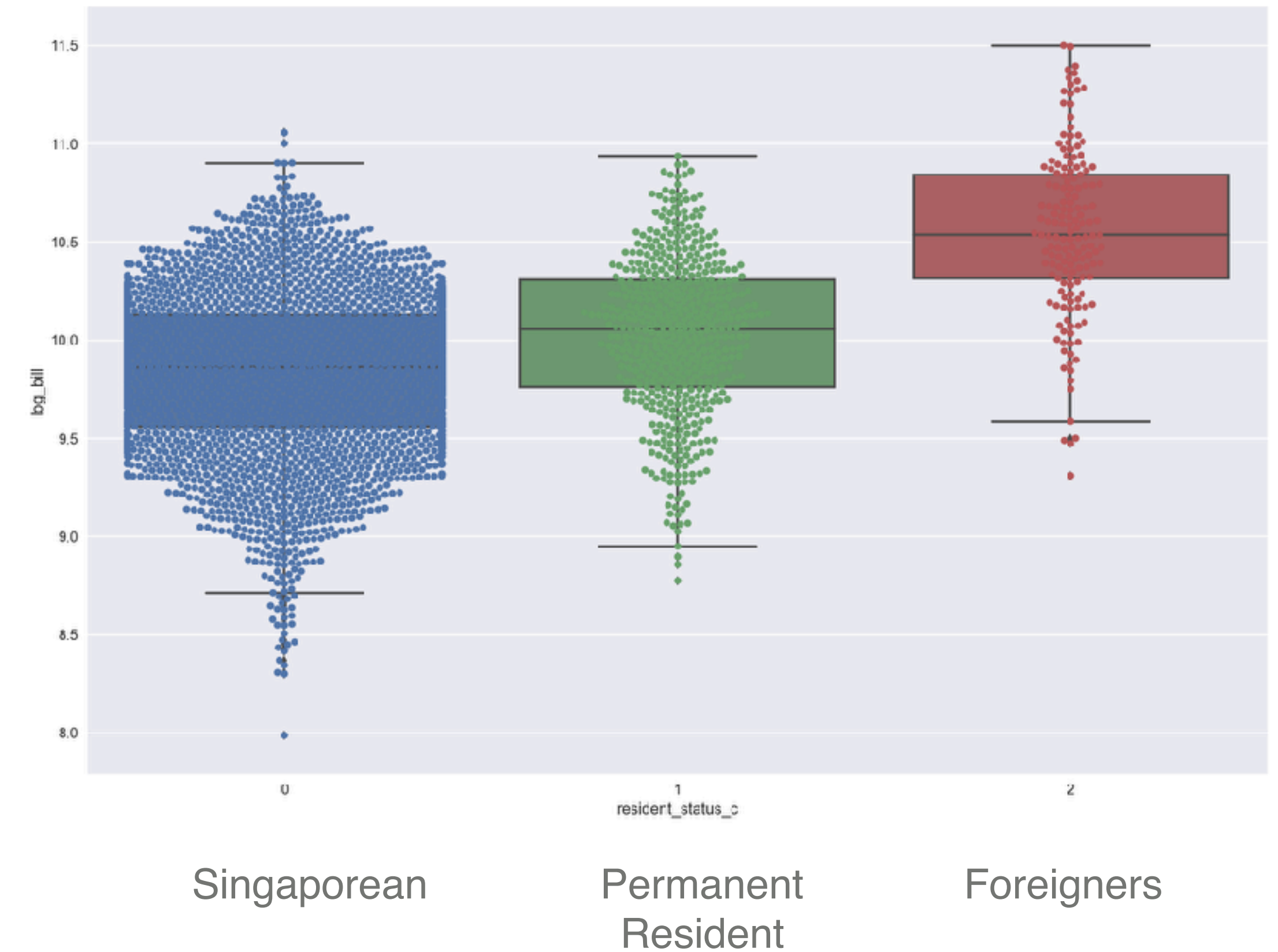
\* Not tested for homogeneity and normality assumptions due to time constraint. One-way ANOVA is used

# Gender



- Almost equal representation between male and female
- Male having 5.5% higher mean than female

