

LUNG CANCER ANALYSIS

by sejal jha

What is the average age of the patients in the dataset?

```
SELECT AVG(age)  
FROM lung_cancer
```

	AVG(age)
	62.6731



HOW MANY PATIENTS ARE MALE VS FEMALE?

```
SELECT count(gender) as count_of_gender , gender  
From lung_cancer  
group by (gender)
```

count_of_gender	gender
162	M
147	F

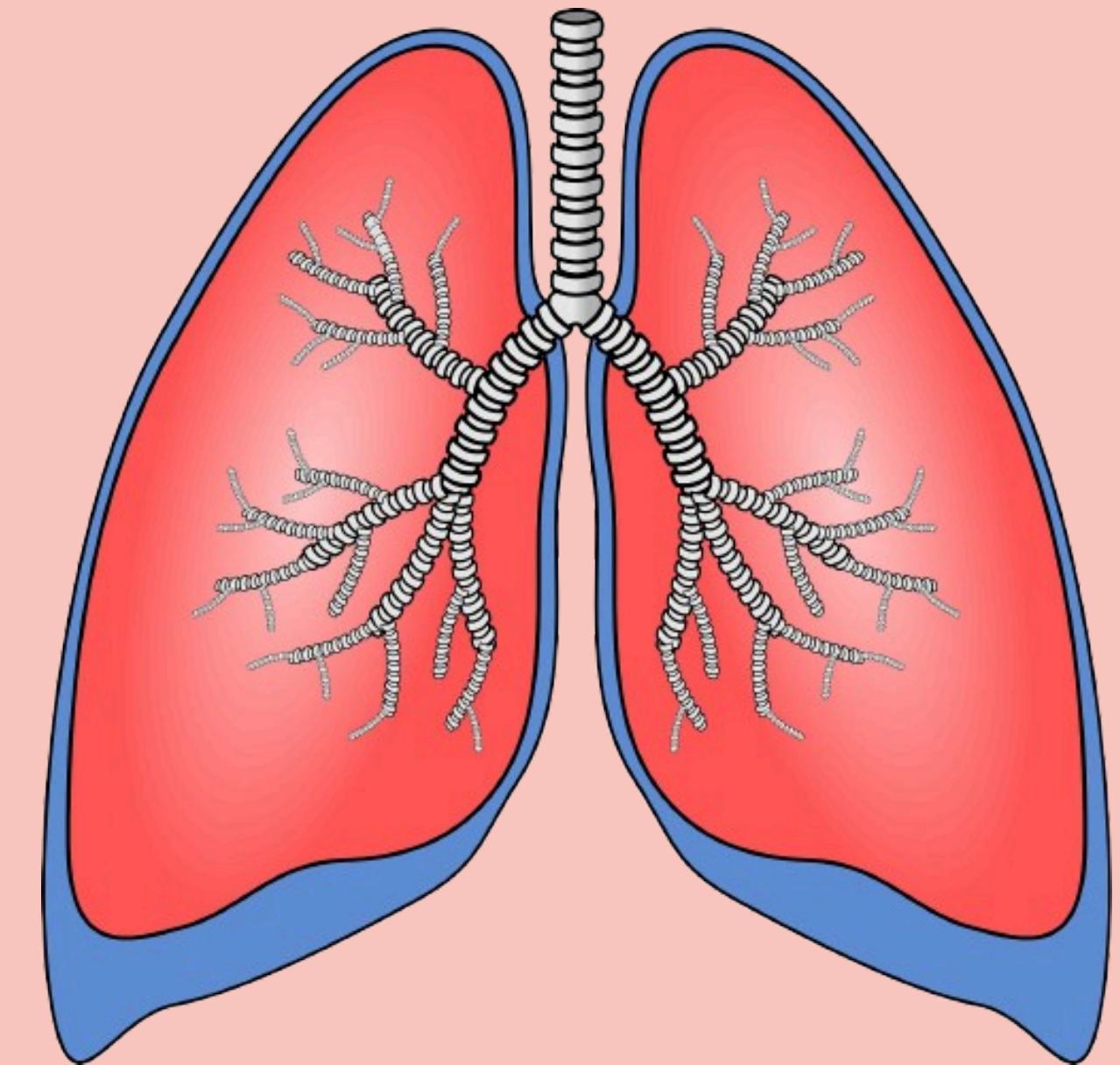
WHAT IS THE PROPORTION OF PATIENTS DIAGNOSED WITH LUNG CANCER VS THOSE NOT DIAGNOSED?

```
SELECT count(lung_cancer) as count_of_lung_cancer, lung_cancer  
FROM lung_cancer  
group by (lung_cancer)
```

