



# COVID-19

...

*Hokage Heroes*

# TABLE OF CONTENTS ...

**1** Problem statement

**2** Dataset and Dashboard

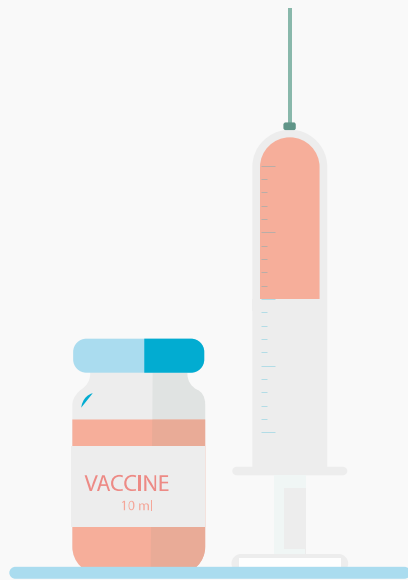
**3** Pig Question



## INTRODUCTION

COVID-19 has presented us with a new way of life, forcing us to adjust to the reality of living with a virus that has affected every corner of the globe.





## Problem statement

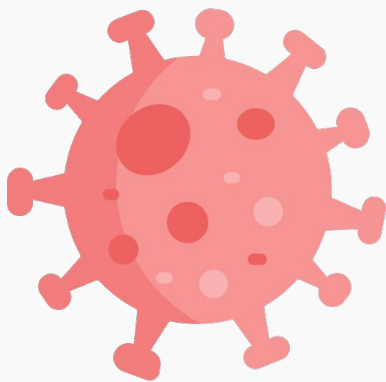
The patient is in high risk or not?



# Problem Statement

The main goal of this project is to build a 10 Pig script that show us the a Covid-19 patient's current symptom, status, and medical history to help us predict whether the patient is in high risk or not.





## Dataset

COVID-19 patient's symptoms, status, and medical history.





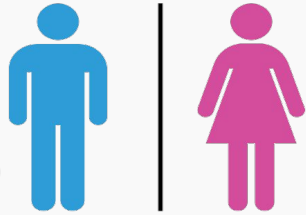
## About the Dataset ...

This dataset consists of 21 unique features and 1,048,576 unique patients.









In the Boolean features, 1 means "yes" and 2 means "no" 1,1 Female ,2 male.

	USMER	MEDICAL_UNIT	SEX	PATIENT_TYPE	DATE_DIED	INTUBED	PNEUMONIA	AGE	PREGNANT	DIABETES	...	ASTHMA	INMSUPR	HIPERTENSION	OTHER_DISEASE	CARDI
0	2		1	1	1	03/05/2020	97	1	65	2	2	...	2	2	1	2
1	2		1	2	1	03/06/2020	97	1	72	97	2	...	2	2	1	2
2	2		1	2	2	09/06/2020	1	2	55	97	1	...	2	2	2	2
3	2		1	1	1	12/06/2020	97	2	53	2	2	...	2	2	2	2
4	2		1	2	1	21/06/2020	97	2	68	97	1	...	2	2	1	2
5 rows x 21 columns																

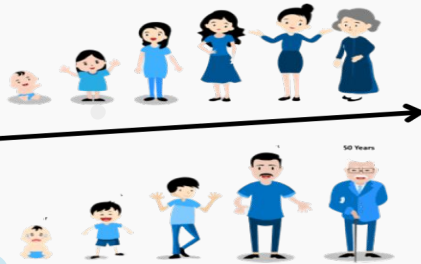
# Columns ...