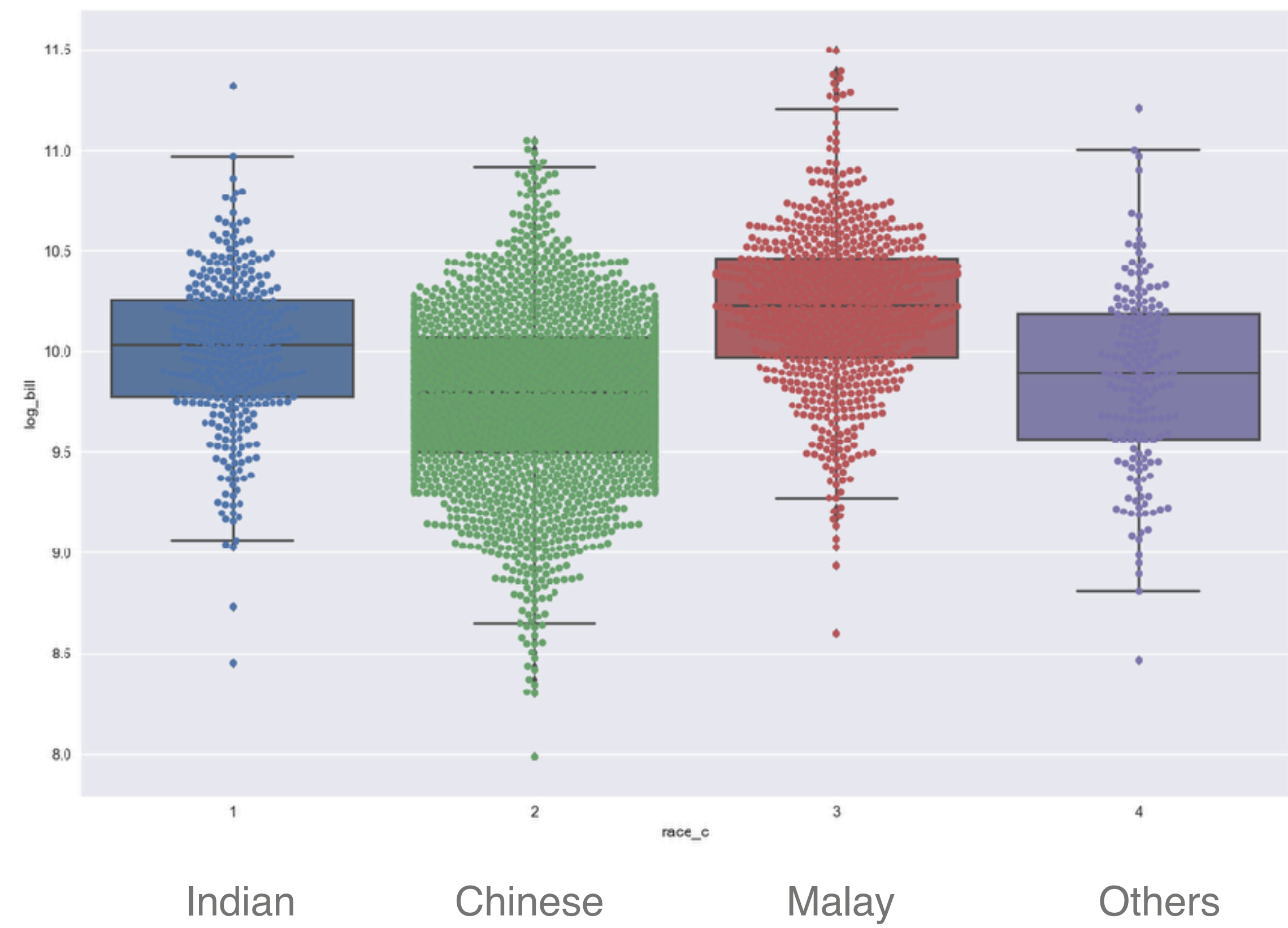
# Resident Status



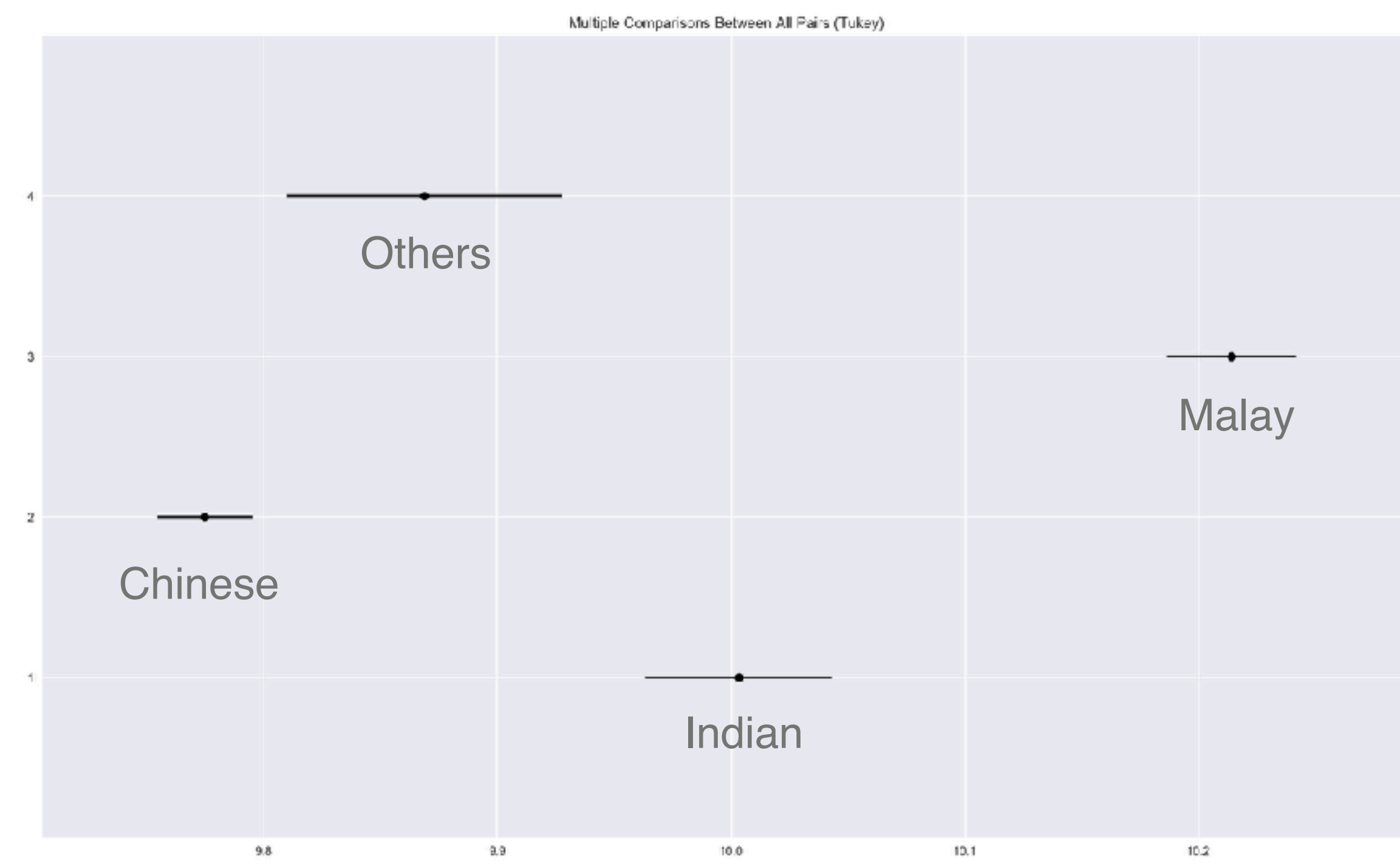
- Singaporean have higher subsidies than PR and foreigners



# Race



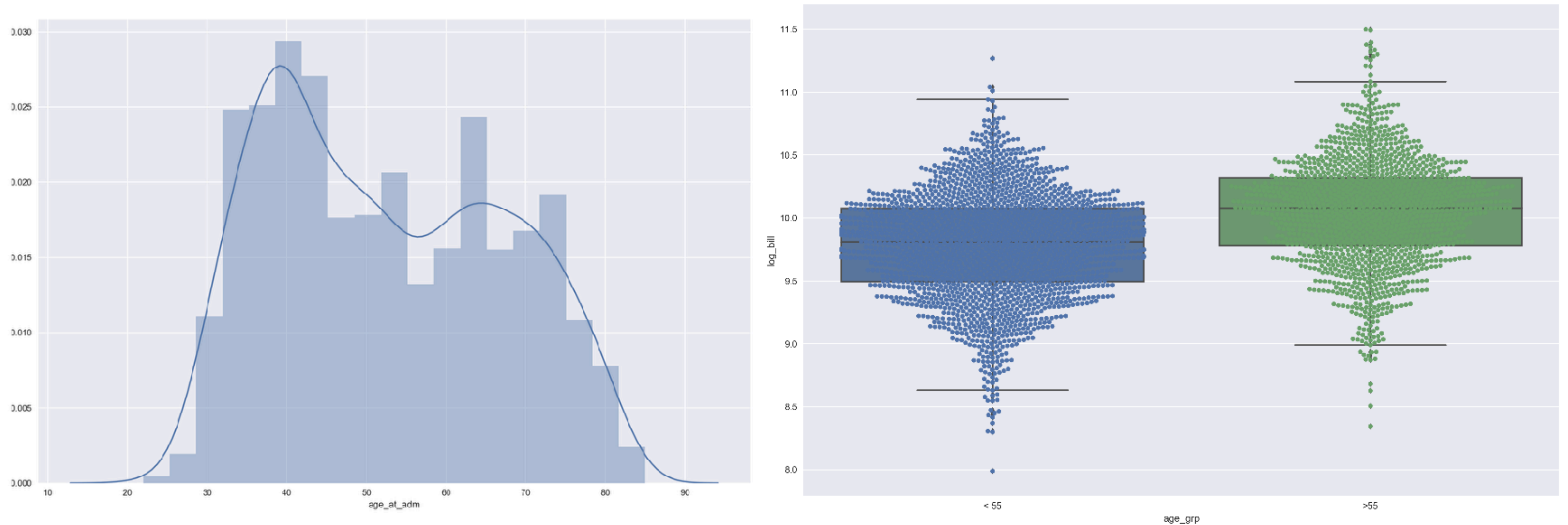
- There are more Chinese represented in our data
- Malays have the highest mean out of all races



Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	lower	upper	reject
1	2	-0.2285	-0.2896	-0.1673	True
1	3	0.2108	0.1415	0.2801	True
1	4	-0.1343	-0.2308	-0.0379	True
2	3	0.4393	0.3936	0.4849	True
2	4	0.0941	0.013	0.1753	True
3	4	-0.3451	-0.4325	-0.2577	True

# Age



- Bimodal Distribution - We have 2 groups - above 55 and below 55
- Patients  $> 55$  years old have ~31% higher mean total bill than those  $< 55$



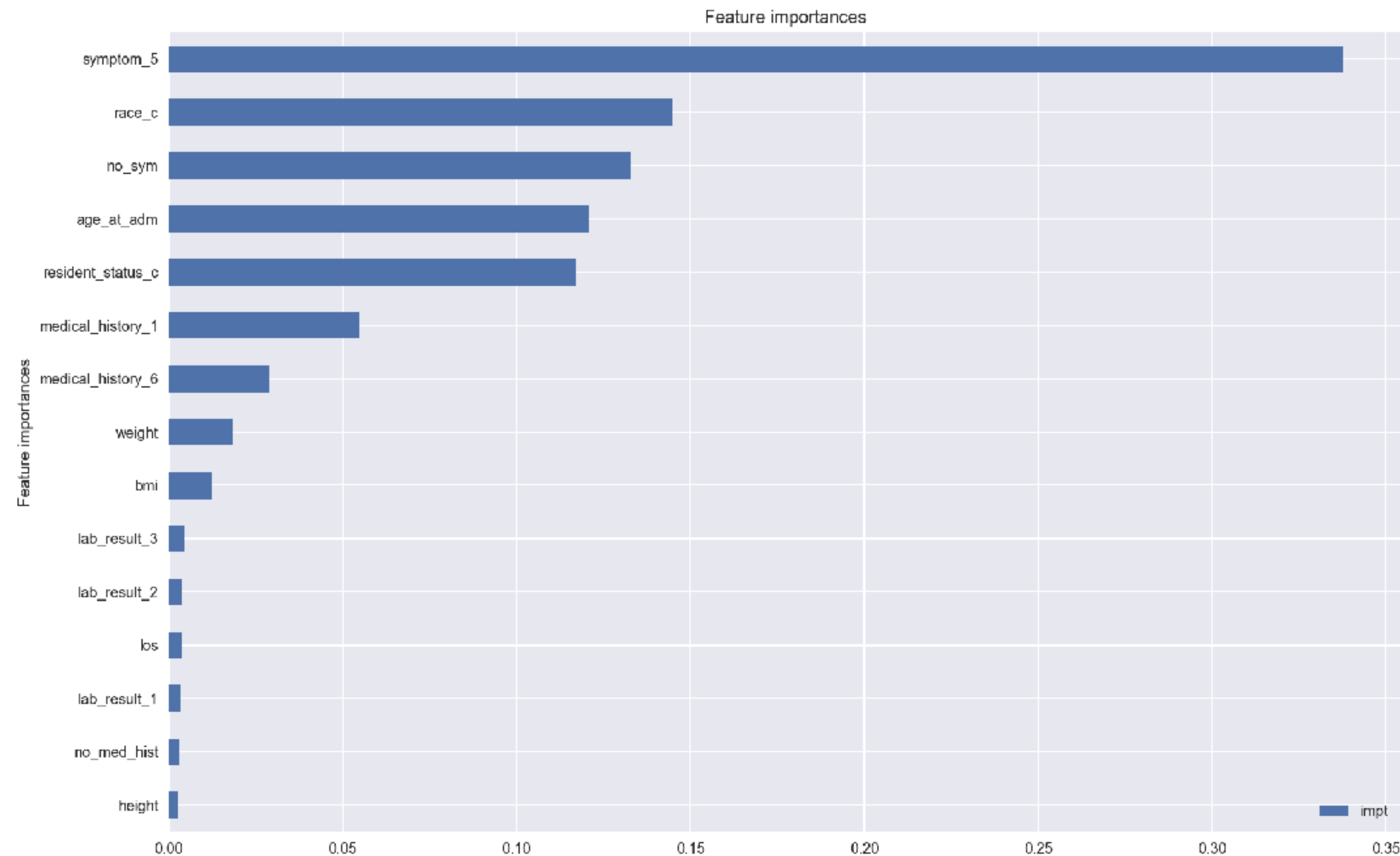


# Which variable matters?

A quick way to help us focus on what matters .

