INTM 572: DATA EXPLOLARION AND PREPARATION

MONKEY POX PATIENTS' DATASET

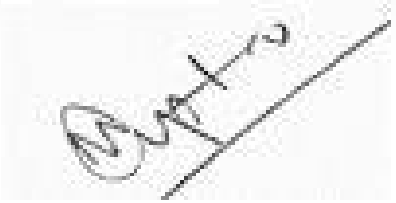
Submitted By: Mehak Gupta

Roll Number: RQ2255B41

Registration Number: 12221647

DECLARATION

I declare that this Assignment is my individual work. I have not copied it from any other student's work or from any other source except where due acknowledgement is made explicitly in the text, nor has any part been written for me by any other person.



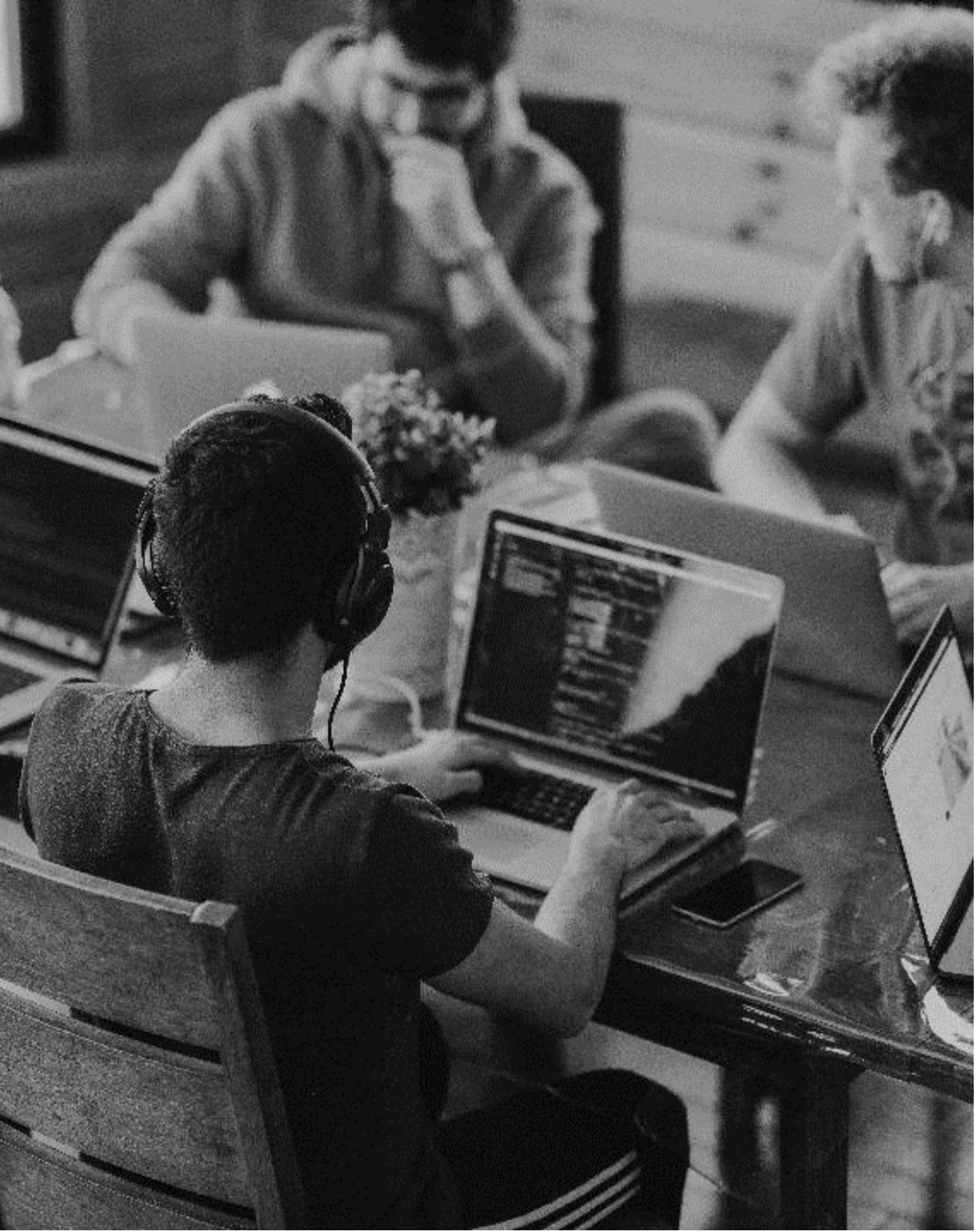
INDEX

INTRODUCTION

UNDERSTANDING THE DATA

DATA VISUALISATION

CONCLUSION



INTRODUCTION

WHAT IS MONKEYPOX?