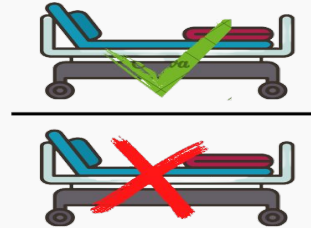
• Sex

Degree 1 	Degree 2 	Degree 3 	Degree 4 
			

• Classification

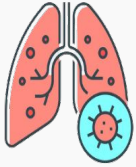


• Age



• Patient type





- **Pneumonia**



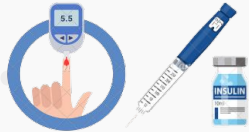
- **Copd**



- **Pregnancy**



- **Asthma**

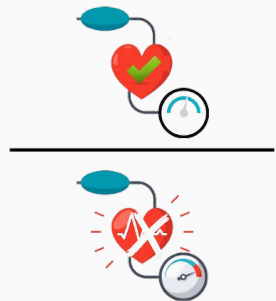


- **Diabetes**



- **Inmsupr**





- Hypertension



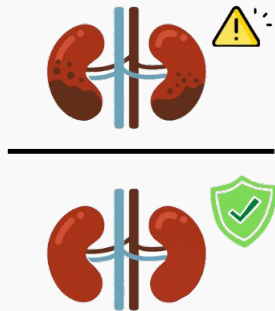
- Other disease