# Decision Tree

A quick dirty way to help zoom into important variables

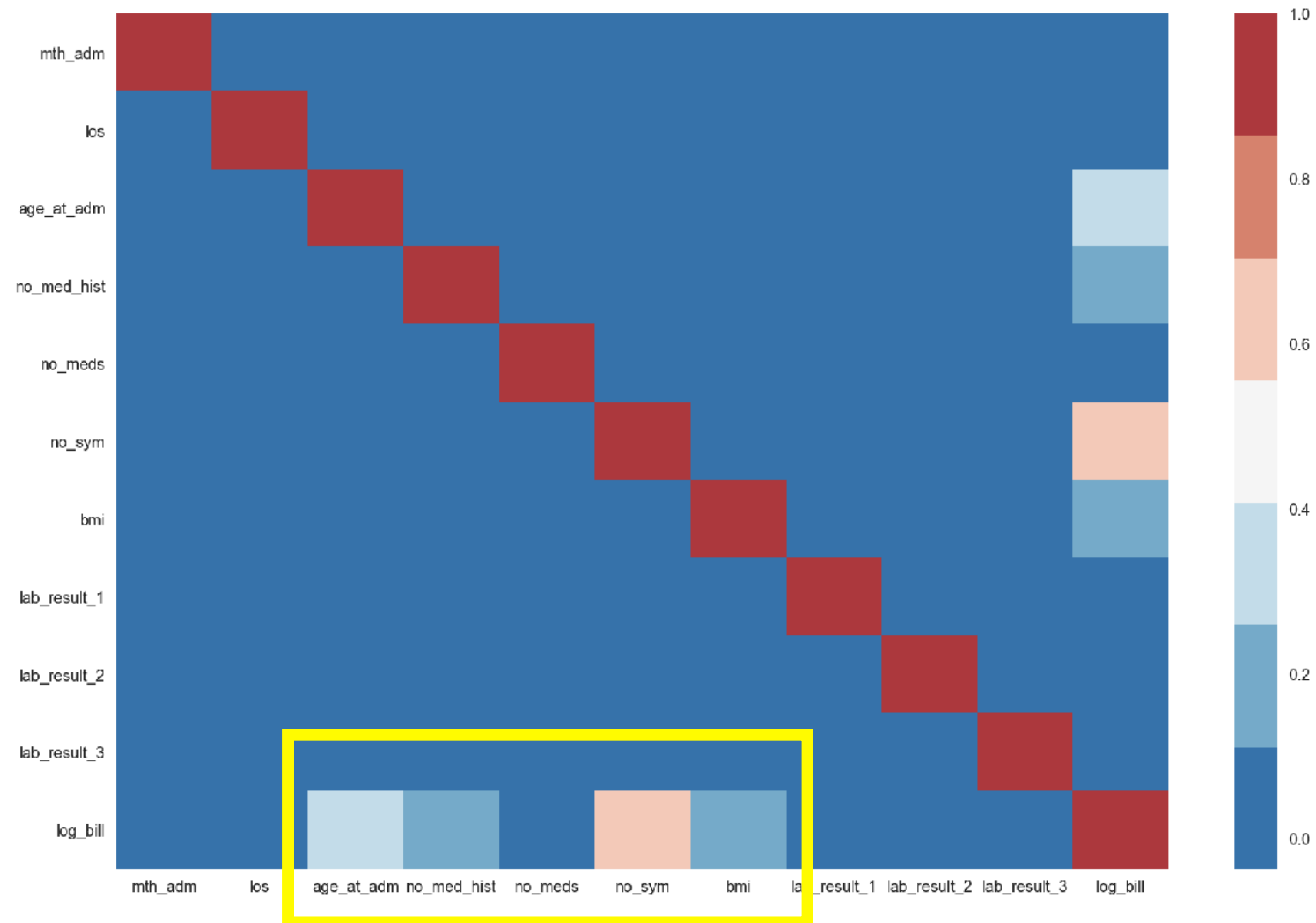


## Top 10

1. Symptom 5
2. Race
3. Number of Symptoms
4. Age
5. Resident Status
6. Medical History 1
7. Medical History 6
8. BMI (Height & Weight)
9. Lab Result 3
10. Lab Result 2

# Correlation Coefficient Matrix

On continuous variables



**Some positive correlation**

- Age
- Number of medical history
- Number of symptoms
- BMI



# Patient Level vs Admission Level

It depends on what we want to achieve .

## Patient

- Race
- Age
- Resident Status
- BMI
- Number of Medical History
- Medical History 1
- Medical History 6

## Clinical (Admission)

- Symptom 5
- Number of Symptoms
- Lab Result 3
- Lab Result 2

## Patient

- Race
- Age
- Resident Status
- BMI
- Number of Medical History
- Medical History 1
- Medical History 6

## Clinical (Admission)

- Symptom 5
- Number of Symptoms
- Lab Result 3
- Lab Result 2



# Our Data

- Clinical and financial data of patients hospitalised for a certain (X) condition
- Per row per patient's total bill per year
- 3,314 rows
- 3,000 unique patients
- 5 year data (2011 - 2015)

# Variables

- **Number of encounter per year** (*continuous*)
- **Readmission (> 1 encounter per year)** (*binary*)
- **Age** (*continuous*)
- **Age (<55, >55)** (*categorical*)
- **BMI** (*continuous*)
- **BMI Risk Level** (*categorical*) ([healthhub.sg](https://www.healthhub.sg))
- **Number of Medical History** (*continuous*)
- **Medical History 1 through 7** (*binary*)