Monkeypox is a viral zoonosis (a virus transmitted to humans from animals) with symptoms similar to smallpox patients in the past, but it is clinically less severe.

Monkeypox has emerged as the most important ortho pox virus for public health since the eradication of smallpox in 1980 and the subsequent cessation of smallpox vaccination.

It is primarily found in central and western Africa, often near tropical rainforests, and is becoming more common in urban areas. A variety of rodents and non-human primates serve as hosts.

UNDERSTANDING THE DATA

LET'S DIVE IN

```
#getting information about the data
print("Data Types of Data:-\n")
print(Mehak_df.info(),'\n')
print('----------'*9)
print("Shape of Data:-\n")
print(Mehak_df.shape,'\n')
print('----------'*9);
```

```
Data Types of Data:-

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 25000 entries, 0 to 24999
Data columns (total 11 columns):
 #   Column                      Non-Null Count  Dtype 
---  ------                      --------------  ----- 
 0   Patient_ID                  25000 non-null  object
 1   Systemic Illness            25000 non-null  object
 2   Rectal Pain                 25000 non-null  bool  
 3   Sore Throat                 25000 non-null  bool  
 4   Penile Oedema               25000 non-null  bool  
 5   Oral Lesions                25000 non-null  bool  
 6   Solitary Lesion             25000 non-null  bool  
 7   Swollen Tonsils             25000 non-null  bool  
 8   HIV Infection               25000 non-null  bool  
 9   Sexually Transmitted Infection  25000 non-null  bool  
 10  MonkeyPox                   25000 non-null  object
dtypes: bool(8), object(3)
memory usage: 781.4+ KB
None

----------------------------------------------------------------------------------------------------
Shape of Data:-

(25000, 11)
```

```
#checking for missing values
print("Checking Null entries & their Percentage in Data")
MissingValues=pd.DataFrame(zip(Mehak_df.isnull().sum(),Mehak_df.isnull().sum()*100/len(Mehak_df)),
                           columns=['Missing Values','Percentage Missing Values'],index=Mehak_df.columns)
MissingValues
```

Checking Null entries & their Percentage in Data

	Missing Values	Percentage Missing Values
Patient_ID	0	0.0
Systemic Illness	0	0.0
Rectal Pain	0	0.0
Sore Throat	0	0.0
Penile Oedema	0	0.0
Oral Lesions	0	0.0
Solitary Lesion	0	0.0
Swollen Tonsils	0	0.0
HIV Infection	0	0.0
Sexually Transmitted Infection	0	0.0
MonkeyPox	0	0.0

```
#checking for duplicate records
Mehak_df[Mehak_df.duplicated()]
```

Patient_ID	Systemic Illness	Rectal Pain	Sore Throat	Penile Oedema	Oral Lesions	Solitary Lesion	Swollen Tonsils	HIV Infection	Sexually Transmitted Infection	MonkeyPox

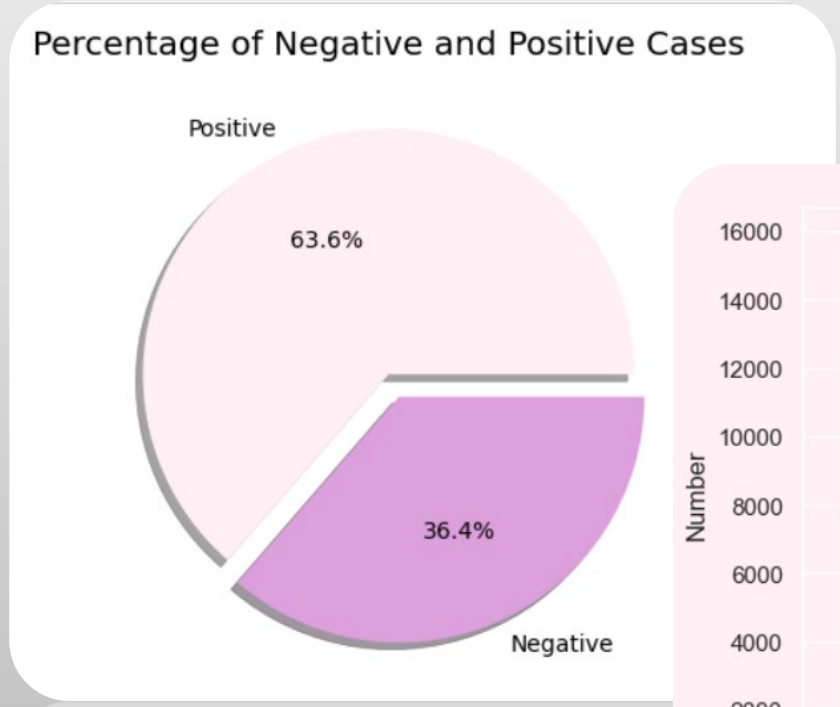
From here, we can understand the following:

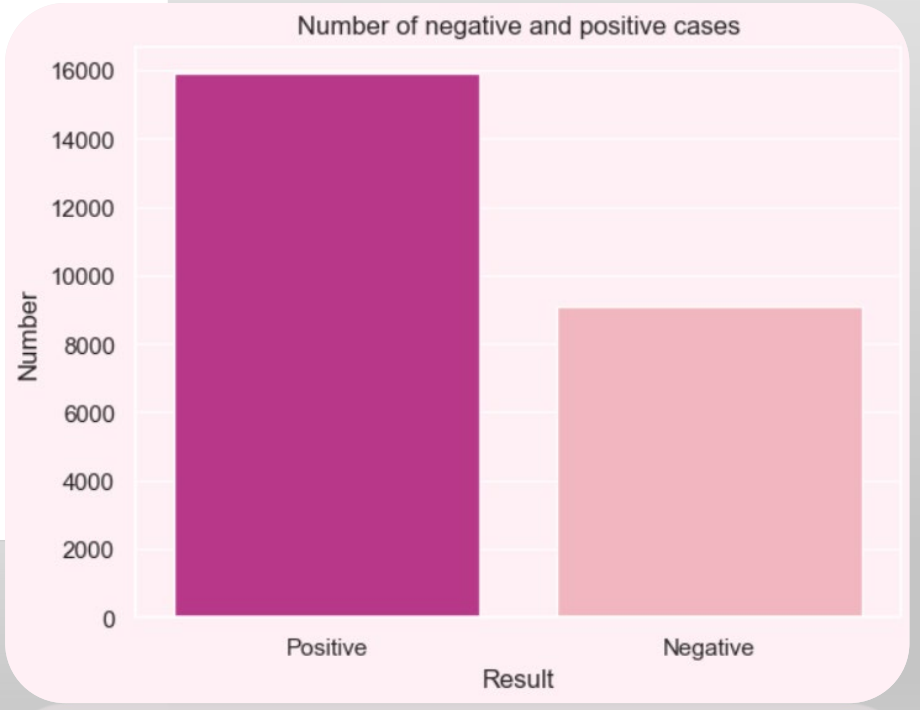
- There are 11 features in the data (i.e., columns)and 25000 sample sets (i.e., rows)
- Three of the features are object while the others are bool.
- There are no duplicate records.
- No missing values or null values have been found in the dataset.

DATA VISUALISATIONS

LET'S DIVE IN

POSITIVE AND NEGATIVE CASES





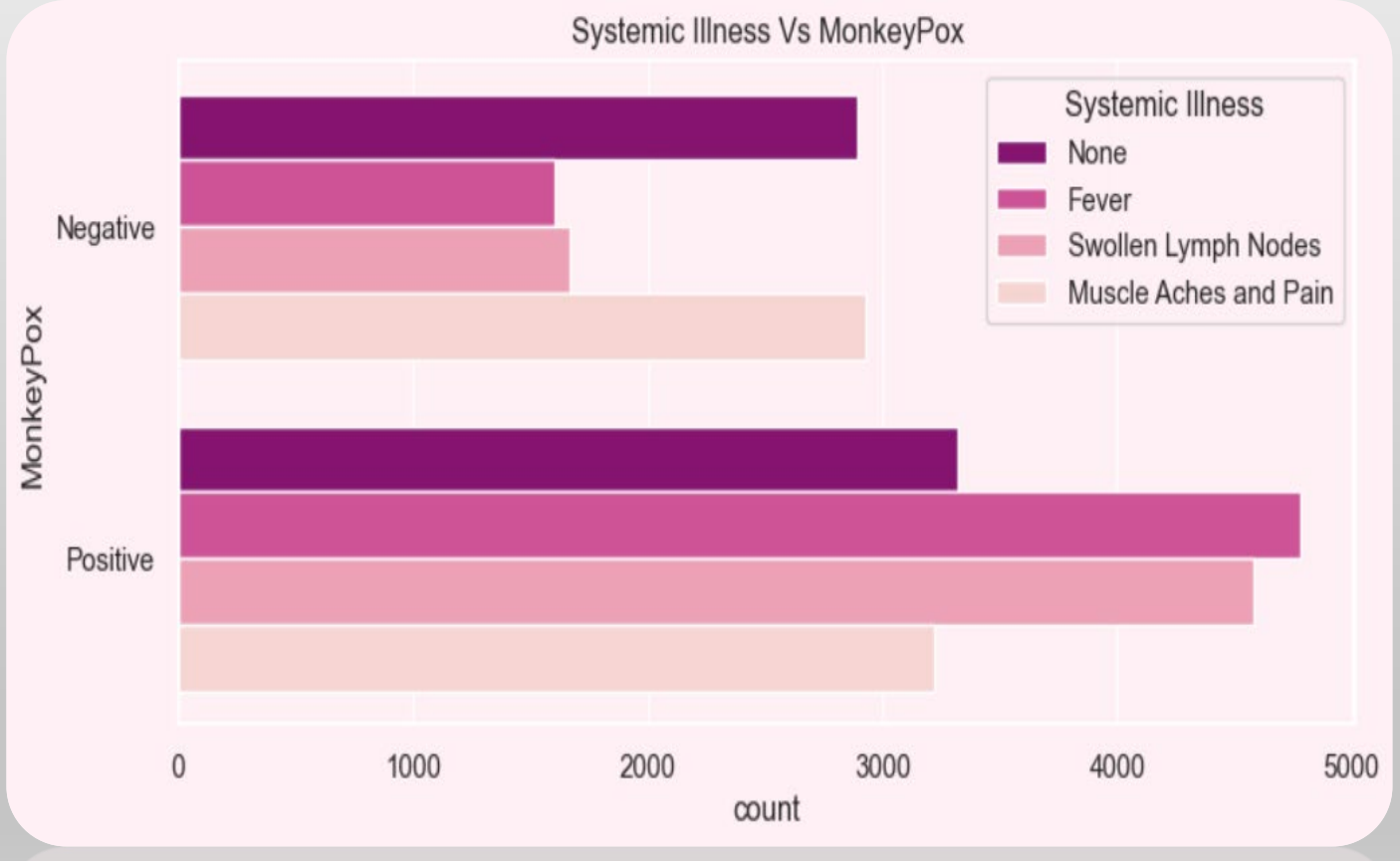
Out of the 25,000 recorded patient details, 63.6% i.e., 15,900 were tested positive for Monkey Pox while 36.4% of them (i.e., 9,100) of the patients tested negative.