- Cardiovascular



- Obesity



- Renal chronic



- Tobacco

First level	Second level	Third level
		

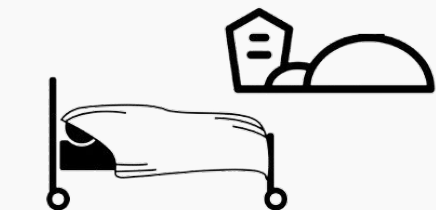
• USMR



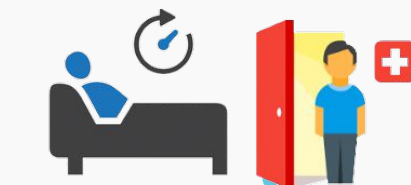
• ICU



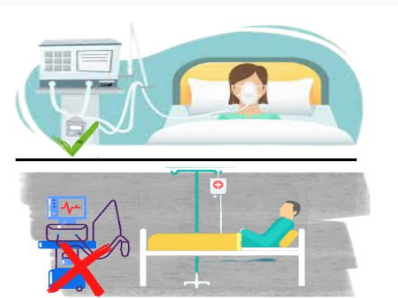
• Medical unit



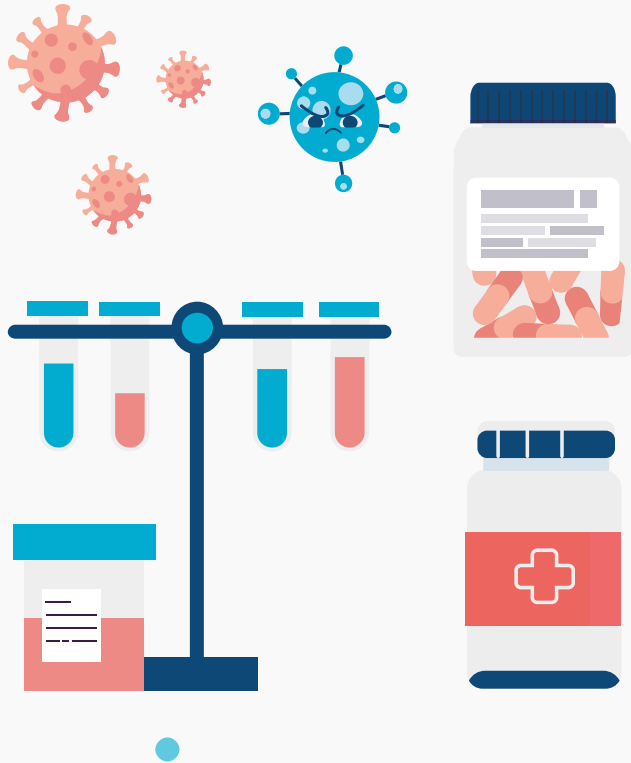
• Death



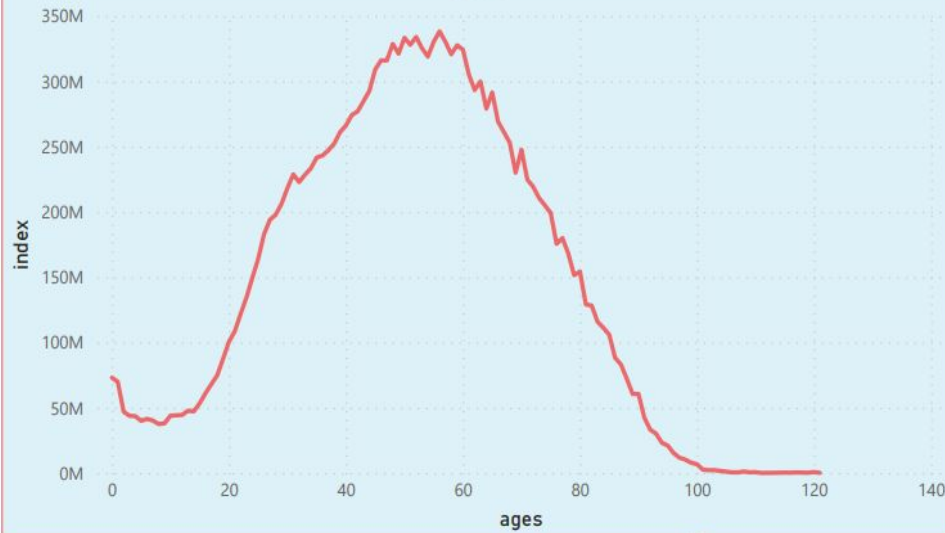
• Intubed



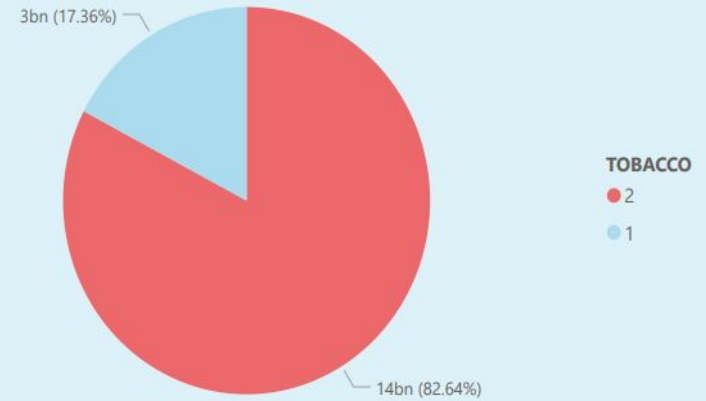
# Dashboard ...