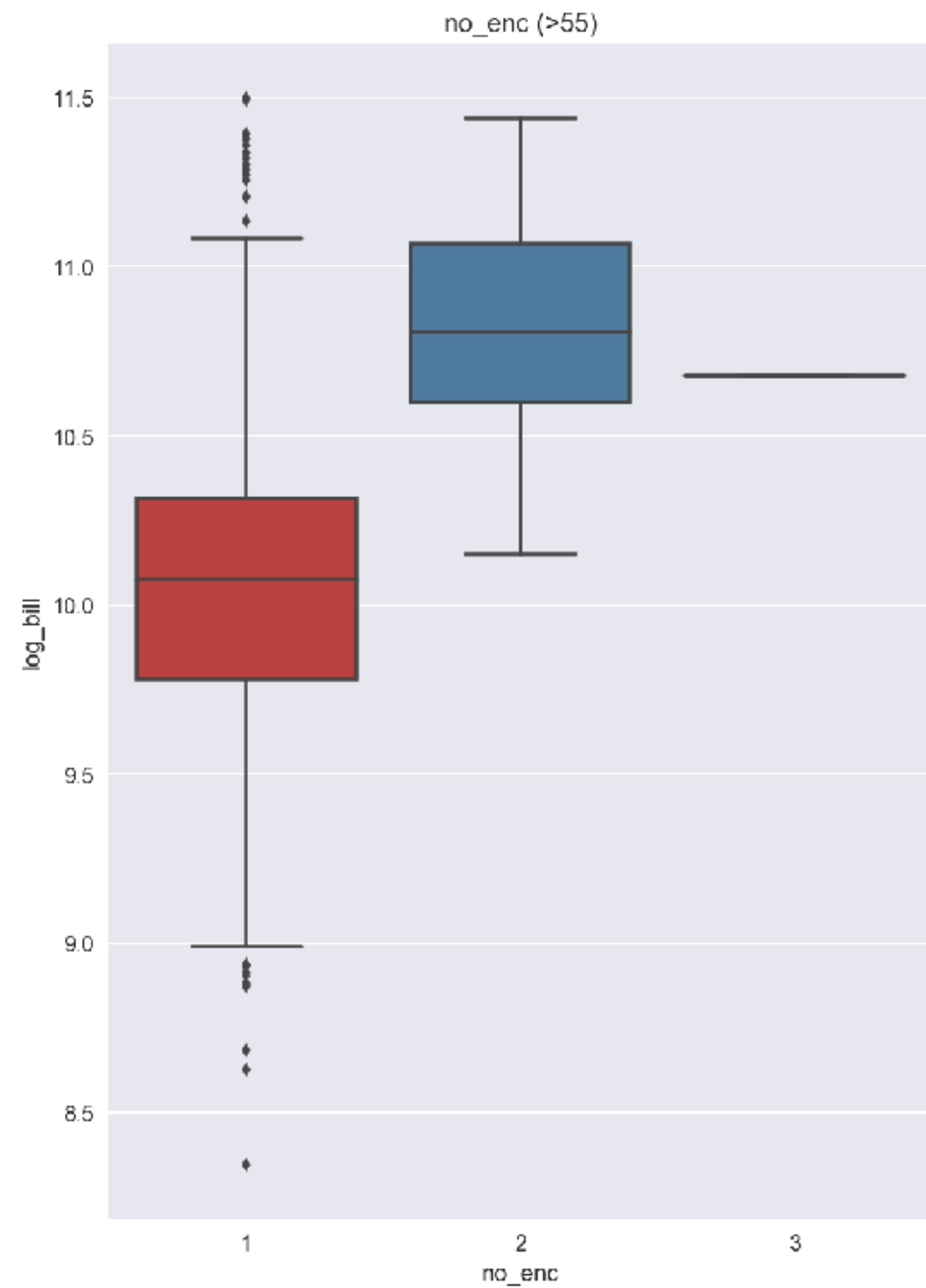
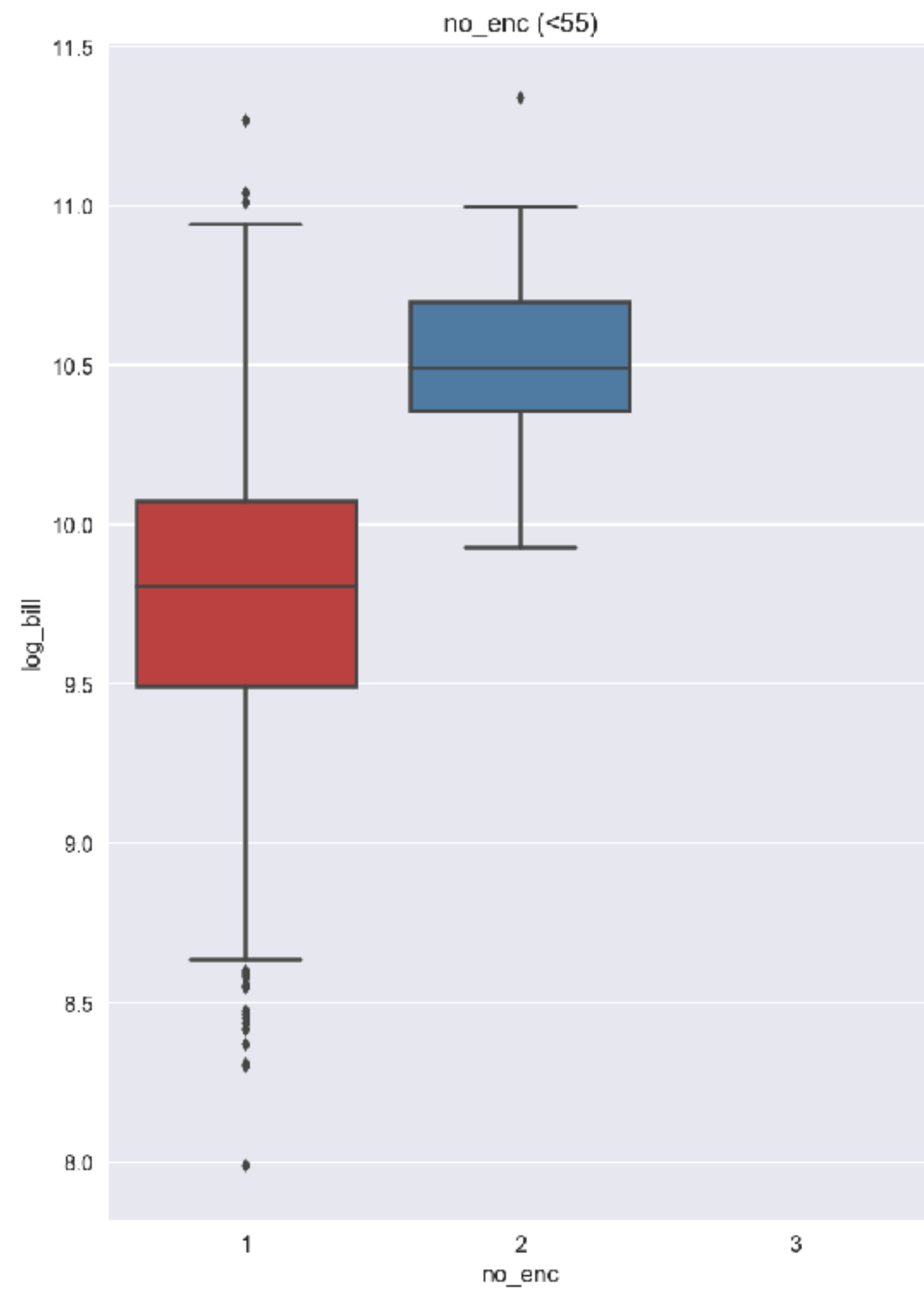
# Target Variable

- **Patient's total bill per year**

# Number of Encounters/Year

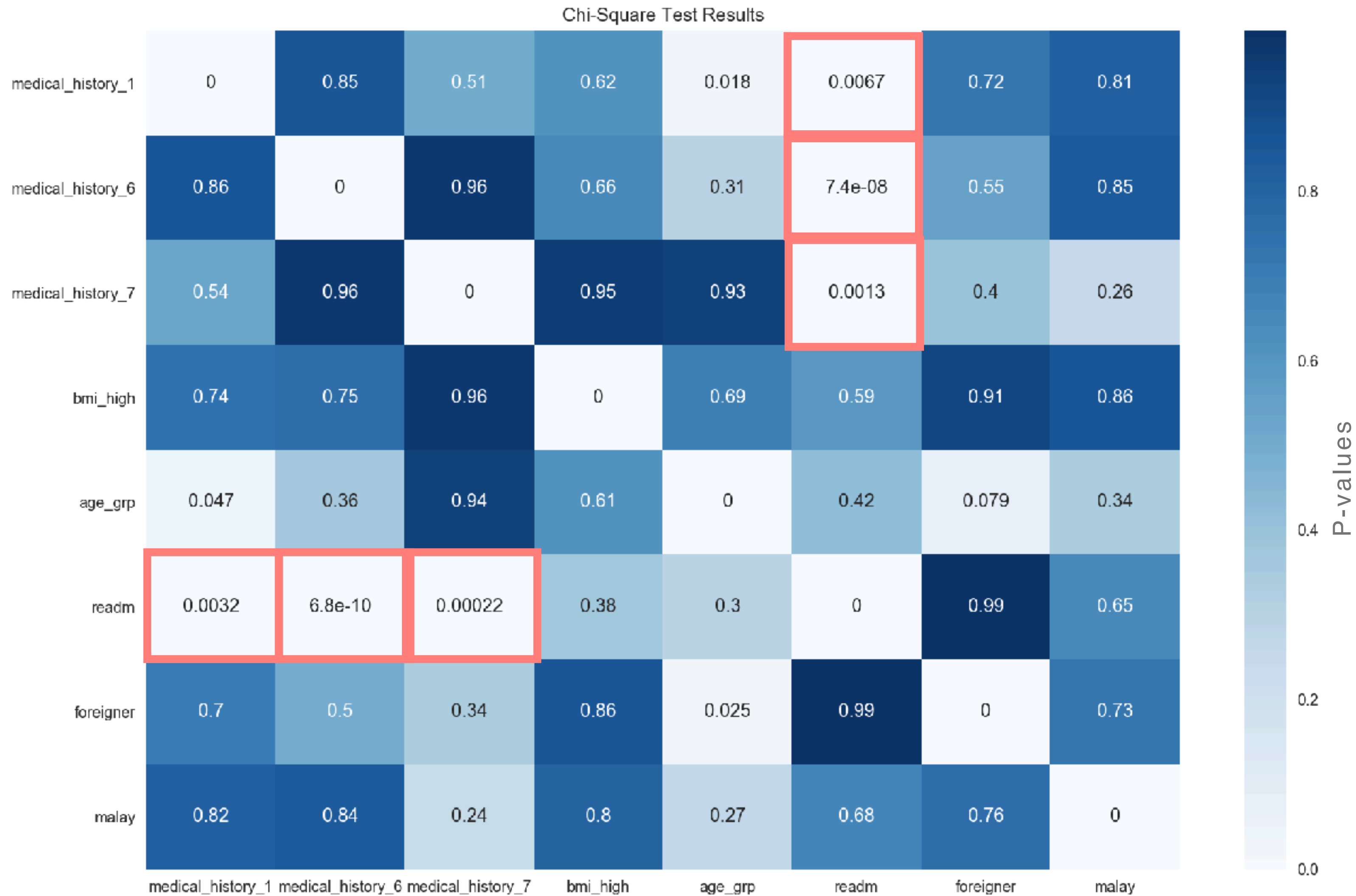
		Count	Total Bill (per yr)	
			Mean	P-Value
Total		3,314		
Number of Encounter / Year	1	2,928	21,828	<0.01*
	2	84	45,140	
	3	1	43,490	