COMPARISON OF SYMPTOMS AND RESULTS

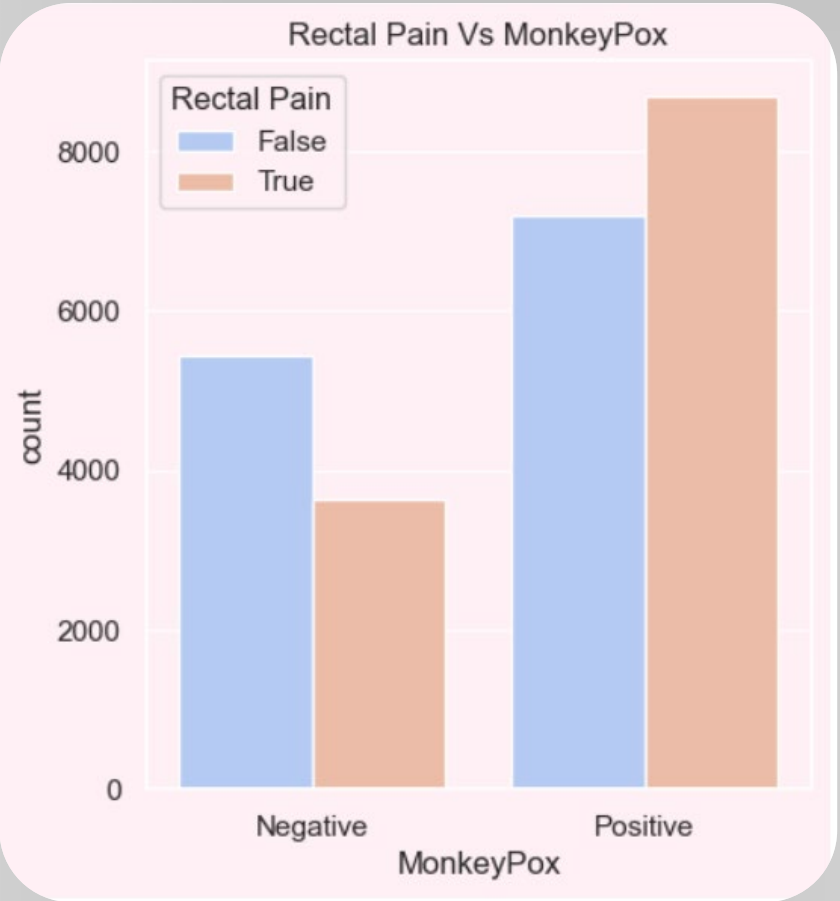
As per dataset, people showing symptoms of fever were most likely to be positive along with the symptoms of Swollen Lymph Nodes.

Most of the patients showing none of the symptoms or muscle aches and pain turned out to be tested negative.

Moreover, over 3000 patients with no systematic illness were tested positive as well.