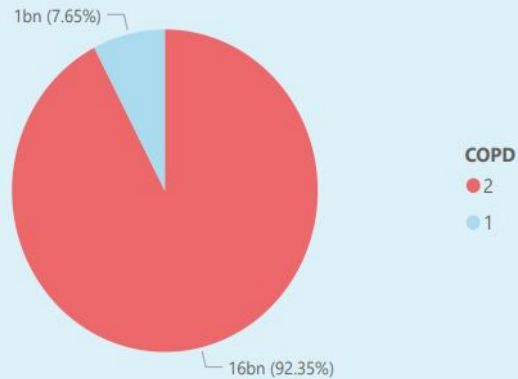
# AGE



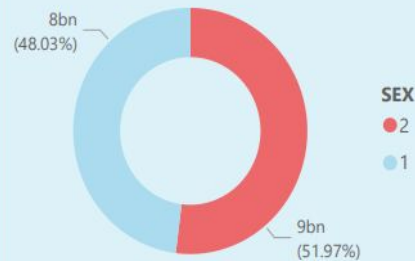
# People if Smoking or not



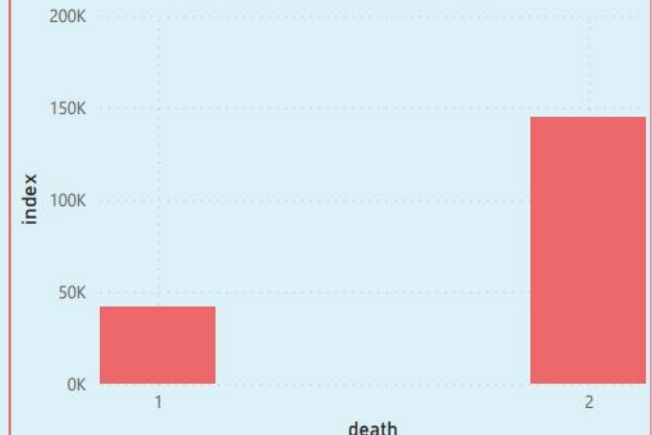
# COPD Percentage

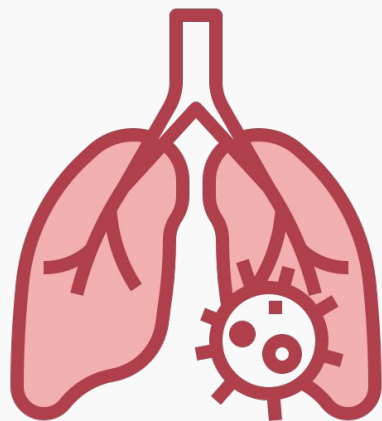


# SEX Percentage



# People if still in Life or not





## Pig Questions



# Question.1

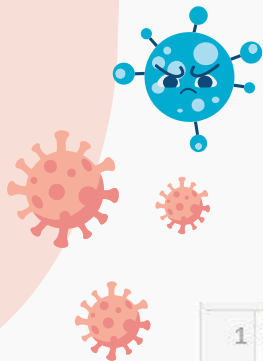
Do you have a covid-19?

```
1 Load_data = LOAD '/user/maria_dev/Covie_19/Covid_df.txt'
2 USING PigStorage(',') as (USMER:int,MEDICAL_UNIT:int,SEX:int,PATIENT_TYPE:chararray,PNEUMONIA:int,
3 AGE:int,PREGNANT:int,DIABETES:int,COPD:int,ASTHMA:int,INMSUPR:int,HIPERTENSION:int,OTHER_DISEASE:int,
4 CARDIOVASCULAR:int,OBESITY:int,RENAL_CHRONIC:int,TOBACCO:int,CLASIFFICATION_FINAL:int,DEATH:int);
5 data_foreach = FOREACH Load_data GENERATE MEDICAL_UNIT,SEX,PATIENT_TYPE,AGE,CLASIFFICATION_FINAL;
6 positive_test = FILTER data_foreach BY CLASIFFICATION_FINAL<=3 ;
```

▼ Results

(12,1,1,59,1)  
(12,1,1,40,1)  
(12,1,1,38,1)  
(12,2,1,68,1)  
(12,2,1,23,1)  
(12,1,1,40,1)  
(12,1,1,25,1)  
(12,1,1,51,1)  
(12,1,1,41,1)  
(12,1,1,50,1)  
(12,2,1,33,1)  
(12,1,1,40,1)  
(12,2,1,64,1)  
(12,1,1,75,1)  
(12,2,1,80,1)  
(12,2,1,23,1)  
(12,1,1,44,1)  
(12,2,1,50,1)  
(12,1,1,55,1)  
(12,1,1,26,1)  
(12,2,1,67,1)  
(12,1,1,22,1)  
(12,1,1,17,1)  
(12,2,1,58,1)  
(12,2,1,50,1)  
(12,1,1,50,1)  
(12,2,1,35,1)  
(12,2,1,47,1)  
(12,1,1,51,1)





## Question.2 ...

Are they still in life?

```
1 Load_data = LOAD '/user/maria_dev/Covie_19/Covid_df.txt'  
2 USING PigStorage(',') as (USMER:int,MEDICAL_UNIT:int,SEX:int,PATIENT_TYPE:chararray,  
3 PNEUMONIA:int,AGE:int,PREGNANT:int,DIABETES:int,COPD:int,ASTHMA:int,INMSUPR:int,  
4 HIPERTENSION:int,OTHER_DISEASE:int,CARDIOVASCULAR:int,OBESITY:int,RENAL_CHRONIC:int,  
5 TOBACCO:int,CLASIFFICATION_FINAL:int,DEATH:int);  
6 data_foreach = FOREACH Load_data GENERATE MEDICAL_UNIT,SEX,PATIENT_TYPE,AGE,DEATH ;  
7 ifstill_alive= FILTER data_foreach BY DEATH ==2;  
8 G_sex= GROUP ifstill_alive BY SEX;  
9 DUMP G_sex;
```

▼ Results

 Download

```
(1,{(8,1,2,88,2),(12,1,1,34,2),(4,1,1,24,2),(12,1,1,65,2),(12,1,1,53,2),(4,1,1,25,2),(12,1,1,40,2),(12,1,1,34,2),(12,1,1,83,2),(8,1,1,  
(2,{(8,2,2,57,2),(1,2,2,24,2),(1,2,2,30,2),(1,2,1,55,2),(1,2,1,61,2),(1,2,1,54,2),(1,2,2,59,2),(1,2,1,30,2),(1,2,1,45,2),(1,2,1,24,2),
```