# Number of Encounters / Year



- P-Value < 0.01 for both age groups

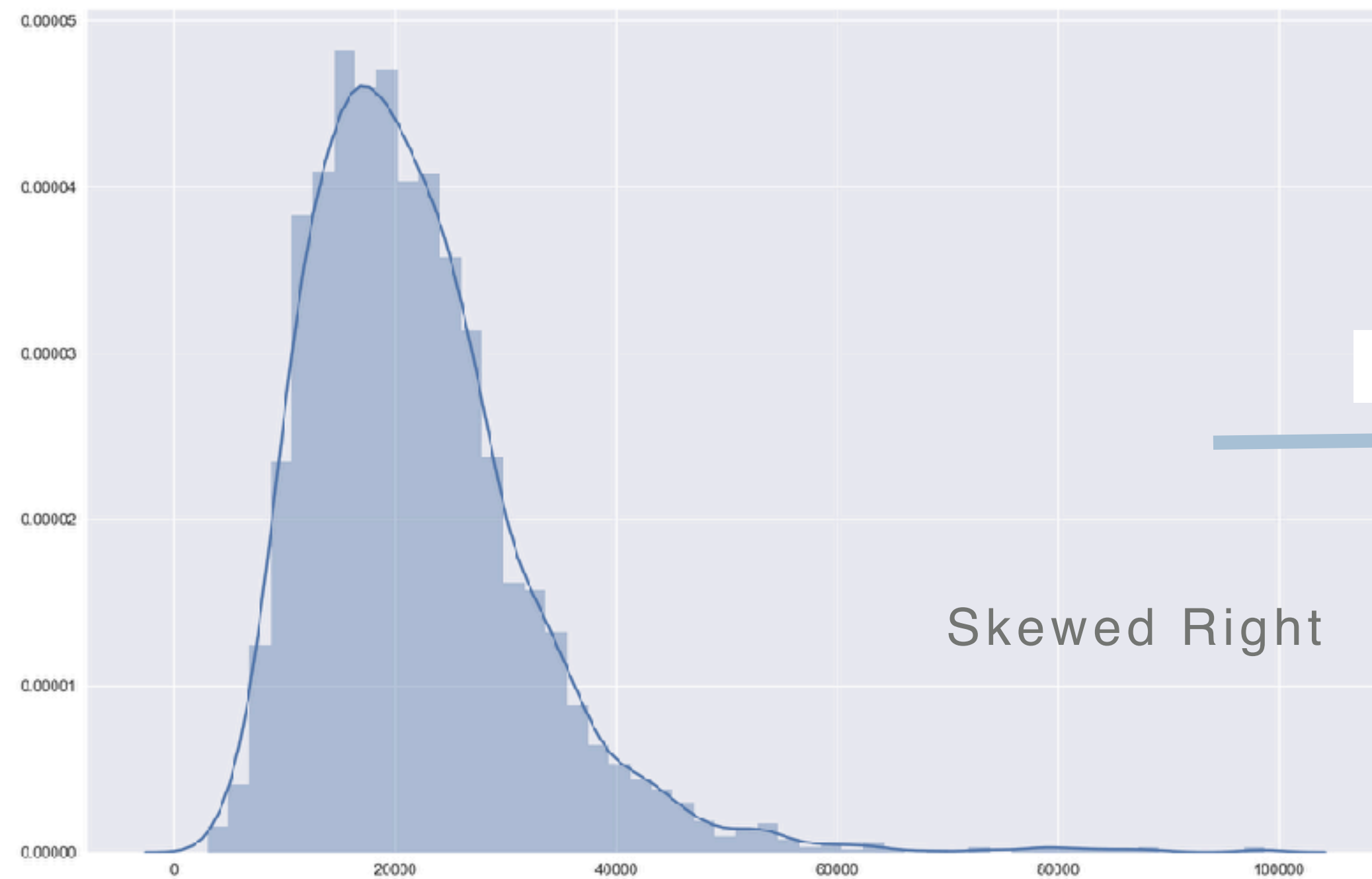
# CHI-Square Test



- The presence of medical history 1, 6 and 7 have high correlation with > 1 encounter per year
- P-values < 0.01

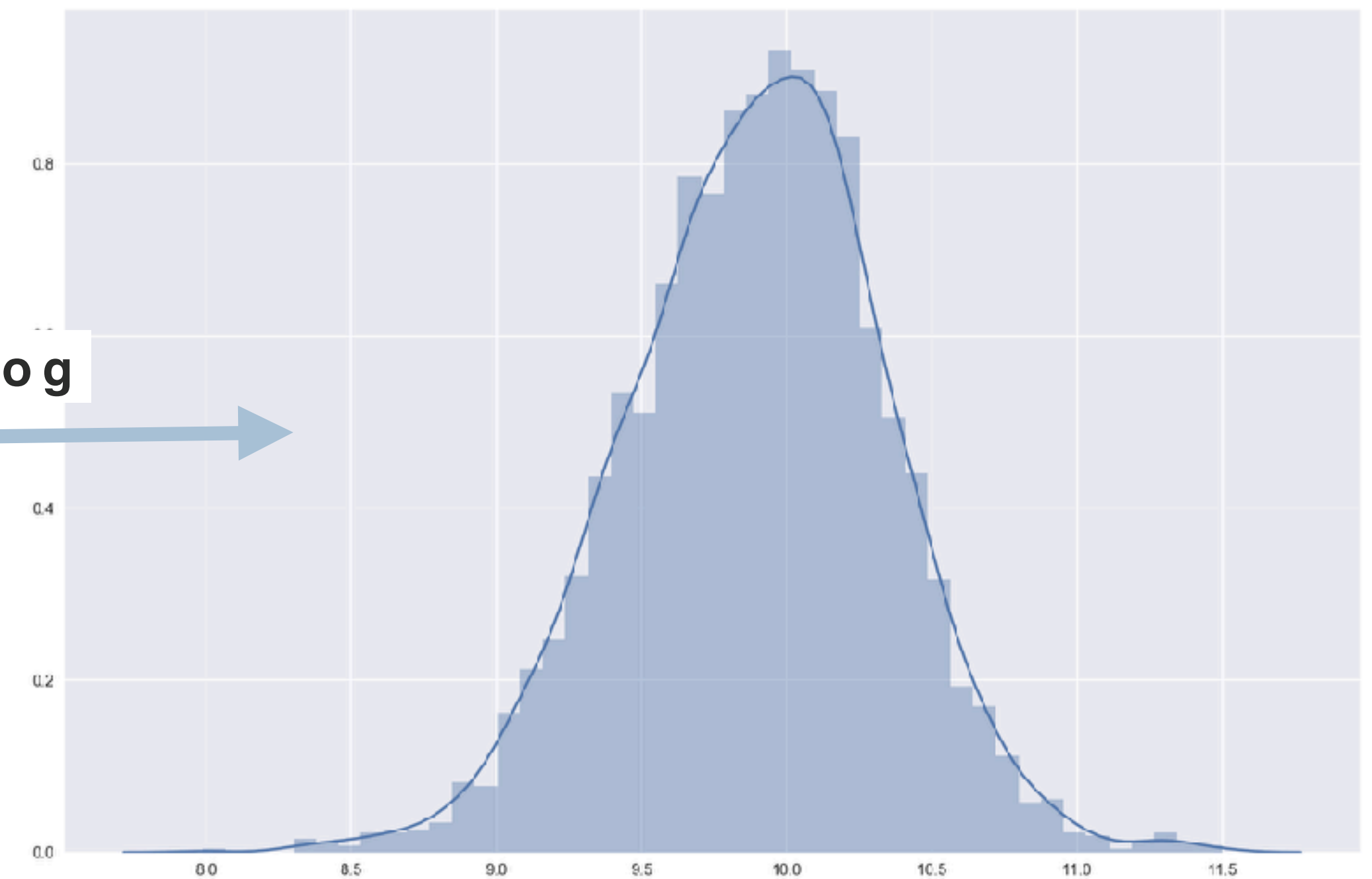
# Target Variable

Distribution



Skewed Right

Natural Log



$\frac{\text{Total Bill / year}}{\text{Number of Encounters / year}} \text{ (Log)}$