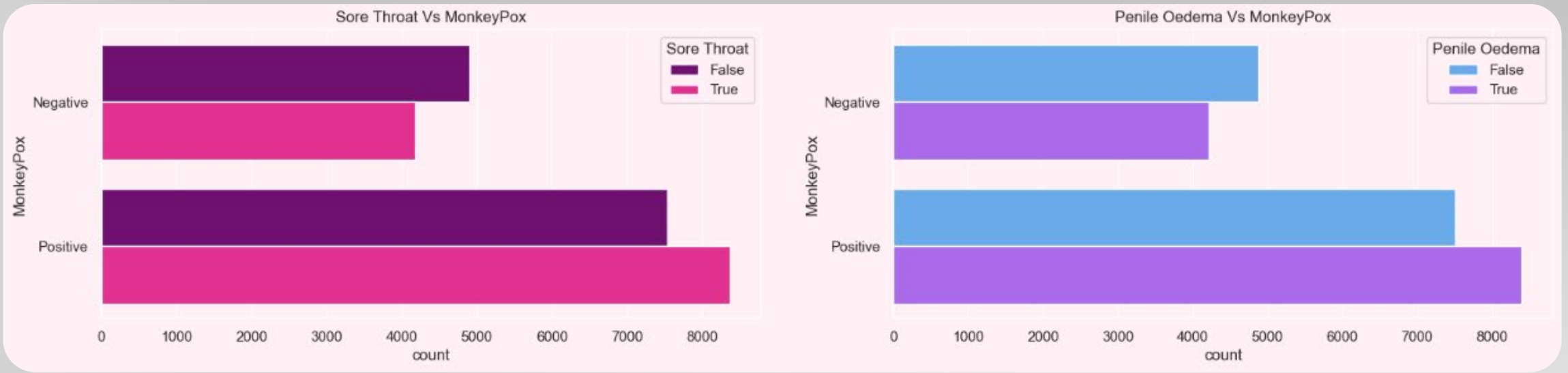
COMPARISON OF SYMPTOMS AND RESULTS



Over 9000 patients with who tested positive showed the symptoms of rectal pain while around 7000 without rectal pain were positive as well.

Moreover, around 4000 patients with rectal pain turned out to be negative.

COMPARISON OF SYMPTOMS AND RESULTS



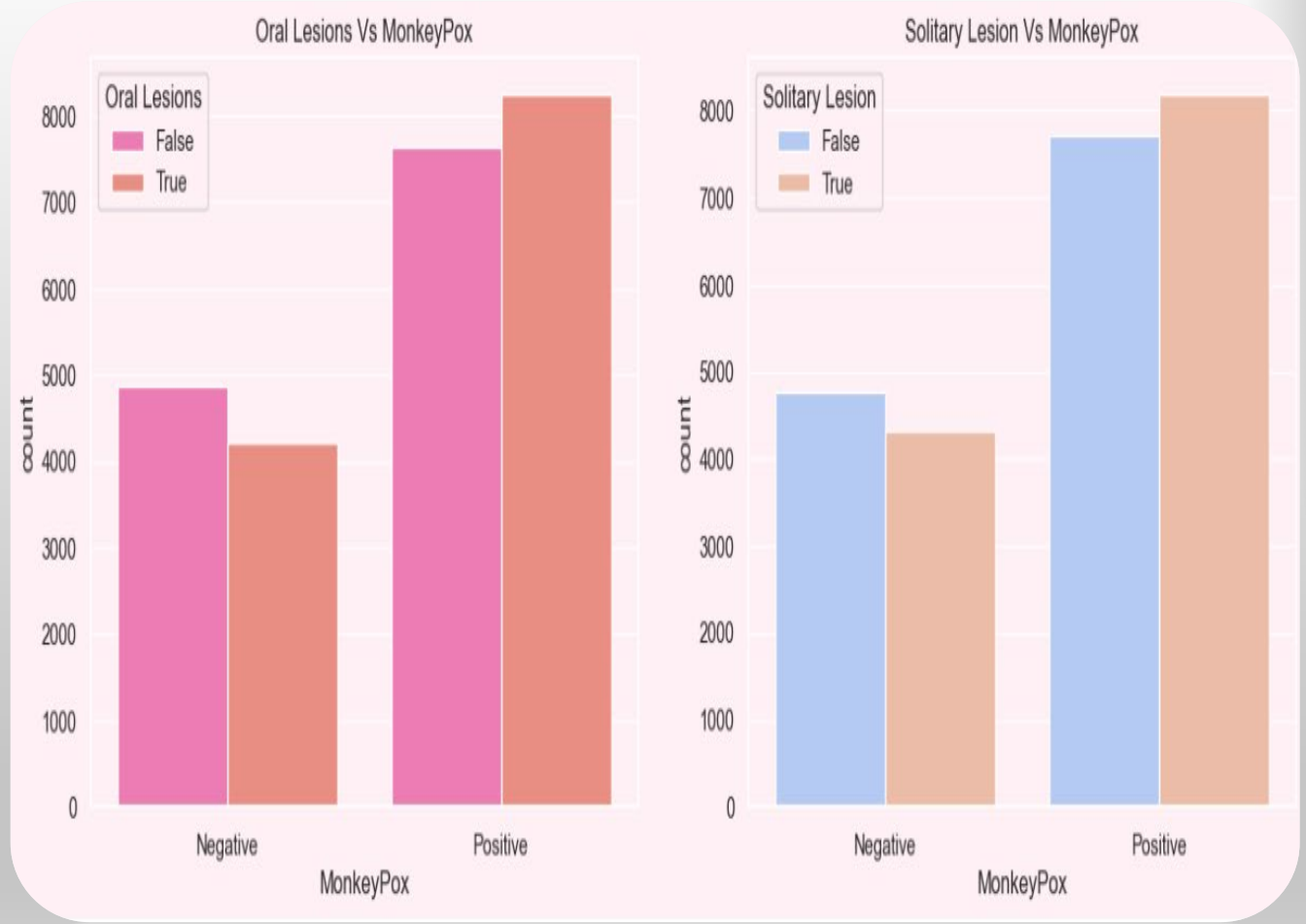
Over 8000 patients with sore throat and penile oedema were positive while around 5000 patients even without these symptoms were tested positive for Monkey Pox as well.