## Question.3 ...

Does anybody from them have diabetes?

```
1 Load_data = LOAD '/user/maria_dev/Covie_19/Covid_df.txt'
2 USING PigStorage(',') as (USMER:int,MEDICAL_UNIT:int,SEX:int,PATIENT_TYPE:chararray,
3 PNEUMONIA:int,AGE:int,PREGNANT:int,DIABETES:int,COPD:int,ASTHMA:int,INMSUPR:int,
4 HIPERTENSION:int,OTHER_DISEASE:int,CARDIOVASCULAR:int,OBESITY:int,RENAL_CHRONIC:int,
5 TOBACCO:int,CLASIFFICATION_FINAL:int,DEATH:int);
6 data_foreach = FOREACH Load_data GENERATE SEX,PATIENT_TYPE,AGE,DIABETES;
7 has_diabete= FILTER data_foreach BY DIABETES == 1;
DUMP has_diabete;
```

▼ Results

(2,2,55,1)  
(2,1,68,1)  
(1,1,64,1)  
(1,2,37,1)  
(1,1,48,1)  
(2,2,59,1)  
(1,1,57,1)  
(1,2,48,1)  
(2,2,30,1)  
(2,1,62,1)  
(2,1,56,1)  
(1,1,53,1)  
(1,1,60,1)  
(1,1,47,1)  
(1,1,53,1)  
(1,1,58,1)  
(1,1,73,1)  
(1,1,17,1)  
(2,1,74,1)  
(1,1,52,1)  
(1,2,47,1)  
(2,2,59,1)  
(2,2,43,1)  
(1,2,81,1)  
(1,2,54,1)  
(1,2,56,1)  
(1,2,67,1)  
(1,2,88,1)  
(1,2,73,1)  
(1,2,87,1)

## Question.4 ...

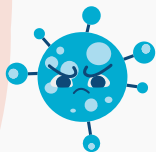
Does anybody from them have blood pressure?

```
1 Load_data = LOAD '/user/maria_dev/Covie_19/Covid_df.txt'
2 USING PigStorage(',') as (USMER:int,MEDICAL_UNIT:int,SEX:int,PATIENT_TYPE:chararray,
3 PNEUMONIA:int,AGE:int,PREGNANT:int,DIABETES:int,COPD:int,ASTHMA:int,INMSUPR:int,
4 HIPERTENSION:int,OTHER_DISEASE:int,CARDIOVASCULAR:int,OBESITY:int,RENAL_CHRONIC:int,
5 TOBACCO:int,CLASIFFICATION_FINAL:int,DEATH:int);
6 data_foreach = FOREACH Load_data GENERATE AGE,HIPERTENSION,CLASIFFICATION_FINAL;
7 has_bloodPressure= FILTER data_foreach BY HIPERTENSION == 1;
8 age= FILTER has_bloodPressure by AGE > 50;
9 DUMP age;
```

▼ Results

(65,1,3)  
(72,1,5)  
(68,1,3)  
(64,1,3)  
(80,1,3)  
(62,1,7)  
(58,1,7)  
(66,1,7)  
(56,1,7)  
(66,1,7)  
(80,1,7)  
(53,1,3)  
(62,1,7)  
(53,1,7)  
(53,1,7)  
(81,1,7)  
(67,1,7)  
(58,1,7)  
(85,1,7)  
(74,1,7)  
(80,1,7)  
(74,1,7)  
(52,1,7)  
(78,1,7)  
(81,1,7)  
(75,1,7)  
(59,1,6)  
(81,1,7)  
(56,1,7)





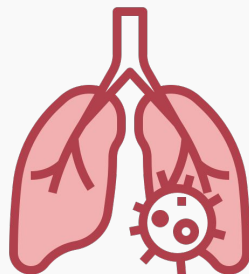
## Question.5 ...