count_of_lung_cancer	lung_cancer
270	YES
39	NO

WHAT IS THE AVERAGE AGE OF SMOKERS COMPARED TO NON-SMOKERS?

```
+ • SELECT avg(age) ,  
5   CASE  
5     WHEN smoking = 1 THEN 'Yes'  
7     WHEN smoking = 2 THEN 'No'  
8     ELSE 'Unknown'  
9   END AS Smoking_Status  
9 FROM lung_cancer  
L group by (smoking);
```



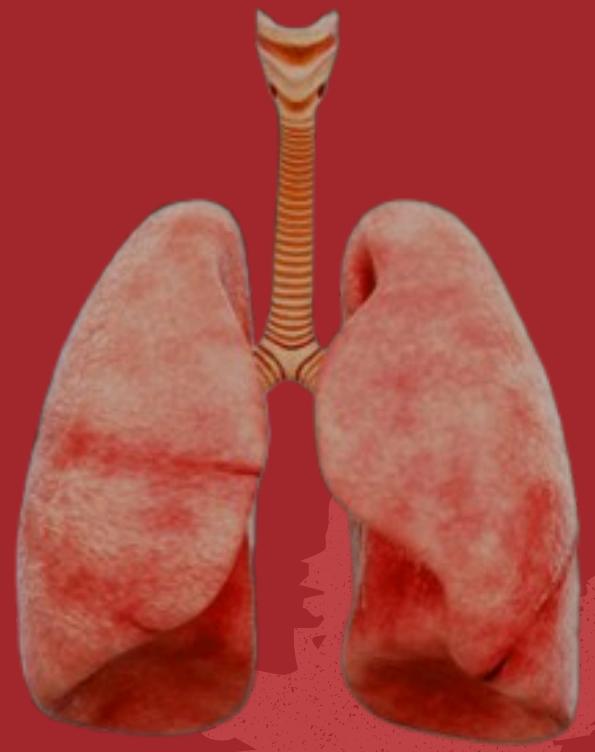
avg(age)	Smoking_Status
63.4593	Yes
62.0632	No

WHAT PERCENTAGE OF LUNG CANCER PATIENTS ARE SMOKERS?

- SELECT

```
100 * count(*) / (SELECT count(*) FROM lung_cancer WHERE smoking = 1) AS Percentage_of_smokers  
FROM lung_cancer  
WHERE smoking = 1 AND lung_cancer = 'YES';
```

Percentage_of_smokers
85.1852



AMONG LUNG CANCER PATIENTS, HOW MANY REPORT YELLOW FINGERS?

```
SELECT count(yellow_fingers) , lung_cancer  
FROM lung_cancer  
where lung_cancer = 'Yes' ;
```