To ensure approximate normality



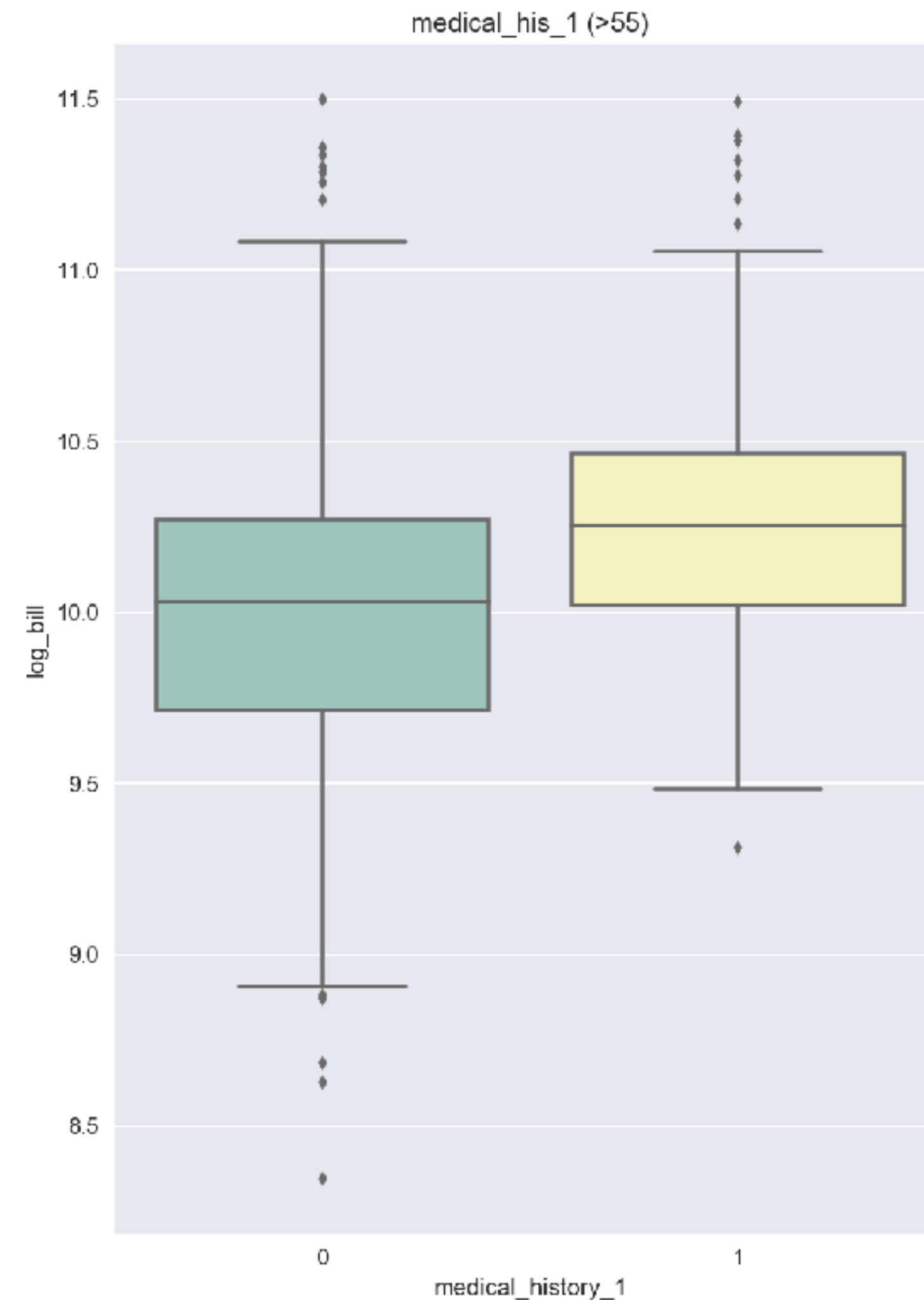
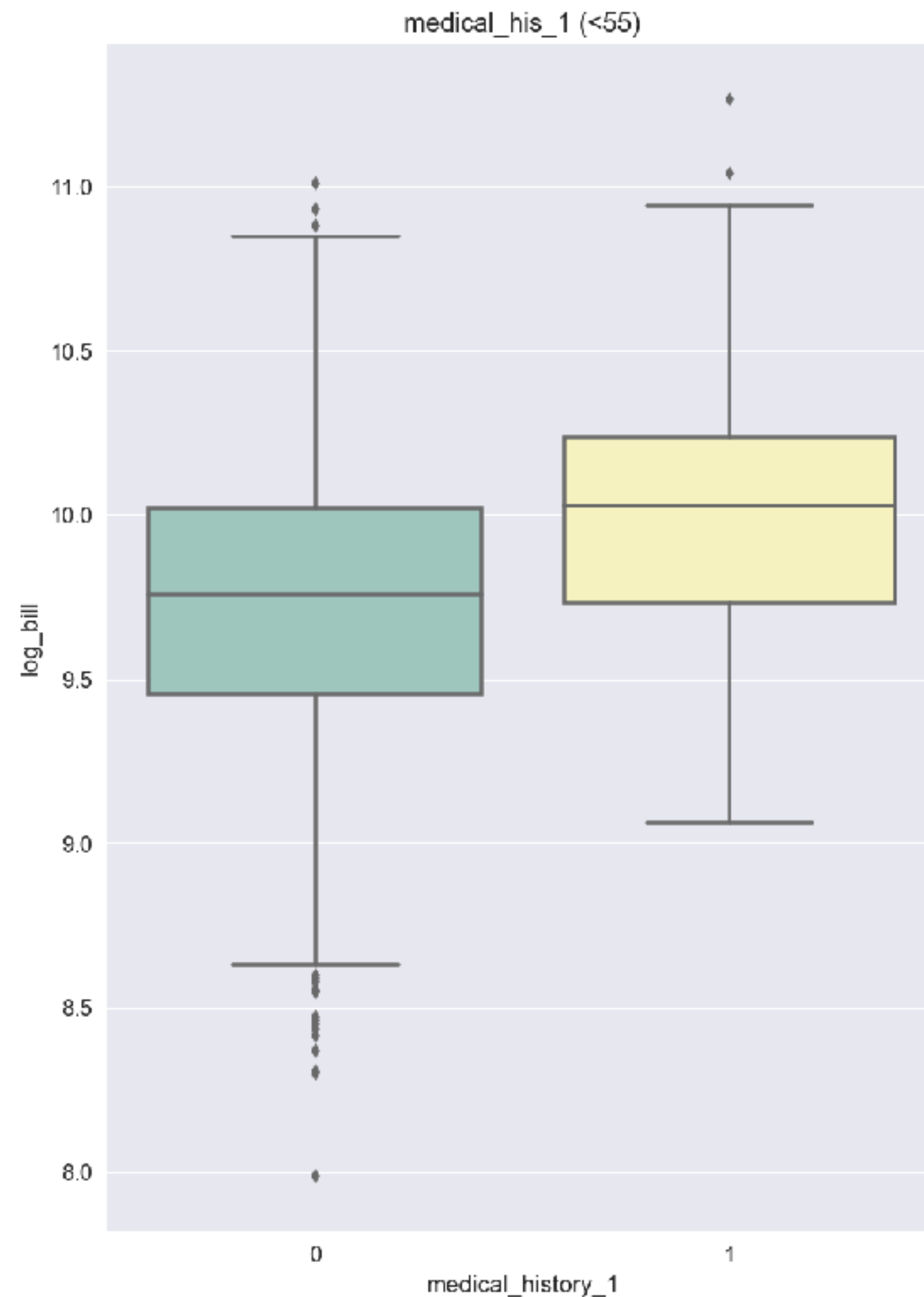
# Medical History

		Count	Avg Total Bill (per enc per yr)	
			Mean	P-Value
Total		3,314		
Number of Med History	0	795	19,497	<b>&lt;0.01*</b>
	1	1,258	21,463	
	2	800	24,696	
	3	310	26,532	
	4	47	31,191	
	5	7	36,853	
Med Hist 1		<b>562</b>	26,850	<b>&lt;0.01*</b>
Med Hist 2		953	22,357	0.086
Med Hist 3		459	22,205	0.023
Med Hist 4		175	21,480	0.943
Med Hist 5		196	23,129	0.018
Med Hist 6		<b>839</b>	24,175	<b>&lt;0.01*</b>
Med Hist 7		842	22,484	0.040

• Not tested for homogeneity and normality assumptions due to time constraint. One-way ANOVA is used

# Medical History 1

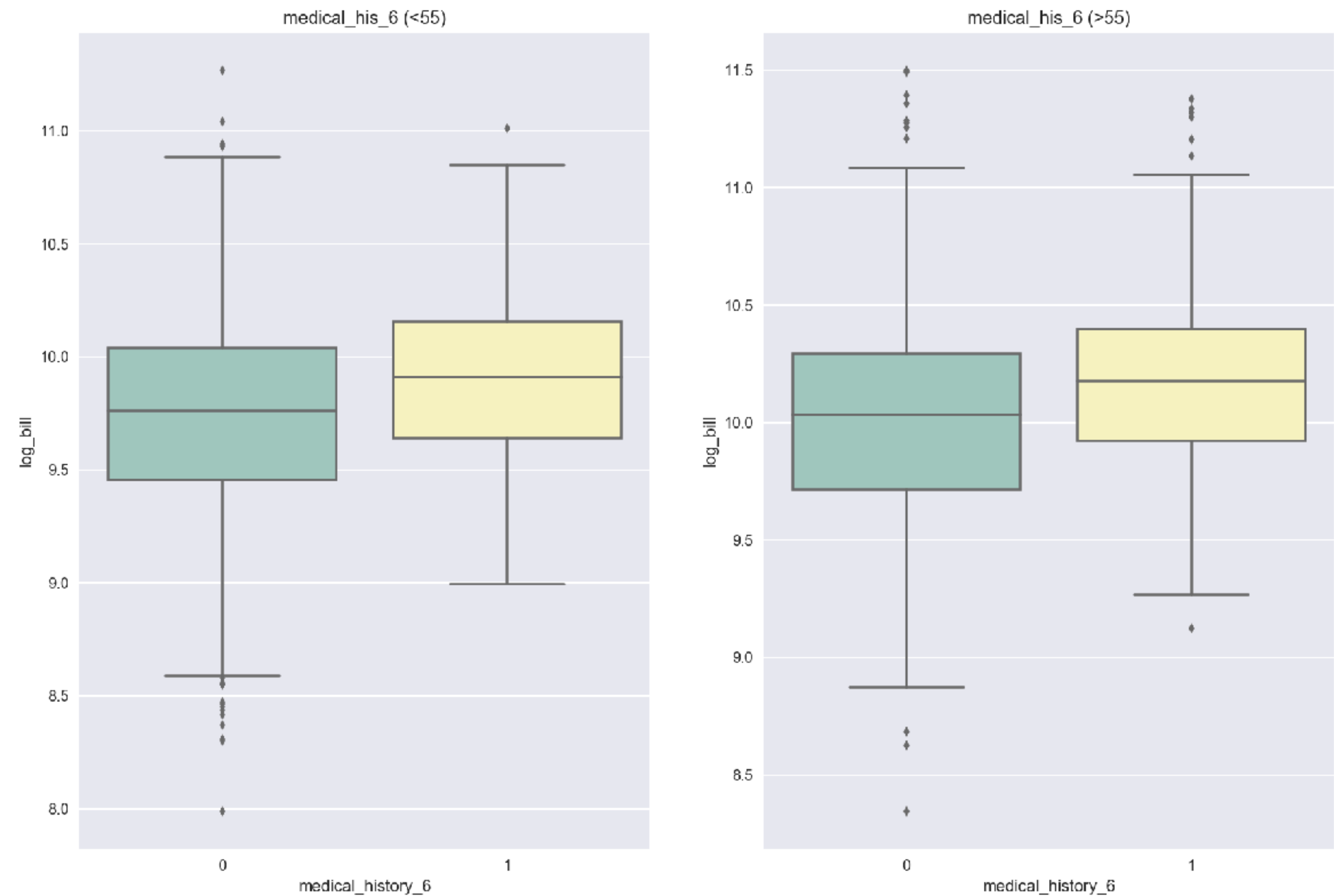
Total Count: 562 / 3,000 (18.7%)