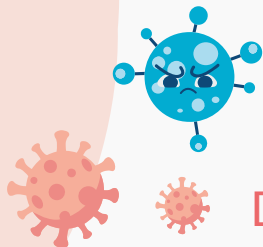
Do you have another disease?

```
1 Load_data = LOAD '/user/maria_dev/CovidProject/Covid_df.txt'
2 USING PigStorage(',') as (USMER:int,MEDICAL_UNIT:int,SEX:int,PATIENT_TYPE:chararray,
3 PNEUMONIA:int,AGE:int,PREGNANT:int,DIABETES:int,COPD:int,ASTHMA:int,INMSUPR:int,HIPERTENSION:int,
4 OTHER_DISEASE:int,CARDIOVASCULAR:int,OBESITY:int,RENAL_CHRONIC:int,TOBACCO:int,CLASIFFICATION_FINAL:int,DEATH:int);
5 data_foreach = FOREACH Load_data GENERATE MEDICAL_UNIT,SEX,AGE,CLASIFFICATION_FINAL,OTHER_DISEASE;
6 other = FILTER data_foreach BY OTHER_DISEASE == 1 ;
7 DUMP other;
```

▼ Results

```
(1,2,47,3,1)
(1,1,30,7,1)
(2,1,10,3,1)
(2,2,1,7,1)
(2,1,15,7,1)
(2,1,2,7,1)
(2,1,10,7,1)
(3,1,64,7,1)
(3,1,49,5,1)
(3,1,49,7,1)
(3,2,63,5,1)
(3,2,79,3,1)
(3,1,74,3,1)
(3,2,77,5,1)
(3,1,83,7,1)
(3,2,60,7,1)
(3,2,55,3,1)
(3,1,61,5,1)
(3,2,51,3,1)
(3,2,20,7,1)
(3,2,67,7,1)
(3,2,78,3,1)
(3,1,70,3,1)
(3,1,58,3,1)
(3,2,53,3,1)
(3,1,81,3,1)
(3,1,81,3,1)
(3,1,85,2,1)
(3,1,83,7,1)
```





## Question.6 ...

Does anybody from them have COPD and ASTHMA?



```
1 Load_data = LOAD '/user/maria_dev/Covie_19/Covid_df.txt'  
2 USING PigStorage(',') as (USMER:int,MEDICAL_UNIT:int,SEX:int,PATIENT_TYPE:chararray,  
3 PNEUMONIA:int,AGE:int,PREGNANT:int,DIABETES:int,COPD:int,ASTHMA:int,INMSUPR:int,  
4 HIPERTENSION:int,OTHER_DISEASE:int,CARDIOVASCULAR:int,OBESITY:int,RENAL_CHRONIC:int,  
5 TOBACCO:int,CLASIFFICATION_FINAL:int,DEATH:int);  
6 data_foreach = FOREACH Load_data GENERATE PATIENT_TYPE,AGE,COPD,ASTHMA,DEATH;  
7 has_copd= FILTER data_foreach BY COPD == 1;  
8 has_ASTHMA= FILTER has_copd BY ASTHMA == 1;
```

COPD: Chronic Obstructive Pulmonary



▼ Results

```
(2,82,1,1,1)  
(2,81,1,1,1)  
(2,84,1,1,1)  
(2,55,1,1,1)  
(2,69,1,1,1)  
(2,55,1,1,1)  
(2,42,1,1,1)  
(2,47,1,1,1)  
(2,75,1,1,1)  
(2,65,1,1,1)  
(2,54,1,1,1)  
(2,64,1,1,1)  
(2,72,1,1,1)  
(2,80,1,1,1)  
(2,64,1,1,1)  
(2,49,1,1,1)  
(2,96,1,1,1)  
(2,80,1,1,1)  
(2,76,1,1,1)  
(2,77,1,1,1)  
(2,77,1,1,1)  
(1,60,1,1,1)  
(2,71,1,1,1)  
(2,67,1,1,1)  
(2,77,1,1,1)  
(2,32,1,1,1)  
(2,74,1,1,1)  
(2,72,1,1,1)  
(2,67,1,1,1)
```

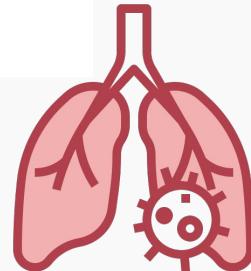
## Question.7 ...

Does anybody from them have been recovered from COVID-19 while the age is under 40 and has a ASTHMA?