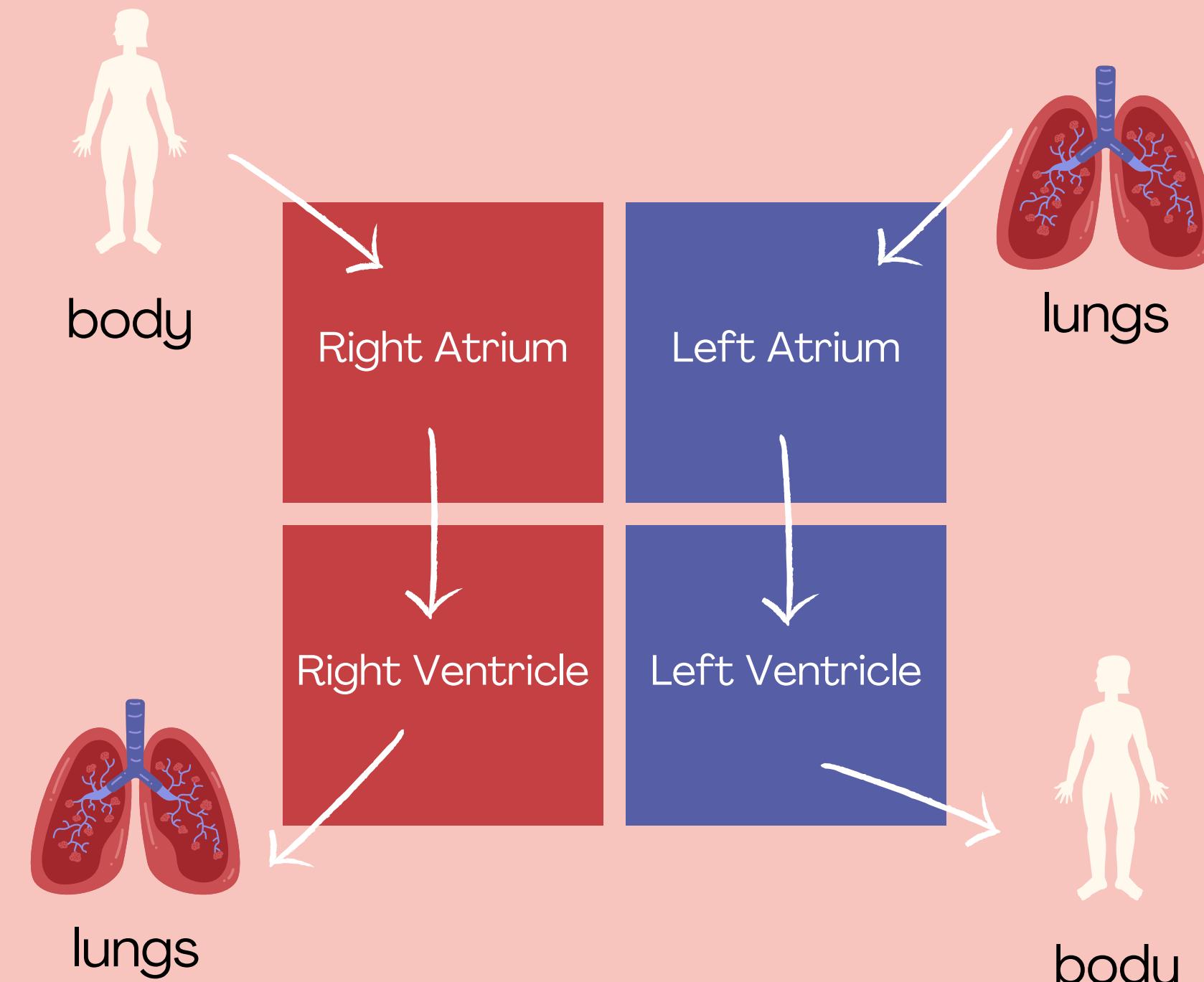
count(yellow_fingers)	lung_cancer
270	YES



WHAT IS THE DISTRIBUTION OF LUNG CANCER PATIENTS BY GENDER AND SMOKING STATUS?

```
SELECT gender , count(gender) as Patient_count ,  
CASE  
    WHEN smoking = 1 THEN 'YES'  
    WHEN smoking = 2 THEN 'NO'  
    ELSE 'Unknown'  
END as Smokers  
FROM lung_cancer  
where lung_cancer = 'yes'
```

gender	Patient_count	Smokers
M	60	YES
M	85	NO
F	55	YES
F	70	NO



AMONG PATIENTS REPORTING CHEST PAIN, HOW MANY HAVE LUNG CANCER?

```
SELECT count(lung_cancer) as count_of_lung_cancer , chest_pain, lung_cancer  
FROM lung_cancer  
Where chest_pain = '1' and lung_cancer = 'yes'  
group by lung_cancer , chest_pain;
```

count_of_lung_cancer	chest_pain	lung_cancer
110	1	YES



FIND THE AVERAGE AGE OF PATIENTS WITH LUNG CANCER WHO DO NOT SMOKE.

```
SELECT avg(age) , lung_cancer , smoking  
FROM lung_cancer  
where lung_cancer = 'yes' and smoking = '2'
```

avg(age)	lung_cancer	smoking
62.3419	YES	2