Moreover, around 4000 patients with the symptoms were tested negative.

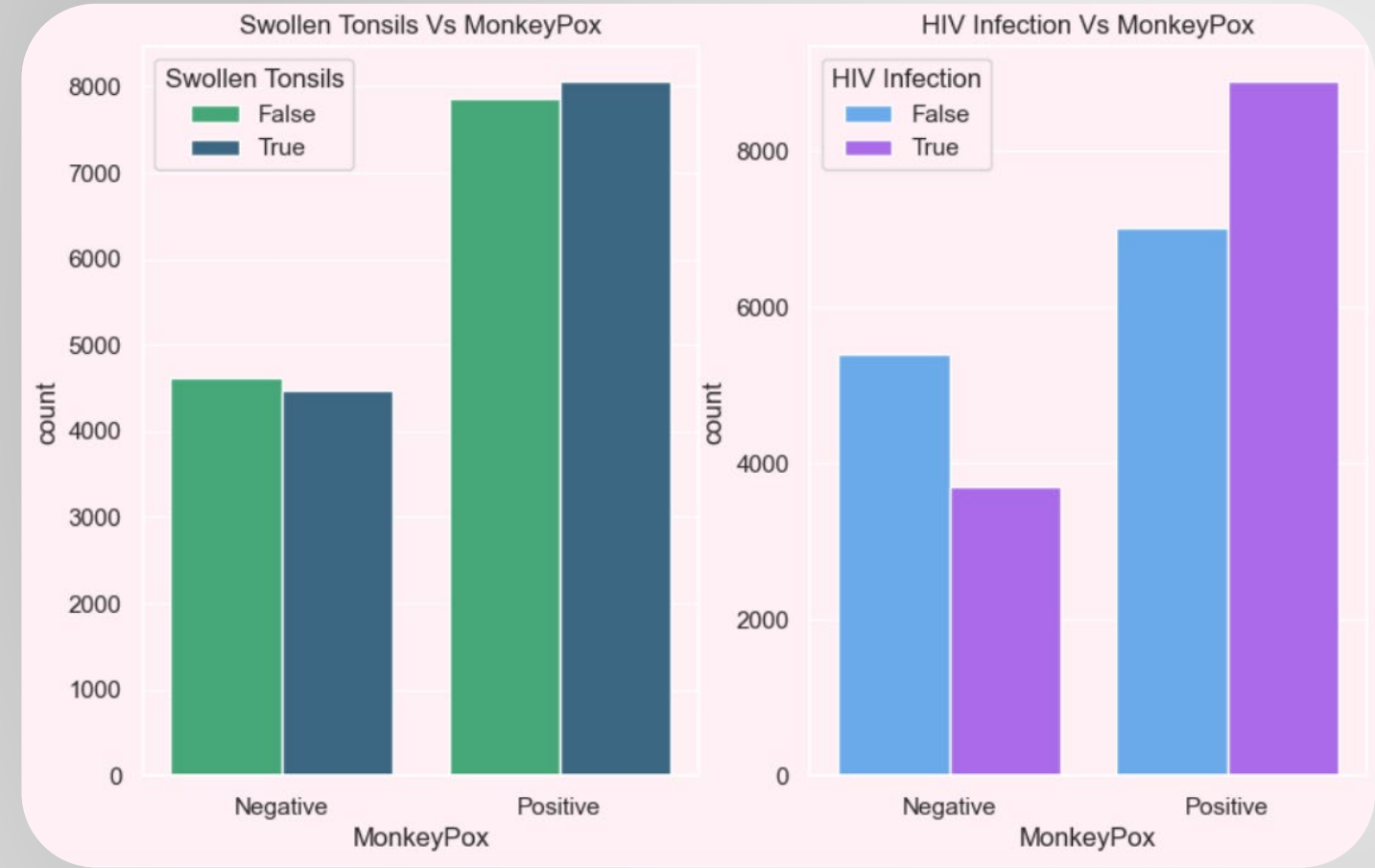
COMPARISON OF SYMPTOMS AND RESULTS

Out of the patients who tested positive for Monkey Pox, around 9000 had oral lesions and solitary lesions while around 5000 with these symptoms turned out to be negative.

In addition to this, over 7000 patients without the symptoms were tested positive.