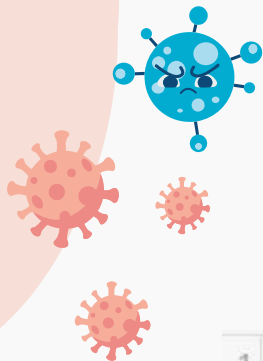
```
1 load_data = LOAD '/user/maria_dev/Covie_19/Covid_df.txt'
2 USING PigStorage(',') as (USMER:int,MEDICAL_UNIT:int,SEX:int,PATIENT_TYPE:chararray,
3 PNEUMONIA:int,AGE:int,PREGNANT:int,DIABETES:int,COPD:int,ASTHMA:int,INMSUPR:int,
4 HIPERTENSION:int,OTHER_DISEASE:int,CARDIOVASCULAR:int,OBESITY:int,RENAL_CHRONIC:int,
5 TOBACCO:int,CLASIFFICATION_FINAL:int,DEATH:int);
6 data_foreach = FOREACH Load_data GENERATE AGE,ASTHMA,DEATH ;
7 has_asthma= FILTER data_foreach BY ASTHMA == 1;
8 is_recovered= FILTER has_asthma BY DEATH == 2;
9 filterPatient = FILTER is_recovered BY AGE < 40;
10 DUMP filterPatient;
```

▼ Results

```
(25,1,2)
(24,1,2)
(8,1,2)
(13,1,2)
(17,1,2)
(27,1,2)
(38,1,2)
(39,1,2)
(26,1,2)
(36,1,2)
(39,1,2)
(36,1,2)
(35,1,2)
(34,1,2)
(37,1,2)
(33,1,2)
(29,1,2)
(25,1,2)
(31,1,2)
(24,1,2)
(34,1,2)
(39,1,2)
(27,1,2)
(15,1,2)
(38,1,2)
(38,1,2)
(28,1,2)
(36,1,2)
(28,1,2)
```







## Question.8 ...

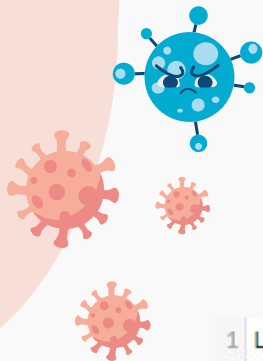
For the women

```
1 Load_data = LOAD '/user/maria_dev/Covie_19/Covid_df.txt'
2 USING PigStorage(',') as (USMER:int,MEDICAL_UNIT:int,SEX:int,PATIENT_TYPE:chararray,
3 PNEUMONIA:int,AGE:int,PREGNANT:int,DIABETES:int,COPD:int,ASTHMA:int,INMSUPR:int,
4 HIPERTENSION:int,OTHER_DISEASE:int,CARDIOVASCULAR:int,OBESITY:int,RENAL_CHRONIC:int,
5 TOBACCO:int,CLASIFFICATION_FINAL:int,DEATH:int);
6 data_foreach = FOREACH Load_data GENERATE SEX,PREGNANT,AGE,CLASIFFICATION_FINAL;
7 she_female= FILTER data_foreach BY SEX == 1;
8 has_PREGNANT= FILTER she_female BY PREGNANT == 1;
9 has_corona= FILTER has_PREGNANT BY CLASIFFICATION_FINAL <= 3;
10 DUMP has_corona;
```

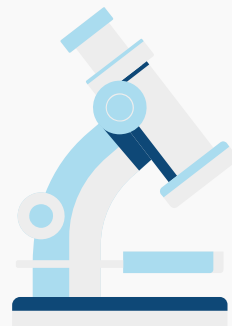
▼ Results

(1,1,23,3)  
(1,1,30,3)  
(1,1,33,3)  
(1,1,57,1)  
(1,1,25,1)  
(1,1,26,3)  
(1,1,25,3)  
(1,1,34,3)  
(1,1,34,3)  
(1,1,22,3)  
(1,1,30,3)  
(1,1,24,3)  
(1,1,27,3)  
(1,1,30,3)  
(1,1,25,3)  
(1,1,34,3)  
(1,1,31,3)  
(1,1,28,3)  
(1,1,30,3)  
(1,1,29,3)  
(1,1,33,3)  
(1,1,28,3)  
(1,1,20,3)  
(1,1,38,3)  
(1,1,31,3)  
(1,1,25,3)  
(1,1,22,3)  
(1,1,21,3)  
(1,1,44,3)