- 18.7% of patients with medical history 1 contributes 21.7% of the avg total bills.
- P-Value < 0.01 for both age groups

# Medical History 6

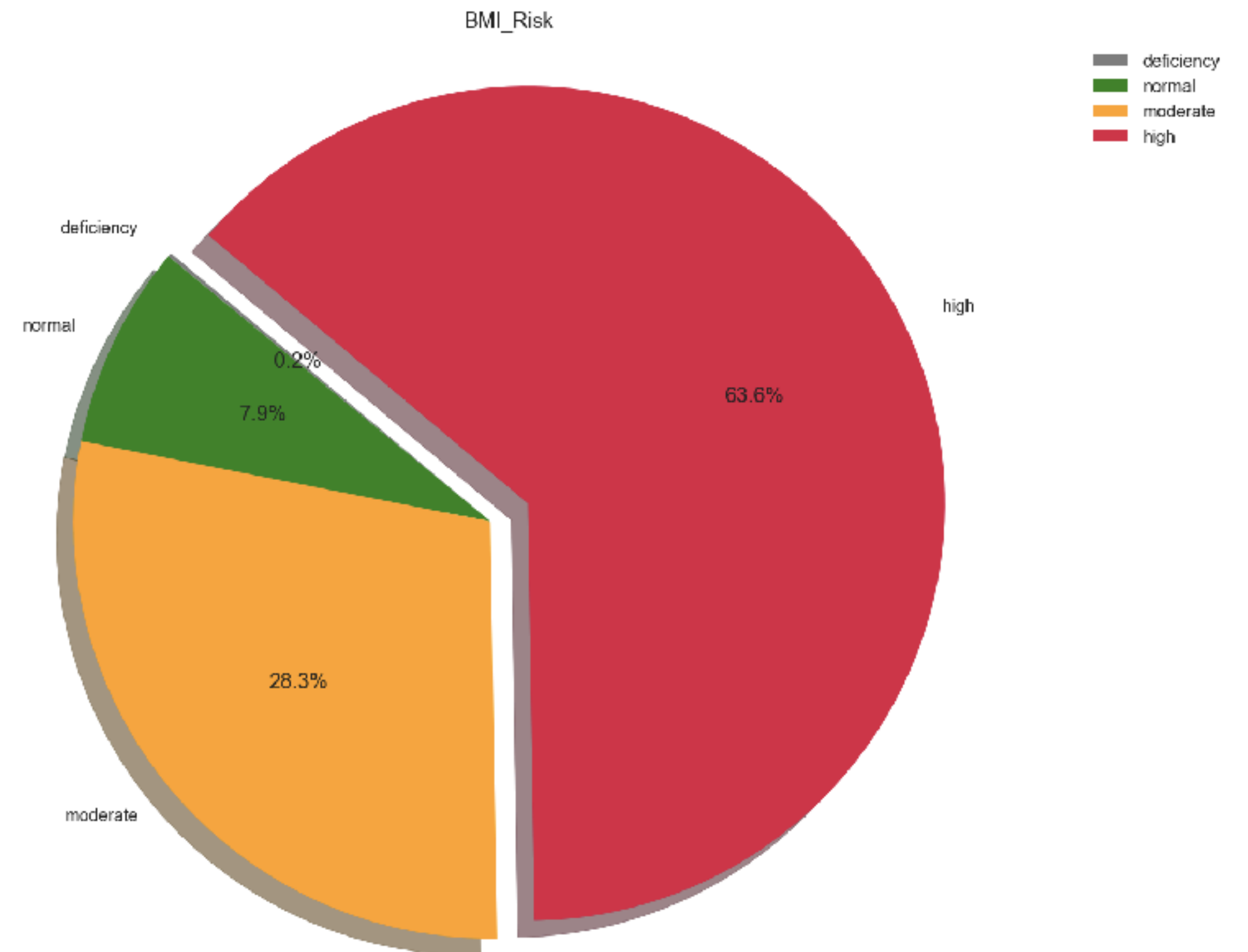
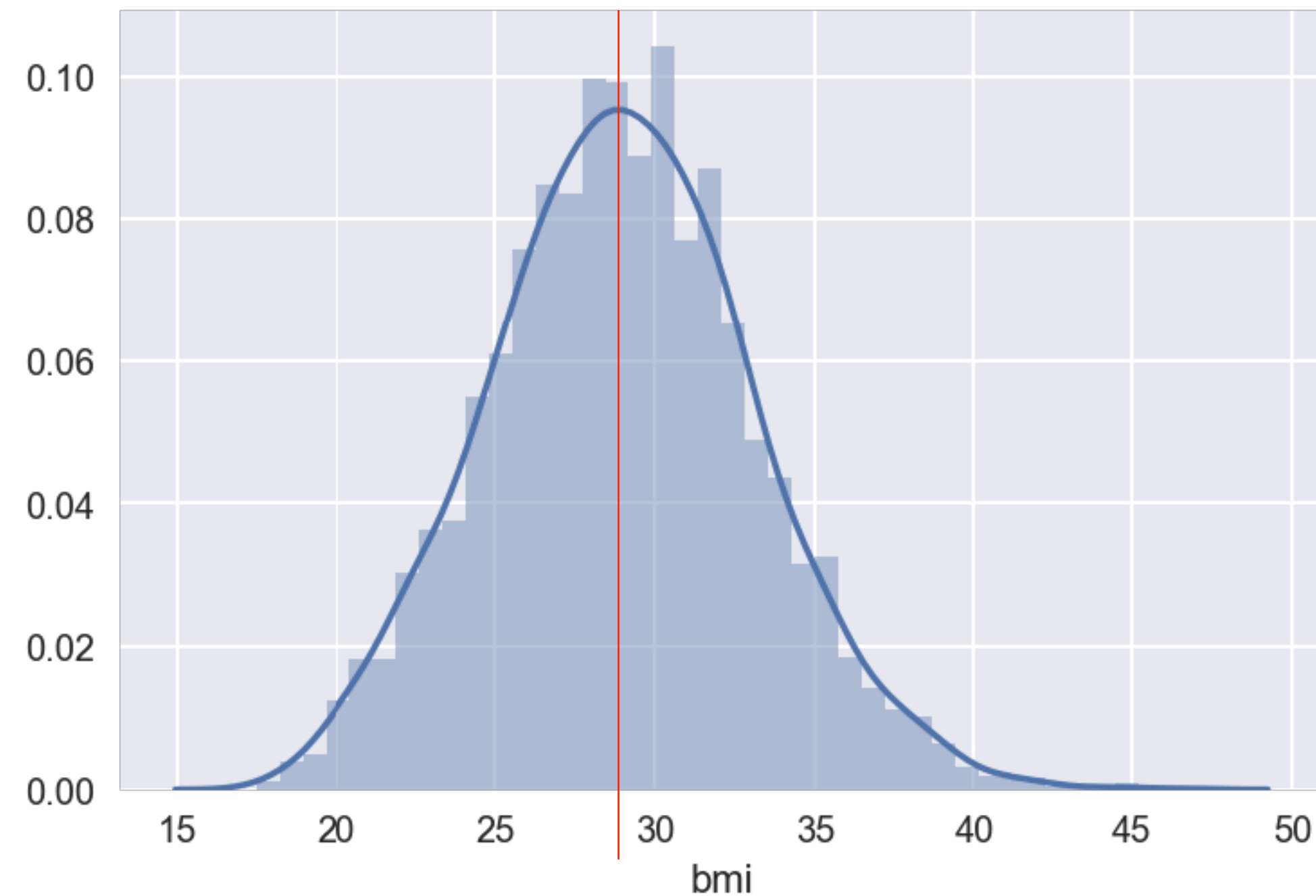
Count: 839/ 3,000 (28.0%)