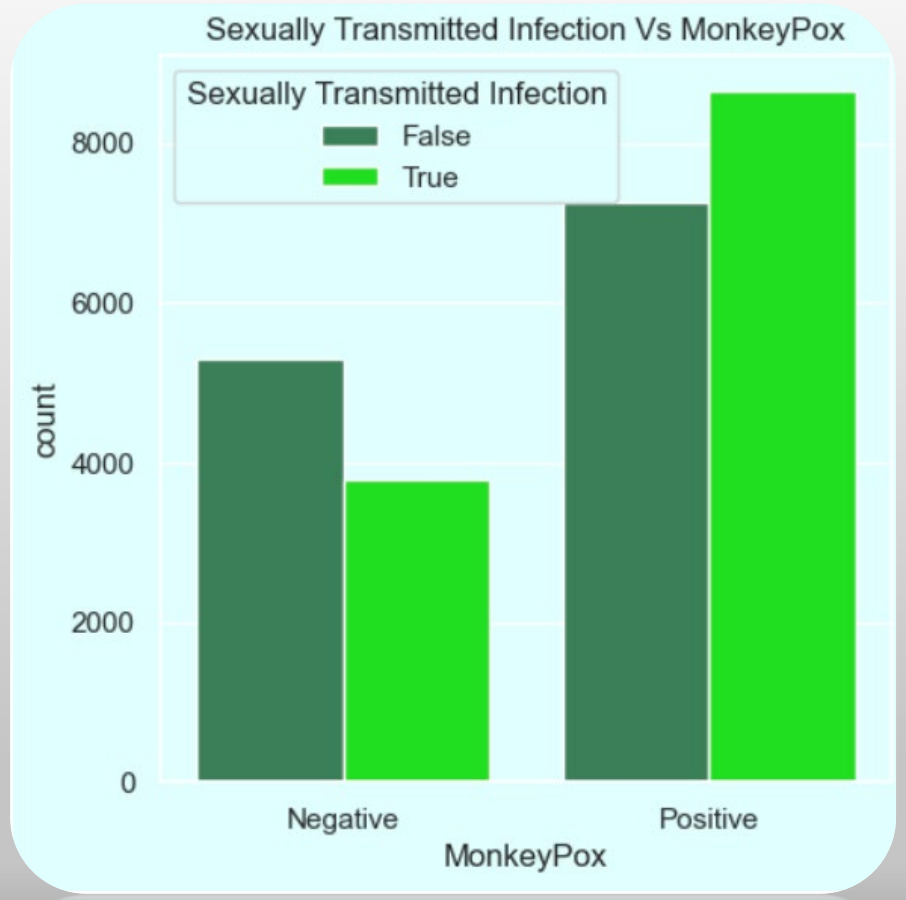
COMPARISON OF SYMPTOMS AND RESULTS



Out of the patients who had HIV Infection and Swollen Tonsils, around 9000 had were tested positive while around 5000 patients with the symptoms did not have monkey pox.

Moreover, around 7000 patients without the symptoms were tested positive.