## Question.9 ...



Are they smoke?

```
1 Load_data = LOAD '/user/maria_dev/CovidProject/Covid_df.txt'
2 USING PigStorage(',') as (USMER:int,MEDICAL_UNIT:int,SEX:int,PATIENT_TYPE:chararray,
3 PNEUMONIA:int,AGE:int,PREGNANT:int,DIABETES:int,COPD:int,ASTHMA:int,INMSUPR:int,HIPERTENSION:int,
4 OTHER_DISEASE:int,CARDIOVASCULAR:int,OBESITY:int,RENAL_CHRONIC:int,TOBACCO:int,CLASIFFICATION_FINAL:int,DEATH:int);
5 data_foreach = FOREACH Load_data GENERATE MEDICAL_UNIT,SEX,PATIENT_TYPE,AGE,TOBACCO,CLASIFFICATION_FINAL;
6 positive_test = FILTER data_foreach BY CLASIFFICATION_FINAL<=3 ;
7 smooking= FILTER positive_test BY TOBACCO == 1;
8 gender_group= group smooking by SEX ;
9 DUMP gender_group;
```

▼ Results

[Download](#)

(1,{(12,1,1,32,1,3),(4,1,1,27,1,3),(4,1,1,26,1,3),(6,1,1,70,1,3),(4,1,1,46,1,3),(4,1,2,61,1,3),(12,1,1,32,1,3),(12,1,1,80,1,3),(4,1,2,50,1,3),(12,1,1,34,1,3),(12,1,1,97,1,3),(4,1,2,56,1,3),(4,1,2,2,59,1,3),(1,2,1,39,1,3),(3,2,2,43,1,3),(3,2,2,86,1,3),(3,2,2,74,1,3),(3,2,2,53,1,3),(3,2,2,57,1,3),(3,2,2,68,1,3),(3,2,2,65,1,3),(3,2,2,56,1,3),(3,2,2,69,1,3),(3,2,2,72,1,3),(3,2,1,53,1,3)

## Question.10 ...

Does anybody set in a hospital during covid-19 affected?

```
1 |load_data = LOAD '/user/maria_dev/CovidProject/Covid_df.txt'
2 |USING PigStorage(',') as (USMER:int,MEDICAL_UNIT:int,SEX:int,PATIENT_TYPE:chararray,
3 |PNEUMONIA:int,AGE:int,PREGNANT:int,DIABETES:int,COPD:int,ASTHMA:int,INMSUPR:int,HIPERTENSION:int,
4 |OTHER_DISEASE:int,CARDIOVASCULAR:int,OBESITY:int,RENAL_CHRONIC:int,TOBACCO:int,CLASIFFICATION_FINAL:int,DEATH:int);
5 |data_foreach = FOREACH Load_data GENERATE MEDICAL_UNIT,SEX,PATIENT_TYPE,AGE,CLASIFFICATION_FINAL;
6 |positive_test = FILTER data_foreach BY CLASIFFICATION_FINAL<=3 ;
7 |patient_type= FILTER positive_test BY PATIENT_TYPE matches '1' ;
8 |DUMP patient_type;
```

▼ Results

(1,1,1,65,3)  
(1,2,1,68,3)  
(1,1,1,64,3)  
(1,1,1,64,3)  
(1,1,1,38,3)  
(1,2,1,55,3)  
(1,1,1,48,3)  
(1,1,1,23,3)  
(1,2,1,61,3)  
(1,2,1,54,3)  
(1,2,1,30,3)  
(1,2,1,45,3)  
(1,1,1,26,3)  
(1,1,1,38,3)  
(1,2,1,24,3)  
(1,2,1,32,3)  
(1,2,1,49,3)  
(1,2,1,39,3)  
(1,2,1,27,3)  
(1,1,1,57,3)  
(1,2,1,25,3)  
(1,2,1,37,3)  
(1,1,1,49,3)  
(1,1,1,20,3)  
(1,1,1,56,3)  
(1,2,1,47,3)  
(1,1,1,50,3)  
(1,1,1,37,3)





# THANKS!



*Do you have any questions?*

*Shouq Alharbi*

*Razan Alajlan*

*Sarah Alrashidi*

*Nada Oteif*

*Hayam Alrashed*