- 28% of patients with medical history 6 contributes 29.44% of the avg total bills.
- P-Value < 0.01 for both age groups

# Body Mass Index (BMI)

$$BMI = \frac{Weight (kg)}{[Height(m)]^2}$$

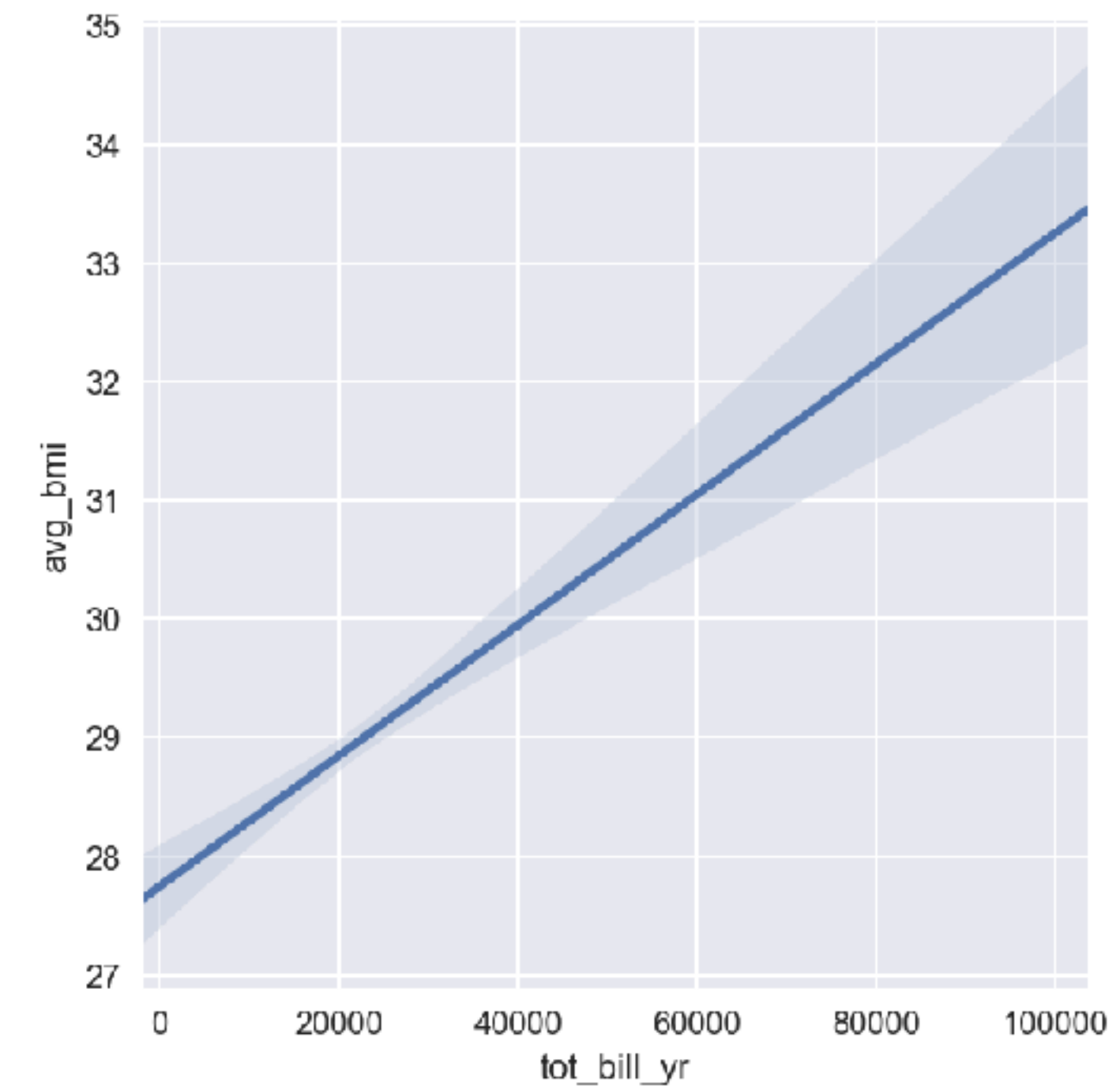


- Mean BMI = 28.95 (**high risk**)
- **63.6%** patient population have **high BMI risk**

# Body Mass Index (BMI)

$$BMI = \frac{Weight (kg)}{[Height(m)]^2}$$

		Total Bill (/enc/yr)	
		Mean	P-Value
BMI Risk Level	Deficiency	11,200	<0.01*
	Normal	19,389	
	Moderate	20,774	
	High	22,651	



R: 0.140, P-value: <0.01\*

- There is significance between BMI Risk and Avg total bill per year at 0.01 significance level
- Positive correlation between BMI and avg total bill



# Model & Interpretation

Explainability matters .



# Variance Inflation Factor

To check for multicollinearity between variables

	VIF Factor	features
0	1.0	medical_history_1
1	1.0	medical_history_6
2	2.0	bmi_high
3	1.0	age_grp
4	1.0	foreigner
5	1.0	malay

No variables have a VIF of more than 5

# Linear Regression Model

$$Y = \beta_0 + \beta_1 \text{ medical history 1} + \beta_2 \text{ medical history 6} + \beta_3 \text{ high bmi risk} \\ + \beta_4 \text{ age} > 55 + \beta_5 \text{ foreigner} + \beta_6 \text{ malay}$$

	Parameter	Estimate	Standard Error	P-Value
Model 2	Intercept	9.51	0.012	<0.01*
	Medical History 1	0.26	0.016	<0.01*
	Medical History 6	0.17	0.014	<0.01*
	High BMI risk	0.11	0.012	<0.01*
	Age > 55	0.25	0.012	<0.01*
	Foreigner	0.68	0.028	<0.01*
	Malay	0.40	0.015	<0.01*

R-squared = 0.415

# Linear Regression Model

$$Y = \beta_0 + \beta_1 \text{ medical history 1} + \beta_2 \text{ medical history 6} + \beta_3 \text{ high bmi risk} \\ + \beta_4 \text{ age} > 55 + \beta_5 \text{ foreigner} + \beta_6 \text{ malay}$$

## Interpretation:

Average total bill per year for patients with condition X and have/is

- medical history 1 is ~ **29.11** % higher when compared to patients without
- medical history 6 is ~ **18.71** % higher when compared to patients without
- high BMI risk is ~ **11.37** % higher when compared to patients doesn't
- age > 55 years old is ~ **27.86** % higher when compared to patients < 55
- a foreigner is ~ **96.83** % higher when compared to patients who are Singaporeans or PR
- malay is ~ **49.05** % higher when compared to patients of other races

# Use Case Discussion

Identify patient population at risk of higher admission cost for condition X for intervention

Findings suggest that admission cost per year are significantly affected by

1. Number of admitted encounters,
2. Medical history,
3. Body Mass Index
4. Age
5. Demographics

# Use Case Discussion

Identify patient population at risk of higher admission cost for condition X for intervention

Findings suggest that admission cost per year are significantly affected by

1. Number of admitted encounters

2. Medical history

3. Body Mass Index

4. Age

5. Demographics

- Cost per year increases with frequent admissions
- Highly correlated to medical history
- Consider continuity of care for these patients after discharge to prevent readmission

# Use Case Discussion

Identify patient population at risk of higher admission cost for condition X for intervention

Findings suggest that admission cost per year are significantly affected by

1. Number of admitted encounters

2. Medical history

3. Body Mass Index

4. Age

5. Demographics

- Increases complexity, more complications, more need for healthcare attention
- Preventive care for patients without existing conditions through regular screenings
- Medical History 1 and 6 have bigger impact than others
- Not enough information

# Use Case Discussion

Identify patient population at risk of higher admission cost for condition X for intervention

Findings suggest that admission cost per year are significantly affected by

1. Number of admitted encounters

2. Medical history

3. Body Mass Index

4. Age

5. Demographics

- Not surprising that patients with high health risk to have more complications and require more interventions
- Better manage patient's BMI by encouraging them to participate in health programmes
- Frequent health screenings



# Use Case Discussion

Identify patient population at risk of higher admission cost for condition X for intervention

Findings suggest that admission cost per year are significantly affected by

1. Number of admitted encounters

2. Medical history

3. Body Mass Index

4. Age

5. Demographics

- Expected that elderly patients would need more medical attention than younger population
- Non-Singaporeans have less subsidies
- Non-modifiable variables, but allows healthcare providers/ministry of health to target population for policy planning and interventions



## Given more time...

- Adjusted for confounders (demographics group) rather than using it as a factor
- Test for homogeneity and normality assumptions
- Look into admission level data for more use cases
  - Symptoms affect cost per admission
  - Useful for pharmaceutical companies
- Interactive dashboards to show how each variable varies
- Improve on visualisation and story-telling