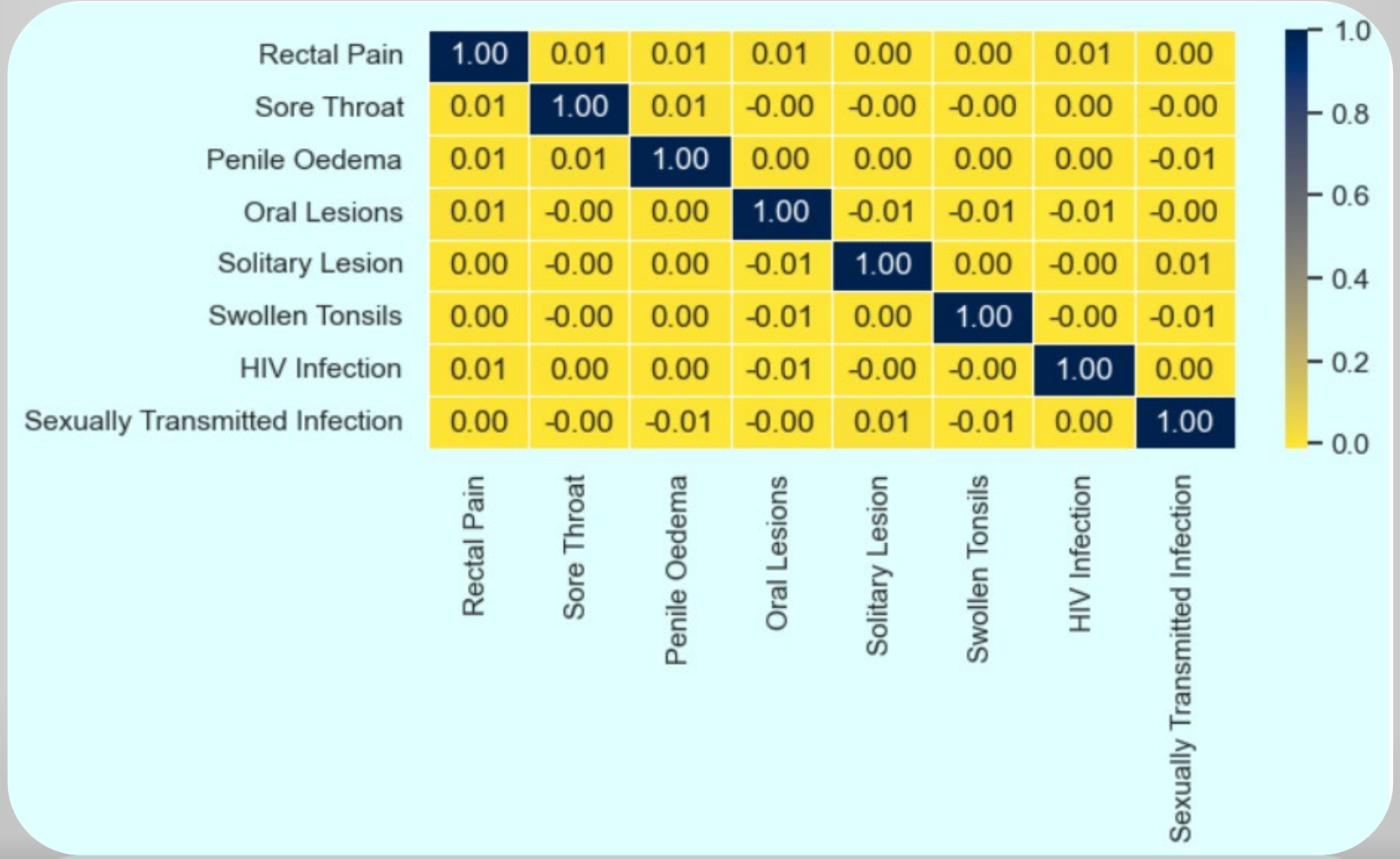
COMPARISON OF SYMPTOMS AND RESULTS



Around 9000 patients with sexually transmitted infection had Monkey Pox while around 7000 of positive patients did not have any sexually transmitted infections.

Over 5000 patients with Sexually Transmitted Diseases did not have Monkey Pox

HEAT MAP


Rectal Pain 1.00 0.01 0.01 0.01 0.00 0.00 0.01 0.00
Sore Throat 0.01 1.00 0.01 -0.00 -0.00 -0.00 0.00 -0.00
Penile Oedema 0.01 0.01 1.00 0.00 0.00 0.00 0.00 -0.01
Oral Lesions 0.01 -0.00 0.00 1.00 -0.01 -0.01 -0.01 -0.00
Solitary Lesion 0.00 -0.00 0.00 -0.01 1.00 0.00 -0.00 0.01
Swollen Tonsils 0.00 -0.00 0.00 -0.01 0.00 1.00 -0.00 -0.01
HIV Infection 0.01 0.00 0.00 -0.01 -0.00 -0.00 1.00 0.00
Sexually Transmitted Infection 0.00 -0.00 -0.01 -0.00 0.01 -0.01 0.00 1.00
Rectal Pain
Sore Throat
Penile Oedema
Oral Lesions
Solitary Lesion
Swollen Tonsils
HIV Infection
Sexually Transmitted Infection
1.0
0.8
0.6
0.4
0.2
0.0

CONCLUSION

A SUMMARY OF ANALYSIS

SUMMARY

- Out of every 250 patients, 159 of them were tested positive for monkey pox (i.e., 63%).
- Around 69-70% of the patients showing the symptoms were actually positive while 57-59% of the patients without the symptoms were tested positive for Monkey Pox.
- There is little to no correlation amongst the symptoms (as shown by the heat map).


THANK YOU

ANNEXURE

THE JUPYTER NOTEBOOK